EPA/600/R-18/247

United States
Environmental Protection
Agency

Report

Watershed Hydrologic and Contaminated Sediment Transport Modeling in the Tri-State Mining District

EPA/600/R-18/247

Watershed Hydrologic and Contaminated Sediment Transport Modeling in the Tri State Mining District

By

Kazi Rahman

Mohamed M. Hantush

Alexander Hall, and

John McKernan

U.S. Environmental Protection Agency

Cincinnati, Ohio 45268

National Risk Management Research Laboratory

Office of Research and Development

U.S. Environmental Protection Agency

Cincinnati, Ohio 45268

Notice

The U.S. Environmental Protection Agency through its Office of Research and Development funded the research described here. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document.

This research was supported in part by an appointment to the Post-Doctoral Research Program at the National Risk Management Research Laboratory, administered by the Oak Ridge Institute for Science and Education through Interagency Agreement No. DW8992433001 between the U.S. Department of Energy and the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

The Spring River Watershed drains most of the Tri-State Mining District that includes parts of southeast Kansas, southwest Missouri, and northeast Oklahoma. The mining activity in the Tri-State District has resulted in considerable historical and ongoing input of cadmium, lead, and zinc to the watershed including Empire Lake in Cherokee County, southeast Kansas. The environmental contamination caused by the decades of mining activity resulted in southeast Cherokee County being listed on the U.S. Environmental Protection Agency's National Priority List as a superfund hazardous waste site in 1983. The mining activities in the TSMD led to a number of health and environmental complications including wide-spread contaminated sediment in floodplains and stream beds, elevated blood Pb levels in surrounding residential areas, Zn poisoning in livestock and wild birds, and elevated contaminate levels in fish and aquatic macro invertebrate.

A semi-distributed hydrologic model was constructed and calibrated to predict streamflow and sediment loading in the Spring River Basin that feeds into Empire Lake, KS. The model was calibrated and evaluated using continuous streamflow measurements and biweekly sampled suspended sediment concentrations (SSC) collected over the period 2014-2016. Over all, simulated flow rates and sediment loadings compared well with the observed values. The calibrated watershed model was used to estimate average annual sediment loading from interior sub-basins and determine percentage contribution of each sub-basin to the total average annual sediment loading to Empire Lake. With the result obtained from sediment loading simulations, hypothetical management scenarios of lake-dredging and sediment filtration were evaluated. This study identified interior sub-basins contributing most of the sediment loading to Empire Lake and can be used to inform management decisions on remediation of metal contamination in the Spring River Watershed and Empire Lake.

This report has been subjected to QA/QC review. The report presented a mathematical framework for modeling hydrology of a watershed and sediment transport processes in the Spring River Watershed.

Executive Summary

The Tri-State Mining District (TSMD) encompassing the Kansas, Missouri and Oklahoma conjunction was the center of historic mining activity, ceasing in 1970. Although mining activity ended almost 50 years ago, its legacy as a source of cadmium, lead, and zinc to the environment continues to this day. This mining activity left 165 million tons of improperly contained piles of mine waste (chat) across the 2,500 sq. mile region. Chat piles were the dominant geographic feature in the TSMD, especially in Short Creek, Center Creek, Turkey Creek, and Shoal Creek, among others. These features, along with waste rock and mine tailings, have contributed to metal contamination of the waterways of the Spring River Watershed (located in the TSMD), and led to the transport of heavy metal-laden (primarily zinc and lead) sediments into the Empire Lake Reservoir in Cherokee County, Kansas. Years of sedimentation have reduced the capacity of the reservoir, leading to the pass-through of contaminated sediments - affecting downstream communities and Indian Tribes.

The Soil and Water Assessment Tool (SWAT) was used to construct a distributed watershed model for streamflow and sediment loading simulations in the Spring River Basin watershed that feeds into Empire Lake, KS. The objective of the watershed model simulations is to provide information on sediment transport and loading from source areas needed to support remediation efforts for the Spring River Watershed and Empire Lake. Geospatial and hydro-climate input data resolution analysis was conducted to identify optimal input data resolution for best model performance in simulating flow and sediment transport within the Spring River Watershed. Input data resolution analysis was conducted prior to model calibration to insure optimal watershed model performance. The SWAT hydrologic model was successfully calibrated and validated both at the monthly and daily time scales using streamflow data downloaded from two USGS gauge stations in the watershed. The flow watershed model at the Spring River and Shoal Creek gauges met the threshold performance statistics and explained more than 67% of the variance in the observed data for both the calibration and validation periods. Wet and relatively drier periods were simulated well by the model. The model reproduced observed low and high streamflows adequately but deviated from middle range observed values.

A sediment transport component of SWAT was constructed and calibrated using three years-worth of biweekly flow and suspended sediment concentration data (2014-2016) sampled from stations in seven different tributaries upstream from Empire Lake. Sediment loading was calibrated at Spring River and Shoal Creek. The model met the threshold performance statistics recommended for sediment and explained 92% of the variance in the observed data at the Spring River Watershed. However, sediment calibration at Shoal Creek was not as good as for Spring River, with the model explaining only 58% of the variance in the observed data. Calibration of sediment loading at smaller tributaries of mainstem Spring River produced R^2 ranging from 0.69 to 0.99, thus explaining more than about 70% of the variance in the observed data. Average annual sediment loading in the watershed were estimated for the period (2010-2016) using the calibrated SWAT model, and areas contributing most of the sediments were identified. The two largest sub-basins, the Spring River and Shoal Creek Watersheds, contributed most of the annual sediment loading (74%), with the former known to be associated with relatively cleaner sediments. While tributaries such as Short Creek, Center Creek, and Turkey Creek contributed an estimated 15% of annual sediment loading over the study period, they drained areas that are substantially affected by historical lead and zinc mining.

Dredging of Empire Lake as a potential remedial measure of contaminated sediments was investigated. Calculations based on SWAT simulated sediment loadings and observed sediment data showed that the time required to fill back the reservoir with a dredged lake sediment mass of 2640 million ibs may exceed 100 years and could be even much longer. Mass balance analysis using suspended sediment concentration data sampled directly downstream from Empire lake reservoir and the calibrated SWAT model indicated net sediment accumulation in 2014 and 2016. However, the mass balance analysis pointed toward a substantial amount of sediment being mobilized from Empire Lake in 2015. It remains to be seen if natural weather events and/or planned reservoir operation may have contributed to the calculated lake sediment removal in 2015.

SWAT computed average annual sediment loading for 2014-2016 and reported studies on historical lead and zinc occurrence within the TSMD were used to make qualitative inferences on efficacy of hypothetical sediment traps as a potential remedial strategy for mining-affected tributaries. While installation of sediment traps in Short, Center, and Turkey Creeks may reduce

less than 14% of annual average sediment loading to Empire Lake (based on 2014-2016 data), these tributaries historically have been associated with highest concentrations of dissolved and sediment-bound zinc and lead. Effectiveness of sediment filtration in reducing heavy metals input to Spring River therefore might be limited by the percentage of fine sediment particles and percentage of total metals in dissolved phase.

These results are useful for identifying critical source areas of sediment and can be used to inform management decisions on lake dredging and sediment traps as viable remedial measures for metal contamination in heavily contaminated tributaries of Spring River and Empire Lake.

Keywords: SWAT, Watershed, Modeling, Hydrology, Sediment Transport, Spring River

Table of Contents

Acknowledgments ix

Quality Assurance 1

1. Introduction 2

1.1 Background 2

1.2 Objectives 4

2. Hydrology Model 6

2.1 Study Area 6

2.2 Watershed Model Development 7

2.3 Data and Sources 8

2.4 Model Performance Statistics 12

2.5 Pre-Calibration Analysis: Significance of Input Data 13

2.6 Streamflow Calibration and Validation 14

2.6.1 Streamflow Calibration (2010-2016) 15

2.6.2 Streamflow Validation (2000-2007) 17

3. Sediment Model 22

3.1 Sediment Transport Model Development 22

3.2 Sediment Model Calibration 24

4. Scenario Analysis 29

4.1 Sediment Source Areas and Annual Yield 29

4.2 Assessment of Potential Remedial Strategies 31

4.2.1 Lake Sediment Dredging Scenario 31

4.2.2 Sediment Trapping Scenario (qualitative assessment) 38

5. Summary and Conclusion 40

5 References 43

List of Figures

Figure 1. Study area map and main features of the Spring River Watershed (upstream from Empire Lake) and tributaries. The map shows a geographic overview of the watershed and the Spring River and Shoal Creek Watersheds. Water quality sampling stations are marked with white circles. The two USGS streamflow gauges on the Spring River and Shoal Creek are marked with yellow circles. Water quality was also sampled at the two USGS gauges. Mining areas are shown in red.

Figure 2. Data quality check. Spring River daily discharge based on 60 years' record.

Figure 3. Rainfall runoff relationship in upper Spring River watershed.

Figure 4. Different data resolution scenarios of geospatial and climatic inputs for SWAT model setup. The circles at the bottom refer to three weather data sources: NOAA, PRISM, and NCEP.

Figure 5. Observed vs. simulated streamflow rates for the calibration period (2010-2016) at Spring River (USGS flow gauge station 07186000). Upper panel depicts daily time steps while the lower panel depicts monthly time steps.

Figure 6. Observed vs. simulated flow rates for the calibration period (2010-2016) at Shoal creek (USGS flow gauge station: 07187000) at daily and monthly time scales.

Figure 7. Observed vs. simulated streamflow rates for the validation period (2000-2007) at Spring River (USGS flow gauge station 07186000) at daily and monthly time scales.

Figure 8. Observed vs. simulated streamflow rates for the validation period (2000-2007) at Shoal Creek (USGS flow gauge station: 07187000) at daily and monthly time scales.

Figure 9. Dry and wet validation for Spring River and Shoal Creek. The left panel for Spring River and the right panel for Shoal Creek.

Figure 10. Flow Duration Curve (FDC) curve for Spring River and Shoal Creek. The upper panel is for Spring River and the lower panel is for Shoal Creek.

Figure 11 [A]. Observed vs. simulated sediment loading in the Spring River Watershed before rainfall data adjustment. [a] Comparison of cumulative loading from observation and simulation. [b] Correlation between observed and simulated sediment loading. [c] Time series of observed and simulated sediment loading.

Figure 11 [B]. Observed vs. simulated sediment loading in the Spring River Watershed after rainfall data adjustment. [a] Comparison of cumulative loading from observation and simulation. [b] Correlation between observed and simulated sediment loading. [c] Time series of observed and simulated sediment loading.

Figure 12. Observed and simulated sediment loading in Shoal Creek after calibration. [A] Comparison of cumulative loading from observation and simulation. [B]

Correlation between observed and simulated sediment loading. [C] Time series of observed and simulated sediment loading.

Figure 13. SWAT computed average annual sediment loading (ton/year) in Spring River Watershed.

Figure 14. Average annual sediment loading (ton/year) and % contribution from individual tributaries.

Figure 15. Annual sediment loading (ton/year) in 2014-2016 from Spring River and Shoal Creek.

Figure 16 [A]. SWAT computed and regressed cumulative sediment accumulation vs. time in years. The lower left corner is SWAT computed values for the period 2010-2016. The inset panel shows refill time as a function of percentage sediment loading retained.

Figure 16 [B]. Refill time as a function of sedimentation rate in units of million ibs/year. Based on the historic sediment accumulation rate of 24 million ibs/year, it takes 110 years to refill the lake back with a dredged sediment mass. The refill time is 98 years based on sediment accumulation rate of 27 million ibs/year obtained from sediment data collected in 2014.

Figure 17. Location of the sediment sampling stations and Empire Lake.

List of Tables

Table 1. Type, area and % of land use in the Spring River Watershed.

Table 2. Data used and their sources for the TSMD study.

Table 3. Number of suspended sediment concentration samples used for model calibration.

Table 4. List of SWAT flow parameters, their ranges and optimized values.

Table 5. Model validation statistics for the two watersheds.

Table 6. List of SWAT sediment parameters, their ranges and optimized values.

List of Abbreviations:

SWAT	Soil and Water Assessment Tool
TSMD	Tri State Mining District
SUFI	Sequential Uncertainties Fitting Algorithm
HRU	Hydrological Response Unit
GIS	Geographic Information Systems
DEM	Digital Elevation Model
NSE	Nash Sutcliffe Efficiency
PBIAS	Percent Bias
NCEP	National Centers for Environmental Protection
NOAA	National Oceanic and Atmospheric Administration
PRISM	Parameter-elevation Regressions on Independent Slopes Model
QAPP	Quality Assurance Project Plan
FDC	Flow Duration Curve
SSC	Suspended Sediment Concentration

Acknowledgments

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded the research described here partly through the Oak Ridge Institute for Science and Education (ORISE) Research Associate Program at the National Risk Management Research Laboratory. This document does not represent and should not be construed to represent any U.S. Environmental Protection Agency determination or policy. The report benefited from the constructive comments of two internal reviews provided by Dr. Michelle Simon and Mr. Bob Lien and two external reviews by Dr. Michele Eddy and Dr. Yusuf Mohamoud.

Quality Assurance

The watershed model was developed and calibrated using the database and performance evaluation measures as outlined in the QAPP G-LRPCD-0018752-QP-1-2 (Watershed-Scale Hydrologic and Contaminated-Sediment Fate and Transport Modeling). For flow calibration and validation, the good-of-fit statistics met the thresholds stated in the QAPP and under Section 3.6 in this report for the two main USGS gauge stations at the Spring River and Shoal Creek during the period in which sediment samples were collected. The sediment data collected according to the QAPP # G-LRPCD-0019809-QP-1-5 “Filed Sampling Plan for Flow and Water Quality Data Collection in the Spring River Watershed” was used alongside the calibrated watershed model to assess two potential remedial strategies.

1. Introduction

1.1 Background

Sediment-based pollutants impact water quality in over 100,000 miles of assessed streams and rivers in the United States (USEPA, 2006). Mining-contaminated sediments in particular play a significant role in the environmental health and biodiversity of affected areas (Angelo et al., 2007; Pope, 2005). The Tri-State Mining District (TSMD), an area of about 2,500 square miles, is a historic lead and zinc mining area located in southwestern Missouri, southeastern Kansas, and northeastern Oklahoma. The TSMD was one of the world's leading zinc and lead mining areas, producing over 400 million tons of crude ore between about 1850 and 1970. Although it is now inactive, the TSMD provides an ongoing source of heavy metals (lead, zinc, and cadmium) to the environment including the US Environmental Protection Agency Superfund site located in Cherokee County, southeast Kansas, USA (Juracek and Drake, 2016; and Barks 1986). The lead and zinc deposits within the TSMD were associated with the Ozark Plateau, a geological region characterized by the presence of Mississippian rocks(Juracek and Drake, 2016; and Brosius and Sawin, 2001). The ore deposits were processed utilizing underground mining systems. The recovered ores were commonly crushed on site and concentrated using gravity separation and/or flotation. These two ore-concentration processes yielded the production of gravel and sand particles called "coarse tailings" or "chat" and sand- and silt-sized particles called "fine tailings". Additional smelting and refining of these ore concentrates were conducted at various locations within or outside the TSMD. These mining activities resulted in contamination of surface water, groundwater, sediments, and flood plain soils in the Spring River basin with lead, zinc, and other heavy metals. Although much of the surface mine wastes has been removed over the last few decades, thousands of acres of wastes (waste rock, chat piles, tailing materials) still remain on the ground surface as a source of heavy metals (e.g., lead and zinc). Over time, trace metals were dispersed over a large area and beyond the original sites of disturbance, mostly in particulate phase. Such areas include streambeds and floodplains in the tributaries and mainstem Spring River downstream from mining affected areas, in gravel paved driveways, landscaped lawns, and rural roads. Remedial measures of areas affected by mining activities hinges upon an understanding of the magnitude and extent of contamination and the environmental fate and transport of mining-contaminated sediment and soil. The latter is imperative because most of the mining-related heavy metals (e.g., zinc and lead) introduced into the environment is in, or becomes associated with, the particulate phase (Juracek and Drake, 2016; Beyer et al. 2004).

The fate and transport pathways of naturally occurring and anthropogenically produced constituents in a watershed are determined by complex interactions of landscape, climate, hydrology, and physical and biochemical processes in the water column and in the sediment bed region. Watershed-scale mathematical models are designed to represent and simulate the hydrology, transport pathways, and fate of contaminants in surface runoff, stream channels, and the subsurface. The models can serve as useful tools in conceptualizing, understanding, and differentiating the relative significance of natural processes and anthropogenic activities on predicting trends in water quality and aquatic ecosystem resources (USEPA, 1995).

Several watershed-scale models have been applied to simulate metal fate and transport (Johnson and Zhong, 2006; Velleux et al., 2006; England et al., 2007; Galvan et. al., 2009). For example, the Two-dimensional Runoff, Erosion, and eXport (TREX) model is perhaps the most comprehensive model for simulating metals transport at the watershed scale; however, it is event-based and data intensive. The model was applied to the California Gulch, Colorado mining-impacted watershed (Velleux et. al., 2006). The study demonstrated the ability of TREX to moderately predict total suspended sediment and metals loadings/concentrations. A second example, the model Contaminant Transport Transformation and Fate (CTT&F) developed by US Army Corp of Engineers, showed a satisfactory agreement between model simulations and experimental data (Johnson and Zhong, 2006). The Soil and Water Assessment Tool (SWAT) has been successfully implemented all over the world to simulate and inform various environmental issues related to water quantity and quality studies (Gassman et al., 2014). The metal loading transported by the Meca River to the Sancho Reservoir (Spain) showed satisfactory agreement between simulated and observed flow data using SWAT (Galvan et. al., 2009).

In recent years, model simulations have become significant in the decision-making process with regard to optimal management of sediment at the watershed scale. The SWAT model along with other hydrological models were applied by various researchers (Tripathi et al., 2003; Phomcha et al., 2012; Mukundan et al., 2015; Liu et al., 2016). Mukundan et al. (2015) calibrated

SWAT using detailed monitoring data to simulate spatial sediment loading in Upper Esopus Creek Watershed (UECW) which is part of the New York City water supply. Their study analyzed the high frequency suspended sediment loading data to assess the inter-annual variability and seasonality in suspended sediment loading in the studied watershed. Identification and prioritization of critical sub-watersheds for soil conservation management using the SWAT model was investigated by Tripathi et al. (2003). In that study, the SWAT model was applied to identify critical sub-watersheds based on estimated sediment yield and nutrient losses of a small agricultural watershed to aid development of an effective management plan. The study established that the Soil and Water Assessment Tool (SWAT) model could accurately simulate runoff, sediment yield, and nutrient losses from the agricultural watershed. Modeling the impacts of alternative soil conservation practices for an agricultural watershed with the SWAT model was studied by Phomcha et al. (2012). They applied SWAT model in The Lam-Sonthi watershed (357 km^2) in central Thailand to identify critical areas and suggest effective soil conservation measures to minimize sediment yield in an agricultural watershed. Briak et al. (2016) used SWAT for sediment yield assessment in Kalaya gauged watershed (Northern Morocco).

In this study, we constructed and calibrated a Soil and Water Assessment Tool (SWAT) model to simulate hydrology and sediment transport within the portion of the Spring River Basin upstream from Empire Lake (Spring River and Shoal Creek Watersheds, Fig. 1). The model was applied to calculated annual sediment loading to Empire Lake and evaluate hypothetical strategies for remediation of contaminated sediments in the lake and mining-affected tributaries.

1.2 Objectives

The overall objective of this study is to evaluate alternative remedial strategies in mining-affected tributaries within the Spring River Basin and Empire Lake using the SWAT model. This report's objectives are threefold:

1. Develop and calibrate a SWAT model for flow and sediment transport in the Spring River Watershed upstream from Empire Lake.
2. Simulate annual sediment loadings from the Spring River and Shoal Creek to Empire Lake.

3. Use the semi-distributed watershed model, observed data, and literature to evaluate two potential remedial management scenarios for contaminated sediments: lake sediment dredging and sediment traps.

