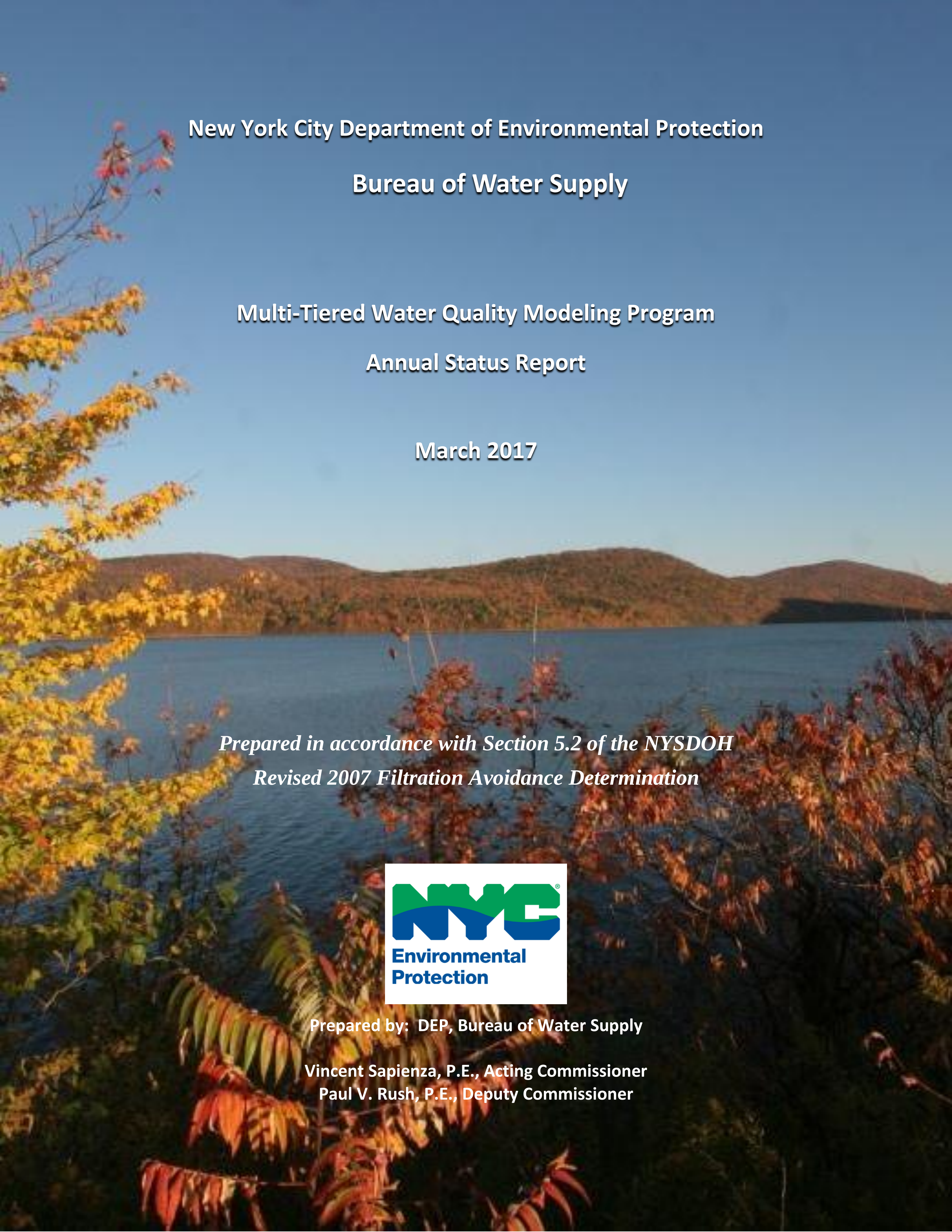


New York City Department of Environmental Protection

Bureau of Water Supply

Multi-Tiered Water Quality Modeling Program

Annual Status Report

March 2017

Prepared in accordance with Section 5.2 of the NYSDOH
Revised 2007 Filtration Avoidance Determination

Prepared by: DEP, Bureau of Water Supply

Vincent Sapienza, P.E., Acting Commissioner
Paul V. Rush, P.E., Deputy Commissioner

Table of Contents

List of Figures

List of Tables

List of Acronyms

AGNPS	Agricultural Non-Point Source Pollution model
AnnAGNPS	Annualized Agricultural Non-Point Source model)
AWBM	Australian Water Balance Model
BLUE	Blue-green light
BMPs	Best management practices
BWS	Bureau of Water Supply
CCIMP	Climate Change Integrated Modeling Project
CFM	Change factor methodology
CFSR	Climate forecast system reanalysis
CN	Curve number
CSSO	Conditional seasonal storage objective
CUNY	City University of New York
DBP	Disinfection by-product
DEM	Digital elevation model
DEP	New York City Department of Environmental Protection
DOC	Dissolved organic carbon
EOH	East of Hudson
EBDR	East Branch Delaware River
EVI	Enhanced vegetation index
GCM	General circulation model
GLUE	Generalized likelihood uncertainty estimation
GWLF	Generalized Watershed Loading Function model
HEC-HMS	Hydrologic Engineering Center Hydrologic Modeling System
HEFS	Hydrologic ensemble forecast service
HRU	Hydrologic response unit
KGE	Kling Gupta efficiency coefficient
LAI	Leaf area index
Landsat-TM	Landsat Thematic Mapper
LIDAR	light detection and ranging
MAE	Mean absolute error
MIR	Mid-infrared light
MT-CLIM	Mountain microclimate simulation model
NCDC	National Climate Data Center
NDMI	Normalized different moisture index
NDVI	Normalized difference vegetation index
NIR	Near infrared light
NLS	Nonlinear least square
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe efficiency
NTU	Nephelometric turbidity unit

OST	Operations Support Tool
PA	Position analysis
PDM	Probability distributed model
PBIAS	Percent bias
PRISM	Parameter-Elevation Relationships on Independent Slopes Model
RED	Red light
RHESSys	Regional Hydro-Ecologic Simulation System
RMSE	Root mean square error
RoboMon	Robotic monitoring
RondoutPA	Rondout Reservoir position analysis water quality model
RSR	RMSE – observations standard deviation ratio
SCS	Soil Conservation Service
SMR	Soil moisture routing
SOC	Soil organic content
SSURGO	Soil Survey Geographic database
STP	Shandaken Tunnel Portal
SWAT	Soil and Water Assessment Tool
SWAT-CUP	SWAT Calibration and Uncertainty Programs
SWMM	Stormwater Management Model
SWG	Stochastic weather generator
TI	Topographic index
WBDR	West Branch Delaware River
WOH	West of Hudson
WQD	Water Quality Directorate
WQSR	Water Quality Science and Research
WSC	Wind sheltering coefficient
USGS	United States Geological Survey
VIC	Variable infiltration capacity
VSA	Variable source area
VIS	Visible light

Acknowledgements

This report was produced by the Water Quality Modeling section of the Division of Water Quality Science and Research (WQSR), Directorate of Water Quality, in the Bureau of Water Supply at DEP.

As in past years, post-doctoral researchers employed by the City University of New York (CUNY) and working as a part of the Water Quality Modeling section contributed significantly to the work described in this report. These individuals, and the report sections which they authored, are: Nachiketa Acharya, Ph.D. (Section 3.2), Linh Hoang, Ph.D. (Section 4.1), and Kyongho Son, Ph.D. (Sections 4.2 through 4.4). Dr. Acharya left his CUNY position in Sept. 2016 and was replaced by Myeong-Ho (Chris) Yeo, Ph.D. in early 2017. In addition, Theodore Kpodonu, Ph.D. also joined the post doc group in early 2017. These researchers are supported by the following faculty advisors: Allan Frei, Ph.D. (Hunter College, CUNY), Tammo Steenhuis, Ph.D. (Cornell University), Paul Hanson, Ph.D. (University of Wisconsin), and Larry Band, Ph.D. (University of North Carolina).

All staff in the Water Quality Modeling section contributed to this report. Jordan Gass, GIS Specialist, was the author of Section 6. Rakesh Gelda, Ph.D., Research Scientist, was the author of Sections 2, 4.5, and 5, while Rajith Mukundan, Ph.D., Research Scientist, was the author of Sections 3.3, 4.6, 4.7, and 4.8. Eliot Schneiderman Ph.D., a former full-time staff, rejoined the section in 2016 on a part-time basis, and reviewed various sections of this report. Emmet Owens, P.E., Section Chief of Water Quality Modeling was responsible for overall report editing and production. Lorraine Janus, Ph.D., Chief of Water Quality Science and Research, and Steven Schindler, Director of Water Quality, provided overall support and guidance for the Water Quality Modeling program and completed final reviews of this report.

Executive Summary

This report has been prepared in accordance with the provisions of the New York City Department of Environmental Protection (NYCDEP) Revised 2007 Filtration Avoidance Determination (FAD). This determination requires NYCDEP to prepare an annual report that summarizes the progress made in water quality modeling and support activities during a particular calendar year. This report covers 2016.

NYCDEP's water quality modeling program is designed to develop, test, validate, and apply mathematical models related to City's drinking water supply. These models are generally in the areas of climate science, watershed and terrestrial modeling, reservoir modeling, and water supply system operations modeling. During the development, testing, and validation phases, the various models are commonly applied and operated individually. Once individual models have proven to accurately represent components of the water supply system, the models are then applied in a multiple-tier or linked manner.

The problem of episodic increases in turbidity in the water supply has been studied in a comprehensive manner in the past, and has led to validated predictive watershed and reservoir water quality models. These models are the Operations Support Tool (OST) and the Roudout Reservoir Position Analysis (RondoutPA) model. In 2016, these models were applied to provide guidance to DEP regarding the operation of the water supply system in response to events or episodes of elevated turbidity. A discussion of model applications for several of the episodes that occurred in 2016 are presented and discussed.

The Bureau of Water Supply's (BWS's) Climate Change Integrated Modeling Project (CCIMP) encompasses an effort to develop and apply a suite or multi-tiered group of models to study the impact of climate change on the water supply. Phase I of CCIMP concluded in 2013 with a number of "first-cut" or screening level modeling analyses. In 2016, progress continued on Phase II of CCIMP, which involves the study of climate impacts using more sophisticated, realistic, and complex modeling approaches and tools. A stochastic weather generator (SWG) has been developed for the City's West of Hudson (WOH) watersheds. This weather model allows prediction of a long synthetic time series of weather (such as precipitation and air temperature) conditions, the statistics of which closely match observed conditions, but which also include infrequently occurring events (e.g. floods and droughts) that have not been captured in monitoring. SWG's offer the promise of a more complete and realistic prediction of the changes that may occur in future decades to the local climate in the watersheds. Section 3.2 of this report is effectively a reprint of a peer-reviewed paper published in the *Journal of Hydrometeorology* in March 2017 describing this effort. In 2016, the first multi-tier application of the SWG was completed by using numerous weather time series generated by the SWG as input to the Generalized Watershed Loading Function (GWLF) model to predict streamflow. An evaluation of the resulting predictions of streamflow, including a comparison of the statistics of predicted and observed flows, is presented.

In addition to more realistic climate predictions, Phase II of CCIMP will also utilize more realistic and complex watershed models. This report contains sections describing the testing and validation of two such watershed models. The Soil and Water Assessment Tool (SWAT) has been applied and tested for the Town Brook watershed, located within the larger Cannonsville Reservoir watershed. Section 4.1 is largely a reprint of a peer-reviewed paper that has been accepted by the journal *Hydrologic Processes,* describing the application of SWAT to Town Brook. SWAT has been shown to be accurate in hindcasting of observed streamflow from the Town Brook watershed on daily, monthly, and seasonal time scales. The particular version of SWAT used in this work was developed by DEP staff in order to allow mechanistic simulation of saturation-excess runoff, a runoff component that is not considered in the standard version of SWAT. Sections 4.2 through 4.4 describe the application of the Regional Hydro-Ecologic Simulation System (RHESSys) to two watersheds (Biscuit Brook and Shelter Creek) in the watershed of Neversink Reservoir. While considerable progress has been made in the testing of RHESSys for these watersheds, more work needs to be done in 2017 before this model can be applied.

The water quality modeling group has begun to extend the application of the two-dimensional model CE-QUAL-W2, including a turbidity sub-model, to Neversink, Cannonsville, and Pepacton Reservoirs. An analysis of meteorological data for these reservoirs and associated watersheds that is presented here as a preliminary step to reservoir model application and testing. Preliminary testing of the CE-QUAL-W2 model was successfully completed in 2016, as described in Section 4.5. As in past applications to Schoharie, Ashokan, Rondout, and Kensico Reservoirs, CE-QUAL-W2 was able to accurately capture the response of Neversink Reservoir to episodes of elevated turbidity loading from surface runoff. Work that supported this turbidity model application for Neversink that was completed in 2016 include the development of empirical temperature and turbidity models for the Neversink River inflow to Neversink Reservoir. These models allow prediction of inflow temperature and turbidity based on streamflow and air temperature.

Work was also begun in 2016 to investigate a method to predict the uncertainty in predictions of turbidity for streams entering West of Hudson reservoirs. For West of Hudson tributary streams, a turbidity rating curve, which relates turbidity to streamflow, has generally been used to quantify turbidity. Application of this approach will predict a single value of turbidity for a given value of streamflow for a particular stream. The current work involves the use of the quantile regression approach, so that a probability distribution for turbidity, rather than a single predicted value, can be generated for a particular streamflow. That distribution can be used to generate inputs to a reservoir turbidity model to generate probabilistic predictions of reservoir and aqueduct turbidity.

Other modeling work completed in 2016 include the testing and application of the watershed model GWLF to watersheds in the East of Hudson (EOH), or Croton, system of reservoirs. Since water quality is of lesser concern in the EOH system due to the filtration plant that serves this system, application of water quality models to the EOH system has not occurred in recent years. However, DEP's Operations Support Tool (OST), which considers water supply from all sources (Catskill, Delaware, and Croton systems), requires stream inflows to EOH reservoirs for historical periods. While historical inflows can be estimated from reservoir water budgets, the data required from such budget calculations is not available for all historic periods of interest. As a result, it would be very useful to have model predictions for such periods. GWLF was applied to gaged EOH watersheds, with calibration of model inputs performed to product good agreement between predicted and measured streamflow. The model was then applied to ungaged watersheds, using model coefficients determined from the application to gaged watersheds. The result is a consistent set of model predictions for the entire EOH system.

The Water Quality Modeling group continues to develop database and data analysis tools to support modeling work. GIS data from a variety of sites and sources continues to be used. Field work for the West of Hudson bathymetry project is complete, and the resulting data will be implemented into DEP models. In addition, a similar project to directly measure new bathymetry for all East of Hudson reservoir basins is scheduled to begin in 2017. The water quality modeling group continues to invest in database to support modeling activities, including a database containing model input datasets, and a database containing time series of predictions from General Circulation Models (GCMs), which will be used together with weather time series resulting from stochastic weather generators to produce synthetic, realistic time series of future weather conditions in the West of Hudson watersheds.

The water quality modeling group continues to be involved with outside groups including the Water Utility Climate Alliance (WUCA). DEP staff attended the WUCA meeting in Boulder CO in May 2016, and participate in regular conference calls with staff from other WUCA members. DEP staff also attended the annual meeting of the Global Lake Ecological Observatory Network (GLEON) in Austria in July 2016. The interaction between post-doctoral researchers employed by the City University of New York (CUNY), university faculty advisors, and DEP staff, all supported by DEP funding, continues to be a major source of ideas, modeling software, modeling products, and reports and publications. The current 4-year contract between the Research Foundation of CUNY and NYCDEP continues until August, 2018.

1. Introduction

This status report describes work completed as a part of the New York City Department of Environmental Protection Multi-Tiered Water Quality Modeling program for the period January through December, 2016. This report was prepared and submitted in accordance with Section 5.2 of the 2007 Revised Filtration Avoidance Determination (NYSDOH, 2014).

A schematic overview of DEP's Water Quality Modeling program is shown in Figure 1.1. The general categories of models used are climate (weather generators), watershed or terrestrial, reservoir, and water supply system operations models. Issues that are considered are water quantity (including droughts), turbidity, disinfection by-products, and eutrophication and harmful algal blooms. Application of these models may lead to changes in watershed regulations and reservoir management, adoption or modification of long-term plans for the water supply, guidance for short term system operations, and evaluation of the adequacy of current infrastructure.

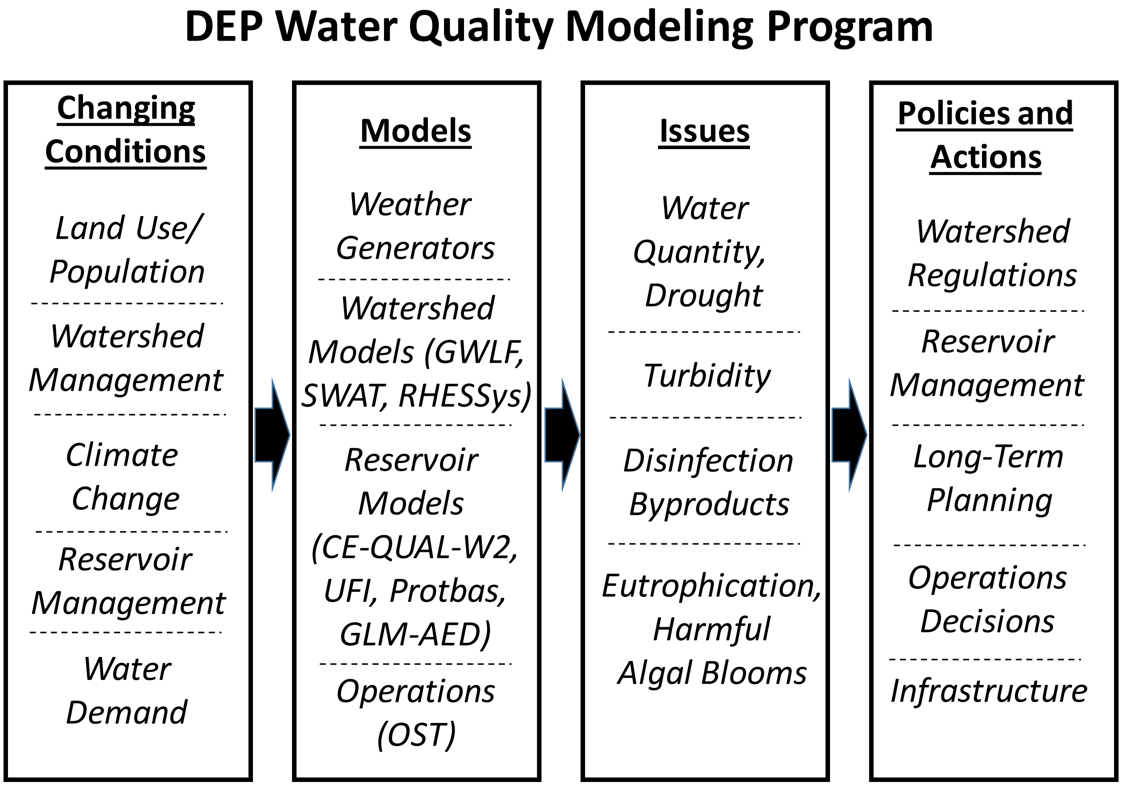


Figure 1.1. Schematic of the conditions considered, models, issues, and policies and actions that are a part of DEP's Water Quality Modeling program.

DEP is developing, testing, and applying a suite of weather, watershed/terrestrial, reservoir, and water system operations models in our water quality modeling program. In 2016, the development and testing of the first stochastic weather generator for use in the watersheds of the Catskill and Delaware system reservoirs was completed. It is expected that application of stochastic weather generators, rather than change factor approaches used in earlier evaluations, will result in more realistic and robust predictions of future conditions, especially future extreme conditions, in the water supply. An initial application of this weather generator was made by generating runoff predictions using the GWLF watershed model. This application is opening the door to the use of "bottom-up" evaluations of the impact of climate change on the water supply system.

The water quality modeling program continued the development and application of three watershed/terrestrial models in 2016. These models are the Generalized Watershed Loading Function (GWLF), the Soil Water Assessment Tool (SWAT), and the Regional Hydro-Ecologic Simulation System (RHESSys). GWLF is a relatively simple, lumped-parameter watershed model that is considered to be fully tested for the West of Hudson watersheds and is now used in applications for those watersheds. During 2016, GWLF was used to generate runoff predictions using weather data predicted by a stochastic weather generator, as described above. In 2016, GWLF was also applied and tested for East of Hudson watersheds to generate runoff predictions that are used in the Operations Support Tool. The first step in the testing and validation of SWAT was completed in 2016, with the finalizing of runoff predictions for the Town Brook watershed located in the drainage area of Cannonsville Reservoir. SWAT proved to be accurate in reproducing measured outflow hydrographs on both daily and monthly time scales from this watershed. Testing of RHESSys for the Biscuit Brook and Shelter Creek watersheds which drain to Neversink Reservoir continued in 2016. This work involved simulating both runoff and nitrate and organic carbon loading from these watersheds.

In 2016, the water quality modeling section began work on the testing and validation of the CE-QUAL-W2 turbidity model, which has been previously tested for Schoharie, Ashokan, Rondout, and Kensico Reservoirs, to the remaining Delaware system reservoirs: Neversink, Cannonsville, and Pepacton. Progress was made on applying this model to Neversink. Empirical models to predict the temperature and turbidity of inflows to Neversink were completed. These models were then used to generate inputs for the testing and validation of the reservoir turbidity model. This work is largely complete and will be finalized in 2017. While the CE-QUAL-W2 turbidity model has been tested and validated for Rondout Reservoir, that model has not been added to the Operations Support Tool (OST). To allow turbidity model runs to be made for Rondout using the position analysis approach as in OST, the Rondout Position Analysis (RondoutPA) software was developed and applied in 2016. This is essentially an interim product that allows turbidity model forecasts to be made for Rondout that utilize multiple forecasts of meteorology and streamflow for days and weeks into the future at the time of forecast. The position analysis approach allows multiple forecasts of future weather and streamflow conditions to be used.

This report describes the major activities of the water quality modeling group in 2016, including the following:

- Use of the Operations Support Tool (OST) and the Rondout Position Analysis software (RondoutPA) to support and guide operations of Schoharie, Ashokan, Rondout, and Kensico Reservoirs during episodes of elevated turbidity
- Testing, documenting, and initial application of a stochastic weather generator (SWG) for simulating long time series of synthetic but statistically-realistic weather data, for use in climate studies
- Application of the Generalized Watershed Loading Function (GWLF) model in various studies including prediction of runoff from watersheds in the East of Hudson water supply system, and predicting runoff from weather time series produced by the SWG
- Completed testing of the Soil Water Assessment Tool (SWAT) for predicting runoff in the Town Brook watershed draining to Cannonsville Reservoir
- Ongoing testing of the Regional Hydro-Ecologic Simulation System (RHESSys) model for both hydrology and water quality, for two watersheds draining to Neversink Reservoir
- Application and testing of DEP's two-dimensional hydrothermal and turbidity model (CE-QUAL-W2) to Neversink Reservoir
- Development of a procedure for introducing estimation of uncertainty in the turbidity-discharge relationships that are used to quantify turbidity from streams to reservoirs
- Data acquisition, organization, and analysis in the support of modeling
- Collaboration of staff in the water quality modeling group with outside organizations
- Papers appearing in peer-reviewed publications and presentations made at professional meetings that describe the work of the water quality modeling group

2. Use of Models for Support of Operational Decisions

Major runoff events in the Catskill and Delaware watersheds can impair water quality in receiving reservoirs. Model simulations to support operational decisions during these events are usually related to elevated turbidity levels that occurred in Cannonsville, Rondout, Schoharie, or Ashokan reservoirs. The models used in these evaluations in 2016 are the Operations Support Tool (OST), which has the capability to simulate turbidity in Schoharie, Ashokan, and Kensico Reservoirs. In addition, the Rondout Position Analysis software (RondoutPA) was used for turbidity simulations of Rondout Reservoir. The simulations usually follow a position analysis strategy that produces an ensemble of probable outcomes and allows managers to make operational decisions in a probabilistic manner. Simulations are intended to answer questions such as: what is the impact of turbid discharge from Cannonsville Reservoir on the withdrawal turbidity from Rondout Reservoir, how should the operations be modified so that the need for adding alum in the Catskill Aqueduct at Kensico Reservoir is reduced or eliminated, what is the optimum turbidity loading to Kensico Reservoir, how should the West Basin of Ashokan Reservoir be drawn down (release channel, dividing weir gates), etc. Selected modeling cases that occurred in 2016 are presented below:

April 7, 2016: With the precipitation moving into the region on April 8, 2016 and the Ashokan West Basin void remaining below 500 MG, there was a question as to whether the dividing weir flow should be increased to 1 BGD to avoid spill, and could this be done without significant impact on the water quality of the East Basin. Also, with Ashokan storage below the CSSO (conditional seasonal storage objective), it would be desirable to increase the flow at Shandaken Tunnel Portal (STP) from 300 MGD to 400 MGD. Any water quality concerns associated with this scenario were to be identified.

State of water quality and issue: The peak flow for Esopus Creek was projected to be about 3500 CFS. On April 7, 2016, both the basins of Ashokan Reservoir were nearly isothermal at about 5.8 °C. The turbidity in the East Basin was 2.3 NTU and 5 NTU in the West Basin. Turbidity in the Schoharie Reservoir withdrawal (Shandaken Tunnel Portal) was 15 NTU. Withdrawal from the East Basin of Ashokan Reservoir was at a rate of 300 MGD. Increasing dividing weir gates could accelerate the transfer of turbid water from the West Basin to the East Basin. What is the relative contribution of STP discharge?

Modeling: Operations alternatives with dividing weir flow at 1000 MGD, 1500 MGD, 2000 MGD; and STP flow at 0, 400 MGD, and 450 MGD were evaluated. OST was setup for Schoharie, Ashokan, and Kensico reservoirs. Simulations were conducted for April 7, 2016 – June 16, 2016 interval using HEFS forecasts. Initial temperatures and turbidities in these reservoirs were approximated from profiles and key point measurements.

Results: For all scenarios, STP turbidity at the outlet to Esopus Creek was projected to trend downward gradually from 15 NTU to 9 NTU over the simulation interval (Figure 2.1) although some traces predicted elevated levels exceeding 20 NTU. The probability of exceeding 10 NTU in the Catskill Aqueduct was predicted to be less than 10% (Figure 2.2). The results also indicated that the benefit of keeping the STP off is very slight and delayed (see the median turbidity plot; Figure 2.2), consistent with the findings of an earlier study that the contribution of STP to the total turbidity loading to Ashokan Reservoir was generally lower during high runoff events (UFI, 2007). At 90th percentile, early benefit of dividing weir flow at 2000 MGD was predicted but it was short-lived. After 4/13/2016, runs with dividing weir flow as 1500 and 2000 MGD predicted higher turbidity.

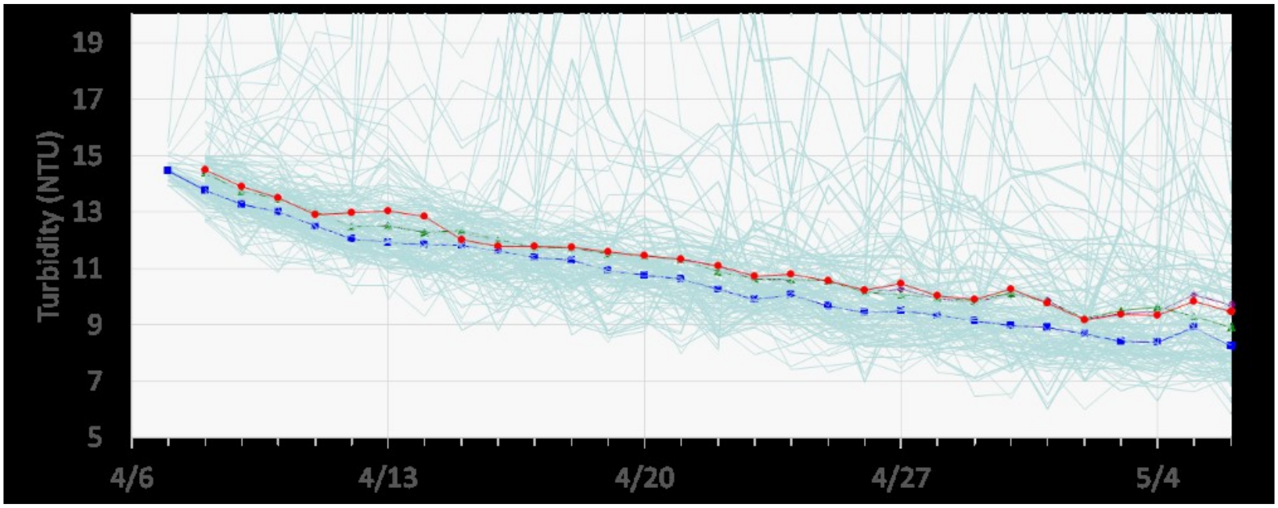


Figure 2.1. Predicted ensemble of withdrawal turbidity in Shandaken Tunnel diversion from Schoharie Reservoir, for three different tunnel flows. Colored lines show median tunnel turbidity for each flow: green – tunnel flow = 450 MGD, blue – tunnel flow=400 MGD, red – tunnel flow = 0.

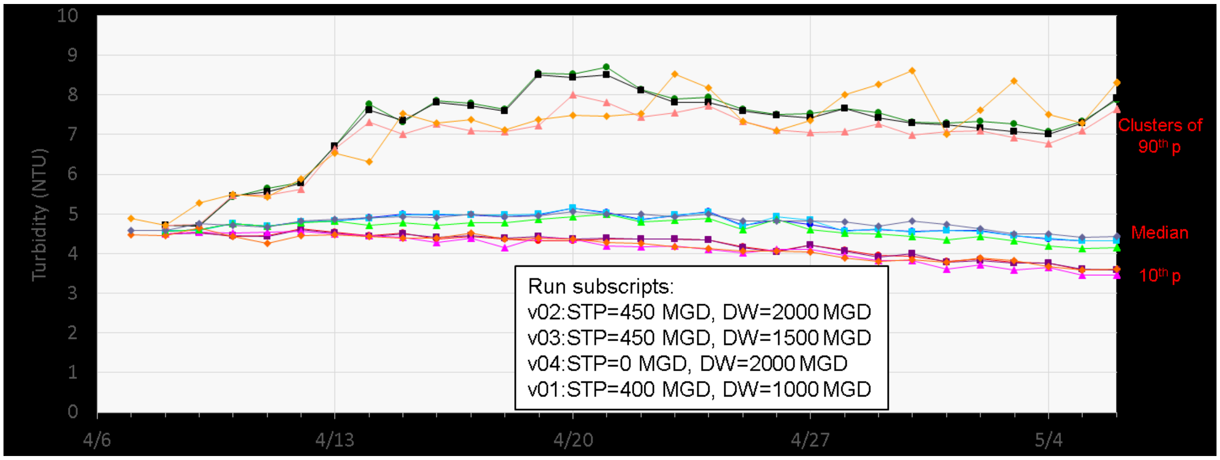


Figure 2.2. Three statistics (10th, 90th percentiles, and median) of predicted withdrawal turbidity for Catskill Aqueduct diversion from Ashokan Reservoir, for each of 4 different operations scenarios shown in the legend: STP is Shandaken Tunnel flow, DW is Ashokan Reservoir dividing weir flow.

June 6, 2016: A turbid plume was observed to be entering into Schoharie Reservoir (Figure 2.3). Modeling was conducted to determine the expected impact on withdrawal turbidity.

Figure 2.3. Schoharie Creek entering Schoharie Reservoir.

Modeling: Simulations with OST were conducted for 30 days and through June 1, 2017. Specifications for these simulations were: no alum application, input forecasts – HEFS, Schoharie diversion at 140 MGD, Ashokan diversion at 500 MGD, Pepacton diversion at 300 MGD during 6/6/2016–6/8/2016; Rondout diversion at 400 MGD for 6/6/2016, 200 MGD during 6/7-6/8/2016, NYC demand as 1015 MGD, dividing weir gates closed, Croton maximum diversion set to 80 MGD, inflow turbidity based on simple regression method, and short-circuiting of West Basin waters not allowed.

Results: The turbid plume in Schoharie Creek was shown to have no impact on SRR2CM turbidity. The two high traces correspond to the two high inflow forecasts but both predict turbidity less than 4 NTU (Figure 2.4).

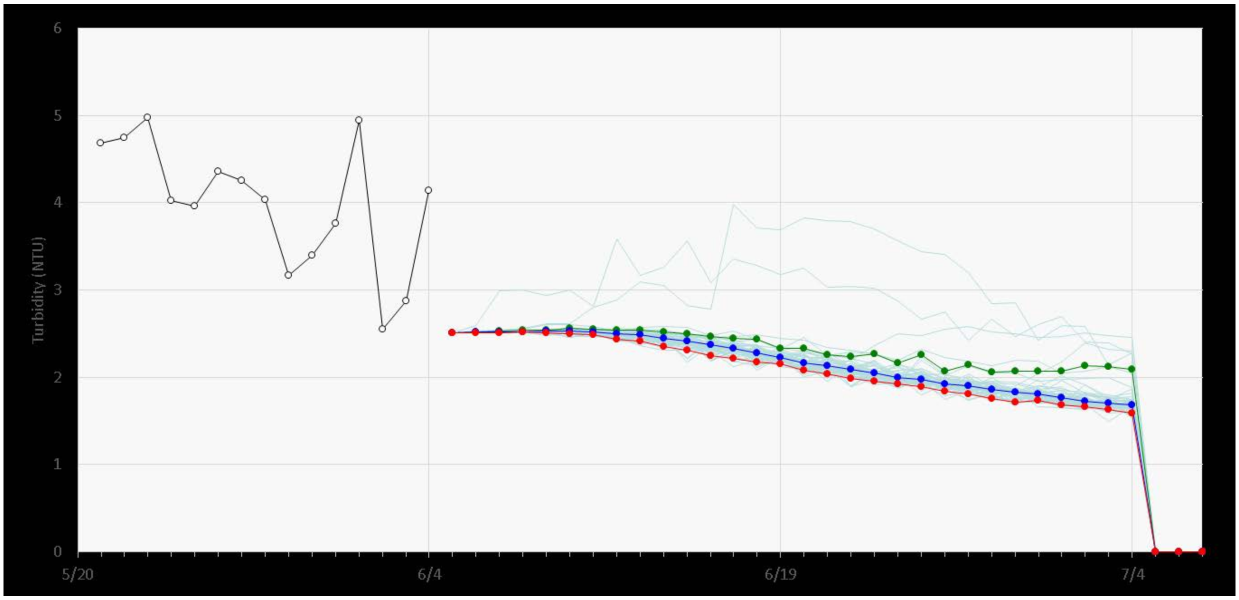


Figure 2.4. Predicted ensemble of Shandaken Tunnel turbidity, May 20 to July 7, 2016. Red, blue, green lines show 10th, 50th (median), and 90th percentiles of predicted withdrawal turbidity, respectively. Grey line: observations for two weeks prior to the start of model run.

Other short-term OST runs to guide system operations included runs conducted on February 22, 2016 to evaluate impact of a storm in the watershed, on March 23, 2016 to guide operation of the dividing weir gates (increase void in the West Basin ahead of a storm or let it spill after the storm) at Ashokan Reservoir, and on December 28, 2016 also, to guide operation of the dividing weir gates (i.e., evaluate impact of different flow rates on withdrawal turbidity at EARCM).

Several additional runs outside OST were conducted with RondoutPA (Rondout Position Analysis water quality model) during 2016 to guide operation of the Delaware Basin reservoirs. A brief summary of selected runs is provided below.

July 11, 2016: Approximately 7 inches of rainfall occurred during July 8-9, 2016 in the East Mountain Brook region of Rondout watershed which triggered localized bank failures resulting in turbid discharge into Rondout Creek (Figure 2.5). Immediately following the storm, turbidity at the mouth of Rondout Creek increased in excess of 1250 NTU, and within two days of the storm, 20–30 NTU in the upstream portion of Rondout Reservoir (Figure 2.6). Model runs were conducted to answer questions such as the timing and magnitude of the peak turbidity at the Rondout Reservoir withdrawal (RDRR) and how long would it take before turbidity level returns to normal (pre-storm) levels.

Results: Predictions of withdrawal turbidity obtained from an initial model run, starting from July 11, 2016 for 30 days, are presented here in a probabilistic format (Figure 2.7). For the first seven days, turbidity was expected to be above 1 NTU (probability of exceedance 100%) but below 2 NTU (probability of exceedance 0%). By early August, it was projected to exceed 3.5 NTU with 25% probability. These results provided information to managers to use in making decisions on whether to adjust operations to minimize the impact. New additional runs were conducted on July 20, 2016 and August 1, 2016, with updated hydrologic forecasts, changed operations of the upstream reservoirs as well as Rondout Reservoir, and updated in-reservoir initial conditions. Results of these runs provided further guidance on the optimum operation strategy for Rondout and its upstream reservoirs (Figure 2.8). Turbidity was projected to remain well within the tolerance limit for this reservoir for the entire duration of these runs and was later confirmed with the observations (Figure 2.8).

Figure 2.5. (a) Rondout Creek upstream of Rondout Reservoir and location of East Mountain Brook, a tributary of Rondout Creek where stream bank erosion occurred, (b) turbid plume from East Mountain Brook entering Rondout Creek on July 9, 2016.

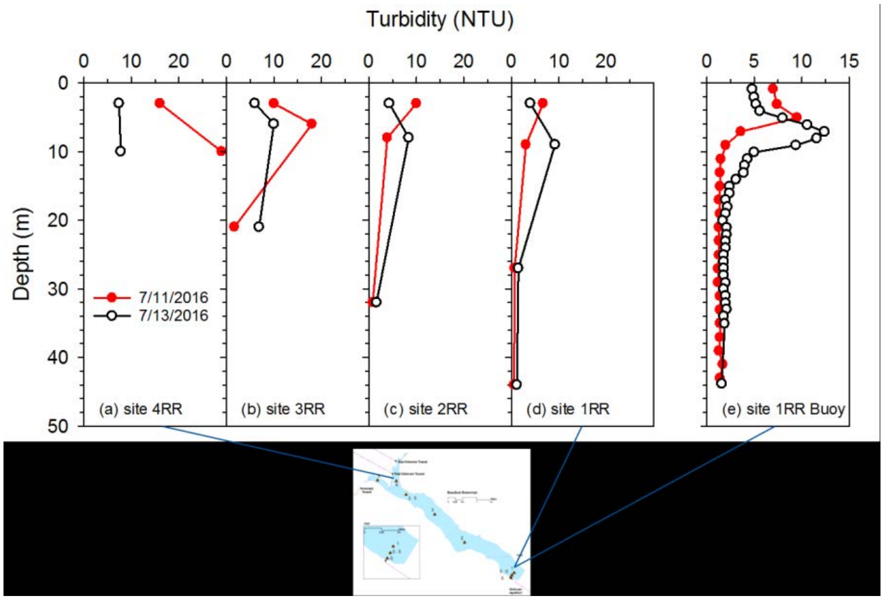


Figure 2.6. Observed depth-profiles of turbidity in Rondout Reservoir for July 11, and July 13, 2016.

