Final Report

Pesticides Fate and Transport Predicted by the Soil and Water Assessment Tool (SWAT)

Atrazine, Metolachlor and Trifluralin in the Sugar Creek Watershed

By
S.L. Neitsch, J.G. Arnold, R. Srinivasan
Grassland, Soil & Water Research Laboratory, USDA-ARS
Blackland Research Center, TAES
Temple, Texas

Submitted to

Office of Pesticide Programs
Environmental Protection Agency

February 27, 2002

BRC Publication #2002-03

Project background

The U.S. Environmental Protection Agency, Office of Pesticide Programs (OPP) is evaluating existing basin-scale models for possible use in drinking water assessments conducted under the Food Quality Protection Act (FQPA) and ecological risk assessments conducted under the Federal Insecticide Fungicide and Rodenticide Act (FIFRA). The Soil and Water Assessment Tool (SWAT) is one of the models being evaluated.

As part of the evaluation process, an application of the model to the Sugar Creek watershed in the White River Basin in central Indiana is being performed. Model outputs are compared to USGS National Water Quality Assessment (NAWQA) field data. Several criteria must be addressed in the field evaluation of SWAT: 1) accuracy of the model in predicting pesticide concentrations at designated locations, 2) ease of model use, 3) cost of simulations in terms of time and/or contractor dollars, and 4) usefulness and relevance of the model to the OPP exposure assessment process.

Field evaluation of SWAT in the Sugar Creek watershed consists of: a) cold simulation (without calibration) for 1992 with report of daily atrazine, metolachlor and trifluralin concentration values at NAWQA sampling point; b) hydrology calibration (flow and sediment where available) for 1992 with report of daily atrazine, metolachlor and trifluralin concentration values at NAWQA sampling point and a summary of the calibration process; c) atrazine calibration for 1992 with report of daily atrazine concentration values at NAWQA sampling point and a summary of the calibration process; d) metolachlor calibration for 1992 with report of daily metolachlor concentration values at NAWQA sampling point and a summary of the calibration process; f) trifluralin calibration for 1992 with report of daily trifluralin concentration values at NAWQA sampling point and a summary of the calibration process; g) report of daily atrazine, metolachlor and trifluralin concentration values at NAWQA sampling point for 1993-1995 period based upon calibrated hydrology and 1992 pesticide calibration with a short discussion on experience of the process. The final report on testing of SWAT in Sugar Creek watershed should include: input parameters for SWAT, results from SWAT modeling of atrazine, metolachlor and trifluralin in the Sugar Creek watershed, comparison of SWAT results to measured concentrations, strengths and weaknesses of SWAT, documentation of ease of use and computational time, and recommendations of the usefulness of SWAT to OPP exposure assessment.

Model Description

SWAT is the acronym for Soil and Water Assessment Tool, a river basin scale model that addresses large area water resource development and management. A complete overview of SWAT is given in the on-line theoretical documentation (Neitsch et al, 2001a). For convenience, a description of the pesticide algorithms is appended to this document.

Description of the Sugar Creek Watershed

The Sugar Creek watershed in central Indiana is a poorly drained agricultural watershed typical of many areas in the Midwestern USA. The Sugar Creek watershed is within the White River Basin, a river basin being studied as part of the USGS National Water-Quality Assessment Program.

Sugar Creek is a perennial stream in the White River Basin that drains 242 km^2 upstream from New Palestine, Indiana (Figure 1). The Sugar Creek watershed is dominated by agricultural land use and has poorly drained soils. Land use in the watershed is 95% row-crop agriculture. The principle crops are corn and soybean with small amounts of alfalfa and wheat. The soils in the watershed were mapped primarily in the Crosby-Brookston soil association. This association is characterized by poorly drained, nearly level, loamy soils developed on Wisconsin glacial till. Tile-drain systems have been installed in areas used for agriculture.

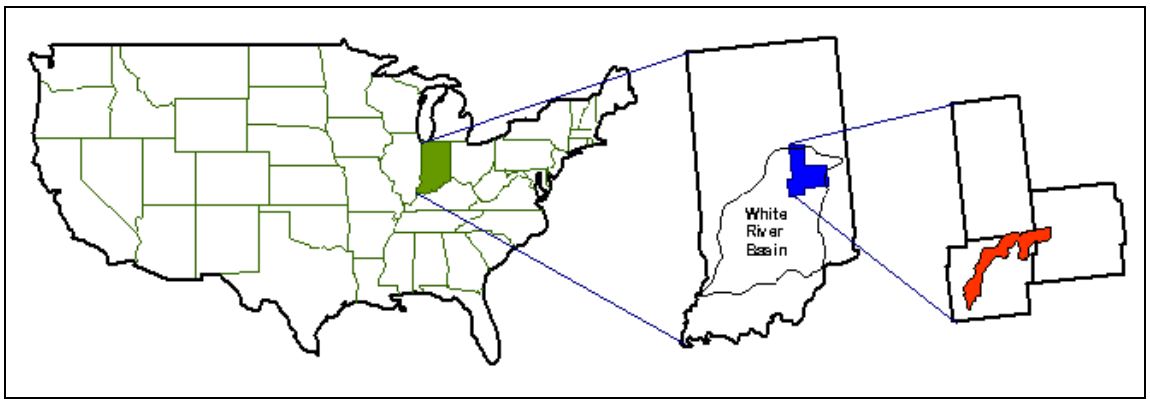


Figure 1: Location of Sugar Creek Watershed

The average annual precipitation in the Sugar Creek watershed is 1000 mm. The precipitation is distributed fairly evenly throughout the year. Storm events in the winter and early spring tend to be of long duration and low intensity while storms in the late spring and summer are short duration, high intensity events.

Model Inputs

SWAT is a physically based model and requires information about weather, soil properties, topography, natural vegetation, and cropping practices. Input data for SWAT was assembled with the SWAT2000 ArcView Interface.

SWAT model input data for topography were extracted from a digital elevation model (DEM) for the contiguous U.S. The DEM was assembled from quadrangles containing 1:250,000 scale USGS 1° by 1°, 3 arc-second data. The horizontal cell size of this data is 100 by 100 meters and the vertical resolution is 1 meter. Figure 2 displays the elevation information for the watershed.

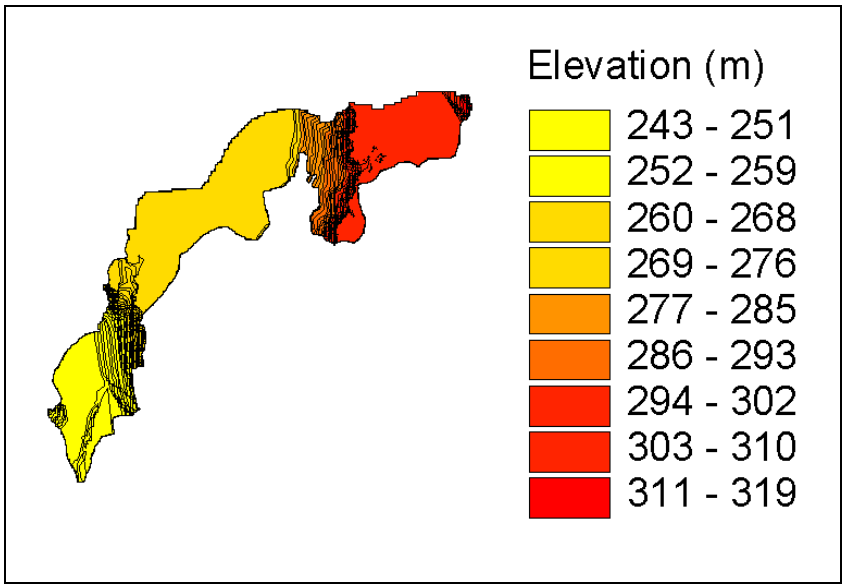


Figure 2: DEM for Sugar Creek watershed

Because the Sugar Creek watershed has very little relief, the interface was unable to correctly predict the stream flow paths using the DEM alone. To obtain a proper stream path delineation, a stream delineation from USGS was overlaid on the DEM and used to burn in the location of the streams in the watershed. Based on the DEM and stream path map themes, the SWAT2000 ArcView interface subdivided the watershed into 23 subbasins (Figure 3). Subbasin delineation was based on natural flow paths and watershed divides.

The watershed outlet coincides with the position of the U.S. Geological Survey streamflow gaging station 03361650 (39°42'51"N 85°53'08"W). The NAWQA water quality sampling station 394340085524601 is located 1.0 river miles upstream from the streamflow gaging station (39°43'40"N 85°52'46"W). Due to the close proximity of the streamflow gage and the water quality sampling station, simulation results calculated for the watershed outlet were used for both the hydrology and water quality components of the study.

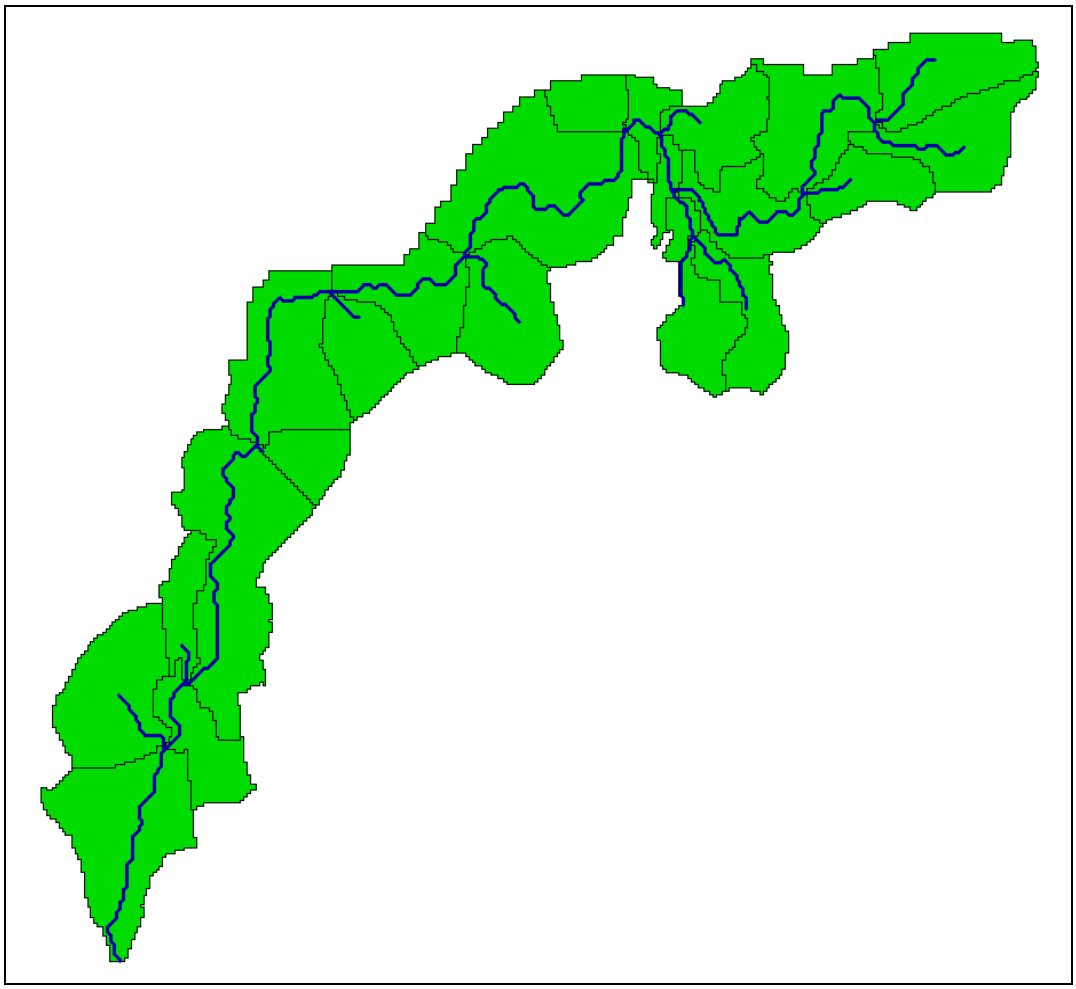


Figure 3: Subbasin delineation

Due to the low relief in the DEM, slope-length values initially assigned by the SWAT2000 ArcView interface were not realistic. Slope length values were obtained from NRI points within the three counties that the Sugar Creek watershed is located. These values were spatially averaged to obtain the following slope-lengths used in the simulation: 72.7 meters for corn/soybean rotation, 65.3 meters for pasture and 71.2 meters for all landuses.

Climatic data are assigned or generated by SWAT at the subbasin level. All HRUs within a subbasin use the same climatic data. Daily measured precipitation data for the simulation were obtained from 5 weather stations located around the watershed (Figure 4). The measured precipitation data were assigned to subbasins based on proximity of the station to the centroid of the subbasin. The precipitation data were used with no modification for effect of distance between the subbasin and the weather station. Daily temperature data from the Greenfield weather station were used for all subbasins in the watershed. Daily solar radiation, relative humidity, and wind speed values were generated

from long-term monthly averages. The long-term monthly average data were acquired from Indianapolis and Hartford City, Indiana.

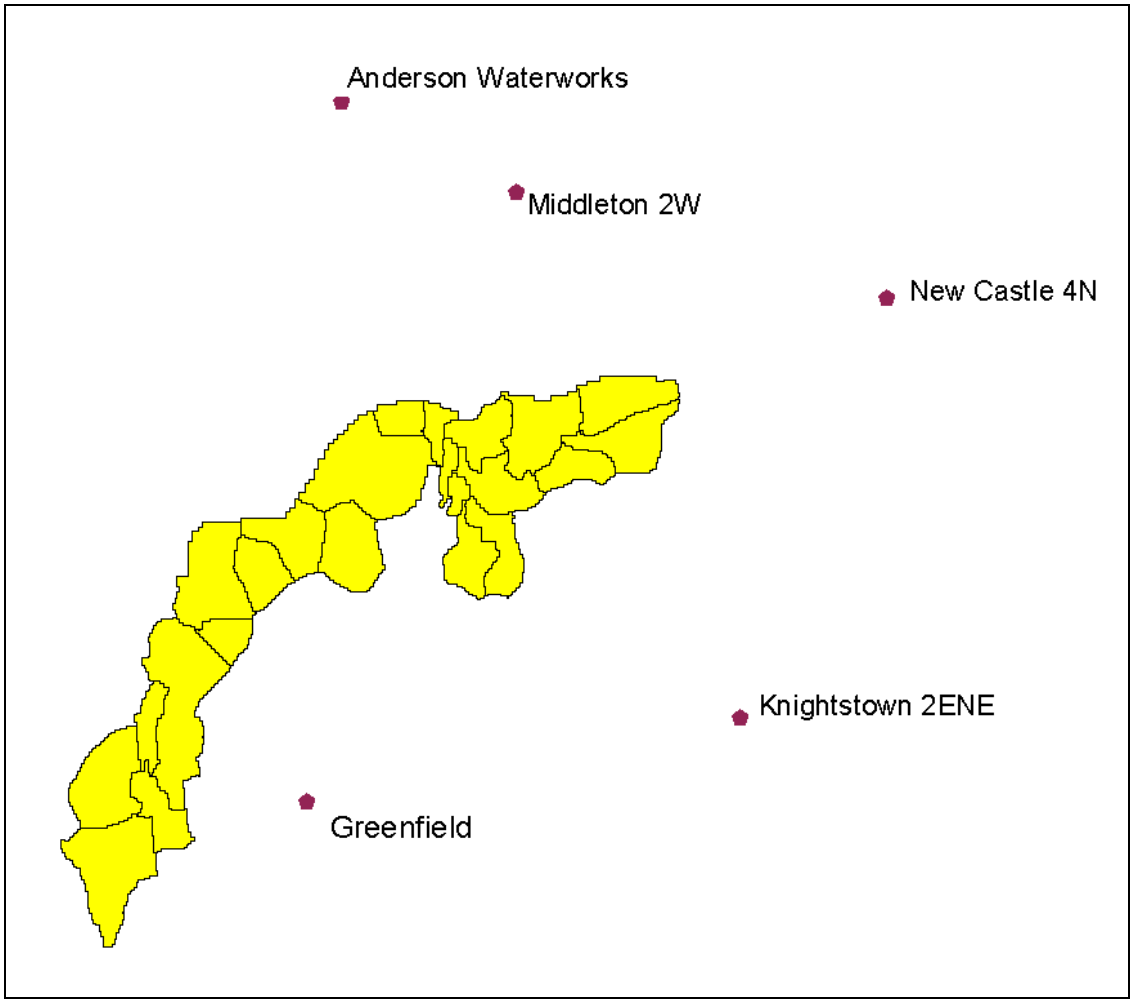


Figure 4: Weather stations used Sugar Creek watershed SWAT simulations

Within each subbasin, SWAT allows hydrologic response units (HRUs) to be defined. HRUs are sets of disconnected units in a subbasin with the same landuse and soil. The SWAT2000 ArcView interface requires land cover and soil maps to define HRUs.

The landuse map used for HRU creation was obtained from the U.S. Geological Survey. The USGS Indiana land cover map was created by Raytheon STX Corporation in 1998 with a 38 meter pixel resolution (Figure 5). The primary data source for the land cover map was leaves-off (spring) Landsat TM data acquired between 1988-1994. Leaves-on (primarily summer) TM data sets were also acquired and referenced.

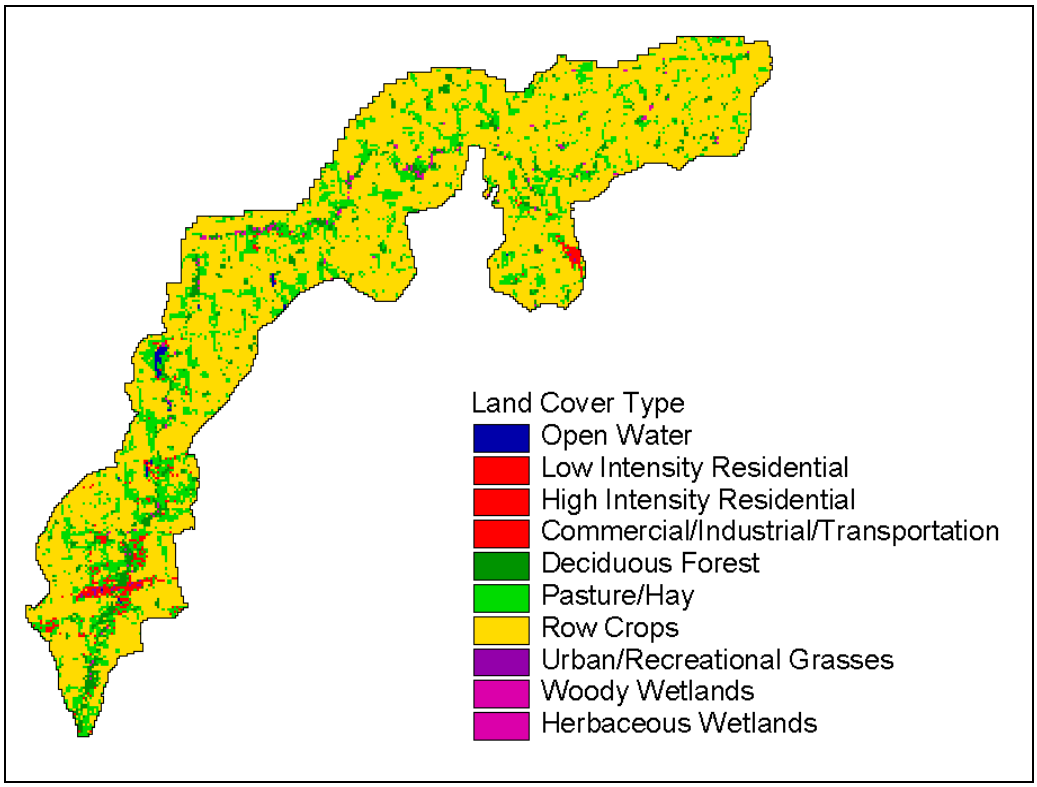


Figure 5: Land cover in Sugar Creek watershed

Land cover in the Sugar Creek watershed is dominated by Row Crops (75.35% of the watershed area) and Pasture/Hay (16.87%) with small areas of Deciduous Forest (4.89%) and Low Intensity Residential (1.48%). All other land cover categories represented less than one percent of the watershed area.

When creating HRUs, the land cover categories had to be linked to land covers defined in the SWAT land cover/plant growth database. The SWAT land cover categories used to model the various map categories are:

Open Water, not modeled as a land area

Low Intensity Residential, modeled as residential-medium/low density with Kentucky bluegrass growing in pervious areas

High Intensity Residential, modeled as residential-high density with Kentucky bluegrass growing in pervious areas

High Intensity Commercial/Industrial/Transportation, modeled as commercial with Kentucky bluegrass growing in pervious areas

Deciduous Forest, modeled as deciduous forest

Pasture/Hay, modeled as alfalfa

Row Crops, modeled as a corn/soybean rotation

Other Grasses (Urban/recreational: parks, lawns, golf courses), modeled as Kentucky bluegrass

Woody Wetlands, modeled as forested wetlands

Herbaceous Wetlands, modeled as non-forested wetlands

Information for soils was obtained from the USDA-NRCS State Soil Geographic Data Base (STATSGO) (USDA Soil Conservation Service, 1992) soil association data set. This data set was assembled from state DLG-3 files. There are 78,863 soil association polygons in the U.S. with up to 21 soil series in each polygon. The pixel resolution of the STATSGO map theme is 250 meters. Figure 6 displays the STATSGO soils map for the Sugar Creek watershed.

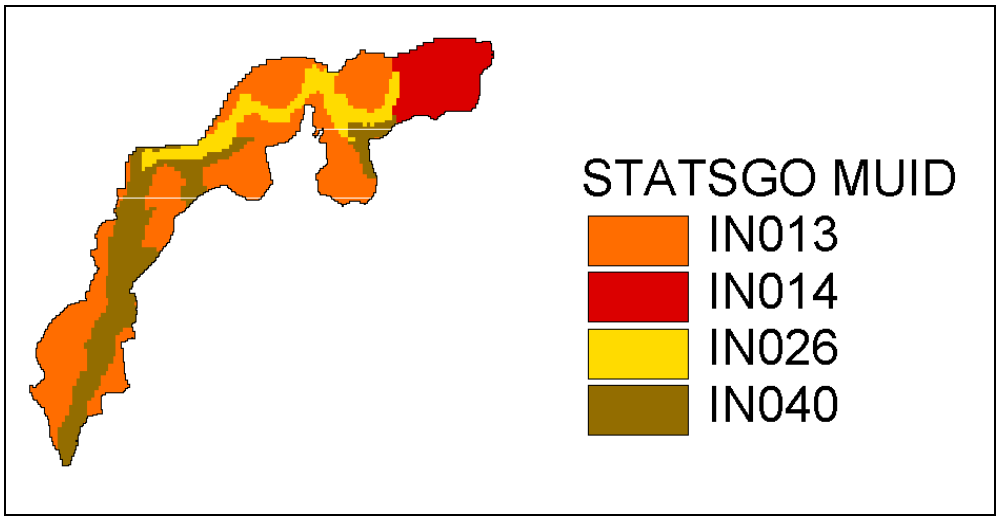


Figure 6: STATSGO soil map for Sugar Creek watershed.

Four soil associations are present in the Sugar Creek watershed. STATSGO polygon IN013 (50.21% of the watershed area) is dominated by the Crosby and Treaty soil series. STATSGO polygon IN014 (13.82%) is comprised primarily of the Crosby and Cyclone soil series. STATSGO polygon IN026 (11.59%) is dominated by the Ockley and Fox soil series while the Miami and Crosby soil series are the dominant soil series found in STATSGO polygon IN040 (24.38%). Data for only one soil series can be used to simulate soil processes in an HRU. The soil series used to model the four different STATSGO polygons are: IN013-Crosby, IN014-Crosby, IN026-Ockley, IN040-Miami.

Input variables for the SWAT model are reviewed in the SWAT2000 User's Manual (Neitsch et al, 2001b). The SWAT2000 ArcView interface populates the input fields with the minimum amount of information needed to run SWAT. The default dataset created by the interface must be modified to more accurately reflect processes occurring in the watershed.

Hydrology

The Sugar Creek watershed dataset was set up to run on a daily time step. Surface runoff is calculated using the SCS curve number method. The

Penman-Monteith method was used to determine potential evapotranspiration. Channel water routing was performed using the Muskingum routing method.

Land Management-General

The management scenarios for HRUs with row crops were modified to simulate a corn/soybean rotation. A conventional tillage schedule developed by Atwood et al (2000) was adopted for use in the simulation. This schedule was:

Tandem disk	2 weeks after harvest
Chisel	3 weeks after harvest
Tandem disk	3 weeks before planting
Tandem disk	2 weeks before planting
Field cultivator	1 week before planting
Plant	
Row cultivator	3 weeks after planting
Row cultivator	5 weeks after planting
Harvest	

Tile drains were simulated in all HRUs with a corn/soybean rotation. Three input variables control the functioning of tile drains in the HRUs. The depth to the tile drain was set to 800 mm. The time to drain the soil profile was set to 24 hours. The time till water enters the channel network after entering the tiles was set to 48 hours.

SWAT was set up to automatically apply fertilizer to HRUs with row crops. The fertilizer application took place any time corn experience a 20% reduction from optimal growth due to nitrogen stress.

Land Management-Pesticide

USGS provided daily amounts of applied pesticides for the entire watershed. The daily pesticide amounts were summed to 5-day totals and applied uniformly over the area simulated in the targeted crop. Multiple applications of the pesticides were simulated in the HRUs to capture the temporal distribution of pesticide application in the watershed. Atrazine was applied only to corn while metolachlor was applied to both corn and soybean. Trifluralin was applied only to soybean.

Pesticide Attribute Data

Pesticide properties that govern pesticide transport and degradation are stored in two input files in SWAT: the pesticide database (pest.dat) and the in-stream water quality file (.swq). Values from the GLEAMS pesticide database were used to populate input fields in the SWAT pesticide database file. For the

in-stream reaction rate, the aerobic soil half-life from the ARS pesticide database was used. The following table summarizes input values for the three pesticides.

Table 1: Pesticide property inputs

Property	Atrazine	Metolachlor	Trifluralin
Soil adsorption coefficient: K_{oc}	100.	200.	8000.
Wash-off fraction	0.45	0.60	0.40
Foliar half-life (days)	5.0	5.0	3.0
Soil half-life (days)	60.0	90.0	60.0
Application efficiency	0.75	0.75	0.75
Water solubility (mg/L)	33.0	530.0	0.3
In-stream/sediment reaction rate (1/day)	.0047	.0267	.0231
Pesticide partition coefficient between sediment and water in reach (m^3/g)	.0000	.0000	.0036

Evaluation of Model Prediction

Descriptive statistics such as the mean, standard error of the mean, median, standard deviation, variance, range, and maximum and minimum values were calculated for each set of comparisons. Three criteria recommended by the ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models (1993) were also included in the statistical analysis for hydrology. These criteria are the deviation of water yields, the Nash-Sutcliffe coefficient, and the coefficient of gain from the daily mean. In addition to these three, the coefficient of determination (R^2) is also calculated as part of the hydrologic analysis.

The deviation of water yields, D_v, quantifies the difference in observed and predicted water volumes and is calculated $D_v = \frac{(V - V')}{V} \cdot 100$ where V is the measured water yield for the period of comparison and V' is the model predicted water yield for the period of comparison.

The Nash-Sutcliffe coefficient, E_{NS}, measures how well the daily simulated and measured flows correspond. This coefficient is calculated $E_{NS} = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2}$ where O_i is the measured daily discharge, P_i is the computed daily discharge, and $\overline{O}$ is the average measured discharge. A Nash-Sutcliffe value can vary between 0.0 and 1.0 where a value of 1.0 indicates a perfect fit while a value of 0.0 indicates that the model is predicting no better than the average of the observed data.

The coefficient of gain from the daily mean, *DG*, compares model results with daily mean discharge values, which vary throughout the year. *DG* can vary between 0 and 1, with 0 being a perfect model. This coefficient is calculated

$DG = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O_i})^2}$ where O_i is the measured daily discharge, P_i is the computed daily discharge, and $\overline{O_i}$ is the average measured daily discharge.

Part 1: Cold Simulation Results

Cold simulation results are results produced by the model before any calibration is performed. This section of the report summarizes cold simulation results for 1992, the period of record used for model calibration, and 1993-1995, the period of record used for model validation.

Hydrology

Cold simulation daily stream flow for 1992 is graphed with measured streamflow in Figure 7. Figure 8 displays the same data for 1993-1995.

Table 2 summarizes the statistical data for the measured flow and SWAT cold simulation flow results. In addition to the time series graphs, box plots and scatter plots for the measured and cold simulation flow are provided in Figures 9-12.

For both periods of comparison, cold simulation results overestimated water yield in the watershed.

Table 2: Statistics for Cold Simulation Stream Flow

	1992 Daily Stream Flow		**1993-1995 Daily Stream Flow**	
Statistic:	**Measured**	**SWAT Cold Run**	**Measured**	**SWAT Cold Run**
Mean	**3.0091 m^3/s**	**3.3420 m^3/s**	**2.8562 m^3/s**	**3.0627 m^3/s**
Std error of mean	.2492	.2439	.1689	.1475
Median	1.5574	1.5450	1.1600	1.4800
Standard deviation	4.7672	4.6660	5.5919	4.8820
Variance	22.7264	21.7711	31.2692	23.8340
Range	36.81	28.43	54.59	46.10
Minimum	.28	.27	.06	.00
Maximum	37.10	28.70	54.65	46.10
D_v	-11.06% difference in water yield		-7.23% difference in water yield	
R^2		0.42		0.62
Nash-Sutcliffe E_{NS}		0.31		0.61
DG		0.33		0.57

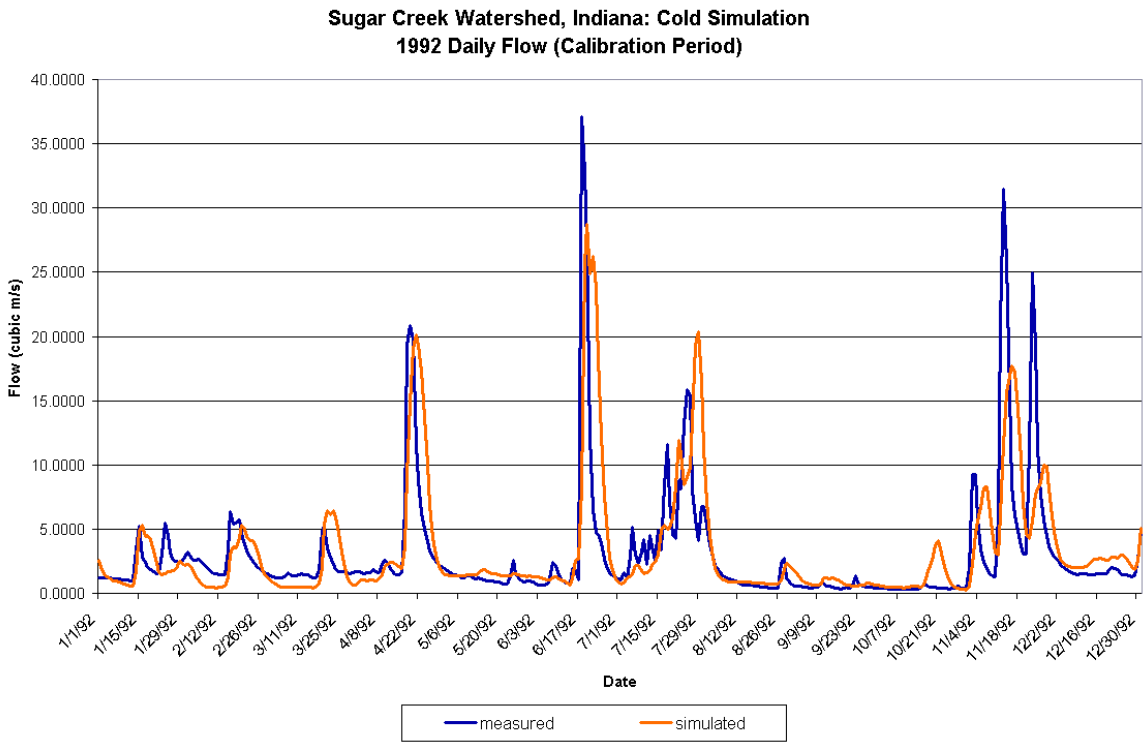


Figure 7: 1992 daily stream flow cold simulation results.