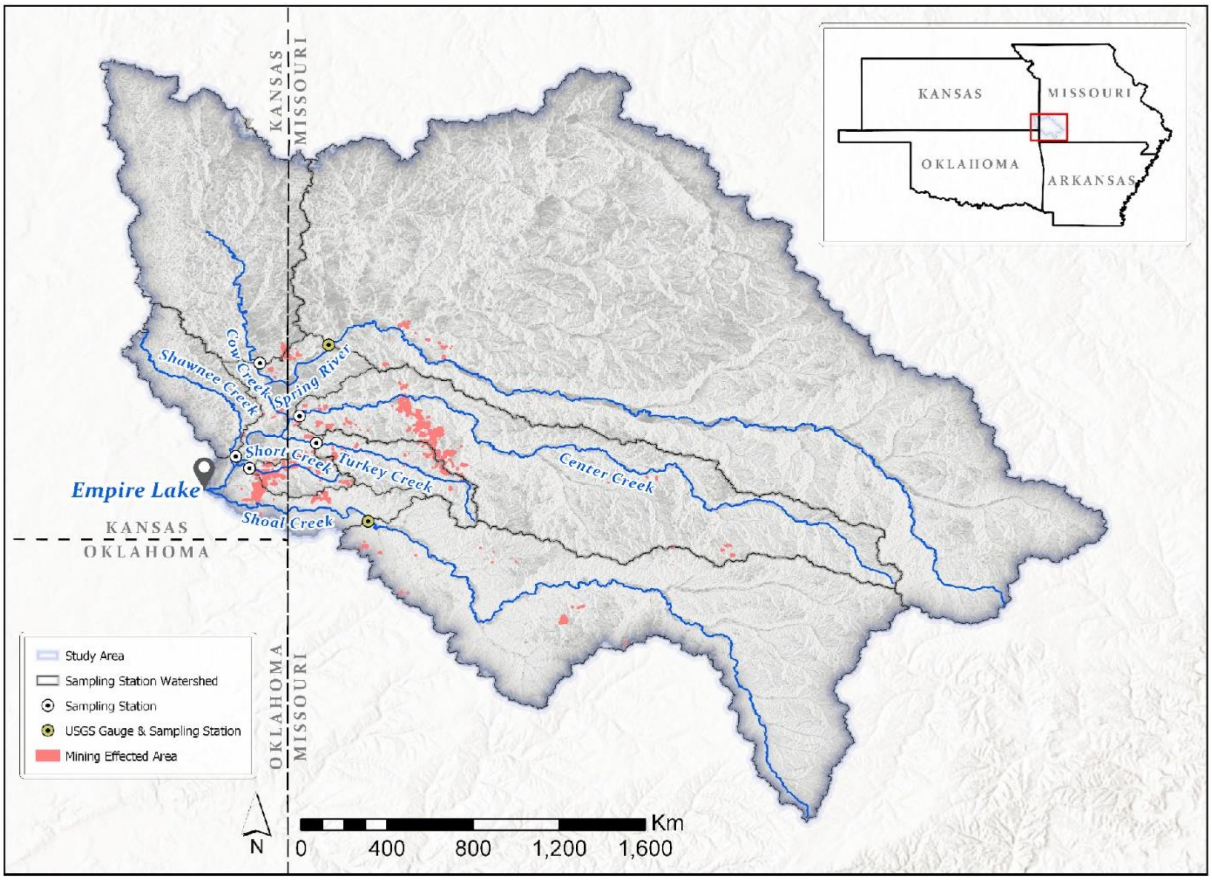

Figure 1. Study area map and main features of the Spring River Watershed (upstream from Empire Lake) and tributaries. The map shows a geographic overview of the watershed and the Spring River and Shoal Creek Watersheds. Water quality sampling stations are marked with white circles. The two USGS streamflow gauges on the Spring River and Shoal Creek are marked with yellow circles. Water quality was also sampled at the two USGS gauges. Mining areas are shown in red.

2. Hydrology Model

2.1 Study Area

The Spring River Basin is located within the TSMD, mostly in southwest Missouri, and encompasses an area of about 2,377 mi^2. The upper portion of the Spring River Basin, which drains to Empire Lake, (henceforth, referred to as the Spring River Watershed) covers a portion of southeast Kansas and northeast Oklahoma including parts of Crawford and Cherokee Counties (Kansas) and Ottawa County (Oklahoma) (Fig. 1) before reaching its confluence with the Neosho River. The Basin also drains Jasper County and portions of Barry, Barton, Lawrence, and Newton Counties in Missouri. Climate of the region is considered temperate, with an average annual temperature of 59^{o}F and average annual precipitation of 40 inches (Adamski et al., 1995). A more recent estimate from data obtained from ground station PGHCNDUSC00232240 for the period (1981-2016) shows an average annual temperature of 58^{o}F and average annual precipitation of 38.40 inches.

Cropland occupies the greatest portion (57%) of the Spring River Watershed, followed by pasture (24%) and forested lands (13%) (Table 1), with forested land occupying most of the Shoal Creek Watershed. The areas in and around cities (e.g., Joplin and Web City) are dominated by high and low-density urban land use (6%). This classification was obtained from USGS land cover map.

Table 1. Type, area and % of land use in the Spring River Watershed.

Land Use	**Area (mi^2)**	**% of Total**
Cropland	1,332	57
Pasture	561	23.6
Forest	301	12.6
Urban	137	5.8
Water	33	1.4
Shrub Land	13	0.6
Total	2,377	100

The Spring River Basin is contained within the Springfield Plain of the Ozark Highlands physiographic region. This region is underlain mostly by sedimentary bedrock including Ordovician-age dolostone and sandstone, Lower Mississippian-age limestone and dolostone, and Pennsylvanian-age sandstone and shale ((USDA, 2006)). The study area has a karst landscape dominated by carbonated water through which dissolution over time created caves and water channels in the Mississippian limestone. Tropical climate caused a massive chemical weathering over time and produced a 400-ft thick shale layer covering 20 square miles containing enough trace elements to account for the Tri-State Minerals (Smith, 2016).

The modelled watershed area of 6,156 km^2 (2,377 mi^2) comprises the majority of the Spring River Watershed (i.e., upstream of Empire Lake) and Shoal Creek Watershed. The area is relatively flat, and the elevation varies from 230 m to 470 m. The Spring River Watershed has 6 different tributaries located within and near the U.S. EPA listed Cherokee County Superfund site (Juracek and Drake, 2016): Center Creek, Turkey Creek, Cow Creek, Shawnee Creek, Shoal Creek, and Short Creek. The Spring River and Shoal Creek discharge into Empire Lake. The lake is a reservoir that was formed at the confluence of Shoal Creek and Spring River, with the completion of a dam on the Spring River at Lowell, Kansas, in 1905 (Jakubauskas, 2008). The surface area is approximately 1 square mile, including the back-water area from Spring River and Shoal Creek. Published studies have shown the chemical composition of Empire Lake sediments is an environmental concern due to high concentration of lead and zin (Jakubauskas, 2008; Juracek and Drake, 2016; Pope, 2005; USEPA, 2006).

2.2 Watershed Model Development

SWAT is a process-based, semi-distributed model that simulates streamflow and water quality (Arnold et al., 1998). To accurately anticipate transport of sediments and dissolved and sediment-bound constituents, the hydrologic cycle as simulated by the model must conform to dominant processes occurring in the watershed. Simulation of the hydrology of the watershed can be divided into two major parts: the land phase, which controls the amount of water and sediment loading into the main channel, and the routing phase, which controls water flow and sediment transport through the channel network, from the watershed headwaters to the outlet (Neitsch et al., 2011).

SWAT partitions watersheds into sub watersheds using the river network and then into smaller units nested within the sub-basins known as Hydrological Response Units (HRUs) (Neitsch et al., 2005). The HRU is the smallest computational unit in a SWAT model within which all combinations of similar land uses, soils, and slopes within a sub-basin are lumped based upon user-defined thresholds (Neitsch, 2005). SWAT's hydrological routine is comprised of discharge, snow melting, and evapotranspiration. For this case study, ArcSWAT version 2012.10.19 was used along with ArcGIS Plugin 10.4.

SWAT model can simulate yearly, monthly, daily and sub daily time steps. Developed model was run on a daily time step incorporating the historic meteorological variables of precipitation, temperature, wind speed, solar radiation, and relative humidity. The USDA's SCS curve number method (reference) was applied for an estimation of surface runoff volume.

2.3 Data and Sources

SWAT uses three types of data: geographic, meteorological, and hydrologic. These data are heterogeneous, typically structured according to several main input data, such as tables, Geographic Information Systems (GIS) raster, GIS vector or multi-dimensional arrays (e.g., NetCDF). Digital Elevation Model (DEM), land use (LU), and soil maps are raster datasets, while river geometry comes typically in vector formats, hydrologic and weather data as tables, and climatic data as arrays of points. For the weather data, the minimum requirements are precipitation and minimum and maximum daily temperatures. Since observed evapotranspiration (ET) was not available for input to SWAT, we used build-in functions in the SWAT model for evapotranspiration (ET) calculation, and Penman Monteith based energy balance method (Allen et al., 1998) was selected among various methods. Hydrologic data include water flow, water quality, and sediment loads. Table 2 presents the data used in developing the watershed model.

To evaluate the performance of modeled watershed hydrology, we used daily stream discharge data from in situ USGS stream gauge stations. Two-gauge stations were used for model calibration and validation: Spring River near Waco, MO (USGS ID: 07186000) and Shoal Creek above Joplin, MO (USGS ID: 07187000). The USGS Waco station has a record of 65 years of daily data and the Shoal Creek station has 41 years of daily record. These stations were selected for their lengths of record, which is essential for SWAT streamflow calibration and performance evaluation.

Table 2: Data used and their sources for the TSMD study.

Data Type	Data Sources	Scale/Resolution	Description
DEM[1]	USGS	10 m, 30 m, 90 m	Elevation
Land use[2]	USGS	30 m	Classified land use such as crop, urban forest water etc.
Soil[3[a,b]]	SSURGO STATSGO	1:12000 1:250000	Classified soil and physical properties such as sand, silt, clay, bulk density.
Hydrological network[4]	NHD	1:24000	River network
River flow[5]	USGS	Daily	Observed streamflow
Weather[6[a,b,c]]	NCDC NOAA PRISM	Daily Daily Daily	Precipitation, Temperature, Wind Speed, Solar radiation

1. https://lta.cr.usgs.gov/NED
2. https://www.mrlc.gov/nlcd2011.php

3[a] https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053628
3[b] https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053629

4. https://nhd.usgs.gov/
5. https://waterdata.usgs.gov/nwis

6[a] https://globalweather.tamu.edu/
6[b] https://www.ncdc.noaa.gov/
6[c] http://prism.oregonstate.edu/

To explore the consistency of the streamflow data, 60 years (1956-2016) of annual and daily records from the gauge station at the Spring River were plotted in a single graph (Figure 2). From this plot, it is evident that Spring River discharge is highest in the spring months (days ~60-150) from February to May each year due to melting snow.

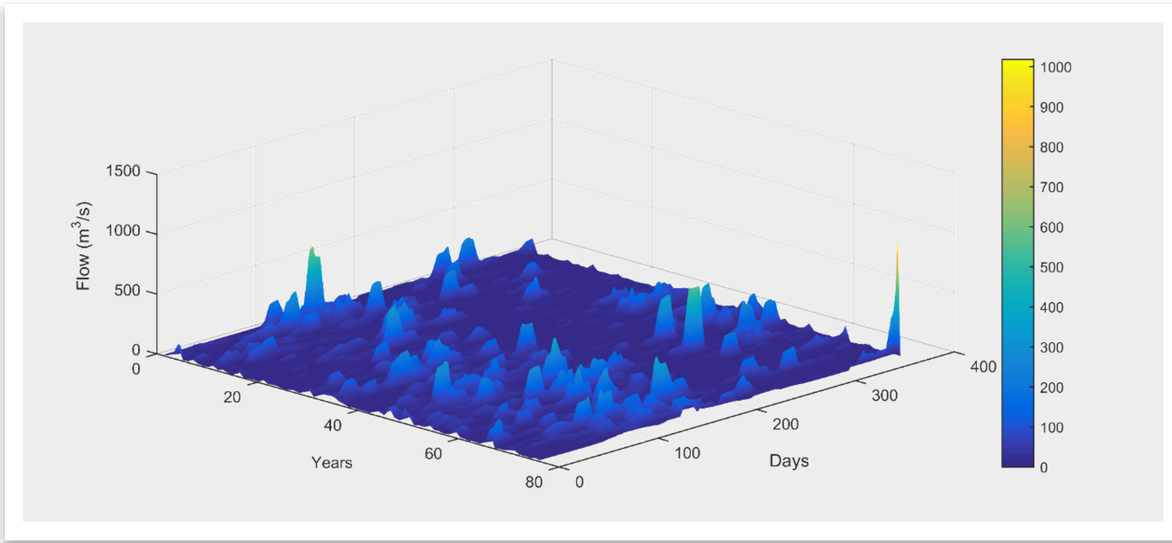

Figure 2. Data quality check. Spring River daily discharge based on 60 years' record.

Rainfall-streamflow relationships were also evaluated in the Spring River and Shoal Creek watersheds using flow records at the corresponding USGS gauge stations and concurrent precipitation records. Figure 3 shows measured flow at the Spring River gauge on the primary axis and measured rainfall depth on the secondary axis.

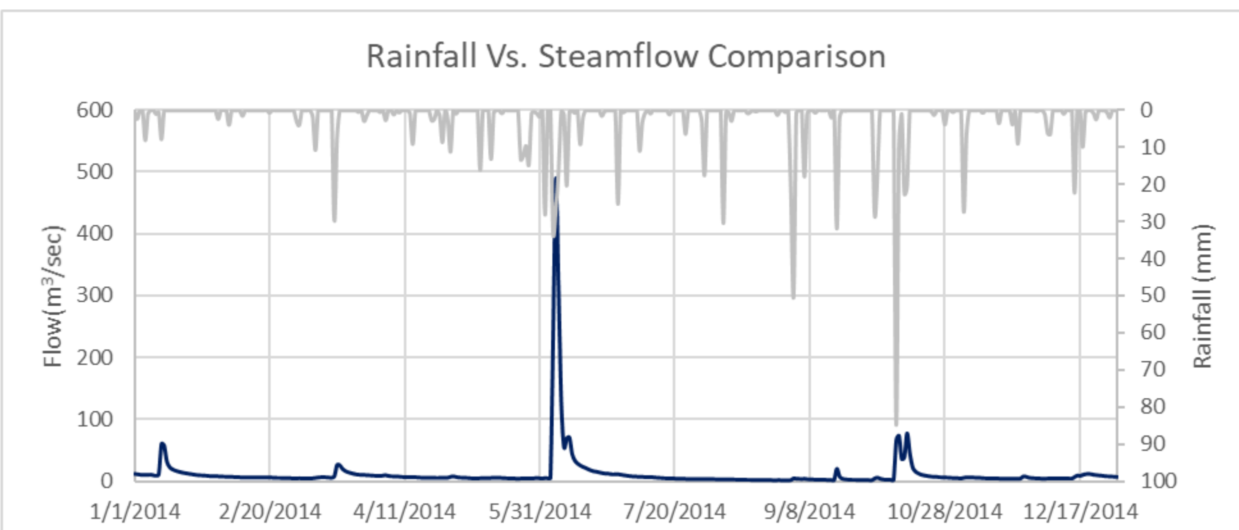

Figure 3. Relationship between rainfall and streamflow in the upper Spring River watershed.

Although most of the rainfall events coincide with the observed streamflow rates, some events do not. For example, during the first week of September 2014 a rainfall event with 50 mm rain was reported, but the concurrent observed flow rate showed little to no response. This discrepancy could be caused by errors in precipitation estimation. The nearest rainfall ground station is located 3.8 miles from the USGS gauging station upstream from Empire Lake in the Spring River Watershed. Six participant precipitation stations were identified in the studied area, which can be considered sparse for precise precipitation estimation over the watershed. To increase the spatial resolution of precipitation data a combination of satellite measured rainfall events along with ground-based measurements were used for this study.

Biweekly observations of sediment and metal concentration data were collected at 7 sampling stations from 2014 to 2016. Table 3 lists the number of samples collected at each site for suspended sediment concentration (SSC).

Table 3. Number of suspended sediment concentration samples used for model calibration.

Year	S1	S2	S3	S4	S5	S6	S7
				SSC [mg/l]			
2014	19	19	19	19	19	15	19
2015	16	16	16	16	16	17	16
2016	18	18	19	19	19	17	19
Total	53	53	54	54	54	49	54

In some events, sediment samples were damaged during shipping. Also, samples were not collected during high flow events (for example October 2015 to February 2016). We analyzed the quality of the sediment concentration measurement with turbidity. A correlation plot is provided in the supplementary section (supplementary figure S12). High correlation with turbidity corroborates the quality of suspended sediment concentration data.

2.4 Model Performance Statistics

Several studies have proposed standard hydrological model performance criteria (Bennett et al., 2013; Ritter and Muñoz-Carpena, 2013). In this study, we used Nash Sutcliffe Efficiency (NSE), PBIAS and Coefficient of Determination (R^2) as goodness-of-fit statistics as reported by Moriasi et al. (2007):

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{T}\left(Q_{m,t} - Q_{s,t}\right)^2}{\sum_{t=1}^{T}\left(Q_{m,t} - \bar{Q}_m\right)^2} \tag{1}$$

$$\mathrm{PBIAS} = \left[\frac{\sum_{t=1}^{T}\left(Q_{s,t} - Q_{m,t}\right)}{\sum_{t=1}^{T} Q_{m,t}}\right] \times 100 \tag{2}$$

$$R^2 = \left[\frac{\sum_{t=1}^{T}\left(Q_{m,t} - \bar{Q}_m\right)\left(Q_{s,t} - \bar{Q}_s\right)}{\left[\sum_{t=1}^{T}\left[\left(Q_{m,t} - \bar{Q}_m\right)^2\right]^{0.5} \sum_{t=1}^{T}\left[\left(Q_{s,t} - \bar{Q}_s\right)^2\right]^{0.5}\right]}\right]^2 \tag{3}$$

NSE is the strength of the relationship between observed and simulated values from the model, where $Q_{m,t}$ is the observed data value at time t and $Q_{s,t}$ is the simulated data value at time t. NSE values vary from -∞ to +1 (Nash and Sutcliffe, 1970). Values of NSE closer to +1 indicate better model performance. NSE is indicative of how well the plot of observed versus simulated values fit the 1:1 line. PBIAS specifies the average tendency of the simulated data to be larger or smaller than their observed values. PBIAS can be used as an indicator of under- or over-estimation between observed and simulated values. Negative PBIAS indicates an underestimation of the observed values. The square of Pearson's product moment correlation R^2 represents the proportion of total variance of observed data that can be explained by the model. Values of R^2 closer to +1 indicate better model performance.

Following Moriasi et al. (2007), we considered that model performance was satisfactory when NSE ≥ 0.6, PBIAS ≤ ± 25% and $R^2 \geq 0.6$ for simulated streamflow and NSE ≥ 0.5, PBIAS ≤ ± 55% and $R^2 \geq 0.6$ for sediment in daily time steps.

2.5 Pre-Calibration Analysis: Significance of Input Data

Geospatial and climatic inputs are essential for distributed watershed models, and they play a significant role in model performance. To increase confidence in the watershed model we conducted pre-calibration input data resolution analysis and examined 18 different hydro-climate and geospatial data resolution input scenarios. This step is intended to examine soundness of model structure and insure an optimal calibrated model using commonly used performance measures. We tested three different resolutions of DEM (10m, 30 m, and 90 m) and two soil data sources, state soil geographic database (STATSGO) and soil survey geographic database (SSURGO) for soil input. Since the land cover did not change much during the simulation period, the USGS land cover map for 2011 was used. Three climate data sources were examined, the National Centers for Environmental Protection (NCEP) based daily observations, the National Oceanic and Atmospheric Administration (NOAA) based ground data, and the reanalysis data from Parameter-elevation Regressions on Independent Slopes Model's (PRISM's) data for climatic input. PRISM dataset provides high resolution [4×4 km] weather data by pooling in and interpolating other data sets on a grid (Supplementary Figure S1). A description of PRISM is provided as a supplementary material. The 18 combinations of input data type and resolution scenarios analyzed are depicted in Figure 4.