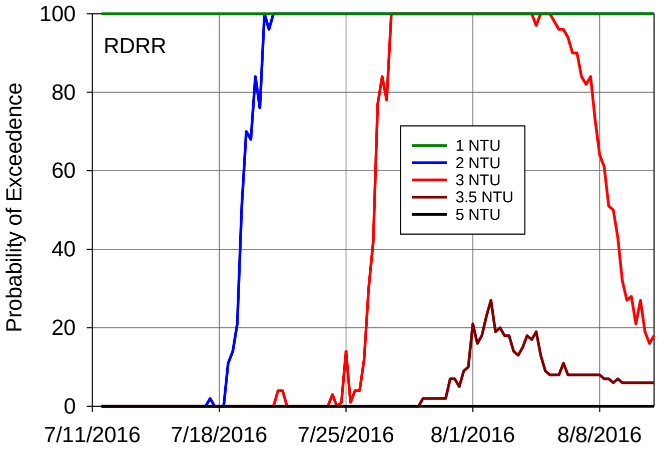


Figure 2.7. Probability of exceeding selected turbidity levels at RDRR obtained from model run starting from 7/11/2016.

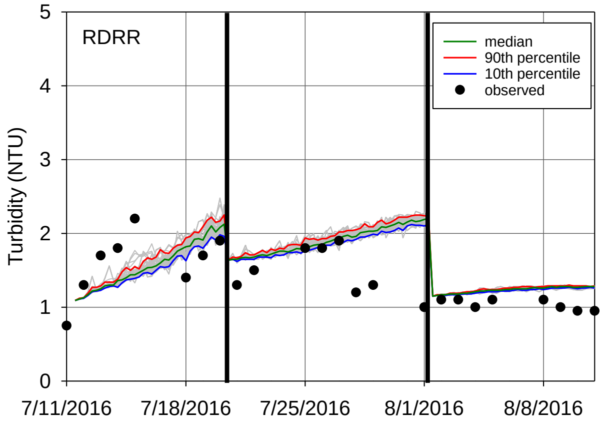


Figure 2.8. Ensemble predictions of withdrawal turbidity obtained from three different model runs starting from 7/11, 7/20, and 8/1, compared with observations.

A graphical summary of other RondoutPA model runs is depicted in Figure 2.9 through Figure 2.11.

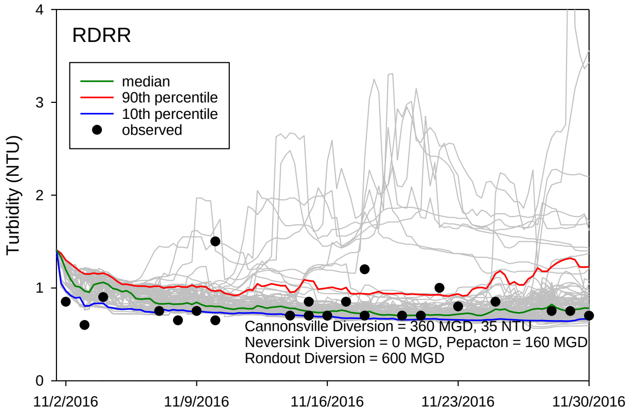


Figure 2.9. Ensemble predictions of turbidity at Rondout Reservoir withdrawal for an operations scenario of increased turbidity loading from Cannonsville Reservoir with discharge rate up from 280 to 340 MGD and turbidity levels 4 NTU. No significant impact on turbidity at RDRR was predicted, which was confirmed later with the observations.

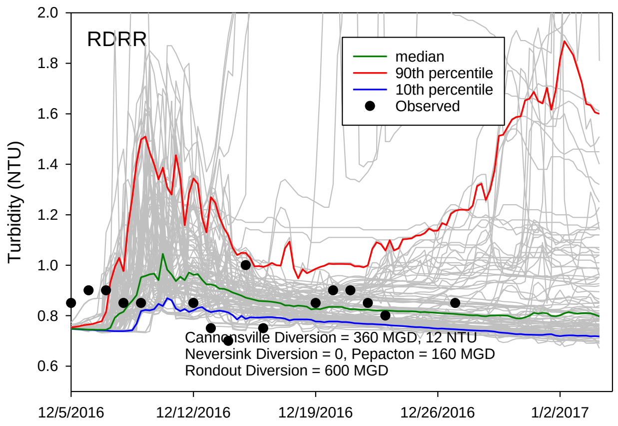


Figure 2.10. Ensemble predictions of turbidity at Rondout Reservoir withdrawal for an operations scenario of bringing Cannonsville Reservoir online (360 MGD, 12 NTU) after the hydroelectric power plant operator Brookfield Renewable Energy completed its work. No significant impact on turbidity at RDRR was predicted, which was confirmed later with the observations. Several other scenarios were also evaluated.

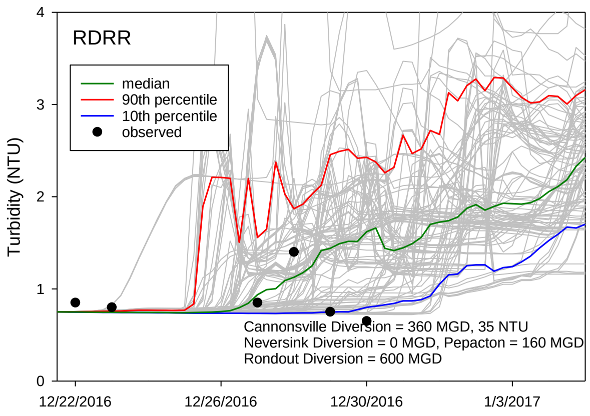


Figure 2.11. Ensemble predictions of turbidity at Rondout Reservoir withdrawal for a scenario of turbid discharge from Cannonsville Reservoir tunnel (turbidity plume from the rain on snow event on December 18, 2016; peak turbidities: West Branch Delaware River in-stream site CBS – 150 NTU, tunnel discharge site WDTO – 42 NTU). Scenarios with WDTO tunnel discharge as 360, 180, 0 MGD and turbidity as 45, 35, 25 NTU were evaluated. With these sustained high level of WDTO turbidity, the withdrawal turbidity at RDRR was projected to rise gradually. Accordingly, the reservoirs operations were adjusted to minimize the impact.

3. Modeling the Impacts of Climate Change

3.1. Climate Change Integrated Modeling Project

The Climate Change Integrated Modeling Project (CCIMP) encompasses the DEP Water Quality Modeling Section's effort to evaluate the effects of future climate change on the quantity and quality of water in the NYC water supply. The CCIMP is designed to address the following major issues: (1) overall quantity of water in the entire water supply; (2) turbidity in the Catskill System of reservoirs, including Kensico; (3) eutrophication in Delaware System reservoirs; and (4) disinfection byproducts in the West of Hudson reservoirs. The first phase of CCIMP was completed in 2013, and work began on the second phase with the signing of a new 4-year research support contract between CUNY and NYCDEP.

Work completed in 2016 was mainly in two areas. A stochastic weather generator for precipitation occurrence and amount, and for minimum and maximum daily air temperature was completed. The technical paper describing this work, which has been accepted by the peer-reviewed *Journal of Hydrometeorology*, is presented in its entirety in section 3.2. The first application of this stochastic weather generator in simulating streamflow in West of Hudson watersheds is presented in Section 3.3.

3.2. Stochastic Weather Generators for Precipitation and Air Temperature

3.2.1. Introduction

The New York City Water Supply System (NYCWSS), one of the largest surface water supply systems in the world, provides on average 1.1 billion gallons (BG) of high quality drinking water each day to almost half the population of the State of New York—over 9 million people in New York City (NYC) and some upstate communities (NYCDEP2016). This system contains 19 interconnected reservoirs, three controlled lakes and connecting aqueducts with a total storage capacity of approximately 580 billion gallon (BG). The New York City Department of Environmental Protection (NYCDEP) is responsible for operation, maintenance, and protection of the NYCWSS. The part of the system located west of the Hudson River (WOH), located up to 125 miles northwest of NYC in the Catskill Mountains, can be further divided into two subsystems, the Delaware and the Catskill. The WOH Catskill/Delaware system supplies approximately 90% of total water demand and is an unfiltered water supply (NYCDEP2016). The Delaware subsystem is comprised four watersheds: Cannonsville (95.7 BG), Neversink (34.9 BG), Pepacton (140.2 BG), and Rondout (49.6 BG); while the Catskill subsystem contains two watersheds: Schoharie (17.6 BG) and Ashokan (122.9 BG) (NYCDEP2016). These watersheds, which are part of the eastern plateau climate region of New York, cover an area of approximately 4103 square kilometers with an elevation range of 125–1275 m (Matonse *et al.* 2011). This region experiences a humid continental climate with cool summers, cold winters, occasionally abundant snowfall, and year round precipitation averaging between 1000 and 1200 mm per year (Anandhi *et al*., 2011a). Precipitation increases along west to east gradient, which is most prominent in winter, a pattern modulated by orographic effects (Thaler, 1996). Since the mid-twentieth century the contribution of snowfall to total annual precipitation is 20% to 30% (Frei *et al.* 2002; Pradhanang *et al.* 2011; Anandhi *et al.* 2011).

Extreme hydrological events are, in general, responsible for a disproportionate transport of nutrients and sediment into the streams and reservoirs. Past studies suggest increasing trends in total precipitation and in the frequency and magnitude of extreme precipitation events in the WOH basins. Burns *et al*. (2007) found that regional mean precipitation for the Catskill Mountain region increased significantly by 136 mm between 1952 and 2005. Matonse and Frei (2013) found that extreme warm season precipitation and streamflow events have been more frequent between 2002 and 2012 than any time during the 20th century. DeGaetano and Castellano (2013) found that the annual frequency of extreme Catskills precipitation (number of events that produce ≥50.8 mm precipitation per year) has an increasing trend over the last 60 years, with the time series dominated by year-to-year and decade-to-decade variability. They also analyzed climate model projections from the North American Regional Climate Change Assessment Program (NARCCAP) which suggests extreme precipitation will increase at a rate of 2–3% per decade through 2069. Thiebeault and Seth (2014) found that for the northeastern US, CMIP5 models predict a significant increase in the frequency and magnitude of extreme precipitation, particularly in mountainous regions such as the Catskills. They also found that models predict a concurrent increase in total precipitation, which could contribute to wetter conditions antecedent to storms and possibly to increased extreme streamflow events. The potential effects of these changes in precipitation include increased sediment erosion, increased nutrient loads, changes in the annual cycle of reservoir thermal structure, and other factors that may pose challenges for water management.

As a part of the NYCDEP's Climate Change Integrated Modeling Project (CCIMP), a series of studies (e.g. Anandhi *et al*., 2011a; Anandhi *et al*., 2011b; Pradhanang *et al*. 2011; Anandhi *et al*., 2013; Matonse *et al*., 2013; Pradhanang *et al*., 2013) have examined the potential impacts of climate change on the availability of high quality water in the WOH. These studies follow the "top-down" approach, using Change Factor Methodology (CFM) downscaled climate scenarios from Global Climate Models (GCMs), to incorporate climate change into vulnerability analyses using NYCDEP's integrated suite of models including watershed hydrology, water quality, water system operations, and reservoir hydrothermal models (Anandhi *et al*., 2011a). As an alternative to the top-down approach, "bottom-up" or vulnerability-based methods to climate change vulnerability analyses have recently been applied to water resources (Wilby and Desai, 2010; Brown *et al*., 2011; Prudhomme *et al.* 2010, 2015). Such approaches can explore the climate vulnerabilities of a system over a wider range of plausible climate change scenarios than the more traditional "top-down" approaches in which GCM projections completely define the parameter space of future scenarios (Wiley and Palmer, 2008; Steinschneider and Brown, 2013). The bottom-up approach can be implemented in a number of ways. For example, Steinschneider and Brown (2013) first determine system vulnerabilities and then assess different adaptation measures to find the most robust measures under future uncertainty. While the bottom-up approach includes the results of GCM simulations, it also enables more flexible definitions of uncertainty. Stochastic weather generators (SWGs) are an integral component of the bottom-up approach.

SWGs are statistical models that produce synthetic weather time series based on observed statistical properties at a particular location, and can include numerous weather variables, although this study is focused solely on precipitation. SWGs can be employed in bottom-up risk assessments to generate multiple scenarios of daily climate variations within which a water resource system model can be tested (Ray and Brown, 2015). A SWG coupled with a single or series of response models facilitates more complete identification of system vulnerabilities, and flexible, quantitative definitions of uncertainty, which can aid in the selection of robust adaption measures (Steinschneider and Brown, 2013). While the potential of SWGs for vulnerability assessments for the NYCWSS has not yet been explored in detail, a recent study (Rossi *et al*., 2015) used a multivariate, multisite weather generator for introducing incremental changes in mean precipitation and temperature to simulate a range of climate change scenarios to study turbidity levels in Ashokan Reservoir. However, the skill of weather generators used to simulate the observed precipitation characteristic is not discussed in detail. There are a number of general categories of SWGs available in the literature with different levels of complexity, the skills of which in simulating observed precipitation characteristics are location specific (Chen and Brissette, 2015). It is therefore essential to examine the model structure and performance in the target watershed before applying the SWG in climate change assessment studies for NYCWSS.

To address this need, the application of a suite of SWGs to each of the WOH watersheds is described. The skill levels of SWG in reproducing the statistical characteristics of observed daily precipitation will provide a useful benchmark for further improvement in climate change impact studies. As the study area has experienced a recent increase in extreme precipitation events, and is forecast to experience increasing trends by climate models, this evaluation of SWGs includes a focus on the large precipitation events which result in water quality management challenges. In order to ensure robust conclusions, we apply several different approaches to evaluating the simulation of extreme precipitation events, including daily and seasonal extremes, extreme events indices, and recurrence intervals of extreme events.

The remainder of this section is organized as follows. Section 3.2.2 provides a brief description of historical datasets used in the study. Section 3.2.3 provides an overview of several SWGs. The framework of calibrating of SWGs is described in Section 3.2.4. Sections 3.2.5 and 3.2.6 present results, and Section 3.2.7 includes discussion and conclusions.

3.2.2. Data

Observed daily precipitation data were obtained from Northeast Regional Climate Center (NRCC) at Cornell University. A total of 18 rain gauge stations of National Climate Data Center (NCDC) stations are non-uniformly distributed across the WOH watersheds (Figure 3.1). As each station has a different record history, precipitation data for the common period of record (1950 to 2009) were used. As the focus of the study is to analyze the average precipitation over each watershed, the weighted mean of nearby stations was calculated to obtain a single time series for each watershed. The Thiessen polygon method (Thiessen, 1911) was used to estimate the weights based on the relative areas of each measurement station in the Thiessen polygon network.

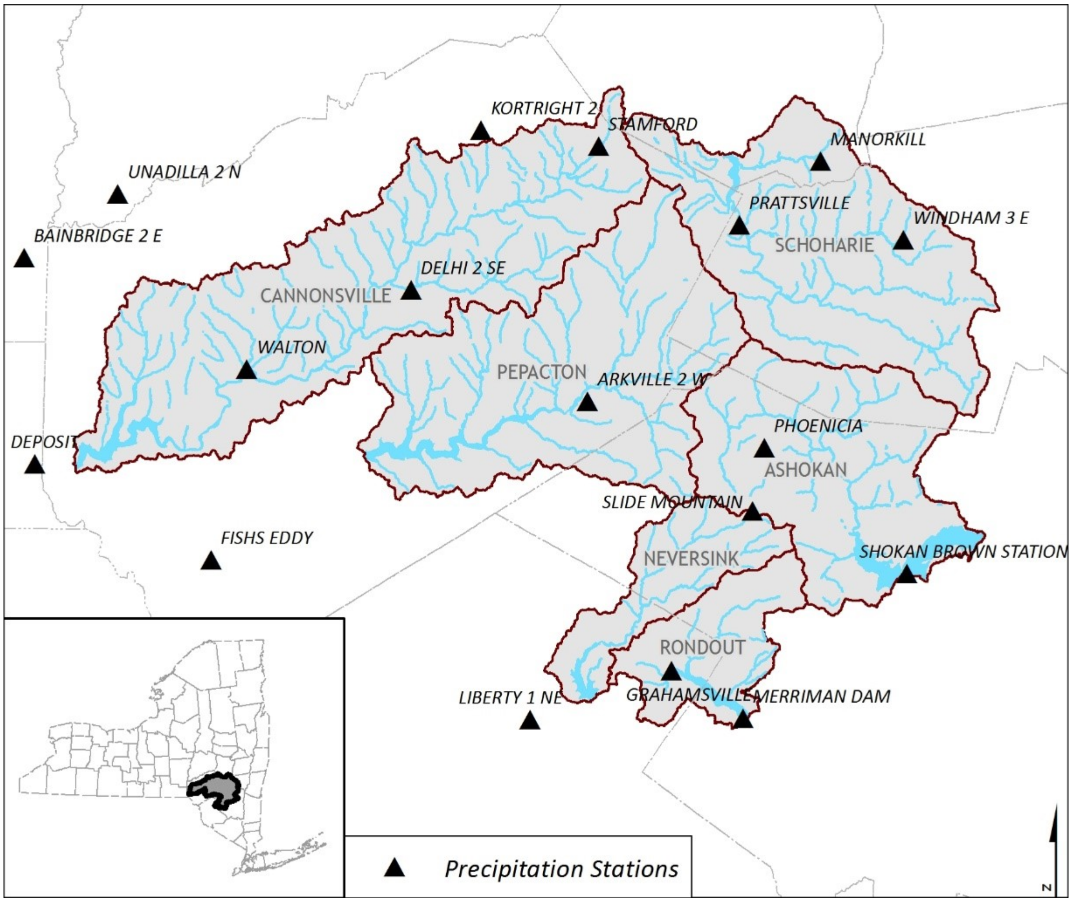


Figure 3.1. A map of the six reservoir watersheds of West of Hudson (WOH) and the precipitation gauge station scattered over the study region.

Table 3.1 describes the stations and their corresponding weights for each watershed. More details can be found in Anandhi *et al*. (2011).

Table 3.1. Description of the nearest stations and associated weights for estimating weighted mean precipitation for each of WOH watersheds.

Watershed	**Stations**
Schoharie	Windham 3E (0.4434), Prattsville (0.2992), Manorkill (0.1694), Stamford (0.0468), Phonecia (0.0389), Shokan Brown (0.0024).
Ashokan	Phonecia (0.4985), Shokan Brown (0.3401), Slide Mountain (0.1413), Windham 3E (0.02)
Cannonsville	Walton (0.3526), Delhi 2SE (0.2948), Kortright 2 (0.1143), Stamford (0.1532),Arkville 2W(0.0068),Bainbridge 2E (0.0085),Deposit (0.0568), Fish Eddy (0.0013), Unadilla 2N (0.0118)
Pepacton	Arkville 2W(0.5564), Delhi 2SE (0.1876), Prattsville (0.1282), Stamford (0.0523), Phonecia (0.0397), Slide Mountain (0.0181), Walton (0.0175), Fish Eddy (0.0002)
Neversink	Slide Mountain (0.5013), Grahamsville (0.3640), Liberty 1 NE (0.1347),
Rondout	Grahamsville (0.5796), Merriman Dam (0.2118), Slide Mountain (0.2084),Shokan Brown (0.0002),
Weights corresponding each station are in parentheses. All the weights add up to 1 for each watershed.	

3.2.3. Overview of Stochastic Weather Generators (SWGs)

The use of SWGs to produce synthetic time series with statistical properties resembling those of an observed time series has wide applications in the modeling of weather and climate-sensitive systems such as crop growth and development, hydrological processes, and ecological systems, where the observed climate records are inadequate in terms of length or completeness (Wilks and Wilby, 1999). A plethora of early studies dedicated to the development and advancement of SWGs (Gabriel and Neumann, 1962; Todorovic and Woolhiser, 1975; Katz, 1977; Richardson, 1981) have been summarized in several review articles (Wilks and Wilby, 1999; Srikanthan *et al.*, 2001; Alliot *et al.*, 2015). SWGs can be broadly classified into four groups: two-part models (the first part is dedicated to precipitation while the second part deals with other meteorological variables such as temperature or solar radiation), resampling models, transition probability models, and auto regressive moving average (ARMA) models (Srikanthan *et al.*, 2001), with the first two approaches being the most common. In this study, we discuss only precipitation generation in two-part models and resampling models.

Precipitation, which is unique among meteorological variables in its mixed nature as both a discrete (occurrence) and continuous (amount) process, has always been a key variable of interest in the construction of weather generators (Wilks and Wilby, 1999). In a two-part model, SWGs analyze precipitation as a chain-dependent model, first simulating precipitation occurrence (wet or dry day) and then precipitation amount. Occurrence is usually simulated using either a Markov Chain (MC) based model or a renewal process, sometime referred to as a spell-length model. Two-state (i.e., precipitation occurs or does not occur) MC models relate the state of precipitation on the current day to the states of precipitation on preceding days where the number of preceding days considered is the order of the MC (Boulanger *et al*., 2007). Although the first-order MC model (precipitation occurrence depends only on the previous day) has been found satisfactory in most of the cases (Katz, 1977; Richardson, 1981, Wilks, 1992), the higher order was found to better simulate long wet and dry spells (Wilks, 1999; Chen and Brissette, 2014). The alternating renewal process, rather than simulating occurrence for each day, fits a probability distribution to the sequence of alternating wet and dry spell lengths which are assumed to be independent (Buishad, 1978; Roldan and Woolhiser, 1982; Semenov and Barrow, 2002). Various probability distributions have been evaluated for the best fit of wet and dry spell lengths such as logarithmic series, truncated negative binomial distribution, truncated geometric distribution and semi-empirical distribution (Wilks and Wilby, 1999). It is typically assumed that either the parameters of the distribution remain constant within seasons, or that they change continually as estimated by a Fourier series. Studies have viewed a drawback of alternating renewal process method to handle the seasonality in the rainfall occurrence process (Sirkanthan and McMahon, 2001).

Given the occurrence of a wet day, the daily precipitation amount is then modeled, typically using a parametric distribution. The distributional pattern of daily precipitation is strongly skewed to the right (i.e., the tail of the distribution is on the right side) as very small daily precipitation events occur frequently, while heavy daily precipitation events are relatively rare (Wilks and Wilby, 1999; Chen and Brissette, 2014). Numerous studies have compared several probability distributions for simulating daily precipitation, including both single and compound distributions such as exponential (Todorovic and Woolhiser, 1975; Roldan and Woolhiser, 1982), Gamma (Ison *et al*., 1971; Richardson and Wright, 1984), Weibull (Stöckle *et al*., 1999), skewed normal (Nicks and Gander, 1994), Mixed Exponential distribution (Roldan and Woolhiser, 1982; Wilks, 1999b) and hybrid exponential and Pareto distributions (Furrer and Katz (2008); Li *et al*., 2012; Chen and Brissette, 2014).

Resampling, a data driven non-parametric method, provides an alternative to parametric techniques. The k-nearest-neighbor (k-NN) conditional bootstrap approach, the most popular resampling scheme for SWG, generates daily weather variables by resampling (with replacement) historical records associated with the wet-dry day series (Rajagopalan and Lall, 1999). The “k - nearest neighbors” for each date are chosen by considering all historical dates within a specified time window. Subsequently, one of the k-nearest neighbors is chosen, with a higher probability given to closer neighbors (King *et al*., 2015). After the pioneering work by Young (1994) and Sharma and Lall (1997), a number of studies extended and improved the k-NN approach (Rajagopalan and Lall; 1999; Buishand and Brandsma, 2001; Yates *et al*., 2003; Sharif and Burn, 2007; King *et al*., 2015). Some studies (Apipattanavis *et al*., 2007, Steinschneider and Brown, 2013) also combine the MC model and k-NN resampling to make a semi-parametric method. Kernel density estimation (KDE) method is also a popular non-parametric method for precipitation generation (Sharma and Lall 1999, Mehrotra and Sharma 2007). The underlying concept of this method based on the counting the relative frequency of the data lying in a local neighborhood which depends on the extent of the kernel functions about the point of estimation (Sharma and Lall 1999).

In addition to the above stated types of SWG, some other theoretical constructs have been applied to generate precipitation amount in some studies. For example, Boulanger *et al.* (2007) introduced a multi-layer perceptron-based neural network to generate synthetic time series of precipitation. Chen *et al.* (2015) proposed the use of a 2^{nd} degree polynomial curve fitting approach to fit a Weibull experimental frequency distribution of observed daily precipitation constrained on the probable maximum precipitation (PMP).

Following the aforementioned concepts, SWGs have been developed and widely used for precipitation generation over the last few decades. A listing of selected, widely-used SWGs is described in Table 3.2.

Table 3.2. Description of some popular SWGs with their precipitation occurrence and amount components and respected reference.

Name	**Precipitation occurrence and amount component**	**Reference**
WGEN	First-order MC for precipitation occurrence and Gamma distribution for precipitation amount	Richardson, 1981; Richardson and Wright, 1984
SIMMETEO	Same as WGEN but use monthly data as input instead of daily	Geng *et al*., 1988; Soltani and Hoogenboom, 2003; Elshamy *et al*., 2006
CLIGEN	First-order MC for precipitation occurrence and skew-normal distribution for precipitation amount	Nicks and Gander (1994)
GEM	First-order MC for precipitation occurrence and mixed exponential distribution for precipitation amount	Hanson and Johnson, 1998
CLIMGEN	Second-order MC for precipitation occurrence and Weibull distribution for precipitation amount	Stockle *et al*., 1999
WGENK	Modification of WGEN by introducing seasonality	Kuchar, 2004
WeaGETS	Third-order MC for precipitation occurrence and mixed exponential distribution for precipitation amount	Chen *et al*., 2012b
LARSWG	semi-empirical distribution to simulate precipitation occurrence and daily precipitation amounts	Semenov and Barrow, 2002
KnnCAD	Precipitation occurrence and amount generated by Resampling the historical data based on k-NN method.	Prodanovic and Simonovic,2008; King *et al*., 2015

3.2.4. Implementation of SWGs

Chain-dependent models using different combinations of MC based models of orders first (MC1), second (MC2), and third (MC3) to generate precipitation occurrence are implemented;

along with several parametric probability distributions, a resampling method, and a curve fitting technique to generate precipitation amount. The overall methodology to implement SWG in this study is outlined using a flow chart shown in Figure 3.2. In this study only two-state (wet or dry day) MC models are applied. A threshold of 0.1 mm is employed to discriminate between wet and dry days at the basin-averaged scale.

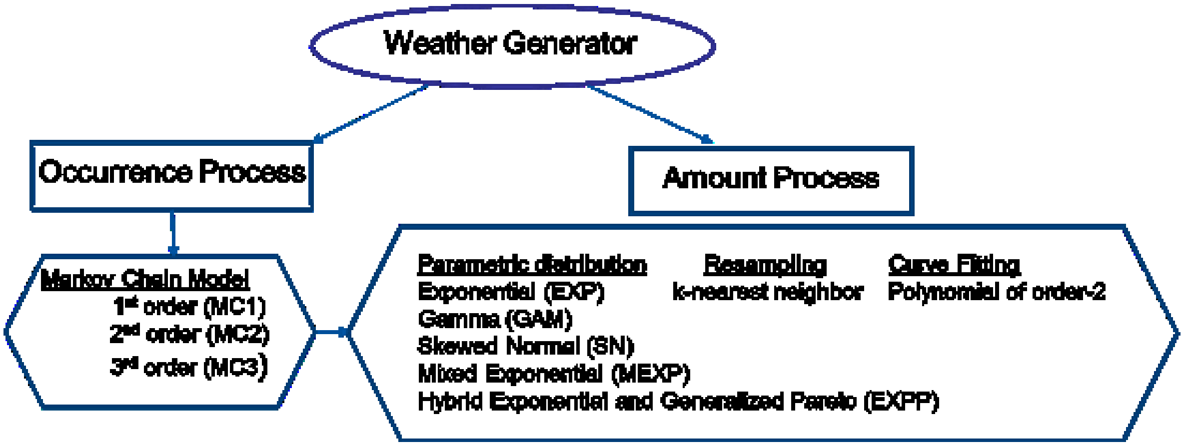


Figure 3.2. Flow chart showing the methodology of calibration of SWG.

3.2.4.1. Generation of the precipitation occurrence

In the first-order MC process, the probability of precipitation on a given day is based on the previous day's condition whether it was wet or dry, which can be defined in terms of transition probabilities, P_{01} and P_{11}:

$$P_{01} = \text{Prob (precipitation on day } t \mid \text{no precipitation on day } t\text{-1)} \tag{3.1}$$

$$P_{11} = \text{Prob (precipitation on day } t \mid \text{precipitation on day } t\text{-1)} \tag{3.2}$$

where the term 'Prob' means probability and the symbol '|' represents 'conditional on'. Since precipitation either occurs or does not occur on a given day, the two complementary transition probabilities are $P_{00} = 1 - P_{01}$ (dry day following a dry day) and $P_{10} = 1 - P_{11}$ (dry day following a wet day).

The first-order MC model can be generalized to higher-orders. Letting $r_t = 0$ if day t is dry, and $r_t = 1$ if day t is wet, Equations (3.1) and (3.2) can be extended to second (Equation (3.3)) and third (Equation (3.4)) order MC processes:

$$P_{ijk} = \text{Prob}\ (r_t = k \mid r_{t-1} = j \mid r_{t-2} = i) \tag{3.3}$$

$$P_{hijk} = \text{Prob}\ (r_t = k \mid r_{t-1} = j \mid r_{t-2} = i \mid r_{t-3} = h) \tag{3.4}$$

where i,j,k,h are values of 0 or 1 representative dry and wet day, respectively. Transition probabilities are estimated from the observed time series by using the maximum likelihood method for each biweekly period. The sequences of wet and dry days in the SWG are produced by applying a random number generator based on the uniform distribution from the interval [0, 1] for each day. The occurrence of precipitation is estimated for each day, in sequence, based on the previous day(s) precipitation state, and a comparison of the random number to the corresponding transitional probability (Wilks and Wilby, 1999).

3.2.4.2. Generation of the precipitation amount

After computing rainfall occurrence, rainfall amounts on wet days are modeled using seven distribution models, including five parametric distributions, one resampling method (k-NN), and one curve fitting method. Parametric distributions include three single distributions - Exponential (EXP), Gamma (GAM), and Skewed-Normal (SN) - and two compound distributions - Mixed Exponential distribution (MEXP) and a Hybrid Exponential and Generalized Pareto (EXPP) distribution. The parameters of the distributions are estimated using the maximum likelihood method for each biweekly time period except for the SN distribution where the method of moments is used. Following Rajagopalan and Lall (1999), we choose k = $\sqrt{N}$ in the k-NN resampling method where N equals the sample size and 7 day time window use to find the "nearest neighbor" for each date considering all historical dates. The 2nd order polynomial-based curve fitting method (PN) to fit the Weibull experimental frequency of daily precipitation used here is similar to Chen *et al*. (2015), except in our case the maximum value is not constrained by the PMP. A summary of each model and associated references are found in Table 3.3. The length of all simulated daily precipitation series are ten times of the observed record length. The generation of such long synthetic series is required to avoid biased simulation of the true climate (Chen and Brissette, 2014).

Table 3.3. Selected 7 models for generating daily precipitation amounts.

Model	Name	Abbreviation	Reference
Parametric	Exponential	EXP	Todorovic and Woolhiser (1975)
	Gamma	GAM	Ison *et al*. (1971), Richardson and Wright (1984)
	Skewed-normal	SN	Nicks and Gander (1994)
	Mixed exponential	MEXP	Woolhiser and Roldán (1982), Wilks (1999b)
	Hybrid exponential and generalized Pareto	EXPP	Li *et al*. (2012)
Resampling	k-nearest-neighbor conditional bootstrap	k-NN	Rajagopalan and Lall (1999)
Curve-fitting	2nd order polynomial without constraint by the probable maximum precipitation (PMP)	PN	Chen *et al*. (2015)

3.2.5. Statistical Evaluation of SWGs

In this section, we evaluate the skill of SWGs in terms of simulating the statistical characteristics of the full distribution of daily precipitation as well as monthly and annual scale for all six watersheds.

3.2.5.1. Daily Precipitation Characteristics

As SWGs simulate the daily precipitation occurrence first and then amount separately, we evaluate each component independently by estimating following statistical metrics for observed and simulated time series for each watershed.

Precipitation Occurrence

Markov chain-based models with different levels of complexity, *viz*. first (MC1), second (MC2) and third order (MC3), are compared with observations with respect to reproducing the frequency of wet days per month and the distributions of wet and dry spell lengths. We compare observed and simulated means and standard deviations of the number of wet days per month for each watershed (Figure 3.3). Wet (dry) spell lengths are defined as the number of consecutive days with precipitation more (less) than the threshold value. Hence, mean, standard deviation and 99th percentile (Q99) of wet and dry spells are evaluated. (Figure 3.4 and Figure 3.5). Based on these metrics, it is found that no significant improvement is achieved by increasing the order of the MC model; the performance of the three MC models is comparable. This result agrees with earlier work. For example, Wilks (1999) found the MC1 broadly appropriate for the central and eastern US. Schoof and Pryor (2008) also found the MC1 outperforms the MC2 and MC3 for the major part of the US including the north-eastern region. Therefore, under the principle of parsimony, only MC1 is used in subsequent sections of this analysis.

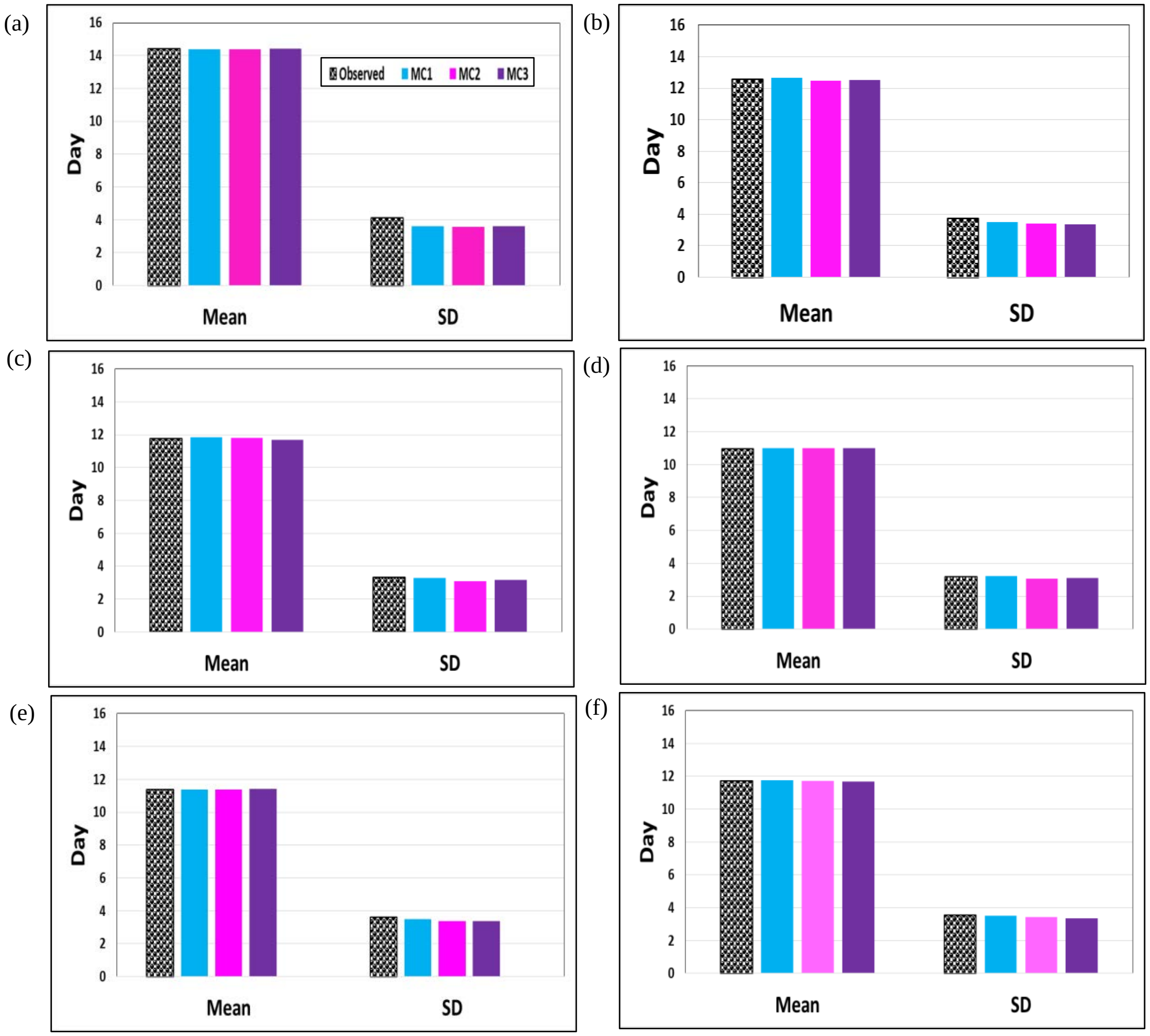


Figure 3.3. Mean and standard deviation (SD) of the observed and generated wet days per month from MC1, MC2 and MC3 for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

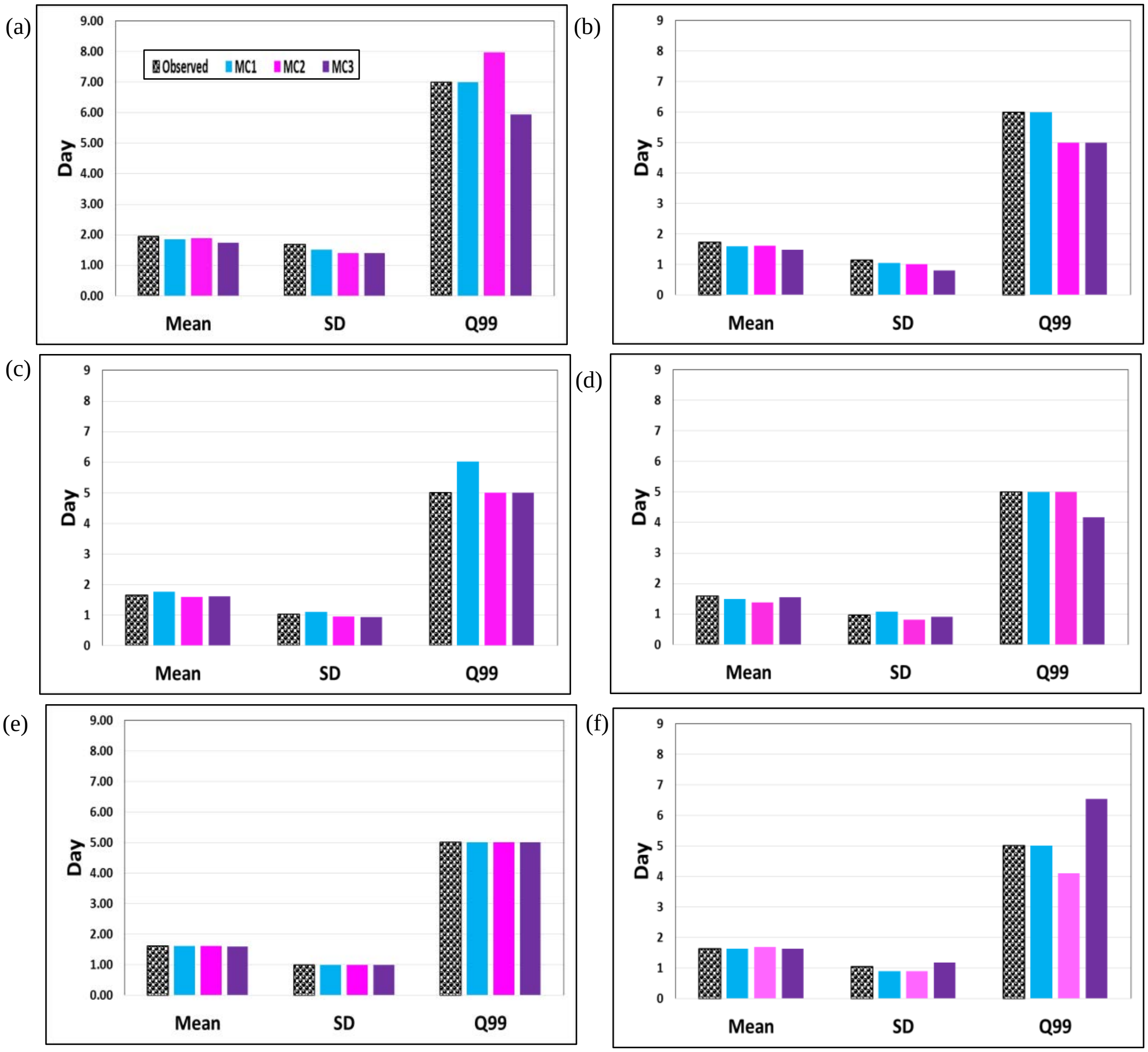


Figure 3.4. Mean and standard deviation of the observed and generated wet spell lengths from MC1, MC2 and MC3 for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

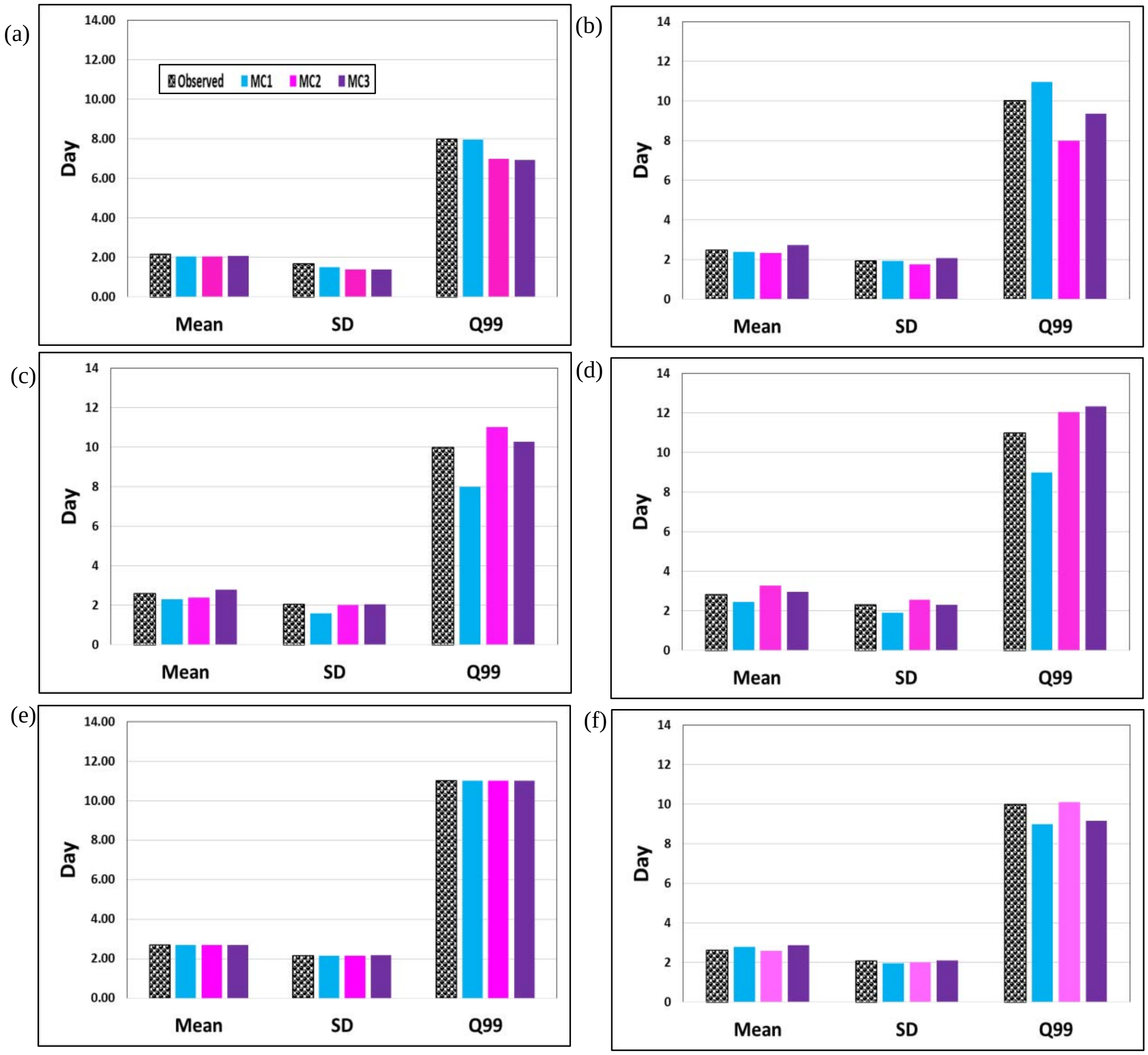


Figure 3.5. Mean and standard deviation of the observed and generated dry spell lengths from MC1, MC2 and MC3 for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

Precipitation Amount

Precipitation amount is evaluated based on mean, median, standard deviation, interquantile range and skewness coefficient of daily precipitation amount (including wet days only) (Figure 3.6). While all seven models adequately simulate mean daily precipitation for all watersheds, their skills vary in reproducing other statistics. For example, the EXP and GAM distributions consistently overestimate the median and underestimate the standard deviation, while EXPP and PN considerably overestimate both median and standard deviation for most watersheds.