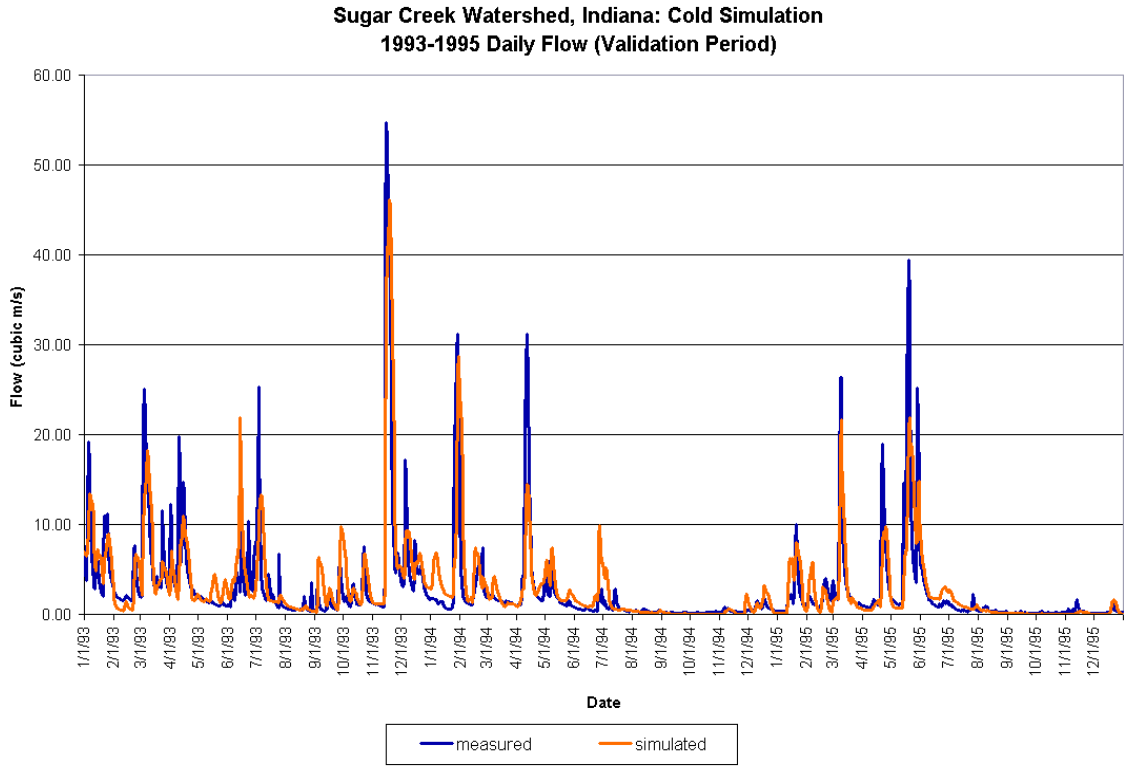


Figure 8: 1993-1995 daily stream flow cold simulation results.

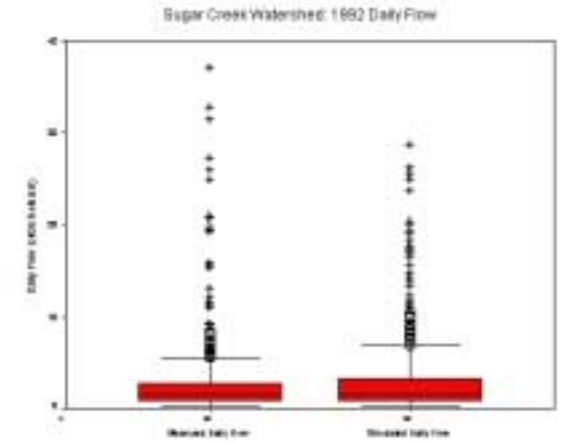


Figure 9: Box plot-1992 cold run stream flow

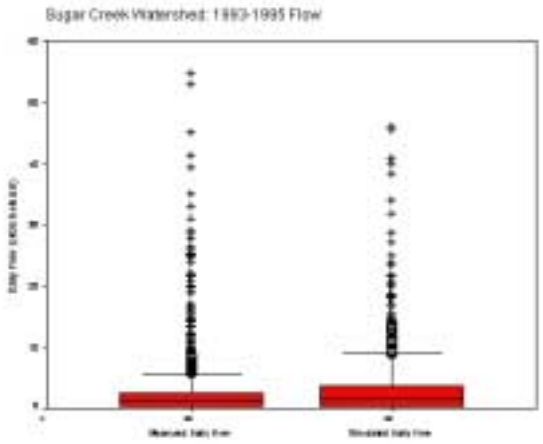


Figure 10: Box plot-1993-1995 cold run stream flow

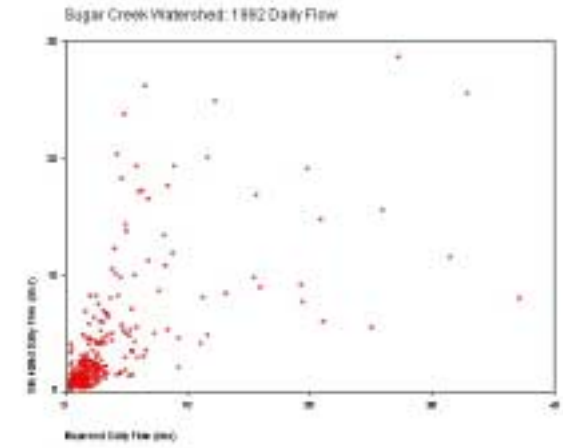


Figure 11: Scatter plot-1992 cold run stream flow

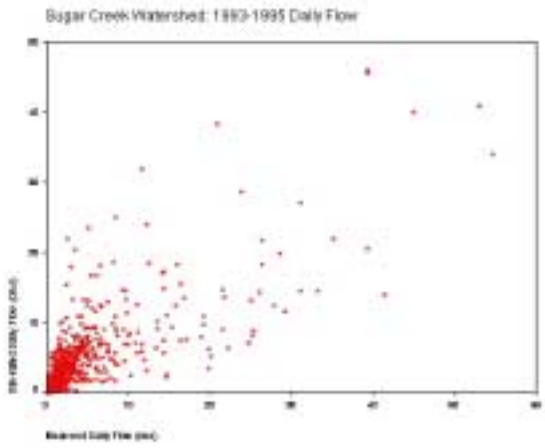


Figure 12: Scatter plot-1993-1995 cold run stream flow

Atrazine

Cold simulation atrazine concentrations for 1992 are graphed with measured grab sample concentrations in Figure 13. Figure 14 displays the same data for 1993-1995.

Table 3 summarizes the R^2 values computed from a regression analysis performed on the pairs of measured and simulated atrazine concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and cold simulation atrazine concentrations are provided in Figures 15-20.

Table 3: Statistics for Cold Simulation-Atrazine

Simulation:	**R^2**	
	conc	**log(conc)**
1992 atrazine, before stream flow calibration	.24	.52
1993-95 atrazine, before stream flow calibration	.31	.29

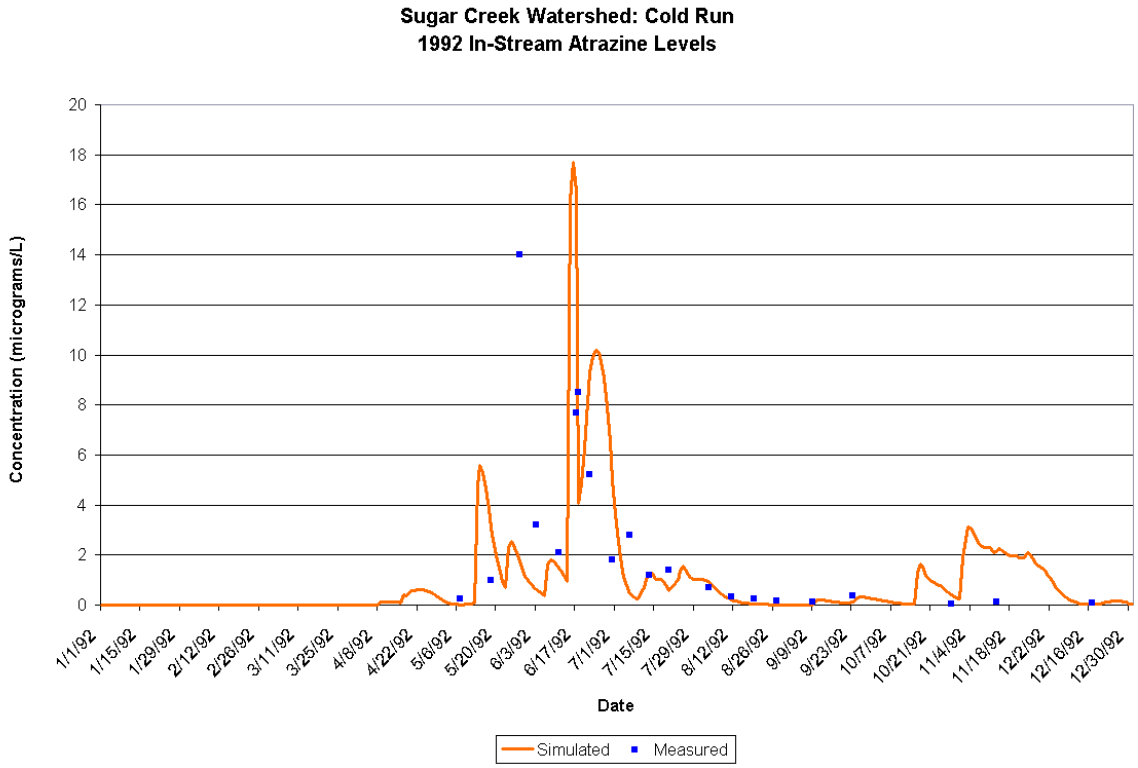


Figure 13: Cold simulation atrazine concentrations for 1992

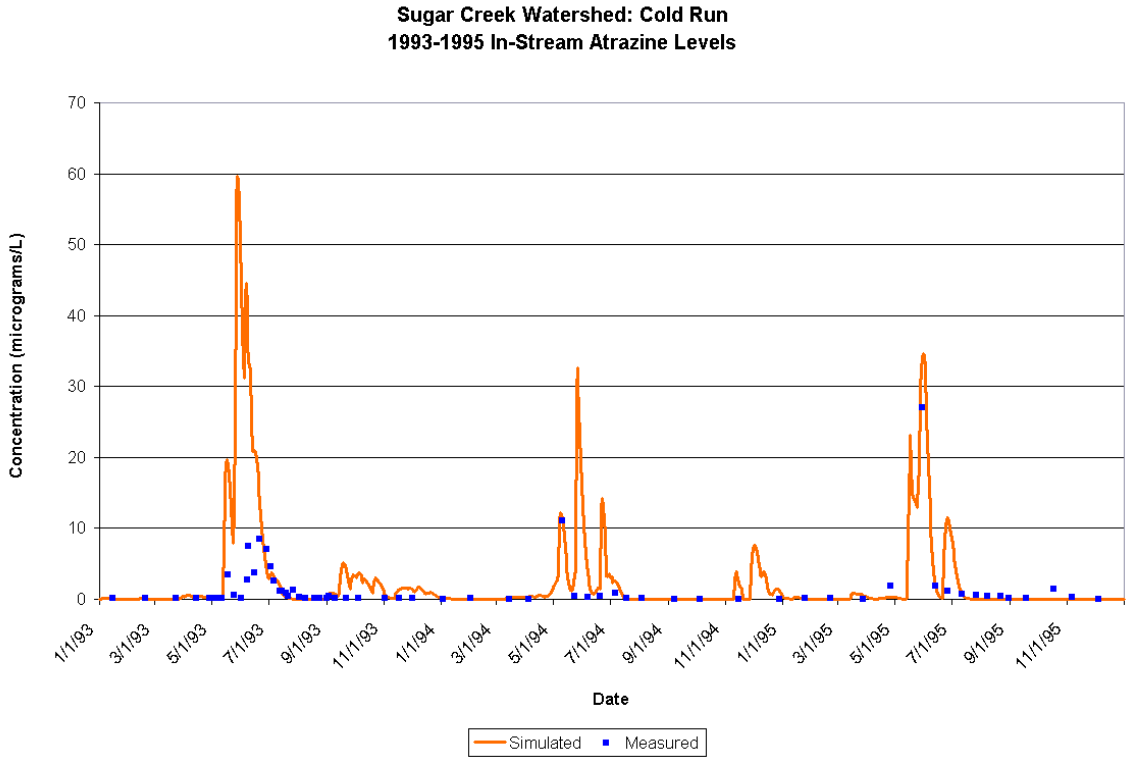


Figure 14: Cold simulation atrazine concentrations for 1993-1995

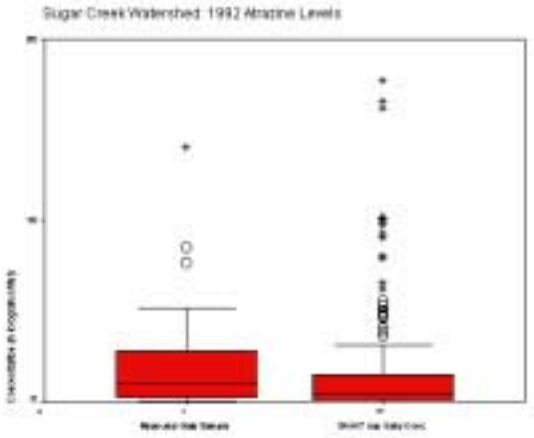


Figure 15: 1992 atrazine results

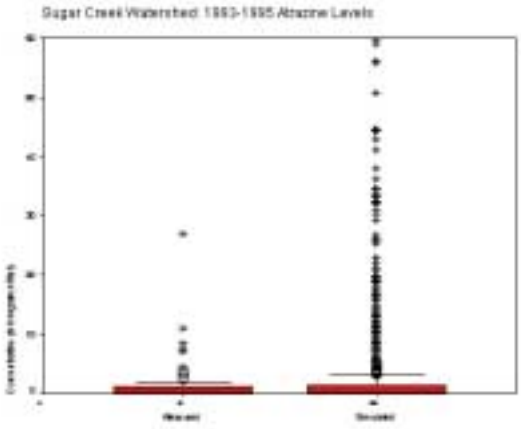


Figure 16: 1993-1995 atrazine results

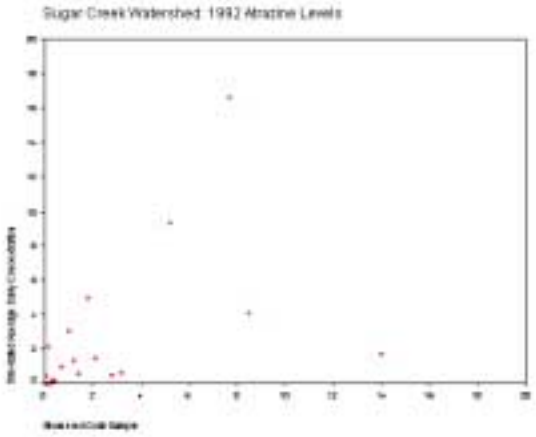


Figure 17: 1992 atrazine results

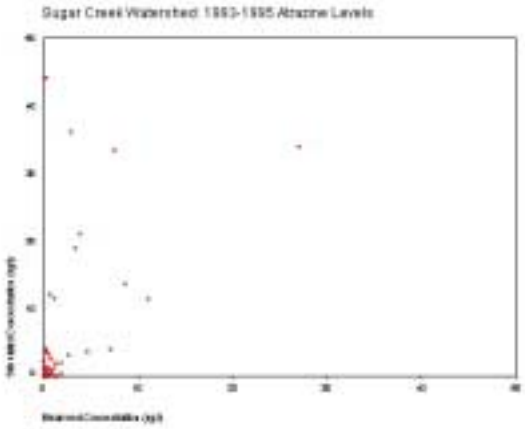


Figure 18: 1993-1995 atrazine results

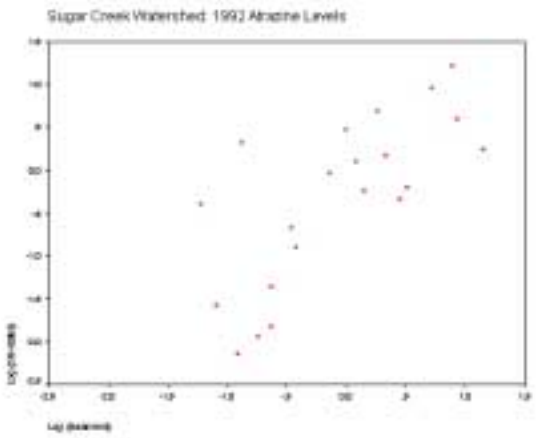


Figure 19: Log-transformed 1992 atrazine results

Figure 20: Log-transformed 1993-1995 atrazine results

Metolachlor

Cold simulation metolachlor concentrations for 1992 are graphed with measured grab sample concentrations in Figure 21. Figure 22 displays the same data for 1993-1995.

Table 4 summarizes the R^2 values computed from a regression analysis performed on the pairs of measured metolachlor grab samples and SWAT cold simulation metolachlor results. In addition to the time series graphs, box plots and scatter plots for the measured and cold simulation atrazine concentrations are provided in Figures 23-28.

Table 4: Statistics for Cold Simulation-Metolachlor

Simulation:	**R^2**	
	conc	**log(conc)**
1992 metolachlor, before stream flow calibration	.51	.47
1993-95 metolachlor, before stream flow calibration	.33	.37

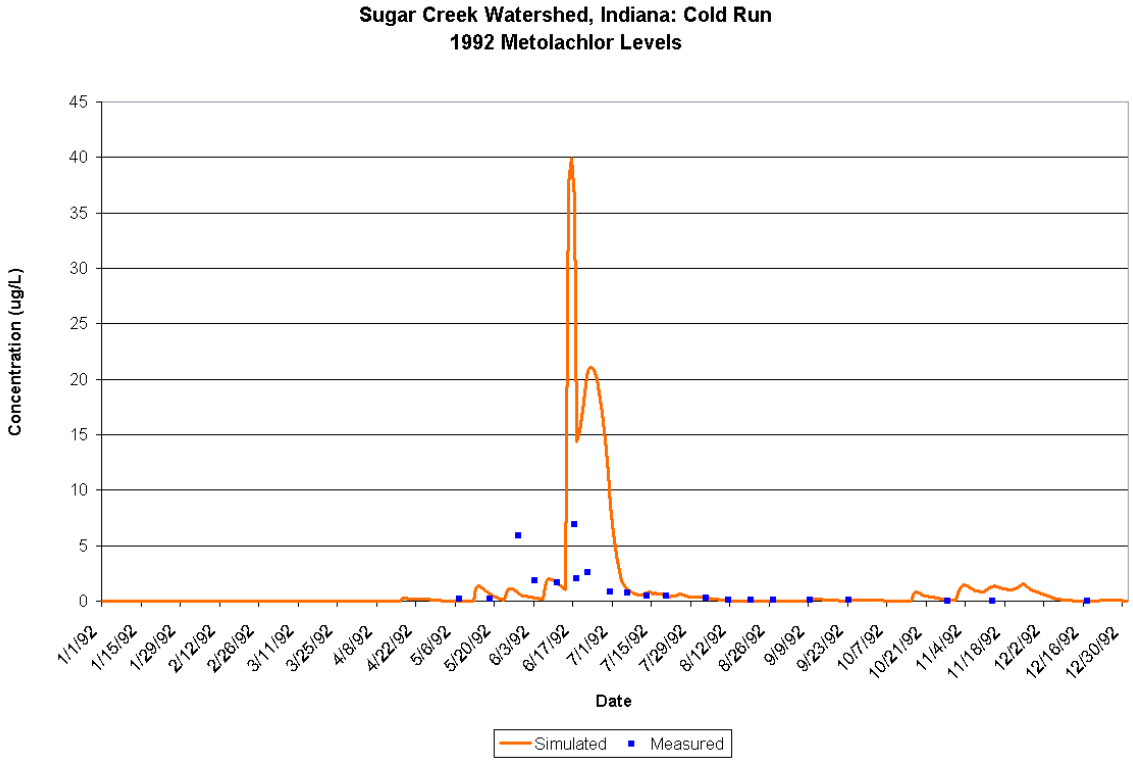


Figure 21: Cold simulation metolachlor concentrations for 1992

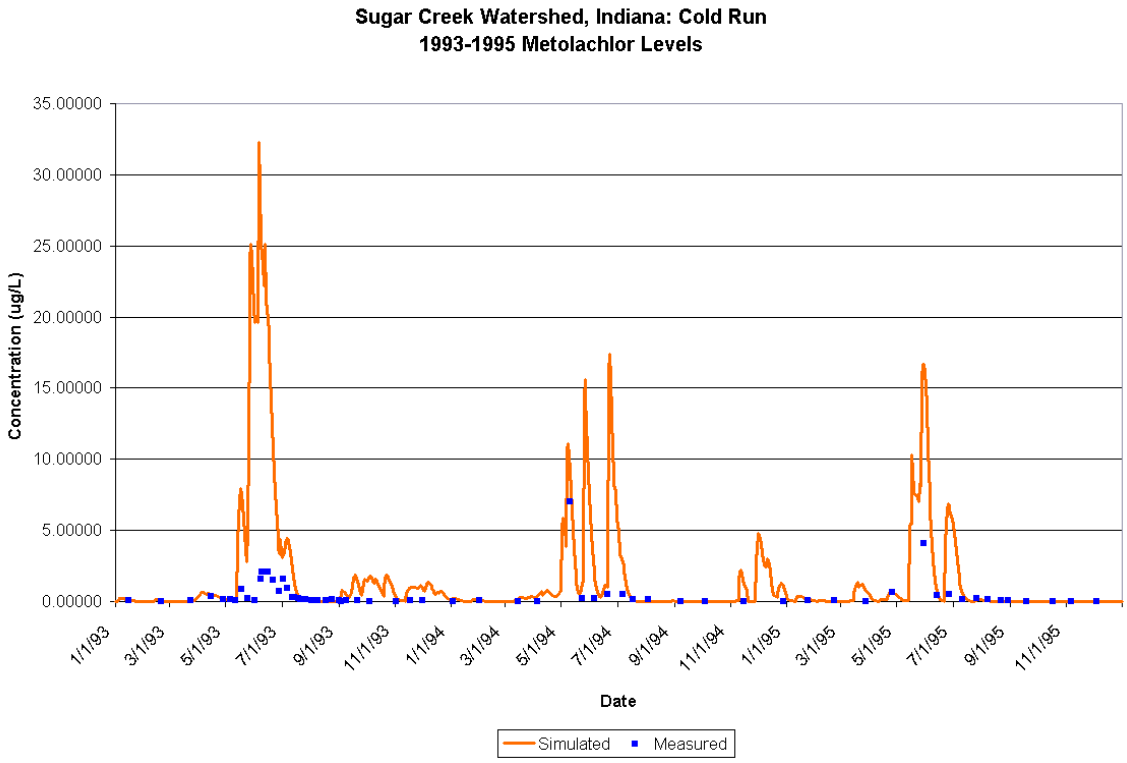


Figure 22: Cold simulation metolachlor concentrations for 1993-1995

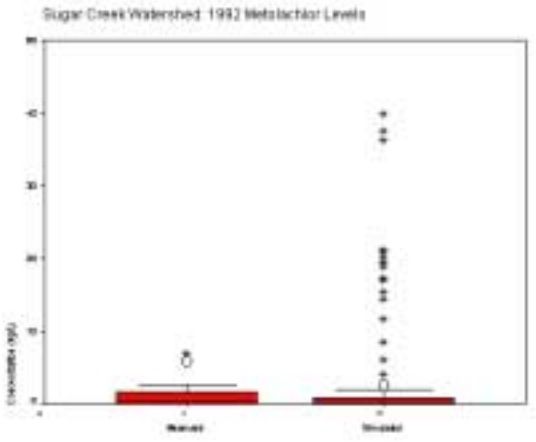


Figure 23: 1992 metolachlor results

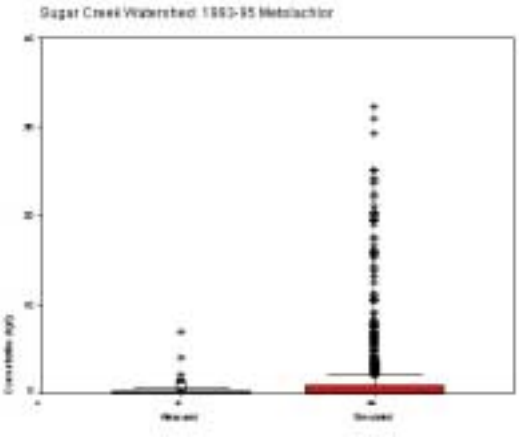


Figure 24: 1993-1995 metolachlor results

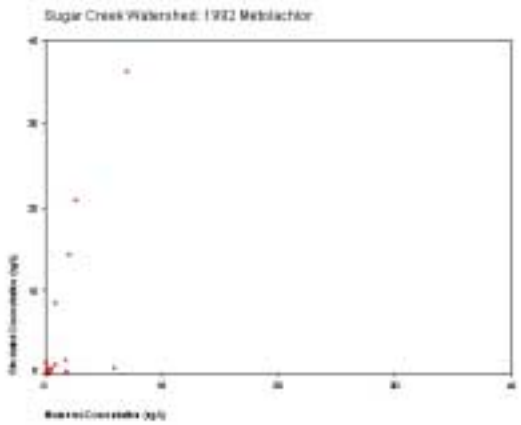


Figure 25: 1992 metolachlor results

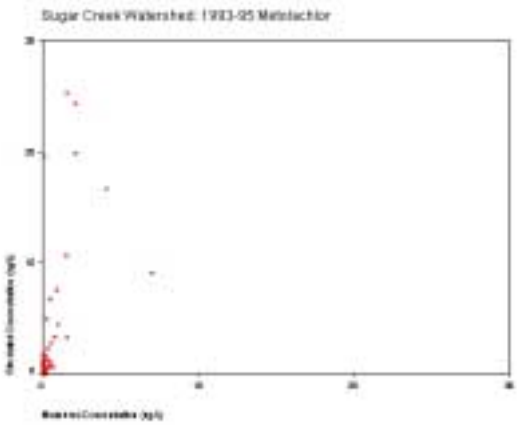


Figure 26: 1993-1995 metolachlor results

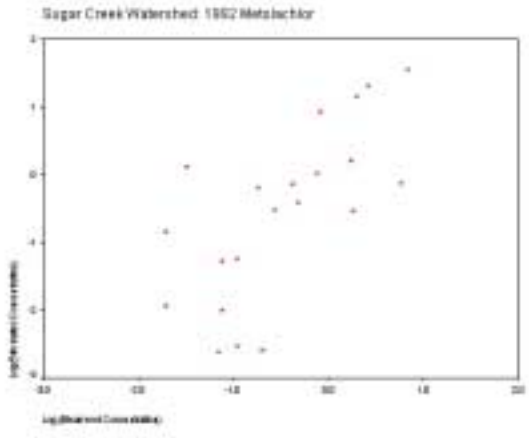


Figure 27: Log-transformed 1992 metolachlor results

Figure 28: Log-transformed 1993-1995 metolachlor results

Trifluralin

Cold simulation trifluralin concentrations for 1992 are graphed with measured grab sample concentrations in Figure 29. Figure 30 displays the same data for 1993-1995.

Table 5 summarizes the R^2 values computed from a regression analysis performed on the pairs of measured trifluralin grab samples and SWAT cold simulation trifluralin results. In addition to the time series graphs, box plots and scatter plots for the measured and cold simulation trifluralin concentrations are provided in Figures 31-36.