STATSGO within the region lists 24 soil types that cover the study area, whereas with the SSURGO dataset there are 377 soil types within the study area. Similarly, there were only 7 observation points from the NOAA based stations compared to 381 grid points from the PRISM dataset. Sub-basin discretization was done to implement all of the small tributaries. For example, definition of the Short Creek sub-basin within the model required the higher resolution DEM and stream network. The watershed was subdivided into 159 sub watersheds to cover all the tributaries. With 159 sub-basins, 2664 Hydrological Response Units (HRU) were modeled for this study.

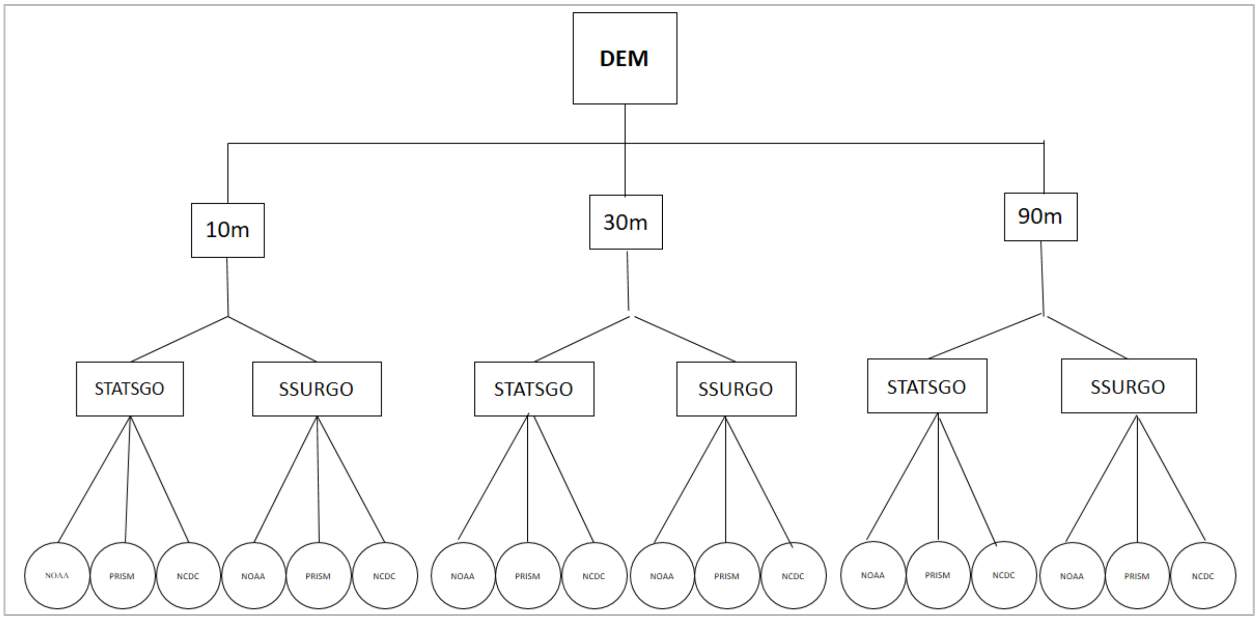

Figure 4. Different data resolution scenarios of geospatial and climatic inputs for SWAT model setup. The circles at the bottom refer to three weather data sources: NOAA, PRISM, and NCEP.

2.6 Streamflow Calibration and Validation

The optimal model was selected as the scenario (Figure 4) that produced the best NSE and R^2 prior to model calibration and using SWAT default parameters. Climate data resolution had greater impact on model performance than other data categories (i.e., DEM and soil data). The scenario corresponding to SSURGO, DEM 10-meter and PRISM as the input data produced the best model performance at the two USGS gauges: NSE= 0.66, R^2 = 0.68 and PBIAS = -16% for Spring River and NSE = 0.61, R^2 = 0.63 and PBIAS = 5.8% for Shoal Creek, both at the daily time scale. It should be noted, these goodness-of-fit statistics are based on the use of SWAT default parameter values and before attempting to calibrate the model. Note that these values satisfy Moriasi et al. (2007) performance threshold values stated above for SWAT flow calibration.

We used manual calibration first to understand parameter sensitivity and physical behavior of the catchment. Later we used auto calibration using AMALGAM (Vrugt and Robinson, 2007) and SWAT-CUP (Abbaspour et al., 2007). The results presented below are based on the best simulation from among 10,000 acceptable (behavioral) SWAT-CUP simulations.

SUFI-2, Sequential Uncertainty Fitting Ver. 2 (Abbaspour et al., 2004), which was interfaced with SWAT using the generic SWAT-CUP program, was used for calibration. In SUFI-2, two measures were used to assess the performance of the calibration: (1) the percentage of data bracketed by the 95% prediction uncertainty calculated at the 2.5 and 97.5 percentiles of the cumulative distribution of the simulated variables, and (2) the d-factor, which is the ratio of the average distance between the above percentiles and the standard deviation of the corresponding measured variable.

Streamflow calibration and validation were carried out for the two USGS gauge stations at Spring River and Shoal Creek. We used 7 years of daily data over the period 2010-2016 for model calibration and the period 2000-2007 for validation. 2008-2009 was considered as a warmup period for the model.

The calibration and validation results at the two USGS gauge stations are discussed in the following subsections.

2.6.1 Streamflow Calibration (2010-2016)

Table 4 list the default values and optimal parameter values corresponding to the best SUFI-2 simulation (i.e., calibrated model). Based on the pre-calibration analysis, we considered the input data resolution scenario which produced best goodness-of-fit statistics (i.e., DEM 10m, SSURGO, and PRISM). Figure 5 compares observed and simulated streamflow in the Spring River Watershed at daily and monthly time scales. No major deviations were found between observed and simulated values. Goodness-of-fit statistics at the daily time scale are NSE =0.77, R^2 =0.78, and PBIAS =-12.16%. For the monthly time scale results are: NSE =0.83, R^2 =0.84, PBIAS =-12.20%. The model can explain 78% and 84% of the variance in the observed data at the daily and monthly time scales, respectively.

Table 4. List of SWAT flow parameters, their ranges and optimized values.

Parameter	Description	Range	Optimal value
SURLAG	Surface runoff lag time	1,4	1.2
CN2	Curve number	10,100	25
ALPHA_BF	Baseflow alpha factor	0,1	0.26
GW_DELAY	Groundwater delay time	0,500	210
SOL_AWC	Soil available water storage capacity	0,1	0.21
CH_N	Manning's n value for the main channel	0.01,3	0.121

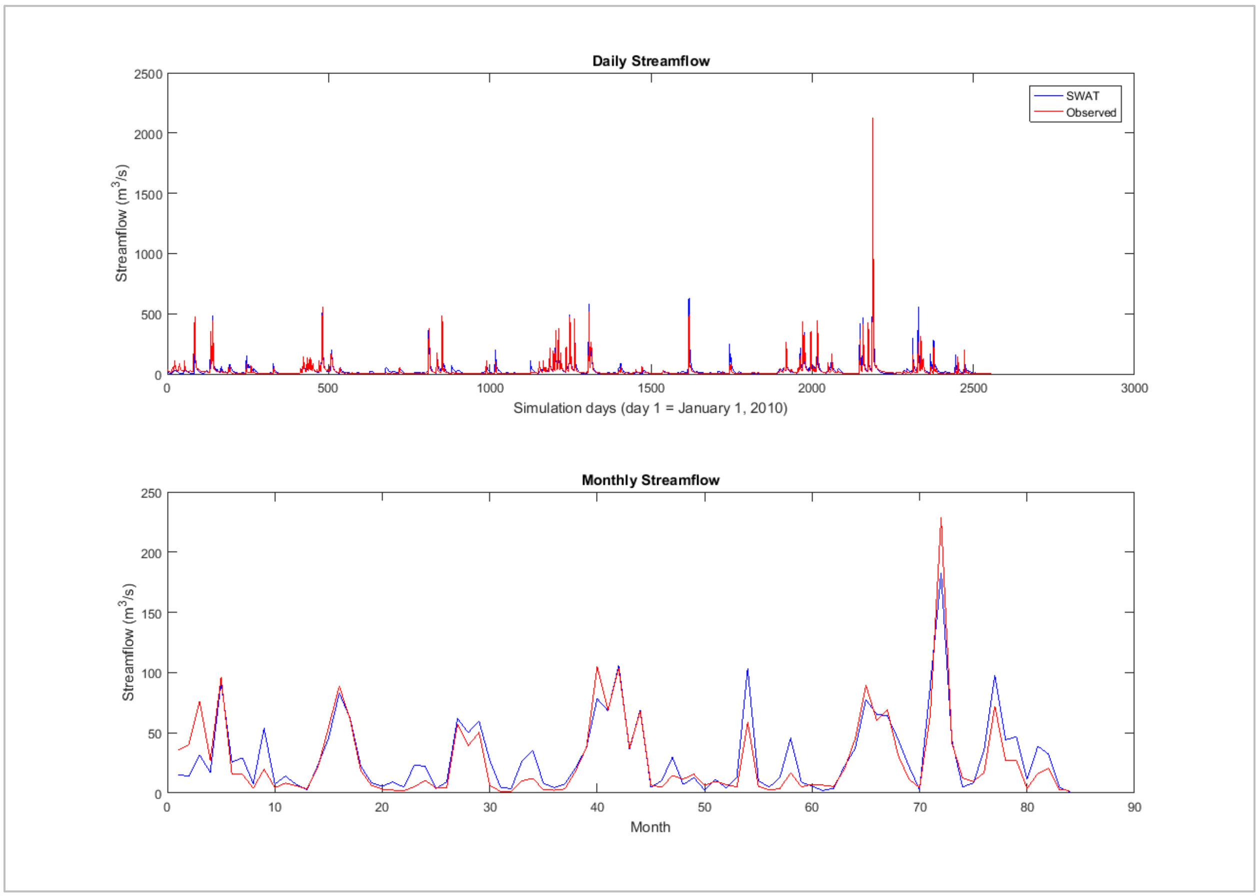

Figure 5. Observed vs. simulated streamflow rates for the calibration period (2010-2016) at Spring River (USGS flow gauge station 07186000). Upper panel depicts daily time steps while the lower panel depicts monthly time steps.

Goodness-of-fit statistics at the daily time scale for Shoal Creek are NSE = 0.67, R^2 = 0.68, PBIAS = 4.45, and for the monthly time scale these are: NSE = 0.81, R^2 = 0.82, PBIAS = 4.56 (Figure 6). The model accounted for 68% and 82% of the variance in the observed data at the daily and monthly time steps, respectively.

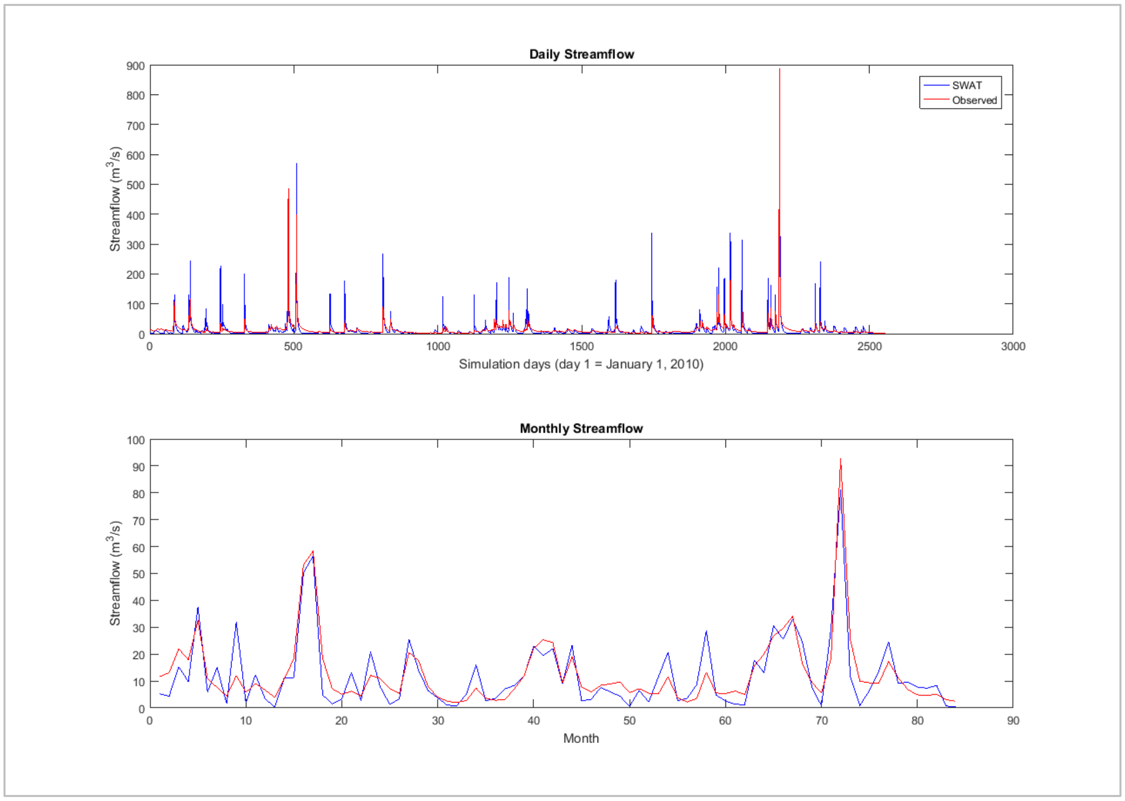

Figure 6. Observed vs. simulated flow rates for the calibration period (2010-2016) at Shoal creek (USGS flow gauge station: 07187000) at daily and monthly time scales.

2.6.2 Streamflow Validation (2000-2007)

Validation of the model with respect to streamflow was performed to test the robustness of the model outside of calibration period (2000-2007). Since suspended sediment concentration was sampled from 2014 to 2016, we opted to use this time as a part of the model calibration period (2010-2016).

Comparison of SWAT predicted streamflow with observed values for Spring River Watershed is shown in Figure 7. Goodness-of-fit statistics at the daily time scale are NSE = 0.66, R^2 = 0.67 PBIAS = -21%, and at the monthly time scale, NSE = 0.67, R^2 = 0.76, PBIAS = -7.7%. As expected, performance at the monthly time scale was better. The model explained 67% and 76% of the variance in the observed data at the daily and monthly time scales, respectively.

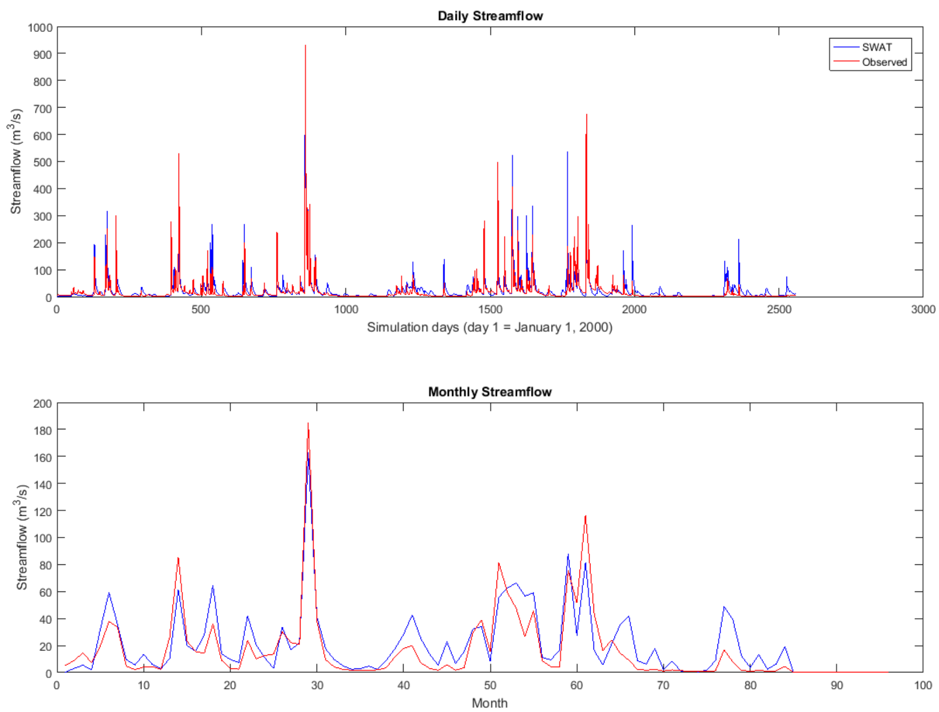

Figure 7. Observed vs. simulated streamflow rates for the validation period (2000-2007) at Spring River (USGS flow gauge station 07186000) at daily and monthly time scales.

Figure 8 shows validation results for the Shoal Creek Watershed at both daily and monthly time scales. Goodness-of-fit statistics at the daily time scale for Shoal Creek are NSE = 0.67, R^2 = 0.68, PBIAS = 4.45%, and for monthly time steps are: NSE = 0.81, R^2 = 0.82, PBIAS = 4.56%. The model explained 68% and 82% of the variance (based on NSE) in the observed data at the daily and monthly time steps, respectively. Table 5 lists performance statistics for the two watersheds.

Table 5. Model validation statistics for the two watersheds.

	Spring River Watershed		Shoal Creek	
Parameter	Daily	Monthly	Daily	Monthly
NSE	0.66	0.67	0.67	0.81
R^2	0.67	0.76	0.68	0.82
PBIAS	-21%	-7.7%	4.45%	4.56%

The calibration and validation goodness-of-fit statistics for both Spring River and Shoal Creek met the threshold performance values recommended by Moriasi et al. (2007).

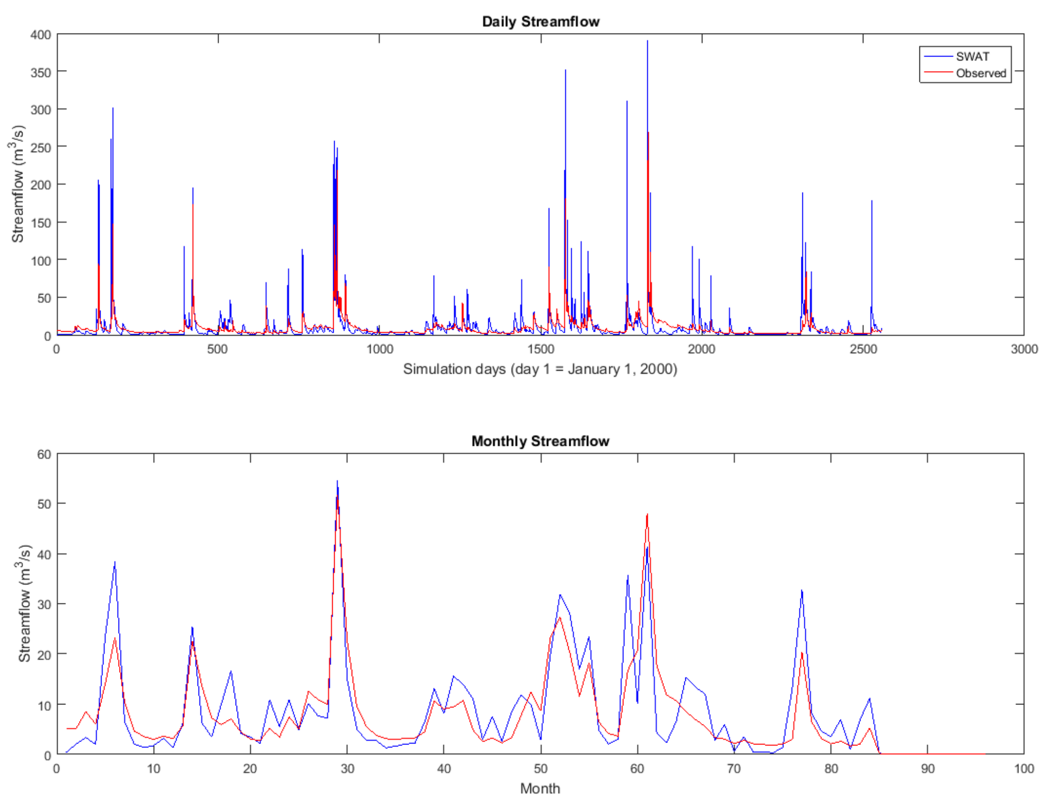

Figure 8. Observed vs. simulated streamflow rates for the validation period (2000-2007) at Shoal Creek (USGS flow gauge station: 07187000) at daily and monthly time scales.