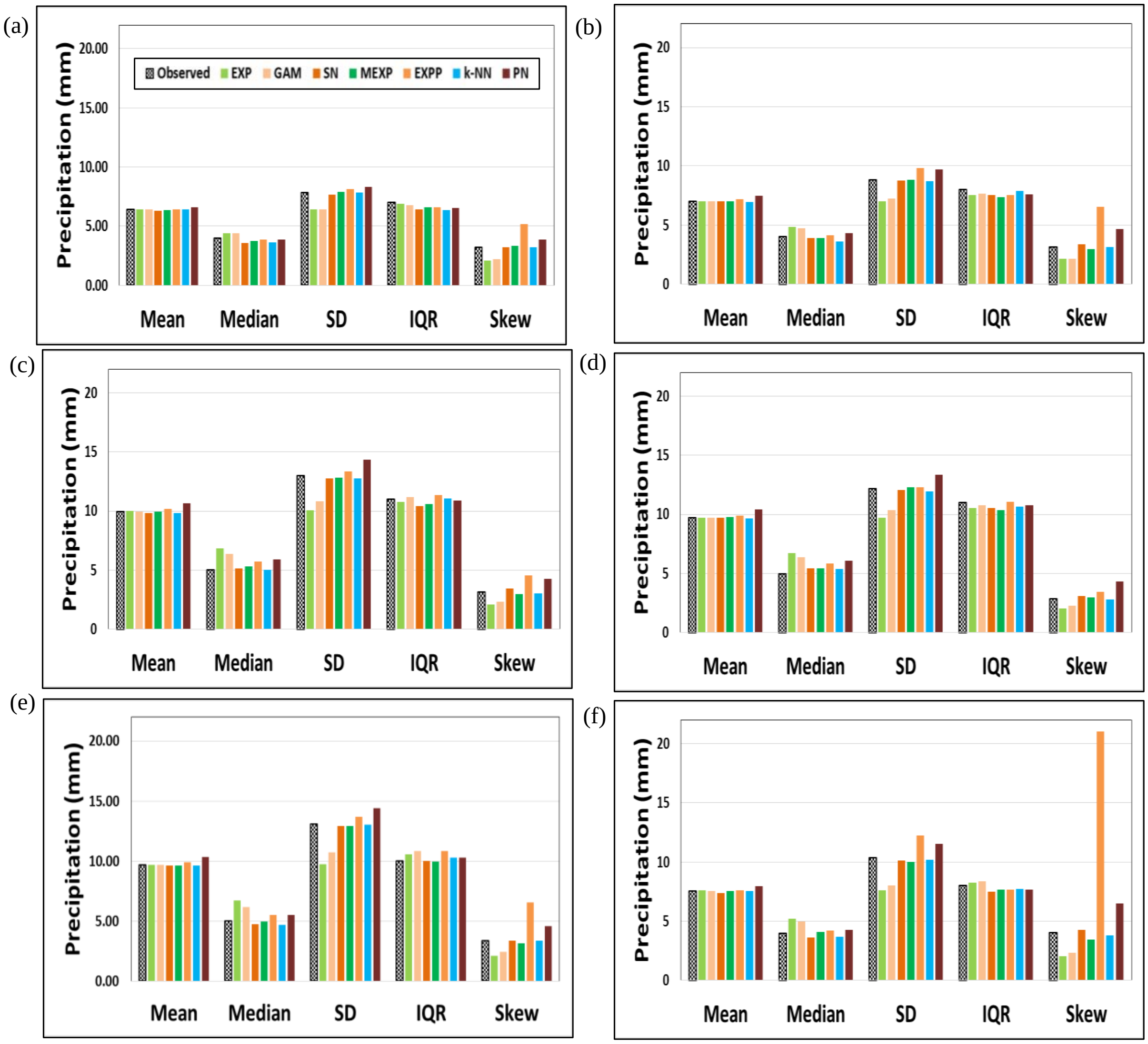


Figure 3.6. Mean, median standard deviation (SD), interquantile range (IQR) and skewness coefficients (Skew) of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

As expected the skewness coefficient of observed daily precipitation is greater than 3 for most watersheds, implying that the distribution of daily precipitation is extremely skewed to the right. Skewness is underestimated by EXP and GAM and overestimated by EXPP and PN (especially for the Schoharie watershed); while SN, MEXP and k-NN adequately simulate the skewness. To understand this issue in more detail, the cumulative density functions (cdf) of observed and simulated time series are estimated. The cumulative probabilities of the observed and simulated daily precipitation are estimated. The cumulative probabilities are plotted in probability scale such that the values at the tails of the distribution can be more easily visually analyzed. The models tend to perform well at smaller precipitation values. In contrast, for the large events the most models perform more poorly: EXP and GAM both underestimate, EXPP and PN generate unreasonably large values and overestimate skewness. This type of overestimation for these models is also reported by other studies for different locations (e.g. Wilks and Wilby, 1999; Chen and Brissette, 2014). However, the SN, MEXP and k-NN capture the most prominent statistical characteristics of the observed cdf for both lower and higher categories (Figure 3.7).

To quantify relative model accuracies, mean absolute percentage error (MAPE, mean(abs(sim-obs)/obs)*100), which is the average of the absolute values of the percentage errors of each watersheds, is calculated for each metric (Table 3.4). On the basis of the magnitude of MAPE, Lewis (1982) developed different classes to judge the accuracy of the model, *viz*., 'Highly Accurate' (less than 10%), 'Good Forecast' (11% to 20%), 'Reasonable Forecast' (21% to 50%), 'Inaccurate Forecast' (51% or more). The model that produced lower MAPE provides a better representation of the rainfall characteristics. It is seen that except for mean and interquartile range, the MAPE values for EXP and GAM shows a 'reasonable forecast'. With values less than 10% for all the statistics, SN, MEXP and k-NN fall in 'highly accurate' category. EXPP and PN also perform 'highly accurate' except for skewness coefficient which MAPE values indicate 'inaccurate forecast'. With regard to skewness, only SN, MEXP, and k-NN fall in the 'highly accurate' category.

Table 3.4. Mean absolute percentage error (MAPE) of generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for all watersheds.

	EXP	GAM	SN	MEXP	EXPP	k-NN	PN
Mean	0.28	0.13	1.02	0.33	1.36	0.95	5.96
Median	27.74	21.91	6.54	6.54	8.83	7.18	11.35
Standard deviation	22.12	17.88	1.52	1.17	6.89	1.15	9.61
Interquantile range	3.8	3.95	4.93	4.67	4.55	3.34	3.59
Skewness coefficients	35.82	30.1	5.07	6.38	124.58	2.07	41.69

Statistics for $n = 60$ years

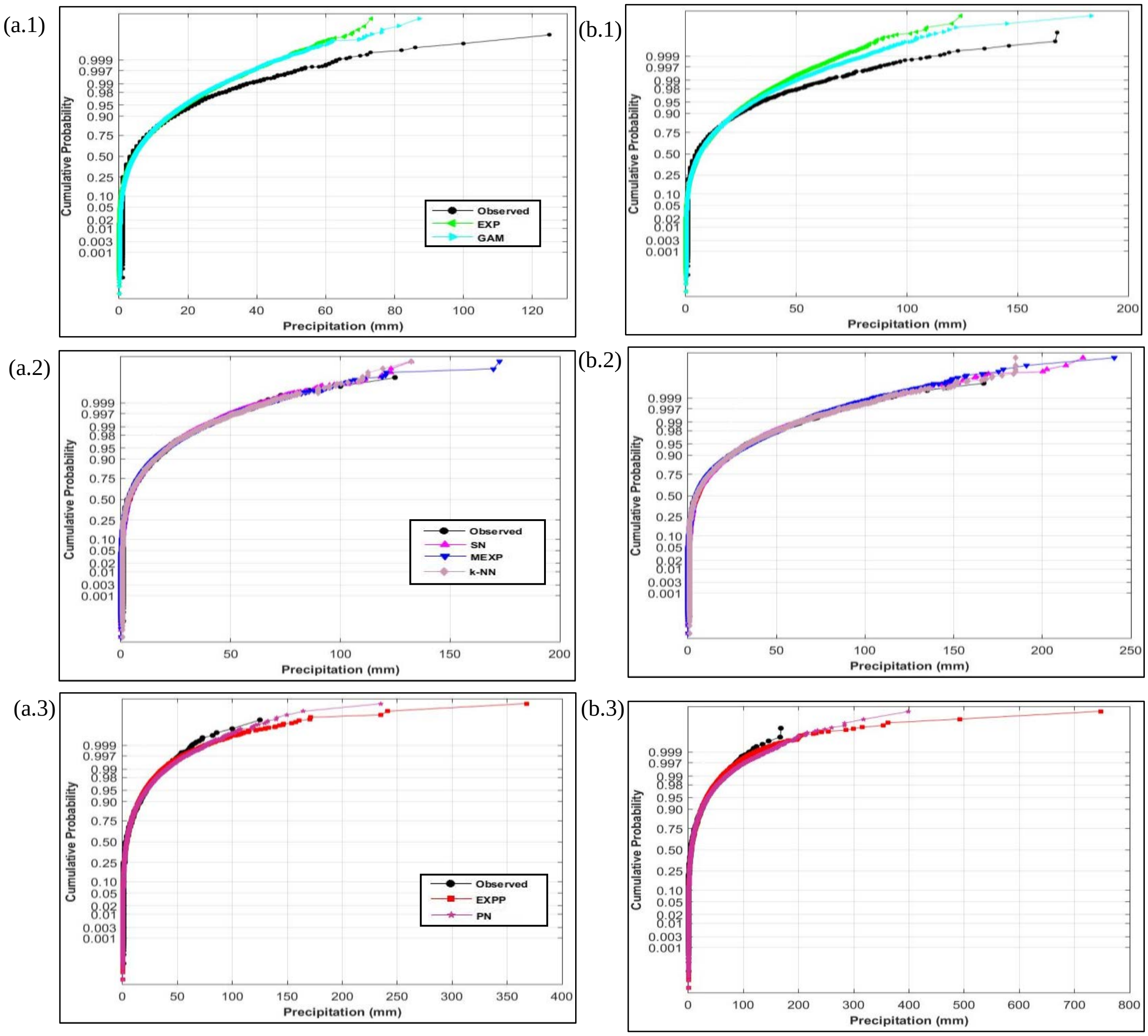


Figure 3.7. Cumulative probability distribution functions of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for (a) Cannonsville and (b) Ashokan watersheds.

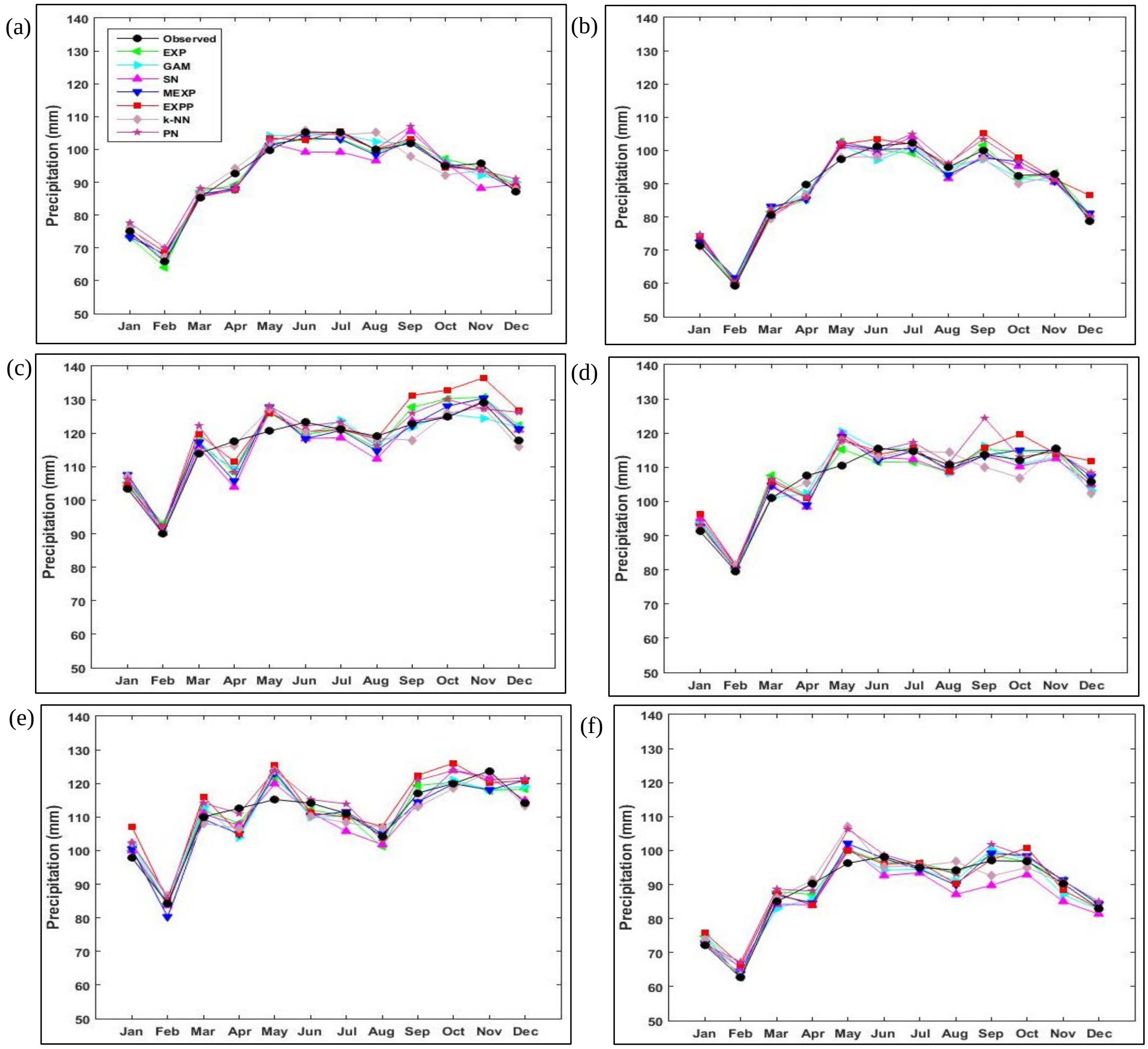


Figure 3.8. Monthly mean precipitation of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

3.2.6. Monthly and Annual Precipitation Characteristics

Most of the models simulate the intra-annual pattern of monthly mean precipitation fairly well (Figure 3.8).While almost all models underestimate the April mean precipitation and overestimate May values, to some extent, MAPE values suggest that all models provide ‘highly accurate’ simulations of monthly mean precipitation (MAPE less than 10%) (Table 3.5). All models accurately reproduce annual averaged precipitation with MAPE less than 2% (Table 3.5).

Table 3.5. Mean absolute percentage error (MAPE) of generated precipitation for each month and annual from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for all watersheds.

	EXP		**GAM**		**SN**		**MEXP**		**EXPP**		**k-NN**		**PN**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Jan	2.52	28.36	2.81	26.08	2.11	19.03	2.08	17.76	4.34	16.85	2.25	5.26	3.01	14.14
Feb	1.17	17.15	1.74	13.55	2.17	7.19	2.80	5.26	2.95	10.17	1.51	1.74	3.26	6.52
Mar	3.11	7.38	1.74	7.30	1.21	7.24	2.31	8.02	3.23	12.55	1.17	11.92	4.46	18.03
Apr	4.89	16.26	5.41	15.26	6.64	5.38	6.80	6.14	5.64	5.62	2.36	6.06	4.39	4.77
May	4.27	18.88	5.60	16.27	4.73	9.79	5.36	11.13	5.43	9.58	5.64	4.62	6.45	7.47
Jun	2.23	25.51	2.82	24.16	3.65	16.20	2.27	13.78	2.13	14.26	2.47	4.46	0.68	9.48
Jul	1.60	11.62	1.04	11.38	2.99	6.37	0.85	4.39	0.67	6.57	0.90	11.33	1.84	5.46
Aug	2.09	21.71	1.83	17.22	4.08	8.89	2.33	7.81	1.71	7.84	2.43	11.03	1.13	6.54
Sep	1.94	18.65	1.53	17.81	2.72	7.28	1.32	9.23	3.35	7.78	3.60	3.34	4.77	7.39
Oct	1.69	28.63	0.96	25.73	2.14	17.87	1.95	15.56	4.77	12.41	2.36	5.35	2.48	10.00
Nov	1.60	7.90	3.39	7.68	3.43	14.47	1.94	15.35	2.58	22.13	0.76	21.28	1.58	21.77
Dec	2.43	20.23	2.41	20.43	1.85	8.61	2.56	9.94	5.18	6.93	1.29	3.06	4.22	6.26
Annual	0.40	29.06	0.32	27.25	1.08	18.95	0.39	17.86	1.67	12.15	0.48	9.05	1.94	13.01

Model simulations of variability as represented by the standard deviation of monthly and annual values are poor (Table 3.5). The underestimation of the inter-annual variability of monthly and annual precipitation by all the models is expected, and can be attributed to the lack of any terms in the models to represent low frequency variation. The inclusion of such low frequency variability has been addressed in the literature (Chen *et al*. 2010), and will be addressed for our study region at a future date. It is our intention here to focus this report on other aspects of the statistical simulation, and to incorporate low frequency variability at a future time.

3.2.7. Performance of SWGs for Extreme Precipitation Events

As the study area has experienced an increased frequency of extreme precipitation events in recent decades, it is important to evaluate the abilities of SWGs to simulate extreme event probabilities. Here we apply a variety of metrics in order to evaluate model simulations in detail.

Daily Extreme Events

The first set of metrics we employ include quantile based metrics of daily extreme event magnitude: the 95th (Q95) and 99th (Q99) percentiles (Figure 3.9). While all models provide reasonable estimates of Q95, Q99 is underestimated by EXP and GAM and tends to be overestimated by PN. In terms of MAPE, all models except EXP and GAM perform ‘highly accurate’ simulation for both Q95 and Q99 (Table 3.5).

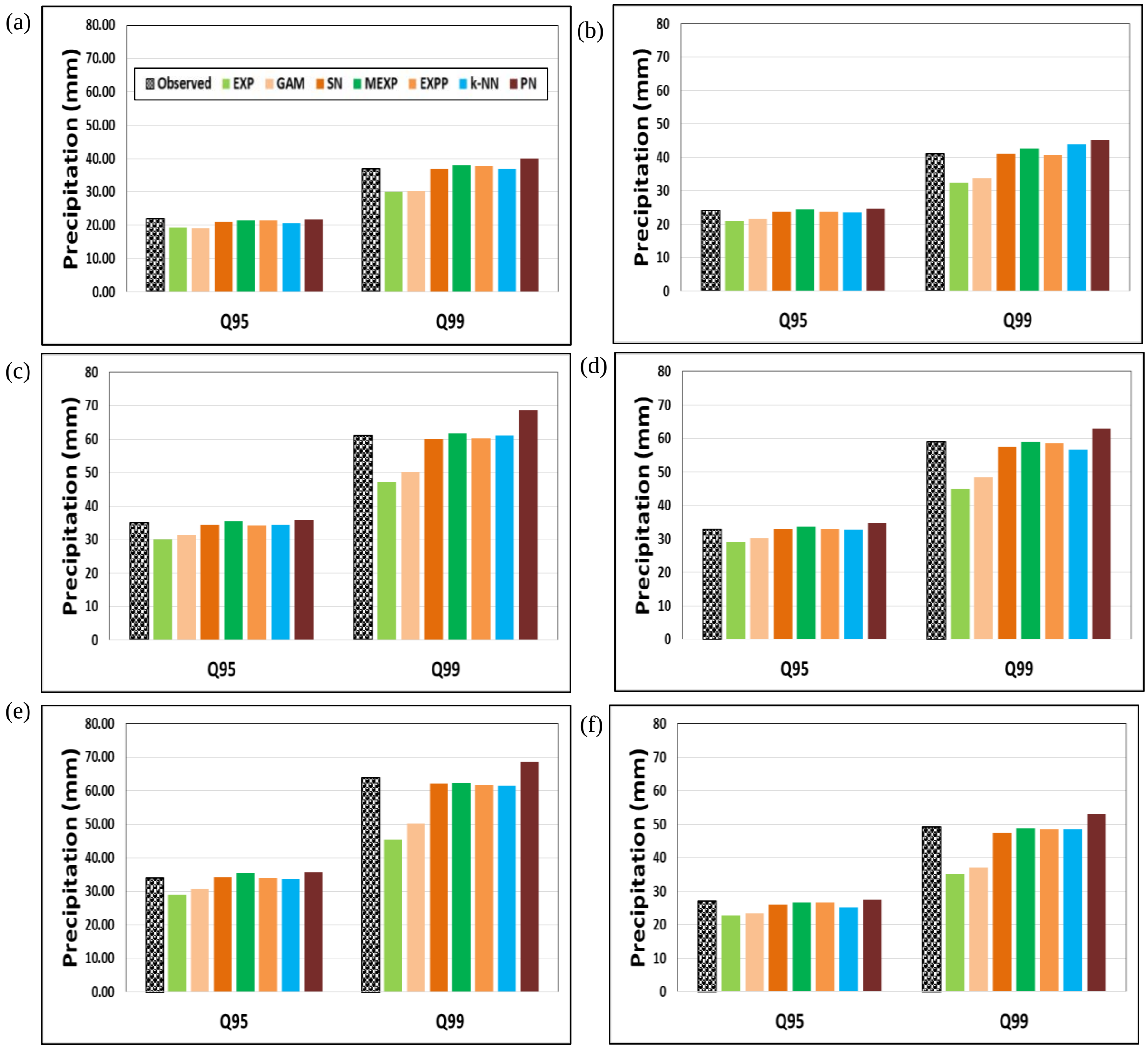


Figure 3.9. 95th (Q95) and 99th (Q99) of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

Seasonal Extreme Events

Matonse and Frei (2013) have studied the nature of the seasonal cycle of extreme precipitation events and found it is unimodal, with peak values occurring between August and October, including but not limited to the direct influence of tropical storms and hurricanes in this region. Following Matonse and Frei (2013), boxplots of all daily precipitation values equal to or greater than the 95th percentile for each watershed are evaluated. Figure 3.10 shows the results for the Cannonsville and Ashokan watersheds. For all the watersheds, EXP and GAM underestimate the median and interquartile range compare to the observed boxplot. The remaining models perform well in capturing the observed pattern, though the SWG have more values bigger than the observed, especially for EXPP and PN. This may be due to their overestimation of high values towards the end of the tail of the distribution.

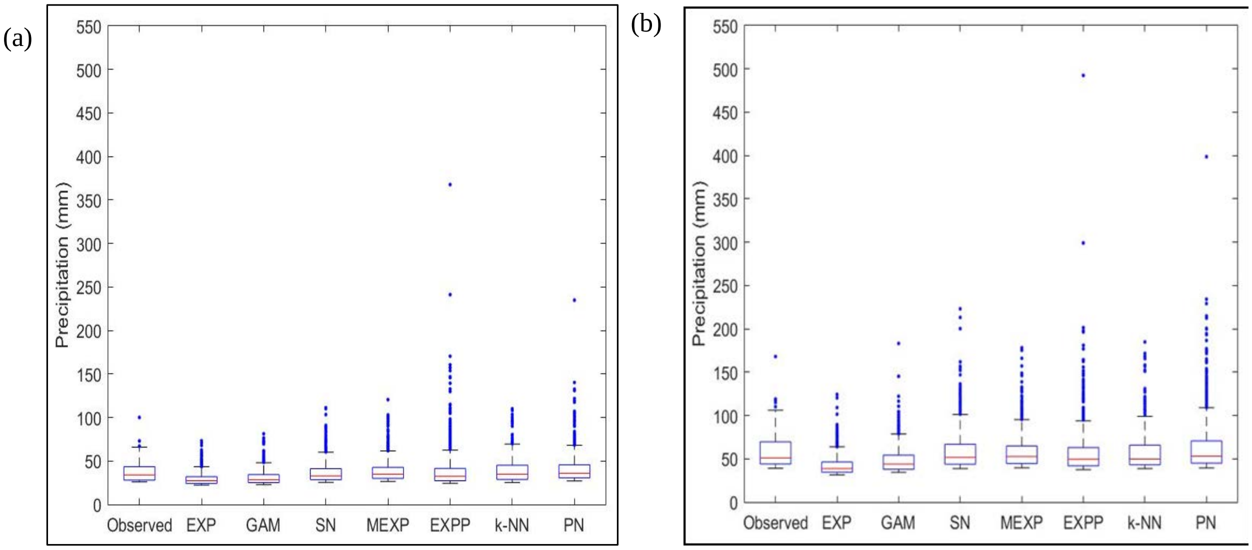


Figure 3.10. Boxplot of daily precipitation values equal to or greater than 95th percentile of observed and generated precipitation for Aug-Sep-Oct from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for (a) Cannonsville and (b) Ashokan watersheds.

Extreme Events Indices

A set of 27 climate extreme indices based on daily temperature and precipitation has been proposed by The Expert Team on Climate Change Detection and Indices (ETCCDI) (Klein Tank *et al*., 2009). Due to their robustness and straightforward calculation and interpretation, these indices are popular in recent years for numerous applications in climate research (Sillmann *et al.* ,2013). Following Zhang *et al*. (2011), SWG performance based on the four extreme event indices associated with large precipitation events is evaluated. These indices are:

- RX1day: Maximum 1-day precipitation per year.
- RX5day: Maximum consecutive 5-day precipitation per year.
- R95p: Annual total precipitation due to events exceeding the 95th percentile of the entire data period (1950-2009).
- R99p: Annual total precipitation due to events exceeding the 99th percentile of the entire data period (1950-2009).

Each index is calculated for each year individually, and yearly results are presented in box and whisker plots. Results for the Cannonsville and Ashokan watersheds are shown in

Figure 3.11 and Figure 3.12; results are similar for the other watersheds. RX1day and RX5day are underestimated by EXP and GAM while other models perform better. For Cannonsville, all models overestimate R95p and R99p for Ashokan results are more mixed.

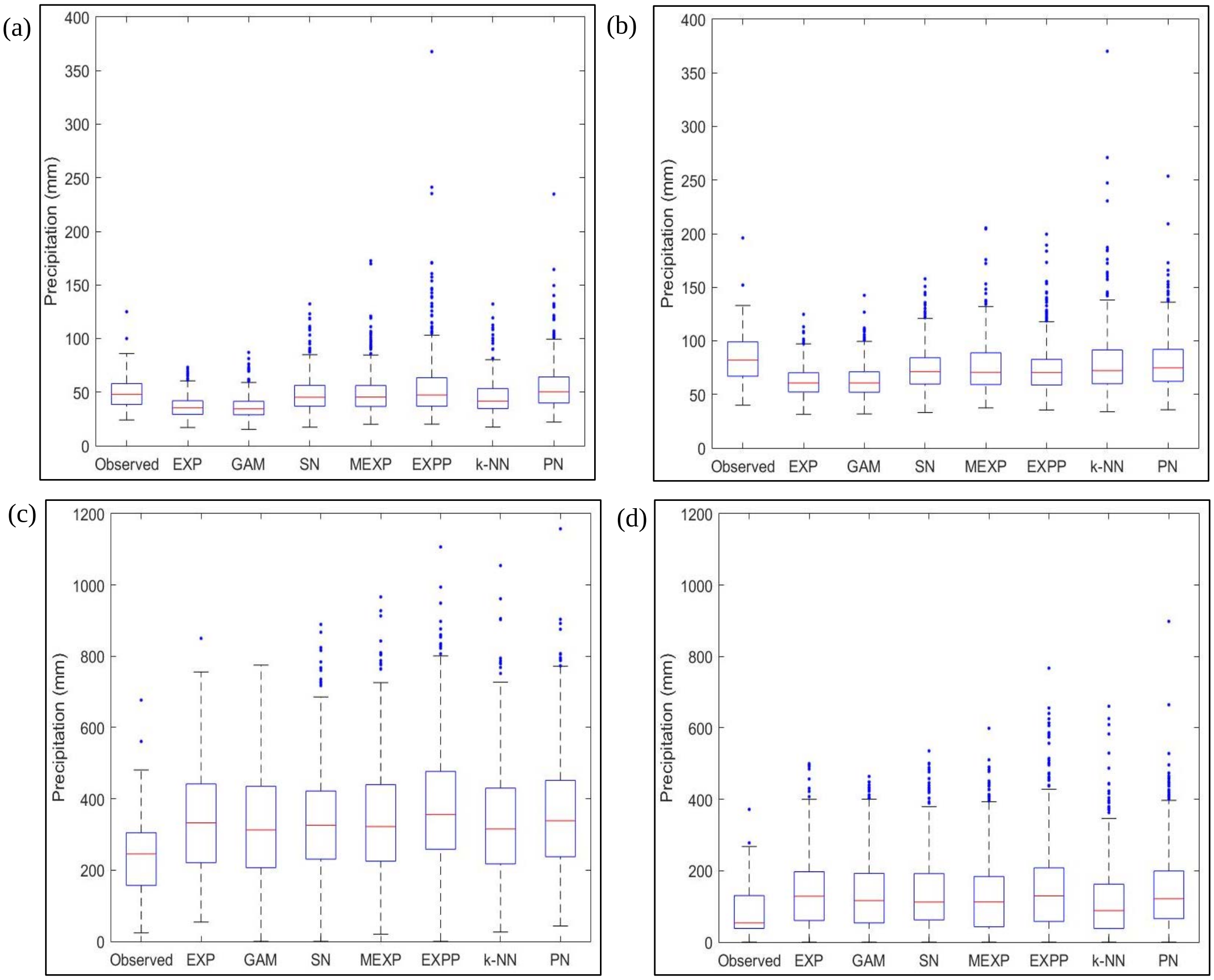


Figure 3.11. Boxplot of (a) RX1day, (b) RX5day, (c) R95p and (d) R99p of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for Cannonsville watershed.

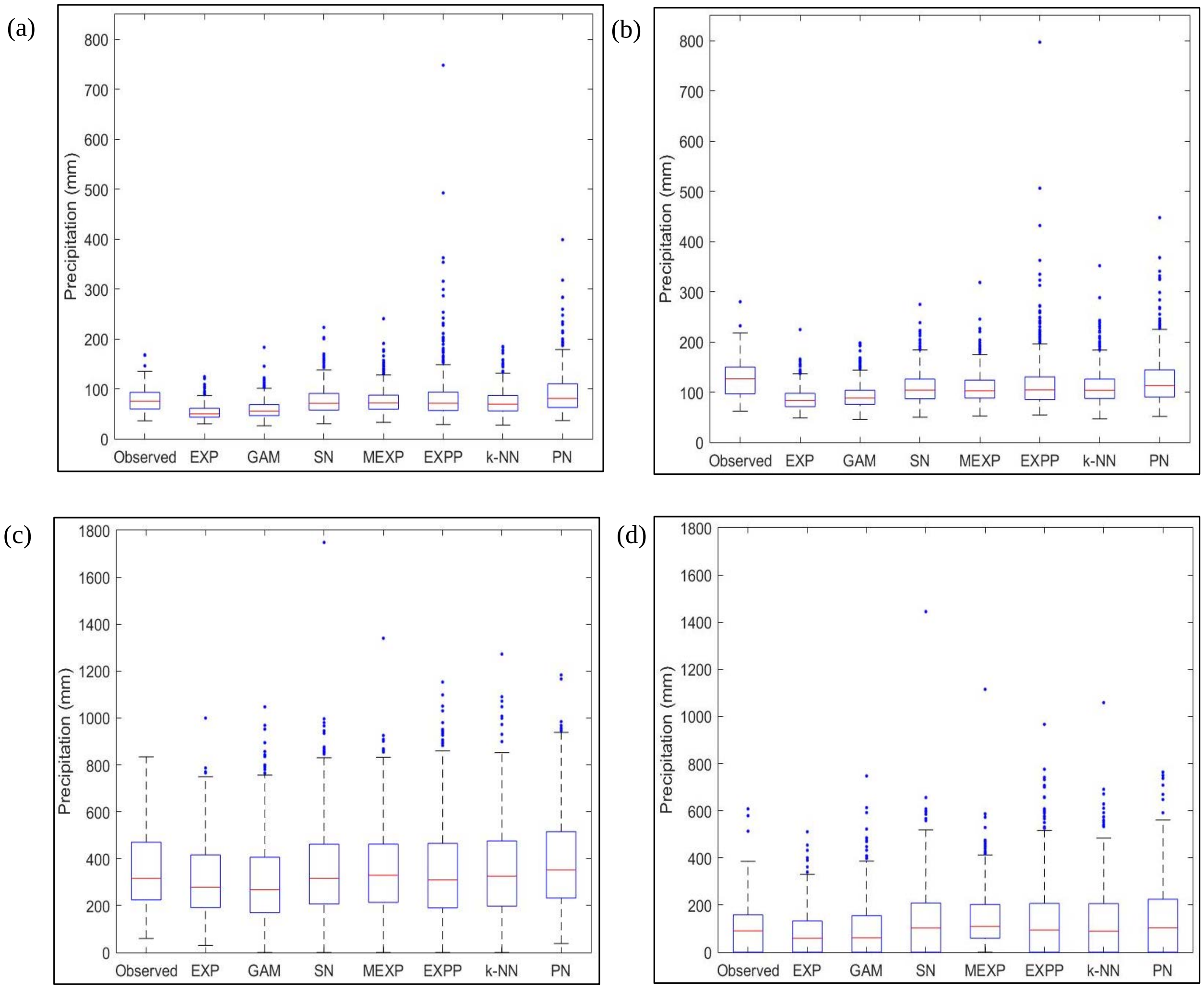


Figure 3.12. Boxplot of (a) RX1day, (b) RX5day, (c) R95p and (d) R99p of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for Ashokan watershed.

Extreme Value Theory Analysis

We applied extreme value theory (EVT) to evaluate the abilities of SWG to simulate the probabilistic structure of observed extreme precipitation events. The general extreme value (GEV) distribution is commonly used to fit block maxima values. The GEV distribution is defined as

$$G\ (x;\mu;\sigma;\xi) = \exp\left[-\left(1+\frac{\xi\ (x-\mu)}{\sigma}\right)^{-1/\xi}\right] \text{where } 1+\frac{\xi\ (x-\mu)}{\sigma} > 0 \qquad (3.5)$$

for the location $-\infty < \mu < \infty$, scale $\sigma > 0$ and shape $-\infty < \xi < \infty$ parameters. The GEV distribution was fitted to the annual maximum precipitation time series to estimate precipitation magnitudes associated with return periods of 50, 75 and 100 years (Figure 3.13). In GEV the return level associated with the extreme events can be calculated in three steps: 1) calculate the annual maximum rainfall; 2) fit a GEV distribution; 3) estimate the return level by calculating the quantiles of the GEV distribution (Coles, 2001).