Table 5: Statistics for Cold Simulation-Trifluralin

Simulation:	**R^2**	
	conc	**log(conc)**
1992 trifluralin, before stream flow calibration	.88	.08
1993-95 trifluralin, before stream flow calibration	.07	.08

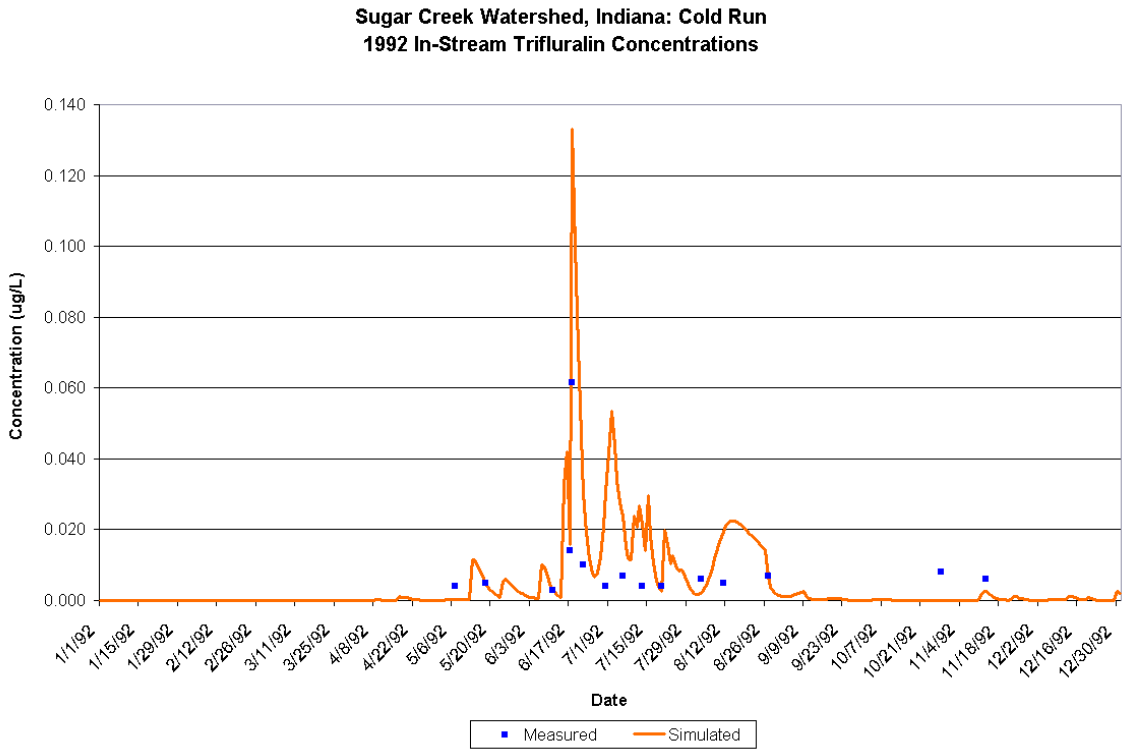


Figure 29: Cold simulation trifluralin concentrations for 1992

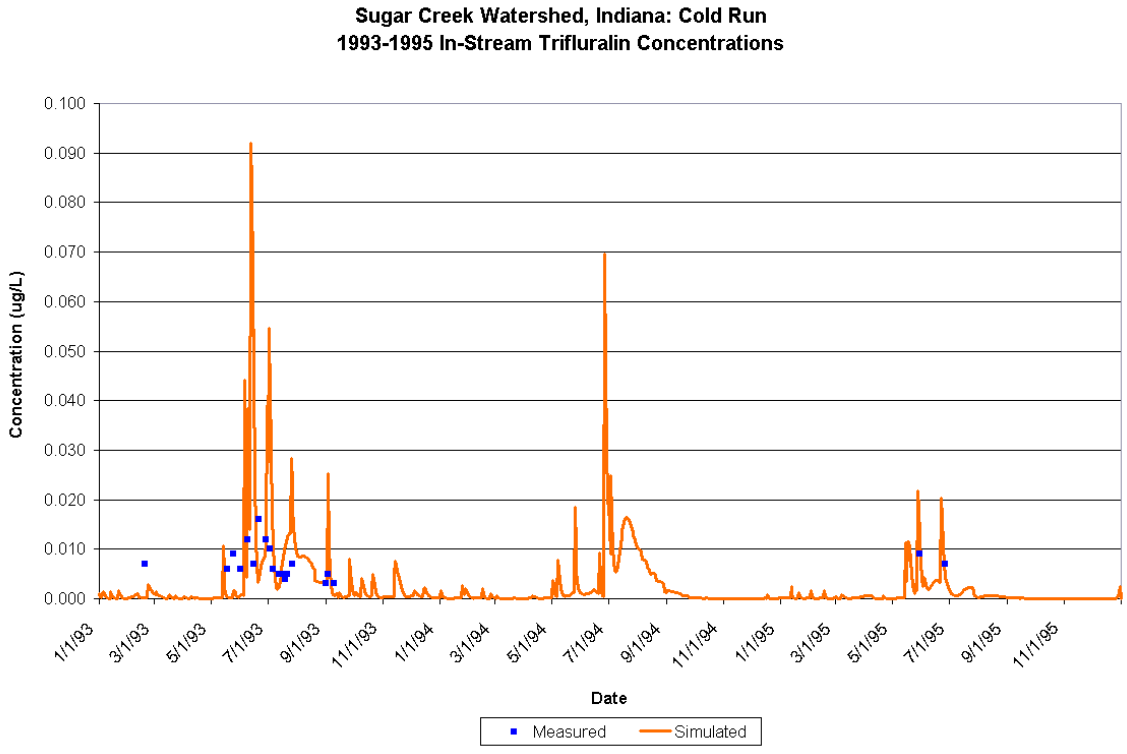


Figure 30: Cold simulation trifluralin concentrations for 1993-1995

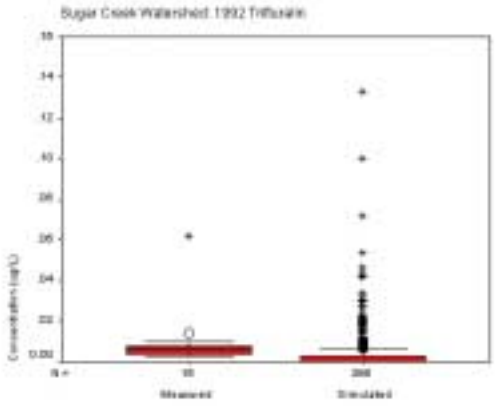


Figure 31: 1992 trifluralin results

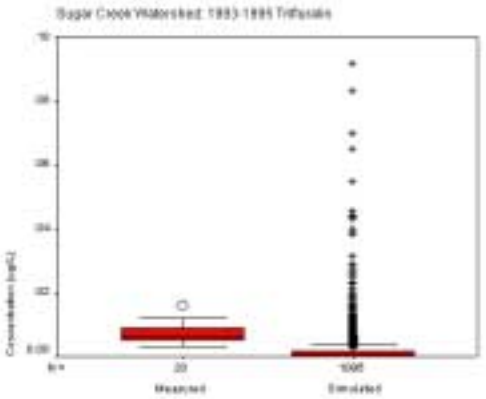


Figure 32: 1993-1995 trifluralin results

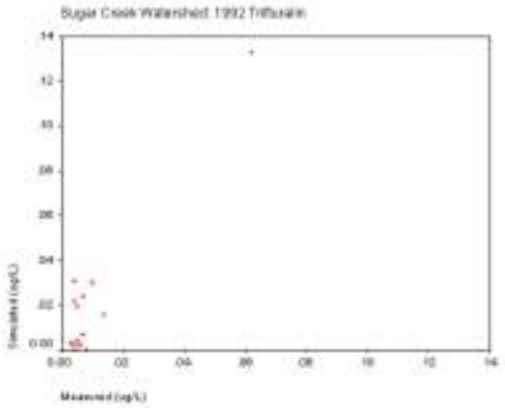


Figure 33: 1992 trifluralin results

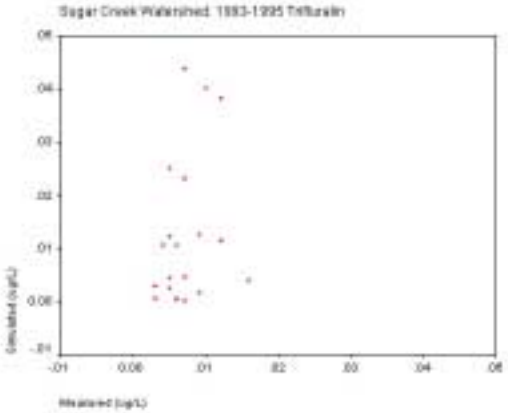


Figure 34: 1993-1995 trifluralin results

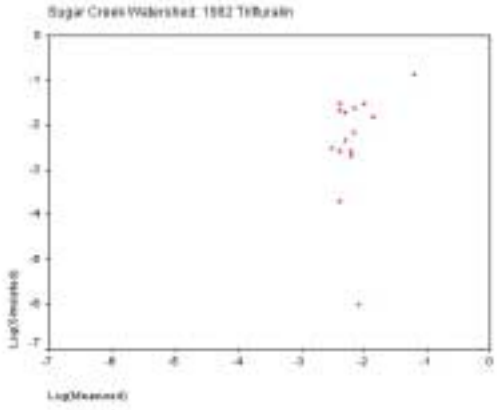


Figure 35: Log-transformed 1992 trifluralin results

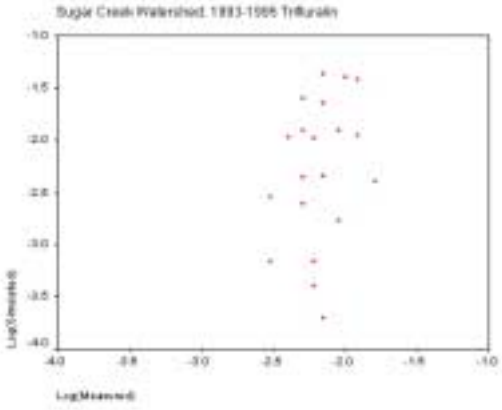


Figure 36: Log-transformed 1993-1995 trifluralin results

Part 2: Calibration—Hydrology

SWAT incorporates algorithms that simulate the physical processes governing the movement of water, nutrients and pesticides within a watershed. One of the principles guiding model development is the use of inputs that are physically based. As noted by Santhi et al (2001), SWAT is not a parametric model with a formal optimization procedure. Instead, variables such as the SCS runoff curve number that are not well defined physically or inputs whose values are assigned with a significant degree of uncertainty may be adjusted to provide a better fit.

Calibration of the water balance in the Sugar Creek watershed was performed in two steps. First, the long-term water balance was calibrated to match up total basin water yield. Second, daily measured and simulated flow data was compared for 1992, the period of record used for model calibration. The focus of the second step was to match the measured and simulated hydrograph shapes.

Long-term water balance

When calibrating hydrology in a watershed, information about the long-term water balance (20-30 year average) is used to ensure that the model calibration encompasses periods with drier than average and wetter than average climatic conditions. This is done to ensure that the model results are not biased towards one type of climatic condition. Another component of the water balance calibration is verifying that the fractions of groundwater and surface water contribution to streamflow are correct. Because surface runoff is the primary pathway by which pollutants enter the stream network, calibrating the model to properly reflect the partitioning between surface runoff and baseflow is very important.

Daily average flow rates were obtained from the U.S.G.S. gaging station at the outlet of the Sugar Creek watershed for October, 1967-October, 1999. Average annual streamflow for the period 10/1967-12/1996 was 2.9455 m^3/s. Expressed as a depth of water over the watershed area, the average annual water yield is 382 mm/yr. A baseflow separation program (Arnold and Allen, 1999) was used to filter the measured daily flow and the baseflow contribution to streamflow was determined to be between 41-57% of total streamflow.

A 30 year simulation was set up to calibrate long-term water yield. The long-term water yield calibration focuses on getting the total water yield and the surface runoff/baseflow values in the general neighborhood of the values determined from observed data. For this component of the simulation, the average annual water balance for the HRUs as well as for the entire watershed were used to make decisions to adjust parameters.

Four parameters were modified or adjusted during the long-term water yield calibration: SCS curve number for moisture condition II (CN2, *.mgt*), maximum canopy water storage (CANMX, *.hru*), available water capacity (SOL_AWC, *.sol*), and saturated hydraulic conductivity (SOL_K, *.sol*). Appendix B lists the modifications made during the calibration process. The long-term water balance for the watershed is listed in Table 6.

Table 6: 1968-1996 Simulated Average Annual Watershed Values: Water Balance for Land Areas

precipitation	1060.5 mm/yr
snow fall	88.02 mm/yr
snow melt	80.79 mm/yr
sublimation	0.64 mm/yr
evapotranspiration	695.5 mm/yr
potential evapotranspiration	1135.5 mm/yr
percolation out of soil profile	104.39 mm/yr
total aquifer recharge	105.70 mm/yr
deep aquifer recharge	4.29 mm/yr
"revap" (shallow aquifer => soil/plants)	8.12 mm/yr
surface runoff contribution to stream	176.79 mm/yr
lateral soil flow contribution to stream	0.26 mm/yr
tile flow contribution to stream	94.96 mm/yr
groundwater contribution to stream	95.77 mm/yr
total water yield	366.02 mm/yr

Daily Flow Calibration

Once the long-term calibration was completed, efforts focused on matching the 1992 simulated daily flow hydrograph to the measured daily flow values recorded for the USGS stream flow gage. Total water yield (expressed as m^3/yr) was calibrated until simulated total water yield for 1992 was within 1% of the measured value. Daily streamflow was calibrated to reach a daily coefficient of determination (R^2) of 0.6 and a Nash-Suttcliffe simulation efficiency (E_{NS}) of 0.5.

Six parameters were modified or adjusted during the daily flow calibration: calibration coefficients for Muskingum channel routing (MSK_X, MSK_CO1, and MSK_CO2, *.bsn*), Manning's n value for the main channels (CH_N, *.rte*), maximum canopy water storage (CANMX, *.hru*), and the soil evaporation compensation factor (ESCO, *.bsn*). Appendix B lists the modifications made during the calibration process.

Table 7 summarizes the statistical data for the measured flow and SWAT calibrated hydrology simulation flow results. Calibrated daily stream flow for 1992 is graphed with measured streamflow in Figure 37. Figure 38 displays the same data for 1993-1995. In addition to the time series graphs, box plots and scatter plots for the measured and calibrated flow are provided in Figures 39-42. Additional statistical analysis is included in Appendix C.

Table 7: Statistics for Daily Stream Flow

	1992 Daily Stream Flow		1993-1995 Daily Stream Flow	
Statistic:	**Measured**	**SWAT Calibration**	**Measured**	**SWAT Validation**
Mean	3.0091	3.0265	2.8562	2.7332
Std error of mean	.2492	.2763	.1689	.1633
Median	1.5574	1.2400	1.1600	1.0700
Mode	.40	1.02	.31	.00
Standard deviation	4.7672	5.2855	5.5919	5.4028
Variance	22.7264	27.9364	31.2692	29.1901
Range	36.81	36.35	54.59	58.20
Minimum	.28	.25	.06	.00
Maximum	37.10	36.60	54.65	58.20
D_v	-0.5783 difference in water yield		4.3092 difference in water yield	
R^2		0.59		0.75
Nash-Sutcliffe, E_{NS}		0.47		0.74
DG		0.50		0.71

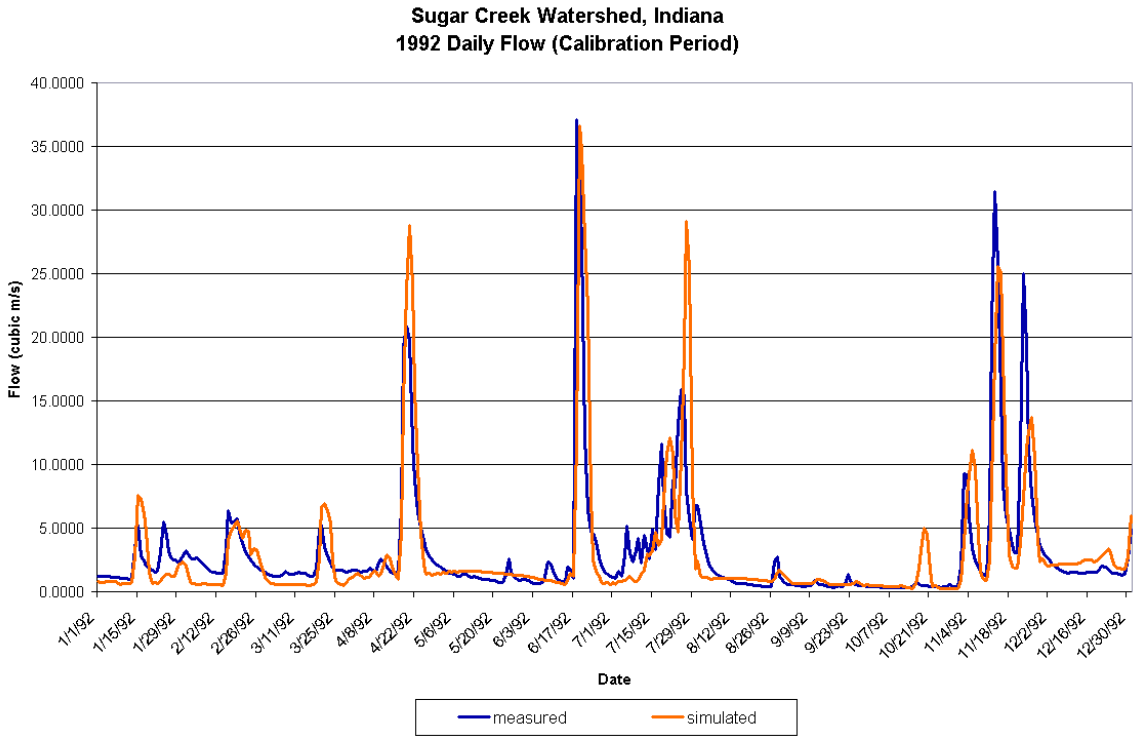


Figure 37: 1992 daily stream flow calibration results

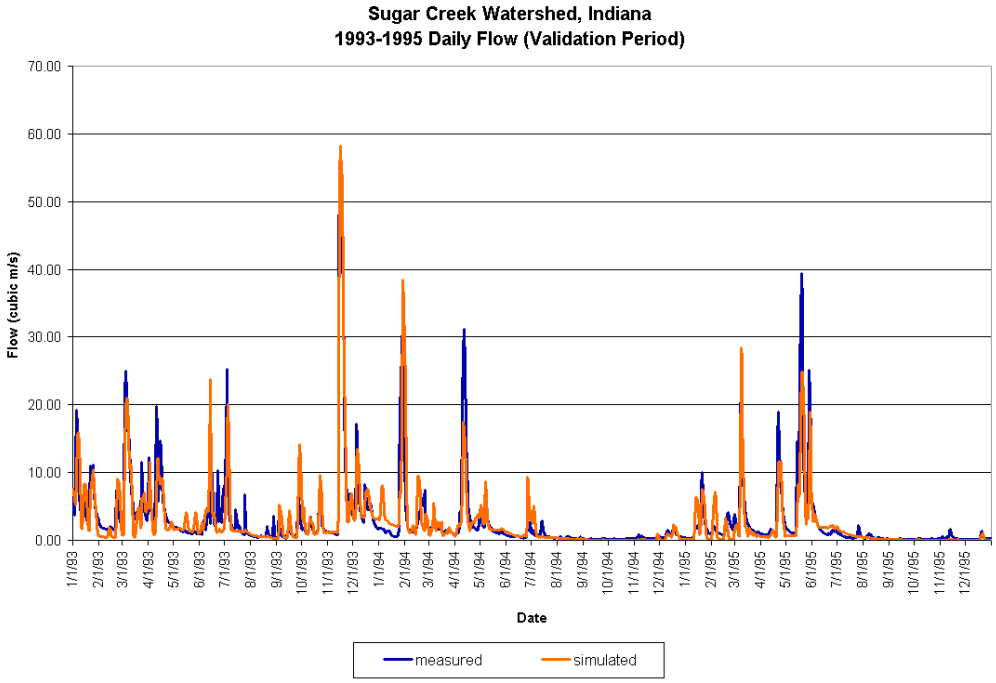


Figure 38: 1993-1995 daily stream flow validation results

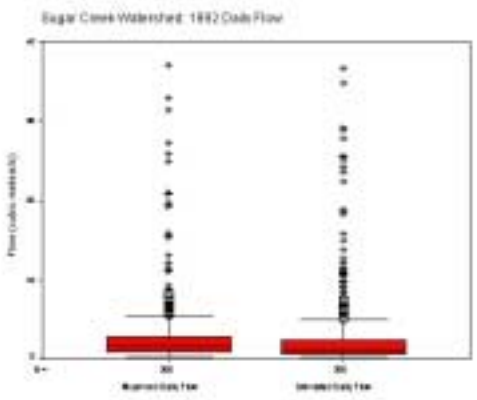


Figure 39: Box plot-1992 calibration results

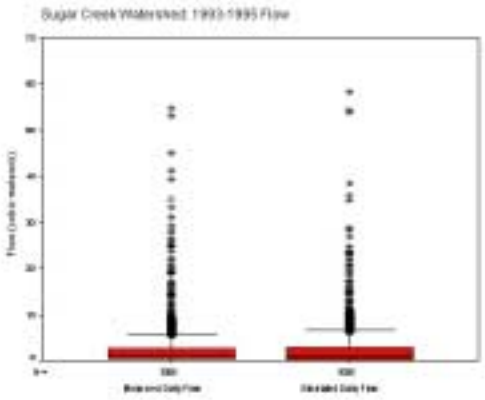


Figure 40: Box plot-1993-1995 validation results

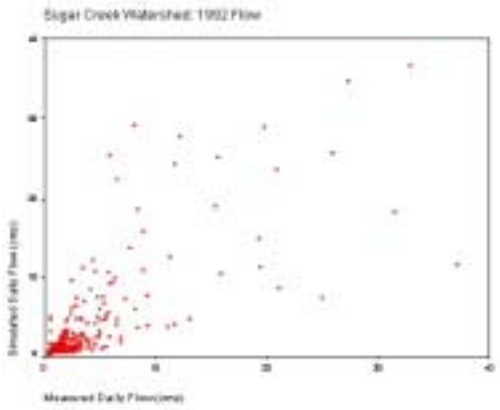


Figure 41: Scatter plot-1992 calibration results

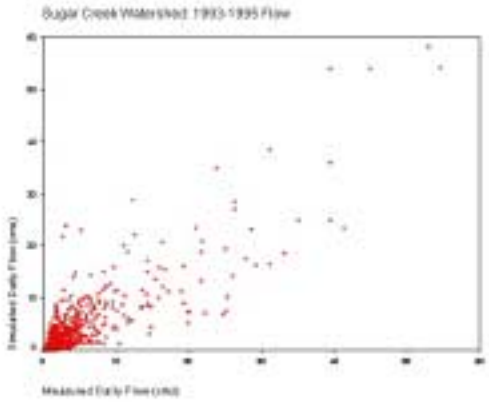


Figure 42: Scatter plot-1993-1995 validation results

Atrazine

Simulation atrazine concentrations after stream flow calibration are graphed with measured grab bag sample concentrations in Figure 43. Figure 44 displays the same data for 1993-1995.

Table 8 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated atrazine concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated atrazine concentrations are provided in Figures 45-50.

Table 8: Statistics for Atrazine

Simulation:	**R^2**	
	conc	**log(conc)**
1992 atrazine, after stream flow calibration	.21	.45
1993-95 atrazine, after stream flow calibration	.41	.32

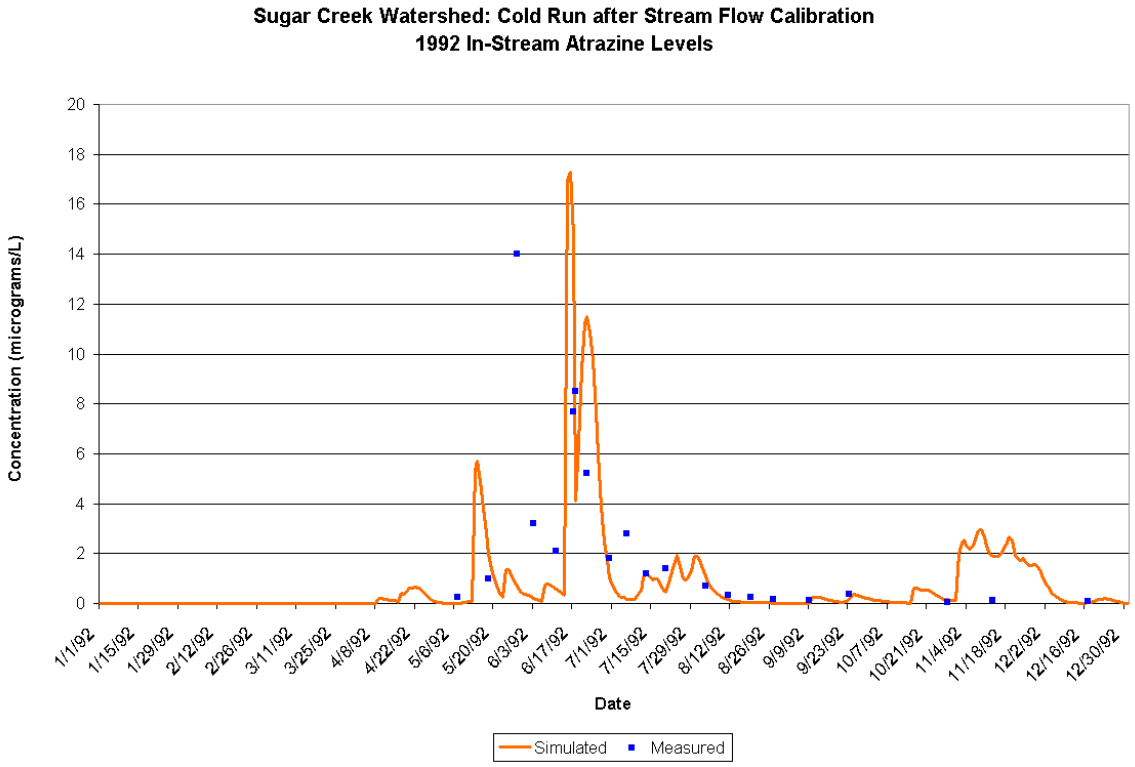


Figure 43: 1992 atrazine concentrations predicted after streamflow calibration

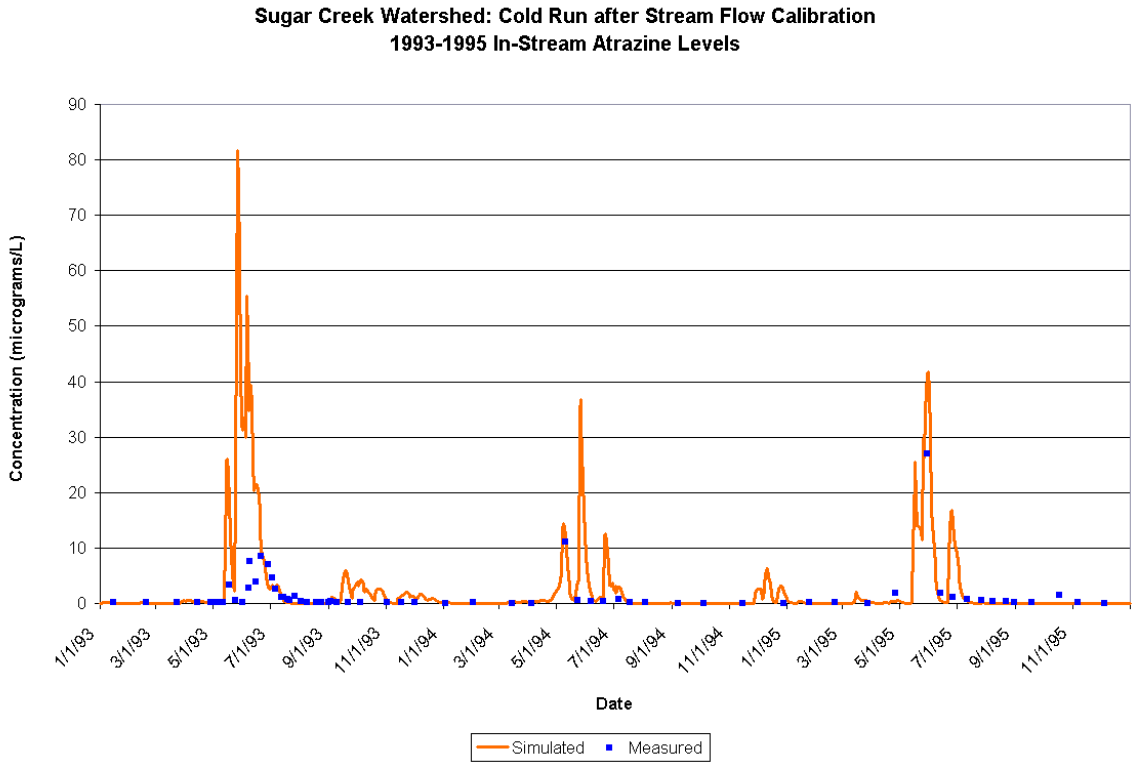


Figure 44: 1993-1995 atrazine concentration predicted after streamflow calibration

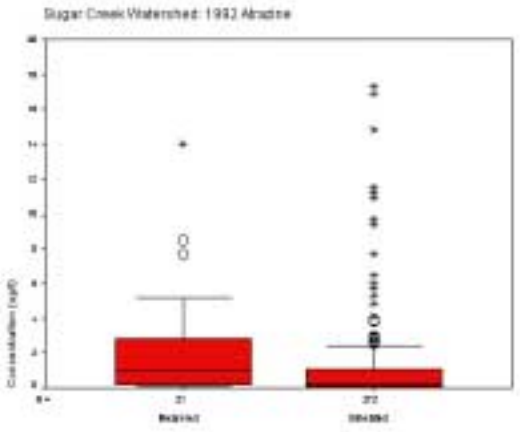


Figure 45: 1992 atrazine results

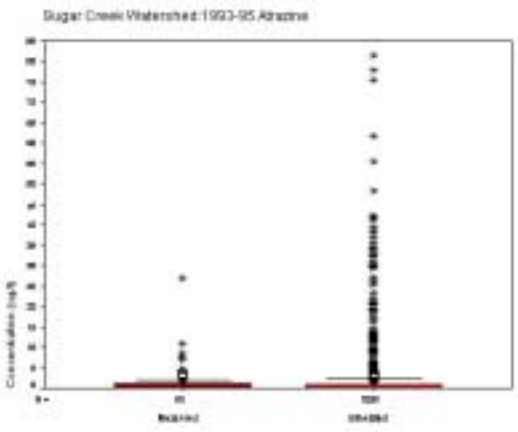


Figure 46: 1993-1995 atrazine results

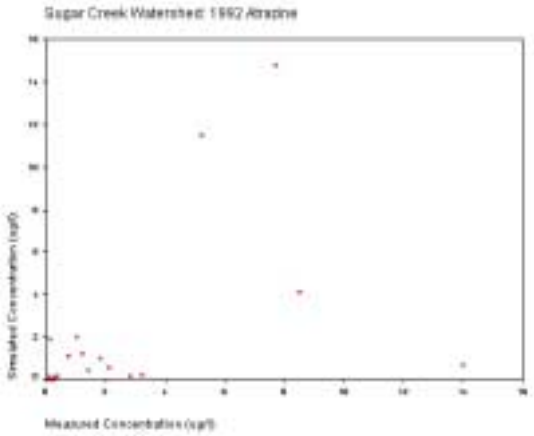


Figure 47: 1992 atrazine results

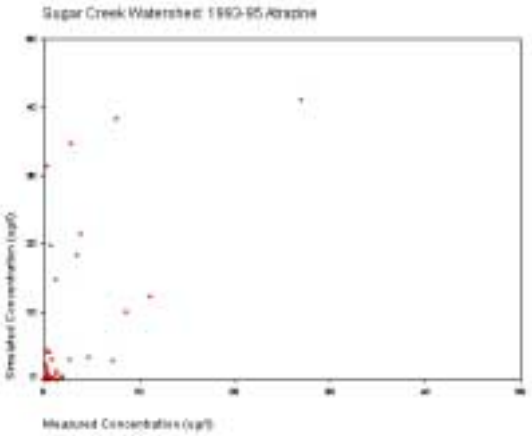


Figure 48: 1993-1995 atrazine results

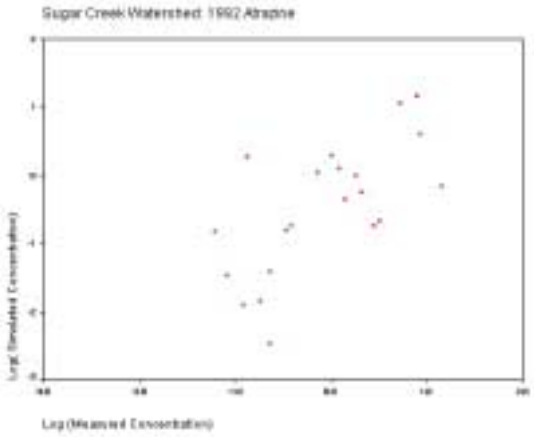


Figure 49: Log-transformed 1992 atrazine results

Figure 50: Log-transformed 1993-1995 atrazine results

Metolachlor

Simulation metolachlor concentrations after stream flow calibration are graphed with measured grab bag sample concentrations in Figure 51. Figure 52 displays the same data for 1993-1995.