To further test the robustness of model, the period 2000-2007 was split into two periods with distinctive precipitation patterns: October-March (low precipitation) and April-September (high rainfall and snowmelt). Figure 9 compares observed to model simulated streamflow values for the two periods in the upper Spring River and Shoal Creek. Scatter plots show good agreement between observed and model simulated flows for both periods. The coefficient of determination R^2 ranged from 0.79 to 0.91, which further reinforced robustness of the calibrated model in simulating two different weather conditions.

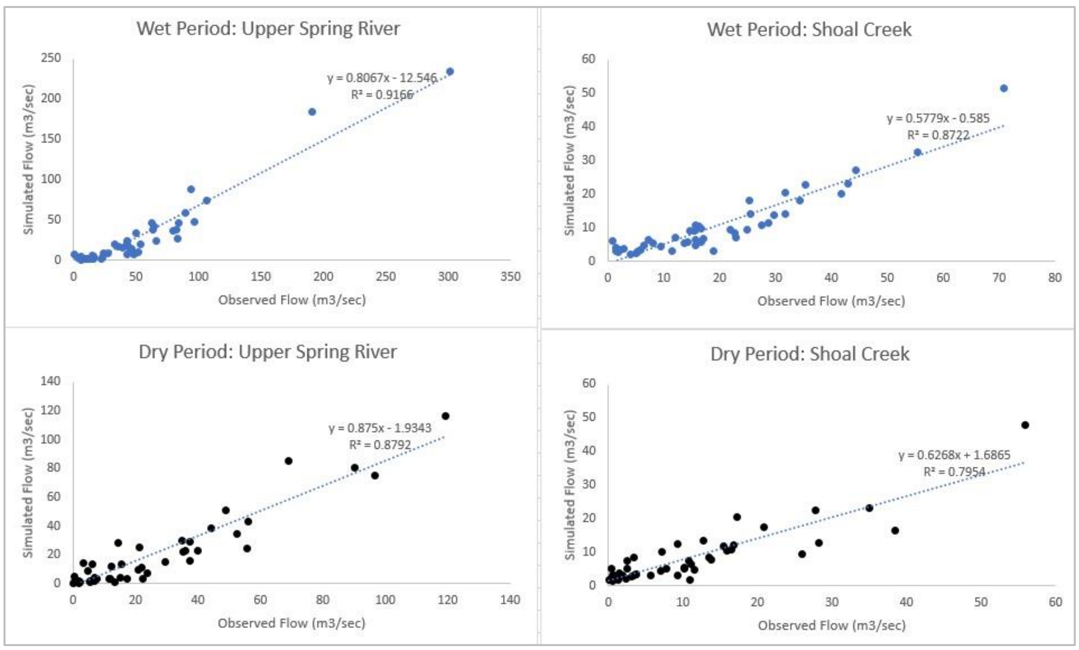

Figure 9. Dry and wet validation for Spring River and Shoal Creek. The left panel for Spring River and the right panel for Shoal Creek.

Further inspection of model performance over the full range of flow rates in the Spring River and Shoal Creek Watersheds is depicted in the flow duration curve (FDC) in Figure 10. FDC curves are most commonly used to depict the temporal variability of flow (Dingman, 2002). FDC is the relation between the magnitudes of streamflow at a gauge (e.g., average daily flow) and the frequency (probability) with which those magnitudes are exceeded over an extended time period;

it is highly informative way to summarize the difference between model simulated and observed flow rates over the full range of the recorded streamflow rates, from low to high flow rates. We used rank-based approach described in Chow et al (1988) for computing exceedance probability of the model simulated flow and observations for both watersheds. Results show that in both cases deviations are mostly in the mid-range flows. Both high and low flow probabilities of exceedance match closely, which indicates the model simulates high and low flow events well. Deviations between model simulated streamflow rates and observed values occur in the range from 50 m^3/sec to 350 m^3/sec for Spring River. For Shoal Creek, the deviations are mostly in the flow range from 20 m^3/sec to 250 m^3/sec. For both watersheds, probability of exceedance of SWAT simulated flows overestimated that generated from the observations.

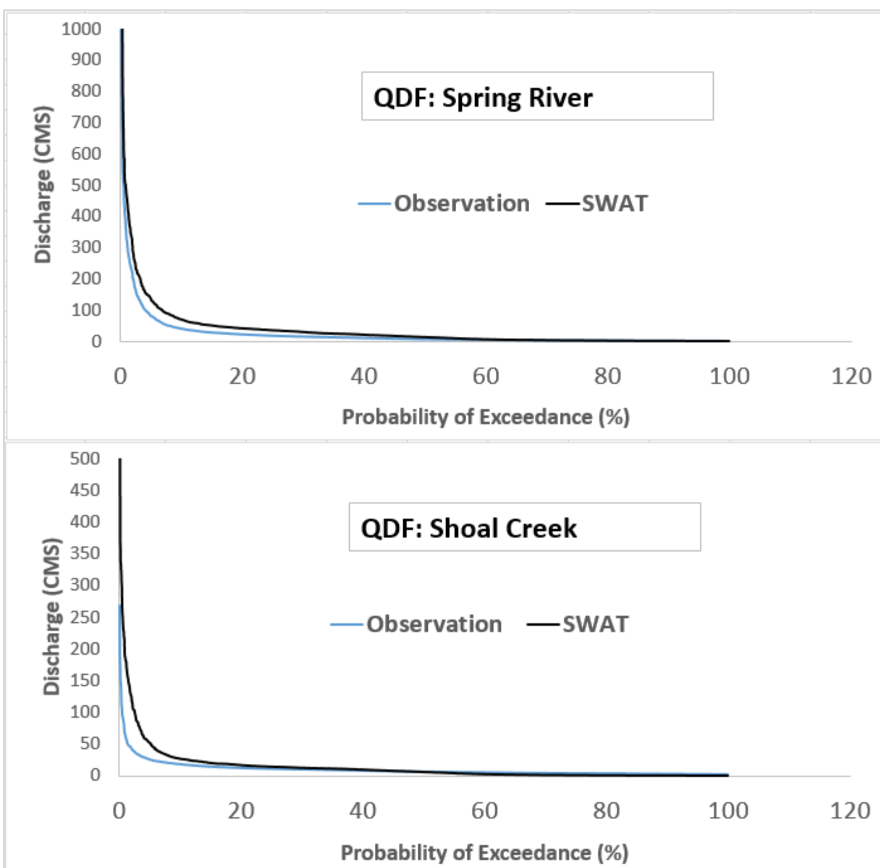

Figure 10. Flow Duration Curve (FDC) curve for Spring River and Shoal Creek. The upper panel is for Spring River and the lower panel is for Shoal Creek.

3. Sediment Model

3.1 Sediment Transport Model Development

The sediment routing model consists of two processes occurring simultaneously: deposition and degradation. Deposition in the channel and floodplain from the sub-watershed to the watershed outlet is based on the sediment particle settling velocity. The settling velocity is determined using Stokes law (Chow et al., 1962) and is calculated as a function of particle diameter squared. The depth of fall through a reach is the effect of settling velocity and the reach travel time. The delivery ratio is estimated for each particle size as a linear function of fall velocity, travel time, and flow depth. Degradation in the channel is based on Bagnold’s stream power concept (Bagnold, 1973).

SWAT uses the Modified Universal Soil Loss Equation (MUSLE) (Williams and Berndt, 1977) to predict sediment generation adopted from FitzHugh and Mackay (2000).

SWAT calculates channel sediment transport using the following equation (Neitsch et al., 2011) :

$$T = a \times V^{b} \tag{5}$$

where T, is the transport capacity (ton/m^3); V, is flow velocity (m/s); and a and b, are constants. Depending on whether the amount of sediment being carried is above or below the transport capacity, SWAT either deposits excess sediment or re-entrains sediment through channel erosion. Flow velocity is computed as:

$$V = \frac{F}{w * d} \tag{6}$$

where F, is the flow volume (m^3/s); w, is channel width (m); and d, is depth of flow (m). For flows below bankfull depth, depth of flow is calculated using Manning's equation, assuming that channel width is much greater than depth:

$$\mathrm{d} = \left(\frac{\mathrm{F} * \mathrm{n}}{\mathrm{w} * \mathrm{cs}^{0.5}}\right)^{0.6} \tag{7}$$

where n, is the Manning's roughness coefficient for the channel and cs, is channel slope (m/m). For flows above bankfull depth, depth of flow is equal to channel depth.

The MUSLE equation used to estimate sediment generation is as follows:

$$Y=11.8(Q\times\mathrm{pr})^{0.56}K\times C\times P\times\mathrm{LS} \tag{8}$$

where Y, is the sediment generation (metric tons); Q, is volume of runoff (m^3); pr, is peak runoff rate (m^3/s); K, is soil erodibility factor; C, is cover and management factor; P, is support practice factor; and LS, is topographic factor. For each day with rainfall and runoff, sediment generation is estimated by applying Eq. (4) for each HRU in the watershed.

Peak runoff rate is calculated using a modified version of the "Rational Equation" (Boughton, 1989):

$$pr = \frac{\alpha * q * A}{360 * tc} \tag{9}$$

where pr, is the peak runoff rate (m^3/s); q, is runoff (mm); A, is HRU area (ha); tc, is time to concentration (h); and α, is a dimensionless parameter that expresses the proportion of total rainfall that occurs during tc. The value of α is calculated as:

$$\alpha = \mathrm{a1} * \left(\frac{\mathrm{tp6}}{\mathrm{tp5}}\right) * \left(\frac{\mathrm{tc}}{6}\right)^{\alpha 2} \tag{10}$$

where $a1$ is the fraction of rainfall that occurs during 0.5 h; tp6 and tp5 are the 10-year frequencies of a 6 and 0.5 h rainfall, respectively, derived from Herschfield (1961) ; and $\alpha 2$ is a constant equal to 0.242 for Dane County, Wisconsin.

Overland time is computed as:

$$\mathrm{ot} = \frac{0.0556(\mathrm{sl} * \mathrm{n})^{0.6}}{\mathrm{s}^{0.3}} \tag{11}$$

where ot, is the overland time to concentration (hours); sl, is average subwatershed slope length (m); *n*, is Manning's overland roughness coefficient for the HRU; and *s*, is overland slope (m/m).

3.2 Sediment Model Calibration

SWAT simulates sediment loading with various temporal scales. For this study we extracted the daily loading values from the model to compare them with the available field measurements. The loading values were calculated by multiplying suspended sediment concentrations (SSC) in milligram per liter (mg/L) with flow rate (m^3/sec) and converting the units to ton per day. Similarly, we used the automatic calibration tool (SWAT-CUP) for sediment calibration. The parameters were selected from published studies in the region and their sensitivities were tested using the SUIFI method (Abbaspour et al., 2007). Five major sediment related parameters were selected in addition to the streamflow parameters. Table 6 lists the five most sensitive parameters and range of values used during SWAT-CUP calibration, starting with CH_COV as the most sensitive parameter and PRF as the least sensitive one among the list.

Table 6. List of SWAT sediment parameters, their ranges and optimized values.

Parameter	Description	Range	Value	Location
CH_COV	Channel Cover Factor	0-1	0.2	*.rte
CH_EROD	Channel Erodibility Factor	0-1	0.06-0.8	*.rte
SP_CON	Liner Transport Capacity Co efficient	0.0001-0.01	0.005	*.bsn
SP_EXP	Exponential Transport Capacity Cofficient	1-2	2.26	*.bsn
PRF	Peak Rate Adjustent Factor	1-2	1.44	*.bsn

Figure 11[A] shows calibration results for the modelled portion of Spring River Watershed (i.e., upstream from Empire Lake). Although SWAT captured most of the events, the comparison of the cumulative loadings showed significant differences between observed and simulated values at the Spring River sampling station. There was an event in December 2014 where SWAT simulated sediment loading rates were higher than corresponding measured values in Spring River.

A close inspection revealed unexpectedly high rainfall rates apparently artificially inferred by the 4 km grid based PRISM data during that event. To remedy this problem, we assimilated ground-based rainfall measurements to corresponding satellite-based, PRISM data entries for SWAT input and obtained an improved calibration as shown in Figure 11[B]. Goodness-of-fit statistics at the daily time scale for Spring River before and after the adjustment were (NSE = 0.65, R^2 = 0.74, PBIAS = 12%) and (NSE = 0.75, R^2 = 0.92, PBIAS = -19%), respectively. The model explained 92% of the variance in the observed data at the daily time scale after the rainfall input data adjustment, a significant increase from 74% before the adjustment. The difference between the sum of observed loadings and sum of corresponding simulated loadings is substantially reduced, comparing panel [a] in both Figures 11[A] and 11[B].

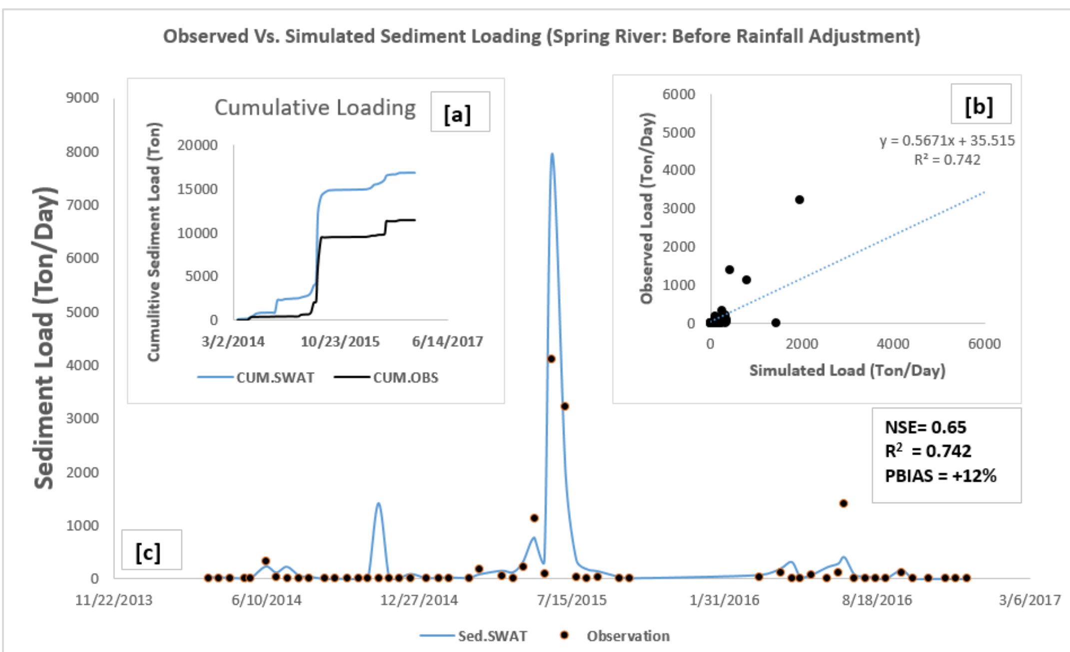

Figure 11[A]. Observed vs. simulated sediment loading in Spring River Watershed before rainfall data adjustment. [a] Comparison of cumulative loading from observation and simulation. [b] Correlation between observed and simulated sediment loading. [c] Time series of observed and simulated sediment loading.

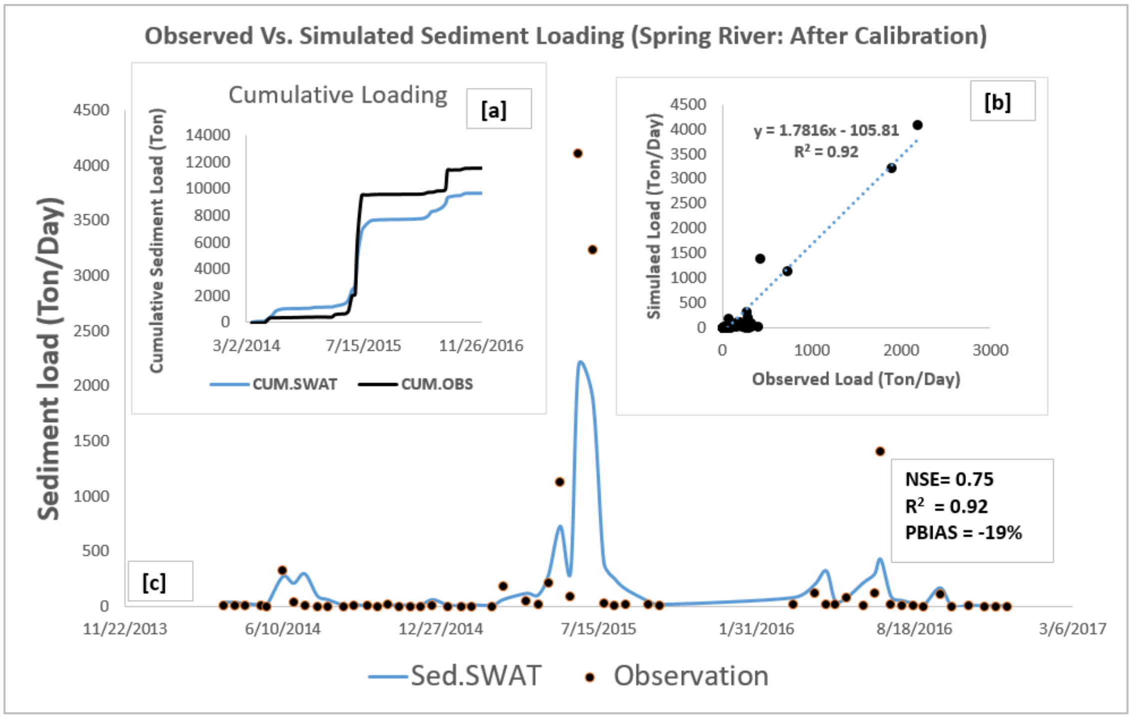

Figure 11 [B]. Observed vs. simulated sediment loading in the Spring River Watershed after calibration. [a] Comparison of sum of observed loadings with sum of corresponding simulated loadings. [b] Correlation between observed and simulated sediment loading. [c] Time series of observed and simulated sediment loading.

Figure 12 compares SWAT simulated loading to observed values at Shoal Creek after calibration and PRISM precipitation data correction. The performance is relatively poorer when compared to calibration at Spring River. Goodness-of-fit statistics at the daily time scale for Shoal Creek were NSE = 0.45, R^2 = 0.58, PBIAS = -50%. The model explained 58% of the variance in the observed data at the daily time scale. Although the NSE is slightly lower than recommended threshold value of 0.5 (Moriasi et al., 2007), nevertheless, R^2 = 0.58 is comparable to the threshold value of 0.6 and PBIAS = -50% is within the limit of ± 55%. Over all, results are satisfactory. The cumulative effect of errors in the simulated sediment loading values over the period 2014-2016, however, is apparent from large difference between SWAT estimated sum of loadings and that based on observations. The 2015 high flow event may have contributed to the significant overestimation by SWAT.