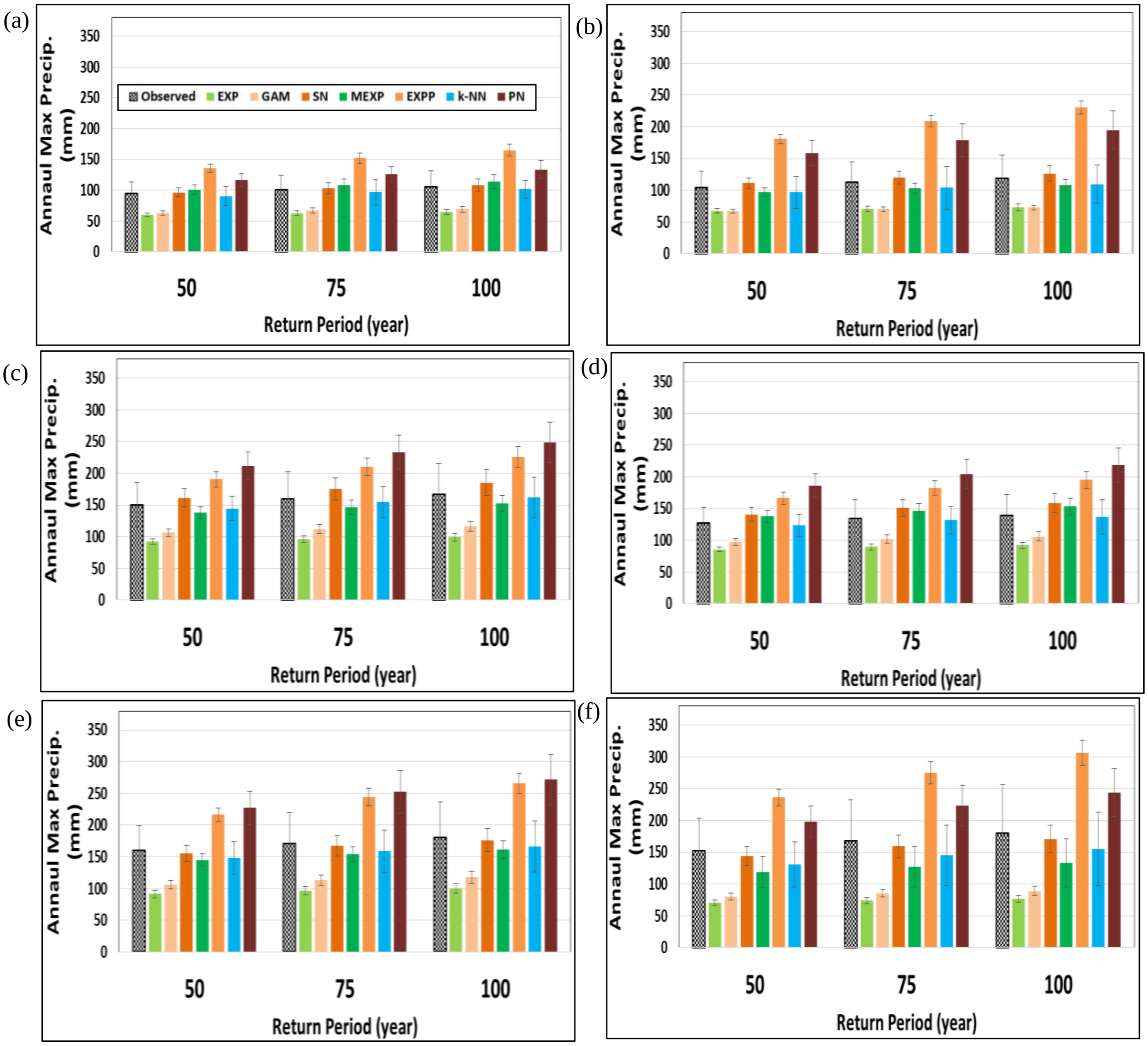


Figure 3.13. Annual maximum daily precipitation levels at the 50, 75 and 100 year return periods of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for a) Cannonsville, b) Pepacton, c) Neversink, d) Rondout, e) Ashokan and f) Schoharie.

SN, MEXP and k-NN are superior to the other distributions in simulating extreme value statistics. EXP and GAM models consistently underestimate precipitation amount corresponding to all return periods. EXPP and PN do not accurately reproduce the upper tail of the observed daily precipitation due to their exceptional overestimation. In view of MAPE, SN, MEXP and k-NN fall in either the ‘highly accurate’ or ‘good forecast’ category (Table 3.6). On the other hand, EXP, GAM, EXPP and PN provide ‘reasonable forecasts’.

Table 3.6. Mean absolute percentage error (MAPE) of 95th (Q95) and 99th (Q99) of observed and generated daily precipitations from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for all watersheds.

	EXP	**GAM**	**SN**	**MEXP**	**EXPP**	**k-NN**	**PN**
Q95	13.82	10.71	2.08	2.25	1.32	3.01	2.83
Q99	24.04	19.61	1.8	1.85	1.71	2.76	8.68

Table 3.7. Mean absolute percentage error (MAPE) of 50, 75 and 100 year return level of observed and generated annual maximum precipitation from seven models (EXP, GAM, SN, MEXP, EXPP, k-NN and PN) for all watersheds.

Return Level	**EXP**	**GAM**	**SN**	**MEXP**	**EXPP**	**k-NN**	**PN**
50 year	40.06	33.76	5.45	9.93	43.86	6.71	38.89
75 year	41.36	34.77	6.23	11.17	51.31	6.5	43.44
100 year	42.27	35.48	6.78	12.07	56.92	6.37	46.78

Summary and Concluding Remarks

This work documents the application and tests the skill level of several SWGs over six watersheds that comprise the primary source for New York City Water Supply System. This study compares the skill of three models for generating precipitation occurrence and seven models for simulating precipitation amount in the study area. Precipitation occurrence is adequately simulated by first, second, and third order MC models for all watersheds. Considering that no significant improvement is achieved when increasing the order of the MC model, and that the number of parameters in an MC model increases exponentially with each increase in order (Wilks, 1999), under the principle of parsimony the first order MC model (MC1) is selected to simulate precipitation occurrence.

Of the seven models evaluated for generating daily precipitation amount (Table 3.3), five are based on parametric distributions, one is based on a semi-parametric k-nearest neighbor bootstrapping technique, and one is based on a 2nd degree polynomial curve-fitting approach. Parametric distribution models include three single distributions (1-parameter EXP, 2-parameter GAM and 3-parameter SN), one compound distribution (3-parameter MEXP) and one hybrid distribution (3-parameter EXP-GP). While all these models reasonably reproduce mean daily precipitation for all the watersheds, their skills differ with regard to standard deviation, skewness coefficient and extreme characteristics of daily precipitation. With the exception of SN, the other two single distributions (EXP and GAM) consistently underestimate the standard deviation and skewness coefficient of daily precipitation for all watersheds because they are unable to preserve the shape of the daily precipitation distribution at the tail. As a result, these two distributions also underestimate the extreme indices and return values of annual maxima. The performance of SN is consistently better than the other single distributions in all respects. Such improved performance may be attributable to the inclusion of additional parameters which increases the model’s flexibility, and also to the use of method of moment rather than maximum likelihood to estimate model parameters (Chen and Brissette, 2014). As in earlier studies (Wilks, 1999; Chen and Brissette, 2014) the compound distribution MEXP performed better than single distributions except SN perhaps because it specifically takes the entire range of precipitation distribution into account, not just the bulk. The EXPP distribution has the unfortunate tendency to overestimate extremes, sometimes by an order of magnitude. This is due to the fact that when the generalized Pareto distribution is used to simulate the tail of the daily precipitation distribution, a few cases of unreasonably large values are typically generated (Li *et al*., 2013). Chen and Brissette (2014) described this issue in detail in their study. The k-NN method consistently performs better at reproducing the observed precipitation characteristics including extreme events for all the watersheds. Like EXPP, PN also overestimates precipitation amount in terms of standard deviation, skewness and extremes. It may be that the 2nd degree polynomial is not the perfect choice to fit the Weibull experimental frequency distribution to observe daily precipitation at this study location, though an evaluation of other polynomials, which have not previously appeared in the literature, is beyond the scope of this analysis. However, for simulating monthly and annual mean precipitation, all models are quite comparable though they are unable to reproduce the variability as they do not take into account the low-frequency component of climate variability.

In conclusion, though our results are fairly typical in the sense that some models perform better with regard to some metrics, while others are superior with regard to other metrics, we are able to identify three models (SN, MEXP, and k-NN) that, when combined with the first order MC model, consistently reproduce the observed statistics of daily precipitation, including statistics of the extreme tail of the distribution. However, as the k-NN method is based on a resampling technique, k-NN is inappropriate for use in climate change impact studies due to its inability to produce any values that are not found in the observations, nor can it extrapolate to produce values that are plausible but outside of the observed range (Chen and Brissette; 2014). Therefore, the identification of two models (SN and MEXP) that are appropriate for use in this region will be valuable in future vulnerability studies.

The application of these models in vulnerability studies will require at least two additional analyses. Several studies have examined the inability of SWGs to represent the low-frequency variability of climatic parameters; an additional analysis will evaluate several known approaches to incorporate such variability into SWGs (Chen *et al*. 2010; Khazaei *et al*. 2013). Subsequently, vulnerability studies will be performed by coupling SWGs with a response model (or a series of response models) to facilitate a more complete documentation of system vulnerabilities. A number of researchers have been developing these so-called “bottom-up” approaches whereby SWGs are used in conjunction with GCM climate scenarios and response models to evaluate water system vulnerabilities to extreme hydrological events including storms, floods, and droughts (Wilby and Desai, 2010; Brown *et al*., 2011; Prudhomme *et al.* 2010, 2015). The advantage of such approaches is that they allow the development of a quantitative and potentially Bayesian approach that includes future scenarios that are outside the range of GCM predictions, allowing water managers to consider potential future conditions when making decisions.

3.3. Evaluating a Stochastic Weather Generator (SWG) using simulated streamflow responses to synthetic weather time series

3.3.1. Introduction

The New York City water supply is currently the largest unfiltered drinking water utility in the US serving over 9 million consumers. Historically about 90% of this water is from the West of Hudson (WOH) watersheds (Figure 3.14). Recent studies have reported increasing trends in total precipitation, and in the frequency and magnitude of extreme precipitation events in WOH watersheds (Matonse and Frei 2013). The potential effects of these trends may pose challenges for both water quality (such as increased sediment and nutrient loading) and quantity (such as reservoir storage and management). NYC DEP’s Climate Change Integrated Modeling Project (CCIMP) is exploring climate impacts on water resources by evaluating the effects of future climate change on the quantity and quality of water in the NYC water supply, and to evaluate how such effects could influence the use and operation of the water supply (DEP 2014). Phase I of CCIMP followed a “top down” approach with the goal of making an initial estimate of climate change impacts using available General Circulation Model (GCM) data sets and DEP’s available suite of watershed, reservoir, and system operation models. The general findings of Phase I included a projected shift in the timing of spring snowmelt from a distinct peak in late March and April to a more consistent distribution throughout the winter and autumn. This projected shift is a function of increased temperatures, which will cause less precipitation to fall as snow and faster melting of the snowpack (Zion *et al*. 2011).

In Phase I, climate change scenarios were derived from GCMs using simple manipulation of current climate observations (e.g., Change Factor Methodology, or CFM, (Anandhi *et al*. 2011)). Using the CFM to downscale GCM output leaves long-term historical patterns of the timing of storm events unchanged, even as the magnitude of these events change based on projected seasonal changes in climate. In addition, the monthly change factor methodology does not account for changes in the frequency or duration of storm events or changes in inter-storm periods. It is widely accepted that extreme hydrologic events, including floods and droughts, are responsible for a vast majority of water resources problems including a disproportionate transport of nutrients and sediments into receiving waterbodies. The ongoing Phase II of CCIMP is exploring the use of “bottom-up” or vulnerability based methods to explore climate impacts on water resources with a focus on extreme hydrologic events. Stochastic weather generators (SWGs) are an integral component of the bottom up approach. The objective of this study was to use the stochastic weather generator model developed for the region (Section 3.2) and a watershed model to simulate streamflow responses to synthetic weather time series in the WOH watersheds.

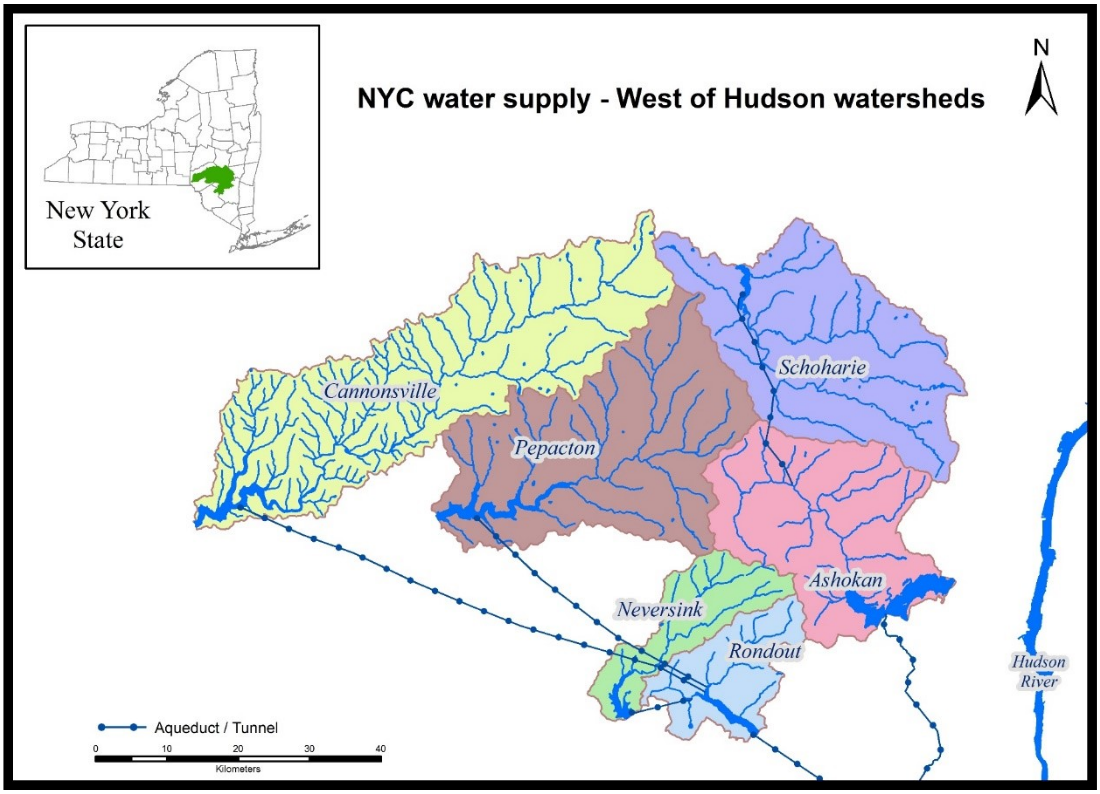


Figure 3.14. Location of six WOH watersheds of NYC water supply.

3.3.2. Methods

Sensitivity Surface

Sensitivity surfaces can be used to depict system responses to a range of climate scenarios. As a proof of concept of this approach, we generated a streamflow sensitivity surface for one of the WOH watersheds by changing the observed precipitation (P) and air temperature (T) values by a certain value (-25% to +25% for P and -1°C to +7°C for T), and using it as input to the calibrated Generalized Watershed Loading Function (GWLF) -Variable Source Area (VSA) watershed model. Specific event periods such as droughts and/or extended flood events can be identified on a response surface in addition to water quality parameters of interest (Figure 3.14).

The GWLF-VSA model

The GWLF-VSA (Schneiderman *et al*. 2007) is a lumped parameter watershed model based on the original GWLF model (Haith and Shoemaker 1987) that simulates daily streamflow and monthly sediment and nutrient loads at a watershed scale. Runoff is calculated using the USDA Soil Conservation Service (SCS) runoff Curve Number (CN) methodology. The CN is an empirical parameter for estimating direct or infiltration excess runoff. The CN values can range from 30 to 100 depending on the hydrologic soil group, land cover type, hydrologic condition, and antecedent runoff condition (Rawls *et al*. 1993). The model simulates streamflow contributions from runoff and baseflow. Daily snowmelt is estimated based on temperature degree-day, with a daily updating of a single watershed-wide snowpack. Inputs to GWLF streamflow model include daily precipitation, air temperature, solar radiation, and relative humidity data. One important modification introduced in the VSA version of GWLF (Schneiderman *et al*. 2007) is accounting for surface runoff distribution across the landscape based on a saturation-excess interpretation of the SCS Curve Number equation and a topographic index of saturation probability. This modification is important since saturation-excess is the dominant process for runoff generation in the study region (Frankenberger *et al*. 1999, Lyon *et al*. 2004, Steenhuis *et al*. 1995, Walter *et al*. 2003). The calibrated model has been applied in previous studies in NYC watersheds (Matonse *et al*. 2011, Mukundan *et al*. 2013).

Streamflow simulation

Calibrated GWLF-VSA watershed model was used to simulate daily streamflow for the predominantly forested six WOH watersheds for the period from 1950 to 2009 (60 years) using observed daily precipitation and air temperature (T_{MAX} and T_{MIN}) data from the cooperator stations recognized by the National Climate Data Center, obtained from the Northeast Regional Climate Center. The daily precipitation station data is averaged for the entire basin using a Thiessen polygon method and is used in the model (Burrough 1986). Potential biases of model input due to spatial averaging of precipitation were corrected with a constant multiplicative factor used in the calibration process. Watershed-specific synthetic weather (precipitation and air temperature) time series for 600 years (10 x observed to avoid biases) was generated using the WeaGETS SWG and was used as input to the GWLF-VSA model to simulate streamflow under a wide range of hydrologic conditions. Other required climate data including solar radiation and relative humidity were computed using a climate interpolation model that uses air temperature and precipitation data (MT-CLIM, (Running *et al*. 1987)). Streamflow simulated using observed weather data was compared with simulations using SWG-based synthetic weather time series for each of six WOH watersheds.

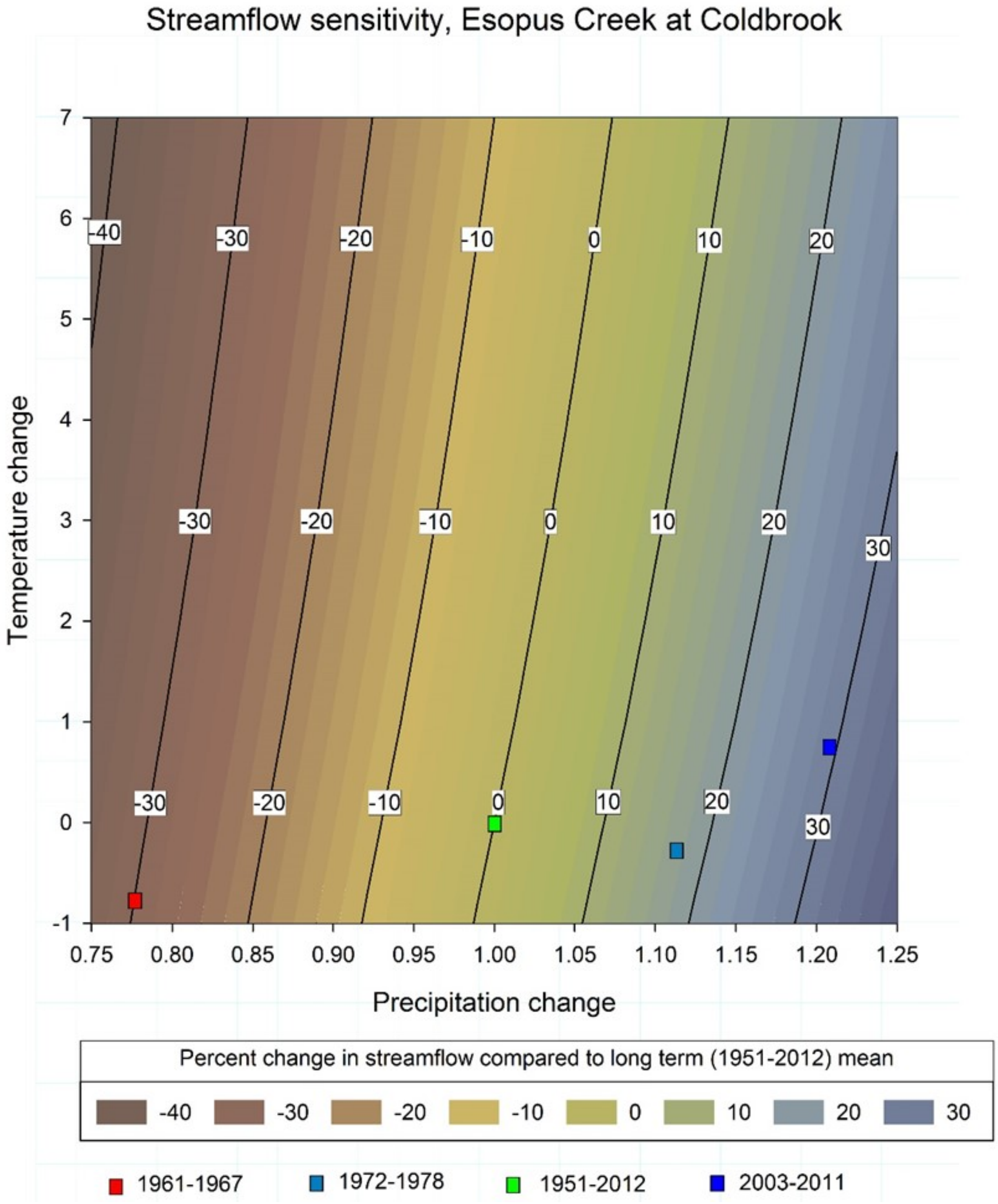


Figure 3.15. A representative sensitivity surface showing changes in streamflow in response to relative changes in precipitation and air temperature. Long term average conditions (green square), drought of 1960s (red square) and periods of high runoff (blue squares) are also marked.

3.3.3. Results

Average daily streamflow simulated using observed weather as input was comparable with streamflow simulated using SWG weather. The seasonal variation in streamflow was well captured by model runs using both observed and SWG derived weather (Figure 3.16). SWG based streamflow showed lesser variability in average daily values, which may be due to averaging of relatively large number of daily values (600 yr. vs. 60 yr. with observed weather). Peak in annual streamflow hydrograph typically observed during winter in the region were consistently lower when using SWG based weather in all six watersheds. Streamflow during this period is highly influenced by snowmelt runoff and temperature plays a greater role during this period than other months. The observed biases in streamflow are possibly due to the stochastic generation of temperatures because different temperature series can alter the hydrologic regime by influencing the snowmelt processes as reported by Li *et al*. (2013).

Long-term median daily streamflow conditions observed in the WOH watersheds were simulated very well using SWG derived weather and was comparable with those derived using observed weather as GWLF model input. There was close agreement between the 50th percentile streamflow simulated by the SWG and observed weather (Figure 3.17), with R^2 and Nash-Sutcliffe coefficient (NSE) ranging from 0.94 to 0.97 and 0.93 to 0.97, respectively (Table 3.8). A similar comparison for the 95th percentile streamflow values representing high streamflow events gave R^2 and NSE between 0.64 to 0.77 and 0.61 to 0.76 respectively (Table 3.9). These performance measures estimated using the two sets of simulations are comparable and acceptable considering the random nature of high streamflow events during a year. Clustering of points around the 1:1 regression line indicates no systematic bias in the magnitude of events for any of the six WOH watersheds (Figure 3.18). The temporal structure of daily streamflow examined using a partial auto-correlation function showed consistent pattern in both observed and SWG based time series with a strong and significant lag of 1-day, and persisting for up to 10 or 12 days.

One of the focus areas in the ongoing CCIMP Phase II “bottom-up” approach is the ability to represent extreme streamflow events of different magnitudes and frequency of occurrence. This represents an improvement over Phase I evaluation of climate change effects using change factor method on GCM outputs, where only magnitude of events changes, leaving the frequency and timing of events unchanged. Figure 3.19 compares the magnitude of annual maximum and minimum daily flows simulated using observed and SWG weather. It is clear that the SWG model is capable of producing weather and associated streamflow events not produced by the observed weather. This was generally consistent for all six WOH watersheds. Figure 3.20 demonstrates the need for generating longer time series of weather (as in this study) using the SWG to capture the observed variability in high streamflow events and produce extreme events that are beyond the observed range. In this example for the Ashokan watershed, it can be seen that maximum variability in annual maximum daily streamflow was seen in the 8th (out of 10) series of 60 yr. SWG weather.

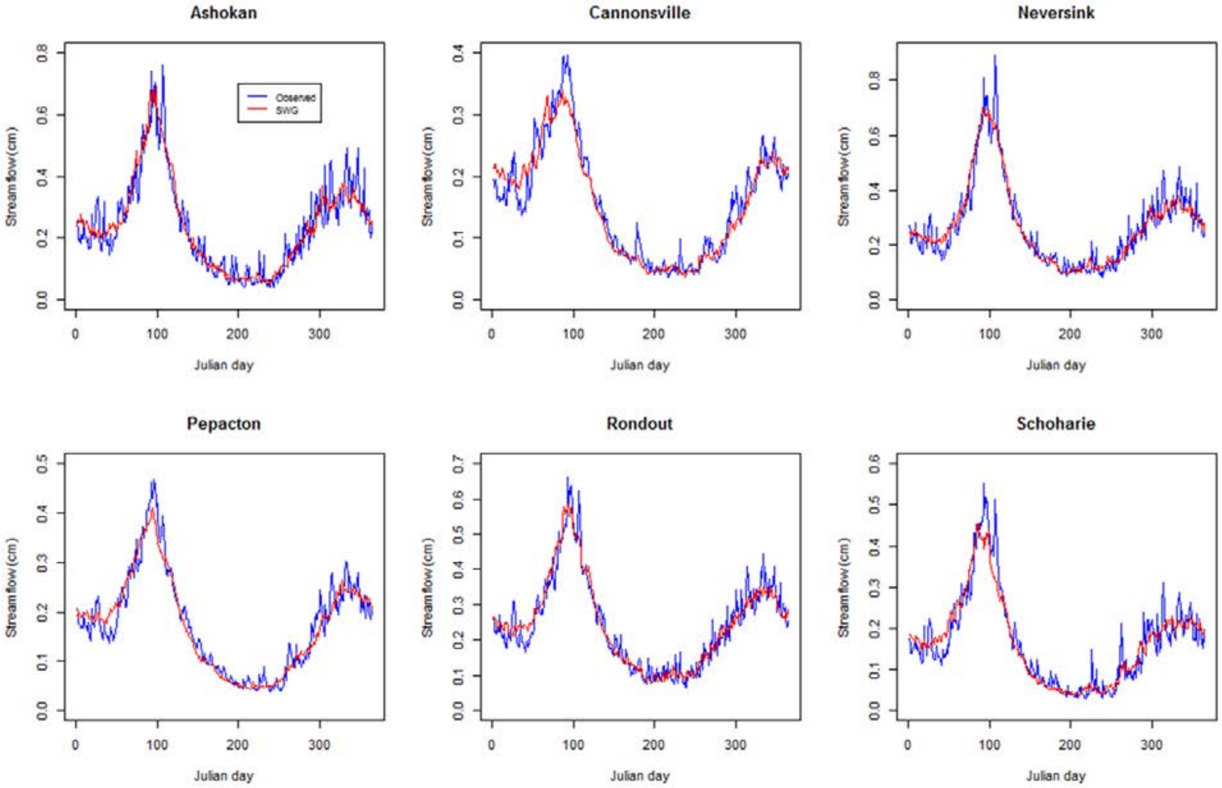


Figure 3.16. Simulated daily average streamflow (cm) using 60 yr. observed weather as input (blue) and using 600 yr. SWG weather as input (red).

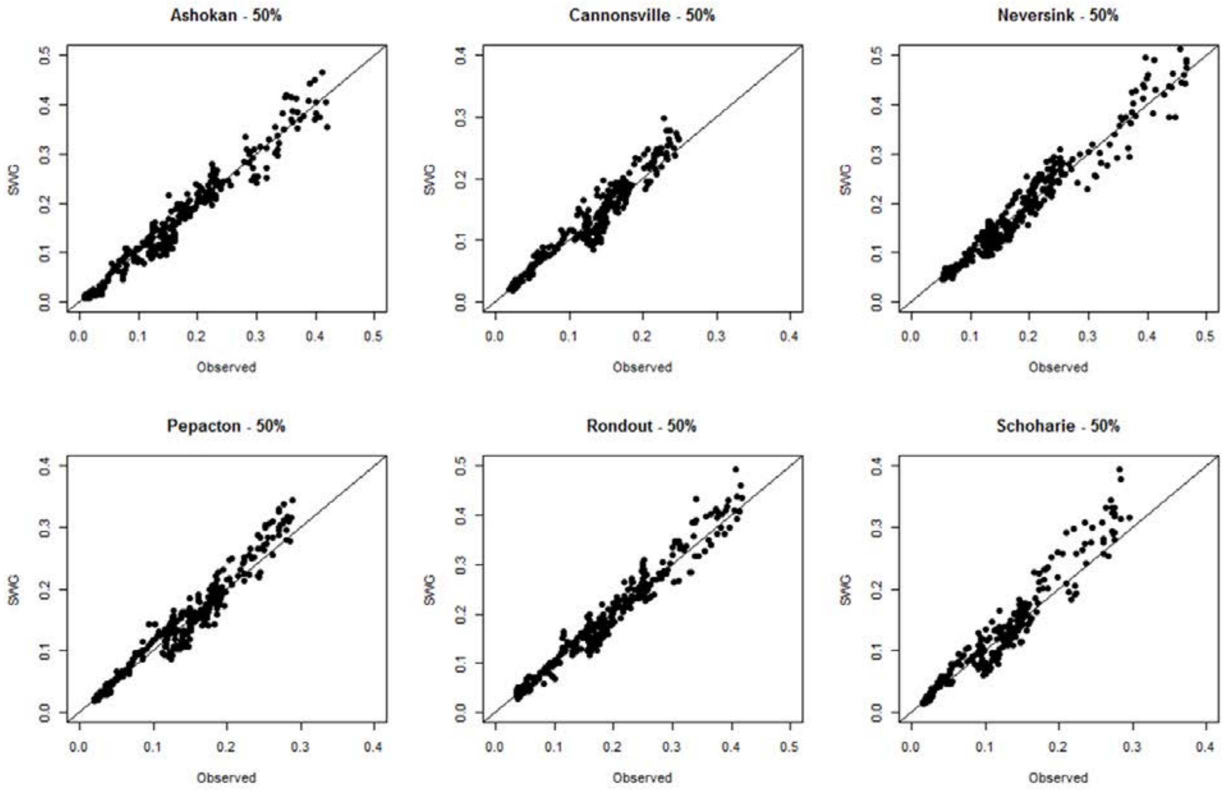


Figure 3.17. Comparison of 50th percentile of simulated daily streamflow (cm) using SWG and observed weather plotted on a 1:1 regression line.

Table 3.8. Performance measures for SWG based weather in reproducing the 50th percentile simulated daily (n=365) streamflows using observed weather.

Watershed	**R^2**	**NSE**
1. Ashokan	0.97	0.97
2. Cannonsville	0.96	0.96
3. Neversink	0.95	0.95
4. Pepacton	0.96	0.95
5. Rondout	0.97	0.96
6. Schoharie	0.94	0.93

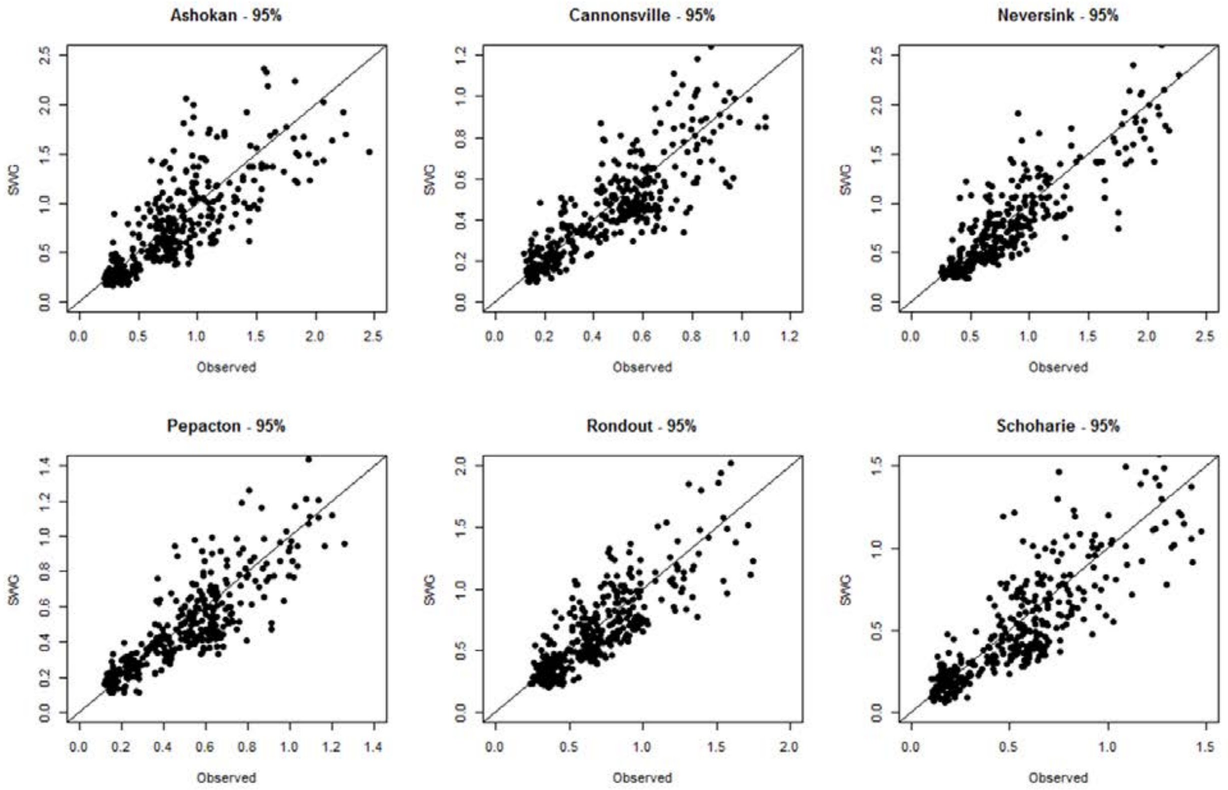


Figure 3.18. Comparison of 95th percentile of simulated daily streamflow (cm) using SWG and observed weather plotted on a 1:1 regression line.

Table 3.9. Performance measures for SWG based weather in reproducing the 95th percentile simulated daily (n=365) streamflows using observed weather.

Watershed	R^2	NSE
1. Ashokan	0.64	0.61
2. Cannonsville	0.72	0.70
3. Neversink	0.77	0.76
4. Pepacton	0.74	0.72
5. Rondout	0.68	0.66
6. Schoharie	0.71	0.69

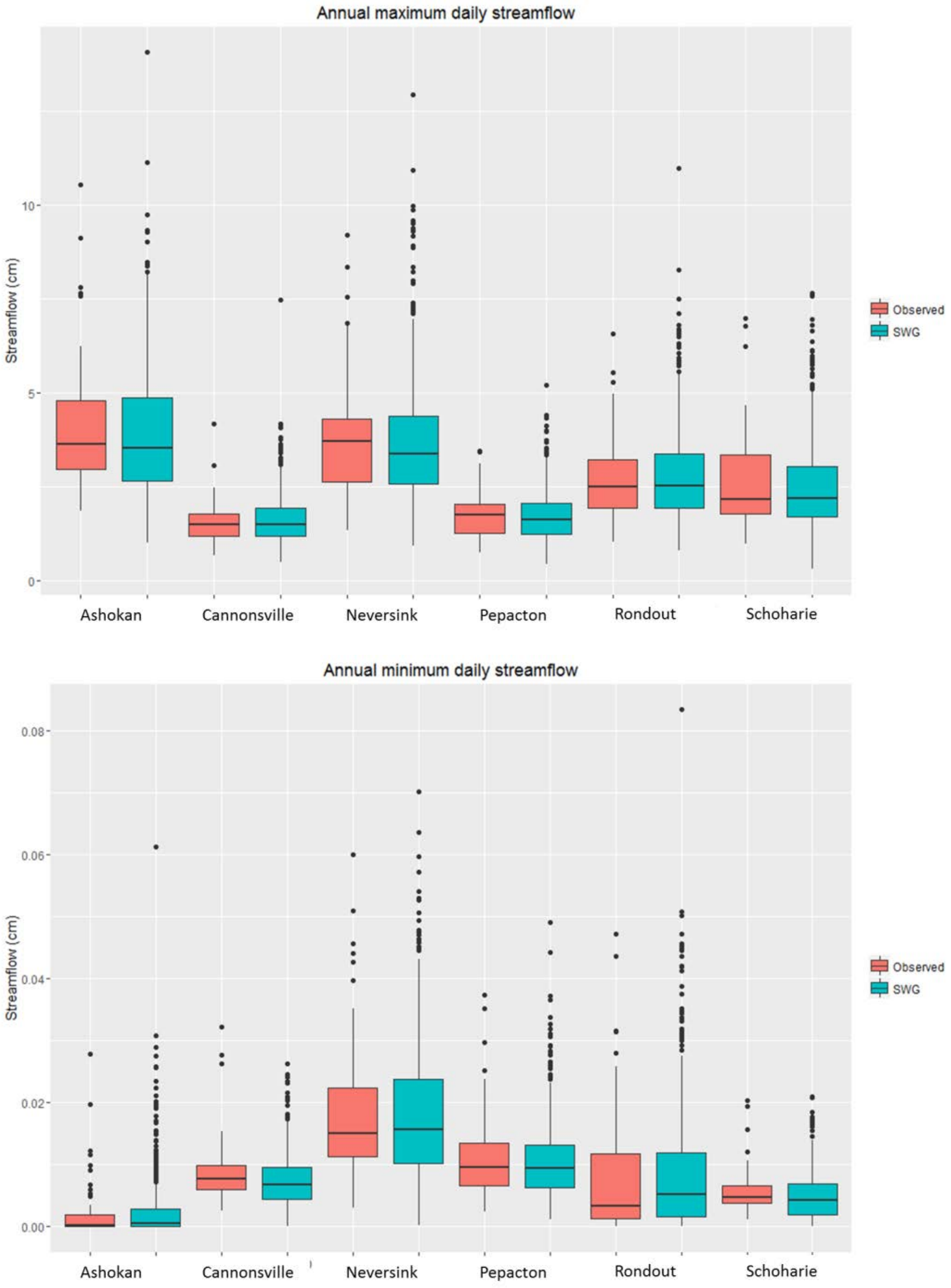


Figure 3.19. Boxplots showing annual maximum (top) and minimum (bottom) daily streamflow (cm) simulated using observed and simulated weather for six WOH watersheds.