Table 9 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated metolachlor concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated metolachlor concentrations are provided in Figures 53-58.

Table 9: Statistics for Metolachlor

Simulation:	**R^2**	
	conc	**log(conc)**
1992 metolachlor, after stream flow calibration	.46	.40
1993-95 metolachlor, after stream flow calibration	.32	.40

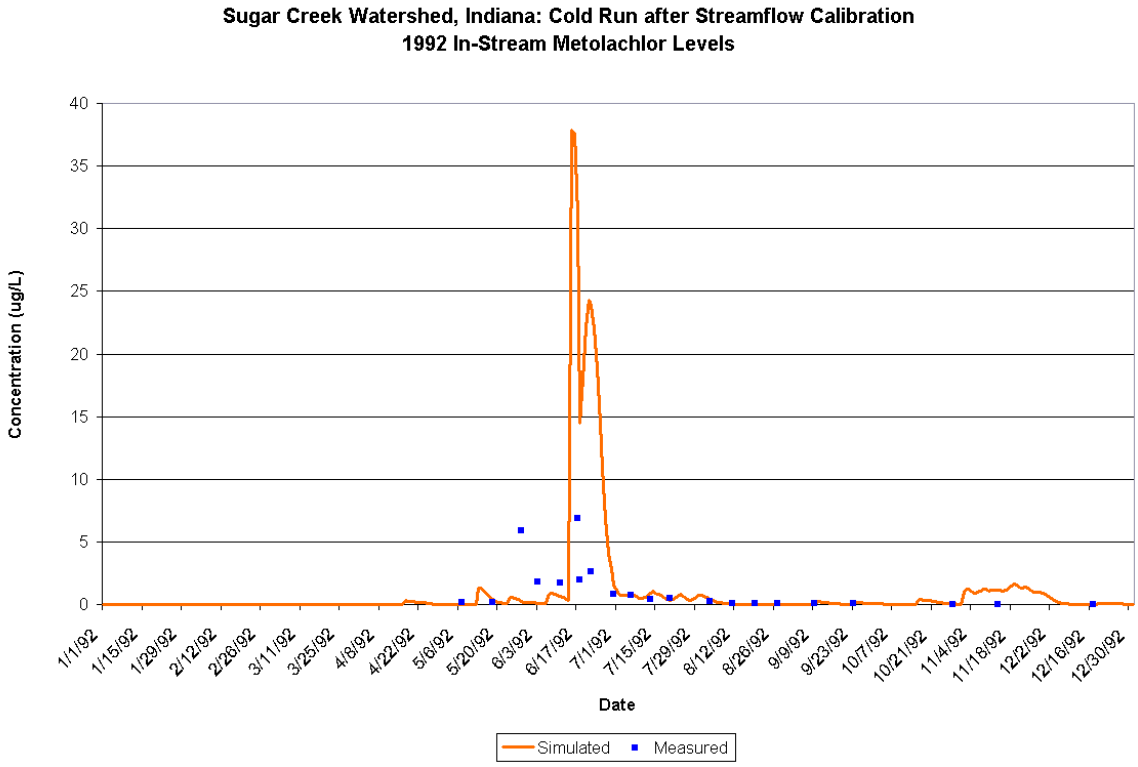


Figure 51: 1992 metolachlor concentration predicted after streamflow calibration

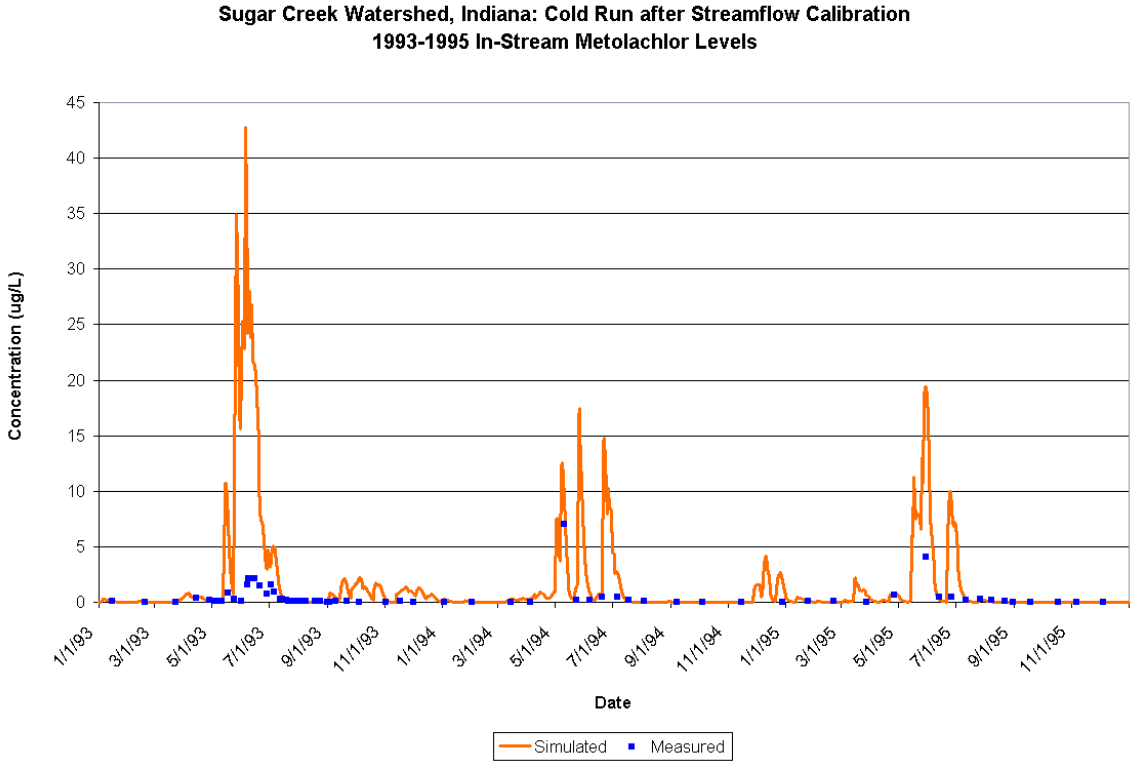


Figure 52: 1993-1995 metolachlor concentration predicted after streamflow calibration

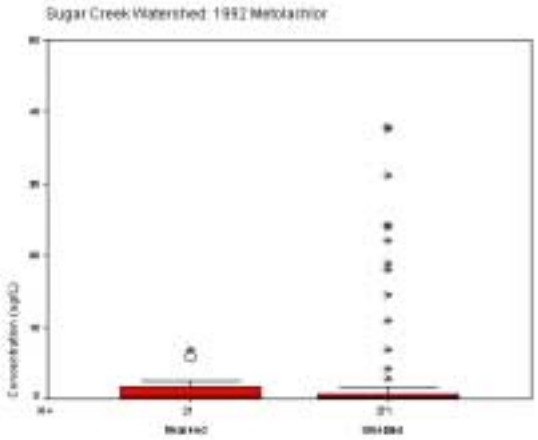


Figure 53: 1992 metolachlor results

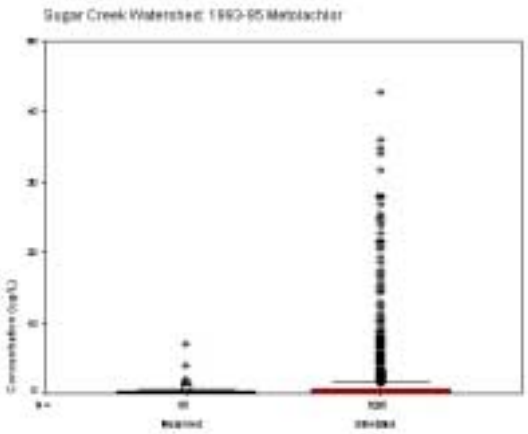


Figure 54: 1993-1995 metolachlor results

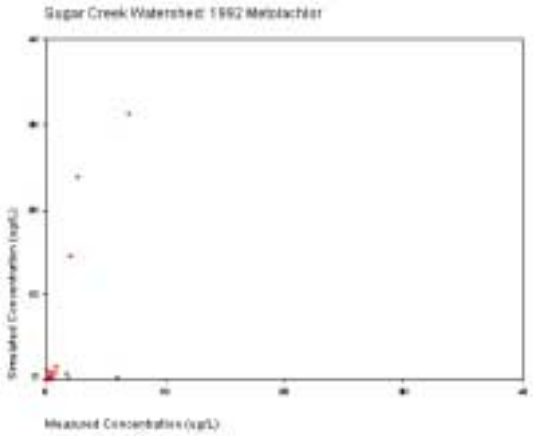


Figure 55: 1992 metolachlor results

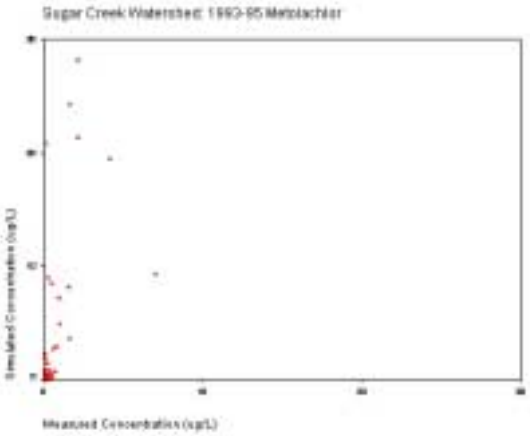


Figure 56: 1993-1995 metolachlor results

Figure 57: Log-transformed 1992 metolachlor results

Figure 58: Log-transformed 1993-1995 metolachlor results

Trifluralin

Simulation trifluralin concentrations after stream flow calibration are graphed with measured grab bag sample concentrations in Figure 59. Figure 60 displays the same data for 1993-1995.

Table 10 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated trifluralin concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated trifluralin concentrations are provided in Figures 61-66.

Table 10: Statistics for Trifluralin

Simulation:	**R^2**	
	conc	**log(conc)**
1992 trifluralin, after stream flow calibration	.48	.06
1993-95 trifluralin, after stream flow calibration	.01	.02

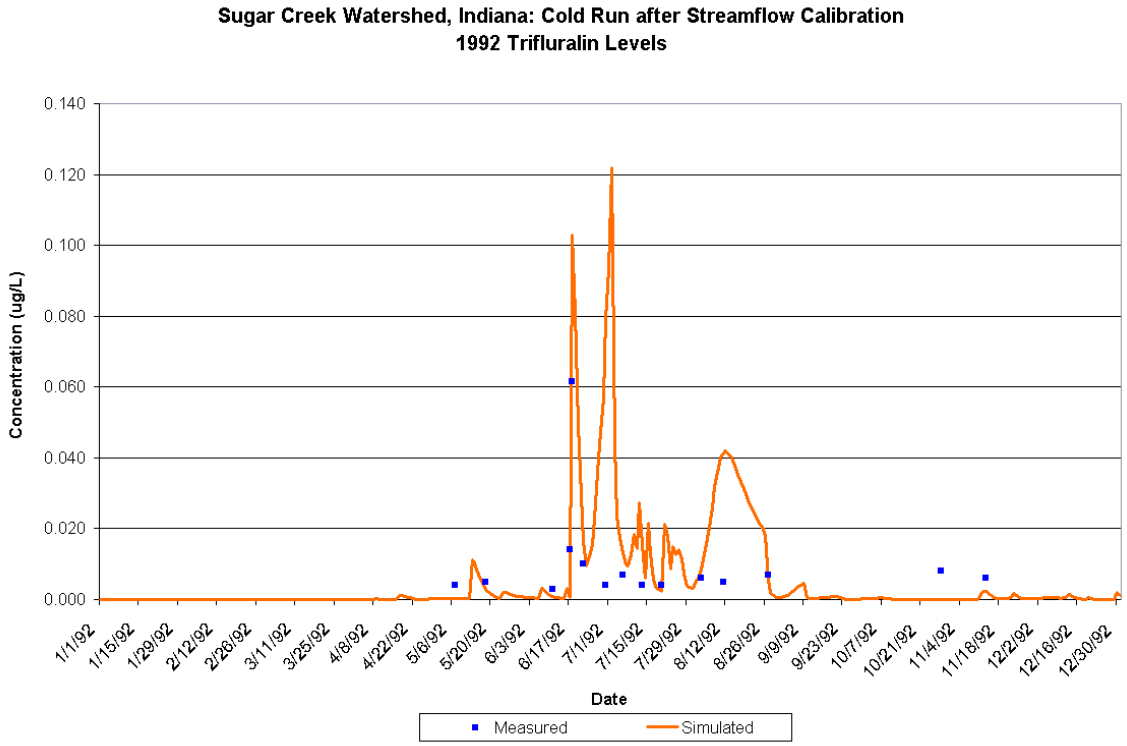


Figure 59: 1992 trifluralin concentration predicted after streamflow calibration

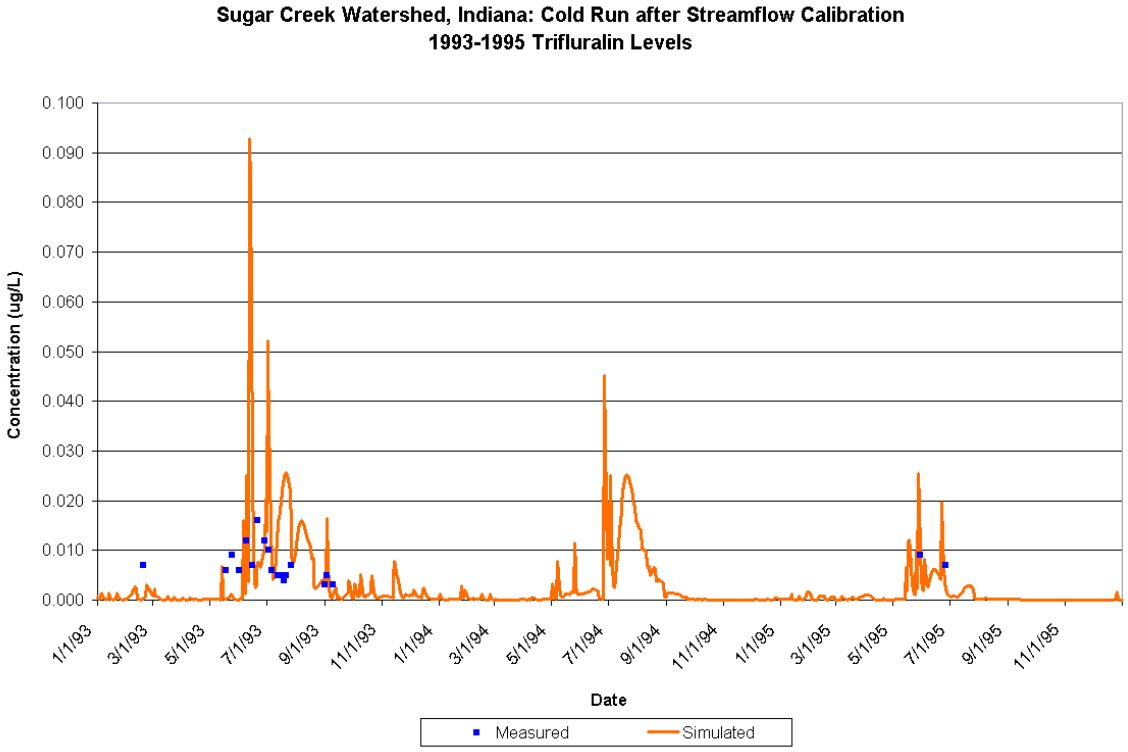


Figure 60: 1993-1995 trifluralin concentration predicted after streamflow calibration

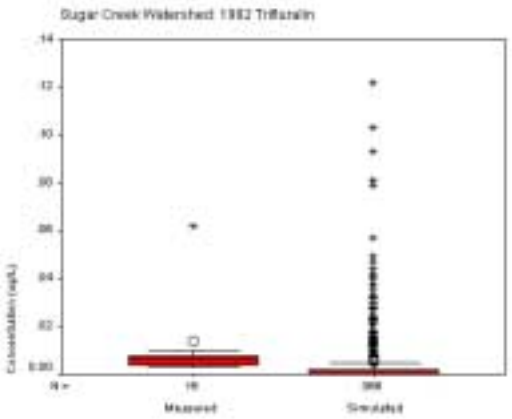


Figure 61: 1992 trifluralin results

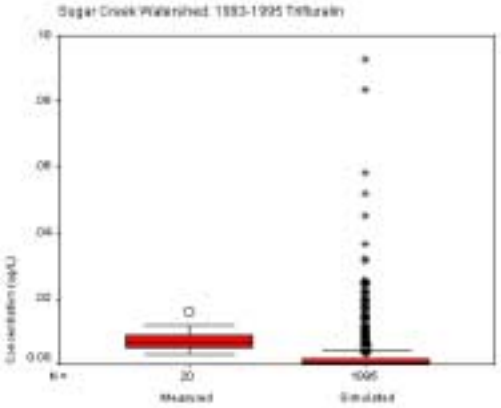


Figure 62: 1993-1995 trifluralin results

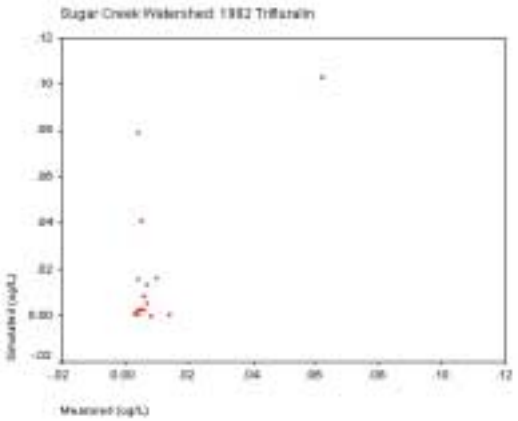


Figure 63: 1992 trifluralin results

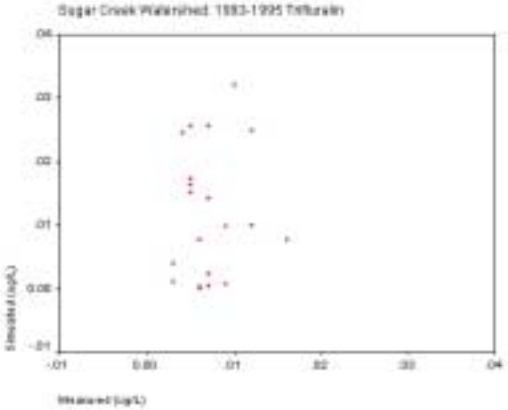


Figure 64: 1993-1995 trifluralin results

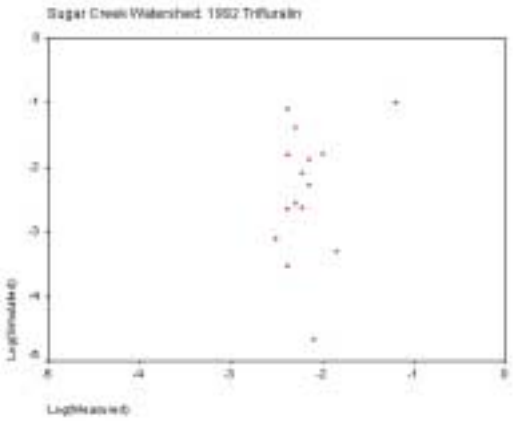


Figure 65: Log-transformed 1992 trifluralin results

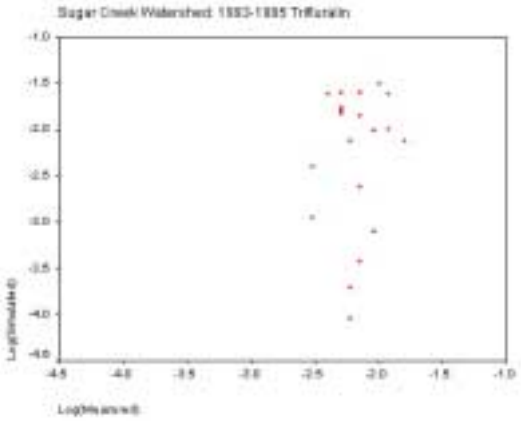


Figure 66: Log-transformed 1993-1995 trifluralin results

Part 3: Calibration—Atrazine

The calibration and validation periods used for hydrology were also used for atrazine. Calibration and validation was performed using atrazine concentrations sampled as part of the U.S.G.S. National Water Quality Assessment Program. Daily in-stream pesticide concentrations simulated by SWAT were compared to grab samples collected at the NAWQA water quality sampling station 39340085524601. Most of the measured concentrations used in the study were obtained from one grab sample (instantaneous measurement) collected on a given day. On a few occasions, more than one grab sample was collected on the same day. When this happened, the concentrations were averaged to obtain a single value for the day.

Looking at the cold run atrazine concentrations (Figure 67), the simulated concentration lagged measured peaks by 2 weeks. The pesticide application data incorporated a 10-day range of error and I attempted to shift the atrazine peak forward a few weeks by shifting the pesticide application dates forward, but was unable to change the timing of the simulated atrazine concentration peaks. Plotting the flow and atrazine data together shows that the high atrazine concentration recorded on May 28, 1992 doesn't coincide with a high flow period (Figure 68).

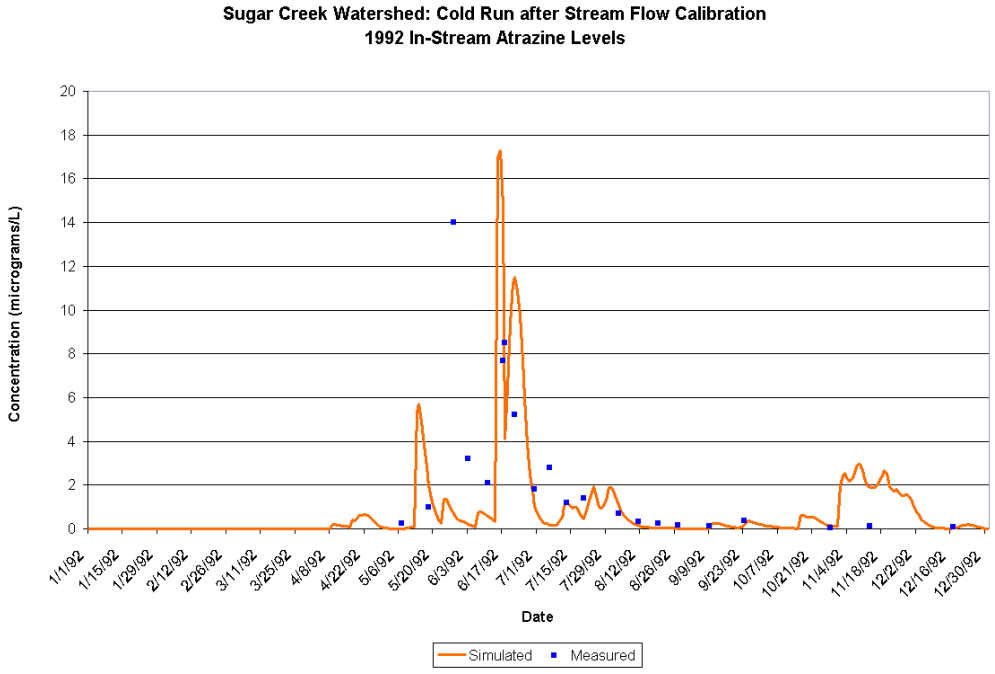


Figure 67: 1992 measured and simulated atrazine concentrations.

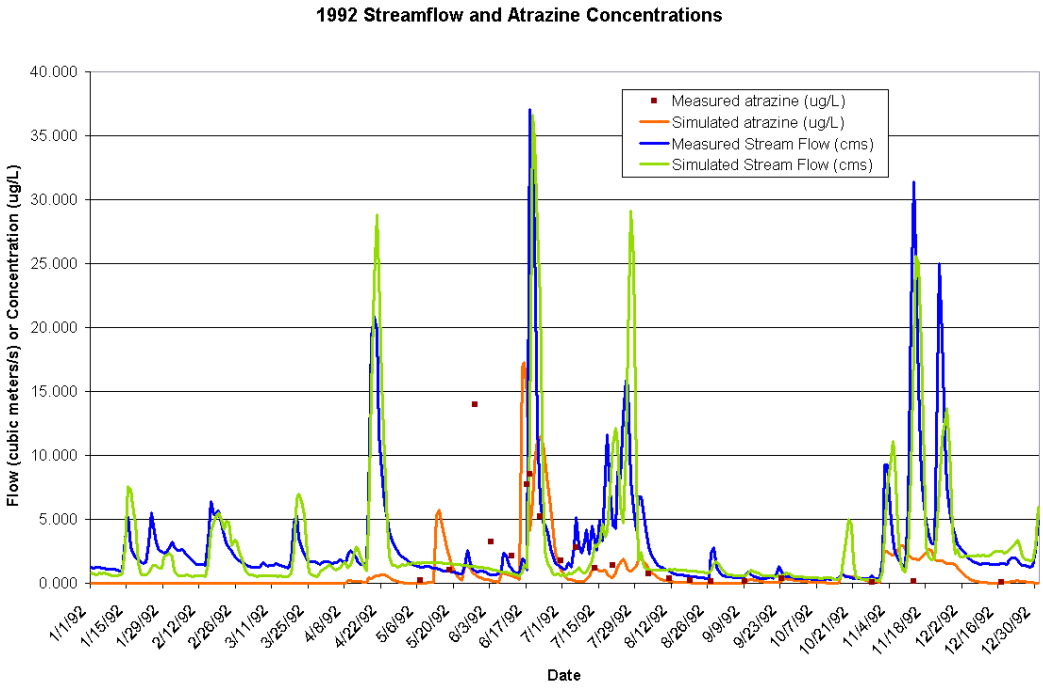


Figure 68: Comparison of flow and atrazine in 1992.

The increase in simulated atrazine concentration during November of 1992 was caused by atrazine moving through the soil profile and entering the stream via tile drainage. While tile drainage does not typically occur in the Sugar Creek watershed at this time of year, there was a considerable amount of rainfall during November in 1992 (177mm/7 inches). By setting the depth of the tile drains to a lower depth, the concentration of atrazine in the stream during November was lowered (more atrazine degraded with the longer residence time in the soil). However, it was decided that a tile drain depth deeper than 800 mm was unrealistic and the peaks in atrazine concentration during November were preferred.

The cold run SWAT atrazine concentrations were found to provide the best fit to measured concentrations. Mass balance values for 1992-1995 averaged over the entire watershed area are provided in Table 11.

Table 11: Atrazine mass balance and fate

Applied	316598 mg/ha/yr	
(excludes fraction lost during application)		
Decayed	295337 mg/ha/yr	93.28%
In Surface Runoff Entering Stream (Soluble)	12048 mg/ha/yr	3.81%
In Surface Runoff Entering Stream (Sorbed)	44 mg/ha/yr	0.01%
Leached out of Profile	744 mg/ha/yr	0.2%
In Lateral Flow/Tile Drainage Entering Stream	2717 mg/ha/yr	0.9%
Final Amount in Ground	22076 mg/ha	1.7%

Atrazine concentrations after calibration are graphed along with measured grab bag sample concentrations in Figure 69. Figure 70 displays the same data for 1993-1995.

Table 12 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated atrazine concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated atrazine concentrations are provided in Figures 71-76. Additional statistical analysis is included in Appendix C.