Worth noting is the relatively short length of the data and uncertainty associated with it. In general, three-year worth of sediment data might be barely enough for model calibration but not long enough to produce a robust model. Also, with all likelihood, the measured SSC may not be representative of the cross-section area-averaged concentration at the sampling station. The latter is what is computed by SWAT rather than sediment concentration at a given point in the sampled cross section. The variability of point concentration from area-averaged concentration can produce measurement errors and contribute to SWAT model uncertainty due to spatial-scale discrepancy between measured SSC and model simulated SSC. Translating a sediment sample to measured SSC may also involve errors and thus contributes to the overall model uncertainty.

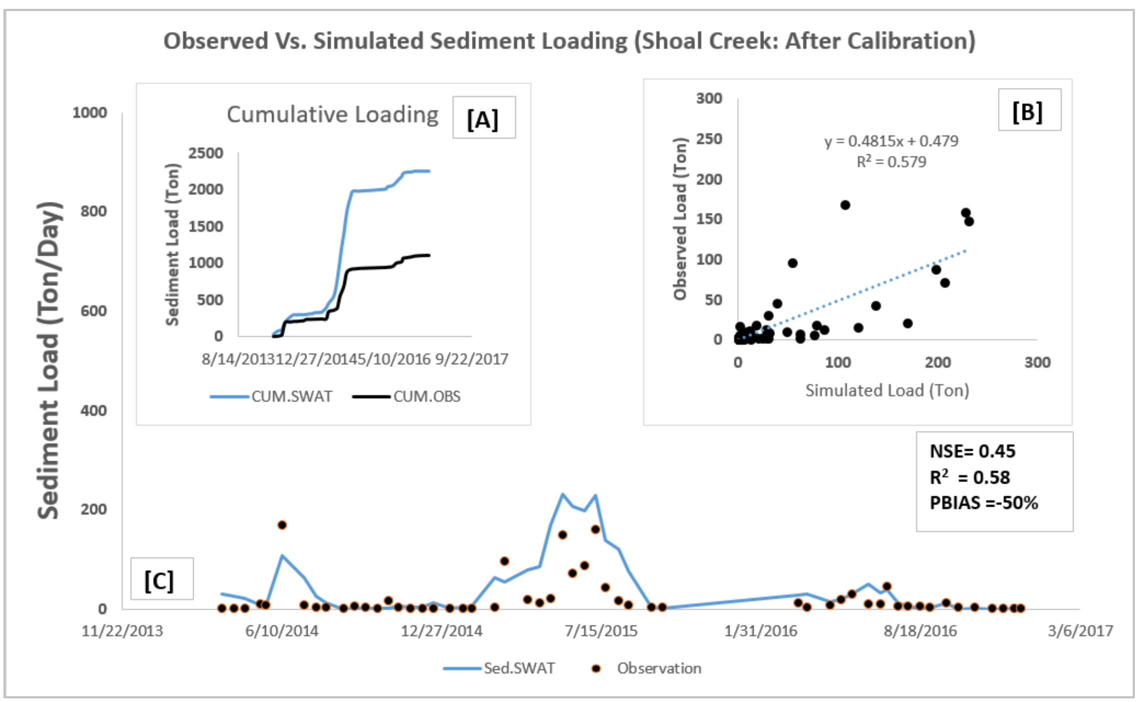

Figure 12. Observed and simulated sediment loading in Shoal Creek after calibration. [A] Comparison of cumulative loading from observation and simulation. [B] Correlation between observed and simulated sediment loading. [C] Time series of observed and simulated sediment loading.

Calibration of sediment loading in tributaries to the mainstem Spring River (Shawnee, Turkey, Cow, and Short Creeks) were overall good and with R^2 values ranging from 0.69 to 0.99 (supplementary figures S2 to S5), thus explaining 69% - 99% of the variance. It should be noted that flow rate in other tributaries was discretely measured and at the same temporal resolution of the sediment data. Even then, in some of these tributaries, measured sediment concentrations lacked corresponding observed streamflow rates which were computed using SWAT model.

4. Scenario Analysis

4.1 Sediment Source Areas and Annual Yield

The average annual sediment loadings from the sub basins in the upper Spring River Watershed simulated for the period 2010-2016 is shown in Figures 13 and 14. Figure 13 is a map depicting in colored-gradation sediment-loading contributions of the all the sub-basins. The darker the shading, the larger the annual loading rate. The magnitude of sediment loading in tons/year averaged over the period is shown in a bar-chart format (Figure 14). The upper Spring River Watershed is the largest contributor of sediment loading (52%) due to size and land use, followed by Shoal Creek (21%), and to a much lesser extent, by Center Creek and Cow Creek, each contributing 12% and 9% of the total sediment loading, respectively (Figure 14). Suspended sediment loading was expected to be higher in Spring River and its tributaries, because of the land use type, which was mostly crop land. The lower part of the watershed (Shoal Creek) was smaller in size and mostly forested; the erosion rate therefore was relatively lower compared to the mostly agricultural upper part of the watershed.

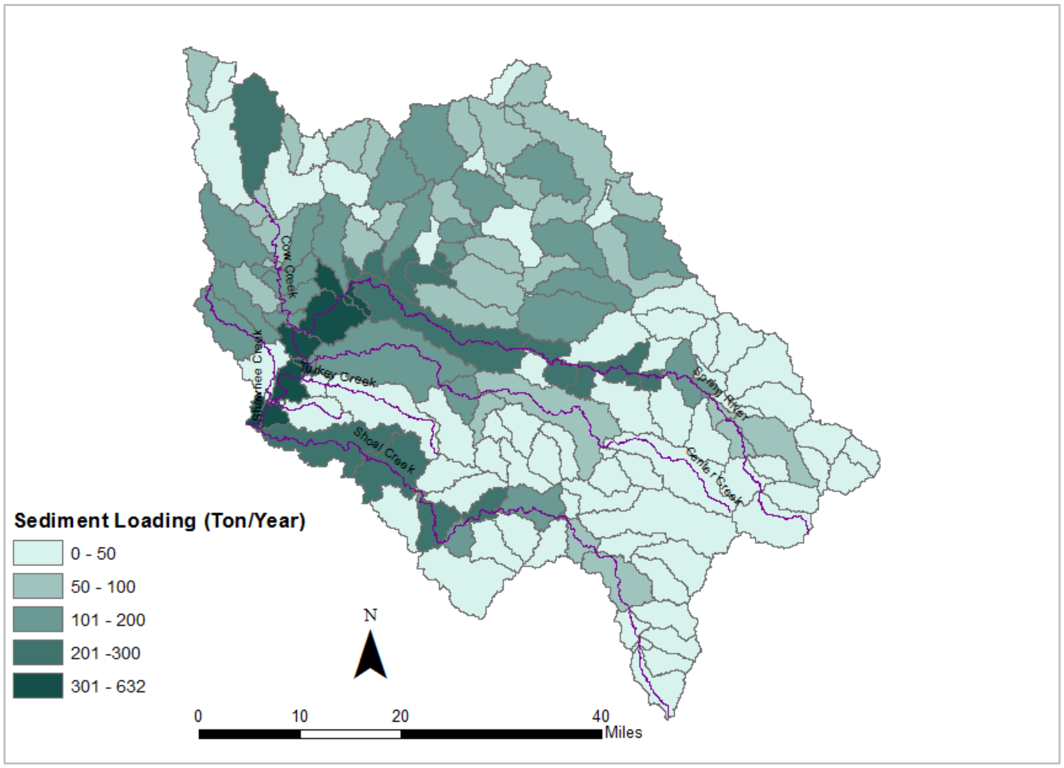

Figure 13. SWAT computed average annual sediment loading (ton/year) in Spring River Watershed.

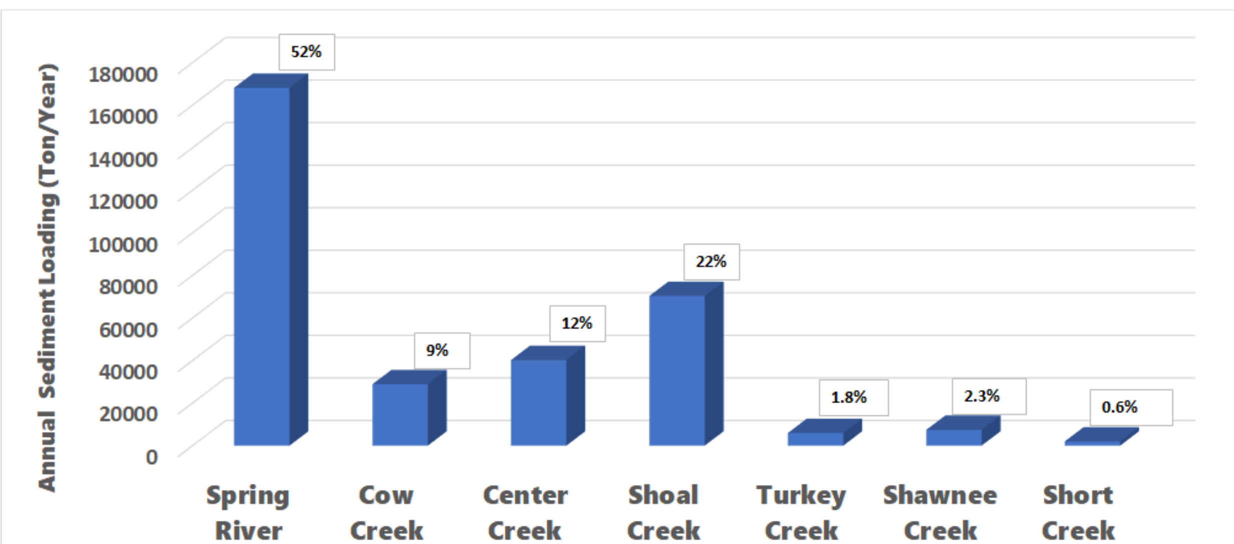

Figure 14. Average annual sediment loading (ton/year) and % contribution from individual tributaries.

Annual sediment loading over the simulation period 2010-2016 from Spring River (upstream of Empire Lake) and Shoal Creek is shown in Figure 14. 2015 had the highest SWAT computed annual loading due to a high flow event that occurred at the end of that year.

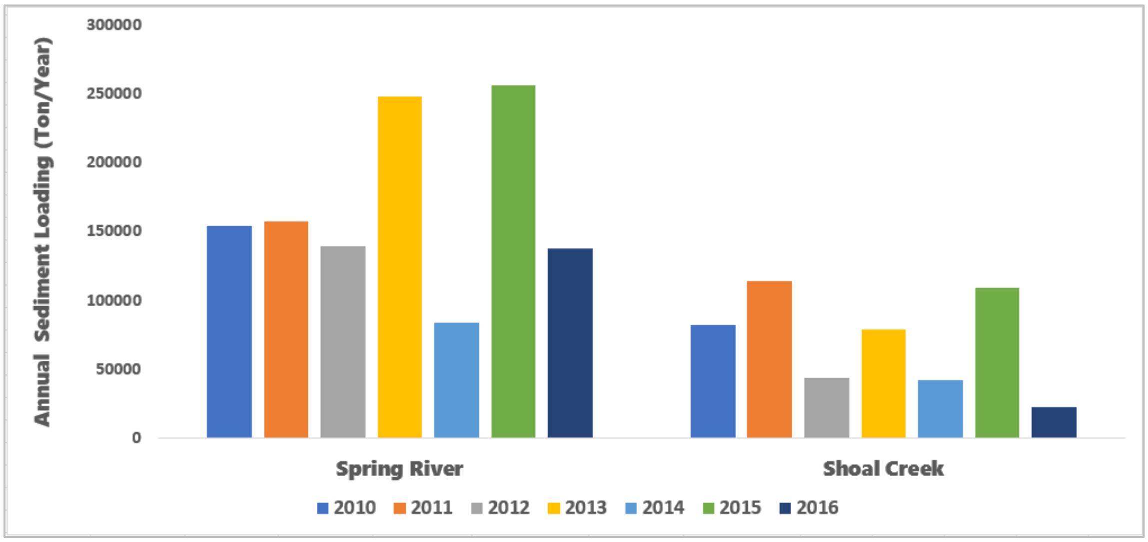

Figure 15. Annual sediment loading (ton/year) in 2010-2016 from Spring River and Shoal Creek.

4.2 Assessment of Potential Remedial Strategies

Two proposed management scenarios were evaluated using SWAT and suspended sediment concentration data: Empire Lake sediment dredging and installation of sediment traps at the outlet of mining affected tributaries (Short Creek, Turkey Creek, Shoal Creek, Center Creek, Cow Creek, and Shawnee Creek).

4.2.1 Lake Sediment Dredging Scenario

Under this scenario, we explored the dredging of Empire Lake as a hypothetical remedial measure for contaminated sediments, and calculated the time required for Empire lake to recover the dredged lake-bottom sediment mass (refill time). It is hypothesized that after being dredged the lake would be filled in time with sediments delivered by Spring River and Shoal Creek. By dividing the dredged lake sediment mass M by the annual average depositional rate S, one can calculate the refill time: $t_r = M/S$. Estimates of Empire Lake bottom sediment mass and annual sedimentation rate in the lake, hence, are needed to calculate the refill time. The annual sedimentation rate to Empire Lake can be obtained from Juracek (2006), and can be computed from the SWAT model and the sampled suspended sediment concentrations directly downstream from the lake. Calculating sediment mass to be dredged, which is key to estimating the refill time, is achieved as follows.

The USGS estimated 44.44 million ft^3 of sediment in the lake as of year 2006, deposited over a period of 100 years (Juracek, 2006). This is equivalent to a volume depositional rate of 0.44 million ft^3/year. The corresponding estimated total mass of the sediments was 2,400 million ibs (Juracek, 2006), which is equivalent to a sedimentation rate of 24 million ibs/year. According to these sedimentation rates, a projection of the total sediment volume and mass in the lake of 48.84 million ft^3 and 2,640 million ibs (1.32 million tons), respectively, can be made by year 2016. This is the sediment mass that would have been dredged assuming hypothetical dredging occurred in 2016 and sediment deposition (retention) at the historic average rate of 24 million ibs/year.

Using geospatial analysis, we carried out an independent estimate of the total volume and the total mass of the bottom sediment in Empire Lake, approximately two miles upstream into Shoal Creek, and approximately five miles into Spring River upstream from the entrance to Empire Lake. The sediment mass and volume were calculated from USGS data of 429 sampling locations in 66 transects (Juracek, 2006). A minimum curvature spline with barriers interpolation method was used for the calculations. This analysis estimated a volume of 49,336,617 ft^3 of sediment in Empire Lake and sections of Spring River and Shoal Creek. While the sediment volume calculation using the geospatial technique is a more complex method for volume estimation than the USGS method, it nevertheless yielded a comparable result: 49.34 million vs. 44.44 million ft^3 of sediment – a 10% difference. The spline interpolation estimated the mass of bottom sediment to be 2,520 million ibs, using a 51.08 lbs/ft^3 bulk density factor based on the average of 26 sediment cores (Juracek, 2006). Between the two methods, there was a difference of 5% in estimated total sediment mass. This independent estimate corroborates the 2,400 million ibs value obtained by the USGS (Juracek, 2006).

A calibrated SWAT model was used to compute sediment loading from Spring River and Shoal Creek to Empire Lake for the period 2010-2016. Cumulative loading to the lake, assuming 95% of sediment entering the lake is discharged, is shown in Figure 16 [A], the plot at the lower left corner. This estimate based on historical data implies 5% of incoming sediment mass was retained during the three-year measurement period. Extrapolation of the cumulative sediment loading to years beyond 2016 was achieved by regressing the SWAT simulated values (Figure 16[A]). The slope of the regression line is the average annual sedimentation rate (27 million ibs/year) corresponding to 5% retention of annual loading to the reservoir. The extrapolated regression line in Figure 16[A] shows that if the sediment mass dredged is 1.32 million tons, the time after dredging required for sediment mass accumulation to recover the dredged mass is about 131 years. We repeated the regression-extrapolation analysis for different % values of annual sediment loading retained, and calculated the sedimentation rate (million ibs/year) for each % value from the slope of the associated regression line. The inset panel in Figure 16[A] depicts estimates of refill time as a function of % sediment loading retained in the reservoir.

Figure 16[B] is the inset plot in Figure 16[A], except here the abscissa corresponds to average annual sedimentation rates (million ibs/year) obtained from the slopes of the regression lines described above. At the historic average sedimentation rate of 24 million ibs/year the refill time is 110 years.

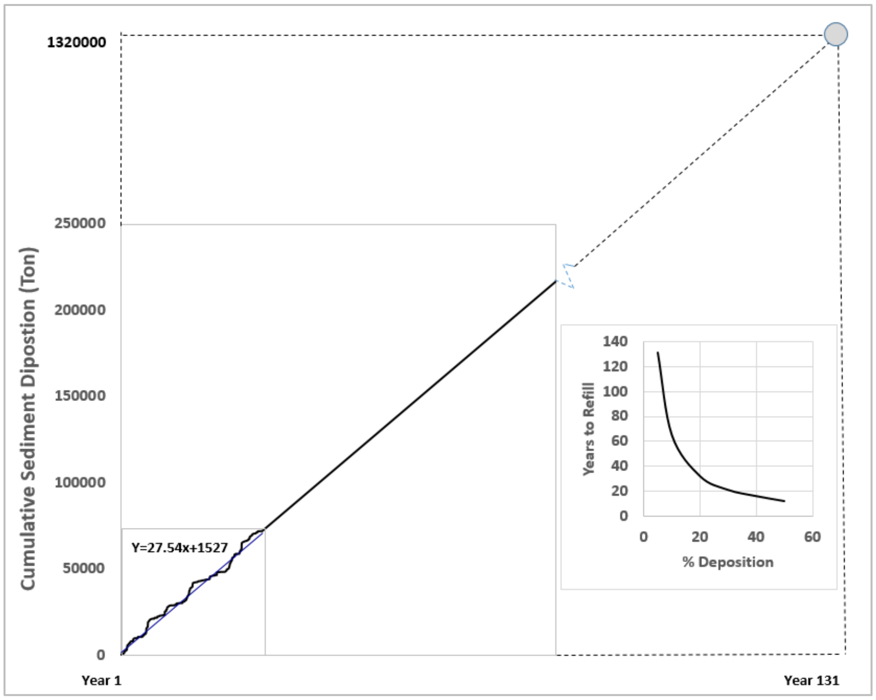

Figure 16 [A]. SWAT computed and regressed cumulative sediment accumulation vs. time in years. The lower left corner is SWAT computed values for the period 2010-2016. The inset panel shows refill time as a function of percentage sediment loading retained.

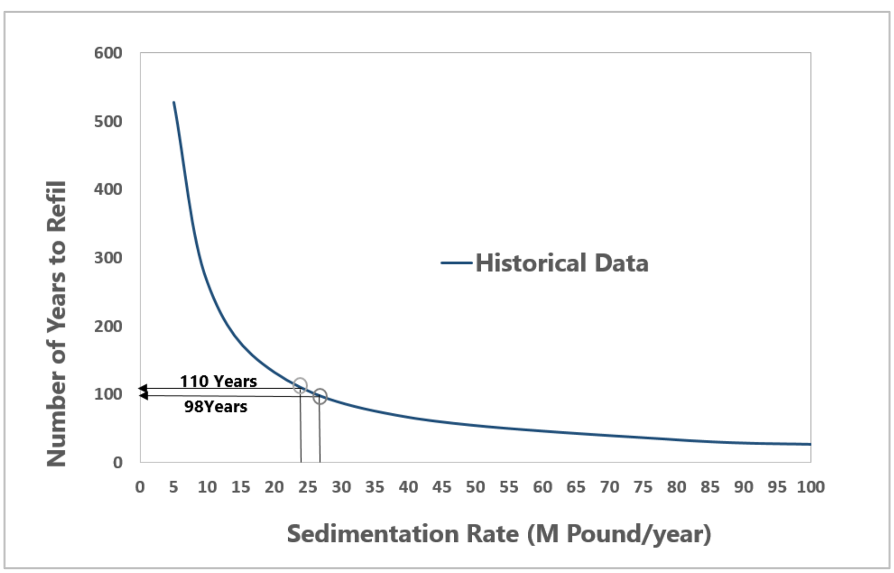

Figure 16 [B]. Refill time as a function of sedimentation rate in units of million ibs/year. Based on the historic sediment accumulation rate of 24 million ibs/year, it

> takes 110 years to refill the lake back with a dredged sediment mass. The refill time is 98 years based on sediment accumulation rate of 27 million ibs/year obtained from sediment data collected in 2014.