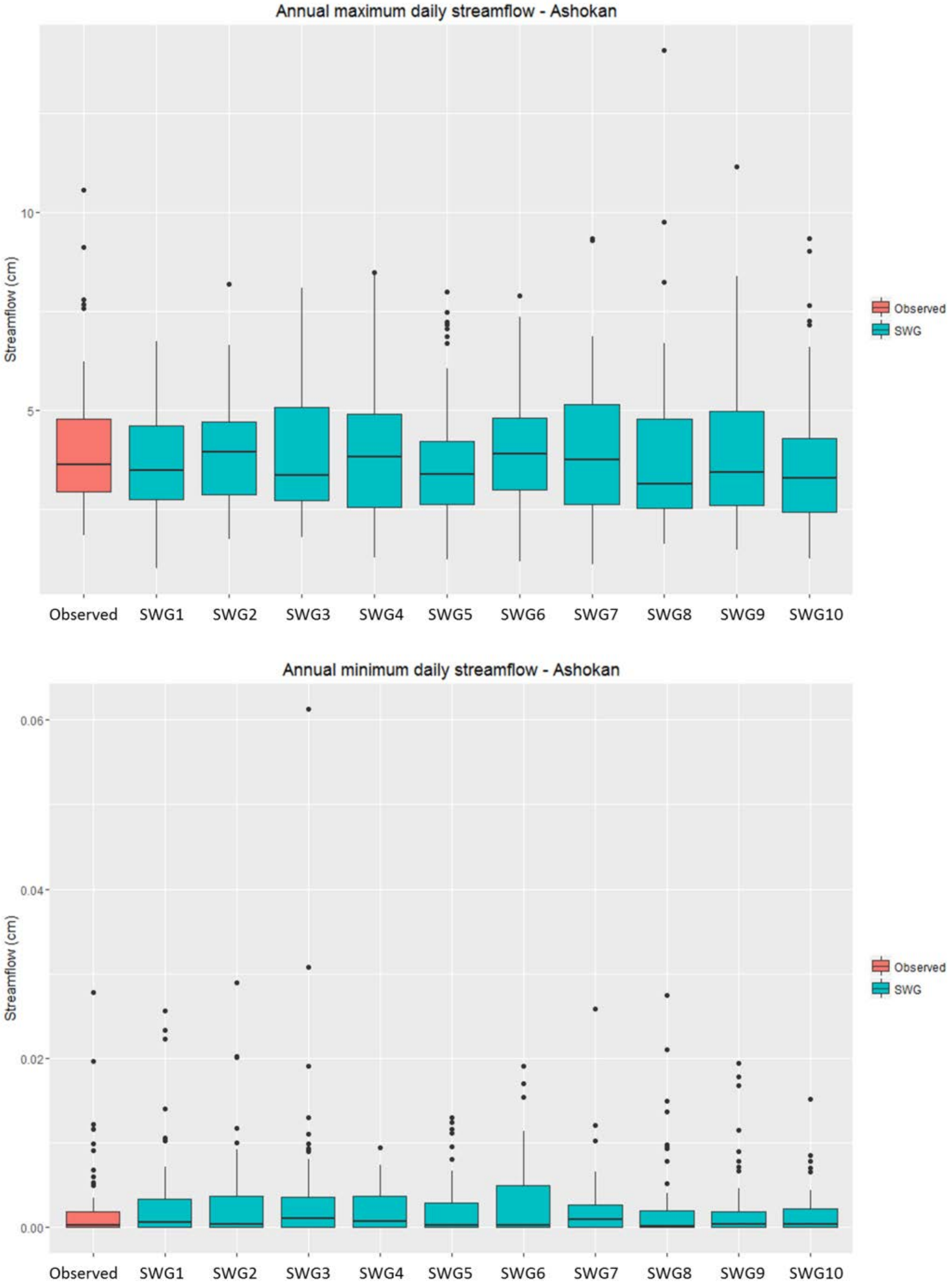


Figure 3.20. Boxplots showing annual maximum (top) and minimum (bottom) daily streamflow (cm) simulated using observed (60 yr.) and simulated (60 yr. x 10 SWG) weather for the Ashokan watershed.

3.3.4. Conclusions and future work

- Long-term average daily streamflow simulated using the synthetic weather time series were comparable to values simulated using observed long-term (1950-2009) weather time series; the seasonal variation in streamflow was captured by the SWG.

- Streamflow values based on SWG were comparable to observed weather based streamflow under all hydrologic conditions (low, average, and high -flows).

- As expected, the synthetic weather produced extreme hydrologic events (floods) not found in streamflow predicted using observed weather; this supports the “bottom-up” approach.

- Longer time series of SWG weather (10 times observed in this study) is needed to capture the required variability in streamflow and generate extreme hydrological events for bottom up analysis.

- In future work we will use GCM output from CMIP-5 (discussed in Section 6.3.2) in conjunction with SWG to generate multiple scenarios of changing climate in the region.

- Sensitivity surfaces based on GCM scenarios and model runs will be used to evaluate water supply system vulnerability to climate change using an integrated suite of watershed-reservoir modeling framework.

4. Model Development and Application

4.1. Predicting saturation-excess runoff distribution with a lumped hillslope model: SWAT-HS

4.1.1. Introduction

Watershed-scale models are important tools in estimating discharge and non-point pollutant loading in reservoirs and river systems. Recent reviews by Borah and Bera (2003), Yang and Wang (2010), and Daniel *et al.* (2011) summarized numerous applications of watershed-scale models for simulating flow, soil erosion, sediment and nutrient transport, and for evaluating the effect of climate change, land use change, and best management practices (BMPs) in watersheds of different sizes with different hydrologic, geologic, and climatic conditions. Since surface runoff is the primary mechanism transporting sediments, agricultural chemicals, and pathogens to receiving waters, it is essential to identify landscape positions of surface runoff correctly in order to estimate pollutant fluxes and properly apply conservation practices to improve water quality (Rode *et al.*, 2010). The prediction of areas generating surface runoff depends on the assumed type of runoff simulated in the model: infiltration-excess runoff and saturation-excess runoff. Infiltration-excess runoff occurs when precipitation intensity exceeds infiltration rate of water in the soil (Horton, 1933; Horton, 1940) and can be observed in arid areas with soil crusts and/or seals, surface irrigated fields, paved urban areas, and during storms with very high rainfall intensities. Saturation-excess runoff is common in humid, well-vegetated areas (Dunne and Leopold, 1978) where soils are well structured and infiltration capacity is high. Saturation-excess runoff is generated in Variable Source Areas (VSAs) where the groundwater table is close to the surface during rainfall periods (USDA-SCS, 1972). The location of VSAs depends on topographic position in the landscape and soil transmissivity, and their extent varies over time with landscape wetness (Dunne and Black, 1970).

Many models are available that simulate watershed hydrology. A list of approximately 100 models is given in http://hydrologicmodels.tamu.edu/models.htm. These models range from very complex distributed models (e.g. MIKE SHE (Refsgaard and Storm, 1995)) to simple and lumped models (e.g. the Generalized Watershed Loading Function model (GWLF) (Haith and Shoemaker, 1987)). They also vary in how they simulate runoff generation. Most models are based on infiltration-excess runoff and some are based on saturated-excess runoff or a combination of the two, as in MIKE-SHE. Infiltration-excess models use land cover and soil types to determine runoff amounts and locations (Bayabil *et al.*, 2010). Since detailed measurements of infiltration rates for various land covers is generally lacking, the SCS curve number (CN) method (USDA-SCS, 1972) is employed. Examples include the Agricultural Non-Point Source Pollution model (AGNPS) (Young *et al.*, 1989) and its upgrade, the Annualized Agricultural Non-Point Source model (AnnAGNPS) (Bingner and Theurer, 2001); the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) (Feldman, 2000); the Generalized Watershed Loading Functions model (GWLF) (Haith and Shoemaker, 1987), the Storm Water Management Model

(SWMM) (Krysanova *et al.*, 1998) and the Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1998).

There are several studies and watershed-scale models that take into account saturation-excess runoff and VSA hydrology using different approaches. Some studies (Steenhuis *et al.* (1995); Lyon *et al.* (2004); Schneiderman *et al.* (2007)) predicted flow generated from VSAs using the re-interpretation of the SCS-CN equation to follow VSA theory. Using a statistical dynamic approach, TOPMODEL (Beven and Kirkby, 1979; Quinn and Beven, 1993), the Variable Infiltration Capacity (VIC) model (Wood *et al.*, 1992; Liang and Lettenmaier, 1994), the Probability Distributed Model (PDM) (Moore, 2007), the Xinanjiang model (Zhao *et al.*, 1995), and the Australian Water Balance Model (AWBM) (Boughton, 2004) simulate saturation-excess runoff by introducing a statistical distribution of soil water storage capacity by different methods. In TOPMODEL, the statistical distribution of storage capacity is a function of topographic index. The VIC, PDM, and Xinanjiang models use a Pareto distribution of storage capacity to derive the saturated fraction of the watershed without simulating the specific spatial locations of saturated areas. Boughton (2004) in the AWBM model divided the watershed into a set of sub-areas of increasing storage capacity with areas and storage capacities are determined by analysis of rainfall and runoff records. A similar variable bucket approach has been proposed in the variable bucket capacity (VBC) model (Sivapalan *et al.*, 1997) as an adaptation of TOPMODEL. In an explicit modeling approach, as in the Distributed Hydrology Soil Vegetation Model (DHSVM) (Wigmosta *et al.*, 1994), the Soil Moisture Routing model (SMR) (Frankenberger *et al.*, 1999) and MIKE SHE (Refsgaard and Storm, 1995; Refsgaard *et al.*, 2010), the watershed is discretized into contiguous grids with water explicitly routed from cell to cell based on topography.

Several efforts have been carried out to include saturation-excess runoff in adaptations of the SWAT model. In SWAT-VSA, Easton *et al.* (2008) used a topographic wetness index to redefine HRUs, which distributed overland flow in ways consistent with VSA hydrology by redefining the definition of CN and available water content. In SWAT-WB (White *et al.*, 2011), saturation deficit was defined based on soil wetness classes determined by the values of soil topographic index, and a water balance was applied instead of the CN method to calculate runoff as water in excess of local soil water storage capacity. Similar to the original version of SWAT, both SWAT-VSA and SWAT-WB do not simulate transfer between HRUs. While both models were found to capture the spatial distribution of saturation-excess runoff, the underlying mechanism of a water table that rises to the surface creating saturated conditions and saturation-excess runoff was not incorporated into either model. Recent efforts to incorporate landscape routing into SWAT may improve SWAT's ability to simulate VSA hydrology. Arnold *et al.* (2010) introduced a hillslope approach for SWAT that allows flow routing between three landscape units, i.e., divide, hillslope, and valley bottom. Bosch *et al.* (2010) tested this SWAT landscape model in a low-gradient coastal plain watershed and showed that average annual surface runoff agreed satisfactorily with observations, but monthly simulated results differed significantly from measurements and estimates of groundwater flow were greater than expected. It was concluded that the landscape model may require additional details to properly describe the interactions between soil surface, vadose zone, and groundwater. Recently Rathjens *et al.* (2015) introduced a fully distributed grid-based version of the SWAT landscape model, which incorporated the hillslope approach of Arnold *et al.* (2010) in simulating landscape flow routing between grids.

Rathjens *et al.* (2015) noted that because the curve number was used, the model seemed unable to represent temporal variations in flow processes related to saturation conditions in the 334 km^2 Little River Watershed near Tifton, Georgia, USA. Because data for operating and evaluating a detailed grid-based model are not always available, the grid-based SWAT landscape model is most efficient for small-scale and data-rich watersheds.

Models such as VSLF (Schneiderman *et al.*, 2007), SWAT-VSA (Easton *et al.*, 2008) and SWAT-WB (White *et al.*, 2011) developed for hilly or mountainous landscapes where saturation-excess runoff is dominant fell short because the uplands were not connected with the valley bottom. The distributed "connected" versions of SWAT (Arnold *et al.*, 2010; Rathjens *et al.*, 2015) use the CN method, and therefore do not predict runoff by saturation excess. The VIC (Wood *et al.*, 1992; Liang and Lettenmaier, 1994), PDM (Moore, 2007), and Xinanjiang models (Zhao *et al.*, 1995) cannot predict the location of saturated areas. These three models, as well as the TOPMODEL focus on flow modeling while their ability to model water quality is limited.

The objective of this study is to develop a computational and data efficient model that can represent saturation-excess runoff and predict locations generating this type of runoff in hilly and mountainous landscapes with restrictive layers and shallow soils. This model can then be the basis for water quality and erosion management models where identifying the location of the runoff is critical. We chose to modify the SWAT model because SWAT is the most widely used watershed model that predicts flow and pollutant loads from non-point sources to receiving waters under varying scenarios of land use, management, and climate change (Gassman *et al.*, 2007; Arnold *et al.*, 2010; Bosch *et al.*, 2010). Its free and open source code give opportunities for users to access, understand and adapt the model to specific case studies as well as improve the model. Although SWAT is a distributed model that simulates hydrological and water quality processes at the level of Hydrological Response Unit (HRU), a unique combination of land use, soil, and slope, SWAT is a computational and data efficient model due to its simple conceptual hydrological approach. Therefore, SWAT has been applied to many watersheds from small to large drainage areas (<10 – 500,000 km^2) (Douglas-Mankin *et al.*, 2010). With the availability of routines modeling management practices, SWAT is a useful tool to evaluate the effect of best management practices in agricultural dominated watersheds.

Our modeling approach was inspired by the work of Dooge (1986), Dooge (2005), Savenije (2010), and McDonnell *et al.* (2007) in that the underlying heterogeneous and complex spatial processes in the landscape are self-organizing. Our landscape methodology is process-based and avoids the fine detail of the Darcy scale (Savenije, 2010; Montanari and Koutsoyiannis, 2012; Beven and Young, 2013). Due to the predominant topographical and gravitational influences on natural slopes (Hewlett, 1961), there is a regularity in the wetting pattern of humid catchments, as shown by other researchers (Boughton, 1989; Grayson and Blöschl, 2001; Sayama *et al.*, 2011) and our own experience (Schneiderman *et al.*, 2007; Bayabil *et al.*, 2010; Tilahun *et al.*, 2014). Based on this evidence, we divided the watershed into conceptual groups of soil moisture storage units with similar wetting up characteristics. These are called wetness classes. In addition, we pose (in agreement with Savenije (2010)) that for shallow soils over a restrictive layer found in most hilly or mountainous landscapes, lateral flow occurs when the moisture content is above field capacity and when macropores are active in moving the water down the hillslope. This flow process is necessarily nonlinear as the conductivity is greatest at the surface and decreases with depth (Brooks *et al.*, 2004; Ameli *et al.*, 2016).

We developed a modified SWAT version, referred as SWAT-Hillslope (SWAT-HS). SWAT-HS improves the simulation of saturation-excess runoff by redefining HRUs based on wetness classes, and by introducing a “surface aquifer” which is a non-linear reservoir with the ability to route interflow from “drier” to “wetter” wetness classes. We tested SWAT-HS in the Town Brook watershed in the upper reaches of the West Branch Delaware River in the Catskill Mountains where rainfall intensities rarely exceed infiltration rates and saturation-excess runoff is common (Walter *et al.*, 2003).

4.1.2. Methodology

4.1.2.1. Description of SWAT-HS

In this section we describe the SWAT-HS model and compare it with the original SWAT (SWAT2012). SWAT-HS modifies primarily the subroutines related to surface and lateral runoff generation in SWAT2012, by changing the definition of HRUs and introducing a surface aquifer. Figure 4.1 illustrates the differences in hydrological approach between SWAT2012 and SWAT-HS. In SWAT2012, surface runoff is either calculated with the curve number or the Green-Ampt method. In SWAT-HS the curve number is replaced by a saturation-excess runoff calculation and the built-in Green-Ampt routine of SWAT is employed for the few cases where the rainfall intensity is greater than the infiltration capacity of the soils.

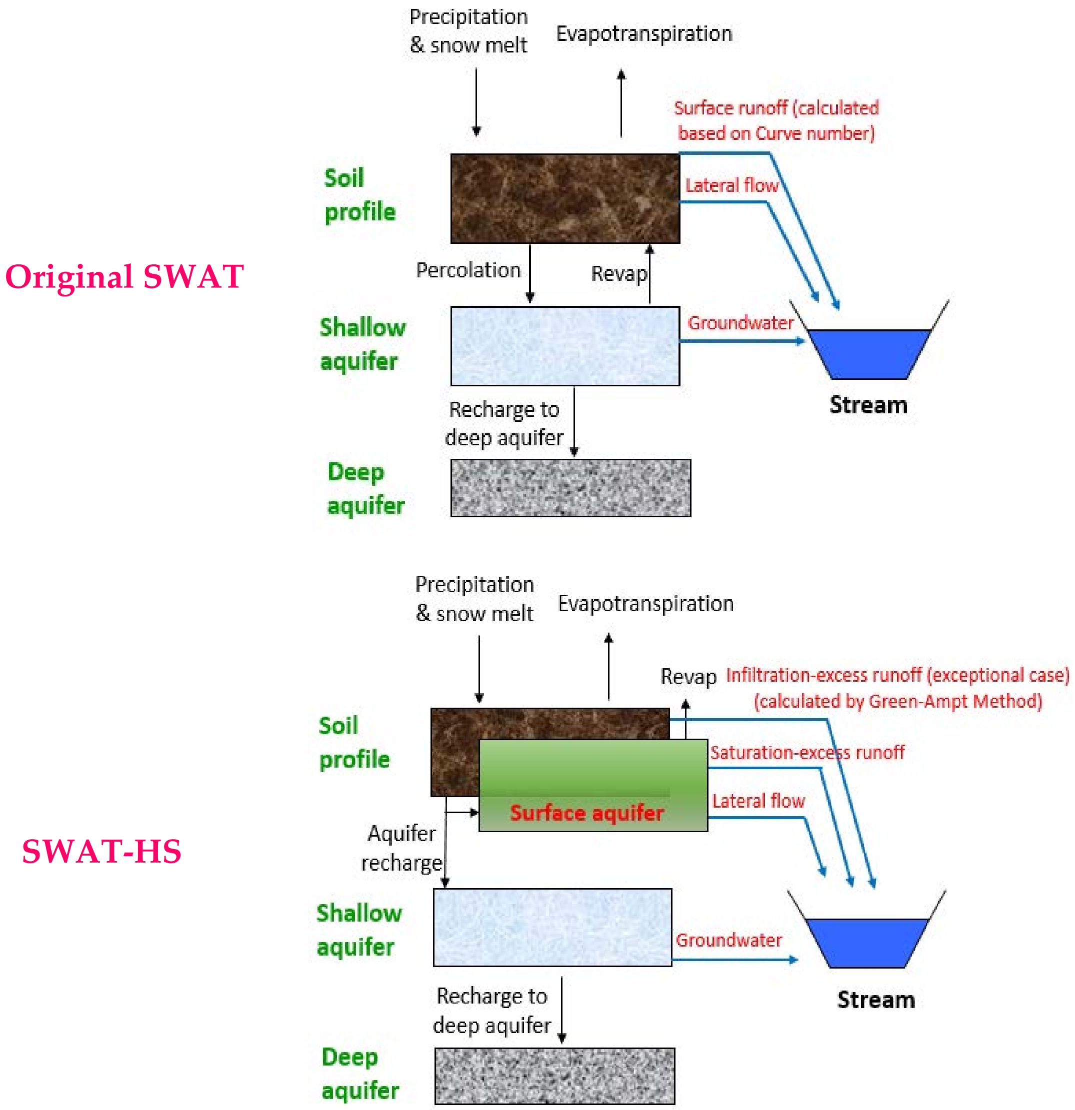


Figure 4.1 Difference in hydrological processes between the original SWAT and SWAT-HS

4.1.2.1.1. Hydrological response units, wetness classes and sub-basins

Both SWAT2012 and SWAT-HS divide a watershed into sub-basins, each of which is then split into multiple Hydrological Response Units (HRUs) (Neitsch *et al.*, 2011). The HRU is the basic modeling unit in the SWAT model. In SWAT2012, HRUs are unique combinations of soil, land use, and slope (Neitsch *et al.*, 2011). In SWAT-HS, a watershed is further divided into multiple '*wetness*' classes that are assigned with increasing soil water storage capacity from downslope to upslope areas. Soil water storage capacity is defined as the amount of water that can be stored above field capacity in the soil before the soil becomes saturated. The downslope wetness classes have lower storage capacities (which means they are wetter) while the further upslope ones have greater capacities to store water before runoff occurs (which mean they are drier). Based on this division, HRUs in SWAT-HS are redefined as unique combinations of soil, land use, slope, and wetness class.

SWAT-HS gives flexibility for users to divide the simulated watershed into wetness classes. In this study, we used topographic index (TI) as one of the possible techniques. The TI value, calculated for each grid cell, is defined as the quotient of the amount of water delivered to the cell in the landscape and the ability to transmit water through the soil at that cell (Beven and Kirkby, 1979; Beven, 1986):

$$TI = ln\left(\frac{\alpha}{tan(\beta)K_s D}\right) \tag{4.1}$$

where *TI* is the soil topographic index [unit: ln(d m $^{-1}$)], α is upslope contributing area per unit contour length (unit: m), $\tan(\beta)$ is the local surface topographic slope, K_s is the mean saturated hydraulic conductivity of the soil (unit: m d^{-1}), and D is the soil depth (unit: m).

The downslope wetness classes correspond to greater TI values while the further upslope ones have lower TI values.

The soil water storage capacity, or *edc*, is assigned for each wetness class. The distribution of the soil water storage capacities is set to follow a Pareto distribution. This Pareto distribution was used first in the Xinanjiang model (Zhao *et al.*, 1980; Zhao *et al.*, 1995), then in the VIC model (Wood *et al.*, 1992; Liang and Lettenmaier, 1994) and the PDM model (Moore, 2007) to define the statistical distribution of storage capacity within a watershed. These three models simulate saturation-excess runoff and estimate the saturated fraction of the watershed, however, they are not able to identify the specific spatial locations of saturated areas. SWAT-HS overcomes this shortcoming by categorizing the watershed into as many as 10 wetness classes, each of which has a specific spatial location, and then applying the Pareto distribution to define the distribution of soil water storage capacities for wetness classes, which is expressed the below equation and Figure 4.2.

$$edc_i = S_{\max}\left[1-\left(1-A_i\right)^{1/b}\right] \tag{4.2}$$

where edc_i is the soil water storage capacity in wetness class *i*, *Smax* is the maximum soil water storage capacity of the watershed, A_i is the fraction of the watershed for which the storage capacity is less than edc_i, and *b* is the shape parameter. The lower *edc* values are assigned to the wetness classes located in downslope areas ("wetter" wetness classes) while higher *edc* values are assigned to wetness classes in upslope areas ("drier" wetness classes).

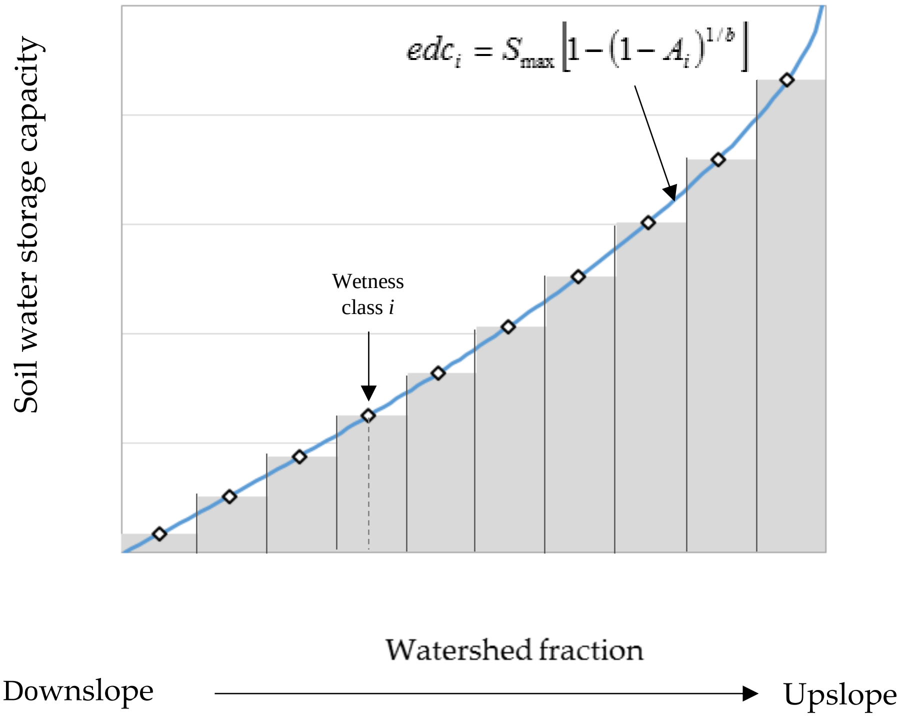


Figure 4.2. Distribution of soil water storage capacity of wetness classes.

4.1.2.1.2. Surface aquifer

In SWAT2012, HRUs in a sub-basin are individual units that are routed to the stream without interaction with each other. In SWAT-HS, interaction is created between wetness classes through a "surface aquifer" so that downslope wetness classes can receive input from upslope wetness classes. The surface aquifer in SWAT-HS represents a part of the soil profile that accumulates water above field capacity, controls soil saturation, provides water for plant use, and influences biogeochemical transformation related to saturated soil conditions. The surface aquifer connects all wetness classes across the hillslope and transmits subsurface flow that is generated from this aquifer (known as lateral flow in SWAT) laterally through the hillslope from "drier" (upslope) to "wetter" (downslope) wetness classes.

The behavior of the surface aquifer in SWAT-HS is based on the assumption that the subsurface flow system is in quasi-equilibrium (as in TOPMODEL and other statistical-dynamical hydrologic models, see Wigmosta and Lettenmaier (1999). By this assumption, the water table as a whole moves up and down as the average depth of water in the aquifer changes, while the shape of the water table in relation to the ground surface is constant over time.

Figure 4.3 illustrates the behavior of the water table in the surface aquifer in a hillslope with 10 wetness classes. The distribution of soil water storage capacities of wetness classes (Figure 4.2) results in the shape of water table across the hillslope shown as the blue line in Figure 4.3. Supposing this is the initial water table, when there is hydrological input, the water table will rise to the red line (water table after rain). The amount of water that is above the soil surface becomes saturation-excess runoff, and the remaining becomes infiltration to the soil. Based on the amount of water stored in the surface aquifer, lateral flow is estimated by a non-linear equation and redistributed to the HRUs in the wetness classes that are saturated.

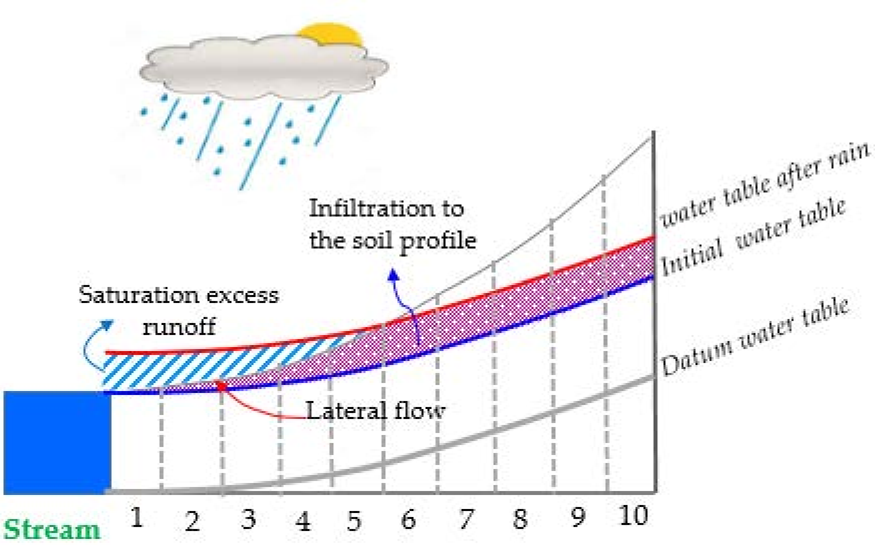


Figure 4.3. Illustration of behavior of water table in surface aquifer in SWAT-HS.

4.1.2.1.3. *Calculation of water fluxes and moisture content*

Similar to SWAT2012, SWAT-HS is run on a daily time step. Most changes in SWAT-HS are in the calculation of lateral fluxes (both interflow and surface runoff) when the soil is above field capacity and the surface aquifer storage is greater than zero. Evapotranspiration is also adjusted in correspondence to the amount of water available in the surface aquifer. SWAT-HS uses the original SWAT2012 equations for calculating vertical fluxes in the soil profile and base flow from the shallow aquifer.

At the daily time step, the various fluxes are calculated as per the following steps. Figure 4.4 provides an illustration of the sequence of surface aquifer and runoff calculations within a time step in SWAT-HS. In this illustration, it is assumed that the watershed is equally divided into 5 wetness classes with increasing soil water storage capacities.

***Step 1**: Checking the saturation status of wetness classes* (Figure 4.4a)

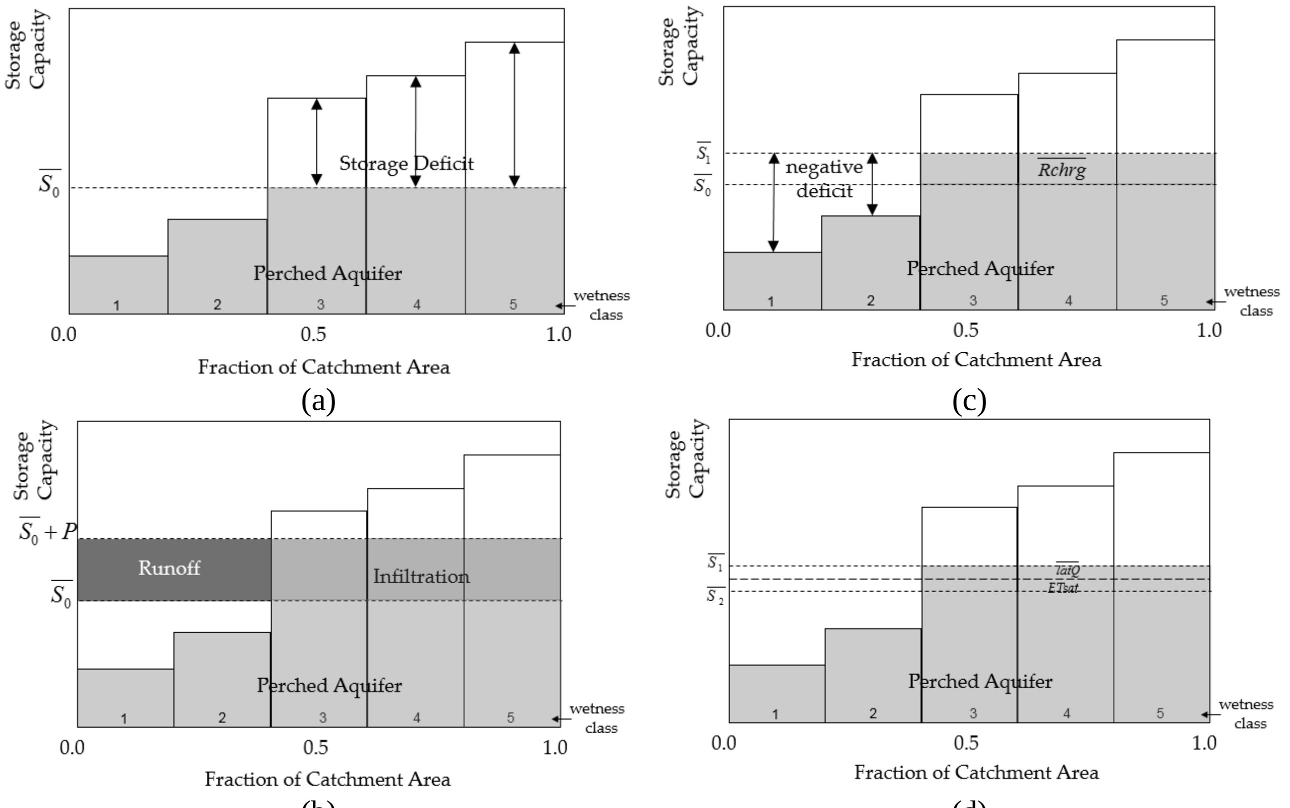
At the beginning of a time step the saturation status of each wetness class is determined by calculating the difference between the mean water depth of the surface aquifer ($\overline{S_0}$) and the soil water storage capacity (*edc*) of the wetness class (Figure 4.4a). Those wetness classes where $\overline{S_0}$ exceeds the storage capacity are saturated (classes 1 and 2) while the remaining classes 3 through 5 have a storage deficit representing the amount of water that is needed to reach saturation.

Step 2: Inputting water from rainfall/snowmelt **(**Figure 4.4b)

Following precipitation or snowmelt inputs, canopy interception is computed and infiltration-excess runoff is calculated by the Green-Ampt method. The remaining water input (*P*), which is precipitation and snowmelt excluding canopy interception and infiltration-excess runoff, increases the storage of the surface aquifer to $\overline{S_0} + P$.

Step 3: Calculating vertical water fluxes and saturated-excess runoff (Figure 4.4b and Figure 4.4c)

For HRU's in wetness classes where $\overline{S_0} + P$ is greater than the soil water storage capacity (*edc*), the excess water becomes saturation-excess runoff and the HRUs are saturated (Figure 4.4b, wetness classes 1 and 2). In the unsaturated areas, the water input becomes infiltration to the soil profile (wetness classes 3, 4, and 5). The infiltrated water first fills up the soil profile to field capacity and the remaining water drains out of the profile. The seepage out of the soil profile is then divided into recharge of the shallow aquifer and adding water to the surface aquifer. The latter raises the mean surface aquifer water depth ($\overline{S_1}$) and reduces wetness class storage deficits calculated as the difference between $\overline{S_1}$ and *edc*. Storage deficits are calculated for wetness classes. Saturated wetness classes have negative storage deficits while the unsaturated ones have positive values. The negative storage deficits will be used in distributing the return flow in the step 4.



(a) (c)

(b) (d)

Figure 4.4. Changes in water content in perched aquifer over a time step as simulated by SWAT-HS (a) maximum storage capacity of 5 wetness classes and the storage capacity $\overline{S_0}$ at the beginning of the time step, (b) following water input P from precipitation or snowmelt, (c) aquifer recharge ($\overline{Rchrg}$) raises the mean perched aquifer water depth to $\overline{S_1}$, and (d) lateral flow ($\overline{latQ}$) and plant water use ($ETsat$) are deducted from storage, resulting in perched aquifer water depth $\overline{S_2}$ at the end of the time step.

Step 4: Calculating lateral flow (Figure 4.4d)

Lateral flow is calculated in two sub-steps. First, the amount of lateral flow is calculated by treating the surface aquifer as a non-linear reservoir to generate lateral flow ($\overline{latQ}$) for the whole sub-basin:

$$\overline{latQ} = latA * \overline{S_1} \wedge latB \tag{4.3}$$

where $latA$ and $latB$ are constant coefficients. Next, lateral flow is assumed to only return to the surface to contribute to streamflow in saturated areas. $\overline{LatQ}$ is apportioned to saturated wetness classes in proportion to their corresponding negative saturation deficit. With this approach, SWAT-HS implicitly routes lateral flow across the hillslope from the upslope dry areas to the downslope wet areas.

Step 5: Calculating evapotranspiration

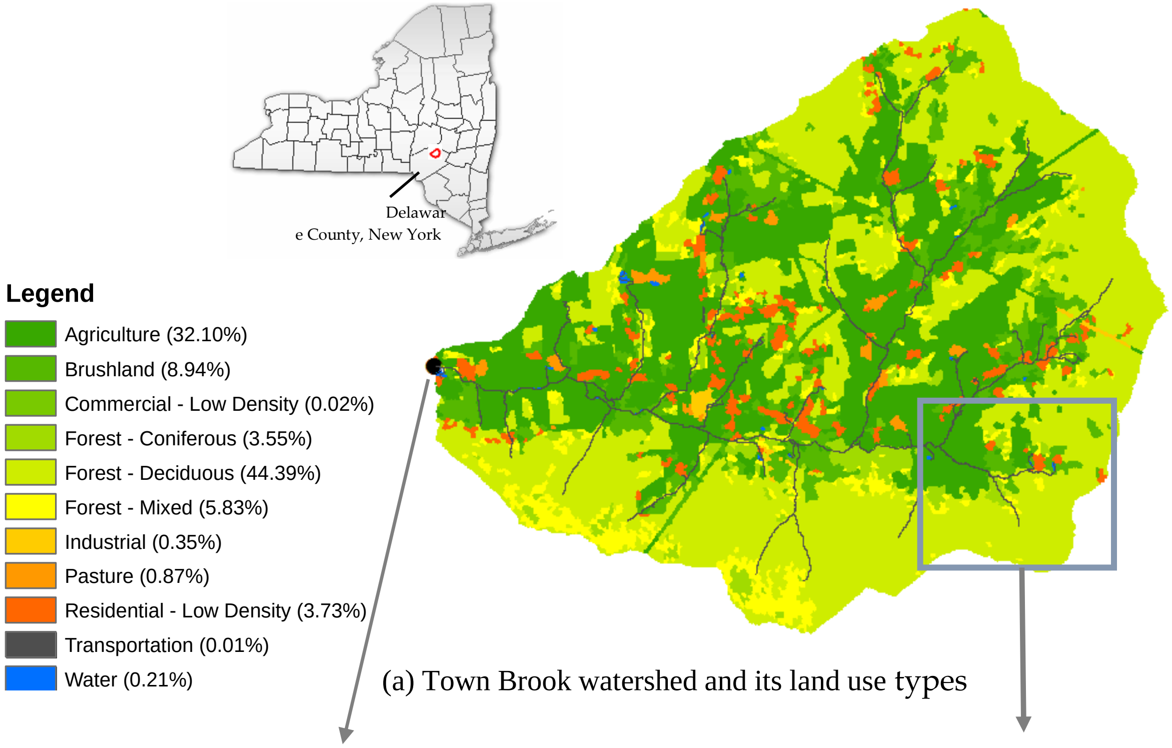
Water stored in the surface aquifer is made available to plants when the water table intersects the root zone. Plant water use is calculated by layers. For layers above the water table, plants use only non-drainable soil water. For saturated layers below the water table, evapotranspiration is extracted from the surface aquifer. For the layer in which the water table in the surface aquifer is located, the layer is split into two parts: (i) above the water table where plants use non-drainable water and (ii) below the water table where plants take the drainable water. At the end of each time step the sub-basin average lateral flow ($\overline{latQ}$) and plant water use (*ETsat*) are deducted from storage (Figure 4.4d) and the mean surface aquifer water depth is updated ($\overline{S_2}$) and becomes the initial water depth for the next time step.