Table 12: Statistics for Atrazine

Simulation:	**R^2**	
	conc	**log(conc)**
1992 atrazine, after calibration	.21	.45
1993-95 atrazine, after calibration	.41	.32

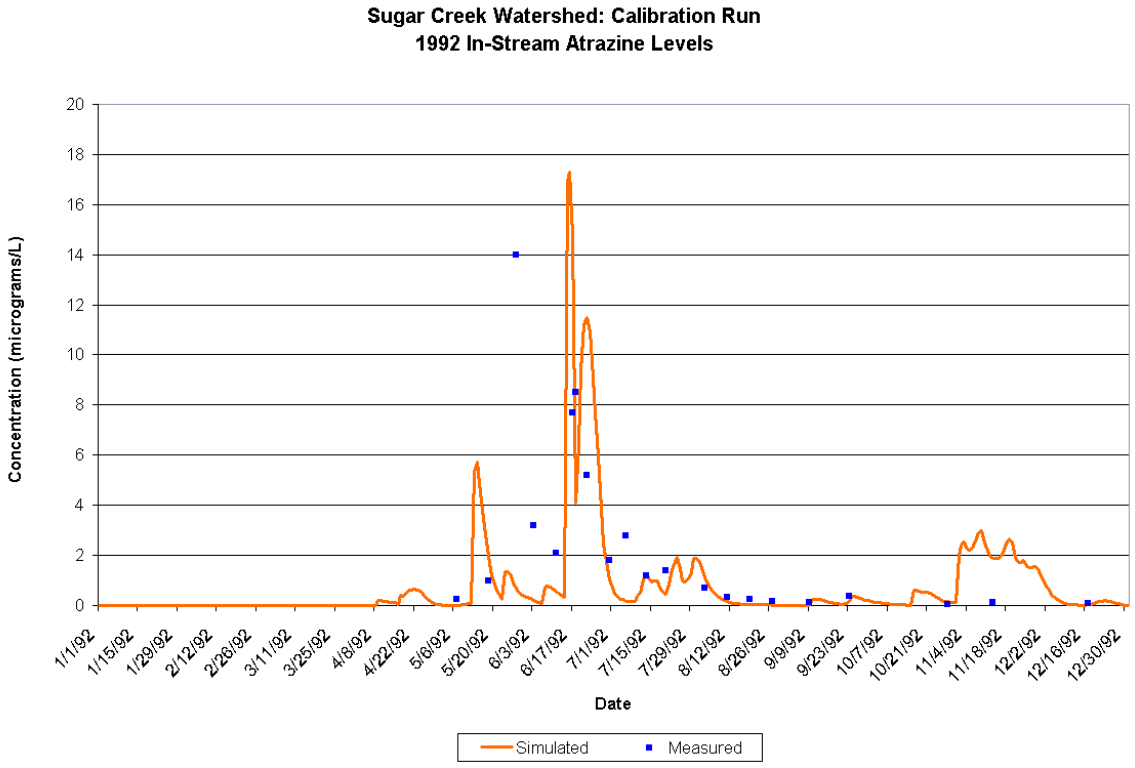


Figure 69: 1992 atrazine concentrations predicted after calibration

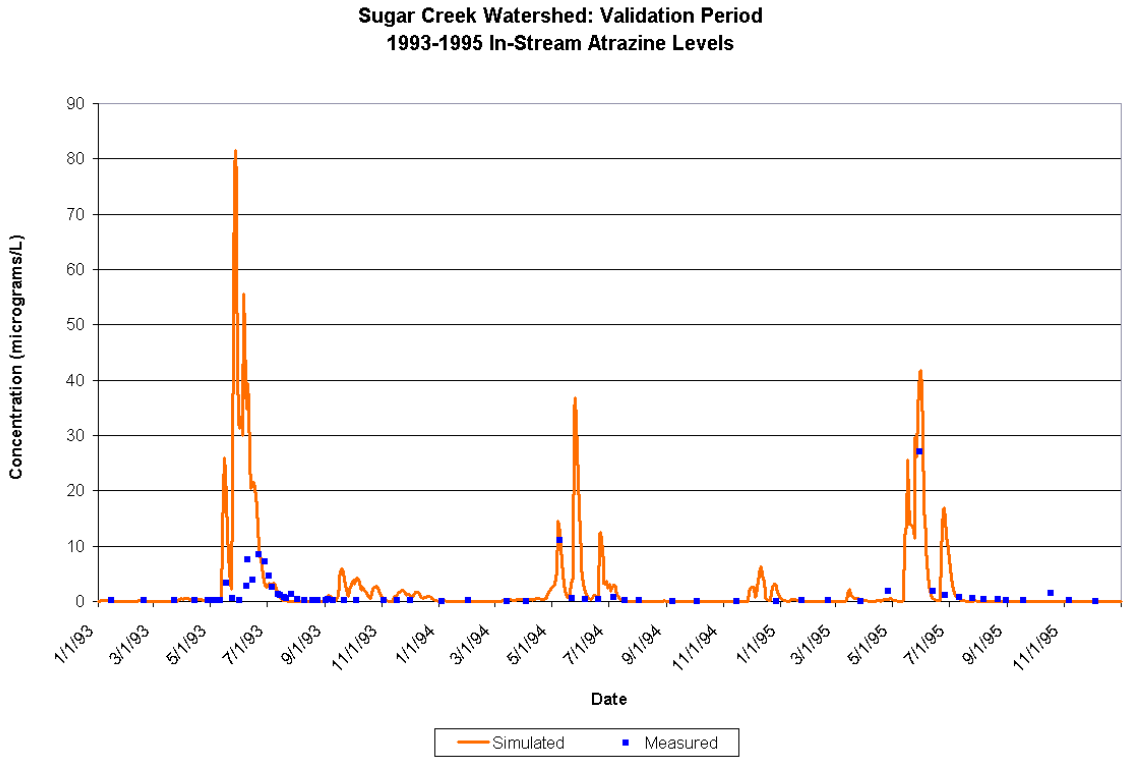


Figure 70: 1993-1995 atrazine concentration predicted after 1992 calibration

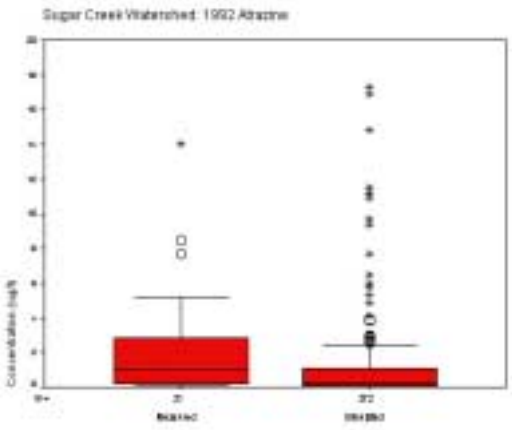


Figure 71: 1992 atrazine results

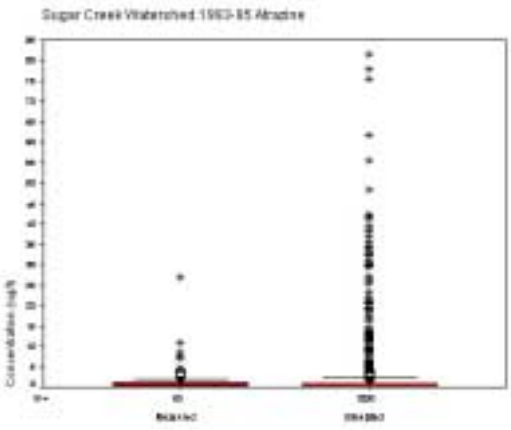


Figure 72: 1993-1995 atrazine results

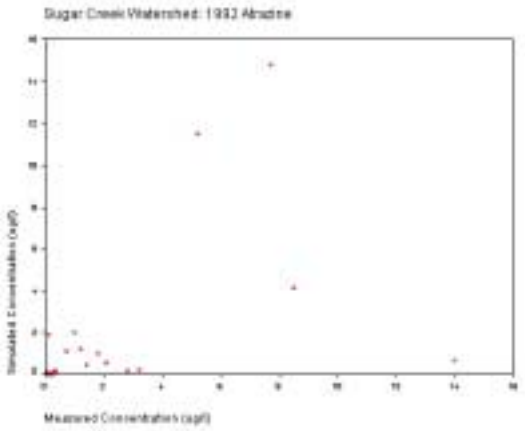


Figure 73: 1992 atrazine results

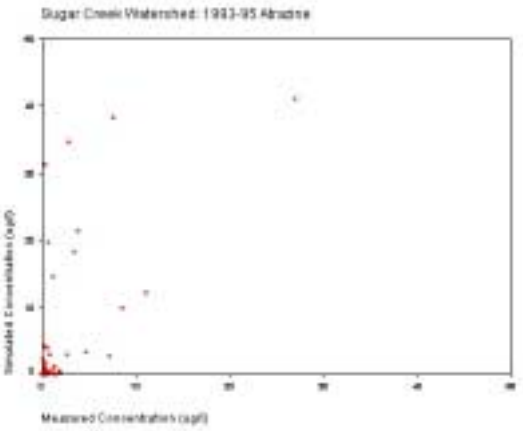


Figure 74: 1993-1995 atrazine results

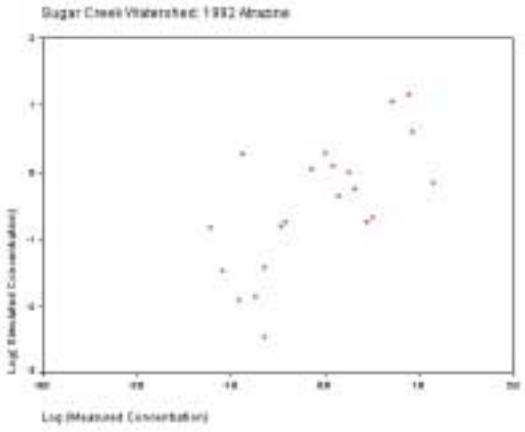


Figure 75: Log-transformed 1992 atrazine results

Figure 76: Log-transformed 1993-1995 atrazine results

Part 4: Calibration—Metolachlor

The calibration and validation periods used for hydrology were also used for metolachlor. Calibration and validation was performed using metolachlor concentrations sampled as part of the U.S.G.S. National Water Quality Assessment Program. Daily in-stream pesticide concentrations simulated by SWAT were compared to grab samples collected at the NAWQA water quality sampling station 39340085524601. Most of the measured concentrations used in the study were obtained from one grab sample (instantaneous measurement) collected on a given day. On a few occasions, more than one grab sample was collected on the same day. When this happened, the concentrations were averaged to obtain a single value for the day.

As with atrazine (Figure 77), the simulated metolachlor concentration missed the first measured peak in 1992 by 2 weeks. The peak concentrations of metolachlor were also overestimated by SWAT in the cold run. The pesticide application data incorporated a 10-day range of error and I shifted the application dates forward by 10 days. While the timing of the simulated metolachlor concentration peaks was not changed, the in-stream peak concentrations of metolachlor dropped due to the longer residence time of the pesticide in the soil prior to removal via runoff. Figure 78 plots the change in in-stream metolachlor concentration with change in time if application.

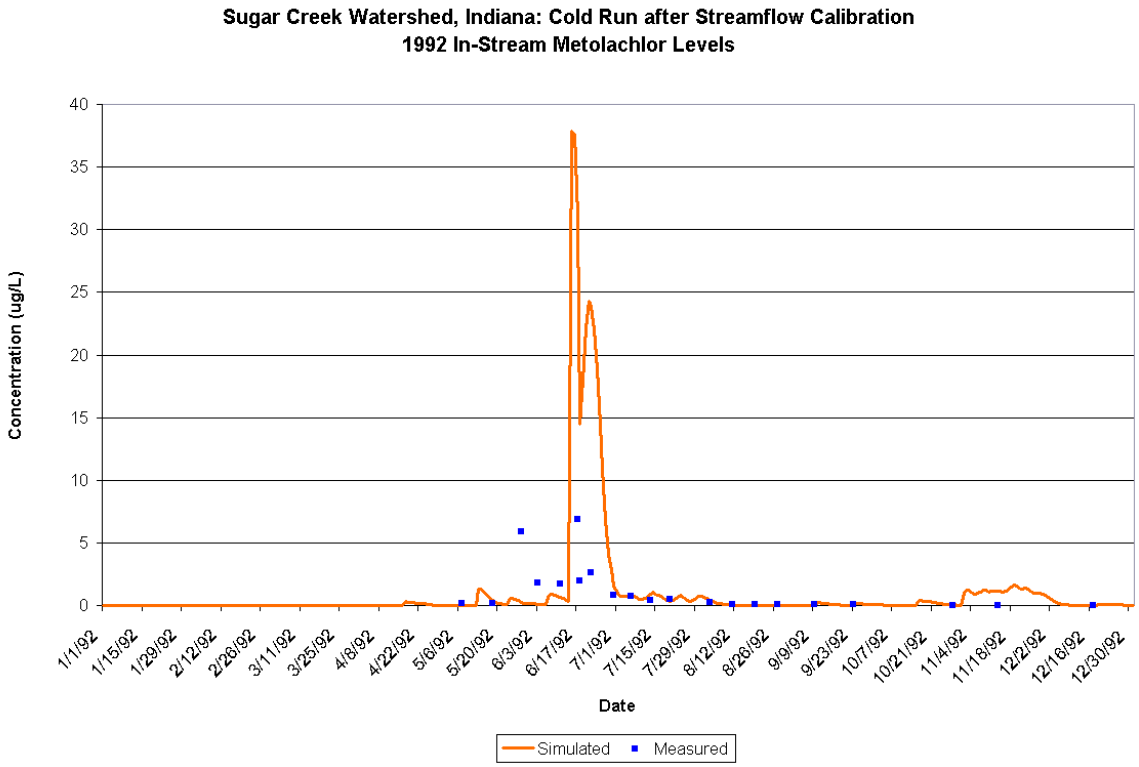


Figure 77: 1992 measured and simulated metolachlor concentrations.

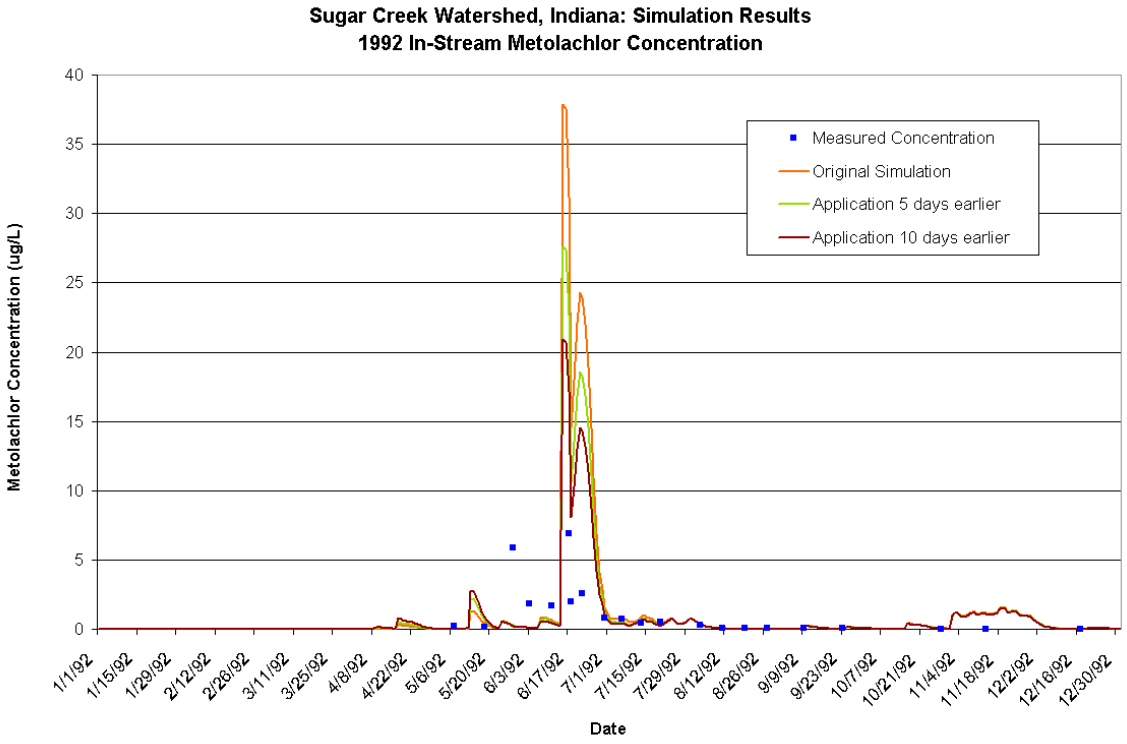


Figure 78: Impact of application timing on in-stream metolachlor concentration.

Visual inspection of the plots of measured and simulated metolachlor concentration was the primary criteria for calibration. One parameter was modified during the metolachlor calibration: PERCOP (*.bsn*). Appendix B lists the modifications made during the calibration process.

Metolachlor concentrations after calibration are graphed along with measured grab bag sample concentrations in Figure 79. Figure 80 displays the same data for 1993-1995. Mass balance values for 1992-1995 averaged over the entire watershed area are provided in Table 13.

Table 13: Metolachlor mass balance and fate

Applied	253264 mg/ha/yr	
(excludes fraction lost during application)		
Decayed	234112 mg/ha/yr	92.44%
In Surface Runoff Entering Stream (Soluble)	6975 mg/ha/yr	2.75%
In Surface Runoff Entering Stream (Sorbed)	113 mg/ha/yr	0.04%
Leached out of Profile	501 mg/ha/yr	0.20%
In Lateral Flow/Tile Drainage Entering Stream	1889 mg/ha/yr	0.75%
Final Amount in Ground	38350 mg/ha	3.79%

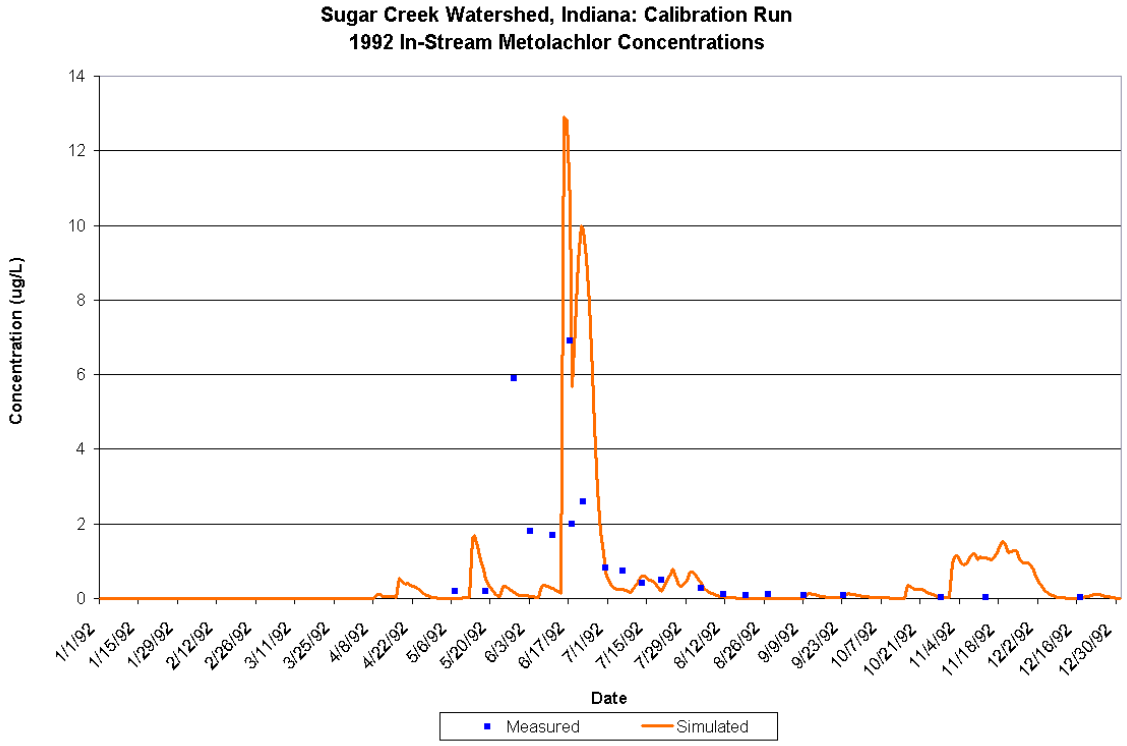


Figure 79: 1992 metolachlor concentrations predicted after calibration

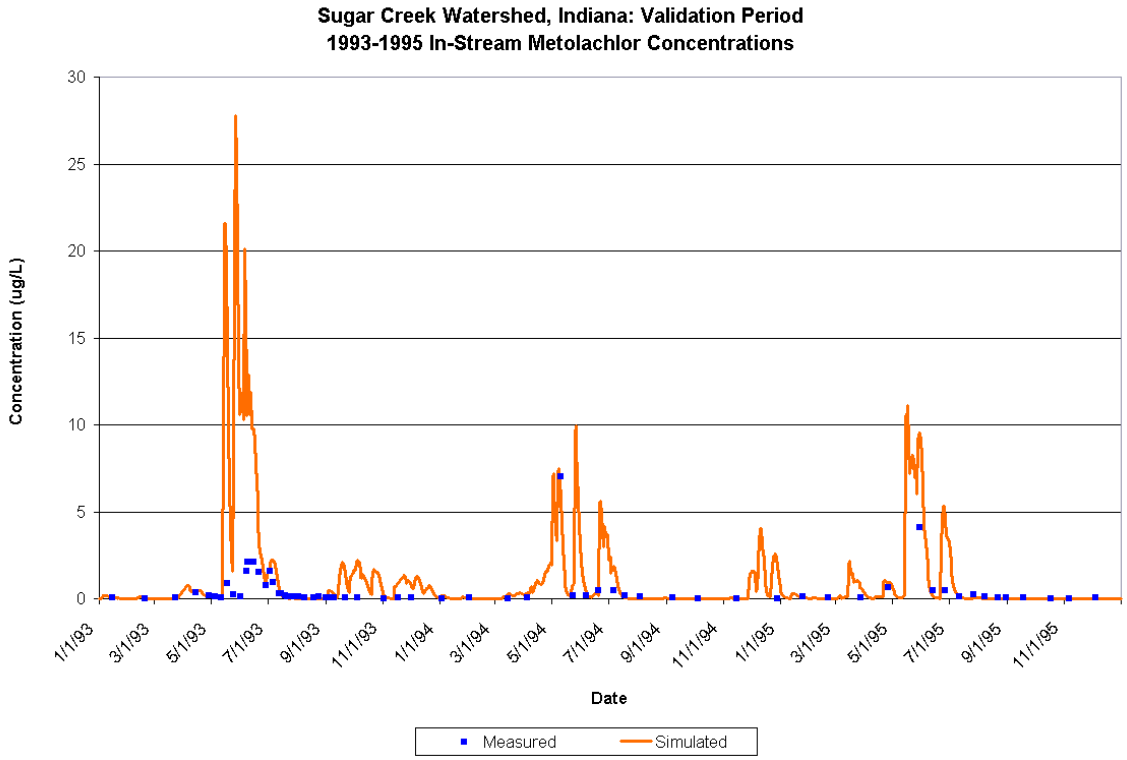


Figure 80: 1993-1995 metolachlor concentration predicted after 1992 calibration

Table 14 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated metolachlor concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated metolachlor concentrations are provided in Figures 81-86. Additional statistical analysis is included in Appendix C.

Table 14: Statistics for Metolachlor

Simulation:	**R^2**	
	conc	**log(conc)**
1992 metolachlor, after calibration	.41	.33
1993-95 metolachlor, after calibration	.28	.34

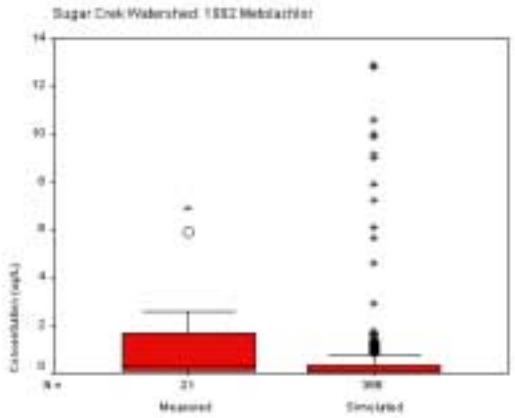


Figure 81: 1992 metolachlor results

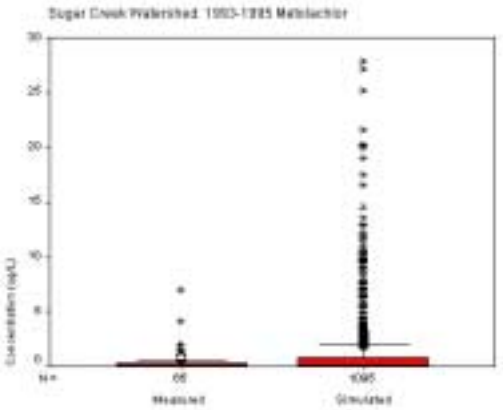


Figure 82: 1993-1995 metolachlor results

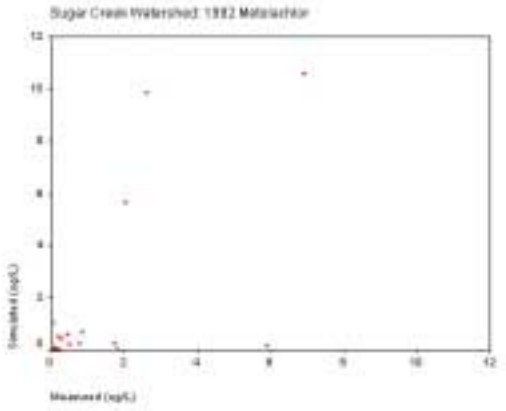


Figure 83: 1992 metolachlor results

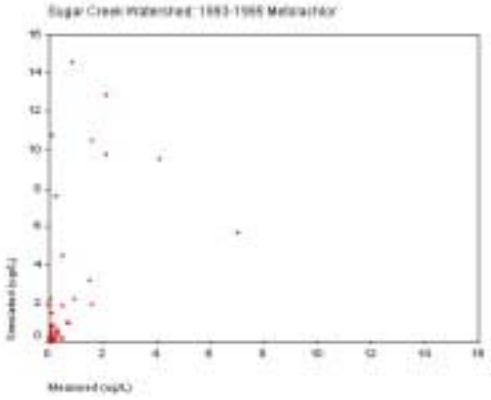


Figure 84: 1993-1995 metolachlor results

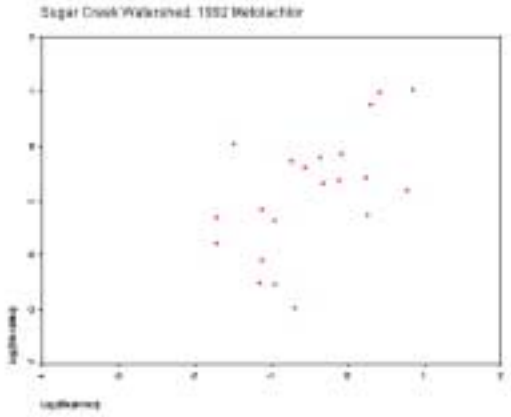


Figure 85: Log-transformed 1992 metolachlor results

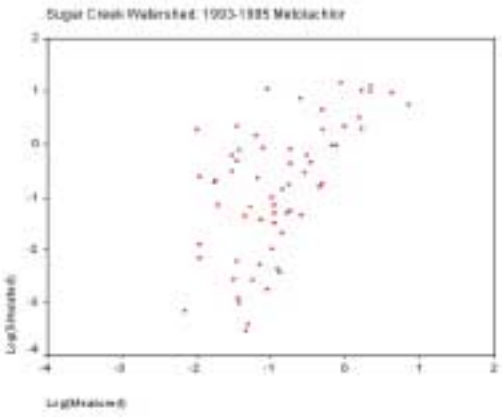


Figure 86: Log-transformed 1993-1995 metolachlor results

Part 5: Calibration—Trifluralin

The calibration and validation periods used for hydrology were also used for trifluralin. Calibration and validation was performed using trifluralin concentrations sampled as part of the U.S.G.S. National Water Quality Assessment Program. Daily in-stream pesticide concentrations simulated by SWAT were compared to grab samples collected at the NAWQA water quality sampling station 39340085524601. Most of the measured concentrations used in the study were obtained from one grab sample (instantaneous measurement) collected on a given day. On a few occasions, more than one grab sample was collected on the same day. When this happened, the concentrations were averaged to obtain a single value for the day.

Unlike atrazine and metolachlor, a significant amount of trifluralin can be transported attached to sediment particles. The values for trifluralin reported by U.S.G.S. are only for trifluralin in solution. It was my understanding that sediment concentration tends to be a highly variable measurement and there is considerable difficulty in obtaining accurate values. To eliminate this source of error, U.S.G.S only reports concentrations for pesticides in solution. Pesticide sorbed to sediment particles is ignored. Figure 87 and 88 plot the measured and simulated sediment concentrations for the calibration and validation periods. Table 15 summarizes the coefficient of determination values for sediment.

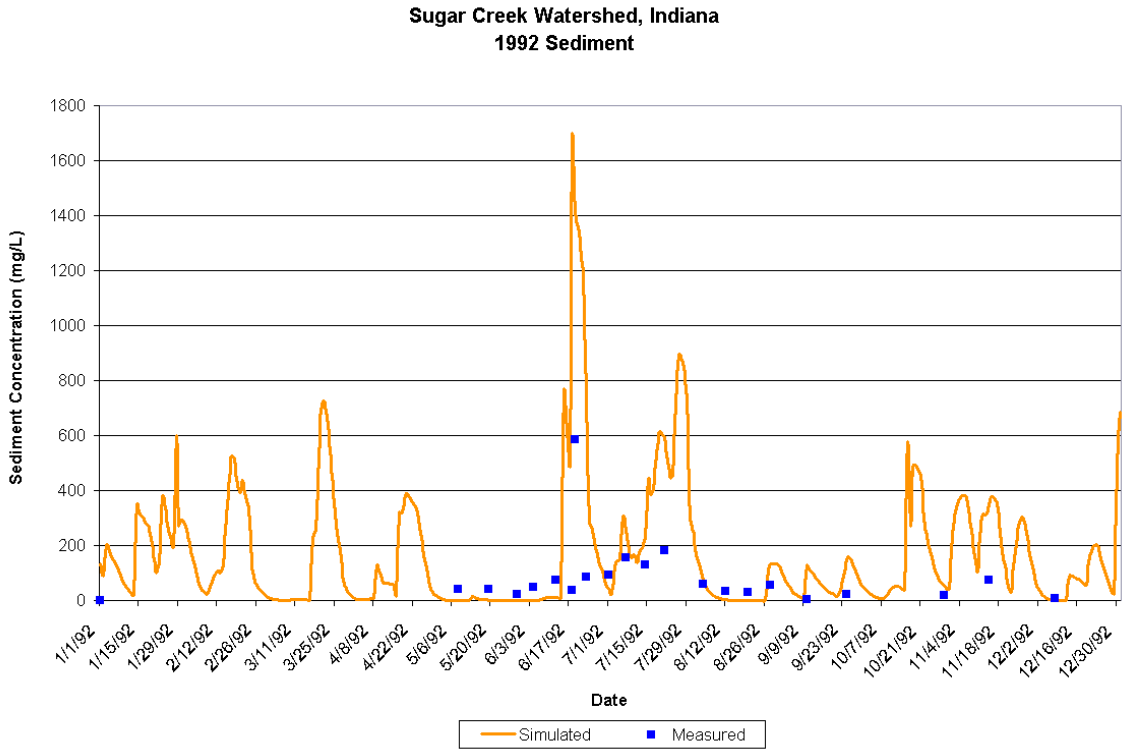


Figure 87: 1992 measured and simulated sediment concentrations.

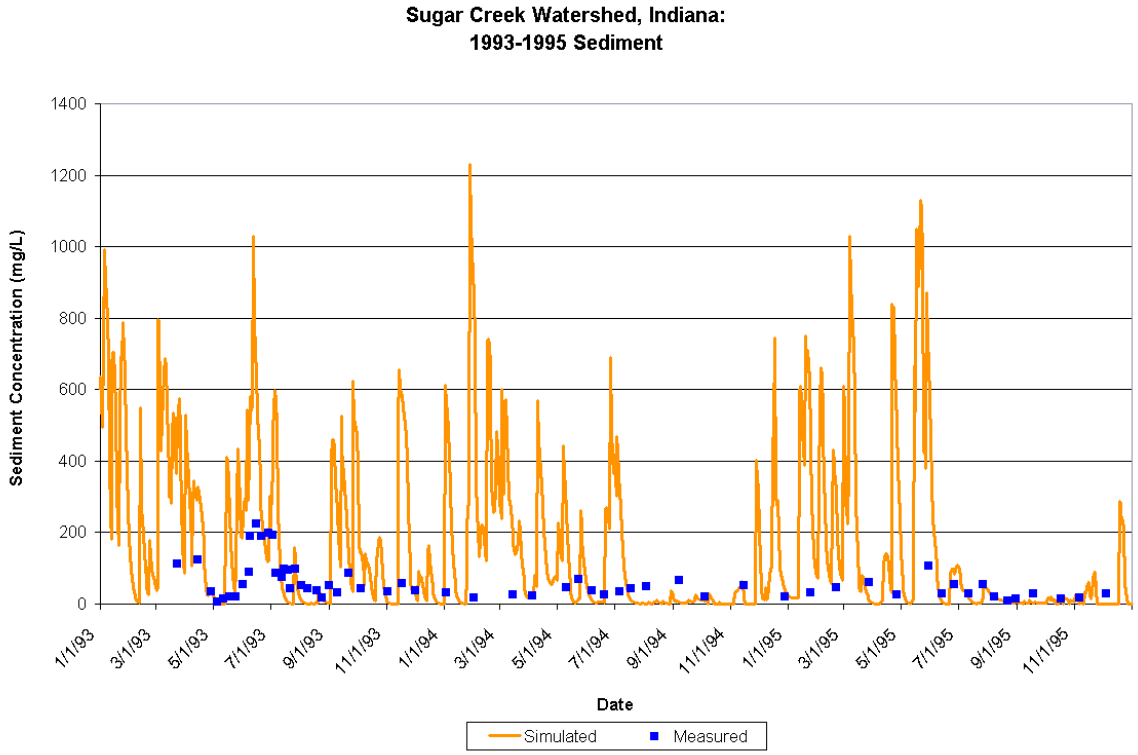


Figure 88: 1993-1995 measured and simulated sediment concentrations.