To gain some insight into the effect of inter-annual variability in the lake sedimentation/retention rate on the refill time, we implemented two approaches for calculating annual sedimentation rate in Empire Lake for the period 2014-2016. In the first approach, we relied on the calibrated SWAT model to simulate sediment loading to the lake for the period (2014-2016) and measured sediment concentrations at two sampling stations directly downstream from Empire Lake (Brush Creek and Baxter Spring, Figure 17). In the second approach, only observed suspended sediment concentrations (SSCs) were used. In both approaches, the annual mass of sediment deposited (S) was computed as the difference between sediment loading into Empire Lake and the sediment loading leaving the lake. The sediment loading into the reservoir is the sum of loadings of Spring River (L_1) and Shoal Creek (L_2). While, sediment loading leaving the reservoir can be estimated as the sediment loading of Baxter Spring (L_4) minus sediment loading of Brush Creek (L_3). Mass balance at the lake requires:

$$L_1 + L_2 - S = L_4 - L_3 \quad (12)$$

where $L_1 + L_2$ is sediment loading rate to the lake; $L_4 - L_3$ is sediment loading rate out of the lake; and S is defined above.

The estimated sediment masses retained in the reservoir during the years 2014, 2015, and 2016 were calculated by SWAT as 27.4 million lbs/year, -1553 million ibs/year, and 162.5 million ibs/year, respectively. A negative sedimentation rate means net removal of sediments. The USGS (Juracek, 2006) sediment mass estimate of 2,400 million ibs translated into the estimate of lake sediment mass left as of 2016: 2,400 + (24×7) + 27.4- 1,553 + 162.5 = 1,205 million ibs, which is less than half the calculated 2,640 million ibs in 2016. In obtaining the latter estimate, we maintained the assumption that the sedimentation rate for the years 2007-2013 was at the historic average rate of 24 million ibs/year.

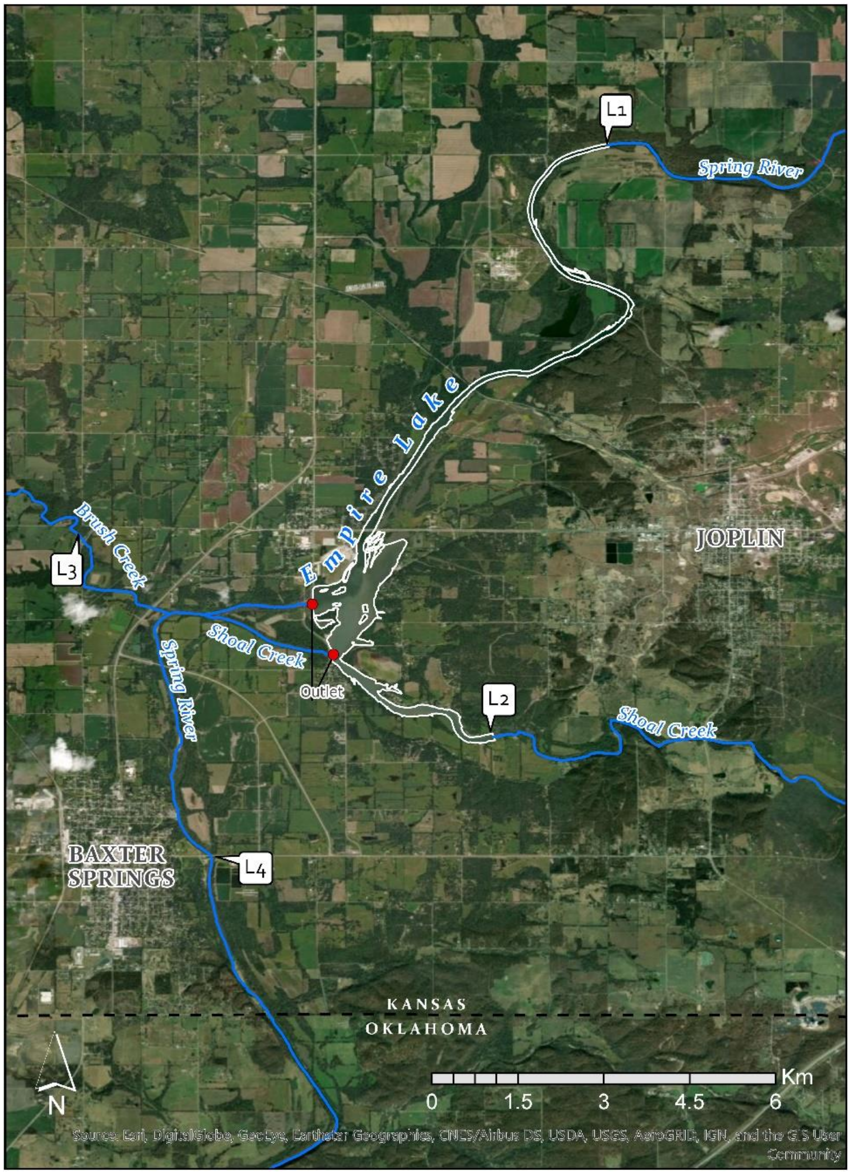

Figure 17. Location of the sediment sampling stations and Empire Lake

In the second approach, where only measured SSCs were used, the estimated sediment masses retained in the reservoir during the years 2014, 2015, and 2016 were 4.9 million ibs/year, -1362.82 million ibs/year, and 66 million ibs/year, respectively. The sediment mass left as of 2016 in this case is 2400 + (24×7) + 4.9 - 1362.82 + 66 = 1276 million ibs.

The estimated sedimentation rates of 27.4 and 4.9 million ibs/year based on 2014 observed data are comparable to the 100-year average of 24 million ibs/year obtained by the USGS (Juracek, 2006), but reflective of the kind of variability expected in the sedimentation rate from year to year. The refill time assuming annual average sedimentation rate of 27.4 million ibs/year is 98 years (Figure 16 [B]), compared to 110 years based on the historical average sedimentation rate (24 million ibs/year).

It should be noted that the streamflow and SSC measurement campaign did not cover the high flow event at the end of 2015 and beginning of 2016, and the impact of the event on lake sediment deposition (or retention), therefore, could not be assessed. The first approach, which is based on SWAT computed sediment loading, yielded loading estimates that are not immune from errors due to model uncertainty and, as described above, potential errors in the measured SSC data. The second approach, while accounting for most of the contributing sub watersheds, did not cover the relatively small watershed area between the two USGS gauge stations and the lake, and the observed sediment loadings were discrete rather than continuous in time as the case for the first approach, wherein continuous daily sediment loading was calculated using SWAT.

It remains to be seen if the estimated net sediment removal from Empire Lake reservoir in 2015, which is reflected by the negative sedimentation rate of magnitude -1553 million ibs and -1363 million ibs represent the true values. But these removal rates manifest the inter-annual variability of the annual lake sediment accumulation caused by climate variability and/or reservoir operation. Is it possible the calculated lake sediment removed in 2015 has resulted in clearing enough storage for more sedimentation in 2016 than the estimated historical rate of 24 million ibs/year? Did the high flow event in 2015 cause substantial removal of sediments from Empire Lake? A new survey of existing sediment volume and mass in the lake as well as more sampling of suspended sediment concentration data of the lake inflows and outflows from similar events may shed light and provide key answers to the above questions.

Considering the limited discretely observed sediment data and potential measurement errors, we acknowledge the uncertainty in the above analysis and estimates of sedimentation rates. As stated above, the estimated values were also subjected to annual variations of hydro-climate and flow conditions in the watershed.

4.2.2 Sediment Trapping Scenario (qualitative assessment)

In this section we make qualitative inferences on the efficacy of sediment traps in mining affected tributaries within and near the Cherokee County Superfund site. This is a mere hypothetical assessment scenario based on SWAT estimated sediment loading and reported historically elevated lead and zinc concentrations in tributaries affected by mining, namely, Short Creek, Shoal Creek, Turkey Creek, and Center Creek. The underlying hypothesis is that trapping sediment in a mining-affected tributary can help reduce discharge of metal contaminated sediment to downstream channel reaches, ultimately to Empire lake. A more objective assessment requires the design and installation of sediment traps and collection of requisite metal data over time. A sediment trap is generally a constructed 'basin' or depression on a watercourse where sediment settles out and accumulates allowing for its removal. The maintenance of the sediment traps (removal of accumulated sediment) is necessary to ensure their proper function (Ciccarello, 2011). It is expected that sediment traps can only be effective in small catchments with relatively low, intermittent flows.

Although the Spring River upstream of mining (Spring River Watershed) contributed an estimated 52% of total sediment loading from 2014-2016 (Figure 14), it historically has contributed relatively clean sediment, potentially diluting contaminated sediment from downstream mining-affected tributaries (Stratus Consulting Inc., 2006; and Juracek and Drake, 2016). On the other hand, Shoal Creek contributed about 21% of total sediment loading over the same period, but historically has been associated with much higher levels of dissolved and sediment bound lead and zinc concentrations than the reported background concentrations (Stratus Consulting Inc., 2006; and Juracek and Drake, 2016). In contrast, Short Creek's share of sediment loading was less than 1%, yet it historically was the largest single source of dissolved zinc to the Spring River (Spruill, 1987; and Davis and Schumacher, 1992). While Center Creek and Turkey Creek accounted up to 14% of total sediment loading and have relatively small catchment areas, these streams drain areas that are substantially affected by historical lead and zinc mining (Juracek and Drake, 2016).

This qualitative assessment shows that for sediment traps to be effective in mitigating zinc and/or lead in downstream reaches and Empire Lake, they should be installed within Short Creek, Turkey Creek, and Center Creek. But even then, their efficacy would be limited by the amount of zinc in the dissolved phase. For example, historical data for Short, Center, and Turkey Creeks indicated relatively high dissolved concentrations of zinc and/or lead (Spruill, 1987; and Davis and Schumacher, 1992). Installation of sediment traps at the mouth of relatively large, high flow catchments, such as the Spring River Watershed and Shoal Creek Watershed is not feasible.

Finally, it may be interesting to see if the installation of sediment traps in the mining-affected tributaries would affect the time to refill. This can be explored by first noting that sediment accumulation in the lake is the product of the average annual sedimentation rate *S* and time *t*. One can easily obtain the following relationship between refill time (t_r) and yearly trapped sediment load (ΔL):

$$t_r = t_{r0} + \frac{M}{S}\left(\frac{\frac{\Delta L}{L}}{1-\frac{\Delta L}{L}}\right) \tag{13}$$

where, t_r is maximum refill time with sediment traps; t_{r0} is time to refill without sediment traps, *M* is dredged sediment mass; *L* is annual sediment loading to the lake; and *S* and ΔL are defined above. Note that ΔL is annual sediment loading from sub-watersheds in which sediment traps would be installed, assuming (hypothetically) complete filtering of the sediments by each sediment trap. $\Delta t = t_r - t_{r0}$ is the maximum increase in refill time due to installation of sediment traps. The above relationship is based on quasi steady-state sediment transport through the lake and assumes that the sediment mass outflow is proportional to the sediment mass inflow to the lake. Since finer sediment particles are likely to pass through traps, complete filtration is not plausible. The estimate of actual refill time given by Eq. (13), therefore, should be viewed as an upper limit to the calculated refill time. In other words, estimate of actual refill time should be less than the value calculated by Eq. (13).

As an example, let's assume sediment traps were installed in Short, Center, and Turkey Creeks in 2016; i.e., $\Delta L/L = 0.14$ since the three creeks contribute about 14% of total sediment loading. Inserting into Eq. (13) the above data and *S* = 24 million Ibs/year, and *M* = 2640 million Ibs, Δt = (2640/24) ×[14/(100-14)] = 18 years at most as the increase in the time for Empire Lake to be filled with sediment back to the pre-dredging level.

5. Summary and Conclusion

This report describes the construction and calibration of SWAT flow and sediment model for a portion of the Spring River Watersheds upstream from Empire Lake. The modeled watershed, comprised of Spring River and the Shoal Creek, is located within the TSMD, which is known for its legacy of mining activity performed for about 100 years (1850-1970). The SWAT watershed flow model was examined for impact of input data resolution (climate, topography, and soil) on its performance prior to calibration. Among the various data categories, climate data resolution had the greatest impact when compared to DEM and soil data.

A combination of 10 m DEM, SSURGO soil data, and PRISM climate data yielded the best performance of SWAT in terms of simulated streamflow rate before attempting to calibrate the model. This step increased our confidence in the model and insured a proper model calibration. A significant change in model performance was observed with the climate data compared to other geospatial data inputs. SWAT flow parameter sensitivity analysis was implemented using the SUFI-2 algorithm, and the model was successfully calibrated and validated at the two USGS gauge stations located upstream from the outlets of the Spring River Watershed and the Shoal Creek Watershed, meeting recommended thresholds of commonly used performance measures. In both watersheds, the model explained more than 67% of the variance in the observed flow data at the daily time scale and more than 76% of the variance in the observed data at the monthly time scale. The model performed well during wet and relatively dry periods. FDCs of SWAT simulated streamflow rates and the observed data showed that the model performed well at low and high flows, with more pronounced deviations in the range 50 to 350 m^3/sec in Spring River and 20 to 250 m^3/sec in Shoal Creek.

A sediment transport model was also constructed and calibrated. Sensitivity analysis and calibration were conducted for both streamflow and sediment transport models using SWAT-CUP SUFI method. The calibration was achieved using three years-worth of biweekly suspended sediment concentration data (2014-2016) sampled from stations in seven different tributaries upstream from Empire Lake. Sediment loading was successfully calibrated at the Spring River Watershed and most of the tributaries, but the relatively short observed sediment record precluded further evaluation of the model. The model explained 92% and 58% of the variances in the observed data at Spring River and Shoal Creek, respectively. Even though overall acceptable, model performance at Shoal Creek was less adequate than at Spring River. The model was adequately calibrated at mainstem Spring River tributaries, with R^2 values ranging from 0.69 and 0.99, explaining more than 70% of the variance in the observed sediment data.

Using the calibrated watershed model, annual average sediment loading from the Spring River and Shoal Creek to Empire Lake and interior sub watersheds were estimated for the period (2010-2016). The two largest sub-watersheds, the Spring River and Shoal Creek, contributed about 74% of the annual sediment loading, with the former delivering 52% of the sediments and the latter contributing 21% of the loading from areas that are substantially affected by historical lead and zinc mining (Juracek and Drake, 2016). While tributaries within (or near) the Cherokee County Superfund site, namely, Short Creek, Center Creek, and Turkey Creek, have contributed 15% of annual sediment loading over the study period, they drain areas that are substantially affected by historical lead and zinc mining (Juracek and Drake, 2016).

Two hypothetical remedial measures of metal contamination were investigated: lake sediment dredging and sediment traps. Calculations based on SWAT simulated sediment loadings and observed sediment data showed that it may take more than 100 years to fill Empire lake with a dredged lake sediment mass of 2640 million ibs. Mass balance analysis using suspended sediment concentration data sampled in 2014-2016 directly downstream from Empire Lake reservoir and SWAT simulated sediment loading to the lake revealed a substantial amount of the sediment being flushed out of the reservoir in 2015, thus, reducing the mass of sediment to be dredged and increasing the capacity for sediment storage in year 2016 and perhaps the following years.

Qualitative assessment of efficiency of sediment traps as a potential remedial strategy for contaminated sediments was explored using SWAT computed annual average sediment loading for 2014-2016 and published literature on historical lead and zinc concentrations in mining-affected tributaries. While installation of sediment traps in Short, Center, and Turkey Creeks may reduce less than 14% of annual average sediment loading to Empire Lake (based on 2014-2016 data), these tributaries historically have been associated with high concentrations of dissolved and sediment-bound zinc and lead. However, efficacy of sediment filtration in reducing metals input to Spring River is limited by the percentage of fine sediment particles and percentage of total lead and zinc in dissolved phase.

It should be noted that the evaluation of the lake dredging scenario is solely based on the 2014-2106 sediment data records and estimates of lake sedimentation rates obtained from a previous USGS study and the data record at hand. The limitation and uncertainty in the results should therefore be recognized. Observed suspended sediment concentrations (SSC) used for model calibration are point estimates which are compared to SWAT computed cross-section area-averaged concentrations during a calibration. This scale discrepancy along with a relatively short observed SSC record (3 years) made SWAT sediment calibration more difficult and may have contributed to uncertainty in model simulated values as well as less than adequate sediment calibration at Shoal Creek. A longer SSC data record (more than 5 years) would be ideal for improved calibration and validation of the sediment model.

Modeling of lake-wide sediment transport in Empire Lake and mass balance (net sedimentation or export) in stream reaches downstream from the mining-affected tributaries, although a formidable undertaking, can further benefit the analysis and provide more insights on the fate and transport of contaminated sediments in the Spring River Watershed. However, the results of this modeling study identified major contributing areas for sediment and with literature reported heavy metal concentrations in the TSMD could be used to inform management decisions on potential remedial measures for clean-up of mining-affected areas and contaminated sediments dispersed in the floodplains of the Spring River Watershed and Empire Lake.

5 References

Abbaspour, K. C., Johnson, C., Van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal 3, 1340-1352.

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R., 2007. Modelling hydrology and water quality in the pre-ailpine/alpine Thur watershed using SWAT. Journal of Hydrology 333, 413-430.

Adamski, J. C., Petersen, J. C., Freiwald, D. A., Davis, J.V., 1995. Environmental and hydrologic setting of the Ozark Plateaus study unit, Arkansas, Kansas, Missouri, and Oklahoma. US Geological Survey Water-Resources Investigations Report, 94: 4022.

Allen, R. G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9): D05109.

Angelo, R.T. 2007. Residual effects of lead and zinc mining on freshwater mussels in the Spring River Basin (Kansas, Missouri, and Oklahoma, USA). Science of the Total Environment, 384(1): 467-496.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., Williams, J. R., 1998. Large area hydrologic modeling and assessment, Part 1: Model Development. JAWRA Journal of the American Water Resources Association, 34(1): 73-89.

Barks, J. H., 1986. Effects of abandoned lead and zinc mines and tailings piles on water quality in the Joplin area, Missouri. Guidebook to the geology and environmental concerns in the Tri-Sate Lead-Zinc District, Missouri Kansas Oklahoma association of Missouri geologists : 60.

Bennett, N. D., Croke, B. F., Guariso, G., Guillaume, J. H., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T., Norton, J. P., Perrin, C., 2013. Characterising performance of environmental models. Environmental Modelling & Software 40, 1-20.

Beyer, W., Dalgarn, J., Dudding, S., French, J., Mateo, R., Miesner, J., Sileo, L., Spann, J., 2004. Zinc and lead poisoning in wild birds in the Tri-State Mining District (Oklahoma, Kansas, and Missouri). Archives of Environmental Contamination and Toxicology 48, 108-117.

Boughton, W., 1989. A review of the USDA SCS curve number method. Soil Research, 27(3): 511-523.

Briak, H., Moussadek, R., Aboumaria, K., Mrabet, R., 2016. Assessing sediment yield in Kalaya gauged watershed (Northern Morocco) using GIS and SWAT model. International Soil and Water Conservation Research 4, 177-185.

Brosius, L., and Sawin, R. S., 2001. Lead and zinc mining in Kansas Geologic Survey Public Information Circular 17.