Table 4.1. New parameters added to SWAT-HS

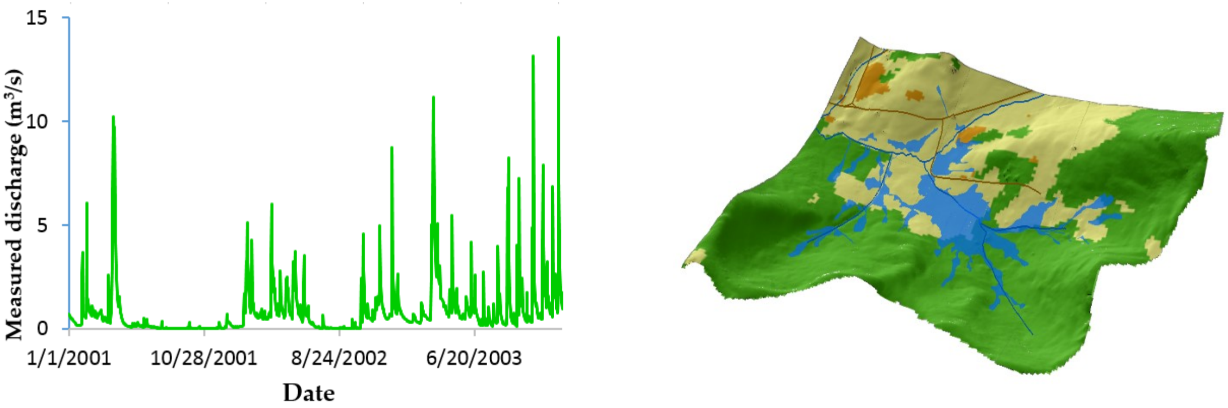
Name	Unit	Definition	Range
Smax	mm	Maximum soil storage capacity in the watershed	100-400
b		Shape parameter defining the distribution of soil storage capacity	0.1-3
effporfactor	-	Fraction of effective porosity that can hold water under saturated conditions	0 - 1
latA		Surface aquifer non-linear reservoir coefficient	0 - 1
latB		Surface aquifer non-linear reservoir exponent	1 - 3
perchst_datum	mm	Mean depth of surface aquifer drawn down just to point where lateral flow from aquifer ceases	
rchrg_paf	-	Fraction of root zone percolation that recharges the surface aquifer	0 - 1

4.1.2.2. Study area: Town Brook watershed, Delaware, New York

The 37 km^2 Town Brook watershed located in the Catskill region of New York State (Figure 4.5) was selected as the test case study for SWAT-HS because input data and observations to evaluate the model performance were sufficient. Observations of temporal streamflow and the spatial extent of saturated areas in 2006 (28-30 April) and 2007 (12 April, 7 June and 2 August) were available from previous studies (Schneiderman *et al.*, 2007; Dahlke *et al.*, 2009; Harpold *et al.*, 2010). The average annual temperature is 8^{o}C and average annual precipitation is 1123 mm/year. The landscape is characterized by steep to moderate hillslopes of glacial origin with shallow permeable soils. Elevation ranges from 493 to 989 m. The majority of the soils are silt loam and silty clay loam overlaying a glacial till. The upper terrain on the north-facing slope of the watershed is characterized by shallow soil (average thickness 80 cm) overlaying fractured bedrock, steep slopes (average slope 29%) and deciduous and mixed forests (60% of the watershed). The south-facing slopes and lowland areas of the watershed have deeper soils (average thickness 180 cm) underlain by a dense fragipan restricting layer, more gentle slopes (average slope 14%), and are predominantly occupied by pasture and row crops (20%) and shrub land (18%). Water and wetland comprise only 2% of the watershed while impervious surface is insignificant. The primary agricultural activity is dairy farming, with pastures, corn, and hay being grown.



(a) Town Brook watershed and its land use types



(b)Temporal observation of discharge at the outlet

(c) Spatial observation of saturated areas in 28-30/04/2006 (shown in blue)

Figure 4.5. Town Brook watershed, Delaware County, New York.

4.1.2.3. SWAT-HS setup for the Town Brook watershed

The study watershed was not split up into multiple sub-basins because of its small size. Instead, the watershed was divided into 124 multiple wetness class-based HRUs. The wetness class indicates the relative position of the HRU in the landscape in respect to surface runoff generation.

HRU definition

Wetness map

We classified the Town Brook watershed into 10 wetness classes, the maximum number that can currently be handled by SWAT-HS, by grouping HRUs with similar Topographic Indices (TI) values and created a wetness map. The TI was calculated using Equation 1 and the 10-meter DEM. The upslope contributing area, α, was determined by the Terrain Analysis Using Digital Elevation Models (TAUDEM) (Tarboton, 1997; Tarboton and Mohammed, 2010). Saturated conductivity and soil depth were extracted from Soil Survey Geographic (SSURGO) Database (USDA-NRCS, 2012).

Agnew *et al.* (2006) developed a relationship between TI and probability of saturation P_{sat} for the Town Brook watershed and suggested that the areas with TI >17.7 is always saturated. Based on this, we grouped the areas with $TI > 17.7$ as the “wettest” wetness class (wetness class 1) with no storage capacity (0 mm). This means that when there is hydrological input, wetness class 1 is always saturated. Wetness class 1 is actually fit to the perennial stream network and occupies a very small fraction of the watershed (0.59%), therefore, it is reasonable to assign no storage capacity to this wetness class. Since all Town Brook studies consisting of observations of saturated areas (Harpold *et al.*, 2010) and predictions by other watershed models (SMR (Agnew *et al.*, 2006), SWAT-VSA (Easton *et al.*, 2008) and SWAT-WB (White *et al.*, 2011) found that saturated areas never exceeded 50% of watershed, we grouped 50% of the watershed with lowest TI values as the “driest” wetness class (wetness 10). Assigning the driest wetness class to half of the watershed allowed us to classify the remaining areas, which are more prone to saturation, into a greater number of wetness classes. We divided the remaining areas into 8 wetness classes (wetness class 2 – 9) with equal areas based on TI values. Wetness classes 2 to 10 have increasing soil water storage capacities that follow a Pareto distribution (Equation 2). The resulting wetness map is shown in Figure 4.6. Table 4.2 shows the final classification of 10 wetness classes in Town Brook and their soil storage capacities based on the optimal parameter set resulting from our calibration. This classification scheme was used in the evaluation of SWAT-HS performance in section 3.1.

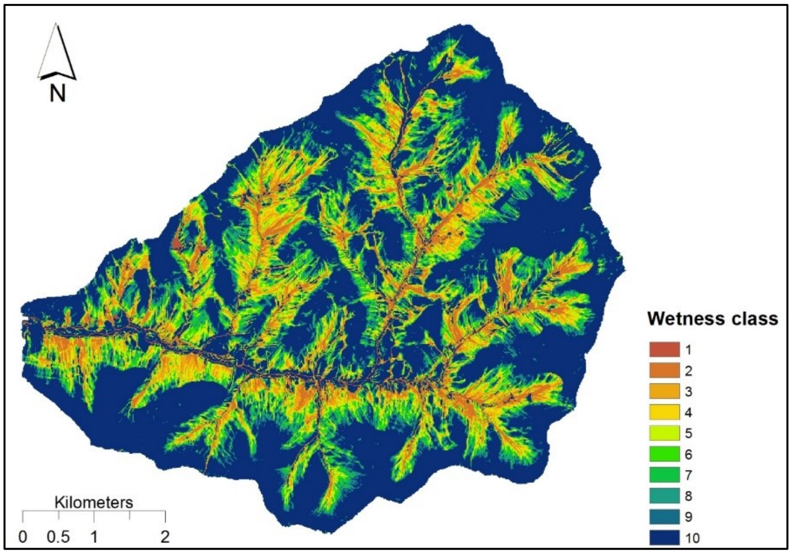


Figure 4.6. Wetness map for the Town Brook watershed.

Table 4.2. Classification of wetness classes for the Town Brook watershed.

Wetness class	TI	% of watershed area	Soil storage capacity (mm)	Characteristics
1	> 17.7	0.59	0	Wet
2	16.73 – 17.70	6.06	5.2	
3	15.77 – 16.73	6.24	14.2	↓
4	14.80 – 15.77	6.18	23.5	
5	13.84 – 14.80	6.37	33.2	
6	12.87 – 13.84	6.20	43.2	
7	11.91 – 12.87	6.90	54.0	
8	10.94 – 11.91	6.67	65.6	
9	9.98 – 10.94	6.04	77.0	Dry
10	< 9.98	48.75	133.6	

Soils

The soil map of Town Brook watershed was extracted from the SSURGO database (USDA-NRCS, 2012). There are 17 defined soil types present in the watershed. The data on soil properties that are necessary for the SWAT model were taken from the SSURGO database and the soil survey of Delaware County (USDA, 2006).

The soil map was then overlaid with the wetness map to create a new soil map in which the same soil types in different wetness classes have different soil names but retain the same soil characteristics. The new soil name reflects both wetness class and soil type. This is an essential step for running SWAT-HS, which assigns wetness class to each HRU in the final discretization.

Land use

The land use map was derived from classified 2009 aerial photography data obtained from New York City Department of Environmental Protection (NYCDEP). Town Brook watershed includes 11 land use types (Figure 4.5). The dominant land use types are forest (54%) and agriculture (32%). Residential area covers a very small part of the watershed (4%).

For simplification purposes, we assumed that slope does not have an impact on HRU discretization. With 10 wetness classes, 17 types of soil and 11 types of land use, 145 HRUs were created for the Town Brook watershed by overlaying the land use map and the redefined soil map. Each HRU is then defined by a unique combination of wetness class, soil type, and land use.

Meteorological inputs

Daily precipitation and temperature data available in a 4km x 4km gridded format, developed using the Parameter-Elevation Relationships on Independent Slopes Model (PRISM) interpolation method (Daly *et al.*, 2008), was accessed through the Applied Climate Information System (http://www.rccacis.org/). Precipitation and temperature data were assumed to be taken from the weather station near the centroid of the watershed. The data at the centroid were interpolated from eight surrounding PRISM grid points by the inverse distance weighting method.

Solar radiation was determined by averaging data from the Albany and Binghamton airports. Relative humidity and wind speed were generated in SWAT.

4.1.2.4. Model calibration/validation and evaluation of model performances

Streamflow calibration was carried out using observed records of streamflow obtained from the USGS gauging station (#01421618) at the outlet of the Town Brook watershed from 1998 to 2007, of which the first three years served as warming up period. The validation period was from 2008 to 2012. In SWAT-HS, three groups of parameters were chosen for calibration (Table 4.3). Group (i) includes 5 parameters related to snowmelt simulation. This group is important to account for in Northeastern US watersheds like Town Brook where snowmelt is significant. Group (ii) related to flow simulation includes 4 new parameters of SWAT-HS (*RCHRG_PAF, latA, latB, EFFPORFACTOR),* controlling the generation of saturation-excess runoff and lateral flow; *SURLAG*, affecting the lag time of surface runoff routing; *ALPHA_BF*, related to groundwater flow; and 2 parameters associated with evapotranspiration (*EPCO*, *ESCO*). Group (iii) including *Smax* and *b* defines the statistical distribution of soil water storage capacity for the watershed.

The SWAT-HS model was calibrated by applying the Monte Carlo sampling method. Ten thousand parameter sets for both snowmelt and streamflow calibration steps were generated, each of which was then run with SWAT-HS. To evaluate the performance of a parameter set, we used the Nash-Sutcliffe efficiency (NSE), a measure of the goodness of fit between simulated streamflow and observations.

Table 4.3. Parameters for streamflow calibration using Monte Carlo sampling method.

Name	**Unit**	**Definition**	**Range**	**Calibrated value**
Group (i): Snowmelt calibration				
SFTMP	ºC	Snowfall temperature	-5 - 5	-0.58
SMTMP	ºC	Snowmelt temperature	-5 - 5	1.10
SMFMX	mm/ ºC	Maximum snowmelt factor	5 - 10	7.62
SMFMN	mm/ ºC	Minimum snowmelt factor	0 - 5	2.68
TIMP	-	Snow pack temperature lag factor	0 – 1	0.022
Group (ii): Flow calibration				
RCHRG_PAF	mm	Fraction of root zone percolation that recharges the surface aquifer	0-1000	0.822
latA		Surface aquifer non-linear reservoir coefficient	0 - 1	0.013
latB		Surface aquifer non-linear reservoir coefficient	1 - 3	1.626
ALPHA_BF	days^{-1}	Base flow recession constant	0 - 1	0.05
EFFPORFACTOR		Fraction of effective porosity that can hold water under saturated conditions	0 - 1	0.877
EPCO		Plant water uptake compensation factor	0 - 1	0.989
ESCO		Soil evaporation compensation factor	0 - 1	0.691
Group (iii): Storage capacity of wetness classes (5 wetness classes were defined in this case study)				
Smax	mm	Maximum soil water storage capacity in the watershed	100-400	227
b		Shape parameter defining the distribution of soil water storage capacity	0.1-3	1.59

The specific steps of the calibration procedure were:

1. Prepare 10,000 parameter sets of group (i) and 10,000 parameter sets of the combination of group (ii) and group (iii) within the default ranges of their corresponding parameters. The 10,000 parameter sets were randomly generated by the Monte Carlo sampling method assuming that all parameters are uniformly distributed.
2. Assign default values for soil water storage capacity in each wetness class: *Smax* at 200mm and *b* at 1; soil water storage capacity, *edc*, was calculated for each wetness classes 2 – 10 and set at 0 for wetness class 1.
3. Calibrate snowmelt: 10,000 simulations corresponding to the 10,000 parameter sets of group (i) were run with SWAT-HS. It may be noted that in this step, other remaining parameters were assigned default values. The best parameter set with the highest NSE was used in step 4.
4. Calibrate streamflow: While snowmelt parameters values were set at best parameter set from step 3, 10,000 simulations corresponding to the 10,000 parameter sets of group (ii) and (iii) were run with SWAT-HS. The parameter sets which result in Nash Sutcliffe Efficiency (NSE) above 0.65 were chosen as good performing parameter sets.

To evaluate SWAT-HS performance for flow modeling, we compared the simulated results with both temporal and spatial observations. Temporally, simulated streamflow at the watershed outlet was compared with measurements. Spatially, the prediction of saturated areas was compared with available observations on 6 observed days: 28, 29, 30 April 2006 (Dahlke *et al.*, 2009), 12 April 2007, 7 June 2007, and 2 August 2007 (Harpold *et al.*, 2010). Among all the parameter sets with good performance from step 4, the parameter set that gave both best fit to observations of streamflow and saturated areas in all six observed days was chosen as the optimal set. This set was used in all evaluations of SWAT-HS performance. Four statistical criteria: Nash Sutcliffe efficiency (*NSE*), Percent bias (*PBIAS*), RMSE – observations standard deviation ratio (*RSR*), and Kling Gupta efficiency coefficient (*KGE*) were used for model evaluation.

✓ *Nash Sutcliffe efficiency (NSE)* (Nash and Sutcliffe, 1970): ranges between -∞ and 1.0, with NSE = 1 being the optimal value.

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs} - Y^{mean})^2} \right] \tag{4.4}$$

where Y_i^{obs} is the i^{th} observation, Y_i^{sim} is i^{th} simulated value, Y_i^{sim} is the mean of observed data and n is the total number of observations.

✓ *Percent bias (PBIAS)*: the optimal value of PBIAS is 0.0, with low magnitude values indicating accurate model simulation. Positive values specify underestimation bias and negative values indicate overestimation bias (Gupta *et al.*, 1999).

$$PBIAS = \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim}) * 100}{\sum_{i=1}^{n} Y_i^{obs}} \right] \tag{4.5}$$

✓ *RMSE – observations standard deviation ratio (RSR)*: RSR standardizes RMSE using the observations standard deviation (Singh *et al.*, 2004). It varies from the optimal value of 0, which indicates zero RMSE to +∞. The lower RSR, the lower RMSE, the better the model performance.

$$RSR = \frac{RMSE}{STDEV_{obs}} = \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs} - Y^{mean})^2} \right] \tag{4.6}$$

✓ *KGE - Kling–Gupta efficiency*: KGE was introduced by Gupta *et al.* (2009). It can be decomposed into a correlation term, a bias term and a variability term. KGE is evaluated to have advantages over NSE. While NSE is found to be sensitive to large runoff values, *KGE* overcomes this problem by giving equal weights for three component measures: correlation, bias and variability. A value of 1 indicates a perfect fit between model predictions and observations.

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\alpha-1)^2} \tag{4.7}$$

$$\beta = \frac{\mu_s}{\mu_0} \tag{4.8}$$

$$\alpha = \frac{CV_s}{CV_0} \tag{4.9}$$

where r is the correlation coefficient, μ is the average, CV is the coefficient of variation, and the subscripts s and o represent simulations and observations.

Moreover, to clarify the effect of modifications made in SWAT-HS, we also compared with SWAT2012. A setup was built with SWAT2012 in the same case study and was calibrated by SWAT-CUP (SWAT Calibration and Uncertainty Programs) using the Parameter Solution (Parasol) algorithm (van Griensven and Meixner, 2006). The comparison of the two models focused on three aspects: the fit of simulated streamflow at the watershed outlet to observations, the spatial distribution of saturated areas and the contributions of different flow components.

Because of the availability of numerous parameter sets that are able to give equally good SWAT-HS performance in step 4, we used the Generalized Likelihood Uncertainty Estimation (GLUE) approach (Beven and Binley, 1992) to estimate the uncertainty of streamflow and saturated area predictions resulting from parameter uncertainty. Following the GLUE approach, a threshold criterion of NSE at 0.65 was set to filter the 10,000 Monte Carlo simulations into behavioral models ($NSE \geq 0.65$) and non-behavioral models ($NSE < 0.65$). The criterion for NSE was set at 0.65 because it is the criterion for “good” model performance according to Moriasi *et al.* (2007). The subset of behavioral models was used for uncertainty estimation while the subset of non-behavioral models was rejected.

4.1.3. Results and discussion

4.1.3.1. SWAT-HS and SWAT2012 performances

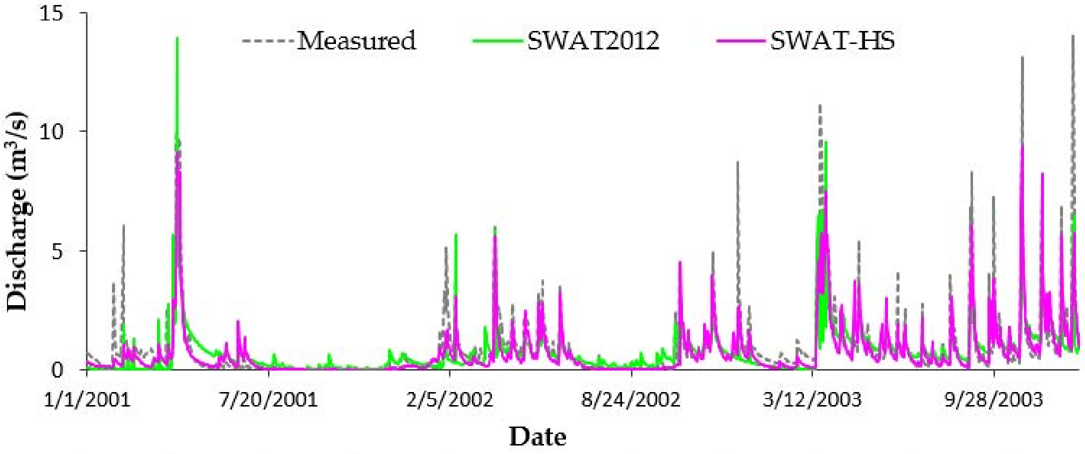
In this section, the performances of SWAT-HS and SWAT2012 are evaluated. We focus on three aspects: comparison of simulated and observed discharge at the outlet, evaluation of simulated spatial distribution of saturated areas, and relative contributions of the different flow components to streamflow.

4.1.3.1.1. Evaluation of discharge at the watershed outlet

The calibrated SWAT-HS and SWAT2012 model simulated discharges were compared with observations at daily and monthly time steps using their optimal parameter sets (Figure 4.7). The calibrated parameter set for SWAT-HS is shown in Table 4.3. The SWAT-HS model captured the observed variation of streamflow in both summer (May to October) and winter (November to April) periods slightly more accurately than SWAT2012. However, SWAT2012 was able to capture the winter peaks more accurately compared to SWAT-HS. Overall, statistical evaluation of both models for observed streamflow at the outlet showed SWAT-HS performing better or comparable to SWAT2012 (Table 4.4). With SWAT-HS, the *NSE* values for daily discharge were 0.68 and 0.52, *PBIAS* values were 7.17 and -3.61, and *RSR* values were 0.57 and 0.69 during the calibration and validation periods, respectively. The *NSE*, *PBIAS*, and *RSR* for daily discharge predictions by SWAT-HS are all within the range of "good" model performance for the calibration period while *NSE* is rated as "satisfactory", *PBIAS* as "very good" and *RSR* as "satisfactory" performance for the validation period, based on the guidelines for evaluation of model performance by Moriasi *et al.* (2007). The *NSE* improved from 0.52 to 0.6 in the validation period if we excluded the year 2011 when extreme events (Hurricane Irene and Tropical Storm Lee) occurred in August 2011 that were not simulated very well by SWAT-HS. We also calculated KGE values (Gupta *et al.*, 2009), which is less sensitive to high streamflow values compared to NSE and gives a better measure of model fit to data at all ranges of streamflow (a value of 1 indicates a perfect fit). Daily KGE values with SWAT-HS were 0.75 and 0.65, respectively during the calibration and validation periods. The SWAT2012 prediction for daily streamflow resulted in *NSE* of 0.52 and 0.37, *PBIAS* of 4.25 and -6.27, *RSR* of 0.65 and 0.75, and *KGE* of 0.65 and 0.60 during the calibration and validation periods, respectively. The *PBIAS* of SWAT2012 is rated as "good" while the *NSE* and *RSR* are within the range of "satisfactory" performance during the calibration period, and are "unsatisfactory" in the validation periods. The values of statistical metrics for both models improved significantly for the monthly predictions (Table 4.4).

Table 4.4: Statistical evaluation for SWAT-HS and SWAT2012 in daily and monthly time steps.

Period	**Time steps**	**SWAT-HS**						**SWAT2012**					
		NSE	*PBIAS*	*RSR*	*KGE*	*Average streamflow (mm/a)*		*NSE*	*PBIAS*	*RSR*	*KGE*	*Average streamflow (mm/a)*	
						Simulated	*Observed*					*Simulated*	*Observed*
Calibration (2001-2007)	Daily	0.68	7.17	0.57	0.75	302.76	326.14	0.52	4.25	0.65	0.65	312.30	326.14
	Monthly	0.87	7.25	0.36	0.87			0.72	4.38	0.53	0.76		
Validation (2008 – 2012)	Daily	0.52	-3.61	0.69	0.65	331.61	320.05	0.37	-6.27	0.75	0.60	340.13	320.05
	Monthly	0.78	-3.38	0.46	0.86			0.62	-6.11	0.61	0.73		
Validation (excluding 2011)	Daily	0.60	0.73	0.63	0.71	232.47	234.18	0.46	-2.71	0.70	0.69	240.54	234.18
	Monthly	0.82	1.03	0.43	0.81			0.65	-2.53	0.59	0.68		



(a) Daily

(b) Monthly

Figure 4.7. Comparison of simulated daily and monthly discharge values between SWAT-HS, SWAT2012, and measured data.

Figure 4.8 presents flow duration curves from daily streamflows simulated by SWAT2012 and SWAT-HS versus observed flows. It again shows the closer fit of simulated streamflow from SWAT-HS to observed flow compared to SWAT2012. SWAT-HS slightly underestimated but showed better predictions of low flows while SWAT2012 overestimated low flows. During high flow periods, SWAT-HS gave a better fit to observed flow than SWAT2012 although it underestimated extreme high peak flows as shown in Figure 4.7a. The six dates with observations of saturated areas, corresponding streamflow values, and their percentages of exceedance are included in Figure 4.8

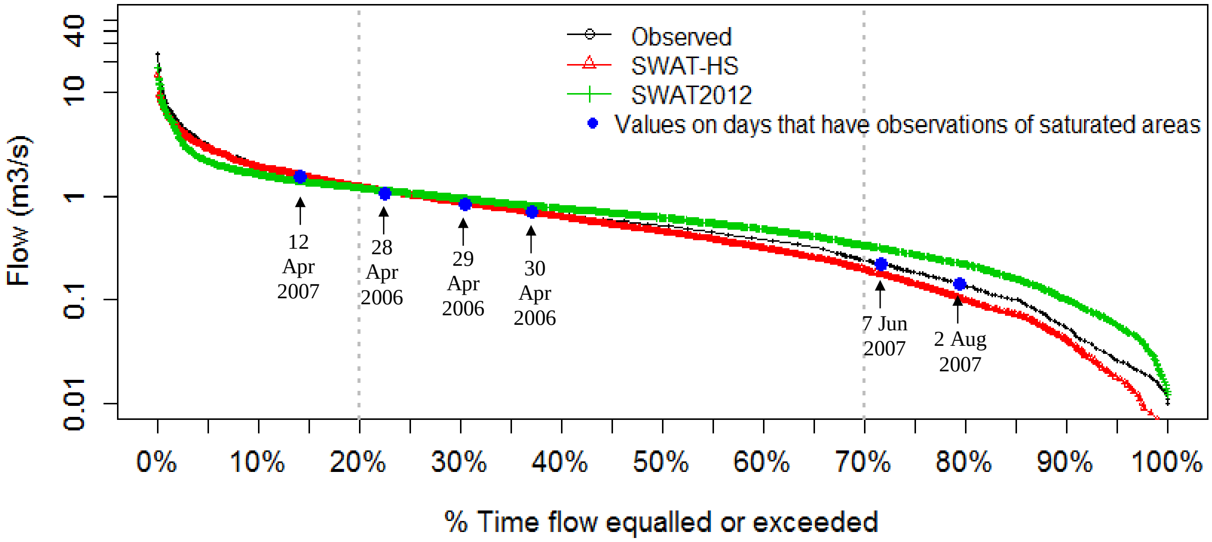


Figure 4.8. Flow duration curves of simulated flow by SWAT-HS and SWAT2012 versus observed flow.

4.1.3.1.2. Evaluation of Spatial Distribution of surface runoff areas

While SWAT-HS was slightly better than SWAT2012 in simulating streamflow at the outlet as per model evaluation statistics, one could argue that these differences could be due to calibration procedures used in the two models. Evaluation of the spatial distribution of runoff-generating areas by both models showed that SWAT-HS gave a better representation of actual locations of runoff generation. The spatial distributions of annual surface runoff during the calibration period generated by the SWAT-HS and SWAT2012 models are shown in Figure 4.9a and Figure 4.9b depict the wetness and land use map of the Town Brook watershed, while Figure 4.9c and Figure 4.9d show differences in the spatial locations of surface runoff generation between SWAT-HS and SWAT2012. The distribution of annual surface runoff predicted by SWAT-HS follows the wetness map. Runoff can only be found in areas corresponding to “wetter” wetness classes with low storage capacity. On the other hand, the distribution of surface runoff predicted by SWAT2012 follows the distribution of land use in which forests contribute the least runoff and agriculture is the most significant contributor. While residential areas have high amounts of surface runoff, they cover a very small area in the watershed.

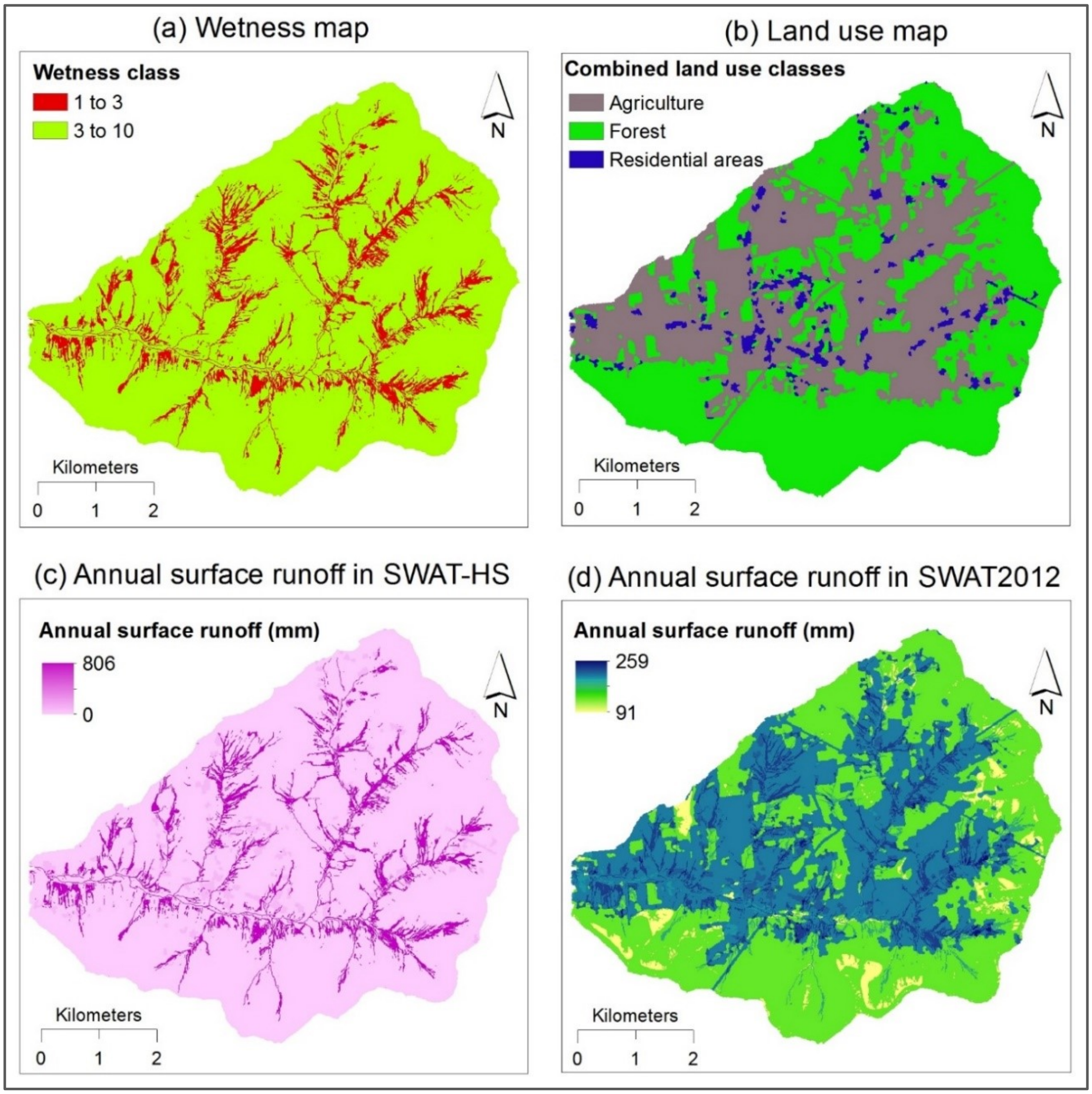


Figure 4.9. Spatial distribution of annual surface runoff simulated by SWAT-HS and SWAT2012.

The spatial distribution of simulated saturated areas by SWAT-HS was compared with observations in a headwater drainage area of Town Brook (Figure 4.5) on 6 observed days: 3 consecutive days from 28-30 April 2006 (Figure 4.10a), 12 April 2007 (Figure 4.10b), 7 June 2007 (Figure 4.10c), and 2 August 2007 (Figure 4.10d). SWAT-HS gave a reasonable fit to observed saturated areas on 28-30 April 2006 and 12 April 2007, but overestimated the saturated areas on June and August 2007. Although in all observed events, the extent of saturated areas simulated by SWAT-HS did not perfectly match observations, the locations where saturation occurred were relatively well predicted. It should be noted here that the extent of saturated areas that were observed in the field using a global positioning system (GPS) unit may have uncertainty and may not have covered the entire extent of saturated areas very accurately. Therefore, although the assessment of the extent of saturated area prediction is still limited, SWAT-HS is useful in predicting areas where saturation-excess runoff occurs, which can also be hot-spots for water quality issues.

Out of the six days with observations, five days including 28-30 April 2006, 7 June 2007 and 2 August 2007 were days with no rainfall while there was small amount of rainfall on 12 April 2007 (Figure 4.10). In these 5 dry days, SWAT2012 predicted no surface runoff (therefore, we did not show the runoff map) due to the use of the curve number concept. In contrast, SWAT-HS predicted the existence of saturated areas, which is comparable with field observations.

We chose the 3-day observation period (28-30 April 2006) to look into details on how the hydrological approaches of the two models differed in their predictions of saturated areas. The relative contribution of different flow components were compared between the two models for April 2006, with the three observed days indicated (Figure 4.11). During the period of 28-30 April 2006, SWAT-HS predicted lateral flow and groundwater contributions and no surface runoff contribution to streamflow (Figure 4.11a). On April 23-24, there was a high rainfall event (60 mm) that caused high lateral flow and high water storage in the surface aquifer. Therefore, although there was no rainfall on April 28-30, lateral flow stored in the surface aquifer was routed from upslope dry areas to downslope wet areas where it surfaced as runoff and created saturated areas. In contrast, SWAT2012 predicted a dominant groundwater contribution, a small lateral flow contribution, and no surface runoff contribution to streamflow during these days (Figure 4.11). Based on the approach of SWAT2012, both lateral flow and groundwater flow cannot reach the ground surface to form saturated areas.

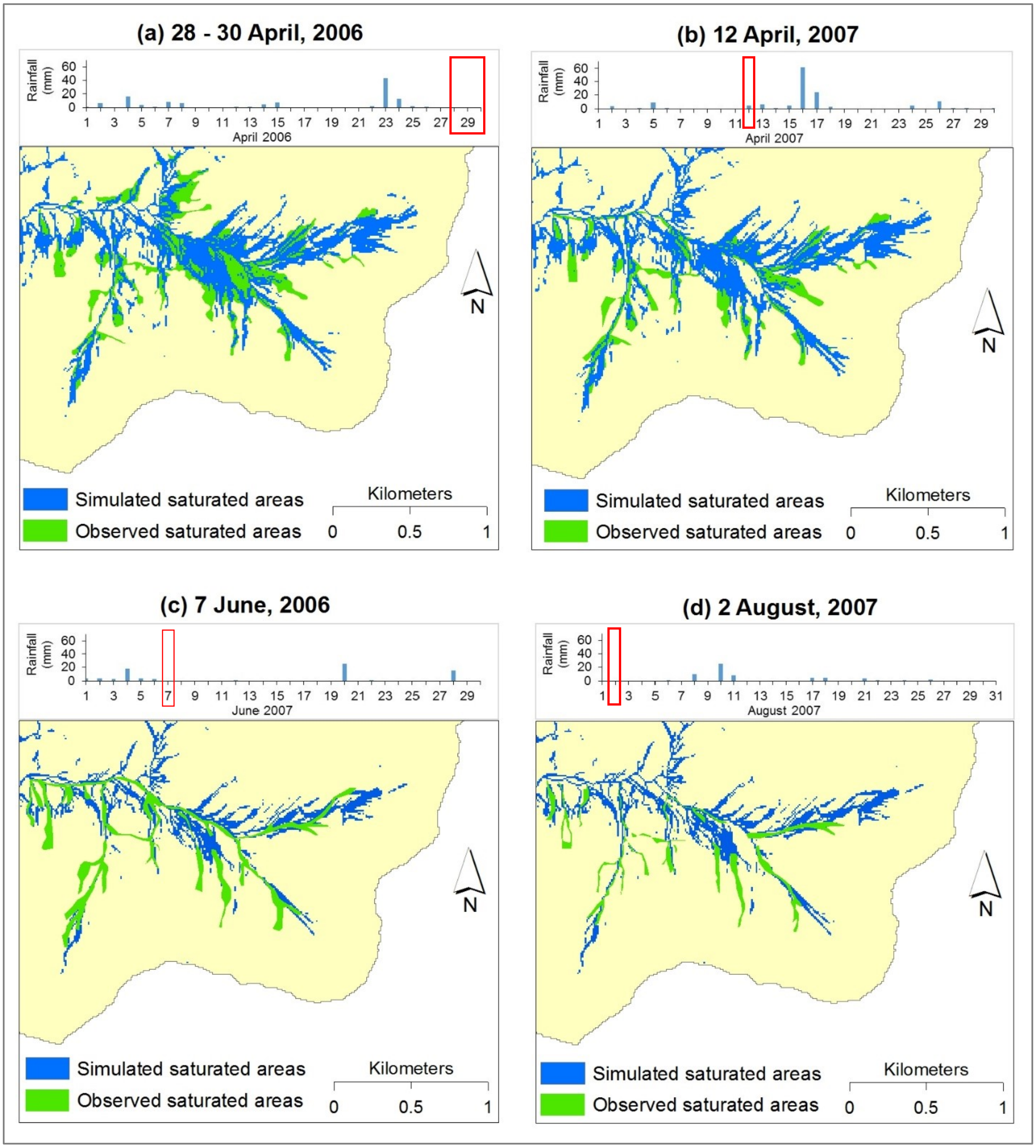
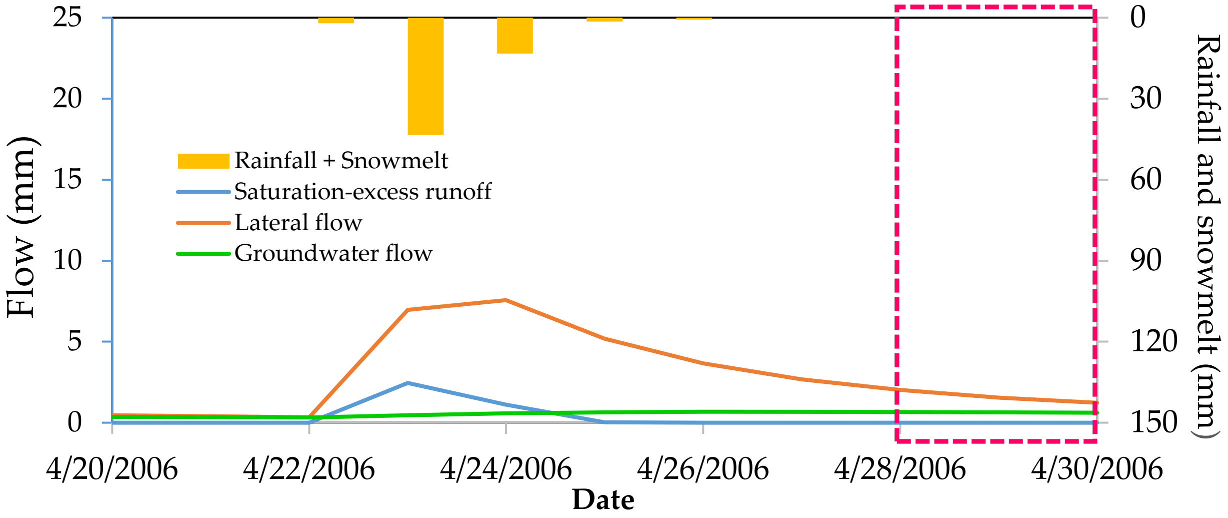
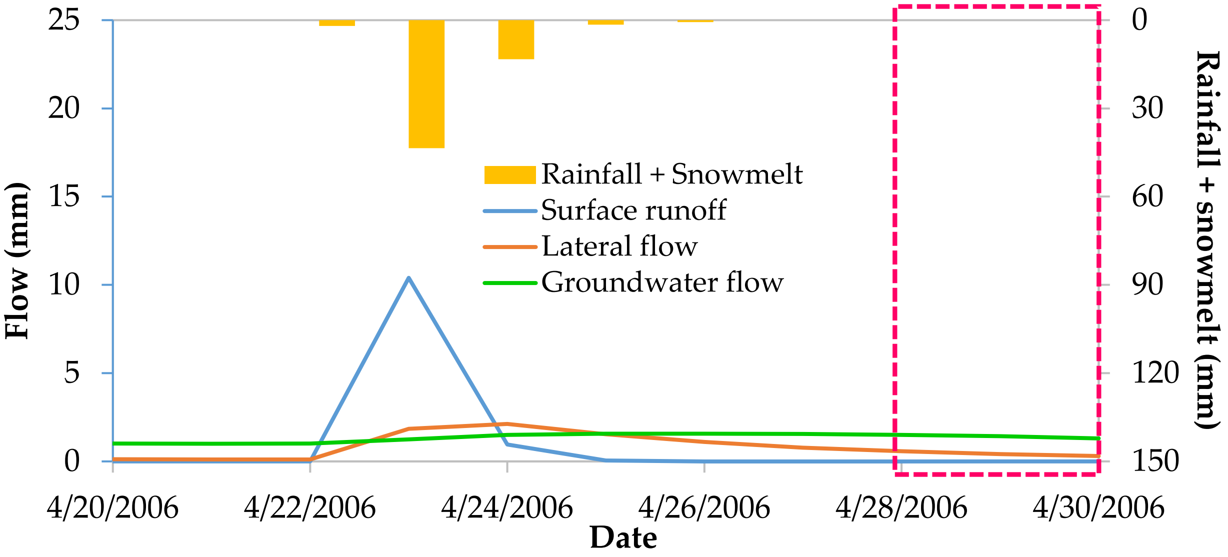


Figure 4.10. Saturated areas simulated by SWAT-HS compared with observations on (a) 28-30 April 2006, (b) 12 April 2007, (c) 7 June 2006 and (d) 2 August 2007. Rainfall variation in the observed months were presented with the red box highlighting the rainfall amount of the observed events.