Table 15: Statistics for Sediment

Simulation:	**R^2**	
	conc	**log(conc)**
1992 sediment, after streamflow calibration	.67	.28
1993-95 sediment, after streamflow calibration	.16	.18

Visual inspection of the plots of measured and simulated trifluralin concentration was the primary criteria for calibration. Four parameters were modified during the trifluralin calibration: AP_EF (*pest.dat*), PERCOP (*.bsn*), CHPST_VOL (*.swq*), and CHPST_STL (*.swq*). Trifluralin is much more volatile than atrazine and metolachlor. To account for the increased amount of volatilization that occurs with this pesticide, the application efficiency and in-stream volatilization coefficient were modified. Appendix B lists all modifications made during the calibration process.

Trifluralin concentrations after calibration are graphed along with measured grab bag sample concentrations in Figure 89. Figure 90 displays the same data for 1993-1995. Mass balance values for 1992-1995 averaged over the entire watershed area are provided in Table 16.

Table 16: Trifluralin mass balance and fate

Applied (excludes fraction lost during application)	20696 mg/ha/yr	
Decayed	20007 mg/ha/yr	96.67%
In Surface Runoff Entering Stream (Soluble)	173 mg/ha/yr	0.84%
In Surface Runoff Entering Stream (Sorbed)	196 mg/ha/yr	0.95%
Leached out of Profile	0 mg/ha/yr	0.00%
In Lateral Flow/Tile Drainage Entering Stream	0.01 mg/ha/yr	0.00%
Final Amount in Ground	1245 mg/ha	1.50%

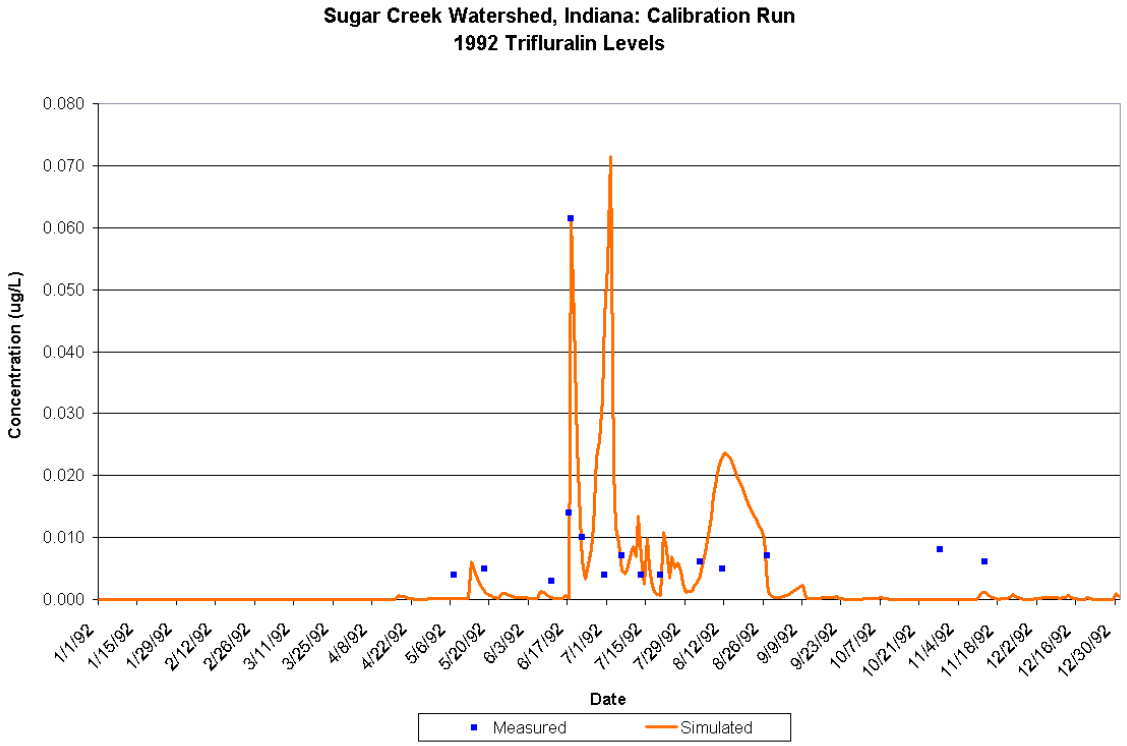


Figure 89: 1992 trifluralin concentrations predicted after calibration

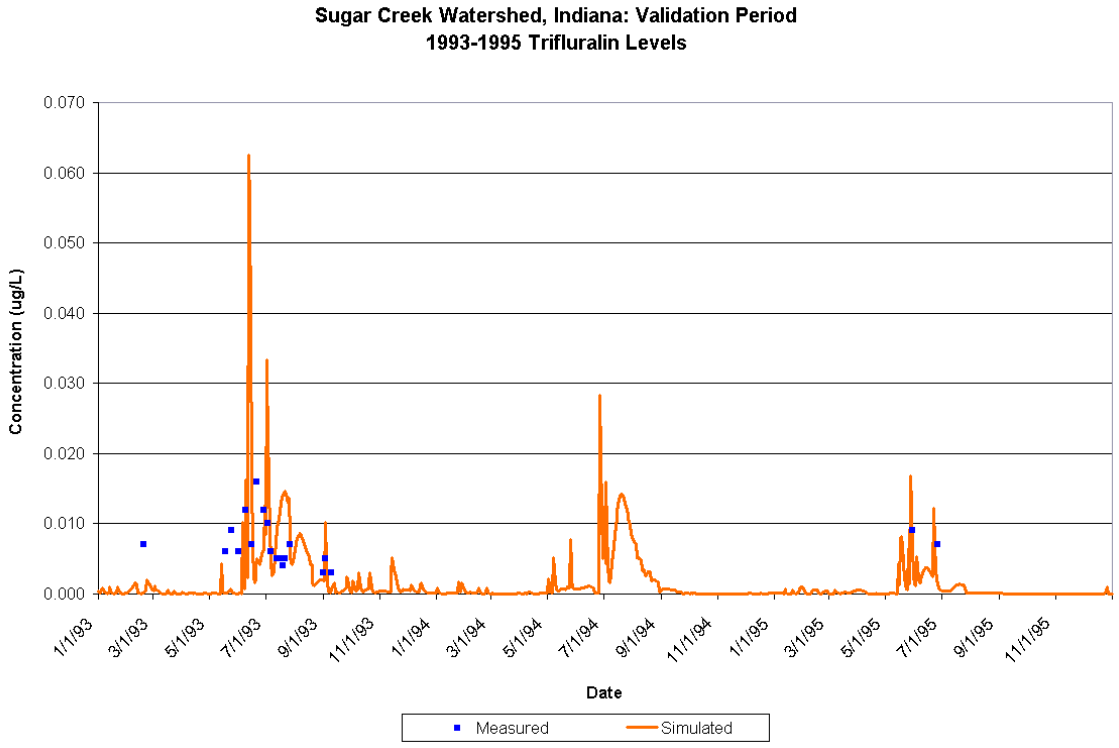


Figure 90: 1993-1995 trifluralin concentrations predicted after 1992 calibration

Table 17 summarizes the coefficient of determination values computed from a regression analysis performed on the pairs of measured and simulated trifluralin concentrations. In addition to the time series graphs, box plots and scatter plots for the measured and simulated trifluralin concentrations are provided in Figures 91-96. Additional statistical analysis is included in Appendix C.

Table 17: Statistics for Trifluralin

Simulation:	**R^2**	
	conc	**log(conc)**
1992 trifluralin, after calibration	.51	.06
1993-95 trifluralin, after calibration	.02	.02

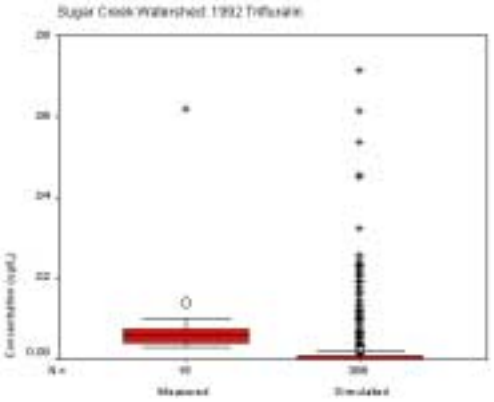


Figure 91: 1992 trifluralin results

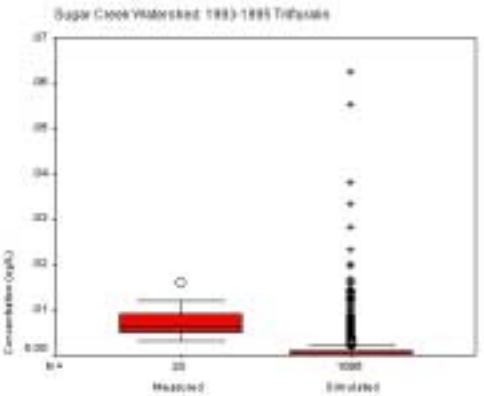


Figure 92: 1993-1995 trifluralin results

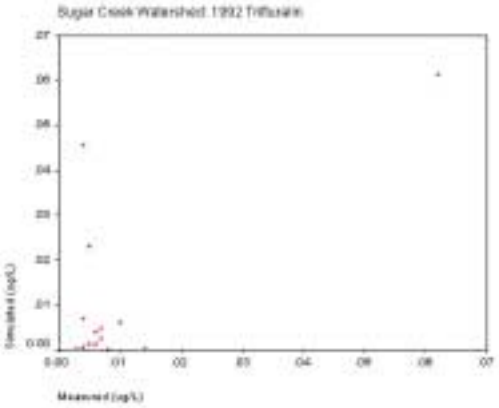


Figure 93: 1992 trifluralin results

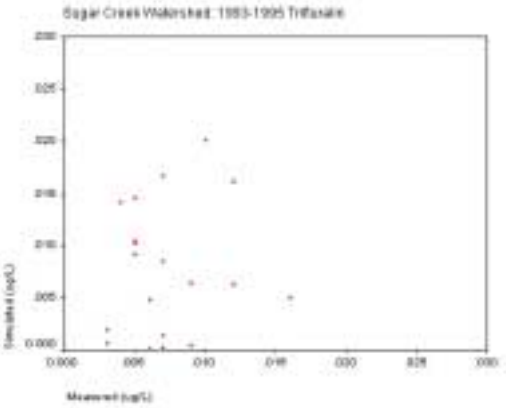
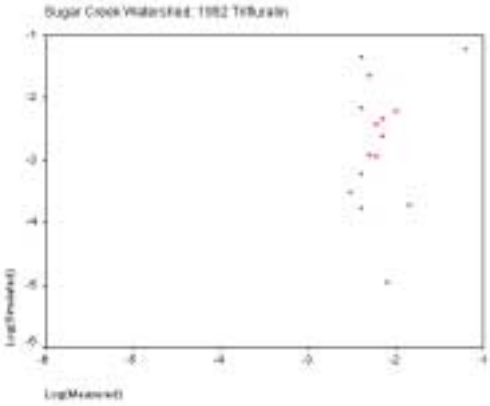


Figure 94: 1993-1995 trifluralin results



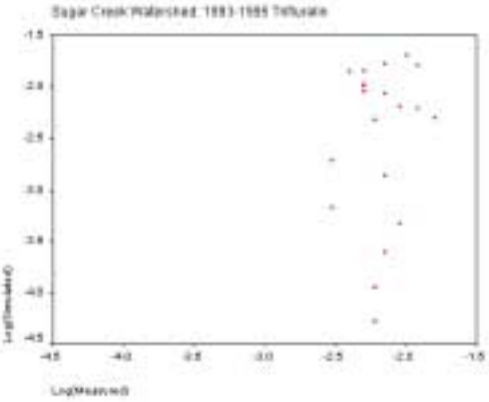


Figure 95: Log-transformed 1992 trifluralin results Figure 96: Log-transformed 1993-1995 trifluralin results

Ease of Use/Computational Time

The SWAT ArcView interface was used to create the dataset used to perform the pesticide simulations. Setting up the watershed configuration, management scenarios, etc. took approximately 2 weeks. Stream flow calibration took 1 week. For each pesticide, calibration took 1 day.

Conclusion

The objective of this study was to calibrate and validate the SWAT model for flow, atrazine, metolachlor and trifluralin in the Sugar Creek watershed, Indiana. Daily simulated flow, and pesticide concentrations were compared with observed values obtained from gaging stations located at the outlet of the watershed. Model calibration was performed with one year of data (1992) while model validation was performed with 3 years of data (1993-1995). Simulated flow and atrazine were closer to the measured values during the validation period. Metolachlor and trifluralin predictions were slightly better in the calibration period than they were in the validation period.

By its very nature, modeling processes at the watershed level requires a large number of assumptions to be made. In addition to the assumptions made during development of the model, we assume the input data (rainfall amounts, pesticide application amounts and timing, etc.) for the watershed is accurate. Each assumption made adds to the amount of uncertainty in modeling results. When analyzing the results of model runs, we make additional assumptions about the reliability of the observed data to which we compare these results. This becomes particularly complicated with pesticides. Unlike streamflow where the flow is measured hourly and 24 values are averaged to obtain the daily flow value, we are lucky to get one pesticide concentration measurement per month. Comparing an instantaneous grab sample with the average daily concentration calculated by the model is a bit like comparing apples to oranges, but that is what is done during model calibration and validation. During the calibration/validation process, the observed pesticide concentrations are assumed to be accurate and no attempt was made to estimate the range in error of the observed values. This assumption is very important to keep in mind when studying results from the pesticide simulations. In particular, the trifluralin simulation was problematic. Most observed values for this pesticide were very close to the minimum detection limit of 0.002 ppb, which means that measurement errors were much more likely.

This study demonstrates the ability of SWAT to realistically predict the movement and transport of pesticides in a watershed. SWAT is particularly able to account for the impact of land management practices on pesticide transport. The effect of pesticide timing on in-stream metolachlor concentration described in

Part 4 is one illustration of the types of evaluations that can be performed with SWAT.

Appendix A-Pesticide Algorithms

Pesticides are toxic by design, and there is a natural concern about the impact of their presence in the environment on human health and environmental quality. The fate and transport of a pesticide are governed by properties such as solubility in water, volatility and ease of degradation. The algorithms in SWAT used to model pesticide movement and fate can be divided into three components: pesticide processes in land areas, transport of pesticide from land areas to the stream network, and in-stream pesticide processes.

Pesticide processes in land areas

The algorithms used to model pesticide movement and fate in land areas are adapted from GLEAMS (Leonard et al., 1987). Figure A1 shows the potential pathways and processes simulated in SWAT.

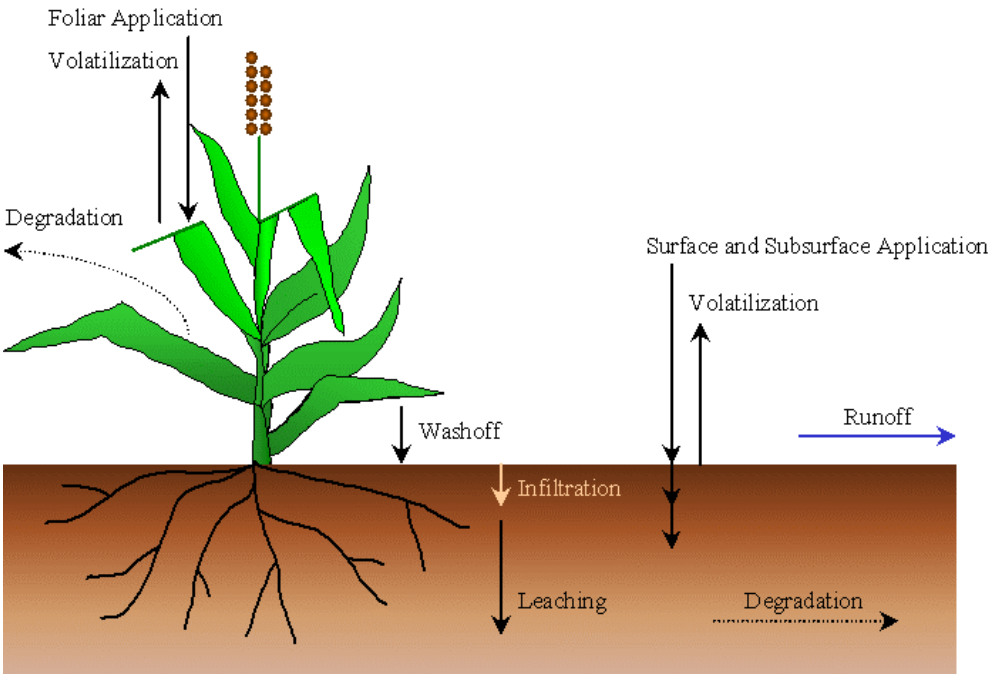


Figure A1: Pesticide transport and fate in land areas.

Pesticide may be aerially applied to a land area with some fraction intercepted by plant foliage and some fraction reaching the soil. The plant leaf-area-index determines what fraction of pesticide reaches the soil surface in an application.

A portion of the pesticide on plant foliage may be washed off during rain events. The fraction washed off is a function of plant morphology, pesticide

solubility, and the timing and intensity of the rainfall event. Wash-off will occur when the amount of precipitation on a given day exceeds 2.54 mm.

The amount of pesticide washing off plant foliage during a precipitation event on a given day is calculated:

$$pst_{f,wsh} = fr_{wsh} \cdot pst_f$$

where $pst_{f,wsh}$ is the amount of pesticide on foliage that is washed off the plant and onto the soil surface on a given day (kg pst/ha), fr_{wsh} is the wash-off fraction for the pesticide, and pst_f is the amount of pesticide on the foliage (kg pst/ha). The wash-off fraction represents the portion of the pesticide on the foliage that is dislodgable.

Degradation is the conversion of a compound into less complex forms. A compound in the soil may degrade upon exposure to light (photo degradation), reaction with chemicals present in the soil (chemical degradation) or through use as a substrate for organisms (biodegradation). Pesticides vary in their susceptibility to degradation. Compounds with chain structures are easier to break apart than compounds containing aromatic rings or other complex structures. The susceptibility of a pesticide to degradation is quantified by the pesticide's half-life.

The half-life for a pesticide defines the number of days required for a given pesticide concentration to be reduced by one-half. The soil half-life entered for a pesticide is a lumped parameter that includes the net effect of volatilization, photolysis, hydrolysis, biological degradation and chemical reactions in the soil. Because pesticide on foliage degrades more rapidly than pesticide in the soil, SWAT allows a different half-life to be defined for foliar degradation.

Pesticide degradation or removal in all soil layers is governed by first-order kinetics:

$$pst_{s,ly,t} = pst_{s,ly,o} \cdot \exp\left[-k_{p,soil} \cdot t\right]$$

where $pst_{s,ly,t}$ is the amount of pesticide in the soil layer at time t (kg pst/ha), $pst_{s,ly,o}$ is the initial amount of pesticide in the soil layer (kg pst/ha), $k_{p,soil}$ is the rate constant for degradation or removal of the pesticide in soil (1/day), and t is the time elapsed since the initial pesticide amount was determined (days). The rate constant is related to the soil half-life as follows:

$$t_{1/2,s} = \frac{0.693}{k_{p,soil}}$$

where $t_{1/2,s}$ is the half-life of the pesticide in the soil (days).

The equation governing pesticide degradation on foliage is:

$$pst_{f,t} = pst_{f,o} \cdot \exp\left[-k_{p,foliar} \cdot t\right]$$

where $pst_{f,t}$ is the amount of pesticide on the foliage at time t (kg pst/ha), $pst_{f,o}$ is the initial amount of pesticide on the foliage (kg pst/ha), $k_{p,foliar}$ is the rate constant for degradation or removal of the pesticide on foliage (1/day), and t is the time elapsed since the initial pesticide amount was determined (days). The rate constant is related to the foliar half-life as follows:

$$t_{1/2,f} = \frac{0.693}{k_{p,foliar}}$$

where $t_{1/2,f}$ is the half-life of the pesticide on foliage (days).

Transport of pesticide from land areas to the stream network

Algorithms governing movement of soluble and sorbed forms of pesticide from land areas to the stream network were taken from EPIC (Williams, 1995).

Pesticide can be transported from land areas to the stream network in solution or attached to sediment. The partitioning of a pesticide between the solution and soil phases is defined by the soil adsorption coefficient for the pesticide. The soil adsorption coefficient is the ratio of the pesticide concentration in the soil or solid phase to the pesticide concentration in the solution or liquid phase:

$$K_p = \frac{C_{solidphase}}{C_{solution}}$$

where K_p is the soil adsorption coefficient ((mg/kg)/(mg/L) or m^3/ton), $C_{solidphase}$ is the concentration of the pesticide sorbed to the solid phase (mg chemical/kg solid material or g/ton), and $C_{solution}$ is the concentration of the pesticide in solution (mg chemical/L solution or g/ton). The definition of the soil adsorption coefficient in the above equation assumes that the pesticide sorption process is linear with concentration and instantaneously reversible.

Because the partitioning of pesticide is dependent upon the amount of organic material in the soil, the soil adsorption coefficient input to the model is normalized for soil organic carbon content. The relationship between the soil adsorption coefficient and the soil adsorption coefficient normalized for soil organic carbon content is:

$$K_p = K_{oc} \cdot \frac{orgC}{100}$$

where K_p is the soil adsorption coefficient ((mg/kg)/(mg/L)), K_{oc} is the soil adsorption coefficient normalized for soil organic carbon content ((mg/kg)/(mg/L) or m^3/ton), and $orgC$ is the percent organic carbon present in the soil.

Pesticide in the soluble phase may be transported with surface runoff, lateral flow or percolation. The change in the amount of pesticide contained in a

soil layer due to transport in solution with flow is a function of time, concentration and amount of flow:

$$\frac{dpst_{s,ly}}{dt} = 0.01 \cdot C_{solution} \cdot w_{mobile}$$

where $pst_{s,ly}$ is the amount of pesticide in the soil layer (kg pst/ha), $C_{solution}$ is the pesticide concentration in solution (mg/L or g/ton), and w_{mobile} is the amount of mobile water on a given day (mm H_2O). The amount of mobile water in the layer is the amount of water lost by surface runoff, lateral flow or percolation:

$$w_{mobile} = Q_{surf} + Q_{lat,surf} + w_{perc,surf} \qquad \text{for top 10 mm}$$

$$w_{mobile} = Q_{lat,ly} + w_{perc,ly} \qquad \text{for lower soil layers}$$

where w_{mobile} is the amount of mobile water in the layer (mm H_2O), Q_{surf} is the surface runoff generated on a given day (mm H_2O), $Q_{lat,ly}$ is the water discharged from the layer by lateral flow (mm H_2O), and $w_{perc,ly}$ is the amount of water percolating to the underlying soil layer on a given day (mm H_2O).

The total amount of pesticide in the soil layer is the sum of the adsorbed and dissolved phases:

$$pst_{s,ly} = 0.01 \cdot \left(C_{solution} \cdot SAT_{ly} + C_{solidphase} \cdot \rho_b \cdot depth_{ly}\right)$$

where $pst_{s,ly}$ is the amount of pesticide in the soil layer (kg pst/ha), $C_{solution}$ is the pesticide concentration in solution (mg/L or g/ton), SAT_{ly} is the amount of water in the soil layer at saturation (mm H_2O), $C_{solidphase}$ is the concentration of the pesticide sorbed to the solid phase (mg/kg or g/ton), ρ_b is the bulk density of the soil layer (Mg/m^3), and $depth_{ly}$ is the depth of the soil layer (mm). Rearranging the equation governing pesticide phase partitioning, $K_p = \frac{C_{solidphase}}{C_{solution}}$, to solve for $C_{solidphase}$ and substituting into the above equation yields:

$$pst_{s,ly} = 0.01 \cdot \left(C_{solution} \cdot SAT_{ly} + C_{solution} \cdot K_p \cdot \rho_b \cdot depth_{ly}\right)$$

which rearranges to

$$C_{solution} = \frac{pst_{s,ly}}{0.01 \cdot \left(SAT_{ly} + K_p \cdot \rho_b \cdot depth_{ly}\right)}$$

Combining the above equation with the pesticide solution transport equation, $\frac{dpst_{s,ly}}{dt} = 0.01 \cdot C_{solution} \cdot w_{mobile}$, yields

$$\frac{dpst_{s,ly}}{dt} = \frac{pst_{s,ly} \cdot w_{mobile}}{\left(SAT_{ly} + K_p \cdot \rho_b \cdot depth_{ly}\right)}$$

Integration of this equation gives

$$pst_{s,ly,t} = pst_{s,ly,o} \cdot \exp\left[\frac{-w_{mobile}}{\left(SAT_{ly} + K_p \cdot \rho_b \cdot depth_{ly}\right)}\right]$$

where $pst_{s,ly,t}$ is the amount of pesticide in the soil layer at time t (kg pst/ha), $pst_{s,ly,o}$ is the initial amount of pesticide in the soil layer (kg pst/ha), w_{mobile} is the amount of mobile water in the layer (mm H_2O), SAT_{ly} is the amount of water in the soil layer at saturation (mm H_2O), K_p is the soil adsorption coefficient ((mg/kg)/(mg/L)), ρ_b is the bulk density of the soil layer (Mg/m^3), and $depth_{ly}$ is the depth of the soil layer (mm).

To obtain the amount of pesticide removed in solution with the flow, the final amount of pesticide is subtracted from the initial amount of pesticide:

$$pst_{flow} = pst_{s,ly,o} \cdot \left(1 - \exp\left[\frac{-w_{mobile}}{\left(SAT_{ly} + K_p \cdot \rho_b \cdot depth_{ly}\right)}\right]\right)$$

where pst_{flow} is the amount of pesticide removed in the flow (kg pst/ha) and all other terms were previously defined.

The pesticide concentration in the mobile water is calculated:

$$conc_{pst,flow} = \min\begin{cases} pst_{flow} / w_{mobile} \\ \\ pst_{sol} / 100. \end{cases}$$

where $conc_{pst,flow}$ is the concentration of pesticide in the mobile water (kg pst/ha-mm H_2O), pst_{flow} is the amount of pesticide removed in the flow (kg pst/ha), w_{mobile} is the amount of mobile water in the layer (mm H_2O), and pst_{sol} is the solubility of the pesticide in water (mg/L).

Pesticide moved to the underlying layer by percolation is calculated:

$$pst_{perc,ly} = conc_{pst,flow} \cdot w_{perc,ly}$$

where $pst_{perc,ly}$ is the pesticide moved to the underlying layer by percolation (kg pst/ha), $conc_{pst,flow}$ is the concentration of pesticide in the mobile water for the layer (kg pst/mm H_2O), and $w_{perc,ly}$ is the amount of water percolating to the underlying soil layer on a given day (mm H_2O).

Pesticide removed in lateral flow is calculated:

$$pst_{lat,surf} = \beta_{pst} \cdot conc_{pst,flow} \cdot Q_{lat,surf} \qquad \text{for top 10 mm}$$

$$pst_{lat,ly} = conc_{pst,flow} \cdot Q_{lat,ly} \qquad \text{for lower layers}$$

where $pst_{lat,ly}$ is the pesticide removed in lateral flow from a layer (kg pst/ha), β_{pst} is the pesticide percolation coefficient, $conc_{pst,flow}$ is the concentration of pesticide in the mobile water for the layer (kg pst/mm H_2O), and $Q_{lat,ly}$ is the water discharged from the layer by lateral flow (mm H_2O). The pesticide percolation coefficient allows the user to set the concentration of pesticide in runoff and lateral flow from the top 10 mm to a fraction of the concentration in percolate.

Pesticide removed in surface runoff is calculated:

$$pst_{surf} = \beta_{pst} \cdot conc_{pst,flow} \cdot Q_{surf}$$

where pst_{surf} is the pesticide removed in surface runoff (kg pst/ha), β_{pst} is the pesticide percolation coefficient, $conc_{pst,flow}$ is the concentration of pesticide in the mobile water for the top 10 mm of soil (kg pst/mm H_2O), and Q_{surf} is the surface runoff generated on a given day (mm H_2O).

Pesticide attached to soil particles may be transported by surface runoff to the main channel. This phase of pesticide is associated with the sediment loading from the HRU and changes in sediment loading will impact the loading of sorbed pesticide. The amount of pesticide transported with sediment to the stream is calculated with a loading function developed by McElroy et al. (1976) and modified by Williams and Hann (1978).

$$pst_{sed} = 0.001 \cdot C_{solidphase} \cdot \frac{sed}{area_{hru}} \cdot \varepsilon_{pst:sed}$$

where pst_{sed} is the amount of sorbed pesticide transported to the main channel in surface runoff (kg pst/ha), $C_{solidphase}$ is the concentration of pesticide on sediment in the top 10 mm (g pst/ metric ton soil), *sed* is the sediment yield on a given day (metric tons), $area_{hru}$ is the HRU area (ha), and $\varepsilon_{pst:sed}$ is the pesticide enrichment ratio.

The total amount of pesticide in the soil layer is the sum of the adsorbed and dissolved phases:

$$pst_{s,ly} = 0.01 \cdot \left(C_{solution} \cdot SAT_{ly} + C_{solidphase} \cdot \rho_b \cdot depth_{ly}\right)$$

where $pst_{s,ly}$ is the amount of pesticide in the soil layer (kg pst/ha), $C_{solution}$ is the pesticide concentration in solution (mg/L or g/ton), SAT_{ly} is the amount of water in the soil layer at saturation (mm H_2O), $C_{solidphase}$ is the concentration of the pesticide sorbed to the solid phase (mg/kg or g/ton), ρ_b is the bulk density of the soil layer (Mg/m^3), and $depth_{ly}$ is the depth of the soil layer (mm). Rearranging the equation governing pesticide partitioning, $K_p = \frac{C_{solidphase}}{C_{solution}}$, to solve for $C_{solution}$ and substituting into the above equation yields:

$$pst_{s,ly} = 0.01 \cdot \left(\frac{C_{solidphase}}{K_p} \cdot SAT_{ly} + C_{solidphase} \cdot \rho_b \cdot depth_{ly}\right)$$

which rearranges to

$$C_{solidphase} = \frac{100 \cdot K_p \cdot pst_{s,ly}}{\left(SAT_{ly} + K_p \cdot \rho_b \cdot depth_{ly}\right)}$$

where $C_{solidphase}$ is the concentration of the pesticide sorbed to the solid phase (mg/kg or g/ton), K_p is the soil adsorption coefficient ((mg/kg)/(mg/L) or m^3/ton) $pst_{s,ly}$ is the amount of pesticide in the soil layer (kg pst/ha), SAT_{ly} is the amount of water in the soil layer at saturation (mm H_2O), ρ_b is the bulk density of the soil layer (Mg/m^3), and $depth_{ly}$ is the depth of the soil layer (mm).