Brown, S., Mahoney, M., Sprenger, M., 2014. A comparison of the efficacy and ecosystem impact of residual-based and topsoil-based amendments for restoring historic mine tailings in the Tri-State mining district. Science of the Total Environment, 485: 624-632.

Chow, V.T., D. R. Maidment, and L. W. Mays. 1988. "*Applied Hydrology*." McGraw-Hill, New York, pp. 572.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International journal of climatology 28, 2031-2064.

Davis, J.V., and Schumacher, J.G. 1992. Water-quality characterization of the Spring River Basin, southwestern Missouri and southeastern Kansas: U.S. Geological Survey Water-Resources Investigations Report 90-4176, 112 p.

Di Luzio, M., Johnson, G.L., Daly, C., Eischeid, J.K., Arnold, J.G., 2008. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. Journal of Applied Meteorology and Climatology 47, 475-497.

Dingman, S.L. 2002. "Physical Hydrology." Second Edition. Waveland Press, Inc. 646 p.

England, J. F. Jr, Velleux, M. L., Julien, P.Y. 2007. Two-dimensional simulations of extreme floods on a large watershed. Journal of Hydrology 347(1–2): 229–241, OI:10·1016/j.jhydrol.2007·09·034.

FitzHugh, T. W., Mackay, D., 2000. Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. Journal of hydrology, 236(1): 35-53.

Galván, L., Olías, M., Fernandez de Villarán, R., Domingo Santos, J. M., Nieto, J. M., Sarmiento, A. M., and Cánovas, C. R., 2009. "Application of the SWAT model to an AMD-affected river (Meca River, SW Spain). Estimation of transported pollutant load." Journal of hydrology 377, no. 3-4 445-454.

Gassman, P. W., Sadeghi, A. M., Srinivasan, R., 2014. Applications of the SWAT model special section: overview and insights. Journal of Environmental Quality 43, 1-8.

Herschfield, D., 1961. Rainfall frequency atlas of the United States. US Weather Bureau Technical Paper, 40.

Jakubauskas, M., 2008. Methods for assessing sedimentation in reservoirs. Sedimentation in Our Reservoirs: Causes and Solutions: 25.

Johnson, B. E., Zhong, Z., 2006. CTT&F: Distributed Sources Chemical Transport, Transformation and Fate Sub-model, ERDC TN-EQT-06-1. U.S. Army Engineer Research and Development Center: Vicksburg, MS; 17.

Juracek, K. E., 2006. Sedimentation and occurrence and trends of selected chemical constituents in bottom sediment, Empire Lake, Cherokee County, Kansas, 1905-2005. US Geological Survey.

Juracek, K. E., Drake, K., 2016. Mining-Related Sediment and Soil Contamination in a Large Superfund Site: Characterization, Habitat Implications, and Remediation. Environmental management, 58(4): 721-740.

Liu, R., Xu, F., Zhang, P., Yu, W., Men, C., 2016. Identifying non-point source critical source areas based on multi-factors at a basin scale with SWAT. Journal of Hydrology 533, 379-388.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and, Veithet, T. L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the Asabe, 50(3): 885-900.

Mukundan, R., Pierson, D. C., Schneiderman, E. M., Zion, M. S., 2015. Using detailed monitoring data to simulate spatial sediment loading in a watershed. Environmental monitoring and assessment, 187(8): 532.

Nash, J. E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models part I -- A discussion of principles. Journal of Hydrology, 10(3): 282-290.

Neitsch, S. L., Arnold, J. G., Kiniry, J., and, Williams, J. R., 2005. Soil and water assessment tool theoretical documentation , USDA Agricultural Research Service and TexasA&MBlackland Research Center, Temple, Texas.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Texas Water Resources Institute.

Phomcha, P., Wirojanagud, P., Vangpaisal, T., Thaveevouthti, T., 2012. Modeling the impacts of alternative soil conservation practices for an agricultural watershed with the SWAT model. Procedia Engineering 32, 1205-1213.

Pope, L. M., 2005. Assessment of contaminated streambed sediment in the Kansas part of the historic Tri-State Lead and Zinc Mining District, Cherokee County, 2004. 2328-0328.

Ritter, A., and R. Muñoz-Carpena, 2013. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. Journal of Hydrology 480, 33-45.

Smith, D. C., 2016. Occurrence, distribution, and volume of metals-contaminated sediment of selected streams draining the Tri-State Mining District, Missouri, Oklahoma, and Kansas, 2011–12: U.S. Geological Survey Scientific Investigations Report 2016–5144, 86 p., http://dx.doi.org/10.3133/sir20165144.

Spruill, T. B., 1987. Assessment of water resources in lead-zinc mined areas in Cherokee County, Kansas, and adjacent areas: U.S. Geological Survey Water-Supply Paper 2268, 68 p.

Stratus Consulting Inc. 2006. Preliminary Evaluation of Mining-Related Injuries in the Cherokee County Superfund Site, Cherokee County, KS. Report prepared for Kansas Department of Health and Environment and U.S. Department of the Interior..

Tripathi, M., Panda, R., and Raghuwanshi, N., 2003. Identification and prioritization of critical sub-watersheds for soil conservation management using the SWAT model. Biosystems Engineering 85, 365-379.

USDA Natural Resources Conservation Service. 2006. “Land Resource Regions and Major Land Resource Areas of the United States, the Caribbean, and the Pacific Basin.” United States Department of Agriculture Handbook 296.

US Environmental Protection Agency (USEPA). 1995. “*Watershed Protection: A Statewide Approach*.” EPA 841-R-95-004. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA, 2006. National water quality inventory report to congress. http://www.epa.gov/305b, accessed June 2017.

Velleux, Mark L., Julien, P. Y., Rojas-Sanchez, R., Clements, W. H., and England, J. F. 2006. "Simulation of metals transport and toxicity at a mine-impacted watershed: California Gulch, Colorado." Environmental science & technology 40, no. 22 (2006): 6996-7004.

Vrugt, J. A., and Robinson, B. A., 2007. Improved evolutionary optimization from genetically adaptive multimethod search. Proceedings of the National Academy of Sciences of the United States of America 104, 708-711.

Williams, J., and Berndt, H., 1977. Sediment yield prediction based on watershed hydrology. Transactions of the ASAE, 20(6): 1100-1104.

Supplementary Information:

This section includes the supplementary information related to the main report.

PRISM Dataset

Parameter-elevation Relationships on Independent Slopes Model (PRISM) data (Daly et al.,2008; Di Luzio et al., 2008) available with a grid size of 4 km with a full spatial extent of the U.S. for the period from 1981 to present (http: //www.prism.oregonstate.edu/). For the TSMD case study time series of daily precipitation and temperature (min and max) from 1981 was extracted and formulated for SWAT input. One of the main reasons to use PRISM data for TSMD is because of its availability for the recent days. The National Centers for Environmental Prediction (NCEP) based Climate Forecast System Reanalysis (CFSR) data are available on the SWAT model website and cover a 36-year period of 1979 through 2014. For suspended sediment and chemical concentration, the sampling record extends from 2014 to the present. Therefore, PRISM is the alternate option from the ground observations. Figure S1 in the supplementary section displays the NOAA based ground observation points against the PRISM grids. 381 grid points covers the entire TSMD study area, while there are only 6 observation points available from NOAA.

PRISM uses a specified interpolation technique called climatologically aided interpolation (CAI). Starting on January 1, 2002, a combination of CAI and Doppler radar data is used in the central and eastern U.S. A number of observer station network data that adhere to the "PRISM day" criterion is included in the PRISM dataset. In PRISM, a climate-elevation regression is calculated for each digital elevation model (DEM) grid cell, and stations entering the regression are assigned weights based on the physiographic similarity of the station to the grid cell. Factors accounted for PRISM based reanalysis are location, elevation, coastal proximity, topographic facet orientation, vertical atmospheric layer, topographic position, and orographic effectiveness of the terrain. A full description of PRISM can be found from (Daly et al.,2008). A function developed for SWAT input from PRISM using R statistical tool which automates the large number of observations to SWAT format.

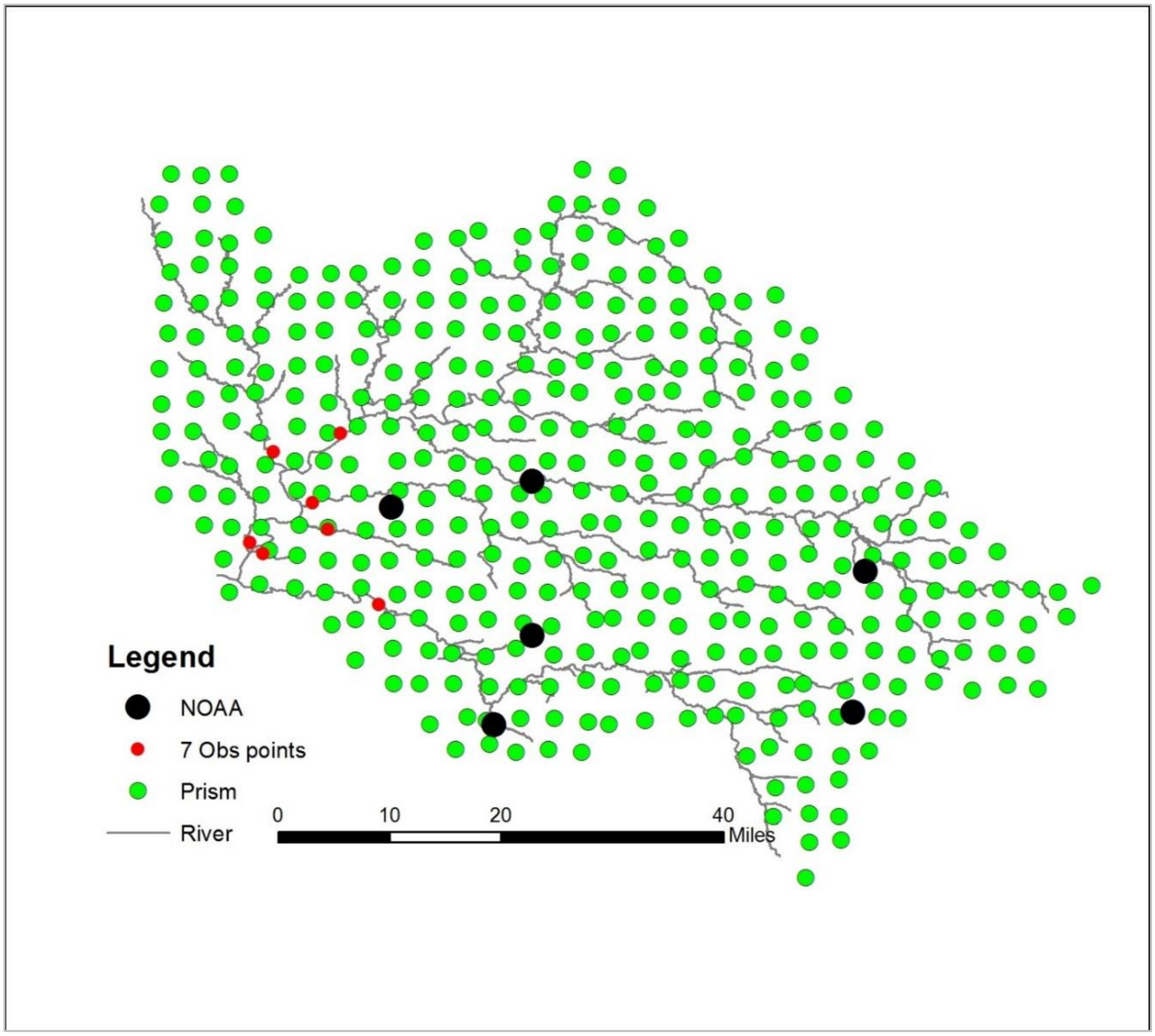

Figure S1: PRISM based 4 km grid points and the NOAA based ground stations.

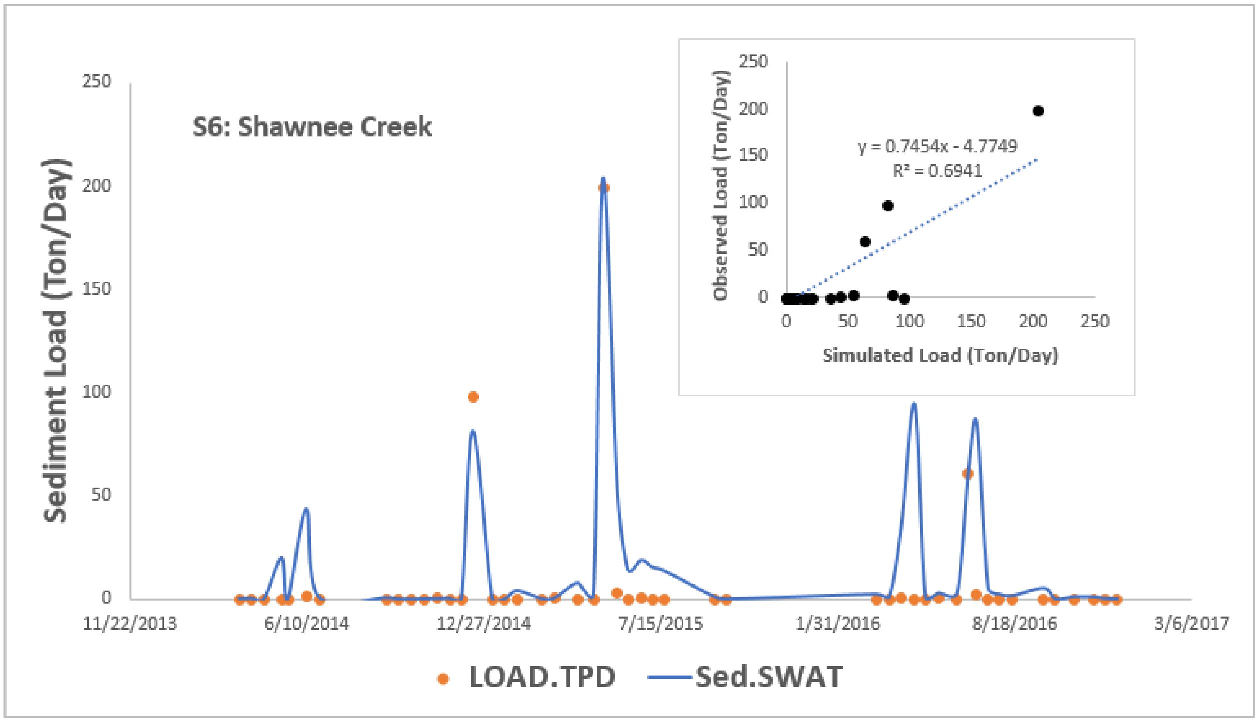

Figure S2: Observed vs. Simulated Sediment loading in the Site 6 (Shawnee Creek).

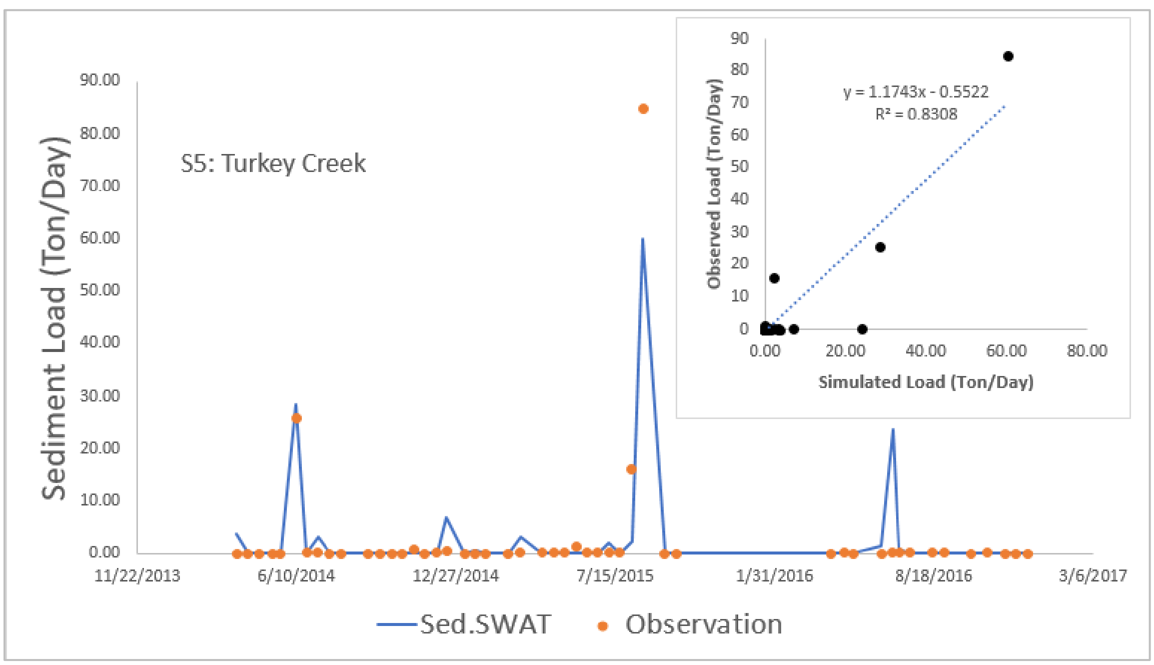

Figure S3: Observed vs. Simulated Sediment loading in the Site 5 (Turkey Creek).

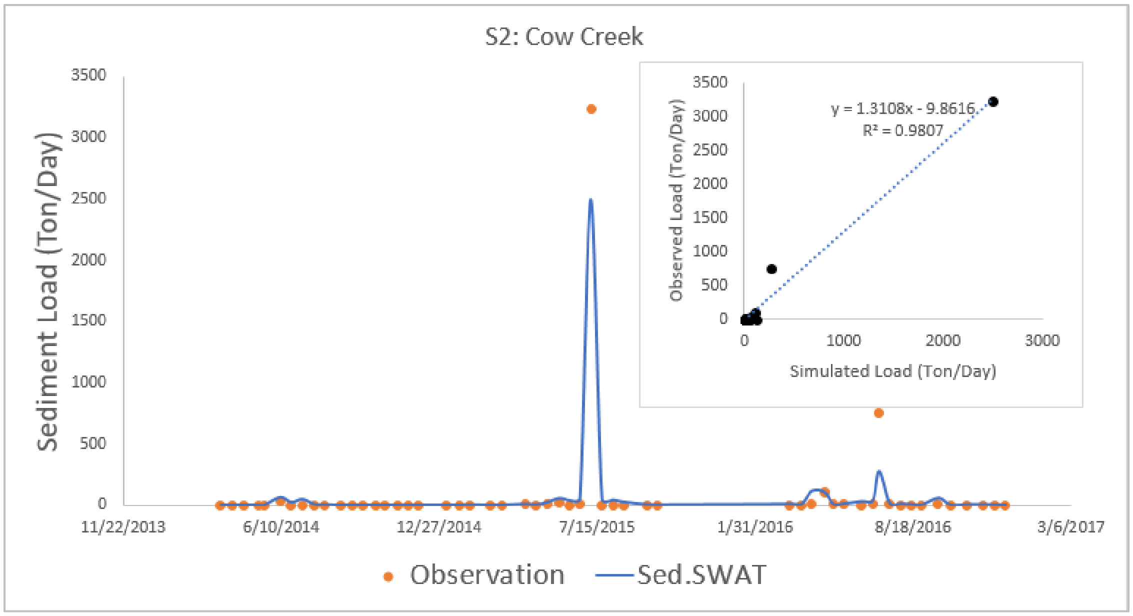

Figure S4: Observed vs. Simulated Sediment loading in the Site 2 (Cow Creek).