(a) SWAT-HS



(b) SWAT2012

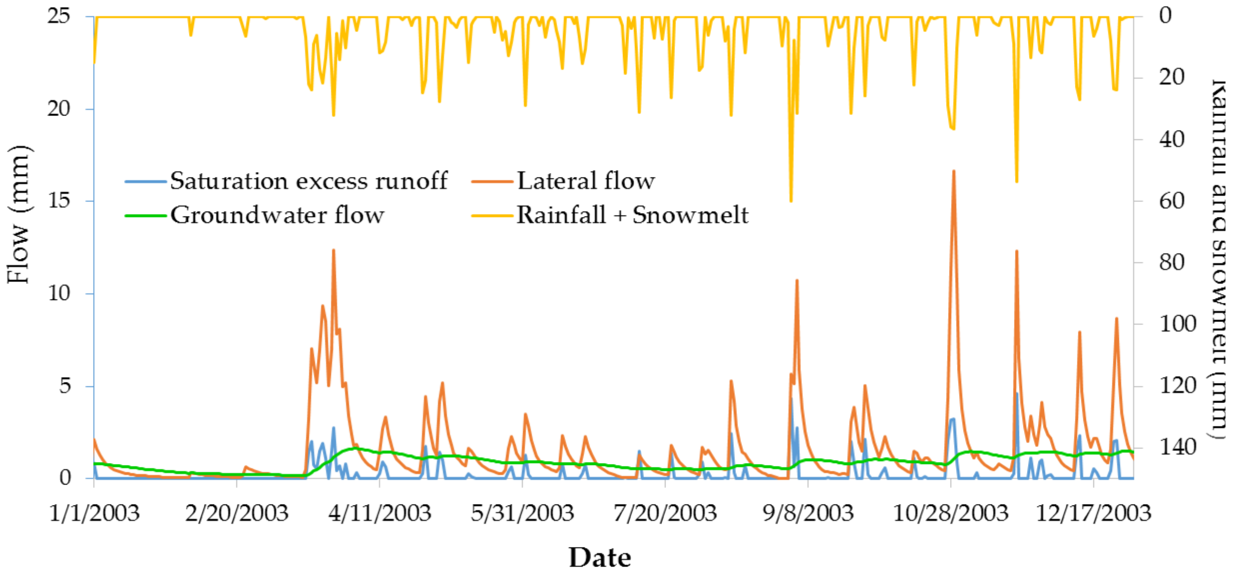
Figure 4.11. Flow components simulated by SWAT-HS and SWAT2012 in April 2006

(The red rectangle focuses the three observed days April 28 – 30, 2006).

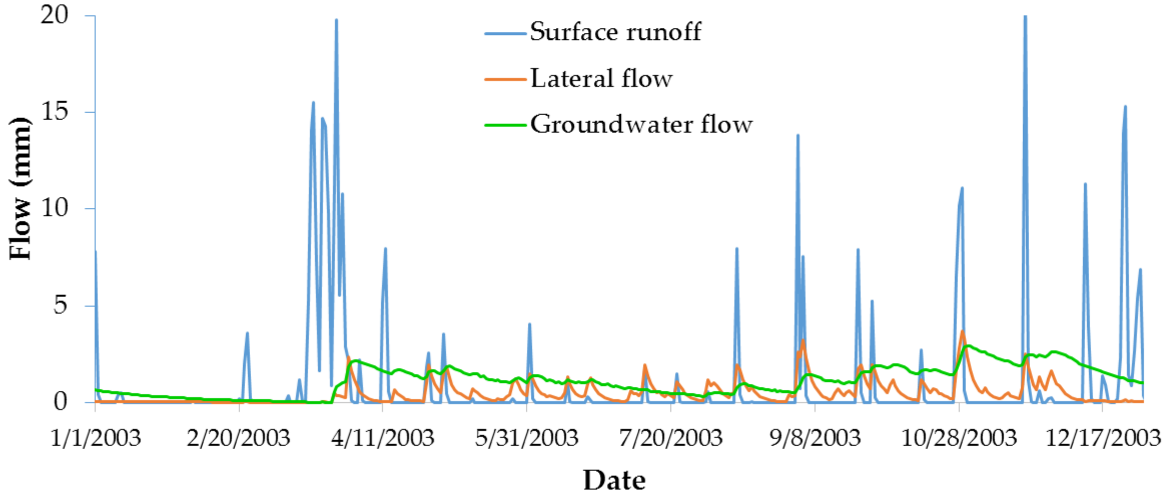
4.1.3.1.3. Simulated flow components

While the discharge values at the outlet did not differ greatly between the two models using their optimized parameters' values, the origin of flow varied greatly as exemplified in 2003 (Figure 4.12). For large rainfall events, SWAT2012 predicted that the main component of discharge at the outlet was surface runoff (Figure 4.12b), while SWAT-HS simulated the largest contribution to streamflow from lateral flow originating in the saturated areas. For the Town Brook watershed, SWAT-HS predicted all surface runoff coming from the saturation-excess process and the infiltration-excess runoff was estimated at 0; this is reasonable for the Catskill region where infiltration rates often exceed rainfall intensity. Lateral flow derived from the surface aquifer in SWAT-HS occurred throughout the year except during the period of snowfall (Jan – Feb). SWAT-HS estimated stable groundwater flow that contributed to streamflow throughout the year, but with small contributions at the beginning of the year when there was snowpack accumulation. In contrast to SWAT-HS, surface runoff in SWAT2012 is implicitly considered as infiltration-excess runoff and occurs with high rainfall except during the snowfall period (Jan – Feb). Lateral flow in SWAT2012, derived from soil moisture in excess of field capacity, occurred throughout the year whenever there was rainfall. Groundwater flow had the same trend in the two models with slightly more fluctuations in the SWAT2012 model because of the difference in calibrated values for the groundwater-related parameters.

Because of the difference in hydrological approach, SWAT-HS and SWAT2012 predicted dissimilar distributions of flow components. It should be noted here that no observations are available to confirm which model is correct and which type of flow is dominant. However, the results from SWAT-HS are in agreement with the conclusion of Walter *et al.* (2003) who indicated a low frequency of Hortonian flow and identified saturation-excess runoff as the primary surface runoff process for generating overland flow in the Catskill region of New York state. The dominant contribution of lateral flow simulated from SWAT-HS may be compatible with the finding of Harpold *et al.* (2010) in an intensive field survey in a 2.5 km^2 headwater watershed of Town Brook watershed where hillside lateral preferential flow paths rapidly transported water to near-stream saturated areas during runoff events under relatively dry antecedent conditions. Harpold *et al.* (2010) also suggested that the lateral redistribution of water from hillside areas reduces the influence of surface topography and channel topology on the sources of stream runoff.



SWAT-HS



(b) SWAT2012

Figure 4.12. Time series of flow components simulated by (a) SWAT-HS compared with (b) SWAT2012.

4.1.3.2. Uncertainty of the predictions of streamflow and saturated areas using SWAT-HS

Similar to other hydrological models, SWAT-HS can use numerous parameter sets that can give equally good fit to the observed data. This is known as the equifinality problem which is due to non-linearity and interdependence of parameters (Bárdossy and Singh, 2008) and results in uncertainty of model predictions. In this section, we estimated the uncertainty of the SWAT-HS predictions of streamflow and saturated areas using the GLUE approach (Beven and Binley, 1992).

Figure 4.13 presents scatter plots of parameter values versus NSE resulting from Monte Carlo simulations. It can be observed that within a value range, each parameter with the same value can result in significantly different NSE values when it is combined with different values of other parameters. Moreover, there are numerous combinations of parameter values that vary in wide ranges that can provide behavioral models (NSE ≥ 0.65) for streamflow. From our 10,000 Monte Carlo parameter sets, the number of behavioral models for streamflow was 995. We compared the performance of 995 behavioral models for their fit to field observations of saturated areas on six observed days. Finally, only 53 parameter sets remained that gave good predictions for both streamflow and saturated areas.

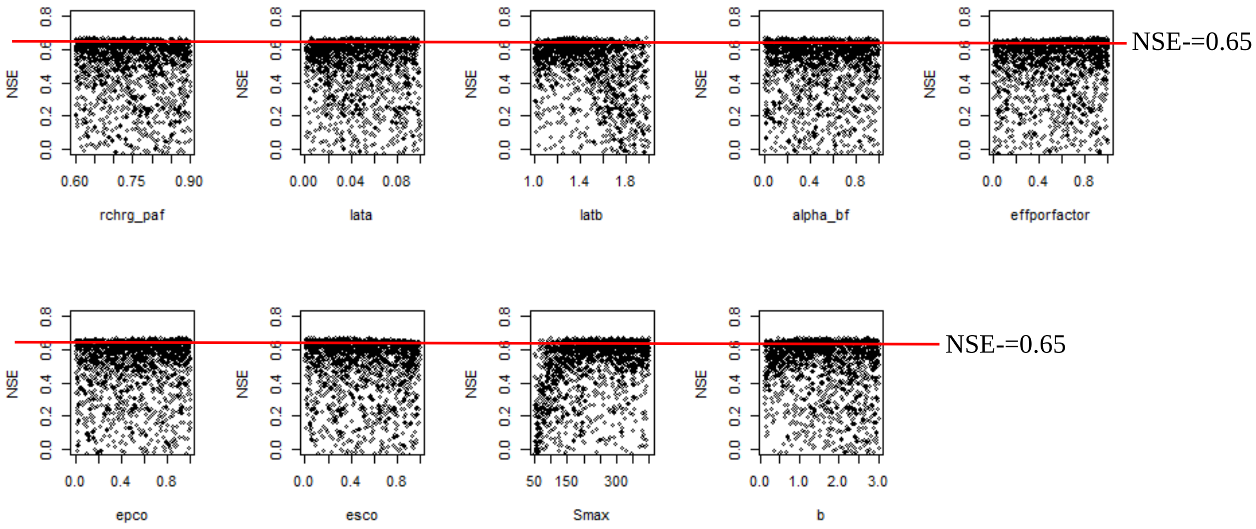


Figure 4.13. Scatter plot of parameter values versus NSE resulting from Monte Carlo simulations

Figure 4.14 shows the distributions of parameter values that gave good predictions for streamflow (in light green, 995 values) and parameter values that gave good predictions for both streamflow and saturated areas (in green, 53 values). It can be seen that for all parameters, the values with good performance are broadly varied in their value ranges, which means that the effect of one parameter can be compensated for other parameters, thus, different combinations of parameter values are able to give similar results. Although the uncertainty ranges of all parameters are still wide and same as the ranges that are used to generate Monte Carlo parameter sets, the distributions of values with good performance for most parameters differ from the uniform distribution which identifies where, within the parameter ranges, a higher probability of good results can be found. If only streamflow is considered, among 995 good models, more good/behavioral models are found when *rchrg_paf* is within 0.7 – 0.75, *latb* from 1.2-1.5, *effporfactor* closer to 1, *epco* is close to 1, and *esco* is close to 0. Soil water storage capacity functions built with *Smax* larger than 200 and *b* larger than 1 have a higher probability of giving good predictions. *Lata* and *alpha_bf* remain uniform in distribution, which means that these two parameters are not as sensitive as other parameters. If both streamflow and six predictions of saturated areas are considered, only 53 parameter sets are qualified for good performance which is a relatively small number to discuss about the parameter distribution. However, what we found here is that including prediction of saturated areas in choosing parameter sets with good performance did help to reduce a significant number of good sets that are based on streamflow prediction alone. We also realized that even though we could restrict good parameter sets to a small number, only two parameters, *Smax* and *b*, had narrow uncertainty ranges (*Smax* values less than 150 and *b* less than 1 are removed) while the uncertainty ranges of other parameters remain wide. *Smax* and *b*, as expected, are main controlling parameters of saturated area prediction, although their effects on predictions are still strongly controlled by other parameters.

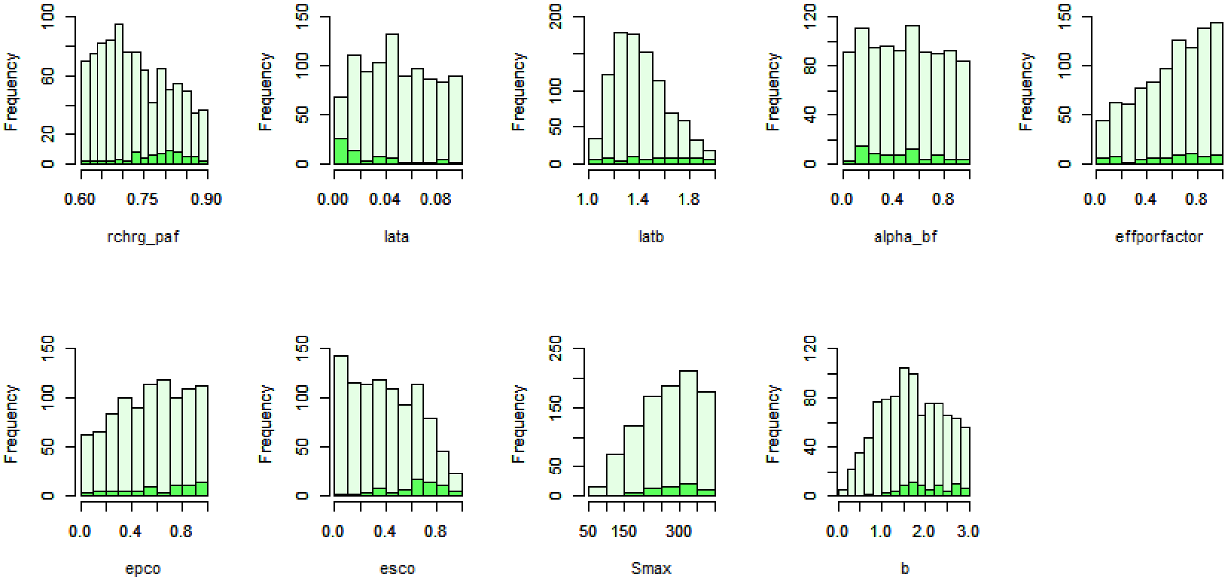
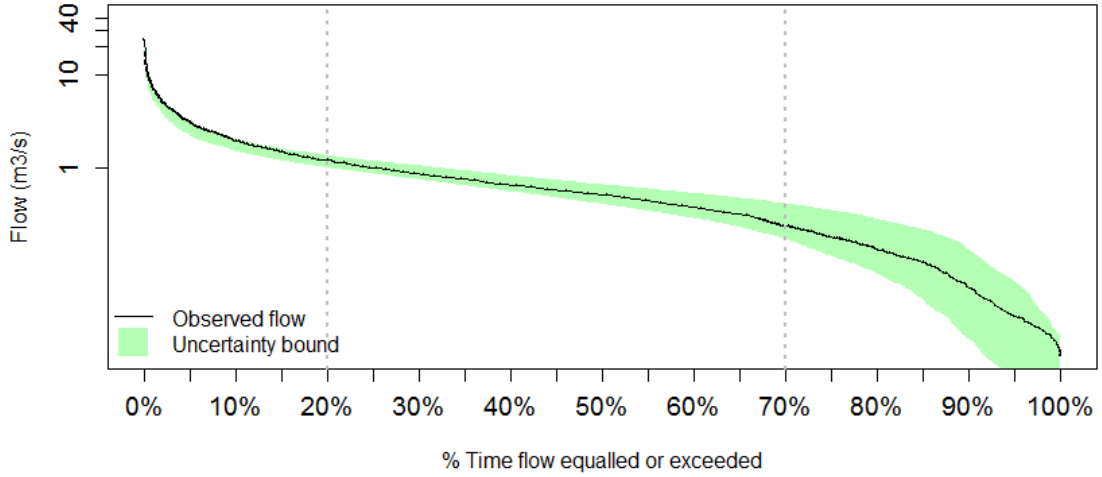
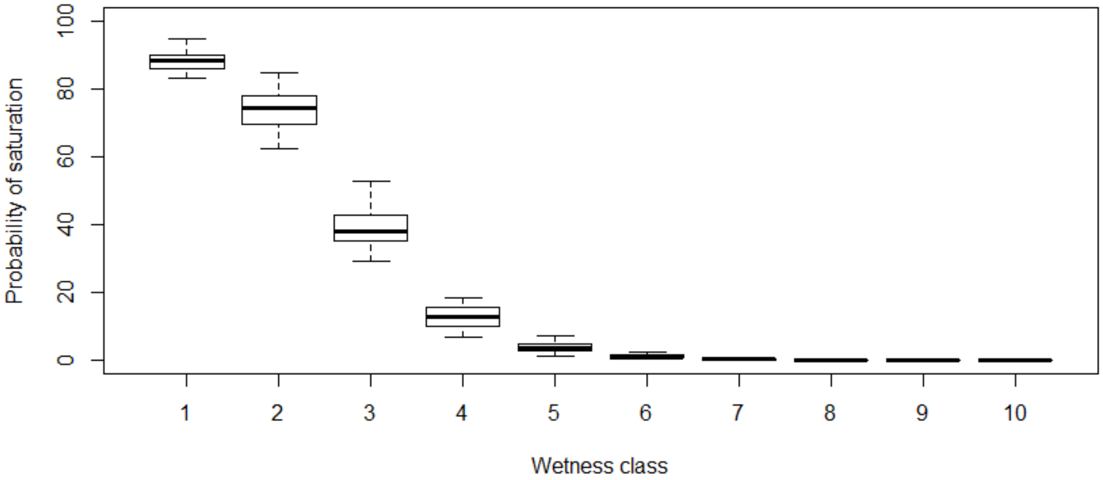


Figure 4.14. Distribution of good performing parameters for streamflow (in light green) and for both streamflow and saturated areas (in green).

Using the 53 good parameter sets identified, we estimated the uncertainty of streamflow prediction (Figure 4.15a) and uncertainty of the probability of saturation in 10 wetness classes (percentage of the days that the wetness class is saturated) (Figure 4.15b). The uncertainty bound of streamflow could capture correctly the statistical distribution of measured streamflow, as shown in Figure 15a. Taking into account parameter uncertainty, SWAT-HS predicted that the probabilities of saturation for wetness classes 7, 8, 9, 10 are almost 0, which means that 64% of the watershed is almost never saturated. Wetness class 1 has the highest probability of saturation (84-100%) because it has the lowest storage capacity i.e., it is always filled when there is rainfall. Wetness class 2 is also prone to saturation with a high probability of saturation varying from 62 – 85%. Wetness class 3 has the largest uncertainty of probability of saturation ranging from 29 – 53%. Wetness classes 4 and 5 have smaller uncertainties because they are only saturated with high rainfall events.



(a) Uncertainty of streamflow prediction in the form of flow duration curve



(b) Uncertainty of probability of saturation in wetness classes

Figure 4.15. Uncertainty of modelled predictions by SWAT-HS: (a) Uncertainty of streamflow prediction and (b) Uncertainty of probability of saturation in wetness classes.

4.1.4. Conclusion

This paper introduces SWAT-HS, which uses a new approach to represent the mechanism of saturation-excess runoff. By redefining HRU to include wetness information and introducing a surface aquifer to route interflow laterally from “drier” (upslope) to “wetter” (downslope) wetness classes, SWAT-HS changed the distribution of surface runoff and predicted soil saturation and excess runoff occurring only in wet downslope areas. Our novel contribution is in calibrating and evaluating a watershed model based on both temporal and spatial observations of streamflow and saturated areas respectively. The SWAT-HS model provided better performance of streamflow simulation and better prediction of saturated areas in the Town Brook watershed, compared to SWAT2012, which uses the curve number approach to compute surface runoff.

Similar to other distributed watershed models, the equifinality problem that causes uncertainty in model predictions is acknowledged in the application of SWAT-HS. However, the use of the combination of temporal and spatial observations helped to reduce a significant number of "good" parameter sets and shorten the uncertainty ranges of several parameters. More observations of saturated areas on high rainfall events with different antecedent soil storage conditions will be useful to identify robust parameter values and test the validity of SWAT-HS in high rainfall conditions.

Finally, the concept of self-organizing wetting patterns in the landscape on which SWAT-HS is based made it possible to identify the dominant response of the watershed (i.e., runoff and location of saturated areas) to climate forcing using the saturated runoff mechanism. Current best management practices in the NYC source watershed in which Town Brook is located are based on targeted application of manure and other fertilizers to avoid area that are periodically saturated. This practice has resulted in greatly reducing phosphorus concentration in the stream and reinforces the validity and usefulness of models with saturation excess and self-organization of moisture patterns for watershed management (Walter *et al.*, 2001; Bryant *et al.*, 2008). This study is the first step towards extending the application of SWAT-HS to larger watersheds for both water quantity and quality evaluations.

4.2. Key controls on spatial and temporal variability of dissolved organic carbon in Neversink River basin

4.2.1. Introduction

Recent studies of long-term trends have shown an increase in DOC concentrations in streams and lakes across northern Europe and North America (Evans *et al.*, 2005). This increase in DOC may be problematic for drinking water supplies due to the potential for a corresponding increase in disinfection by-products (DBPs) following chlorination. Here we describe analysis of the spatial and temporal variability of DOC within the Neversink River Basin in order to understand sources of variability and change in this important water quality parameter.

4.2.2. Study sites and methodology

This analysis used DOC data from 12 USGS stream gauging stations within the Neversink River Basin, located in the Catskill Mountains, New York and included an analysis of the changes in DOC concentration at daily, seasonal, and annual scales. We examined the relationships between DOC concentrations observed at these 12 sites and watershed characteristics. The watershed characteristics included topography (slope, wetness index), soil properties (soil carbon content), and vegetation density (leaf area index). Equation (4.10) lists the key variables that were used to explain observed DOC concentrations in streams.

$$DOC = f(Top, V, S, H) \tag{4.10}$$

where Top is topographic parameters (slope, topographic wetness index), V is vegetation density (leaf area index), S is soil properties (soil organic matter), and H is hydrologic flow partition.

Figure 4.16 and Table 4.5 show the name and location of the 12 sub-watersheds in Neversink River Basin and their size. DOC data is available for all 12 sub-watersheds for 1992-94. This study focused on analysis of the spatial variability of DOC between the 12 sub-watersheds. We identified the properties of 12 sub-watersheds, including slope, topographic wetness index, leaf area index (LAI), baseflow ratio, and soil organic content (SOC) (Figure 4.17). SOC values are based on SSURGO soil maps. Baseflow was calculated using the baseflow separation method of Nathan and McMahon (1990). Landsat-TM images (06/21/1991) were used to derive normalized difference vegetation index (NDVI), applying an empirical calibrated equation of NDVI and LAI developed in other studies (Hwang, unpublished data) to derive the LAI values for the 12 sub-watersheds. Based on these data, it was found that watersheds W2 and W12 have the lowest slopes, while W7 has the highest slope, the lowest TWI and the lowest LAI. Base flow ratio (base flow/ total flow) has a consistent pattern among years. WY 1992 has the highest base flow ratio. WY 1993 has lower base flow ratios than WY 1994, while WY 1994 has higher annual precipitation than WY 1993. Among the watersheds, W-3 has the lowest base flow ratio, while W-6 has the highest.

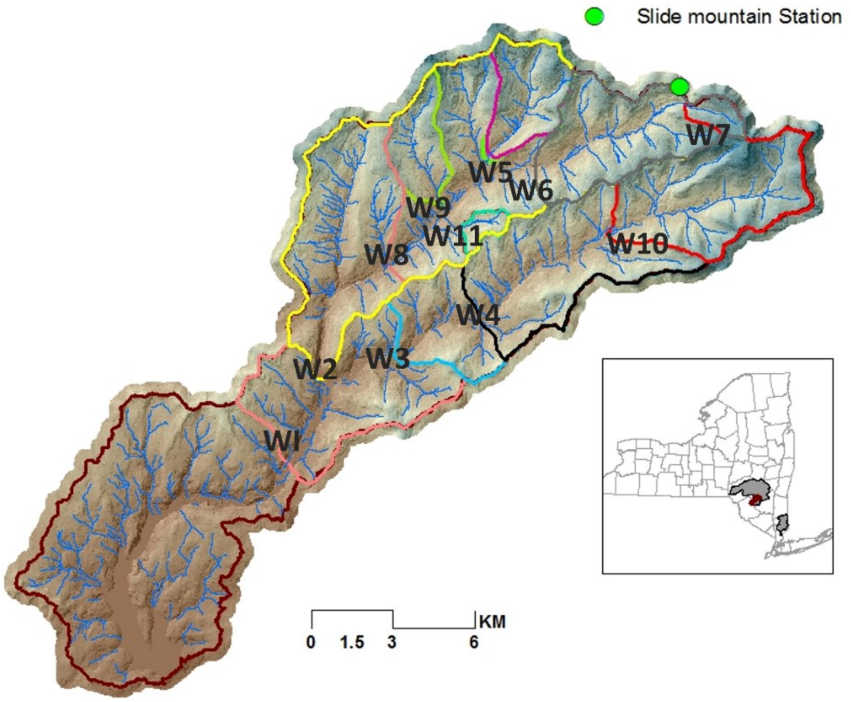


Figure 4.16. Study sites: USGS gauging stations and Slide mountain climate station.

Table 4.5. Location of streamflow and DOC sampling sites.

ID	Watershed name	Gauge ID	Area(mi^2)	Period of record USGS
W1	Main Branch (MB)	01435000	66.6	1954-04-06 to 2016-07-13
W2	West Branch (WB)	01434498	33.8	1985-09-04 to 2014-08-27
W3	East Branch (EB)	01434017	22.9	1991-07-02 to 2014-08-20
W4	New Hill (NH)	01434013	18.6	1991-07-02 to 2014-07-08
W5	Biscuit Brook (BB)	01434025	3.72	1983-04-10 to 2015-03-31
W6	Wild Cat (WC)	0143402265	7.89	1991-05-01 to 1994-07-19
W7	Winnisook (WS)	01434021	0.77	1991-01-01 to 2013-03-13 (water quality data is 1985-2015)
W8	Otter Pool (OP)	01434176	25.3	1991-07-01 to 1994-07-19
W9	High Falls (HF)	01434105	2.74	1990-10-01 to 1995-09-30 (water quality 1983-1999)
W10	Tisons (TS)	0143400680	8.93	1990-10-01 to 2013-03-31
W11	Shelter Creek (SC20)	01434092	0.62	1992-10-01 to 2007-09-30
W12	Clear Creek (CL25)	01434112	0.20	water quality:1992-10-01 to 2010-03-17

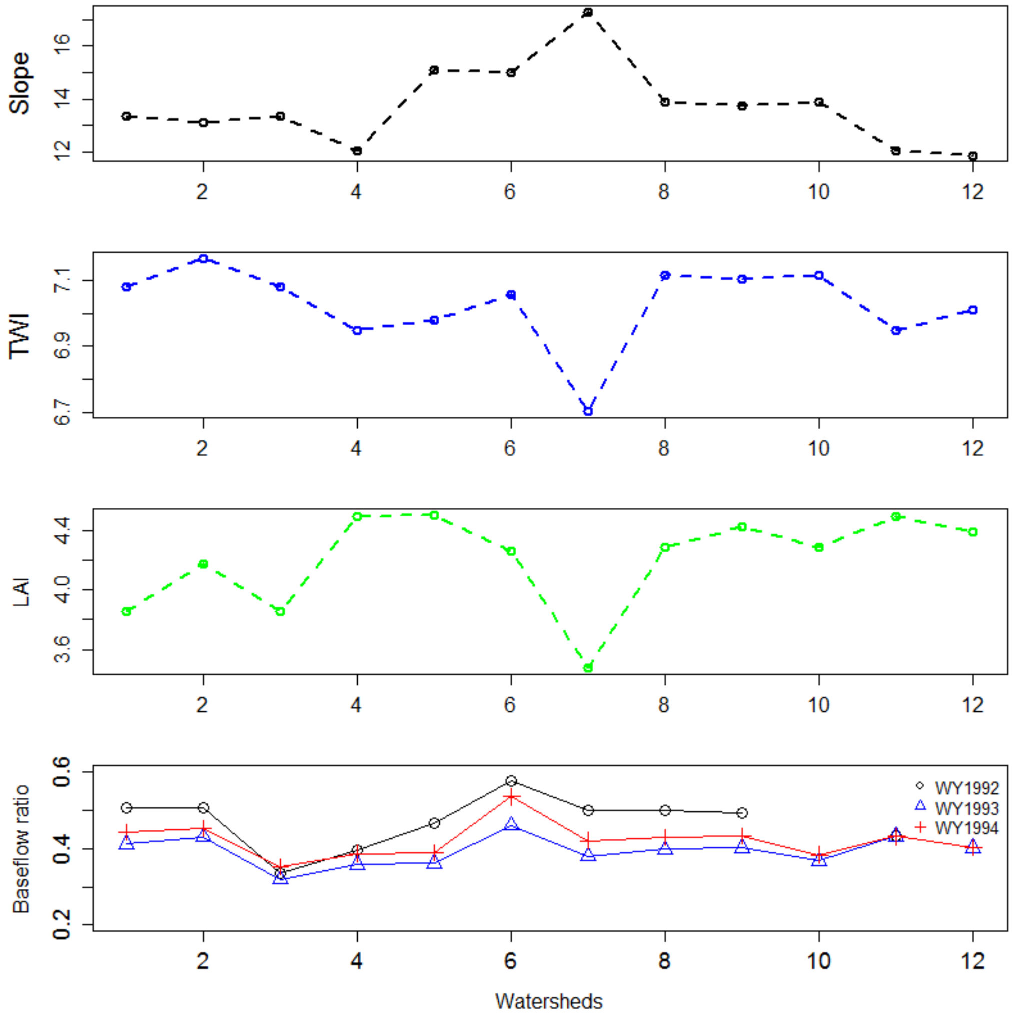


Figure 4.17. Catchment properties: (a) Slope, (b) TWI, (c) LAI and (d) Baseflow ratio.

4.2.3. Temporal variability of dissolved organic carbon

To analyze the temporal variability of DOC in the 12 sub-watersheds, we used DOC concentration data in water years (WY) 1992, 1993 and 1994. Figure 4.18 shows the daily averaged DOC concentration data of the 12 sub-watersheds in 1992, 1993 and 1994. WY 1994 has higher DOC concentrations for most of the sub-watersheds as compared to WY 1992, and 1993. Figure 4.19a, and Figure 4.19b show the annual precipitation and annual maximum daily temperature in WY 1992, 1993 and 1994. The difference in annual maximum daily temperature among the three years is minimal but the annual precipitation of WY 1994 is greater than those of WY 1992 and 1993. The relationship between annual precipitation and annual median DOC concentration was analyzed using WY1992 to WY 2014 data in Biscuit Brook (Figure 4.19c). Stream DOC concentrations at Biscuit Brook are positively correlated with the annual precipitation. The coefficient of determination (R^2) between annual precipitation and annual median daily stream DOC concentration at the Biscuit Brook is 0.22. These results suggest that in higher precipitation year stream DOC tends to increase. We also found that annual streamflow is positively correlated with stream DOC concentrations (not shown here).

In addition to annual DOC variability, we also explored the relationship between DOC concentration and daily average streamflow. Figure 4.20 shows the relationship between storm flow (and baseflow) and stream DOC concentration for Biscuit Brook in the WY 1993, and 1994. WY 1993 has stronger relationship (R^2=0.42, 0.18) between storm flow (and baseflow) and stream DOC concentration than WY 1994 (R^2=0.33, 0.03). Storm flow can better explain the variance of measured stream DOC concentration than baseflow. These patterns are also shown in other sub-watersheds. The statistical relationship between streamflow and stream DOC concentration is summarized in Table 4.6.

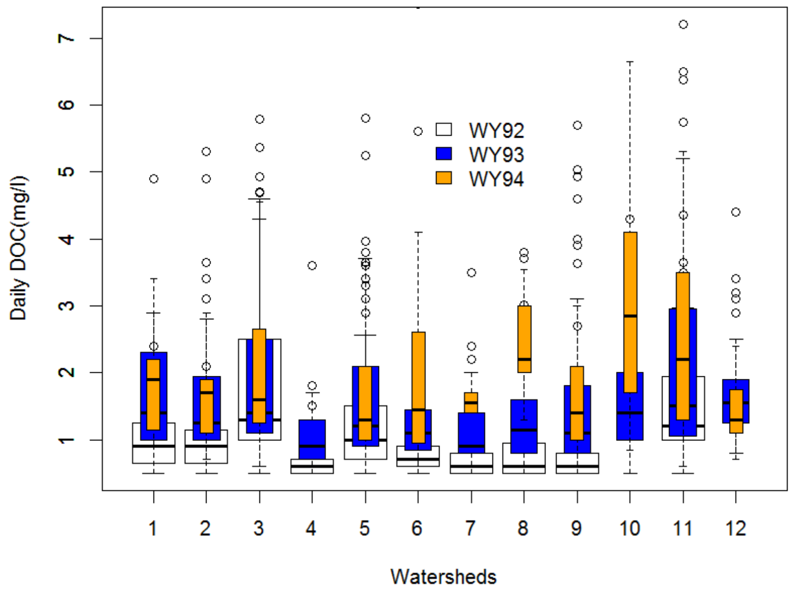


Figure 4.18. Daily DOC concentration variations among watersheds in water year (WY) 1992, 1993 and 1994:Sampled DOC data is averaged per day. Middle line in the box is the median; upper and lower ends of the box are 75th and 25th percentiles; thin lines are the data range; circles are considered outliers.

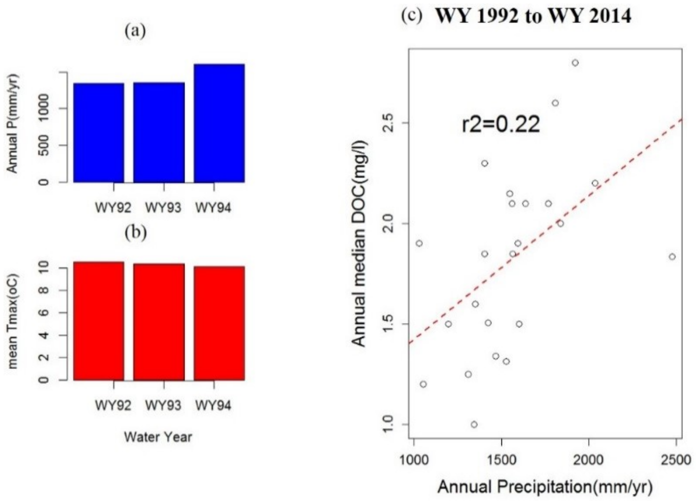


Figure 4.19. Climate variability in WY 1992, 1993 and 1994 and the relationship between annual DOC concentration and annual precipitation of Biscuit Brook: (a) Annual precipitation, (b) annual maximum daily temperature, (c) the relationship between annual median DOC and annual precipitation.

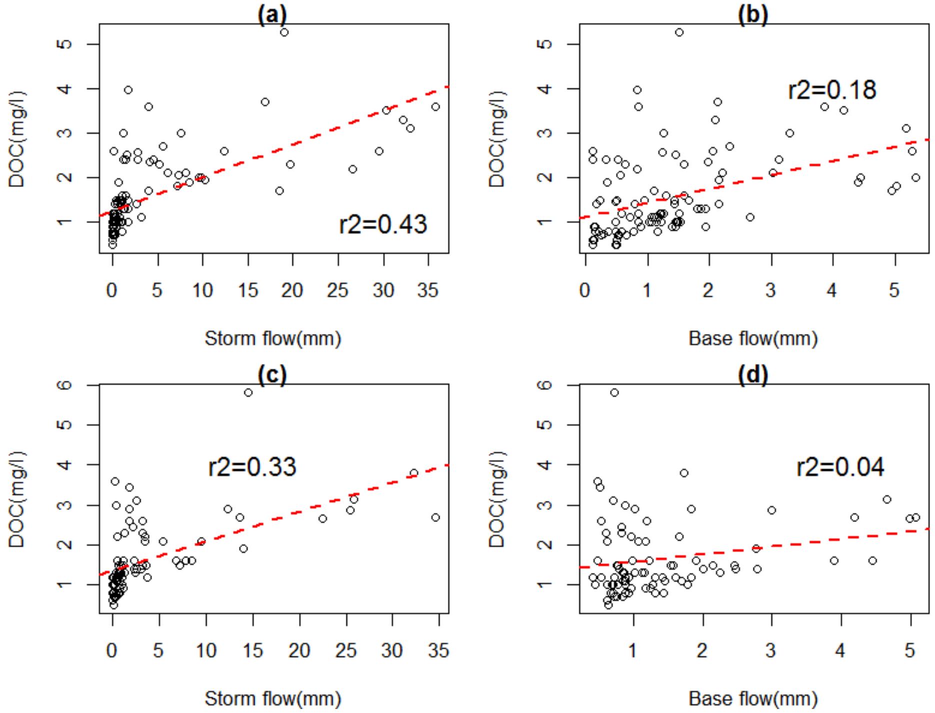


Figure 4.20. Relationship between daily storm/baseflow and daily DOC concentration of Biscuit Brook in the period of 1993 and 1994: (a) storm flow vs DOC in year 1993, (b) baseflow vs DOC in year 1993, (c) storm flow vs DOC in year 1994, and (d) baseflow vs DOC in year 1994.

Table 4.6. Statistical relationship (R^2)[1] between daily streamflow and DOC concentration in Neversink River Basin

ID	Watershed name	WY 1993	WY 1994
W1	Main Branch (MB)	0.45	0.58
W2	West Branch (WB)	0.55	-
W3	East Branch (EB)	0.65	0.42
W4	New Hill (NH)	0.55	0.43
W5	Biscuit Brook (BB)	0.43	0.30
W6	Wild Cat (WC)	0.66	-
W7	Winnisook (WS)	0.56	0.67
W8	Otter Pool (OP)	0.68	-
W9	High Falls (HF)	0.40	0.41
W10	Tisons (TS)	0.59	0.26
W11	Shelter Creek (SC20)	0.81	0.05
W12	Clear Creek (CL25)	0.05	0.08

[1]: R^2 is coefficient of determination.

We also explored how the relationship between storm flow and DOC concentration varies among different seasons. Figure 4.21 shows the scatter plots between calculated storm flow and measured stream DOC concentration in spring, summer, fall, and winter. Data from WY 1992 to 2004 were used to calculate these relationships. We used R^2 (coefficient of determination) of as a measure of how well storm flow explains the variance of stream DOC concentrations, and applied a linear power regression model to explain the relationship between streamflow and stream DOC. Stream DOC is positively correlated with streamflow in spring, summer, fall and winter seasons, and the relationship between the two variables in all seasons has a power form ($DOC=aQ^b$). Calculated R^2 varies between seasons; R^2 values of spring, summer, fall and winter seasons are 0.40, 0.51, 0.54, and 0.42 respectively. Storm runoff has a higher correlation with stream DOC concentration than baseflow (not shown here), suggesting that most of the organic matter is transported to the stream via surface and shallow subsurface runoff that is in contact with organic rich soil horizons.

We also used the relationship between flow and DOC fluxes to characterize DOC export patterns as source-limited or transport-limited. The relationship between flow and DOC fluxes also has a power form. Figure 4.22 shows the scatter plots between measured flow and measured stream DOC fluxes. Linear power regression models had robust goodness of fit for the four seasons (R^2>0.95). The estimated ‘b’ (exponent) values are larger than 1, which suggests that DOC export in Biscuit Brook is transport limited rather than source-limited. These results are consistent with the results of Raymond and Saiers, (2010) showing that most of DOC is exported during large events.