As surface runoff flows over the soil surface, part of the water's energy is used to pick up and transport soil particles. The smaller particles weigh less and are more easily transported than coarser particles. When the particle size distribution of the transported sediment is compared to that of the soil surface layer, the sediment load to the main channel has a greater proportion of clay sized particles. In other words, the sediment load is enriched in clay particles. The sorbed phase of pesticide in the soil is attached primarily to colloidal (clay) particles, so the sediment load will also contain a greater proportion or concentration of pesticide than that found in the soil surface layer.

The enrichment ratio is defined as the ratio of the concentration of sorbed pesticide transported with the sediment to the concentration in the soil surface layer. SWAT will calculate an enrichment ratio for each storm event, or allow the user to define a particular enrichment ratio for sorbed pesticide that is used for all storms during the simulation. To calculate the enrichment ratio, SWAT uses a relationship described by Menzel (1980) in which the enrichment ratio is logarithmically related to sediment concentration. The equation used to calculate the pesticide enrichment ratio, $\varepsilon_{pst:sed}$, for each storm event is:

$$\varepsilon_{pst:sed} = 0.78 \cdot \left(conc_{sed,surq}\right)^{-0.2468}$$

where $conc_{sed,surq}$ is the concentration of sediment in surface runoff (Mg sed/m^3 H_2O). The concentration of sediment in surface runoff is calculated:

$$conc_{sed,surq} = \frac{sed}{10 \cdot area_{hru} \cdot Q_{surf}}$$

where *sed* is the sediment yield on a given day (metric tons), $area_{hru}$ is the HRU area (ha), and Q_{surf} is the amount of surface runoff on a given day (mm H_2O).

In-stream pesticide processes

SWAT incorporates a simple mass balance developed by Chapra (1997) to model the transformation and transport of pesticides in streams. The model assumes a well-mixed layer of water overlying a homogenous sediment layer. Only one pesticide can be routed through the stream network in a given simulation.

Pesticide in a reach segment is increased through addition of mass in inflow as well as resuspension and diffusion of pesticide from the sediment layer. The amount of pesticide in a reach segment is reduced through removal in outflow as well as degradation, volatilization, settling and diffusion into the underlying sediment.

Pesticides will partition into particulate and dissolved forms. The fraction of pesticide in each phase is a function of the pesticide's partition coefficient and the reach segment's suspended solid concentration:

$$F_d = \frac{1}{1 + K_d \cdot conc_{sed}}$$

$$F_p = \frac{K_d \cdot conc_{sed}}{1 + K_d \cdot conc_{sed}} = 1 - F_d$$

where F_d is the fraction of total pesticide in the dissolved phase, F_p is the fraction of total pesticide in the particulate phase, K_d is the pesticide partition coefficient (m^3/g), and $conc_{sed}$ is the concentration of suspended solids in the water (g/m^3).

The pesticide partition coefficient can be estimated from the octanol-water partition coefficient (Chapra, 1997):

$$K_d = 3.085 \times 10^{-8} \cdot K_{ow}$$

where K_d is the pesticide partition coefficient (m^3/g) and K_{ow} is the pesticide's octanol-water partition coefficient ($\mathrm{mg\,m^{-3}_{octanol}}\left(\mathrm{mg\,m^{-3}_{water}}\right)^{-1}$). Values for the octanol-water partition coefficient have been published for many chemicals. If a published value cannot be found, it can be estimated from solubility (Chapra, 1997):

$$\log(K_{ow}) = 5.00 - 0.670 \cdot \log(pst'_{sol})$$

where pst'_{sol} is the pesticide solubility (µmoles/L). The solubility in these units is calculated:

$$pst'_{sol} = \frac{pst_{sol}}{MW} \cdot 10^3$$

where pst'_{sol} is the pesticide solubility (µmoles/L), pst_{sol} is the pesticide solubility (mg/L) and MW is the molecular weight (g/mole).

Pesticides in both the particulate and dissolved forms are subject to degradation. The amount of pesticide that is removed from the water via degradation is:

$$pst_{deg,wtr} = k_{p,aq} \cdot pst_{rchwtr} \cdot TT$$

where $pst_{deg,wtr}$ is the amount of pesticide removed from the water via degradation (mg pst), $k_{p,aq}$ is the rate constant for degradation or removal of pesticide in the water (1/day), pst_{rchwtr} is the amount of pesticide in the water at the beginning of the day (mg pst), and TT is the flow travel time (days). The rate constant is related to the aqueous half-life:

$$k_{p,aq} = \frac{0.693}{t_{1/2,aq}}$$

where $k_{p,aq}$ is the rate constant for degradation or removal of pesticide in the water (1/day), and $t_{1/2,aq}$ is the aqueous half-life for the pesticide (days).

Pesticide in the dissolved phase is available for volatilization. The amount of pesticide removed from the water via volatilization is:

$$pst_{vol,wtr} = \frac{v_v}{depth} \cdot F_d \cdot pst_{rchwtr} \cdot TT$$

where $pst_{vol,wtr}$ is the amount of pesticide removed via volatilization (mg pst), v_v is the volatilization mass-transfer coefficient (m/day), *depth* is the flow depth (m), F_d is the fraction of total pesticide in the dissolved phase, pst_{rchwtr} is the amount of pesticide in the water (mg pst), and *TT* is the flow travel time (days).

The volatilization mass-transfer coefficient can be calculated based on Whitman's two-film or two-resistance theory (Whitman, 1923; Lewis and Whitman, 1924 as described in Chapra, 1997). While the main body of the gas and liquid phases are assumed to be well-mixed and homogenous, the two-film theory assumes that a substance moving between the two phases encounters maximum resistance in two laminar boundary layers where transfer is a function of molecular diffusion. In this type of system the transfer coefficient or velocity is:

$$v_v = K_l \cdot \frac{H_e}{H_e + R \cdot T_K \cdot \left(K_l / K_g\right)}$$

where v_v is the volatilization mass-transfer coefficient (m/day), K_l is the mass-transfer velocity in the liquid laminar layer (m/day), K_g is the mass-transfer velocity in the gaseous laminar layer (m/day), H_e is Henry's constant (atm m^3 $mole^{-1}$), R is the universal gas constant (8.206×10^{-5} atm m^3 $(K\ mole)^{-1}$), and T_K is the temperature (K).

For rivers where liquid flow is turbulent, the transfer coefficients are estimated using the surface renewal theory (Higbie, 1935; Danckwerts, 1951; as described by Chapra, 1997). The surface renewal model visualizes the system as consisting of parcels of water that are brought to the surface for a period of time. The fluid elements are assumed to reach and leave the air/water interface randomly, i.e. the exposure of the fluid elements to air is described by a statistical distribution. The transfer velocities for the liquid and gaseous phases are calculated:

$$K_l = \sqrt{r_l \cdot D_l} \qquad K_g = \sqrt{r_g \cdot D_g}$$

where K_l is the mass-transfer velocity in the liquid laminar layer (m/day), K_g is the mass-transfer velocity in the gaseous laminar layer (m/day), D_l is the liquid molecular diffusion coefficient (m^2/day), D_g is the gas molecular diffusion coefficient (m^2/day), r_l is the liquid surface renewal rate (1/day), and r_g is the gaseous surface renewal rate (1/day).

O'Connor and Dobbins (1958) defined the surface renewal rate as the ratio of the average stream velocity to depth.

$$r_l = \frac{86400 \cdot v_c}{depth}$$

where r_l is the liquid surface renewal rate (1/day), v_c is the average stream velocity (m/s) and *depth* is the depth of flow (m).

Pesticide in the particulate phase may be removed from the water layer by settling. Settling transfers pesticide from the water to the sediment layer. The amount of pesticide that is removed from the water via settling is:

$$pst_{stl,wtr} = \frac{v_s}{depth} \cdot F_p \cdot pst_{rchwtr} \cdot TT$$

where $pst_{stl,wtr}$ is the amount of pesticide removed from the water due to settling (mg pst), v_s is the settling velocity (m/day), *depth* is the flow depth (m), F_p is the fraction of total pesticide in the particulate phase, pst_{rchwtr} is the amount of pesticide in the water (mg pst), and *TT* is the flow travel time (days).

Pesticide is removed from the reach segment in outflow. The amount of dissolved and particulate pesticide removed from the reach segment in outflow is:

$$pst_{sol,o} = Q \cdot \frac{F_d \cdot pst_{rchwtr}}{V}$$

$$pst_{sorb,o} = Q \cdot \frac{F_p \cdot pst_{rchwtr}}{V}$$

where $pst_{sol,o}$ is the amount of dissolved pesticide removed via outflow (mg pst), $pst_{sorb,o}$ is the amount of particulate pesticide removed via outflow (mg pst), Q is the rate of outflow from the reach segment (m^3 H_2O/day), F_d is the fraction of total pesticide in the dissolved phase, F_p is the fraction of total pesticide in the particulate phase, pst_{rchwtr} is the amount of pesticide in the water (mg pst), and V is the volume of water in the reach segment (m^3 H_2O).

Pesticide in the sediment layer underlying a reach segment is increased through addition of mass by settling and diffusion from the water. The amount of pesticide in the sediment layer is reduced through removal by degradation, resuspension, diffusion into the overlying water, and burial.

As in the water layer, pesticides in the sediment layer will partition into particulate and dissolved forms. Calculation of the solid-liquid partitioning in the sediment layer requires a suspended solid concentration. The “concentration” of solid particles in the sediment layer is defined as:

$$conc^*_{sed} = \frac{M_{sed}}{V_{tot}}$$

where $conc^*_{sed}$ is the “concentration” of solid particles in the sediment layer (g/m^3), M_{sed} is the mass of solid particles in the sediment layer (g) and V_{tot} is the total volume of the sediment layer (m^3).

Mass and volume are also used to define the porosity and density of the sediment layer. In the sediment layer, porosity is the fraction of the total volume in the liquid phase:

$$\phi = \frac{V_{wtr}}{V_{tot}}$$

where ϕ is the porosity, V_{wtr} is the volume of water in the sediment layer (m^3) and V_{tot} is the total volume of the sediment layer (m^3). The fraction of the volume in the solid phase can then be defined as:

$$1-\phi = \frac{V_{sed}}{V_{tot}}$$

where ϕ is the porosity, V_{sed} is the volume of solids in the sediment layer (m^3) and V_{tot} is the total volume of the sediment layer (m^3).

The density of sediment particles is defined as:

$$\rho_s = \frac{M_{sed}}{V_{sed}}$$

where ρ_s is the particle density (g/m^3), M_{sed} is the mass of solid particles in the sediment layer (g), and V_{sed} is the volume of solids in the sediment layer (m^3).

Solving $1-\phi = \frac{V_{sed}}{V_{tot}}$ for V_{tot} and $\rho_s = \frac{M_{sed}}{V_{sed}}$ for M_{sed} and substituting into $conc^*_{sed} = \frac{M_{sed}}{V_{tot}}$ $conc^*_{sed} = \frac{M_{sed}}{V_{tot}}$ yields:

$$conc^*_{sed} = (1-\phi)\cdot \rho_s$$

where $conc^*_{sed}$ is the "concentration" of solid particles in the sediment layer (g/m^3), ϕ is the porosity, and ρ_s is the particle density (g/m^3).

Assuming ϕ = 0.5 and ρ_s = 2.6×10^6 g/m^3, the "concentration" of solid particles in the sediment layer is 1.3×10^6 g/m^3.

The fraction of pesticide in each phase is then calculated:

$$F_{d,sed} = \frac{1}{\phi + (1-\phi)\cdot \rho_s \cdot K_d}$$

$$F_{p,sed} = 1 - F_{d,sed}$$

where $F_{d,sed}$ is the fraction of total sediment pesticide in the dissolved phase, $F_{p,sed}$ is the fraction of total sediment pesticide in the particulate phase, ϕ is the porosity, ρ_s is the particle density (g/m^3), and K_d is the pesticide partition coefficient (m^3/g). The pesticide partition coefficient used for the water layer is also used for the sediment layer.

Pesticides in both the particulate and dissolved forms are subject to degradation. The amount of pesticide that is removed from the sediment via degradation is:

$$pst_{deg,sed} = k_{p,sed} \cdot pst_{rchsed}$$

where $pst_{deg,sed}$ is the amount of pesticide removed from the sediment via degradation (mg pst), $k_{p,sed}$ is the rate constant for degradation or removal of pesticide in the sediment (1/day), and pst_{rchsed} is the amount of pesticide in the sediment (mg pst). The rate constant is related to the sediment half-life:

$$k_{p,sed} = \frac{0.693}{t_{1/2,sed}}$$

where $k_{p,sed}$ is the rate constant for degradation or removal of pesticide in the sediment (1/day), and $t_{1/2,sed}$ is the sediment half-life for the pesticide (days).

Pesticide in the sediment layer is available for resuspension. The amount of pesticide that is removed from the sediment via resuspension is:

$$pst_{rsp,wtr} = \frac{v_r}{depth} \cdot pst_{rchsed} \cdot TT$$

where $pst_{rsp,wtr}$ is the amount of pesticide removed via resuspension (mg pst), v_r is the resuspension velocity (m/day), *depth* is the flow depth (m), pst_{rchsed} is the amount of pesticide in the sediment (mg pst), and *TT* is the flow travel time (days). Pesticide removed from the sediment layer by resuspension is added to the water layer.

Pesticide in the dissolved phase is available for diffusion. Diffusion transfers pesticide between the water and sediment layers. The direction of movement is controlled by the pesticide concentration. Pesticide will move from areas of high concentration to areas of low concentration. The amount of pesticide that is transferred between the water and sediment by diffusion is:

$$pst_{dif} = \left| \frac{v_d}{depth} \cdot \left(F_{d,sed} \cdot pst_{rchsed} - F_d \cdot pst_{rchwtr} \right) \cdot TT \right|$$

where pst_{dif} is the amount of pesticide transferred between the water and sediment by diffusion (mg pst), v_d is the rate of diffusion or mixing velocity (m/day), *depth* is the flow depth (m), $F_{d,sed}$ is the fraction of total sediment pesticide in the dissolved phase, pst_{rchsed} is the amount of pesticide in the sediment (mg pst), F_d is the fraction of total water layer pesticide in the dissolved phase, pst_{rchwtr} is the amount of pesticide in the water (mg pst), and *TT* is the flow duration (days). If $F_{d,sed} \cdot pst_{rchsed} > F_d \cdot pst_{rchwtr}$, pst_{dif} is transferred from the sediment to the water layer. If, $F_{d,sed} \cdot pst_{rchsed} < F_d \cdot pst_{rchwtr}$, pst_{dif} is transferred from the water to the sediment layer.

The diffusive mixing velocity, v_d, can be estimated from the empirically derived formula (Chapra, 1997):

$$v_d = \frac{69.35}{365} \cdot \phi \cdot MW^{-2/3}$$

where v_d is the rate of diffusion or mixing velocity (m/day), ϕ is the sediment porosity, and *MW* is the molecular weight of the pesticide compound.

Pesticide in the sediment layer may be lost by burial. The amount of pesticide that is removed from the sediment via burial is:

$$pst_{bur} = \frac{v_b}{D_{sed}} \cdot pst_{rchsed}$$

where pst_{bur} is the amount of pesticide removed via burial (mg pst), v_b is the burial velocity (m/day), D_{sed} is the depth of the active sediment layer (m), and pst_{rchsed} is the amount of pesticide in the sediment (mg pst).

The processes described above can be combined into mass balance equations for the well-mixed reach segment and the well-mixed sediment layer:

$$\Delta pst_{rchwtr} = pst_{in} - \left(pst_{sol,o} + pst_{sorb,o}\right) - pst_{deg,wtr} - pst_{vol,wtr} - pst_{stl,wtr} + pst_{rsp,wtr} \pm pst_{dif}$$

$$\Delta pst_{rchsed} = -pst_{deg,sed} + pst_{stl,wtr} - pst_{rsp,wtr} - pst_{bur} \pm pst_{dif}$$

where Δpst_{rchwtr} is the change in pesticide mass in the water layer (mg pst), Δpst_{rchsed} is the change in pesticide mass in the sediment layer (mg pst), pst_{in} is the pesticide added to the reach segment via inflow (mg pst), $pst_{sol,o}$ is the amount of dissolved pesticide removed via outflow (mg pst), $pst_{sorb,o}$ is the amount of particulate pesticide removed via outflow (mg pst), $pst_{deg,wtr}$ is the amount of pesticide removed from the water via degradation (mg pst), $pst_{vol,wtr}$ is the amount of pesticide removed via volatilization (mg pst), $pst_{stl,wtr}$ is the amount of pesticide removed from the water due to settling (mg pst), $pst_{rsp,wtr}$ is the amount of pesticide removed via resuspension (mg pst), pst_{dif} is the amount of pesticide transferred between the water and sediment by diffusion (mg pst), $pst_{deg,sed}$ is the amount of pesticide removed from the sediment via degradation (mg pst), pst_{bur} is the amount of pesticide removed via burial (mg pst)

Appendix B-Model Inputs Modified during Calibration

Table B1 lists input parameters used to calibrate long-term stream flow in the Sugar Creek Watershed.

Table B1: Long-term water yield calibration parameters

Variable Name	Description	Final Value
CN2	SCS runoff curve number for moisture condition II	lowered CN2 for HRUs with row crops by 7
CANMX	maximum water storage in canopy	= 40 mm for HRUs with deciduous forest and forested wetland
SOL_AWC	available water capacity (by soil layer)	increased by 0.02 for HRUs with row crops
SOL_K	saturated hydraulic conductivity (by soil layer)	increased by 10% for HRUs with row crops

Table B2 lists input parameters used to calibrate 1992 daily stream flow in the Sugar Creek Watershed.

Table B2: Daily stream flow calibration parameters

Variable Name	Description	Initial/Final Value
MSK_X	weighting factor controlling relative importance of inflow and outflow in determining water storage in reach	0.2 => 0.3
MSK_CO1	Coefficient controlling impact of normal flow storage time constant on water routing	0.0 => 0.5
MSK_CO2	Coefficient controlling impact of low flow storage time constant on water routing	3.5 => 3.0
CH_N	Manning's n value for the main channels	0.014 => 0.009
CANMX	maximum water storage in canopy	= 50 mm for HRUs with deciduous forest and forested wetland
ESCO	Soil evaporation compensation factor	0.95 => 0.90

Table B3 lists input parameters used to calibrate metolachlor concentration in the Sugar Creek Watershed.

Table B3: Metolachlor calibration parameters

Variable Name	Description	Initial/Final Value
MON/DAY	Timing of pesticide applications	shifted application dates forward by 10 days in all years
PERCOP	parameter controlling partitioning of soluble pesticide between percolate and surface runoff	0.50 => 0.30

Table B4 lists input parameters used to calibrate trifluralin concentration in the Sugar Creek Watershed.

Table B4: Trifluralin calibration parameters

Variable Name	Description	Initial/Final Value
AP_EF	application efficiency	0.75 => 0.60
PERCOP	parameter controlling partitioning of soluble pesticide between percolate and surface runoff	0.50 => 0.30
CHPST_VOL	pesticide volatilization coefficient in reach (m/day)	0.010 => 0.020
CHPST_STL	Settling velocity for pesticide sorbed to sediment (m/day)	1.0 => 1.2

Appendix C-Additional Statistical Analysis

In an unpublished article, Burns proposed a methodology for evaluating the accuracy of environmental models. This appendix presents the results of this statistical analysis for SWAT predictions in the Sugar Creek watershed.

The methodology is comprised of 8 steps. Steps 1-3 are universal to all datasets tested and are summarized prior to presenting results for the individual datasets.

Step 1:

Select appropriate validity criteria and establish acceptable levels of user and modeler risk.

Will test all combinations in the following table:

Model predictions and observed data must differ by no more than a factor of	Percent of pairs that must meet difference criteria	User error or uncertainty (α) / modeler error or uncertainty (β)
2x	99%	1%
5x	95%	2%
10x	90%	5%
20x	80%	
100x		

Step 2:

Develop an appropriate error measure for

H_0: the model is invalid, vs.

H_a: the model is valid

$$D = \log(P) - \log(O)$$

D^*, the limit of the acceptable difference, is

Acceptable difference	D^*
2x	log(2.0) = 0.301
5x	log(5.0) = 0.699
10x	log(10.0) = 1.00
20x	log(20.0) = 1.301
100x	log(100.0) = 2.0

σ^*, the standard deviation for the percentage that must meet the acceptable range is $\sigma^* = D^* / z_{freq}$ where z_{freq} is the value corresponding to the upper and lower limits of the range containing the area under the curve (Table 1: Normal area curves in stat book)

2x	99%	$0.301/z_{0.01} = 0.301/2.576 = 0.1168$
	95%	$0.301/z_{0.05} = 0.301/1.96 = 0.1536$
	90%	$0.301/z_{0.10} = 0.301/1.645 = 0.1830$
	80%	$0.301/z_{0.20} = 0.301/1.285 = 0.2342$
5x	99%	$0.699/ z_{0.01} = 0.2714$
	95%	$0.699/ z_{0.05} = 0.3566$
	90%	$0.699/ z_{0.10} = 0.4249$
	80%	$0.699/ z_{0.20} = 0.5440$
10x	99%	$1.0/ z_{0.01} = 0.3882$
	95%	$1.0/ z_{0.05} = 0.5102$
	90%	$1.0/ z_{0.10} = 0.6079$
	80%	$1.0/ z_{0.20} = 0.7782$
20x	99%	$1.301/ z_{0.01} = 0.5050$
	95%	$1.301/ z_{0.05} = 0.6638$
	90%	$1.301/ z_{0.10} = 0.7909$
	80%	$1.301/ z_{0.20} = 1.0125$
100x	99%	$2.0/ z_{0.01} = 0.7764$
	95%	$2.0/ z_{0.05} = 1.0204$
	90%	$2.0/ z_{0.10} = 1.2158$
	80%	$2.0/ z_{0.20} = 1.5564$

Step 3:

Determine the minimum sampling size requirements needed to control risks at acceptable levels and to provide sufficient data for tests of underlying assumptions.

$N_t = (t_\alpha + t_\beta)^2 \cdot \frac{\sigma^2}{\delta^2}$ where $\delta = z_{freq} \cdot \sigma$ (use Table 2-upper tail areas for the normal curve)

and the degrees of freedom are calculated: *N* - 1 where $N = (U_\alpha + U_\beta)^2 \cdot \frac{\sigma^2}{\delta^2}$

% that must meet difference criteria		user/modeler error	N	df
99%	($z_{0.01}$=2.326)	1%	$N = (2.326+2.326)^2/2.326^2 = 4$	3
		2%	$N = (2.054+2.054)^2/2.326^2 = 3.2$	3
		5%	$N = (1.645+1.645)^2/2.326^2 = 2$	1
95%	($z_{0.05}$=1.645)	1%	$N = (2.326+2.326)^2/1.645^2 = 8$	7
		2%	$N = (2.054+2.054)^2/1.645^2 = 6.2$	6
		5%	$N = (1.645+1.645)^2/1.645^2 = 4$	3

90%	($z_{0.10}$=1.282)	1%	$N = (2.326+2.326)^2/1.282^2 = 13.2$	13
		2%	$N = (2.054+2.054)^2/1.282^2 = 10.3$	10
		5%	$N = (1.645+1.645)^2/1.282^2 = 6.6$	6
80%	($z_{0.20}$=0.842)	1%	$N = (2.326+2.326)^2/0.842^2 = 30.5$	30
		2%	$N = (2.054+2.054)^2/0.842^2 = 23.8$	23
		5%	$N = (1.645+1.645)^2/0.842^2 = 15.3$	15

% that must meet difference criteria		user/modeler error	df		N_t
99%	($z_{0.01}$=2.326)	1%	3	$t_{0.01,3} = 4.541$	$(4.541+4.541)^2/2.326^2 = 16$
		2%	3	$t_{0.02,3} = 3.635$	$(3.635+3.635)^2/2.326^2 = 10$
		5%	1	$t_{0.05,1} = 6.314$	$(6.314+6.314)^2/2.326^2 = 30$
95%	($z_{0.05}$=1.645)	1%	7	$t_{0.01,7} = 2.998$	$(2.998+2.998)^2/1.645^2 = 14$
		2%	6	$t_{0.02,6} = 2.679$	$(2.679+2.679)^2/1.645^2 = 11$
		5%	3	$t_{0.05,3} = 2.353$	$(2.353+2.353)^2/1.645^2 = 9$
90%	($z_{0.10}$=1.282)	1%	13	$t_{0.01,13} = 2.650$	$(2.650+2.650)^2/1.282^2 = 18$
		2%	10	$t_{0.02,10} = 2.407$	$(2.407+2.407)^2/1.282^2 = 15$
		5%	6	$t_{0.05,6} = 1.943$	$(1.943+1.943)^2/1.282^2 = 10$
80%	($z_{0.20}$=0.842)	1%	30	$t_{0.01,30} = 2.457$	$(2.457+2.457)^2/0.842^2 = 35$
		2%	23	$t_{0.02,23} = 2.213$	$(2.213+2.213)^2/0.842^2 = 28$
		5%	15	$t_{0.05,15} = 1.753$	$(1.753+1.753)^2/0.842^2 = 18$

1992 Daily Flow Values-Calibrated Run

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .768$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{366-3}}{2} \cdot \ln\left(\frac{1.768}{.232}\right) = 19.35$$

$$z_{0.01} = 2.326$$

since $z > z_{0.01}$, reject null hypothesis

conclusion: model predictions and observations are positively correlated at all test levels

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.044, σ=0.2836)*	Expected Prop	Expected Counts	Obs Counts
< -.2354	25%	91.5	100
-.2354 to -.044	25%	91.5	73
-.044 to .1474	25%	91.5	92
> .1474	25%	91.5	101

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where z = 0.675 for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = .7896 + 3.7404 + .0055 + .9863 = 5.5218$$

4 categories - 2 parameters -1=> 1 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

5.52 > 2.70 so ***D* is not normally distributed** (same result with 30 categories)

1993-1995 Daily Flow Values-after 1992 Flow Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .865$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{1095-3}}{2} \cdot \ln\left(\frac{1.865}{.135}\right) = 43.38$$

$$z_{0.01} = 2.326$$

since $z > z_{0.01}$, reject null hypothesis

conclusion: model predictions and observations are positively correlated at all test levels

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.1931, σ=0.4852)*	Expected Prop	Expected Counts	Obs Counts
< -.5206	25%	273.75	245
-.5206 to -.1931	25%	273.75	231
-.1931 to .1344	25%	273.75	379
> .1344	25%	273.75	240

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where *z* = 0.675 for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = 3.0194 + 6.6760 + 40.4660 + 4.1610 = 54.3224$$

4 categories - 2 parameters -1 => 1 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

54.3224 > 2.70 so ***D* is not normally distributed**

1992 Atrazine Concentration Values-after Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .460$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{21-3}}{2} \cdot \ln\left(\frac{1.460}{.540}\right) = 2.11$$

$z_{0.01} = 2.326$

$z_{0.02} = 2.054$

$z_{0.05} = 1.645$

$z < z_{0.01}$ so accept null hypothesis at 1% user error level

$z > z_{0.02}$ reject null hypothesis at 2% user error level

conclusion: model predictions and observations are positively correlated at user error levels of 2% or greater. At user error level of 1%, model predictions and observations are not positively correlated.

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.3959, σ=0.7145)*	Expected Prop	Expected Counts	Obs Counts
< -.8782	25%	5.25	6
-.8782 to -.3959	25%	5.25	3
-.3959 to .0864	25%	5.25	6
> .0864	25%	5.25	6

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where *z* = 0.675 for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = .1071 + .9643 + .1071 + .1071 = 1.2856$$

4 categories - 2 parameters -1 => 1 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

1.29 < 2.70 so ***D* is normally distributed**

Step 7:

Test the null hypothesis

H_0: $\sigma = \sigma^*$, the model is unacceptably imprecise, vs. the one-sided composite alternative

H_a: $\sigma < \sigma^*$, the model precision is adequate

where σ^* is the critical maximum permissible variance

Reject the null hypothesis if the computed χ^2 statistic is less than the value of χ^2 for the (1-α) level of significance with $v = n$-1 degrees of freedom

2x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.1168^2} = 748.4$	$\chi^2_{.99,20} = 37.6$	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.1536^2} = 432.8$	$\chi^2_{.95,20} = 31.4$	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.1830^2} = 304.9$	$\chi^2_{.90,20} = 28.4$	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.2342^2} = 186.1$	$\chi^2_{.80,20} = 25.4$	precision inadequate
5x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.2714^2} = 138.6$	$\chi^2_{.99,20} = 37.6$	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.3566^2} = 80.3$	$\chi^2_{.95,20} = 31.4$	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.4249^2} = 56.6$	$\chi^2_{.90,20} = 28.4$	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.5440^2} = 34.5$	$\chi^2_{.80,20} = 25.4$	precision inadequate
10x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.3882^2} = 67.8$	$\chi^2_{.99,20} = 37.6$	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.5102^2} = 39.2$	$\chi^2_{.95,20} = 31.4$	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.6079^2} = 27.6$	$\chi^2_{.90,20} = 28.4$	**precision adequate**
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.7782^2} = 16.9$	$\chi^2_{.80,20} = 25.4$	**precision adequate**

20x 99% $\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.5050^2} = 40.0$ $\chi^2_{.99,20}$ = 37.6 precision inadequate

95% $\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.6638^2} = 23.2$ $\chi^2_{.95,20}$ = 31.4 **precision adequate**

90% $\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{0.7909^2} = 16.3$ $\chi^2_{.90,20}$ = 28.4 **precision adequate**

80% $\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .7145^2}{1.0125^2} = 10.0$ $\chi^2_{.80,20}$ = 25.4 **precision adequate**

precision adequate at 100x for all levels of frequency meeting range.