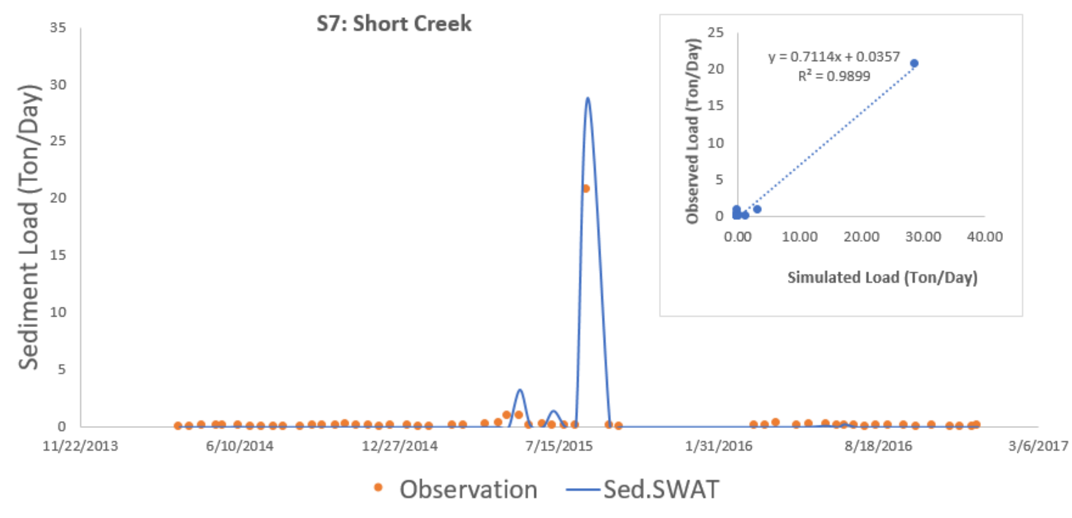

Figure S5: Observed vs. Simulated Sediment loading in the Site 7 (Short Creek).

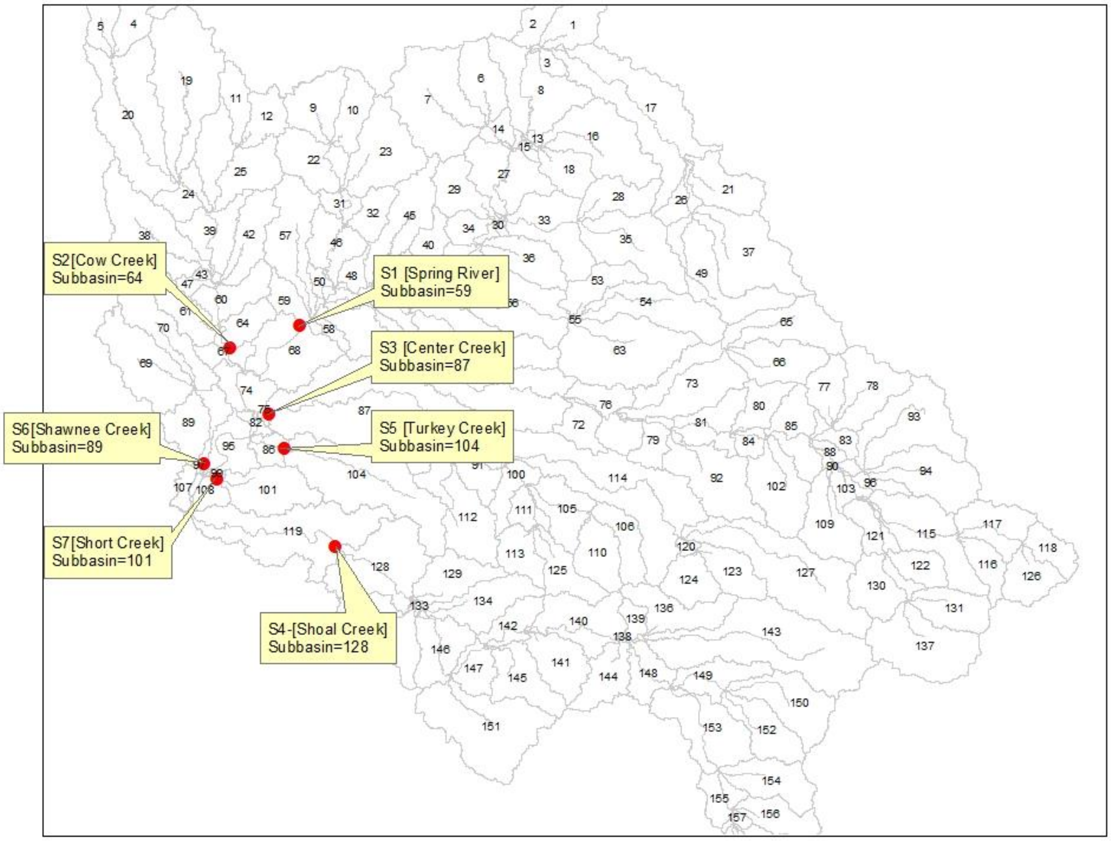

Figure S6: 159 Sub watersheds used for this study. Red dot points describe the observation points.

Figure S7 depicts the role of climatic input (for example precipitation) in sediment loading and flow generation process. The upper panel represents the suspended sediment concentration measurement and the lower panel represents flow in primary axis. In both panels, the y-axis on the right shows rainfall depth expressed in units of mm. Blue lines represent PRISM based satellite rainfall estimates and yellow lines represents ground data. We can see in some events PRISM estimates higher rainfall than observations. But, observed flow and sediment data do not corroborate a rainfall event. In another event (07.15.2015), PRISM produces better results than ground observations. This is due to the distance of the ground observation data from the measurement point.

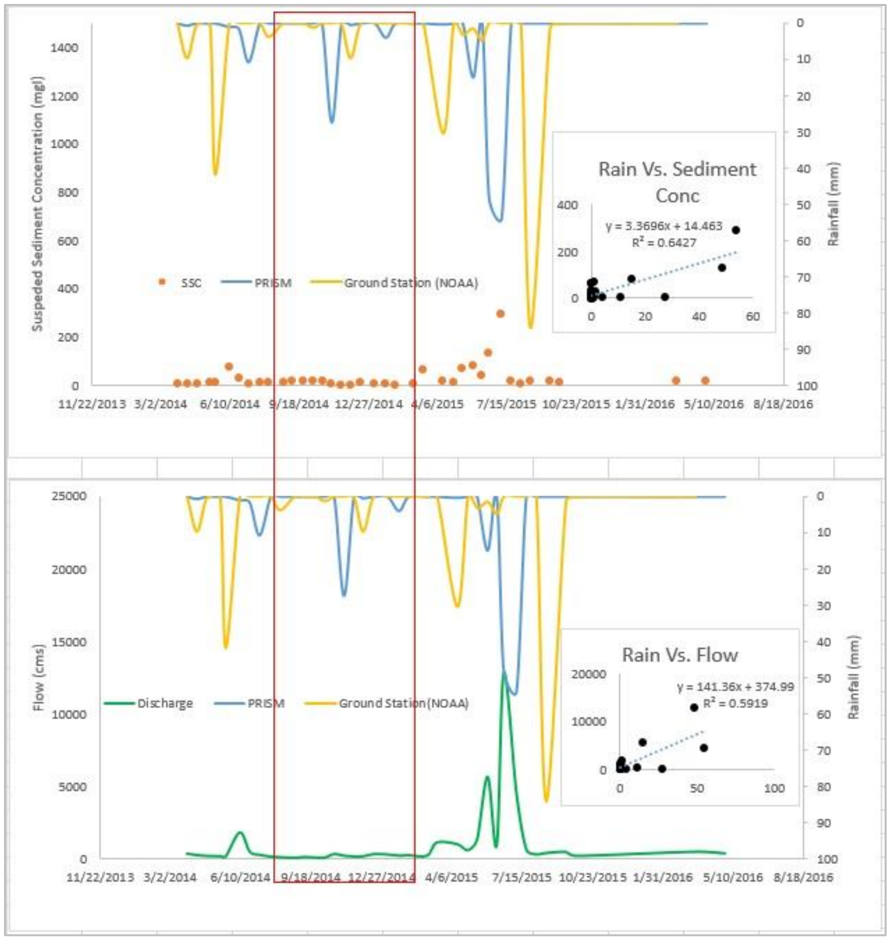

Figure S7: Importance of climatic input on flow and sediment loading.

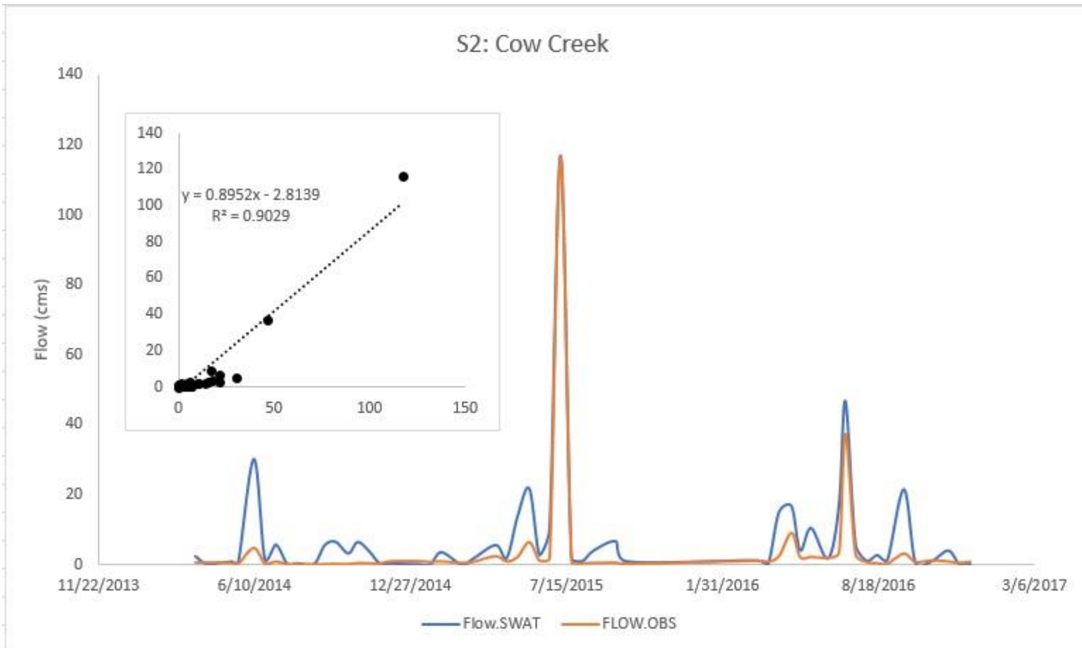

Figure S8: Cross validation of the flow in Cow Creek.

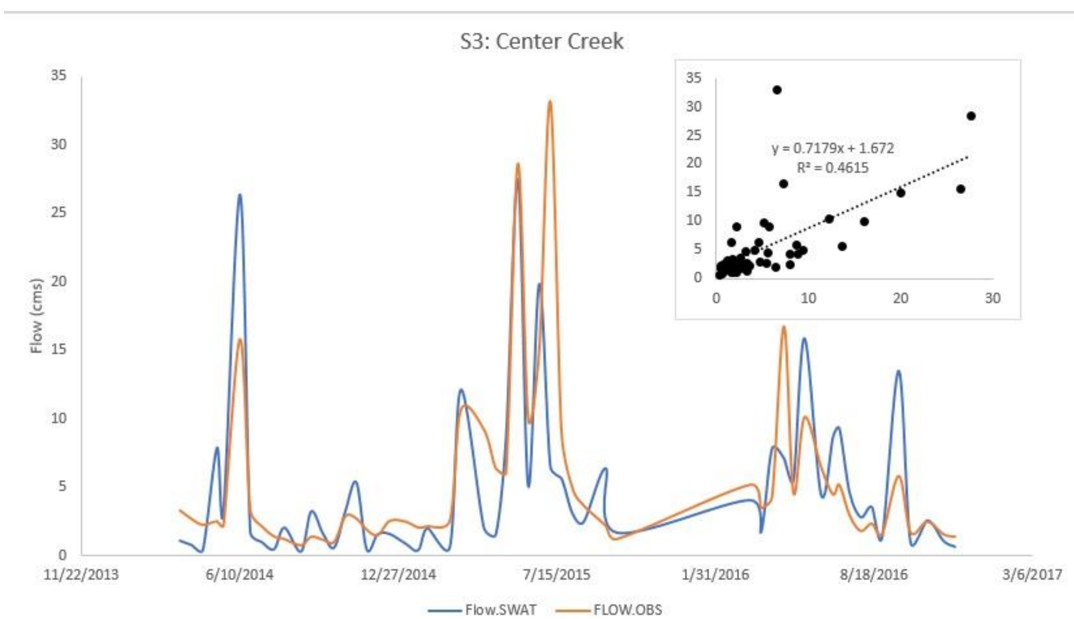

Figure S9: Cross validation of the flow in Center Creek.

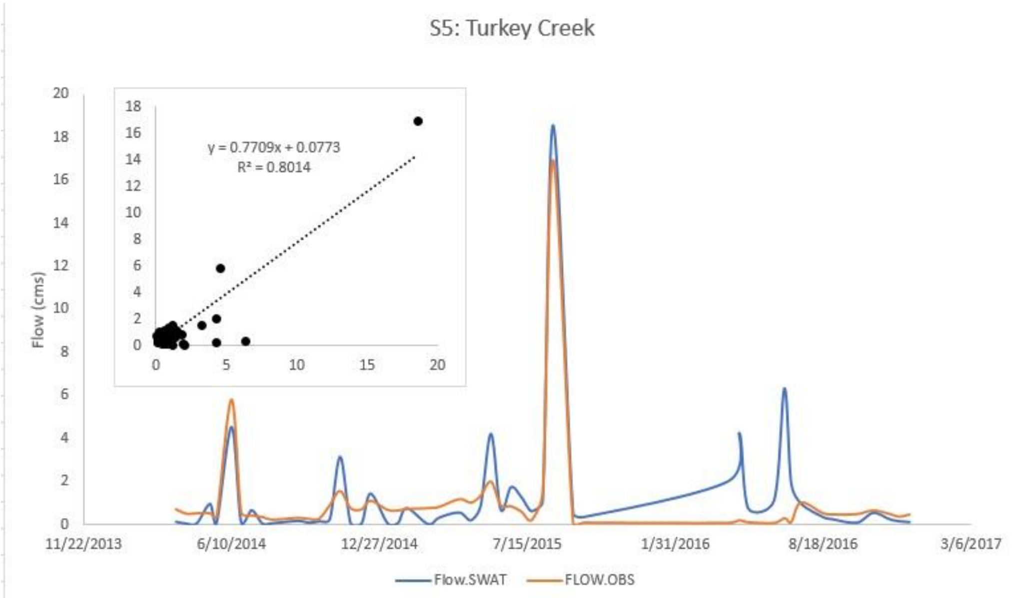

Figure S10: Cross validation of the flow in Turkey Creek.

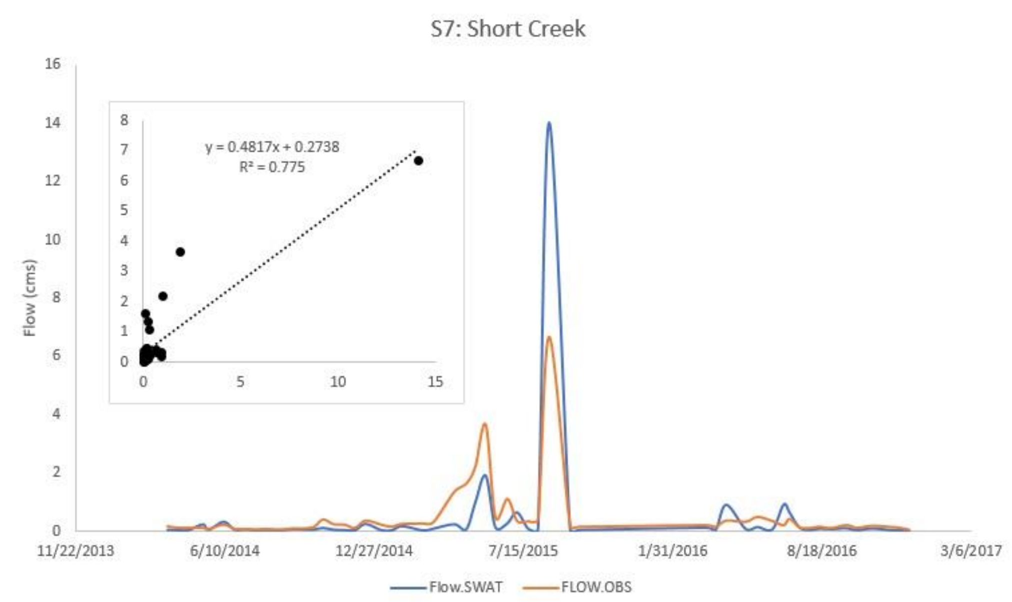

Figure S11: Cross validation of the flow in Short Creek.

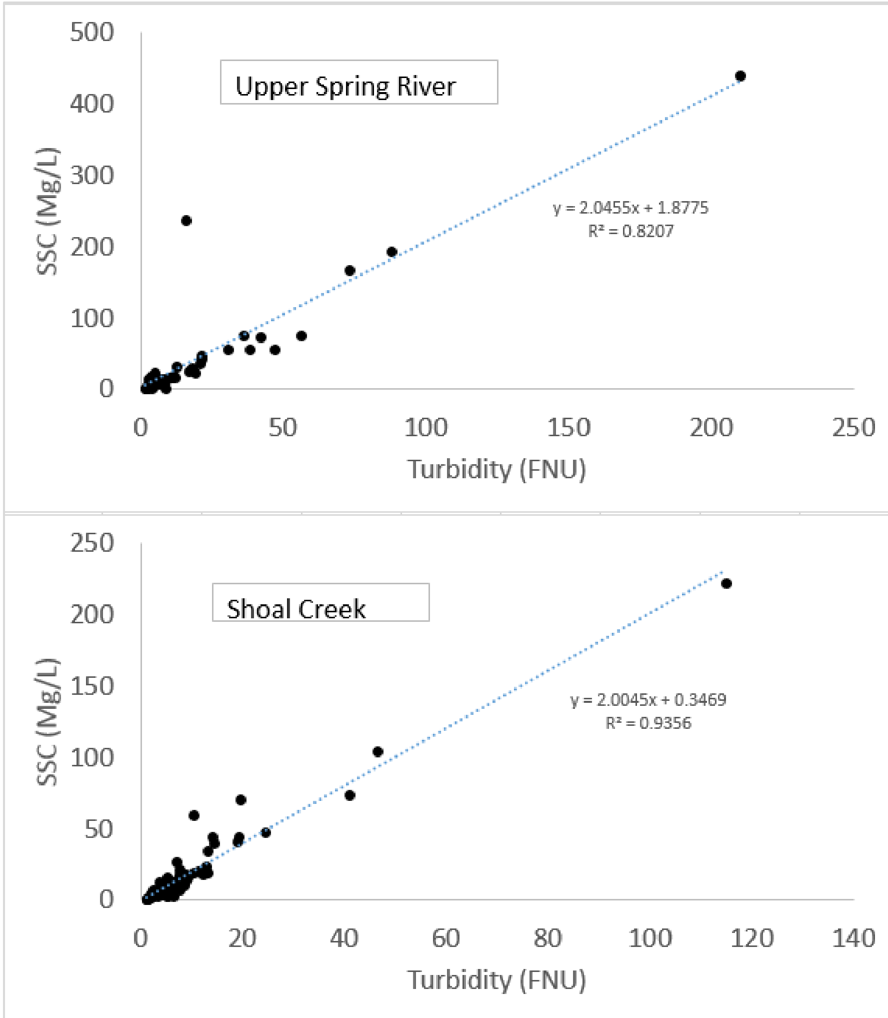

Figure S12. Relationship between turbidity and suspended sediment concentration in upper Spring River and Shoal Creek.

Turbidity, measured in Formazan Nephelometric Unit (FNU) has a high correlation with suspended sediment concentration. The upper panel describe the correlation between turbidity and suspended sediment concentration in upper Spring River and the lower panel describes the correlation at Shoal Creek. A higher correlation indicates the strength of the correlation of the variables. Usually turbidity and suspended sediment has linear correlation.