Step 8:

Test the full validity of the simulation model by comparison of the computed value of the *t*-statistic to the appropriate (α, *v*) point on the *t* distribution, *v* degrees of freedom.

since the mean of *D* is < 0

H_0: $\mu_D = (-D^* + \delta)$

H_a: $\mu_D > (-D^* + \delta)$

where $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}}$ (*P* is the level of precision)

when $\overline{D} < 0.0$, the sample test statistic is computed

$$t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right)\cdot \sqrt{n}}{S}$$

and H_0 is rejected if $t > t_c$ where $t_c = t_{\alpha,n-1}$ where α is the level of user risk. (want to reject)

values for δ

99% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{2.528 \cdot 0.7145}{\sqrt{21}} = .3942$

95% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.725 \cdot 0.7145}{\sqrt{21}} = .2690$

90% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.325 \cdot 0.7145}{\sqrt{21}} = .2066$

80% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{0.8997 \cdot 0.7145}{\sqrt{21}} = .1403$

2x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.301 + .3942\right)\right) \cdot \sqrt{21}}{0.7145} = -3.1369$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.301 + .2690\right)\right) \cdot \sqrt{21}}{0.7145} = -2.3339$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.301 + .2066\right)\right) \cdot \sqrt{21}}{0.7145} = -1.9337$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.301 + .1403\right)\right) \cdot \sqrt{21}}{0.7145} = -1.5085$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

5x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.699 + .3942\right)\right) \cdot \sqrt{21}}{0.7145} = -0.5843$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.699 + .2690\right)\right) \cdot \sqrt{21}}{0.7145} = 0.2187$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.699 + .2066\right)\right) \cdot \sqrt{21}}{0.7145} = 0.6189$

user error = 1% $t_{0.01,20} = 2.528$
user error = 2% $t_{0.02,20} = 2.233$
user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-0.699 + .1403\right)\right) \cdot \sqrt{21}}{0.7145} = 1.0441$

user error = 1% $t_{0.01,20}$ = 2.528
user error = 2% $t_{0.02,20}$ = 2.233
user error = 5% $t_{0.05,20}$ = 1.725

$t < t_c$ in all cases: **model is not valid**

10x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-1.00 + .3942\right)\right) \cdot \sqrt{21}}{0.7145} = 1.3462$

user error = 1% $t_{0.01,20}$ = 2.528
user error = 2% $t_{0.02,20}$ = 2.233
user error = 5% $t_{0.05,20}$ = 1.725

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-1.00 + .2690\right)\right) \cdot \sqrt{21}}{0.7145} = 2.1492$

user error = 1% $t_{0.01,20}$ = 2.528 not valid
user error = 2% $t_{0.02,20}$ = 2.233 not valid
user error = 5% $t_{0.05,20}$ = 1.725 **valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-1.00 + .2066\right)\right) \cdot \sqrt{21}}{0.7145} = 2.5494$

user error = 1% $t_{0.01,20}$ = 2.528
user error = 2% $t_{0.02,20}$ = 2.233
user error = 5% $t_{0.05,20}$ = 1.725

$t > t_c$ in all cases: **model is valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3959 - \left(-1.00 + .1403\right)\right) \cdot \sqrt{21}}{0.7145} = 2.9747$

user error = 1% $t_{0.01,20}$ = 2.528
user error = 2% $t_{0.02,20}$ = 2.233
user error = 5% $t_{0.05,20}$ = 1.725

$t > t_c$ in all cases: **model is valid**

Model is valid for all levels of precision and user error at 20x and 100x

1993-1995 Atrazine Concentration Values-after 1992 Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .640$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{65-3}}{2} \cdot \ln\left(\frac{1.640}{.360}\right) = 5.970$$

$z_{0.01} = 2.326$

$z > z_{0.01}$ so reject null hypothesis at 1% user error level

conclusion: model predictions and observations are positively correlated at all user error levels

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.3205, σ=1.3110)*	Expected Prop	Expected Counts	Obs Counts
< -1.2054	25%	16.25	13
-1.2054 to -.3205	25%	16.25	16
-.3205 to .5644	25%	16.25	17
> .5644	25%	16.25	19

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where *z* = 0.675 for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = .6500 + .0038 + .0346 + .4654 = 1.1538$$

4 categories => 3 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

1.15 < 2.70 so ***D* is normally distributed**

Step 7:

Test the null hypothesis

H_0: $\sigma = \sigma^*$, the model is unacceptably imprecise, vs. the one-sided composite alternative

H_a: $\sigma < \sigma^*$, the model precision is adequate

where σ^* is the critical maximum permissible variance

Reject the null hypothesis if the computed χ^2 statistic is less than the value of χ^2 for the (1-α) level of significance with $v = n$-1 degrees of freedom

2x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.1168^2} = 8063.1$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.1536^2} = 4662.3$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.1830^2} = 3284.6$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.2342^2} = 2005.4$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
5x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.2714^2} = 1493.4$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.3566^2} = 865.0$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.4249^2} = 609.3$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.5440^2} = 371.7$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
10x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.3882^2} = 729.9$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.5102^2} = 422.6$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.6079^2} = 297.7$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.7782^2} = 110.0$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
20x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.5050^2} = 431.3$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.6638^2} = 249.6$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.7909^2} = 175.8$	$\chi^2_{.90,64}$ = 78.8	precision inadequate

	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{1.0125^2} = 107.3$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
100x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{0.7764^2} = 182.5$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{1.0204^2} = 105.6$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{1.2158^2} = 74.4$	$\chi^2_{.90,64}$ = 78.8	**precision adequate**
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.311^2}{1.5564^2} = 45.4$	$\chi^2_{.80,64}$ = 73.8	**precision adequate**

Step 8:
Test the full validity of the simulation model by comparison of the computed value of the *t*-statistic to the appropriate (α, *v*) point on the *t* distribution, *v* degrees of freedom.

since the mean of *D* is < 0
H_0: $\mu_D = (-D^* + \delta)$
H_a: $\mu_D > (-D^* + \delta)$

where $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}}$ (*P* is the level of precision)
when $\overline{D} < 0.0$, the sample test statistic is computed
$t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S}$
and H_0 is rejected if $t > t_c$ where $t_c = t_{\alpha,n-1}$ where α is the level of user risk. (want to reject)

values for δ

99% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{2.388 \cdot 1.311}{\sqrt{65}} = .3883$

95% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.670 \cdot 1.311}{\sqrt{65}} = .2716$

90% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.296 \cdot 1.311}{\sqrt{65}} = .2107$

80% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{0.885 \cdot 1.311}{\sqrt{65}} = .1439$

2x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.301 + 0.3883\right)\right) \cdot \sqrt{65}}{1.311} = -2.5078$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.301 + .2716\right)\right) \cdot \sqrt{65}}{1.311} = -1.7902$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.301 + .2107\right)\right) \cdot \sqrt{65}}{1.311} = -1.4157$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.301 + .1439\right)\right) \cdot \sqrt{65}}{1.311} = -1.0049$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

5x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.699 + .3883\right)\right) \cdot \sqrt{65}}{1.311} = -0.0603$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.699 + .2716\right)\right) \cdot \sqrt{65}}{1.311} = 0.6574$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.699 + .2107\right)\right) \cdot \sqrt{65}}{1.311} = 1.0319$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-0.699 + .1439\right)\right) \cdot \sqrt{65}}{1.311} = 1.4427$

user error = 1% $t_{0.01,64} = 2.388$

user error = 2% $t_{0.02,64} = 2.129$

user error = 5% $t_{0.05,64} = 1.670$

$t < t_c$ in all cases: **model is not valid**

10x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-1.00 + .3883\right)\right) \cdot \sqrt{65}}{1.311} = 1.7908$

user error = 1% $t_{0.01,64} = 2.388$ not valid

user error = 2% $t_{0.02,64} = 2.129$ not valid

user error = 5% $t_{0.05,64} = 1.670$ **valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-1.00 + .2716\right)\right) \cdot \sqrt{65}}{1.311} = 2.5085$

user error = 1% $t_{0.01,64} = 2.388$

user error = 2% $t_{0.02,64} = 2.129$

user error = 5% $t_{0.05,64} = 1.670$

$t > t_c$ in all cases: **model is valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-1.00 + .2107\right)\right) \cdot \sqrt{65}}{1.311} = 2.8830$

user error = 1% $t_{0.01,64} = 2.388$

user error = 2% $t_{0.02,64} = 2.129$

user error = 5% $t_{0.05,64} = 1.670$

$t > t_c$ in all cases: **model is valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3205 - \left(-1.00 + .1439\right)\right) \cdot \sqrt{65}}{1.311} = 3.2938$

user error = 1% $t_{0.01,64} = 2.388$

user error = 2% $t_{0.02,64} = 2.129$

user error = 5% $t_{0.05,64} = 1.670$

$t > t_c$ in all cases: **model is valid**

Model is valid for all levels of precision and user error at 20x and 100x

1992 Metolachlor Concentration Values-after Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .641$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{21-3}}{2} \cdot \ln\left(\frac{1.641}{.359}\right) = 3.2239$$

$z_{0.01} = 2.326$

$z > z_{0.01}$ so reject null hypothesis at 1% user error level

conclusion: model predictions and observations are positively correlated at user error levels of 1% or greater.

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.3593, σ=0.9218)*	Expected Prop	Expected Counts	Obs Counts
< -.9815	25%	5.25	5
-.9815 to -.3593	25%	5.25	5
-.3593 to .2629	25%	5.25	6
> .2629	25%	5.25	5

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where *z* = 0.675 for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = .0119 + .0119 + .1071 + .0119 = .1428$$

4 categories - 2 parameters -1 => 1 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

0.14 < 2.70 so ***D* is normally distributed**

Step 7:

Test the null hypothesis

H_0: $\sigma = \sigma^*$, the model is unacceptably imprecise, vs. the one-sided composite alternative

H_a: $\sigma < \sigma^*$, the model precision is adequate

where σ^* is the critical maximum permissible variance

Reject the null hypothesis if the computed χ^2 statistic is less than the value of χ^2 for the (1-α) level of significance with $v = n$-1 degrees of freedom

2x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.1168^2} = 1245.7$	$\chi^2_{.99,20}$ = 37.6	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.1536^2} = 720.3$	$\chi^2_{.95,20}$ = 31.4	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.1830^2} = 507.5$	$\chi^2_{.90,20}$ = 28.4	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.2342^2} = 309.8$	$\chi^2_{.80,20}$ = 25.4	precision inadequate
5x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.2714^2} = 230.7$	$\chi^2_{.99,20}$ = 37.6	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.3566^2} = 133.6$	$\chi^2_{.95,20}$ = 31.4	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.4249^2} = 94.1$	$\chi^2_{.90,20}$ = 28.4	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.5440^2} = 57.4$	$\chi^2_{.80,20}$ = 25.4	precision inadequate
10x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.3882^2} = 112.8$	$\chi^2_{.99,20}$ = 37.6	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.5102^2} = 65.3$	$\chi^2_{.95,20}$ = 31.4	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.6079^2} = 46.0$	$\chi^2_{.90,20}$ = 28.4	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.7782^2} = 28.1$	$\chi^2_{.80,20}$ = 25.4	precision inadequate
20x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.5050^2} = 66.6$	$\chi^2_{.99,20}$ = 37.6	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.6638^2} = 38.6$	$\chi^2_{.95,20}$ = 31.4	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{0.7909^2} = 27.2$	$\chi^2_{.90,20}$ = 28.4	**precision adequate**

80% $\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{20 \cdot .9218^2}{1.0125^2} = 16.6$ $\chi^2_{.80,20}$ = 25.4 **precision adequate**

precision adequate at 100x for all levels of frequency meeting range.

Step 8:

Test the full validity of the simulation model by comparison of the computed value of the *t*-statistic to the appropriate (α, *v*) point on the *t* distribution, *v* degrees of freedom.

since the mean of *D* is < 0

H_0: $\mu_D = (-D^* + \delta)$

H_a: $\mu_D > (-D^* + \delta)$

where $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}}$ (*P* is the level of precision)

when $\overline{D} < 0.0$, the sample test statistic is computed

$$t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right)\cdot \sqrt{n}}{S}$$

and H_0 is rejected if $t > t_c$ where $t_c = t_{\alpha,n-1}$ where α is the level of user risk. (want to reject)

values for δ

99% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{2.528 \cdot 0.9218}{\sqrt{21}} = .5085$

95% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.725 \cdot 0.9218}{\sqrt{21}} = .3470$

90% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.325 \cdot 0.9218}{\sqrt{21}} = .2665$

80% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{0.8997 \cdot 0.9218}{\sqrt{21}} = .1810$

2x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right)\cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-0.301 + .5085\right)\right)\cdot \sqrt{21}}{0.9218} = -2.8178$

user error = 1% $t_{0.01,20}$ = 2.528

user error = 2% $t_{0.02,20}$ = 2.233

user error = 5% $t_{0.05,20}$ = 1.725

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.301 + .3470)) \cdot \sqrt{21}}{0.9218} = -2.0149$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.301 + .2665)) \cdot \sqrt{21}}{0.9218} = -1.6147$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.301 + .1810)) \cdot \sqrt{21}}{0.9218} = -1.1896$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

5x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.699 + .5085)) \cdot \sqrt{21}}{0.9218} = -0.8392$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.699 + .3470)) \cdot \sqrt{21}}{0.9218} = -0.0363$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.699 + .2665)) \cdot \sqrt{21}}{0.9218} = 0.3639$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{(-0.3593 - (-0.699 + .1810)) \cdot \sqrt{21}}{0.9218} = 0.7890$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t < t_c$ in all cases: **model is not valid**

10x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.00 + .5085\right)\right) \cdot \sqrt{21}}{0.9218} = 0.6572$

user error = 1% $t_{0.01,20}$ = 2.528

user error = 2% $t_{0.02,20}$ = 2.233

user error = 5% $t_{0.05,20}$ = 1.725

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.00 + .3470\right)\right) \cdot \sqrt{21}}{0.9218} = 1.4601$

user error = 1% $t_{0.01,20}$ = 2.528

user error = 2% $t_{0.02,20}$ = 2.233

user error = 5% $t_{0.05,20}$ = 1.725

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.00 + .2665\right)\right) \cdot \sqrt{21}}{0.9218} = 1.8603$

user error = 1% $t_{0.01,20}$ = 2.528 not valid

user error = 2% $t_{0.02,20}$ = 2.233 not valid

user error = 5% $t_{0.05,20}$ = 1.725 **valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.00 + .1810\right)\right) \cdot \sqrt{21}}{0.9218} = 2.2853$

user error = 1% $t_{0.01,20}$ = 2.528 not valid

user error = 2% $t_{0.02,20}$ = 2.233 **valid**

user error = 5% $t_{0.05,20}$ = 1.725 **valid**

20x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.301 + .5085\right)\right) \cdot \sqrt{21}}{0.9218} = 2.1536$

user error = 1% $t_{0.01,20}$ = 2.528 not valid

user error = 2% $t_{0.02,20}$ = 2.233 not valid

user error = 5% $t_{0.05,20}$ = 1.725 **valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.301 + .3470\right)\right) \cdot \sqrt{21}}{0.9218} = 2.9565$

user error = 1% $t_{0.01,20}$ = 2.528

user error = 2% $t_{0.02,20}$ = 2.233

user error = 5% $t_{0.05,20}$ = 1.725

$t > t_c$ in all cases: **model is valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.301 + .2665\right)\right) \cdot \sqrt{21}}{0.9218} = 3.3566$

user error = 1% $t_{0.01,20}$ = 2.528

user error = 2% $t_{0.02,20}$ = 2.233

user error = 5% $t_{0.05,20}$ = 1.725

$t > t_c$ in all cases: **model is valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.3593 - \left(-1.301 + .1810\right)\right) \cdot \sqrt{21}}{0.9218} = 3.7817$

user error = 1% $t_{0.01,20} = 2.528$

user error = 2% $t_{0.02,20} = 2.233$

user error = 5% $t_{0.05,20} = 1.725$

$t > t_c$ in all cases: **model is valid**

Model is valid for all levels of precision and user error at 100x

1993-1995 Metolachlor Concentration Values-after 1992 Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .529$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{65-3}}{2} \cdot \ln\left(\frac{1.529}{.471}\right) = 4.6359$$

$z_{0.01} = 2.326$

$z > z_{0.01}$ so reject null hypothesis at 1% user error level

conclusion: model predictions and observations are positively correlated at all user error levels

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use χ^2 test of normality for error measure *D*

Intervals of 25% on *N(D; μ=-0.069, σ=1.2676)*	Expected Prop	Expected Counts	Obs Counts
< -.9246	25%	16.25	14
-.9246 to -.069	25%	16.25	13
-.069 to .7866	25%	16.25	21
> .7866	25%	16.25	17

ranges calculated with equation $z = \frac{y-\mu}{\sigma}$ where $z = 0.675$ for 25% area

$$\chi^2 = \sum \frac{(O-E)^2}{E} = 0.3115 + 0.6500 + 1.3885 + 0.0346 = 2.3846$$

4 categories => 3 degrees of freedom

$\chi^2_{0.10,1} = 2.70$

2.38 < 2.70 so ***D* is normally distributed**

Step 7:

Test the null hypothesis

H_0: $\sigma = \sigma^*$, the model is unacceptably imprecise, vs. the one-sided composite alternative

H_a: $\sigma < \sigma^*$, the model precision is adequate

where σ^* is the critical maximum permissible variance

Reject the null hypothesis if the computed χ^2 statistic is less than the value of χ^2 for the (1-α) level of significance with $v = n$-1 degrees of freedom

2x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.1168^2} = 7538.0$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.1536^2} = 4358.8$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.1830^2} = 3070.7$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.2342^2} = 1874.9$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
5x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.2714^2} = 1396.1$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.3566^2} = 808.7$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.4249^2} = 569.6$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.5440^2} = 347.5$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
10x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.3882^2} = 682.4$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.5102^2} = 395.1$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.6079^2} = 278.3$	$\chi^2_{.90,64}$ = 78.8	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.7782^2} = 169.8$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
20x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.5050^2} = 403.2$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.6638^2} = 233.4$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.7909^2} = 164.4$	$\chi^2_{.90,64}$ = 78.8	precision inadequate

	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{1.0125^2} = 100.3$	$\chi^2_{.80,64}$ = 73.8	precision inadequate
100x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{0.7764^2} = 170.6$	$\chi^2_{.99,64}$ = 93.2	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{1.0204^2} = 98.8$	$\chi^2_{.95,64}$ = 83.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{1.2158^2} = 69.6$	$\chi^2_{.90,64}$ = 78.8	**precision adequate**
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{64 \cdot 1.2676^2}{1.5564^2} = 42.5$	$\chi^2_{.80,64}$ = 73.8	**precision adequate**

Step 8:
Test the full validity of the simulation model by comparison of the computed value of the *t*-statistic to the appropriate (α, *v*) point on the *t* distribution, *v* degrees of freedom.

since the mean of *D* is < 0
H_0: $\mu_D = (-D^* + \delta)$
H_a: $\mu_D > (-D^* + \delta)$

where $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}}$ (*P* is the level of precision)
when $\overline{D} < 0.0$, the sample test statistic is computed
$t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S}$
and H_0 is rejected if $t > t_c$ where $t_c = t_{\alpha,n-1}$ where α is the level of user risk. (want to reject)

values for δ
99% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{2.388 \cdot 1.2676}{\sqrt{65}} = ..3755$
95% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.670 \cdot 1.2676}{\sqrt{65}} = .2626$
90% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.296 \cdot 1.2676}{\sqrt{65}} = .2038$
80% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{0.885 \cdot 1.2676}{\sqrt{65}} = .1391$

2x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.301 + 0.3755\right)\right) \cdot \sqrt{65}}{1.2676} = -0.9127$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.301 + .2626\right)\right) \cdot \sqrt{65}}{1.2676} = -0.2264$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.301 + .2038\right)\right) \cdot \sqrt{65}}{1.2676} = 0.1794$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.301 + .1391\right)\right) \cdot \sqrt{65}}{1.2676} = 0.5909$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

5x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.699 + .3755\right)\right) \cdot \sqrt{65}}{1.2676} = 1.6187$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.699 + .2626\right)\right) \cdot \sqrt{65}}{1.2676} = 2.3368$

user error = 1% $t_{0.01,64}$ = 2.388 not valid
user error = 2% $t_{0.02,64}$ = 2.129 **valid**
user error = 5% $t_{0.05,64}$ = 1.670 **valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.699 + .2038\right)\right) \cdot \sqrt{65}}{1.2676} = 2.7107$

user error = 1% $t_{0.01,64}$ = 2.388
user error = 2% $t_{0.02,64}$ = 2.129
user error = 5% $t_{0.05,64}$ = 1.670

$t > t_c$ in all cases: **model is valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.069 - \left(-0.699 + .1391\right)\right) \cdot \sqrt{65}}{1.2676} = 3.1222$

user error = 1% $t_{0.01,64} = 2.388$

user error = 2% $t_{0.02,64} = 2.129$

user error = 5% $t_{0.05,64} = 1.670$

$t > t_c$ in all cases: **model is valid**

Model is valid for all levels of precision and user error at 10x, 20x and 100x

1992 Trifluralin Concentration Values-after Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .711$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{15-3}}{2} \cdot \ln\left(\frac{1.711}{.289}\right) = 3.0803$$

$z_{0.01} = 2.326$

$z > z_{0.01}$ so reject null hypothesis at 1% user error level

conclusion: model predictions and observations are positively correlated at user error levels of 1% or greater.

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use Shapiro-Wilk test of normality for error measure *D*

n = 15, W = 0.961, significance = 0.673

significance level is greater than 0.10 so **D is normally distributed**

Step 7:

Test the null hypothesis

H_0: $\sigma = \sigma^*$, the model is unacceptably imprecise, vs. the one-sided composite alternative

H_a: $\sigma < \sigma^*$, the model precision is adequate

where σ^* is the critical maximum permissible variance

Reject the null hypothesis if the computed χ^2 statistic is less than the value of χ^2 for the (1-α) level of significance with $v = n$-1 degrees of freedom

2x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.1168^2} = 982.4$	$\chi^2_{.99,14}$ = 29.1	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.1536^2} = 568.0$	$\chi^2_{.95,14}$ = 23.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.1830^2} = 400.2$	$\chi^2_{.90,14}$ = 21.1	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.2342^2} = 244.3$	$\chi^2_{.80,14}$ = 18.4	precision inadequate
5x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.2714^2} = 181.9$	$\chi^2_{.99,14}$ = 29.1	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.3566^2} = 105.4$	$\chi^2_{.95,14}$ = 23.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.4249^2} = 74.2$	$\chi^2_{.90,14}$ = 21.1	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.5440^2} = 45.3$	$\chi^2_{.80,14}$ = 18.4	precision inadequate
10x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.3882^2} = 88.9$	$\chi^2_{.99,14}$ = 29.1	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.5102^2} = 51.5$	$\chi^2_{.95,14}$ = 23.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.6079^2} = 36.3$	$\chi^2_{.90,14}$ = 21.1	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.7782^2} = 22.1$	$\chi^2_{.80,14}$ = 18.4	precision inadequate
20x	99%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.5050^2} = 52.6$	$\chi^2_{.99,14}$ = 29.1	precision inadequate
	95%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.6638^2} = 30.4$	$\chi^2_{.95,14}$ = 23.7	precision inadequate
	90%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{0.7909^2} = 21.4$	$\chi^2_{.90,14}$ = 21.1	precision inadequate
	80%	$\chi^2 = \frac{(n-1)\cdot S^2}{\sigma^2} = \frac{14 \cdot .9784^2}{1.0125^2} = 13.1$	$\chi^2_{.80,14}$ = 18.4	**precision adequate**

precision adequate at 100x for all levels of frequency meeting range.

Step 8:

Test the full validity of the simulation model by comparison of the computed value of the *t*-statistic to the appropriate (α, *v*) point on the *t* distribution, *v* degrees of freedom.

since the mean of *D* is < 0

H_0: $\mu_D = (-D^* + \delta)$

H_a: $\mu_D > (-D^* + \delta)$

where $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}}$ (P is the level of precision)

when $\overline{D} < 0.0$, the sample test statistic is computed

$$t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S}$$

and H_0 is rejected if $t > t_c$ where $t_c = t_{\alpha,n-1}$ where α is the level of user risk. (want to reject)

values for δ

99% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{2.624 \cdot 0.9784}{\sqrt{15}} = .6629$

95% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.761 \cdot 0.9784}{\sqrt{15}} = .4449$

90% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{1.345 \cdot 0.9784}{\sqrt{15}} = .3398$

80% $\delta = \frac{t_{P,v} \cdot S}{\sqrt{n}} = \frac{0.9097 \cdot 0.9784}{\sqrt{15}} = .2298$

2x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.301 + .6629\right)\right) \cdot \sqrt{15}}{0.9784} = -3.6200$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.301 + .4449\right)\right) \cdot \sqrt{15}}{0.9784} = -2.7571$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.301 + .3398\right)\right) \cdot \sqrt{15}}{0.9784} = -2.3410$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.301 + .2298\right)\right) \cdot \sqrt{15}}{0.9784} = -1.9056$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

5x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.699 + .6629\right)\right) \cdot \sqrt{15}}{0.9784} = -2.0446$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.699 + .4449\right)\right) \cdot \sqrt{15}}{0.9784} = -1.1816$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.699 + .3398\right)\right) \cdot \sqrt{15}}{0.9784} = -0.7656$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-0.699 + .2298\right)\right) \cdot \sqrt{15}}{0.9784} = -0.3301$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

10x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.00 + .6629\right)\right) \cdot \sqrt{15}}{0.9784} = -0.8531$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.00 + .4449\right)\right) \cdot \sqrt{15}}{0.9784} = 0.0099$

user error = 1% $t_{0.01,14} = 2.624$

user error = 2% $t_{0.02,14} = 2.305$

user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.00 + .3398\right)\right) \cdot \sqrt{15}}{0.9784} = 0.4259$

user error = 1% $t_{0.01,14} = 2.624$
user error = 2% $t_{0.02,14} = 2.305$
user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.00 + .2298\right)\right) \cdot \sqrt{15}}{0.9784} = 0.8614$

user error = 1% $t_{0.01,14} = 2.624$
user error = 2% $t_{0.02,14} = 2.305$
user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

20x 99% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.301 + .6629\right)\right) \cdot \sqrt{15}}{0.9784} = 0.3385$

user error = 1% $t_{0.01,14} = 2.624$
user error = 2% $t_{0.02,14} = 2.305$
user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

95% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.301 + .4449\right)\right) \cdot \sqrt{15}}{0.9784} = 1.2014$

user error = 1% $t_{0.01,14} = 2.624$
user error = 2% $t_{0.02,14} = 2.305$
user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

90% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.301 + .3398\right)\right) \cdot \sqrt{15}}{0.9784} = 1.6174$

user error = 1% $t_{0.01,14} = 2.624$
user error = 2% $t_{0.02,14} = 2.305$
user error = 5% $t_{0.05,14} = 1.761$

$t < t_c$ in all cases: **model is not valid**

80% $t = \frac{\left(\overline{D} - \left(-D^* + \delta\right)\right) \cdot \sqrt{n}}{S} = \frac{\left(-0.5526 - \left(-1.301 + .2298\right)\right) \cdot \sqrt{15}}{0.9784} = 2.0529$

user error = 1% $t_{0.01,14} = 2.624$ not valid
user error = 2% $t_{0.02,14} = 2.305$ not valid
user error = 5% $t_{0.05,14} = 1.761$ **valid**

Model is valid for all levels of precision and user error at 100x

1993-1995 Trifluralin Concentration Values-after Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .138$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{20-3}}{2} \cdot \ln\left(\frac{1.138}{.862}\right) = 0.5726$$

$z_{0.01} = 2.326$

$z_{0.02} = 2.054$

$z_{0.05} = 1.645$

$z < z_{0.01}, z_{0.02}, z_{0.05}$ so cannot reject null hypothesis at user error level

conclusion: model predictions and observations are not positively correlated at user error levels of 1%, 2%, or 5%. model predictions and observations are positively correlated at user error level of 28% or greater.

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use Shapiro-Wilk test of normality for error measure *D*

n = 20, W = 0.878, significance = 0.016

significance level is less than 0.10 so **D is not normally distributed**

1992-1995 Trifluralin Concentration Values-after Calibration

Step 4:

Collect a set of paired samples of simulation model predictions and observations

Step 5:

Test the paired data for significant correlation; reject model if they are not positively correlated

H_0: $\rho = 0$ vs H_a: $\rho > 0$

where ρ is the population correlation coefficient and testing is conducted at a significance level of α where α is the user error

Correlation of the O and P datasets yields a correlation coefficient value $r = .657$

The z statistic associated with this value is

$$z = \frac{\sqrt{n-3}}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) = \frac{\sqrt{35-3}}{2} \cdot \ln\left(\frac{1.657}{.343}\right) = 4.4549$$

$z_{0.01} = 2.326$

$z > z_{0.01}$ so reject null hypothesis at 1% user error level

conclusion: model predictions and observations are positively correlated at user error levels of 1% or greater.

Step 6:

Test that the error measure, *D*, is normal at a level of significance suitably less restrictive that the α risk level in use to control user's risk.

H_0: *D* is approximately normal vs

H_a: *D* is not normally distributed

use Shapiro-Wilk test of normality for error measure *D*

n = 35, W = 0.944, significance = 0.099

significance level is less than 0.10 so **D is not normally distributed**

References

Arnold, J.G. and P.M. Allen. 1999. Automated methods for estimating baseflow and groundwater recharge from streamflow records. JAWRA 35(2): 411-424.

ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, Irrigation and Drainage Division. 1993. Criteria for evaluation of watershed models. Journal of Irrigation and Drainage Engineering. 119(3): 429-442.

Atwood, J.D., D.W. Goss, R.L. Kellogg, T.A. Pitts, S.R. Potter, S. Wallace. 2000. The NRCS national nutrient loss modeling project: preliminary results for corn east of the Rocky Mountains. Unpublished.

Chapra, S.C. 1997. Surface water-quality modeling. WCB/McGraw-Hill, Boston, MA.

Danckwerts, P.V. 1951. Significance of liquid-film coefficients in gas absorption. Ind. Eng. Chem. 43:1460-1467.

Higbie, R. 1935. The rate of adsorption of a pure gas into a still liquid during short periods of exposure. Trans. Amer. Inst. Chem. Engin. 31:365-389.

Leonard, R.A., W.G. Knisel., and D.A. Still. 1987. GLEAMS: Groundwater loading effects of agricultural management systems. Trans. ASAE. 30:1403-1418.

Lewis, W.K. and W.G. Whitman. 1924. Principles of gas absorption. Ind. Eng. Chem. 16:1215-1220.

McElroy, A.D., S.Y. Chiu, J.W. Nebgen, A. Aleti, and F.W. Bennett. 1976. Loading functions for assessment of water pollution from nonpoint sources. Environ. Prot. Tech. Serv., EPA 600/2-76-151.

Menzel, R.G. 1980. Enrichment ratios for water quality modeling. p. 486-492. *In* W.G. Knisel (ed.) CREAMS, A field scale model for chemicals, runoff, and erosion from agricultural management systems. U.S. Dept. Agric. Conserv. Res. Rept. No. 26.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry and J.R. Williams. 2001a. Soil and Water Assessment Tool: Theoretical Documentation Version 2000. available via the web at ftp://ftp.brc.tamus.edu/pub/swat/doc/swat2000theory.pdf.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2001b. Soil and Water Assessment Tool: User's Manual Version 2000. available via the web at ftp://ftp.brc.tamus.edu/pub/swat/doc/swatuserman.pdf.

O'Connor, D.J. and Dobbins, W.E. 1958. Mechanisms of reaeration in natural streams. Trans ASAE 123: 641-666.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, and L. Hauck. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. JAWRA 37(5): 1169-1188.

Whitman, W.G. 1923. The two-film theory of gas adsorption. Chem. Metallurg. Eng. 29:146-148.

Williams, J.R. 1995. Chapter 25: The EPIC model. p. 909-1000. *In* V.P. Singh (ed.). Computer models of watershed hydrology. Water Resources Publications.

Williams, J.R. and R.W. Hann. 1978. Optimal operation of large agricultural watersheds with water quality constraints. Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.