The difference in R^2 between seasons may be attributed to the difference of the antecedent moisture conditions, and available DOC variability within the watershed. The study of Brown *et al*. (1998) in Shelter Creek showed that summer DOC concentrations are lower in the rising limb of the hydrograph than those on the recession limb. In the small forested watershed in the Rocky Mountains of Colorado, spring stream DOC concentration were found to peak on the rising limb of the snowmelt hydrograph, prior to peak snowmelt (Boyer *et al.*,1997). Since seasonal soil moisture (or groundwater level) and organic matter data is not available in the watershed, application of a hydrologic model can be an alternative in order to explore the changes of soil moisture and DOC availability within a watershed. Boyer *et al.* (1996) combined a hydrologic model (TOPMODEL) and a simple DOC model to explain the flushing mechanism of DOC in streams of Deer Creek catchment, Colorado.

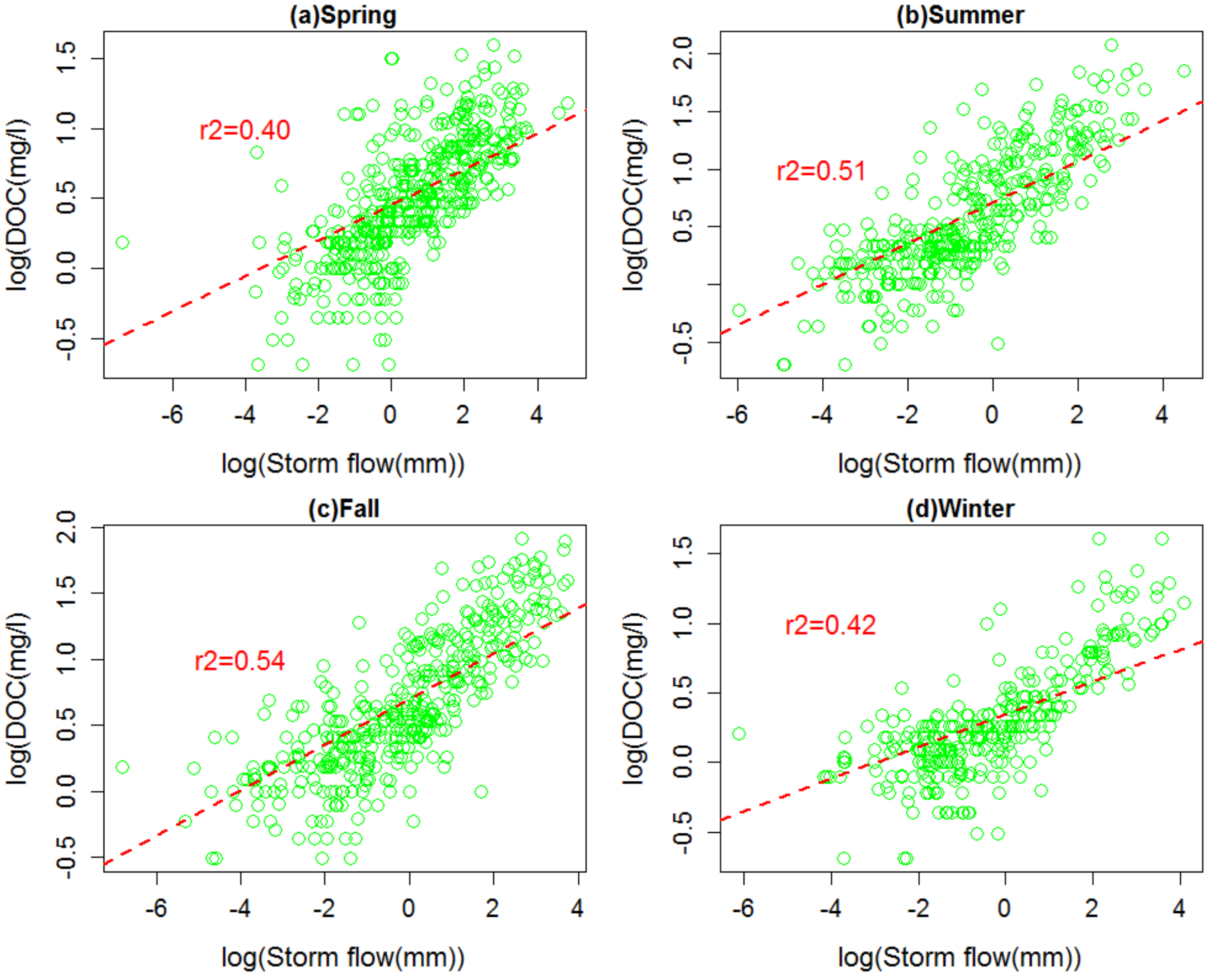


Figure 4.21. Relationship between log daily storm flow and log daily averaged DOC concentration of Biscuit Brook in different seasons, 1992-2014.

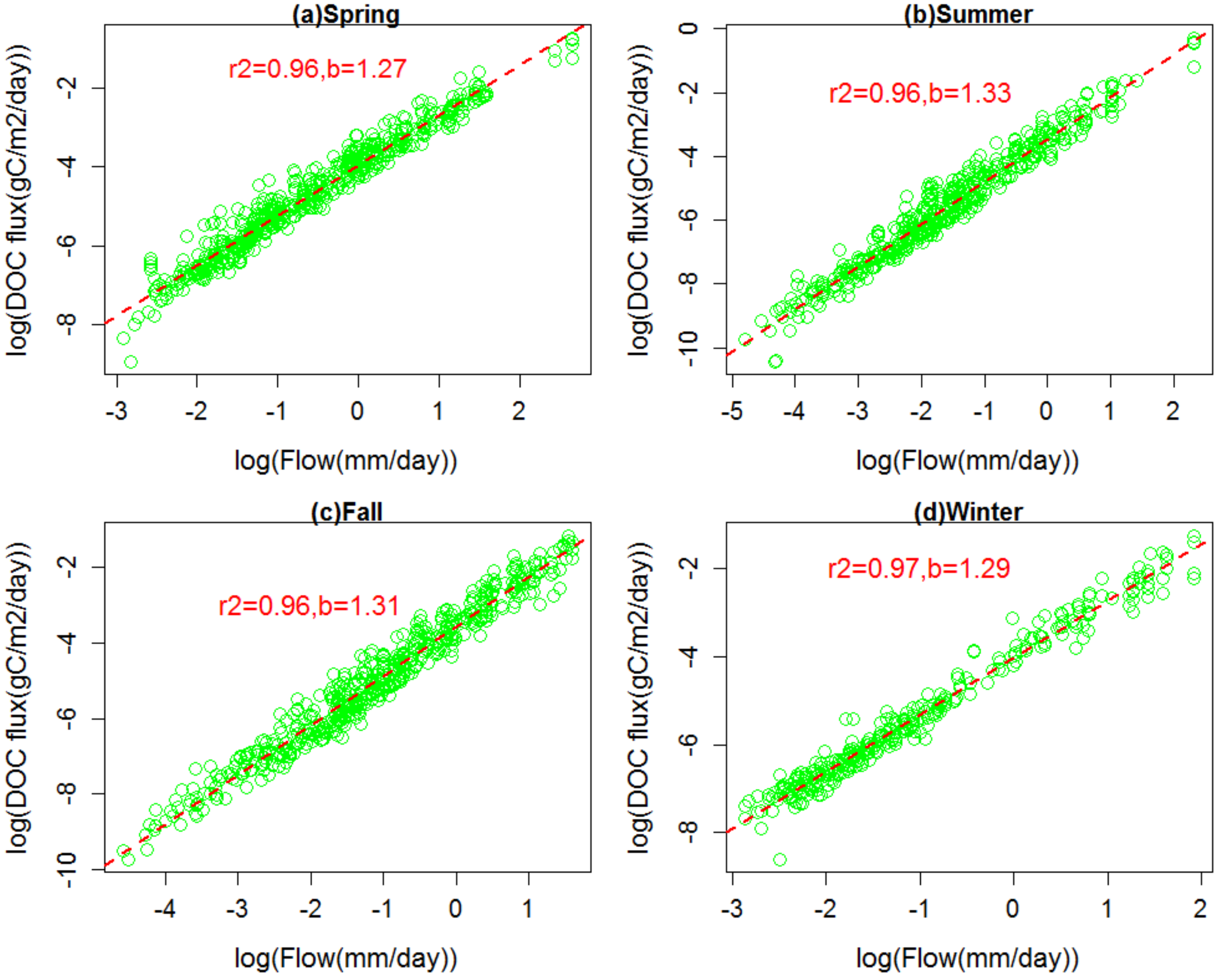


Figure 4.22. Relationship between log daily flow and log daily DOC fluxes of Biscuit Brook in different seasons, 1992-2014.

4.2.4. Spatial variability of dissolved organic carbon in Neversink River Basin

We analyzed the relationship between median annual DOC concentration and watershed properties (Figure 4.23). Slope was the only characteristic that has a statistically significant relationship (p-value<0.05) with DOC in these watersheds. Steeper sloped watersheds tend to have lower stream DOC concentrations than milder sloped watersheds. Other watershed properties, including TWI, LAI, baseflow ratio and SOC do not show statistically significant relationship with the difference of stream DOC concentrations between the watersheds.

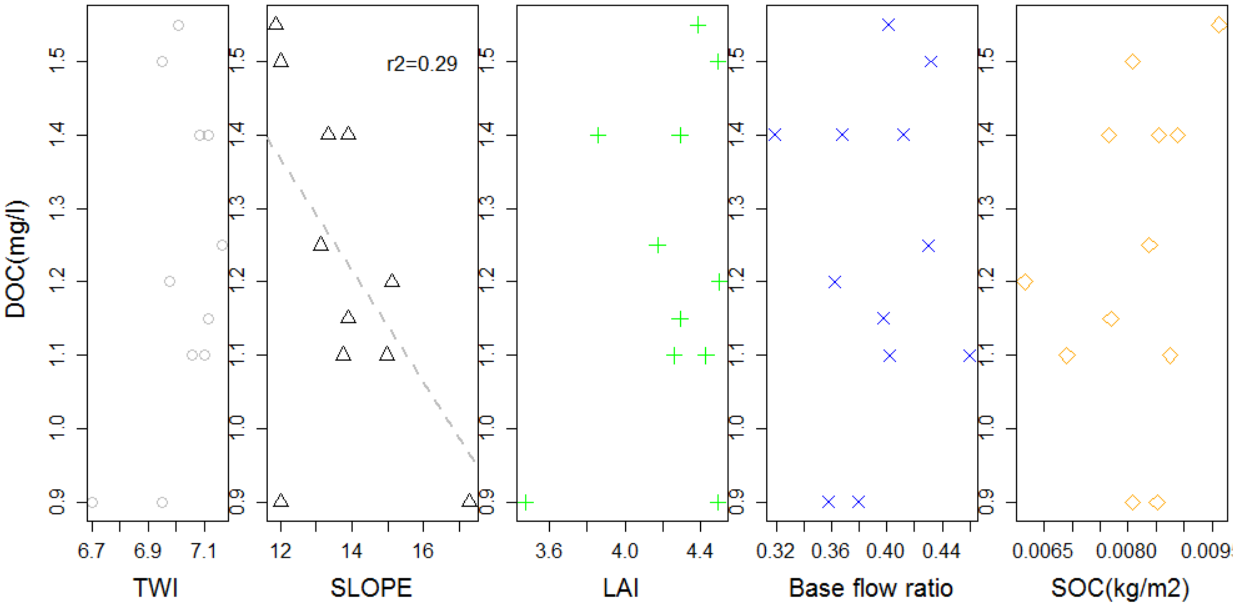


Figure 4.23. Relationship between catchment properties and annual median daily DOC: (a) TWI (b) slope, (c) LAI, (d) baseflow ratio and (e) soil organic carbon

4.2.5. Developing a conceptual model for explaining the variability of dissolved organic carbon

We analyzed the spatial and temporal variability of DOC concentration within Neversink River basin. The analysis is based on watershed properties (topography, soil and vegetation and hydrologic flow partition) and stream DOC concentration variation. The analysis showed that flatter sloped watersheds tend to produce higher DOC concentration than steeper sloped watersheds. Flatter slope can be related to the probability of having saturated areas, and wetlands. Webster *et al.*(2011) present topographic templates that estimates the soil carbon pools in the forested catchments, and the template demonstrates that larger soil carbon pools are observed in the footslope, toeslope and wetland features. Since soil carbon pool tends to positively correlated with stream DOC, we can infer that flatter sloped watershed will have higher accumulated soil carbon and its higher soil carbon leads to transporting higher DOC concentration in streams.

The stream DOC concentration is also positively correlated with storm flow (Figure 4.20 and Figure 4.24). A study by Brown *et al*. (1998) in Shelter Creek watersheds showed that Soil B and C horizon carbon levels are too low to explain the higher DOC concentration in the stream during peak events. Through end-member mixing analysis, Brown *et al*. (1998) found that at the start of an event, stream DOC concentrations were close to groundwater DOC levels. At the peak of event, stream concentrations were close to through fall (under canopy) samples, while after the peak of the hydrograph, stream concentrations became close to that observed in the O horizon. These results suggest that the higher surface organic layer is a major contributor to stream DOC, when groundwater level is close to surface. These finding are similar to the study of Boyer *et al.* (1997) that suggest the flushing mechanism explained the higher DOC transport in the snowmelt periods in the Deer Creek watershed in Colorado.

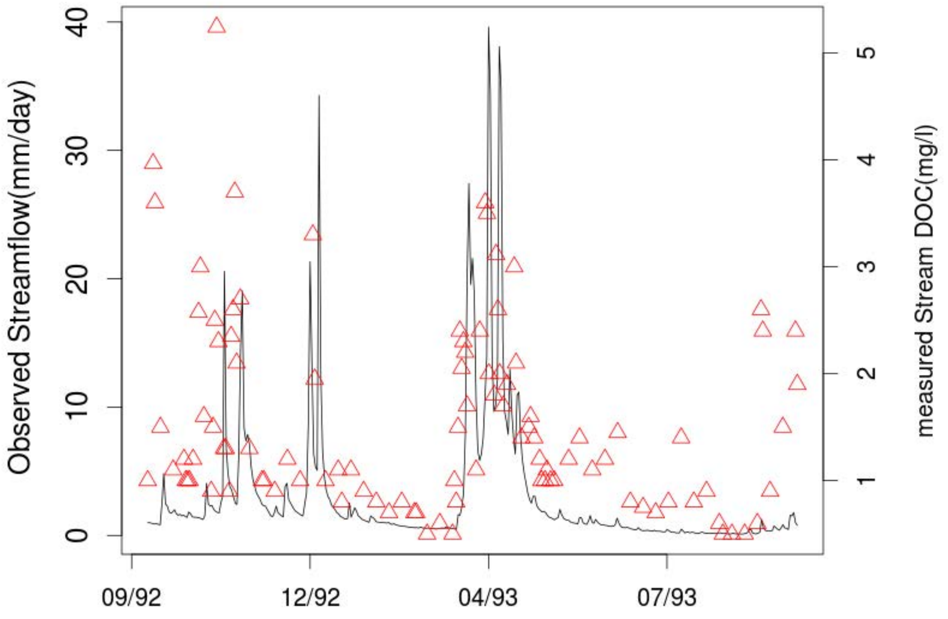


Figure 4.24. Daily time series of streamflow with sampled DOC concentration data in water year 1993, Biscuit Brook

4.2.6. Summary and conclusion

To understand the key control(s) on stream DOC processes, we explored the DOC data from 12 USGS stream gauging stations, and examined the relationship between DOC concentrations and watershed characteristics at the 12 sites. These characteristics included topography (slope and topographic wetness index), soil properties (soil carbon) and vegetation density (leaf area index). We found that only slope has a significant relationship with observed DOC concentration. In addition to spatial variation of DOC concentration, we also explored the temporal variability of DOC. At an annual time scale, precipitation has a positive relationship with annual median DOC concentration. Daily DOC concentration is positively correlated with daily streamflow as well. In particular, storm runoff has higher correlation with daily DOC concentration than baseflow. Similar results were found for seasonal DOC patterns; spring, summer, fall and winter DOC concentrations are positively correlated with storm runoff. These results suggest that most organic matter is transported to stream via surface and shallow subsurface runoff that are in contact with organic rich soil horizons. In addition, high slopes (b>1) of the power-law relationship between flow and DOC fluxes for four seasons suggest that DOC export is characterized as transport-limited rather than source-limited. However, the relationship between storm flow and stream DOC concentration varied between seasons. This difference may be attributed to variation in antecedent moisture conditions, and available DOC variability. Current data analysis is solely based on instream DOC and flow. Therefore, the analysis is limited to explaining spatial and temporal distribution of DOC sources areas, and associated water movements within the watershed. In the next section, an ecohydrologic model RHESSys (Tague and Band, 2004) is used to investigate how DOC sources vary spatially and temporally within a watershed, and how available DOC is transported to stream.

4.3. RHESSys model predictions of streamflow, nitrate and dissolved organic carbon in two Neversink River watersheds

4.3.1. Introduction

Catskill Mountain watersheds in New York State have historically experienced soil and stream acidification due to atmospheric deposition of SO_2 and NO_x. Change in atmospheric deposition has altered the biogeochemistry in soil and stream, as evidenced by increases in dissolved organic carbon (DOC) (Findlay, 2005) and nitrate (Murdoch and Stoddard, 1992). These DOC and nitrate increases are key questions for watershed and water supply management in this region. This study applies a process-based ecohydrologic model (RHESSys) to two watersheds in Neversink River Basin to improve the understanding of the sources, fate and transport mechanisms, and export patterns for nitrate and DOC.

4.3.2. Study sites and methodology

This study focus on the two headwater watersheds, Shelter Creek (1.6 km^2) and Biscuit Brook (9.2 km^2), located in Neversink River Basin. Figure 4.25 shows the location and boundary of each watershed. The major land cover type is forest, and mixed northern hardwood forest is the dominant forest type, consisting of American beech, red maple, sugar maple and yellow birch (Harpold, *et al*., 2010). Balsam fir is found above 1000 m elevation, and hemlock is found at the ridge and near streams. The major soil type is Inceptisols, and its depths range from 0.5 - 1 m. Soils have coarse texture, and are well drained. The major surface geologic type is sandstone (60%). Underlying bedrock type is sedimentary rock, and the bedrock has high percentages of fractures. Bedrock mineralogy is spatially uniform. The till depth is shallow (<1 m), and in the valley bottoms, the coarse sand and gravel is formed by the glacial outwash and recent alluvium deposits. The average precipitation (based on Slide Mountain climate station, located in the summit of Neversink River Basin) is 1570 mm, and mean air temperature is 5.2℃. Snow as a percentage of the total precipitation ranges from 20 to 35% (Harpold *et al.*, 2010).

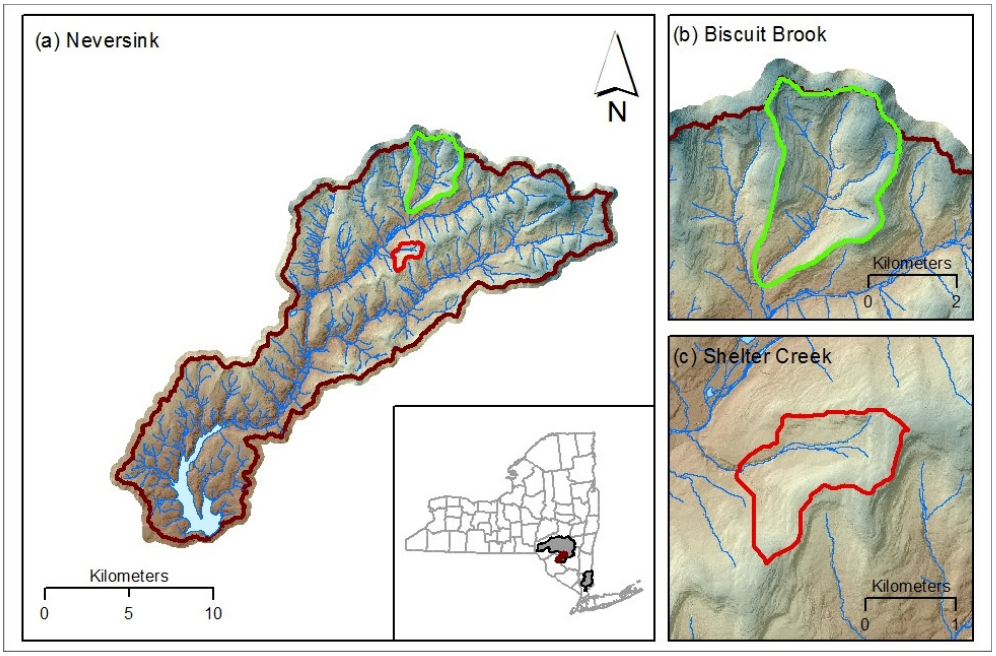


Figure 4.25. Study sites: (a) Neversink River Basin, (b) Biscuit Brook and (c) Shelter Creek.

The model used for this study is RHESSys, a physically based, spatially explicit model that has the capability to simulate climate and vegetation change impacts on hydrology and ecosystem carbon and nutrient cycling. Figure 4.26 shows RHESSys modeling framework, including the model structure, and detailed vertical and horizontal hydrological processes. This model was applied to the Shelter Creek and the Biscuit Brook watersheds.

RHESSys requires various input data; climate inputs (e.g. daily precipitation and daily minimum and maximum temperature), topography (i.e. Digital Elevation Model), soil and vegetation data. Figure 4.27 shows some of the RHESSys input data for Shelter Creek. Similar level of model input data was used for Biscuit Brook. NYC DEP produced 1m Digital Elevation Model (DEM) using LIDAR point data. This study used resampled 10m DEM (for Shelter Cr.) and 30m DEM (for Biscuit Br.) to generate the topographic parameters (elevation, slope, aspect, flow drainage maps, watershed boundary maps, etc.). Soils parameters are initially estimated based on Soil Survey Geographic (SSURGO, USDA Natural Resources Conservation Service) and existing RHESSys parameter library (available at https:www.github.com/RHESSys/ParamDB). We used a detailed vegetation species map (Driese *et al.*, 2004) for characterizing the different vegetation functional types, and associate vegetation parameters for each vegetation type was based on the RHESSys parameter library. Current model setups used Landsat TM-based LAI estimates for Biscuit Brook and Shelter Creek. Landsat-TM image (06/21/1991) were used to derive normalized difference vegetation index (NDVI). LAI values were estimated using an empirical calibrated equation of NDVI and

LAI developed in other studies (Hwang, unpublished data) because measured LAI data are not available in our sites.

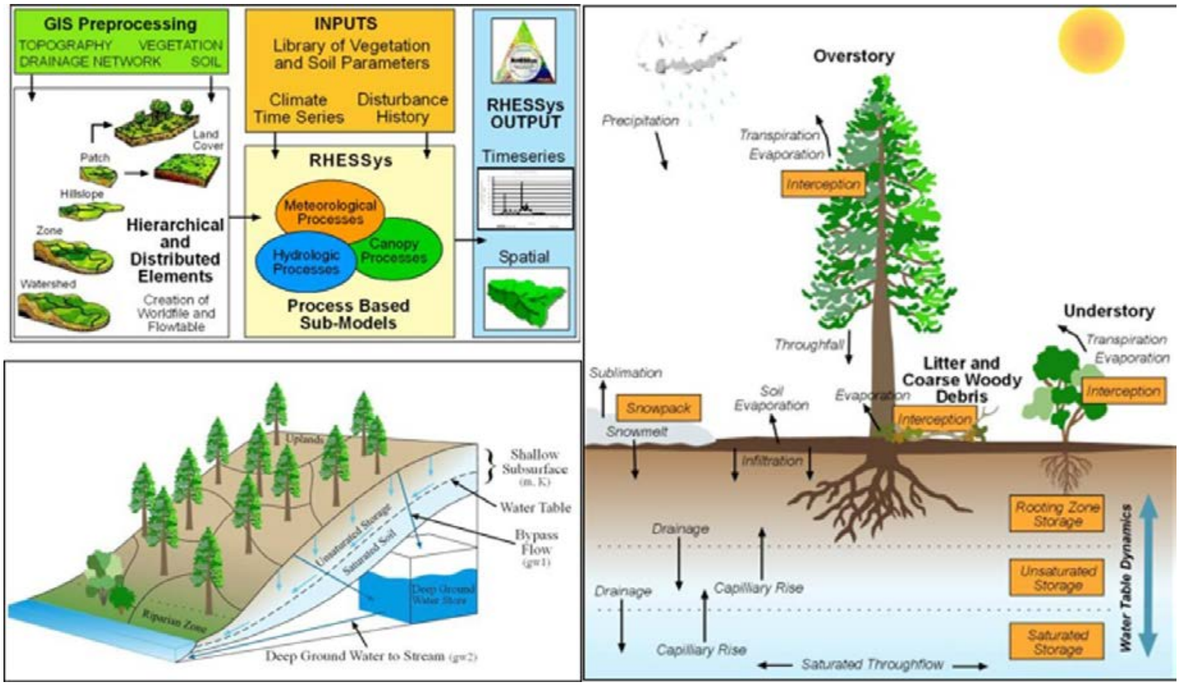


Figure 4.26. RHESSys modeling framework.

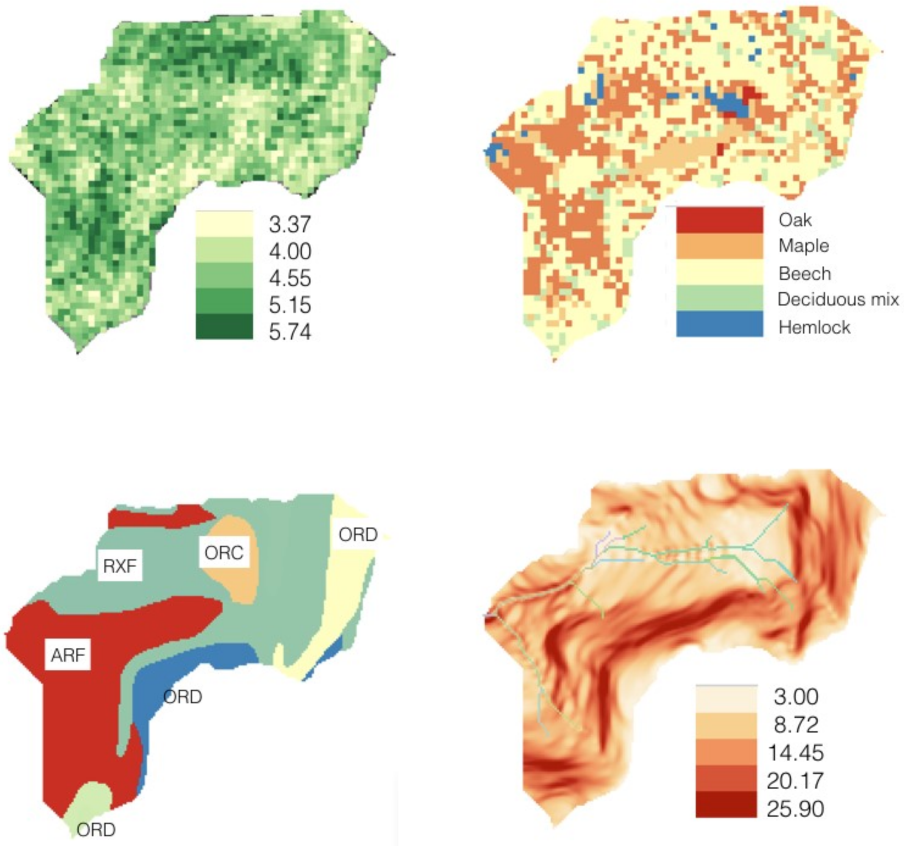


Figure 4.27. RHESSys input data for Shelter Creek: (a) LAI, (b) vegetation species, (c) soil types, and (d) slope.

For climate input data, this study used precipitation and air temperature data from the Slide Mountain climate station, located about 9 km and 6 km from Shelter Creek and Biscuit Brook, respectively. Other climate variables (e.g. solar radiation, saturation vapor pressure, and relative humidity) are computed using a climate interpolation model (MT-CLIM, Running et al., 1987).

RHESSys is initially calibrated with measured streamflow, with the focus on calibrating soil parameters. Six soil parameters were varied in the model calibration described below; these parameters are vertical and horizontal saturated hydraulic conductivities (Ksat), decay coefficient of Ksat with depth (m), soil depth (sd), the percentage of preferential flow from the soil surface to deep groundwater storage (gw1), and the linear coefficient of groundwater storage (gw2). The predictive performance of the model is evaluated using a combination of three accuracy measures (Equations (4.11, (4.12, and (4.13); a) Nash-Sutcliffe efficiency coefficient, NSE (Nash and Sutcliffe, 1970), b) Nash-Sutcliff efficiency coefficient with logarithmic values, logNSE and c) the percent volume error (PerErr). The value of combining the three accuracy (Equation (4.14)) measures range from 0 to 1 with the perfect fit at 1.

$$NSE = 1 - \frac{\sum_i (Q_{obs,i} - Q_{sim,i})^2}{\sum_i (Q_{sim,i} - \bar{Q}_{obs,i)}^{2}} \qquad (4.11)$$

$$logNSE = 1 - \frac{\sum_i (\log(Q_{obs,i)} - \log(Q_{sim,i}))^2}{\sum_i (\log(Q_{sim,i}) - \log(\bar{Q}_{obs}))^2} \qquad (4.12)$$

$$PerErr = \frac{(\bar{Q}_{sim} - \bar{Q}_{obs})}{\bar{Q}_{obs}} \times 100 \qquad (4.13)$$

$$Accuracy = NSE \times logNSE \times (1 - \left|\frac{PerErr}{100}\right|) \qquad (4.14)$$

where $Q_{obs,i}$ is the observed streamflow and $Q_{sim,i}$ is the simulated flow at any given time step (i), and [redacted] and [redacted] are the average of daily streamflow. Following calibration for streamflow, the model was used to predict stream nitrate and DOC for the two watersheds. In addition, LAI predicted by RHESSys is compared with LAI based on Landsat TM imagery.

4.3.3. Prediction of flow, nitrate, and DOC in Shelter Creek

RHESSys was applied to the Shelter Creek watershed (SC20), and the model was calibrated based on the measured streamflow. The calibration period of the model is from 10/1/1992 to 09/30/1995. Figure 4.28 shows the calibration results of modeled monthly flow. RHESSys failed to capture the timing of the peak flow, but the low flow prediction is better; (logNSE = 0.32). RHESSys also underestimated the annual flow (PerErr = -16%). However, RHESSys was able to reproduce the measured monthly flow; NSE and logNSE equal to 0.80 and 0.6, respectively. The poor prediction of daily streamflow may be attributed to high uncertainty in climate input data since this study used meteorological data from the Slide Mountain station, located about 9 km from Shelter Creek.

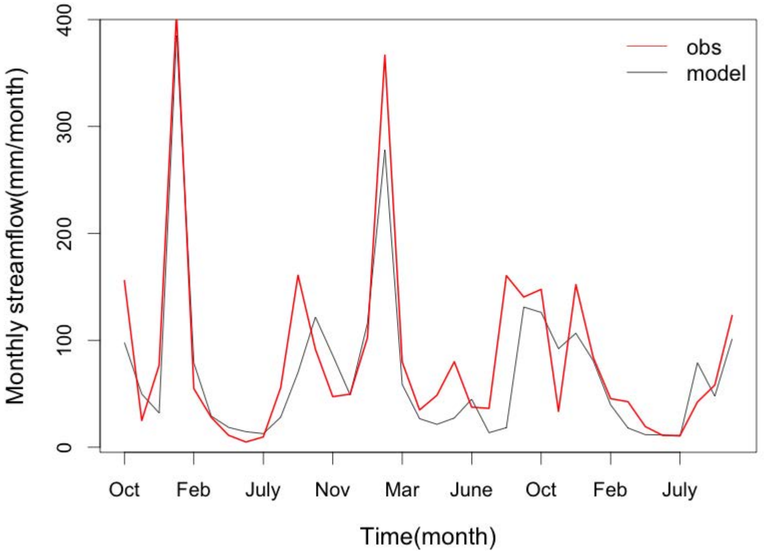


Figure 4.28. RHESSys streamflow predictions: monthly modeled flow vs measured flow, obs is measured streamflow, and model is modeled streamflow.

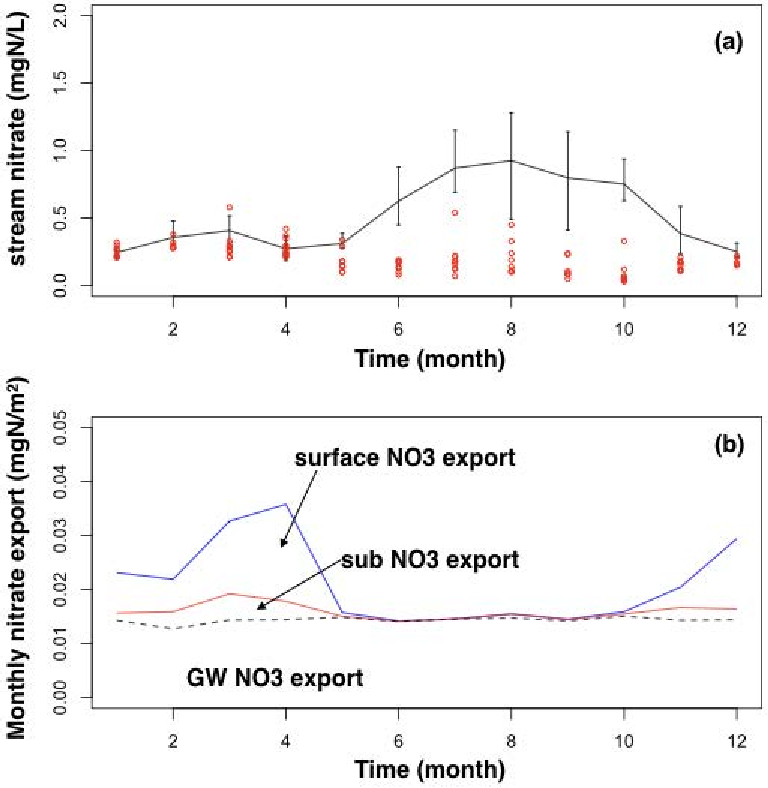


Figure 4.29. RHESSys nitrate simulations for Shelter Creek (SC20): (a) comparison of simulated and observed stream nitrate concentration, grouped by month, 1992-1995 in Shelter Creek (SC20), and (b) model estimates of stream nitrate sources.

The calibrated model was used to predict stream nitrate and DOC concentrations and fluxes. The diagnosis of the model performance regarding nitrate and DOC is limited to the monthly scale. Figure 4.29 shows RHESSys nitrate simulation results for Shelter Creek. RHESSys tends to overestimate the stream nitrate particularly in summer (Figure 4.29a). The overestimation of summer nitrate may result from underprediction of denitrification in groundwater storage because simulated NO_3 export is dominated by the deep groundwater N flux in summer (Figure 4.29b). Ashby *et al*. (1998) showed high denitrification rates in the groundwater seepage and near-stream areas in the Shelter Creek watersheds. Improving the denitrification processes in those areas will be necessary. We also found that stream nitrate estimates are sensitive to N deposition rates with RHESSys (not shown here). Current model implementation assumes a steady N deposition rate. Therefore, accounting for time-variable N deposition in RHESSys modeling will be necessary to improve the stream nitrate predictions.

Figure 4.30 shows the monthly DOC prediction for Shelter Creek. The model overestimates winter DOC and underestimates summer DOC; however, the ranges of modelled stream DOC concentration are within the measured values. The model estimate has very low seasonality of available soil DOC (Figure 4.30c). These results suggest that stream DOC processes is dominated by hydrologic transport rather than by source magnitude. It should be noted that this current model prediction is only based on soil parameters calibration. Future research requires adjusting DOC decay and adsorption parameters to improve the stream DOC predictions.

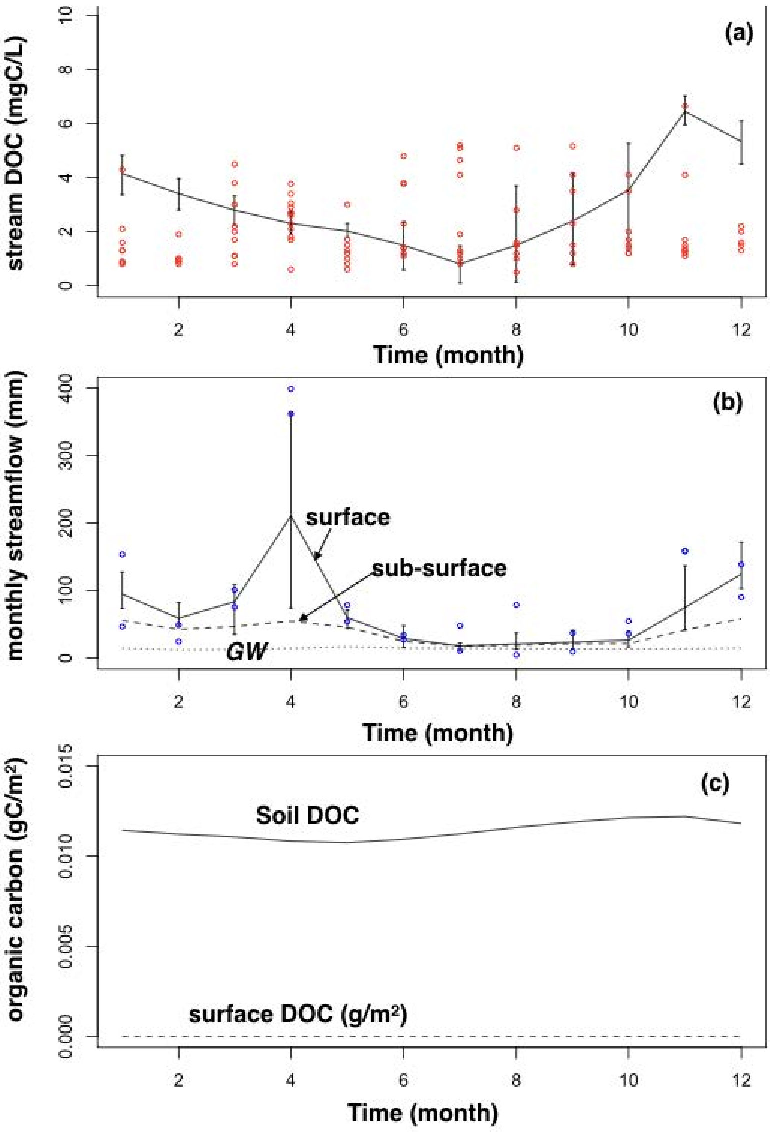


Figure 4.30. RHESSys stream DOC simulations in Shelter Creek (SC20): (a) Comparison of simulated and observed stream DOC concentration, grouped by month, 1992-1995, and (b) modeled estimates of flow components, (c) model estimates of soil DOC and surface DOC.

4.3.4. Prediction of flow, nitrate and DOC in Biscuit Brook

RHESSys was applied to Biscuit Brook, and the model was calibrated with measured streamflow. The calibration period of the model is from 10/1/1992 to 09/30/1995. As in the Shelter Creek application, six soil parameters were calibrated. Figure 4.31 compares RHESSys streamflow predictions with observations. Predicted daily streamflow is underpredicted in fall and early winter flow, with large error in the timing and magnitude of peak flow (R^2 value is 0.3

and daily NSE is 0.25). The error in early winter may be due to missing snow-rain transition, and the partition of total precipitation in rain and snow. Due to climate inputs from outside the watershed and lack of snow data, it may be difficult to correct these model errors. In addition, the underestimated peak flow in fall may be related to poor parameterization of soil depths. Estimated soil depth may be too deep to generate storm runoff when rain occurs. The predictions for low flow (logNSE=0.47) were better than those for high flow (NSE=0.25). Predictions of the seasonal pattern of streamflow (not shown here) were better; (NSE = 0.76 for monthly average streamflow).

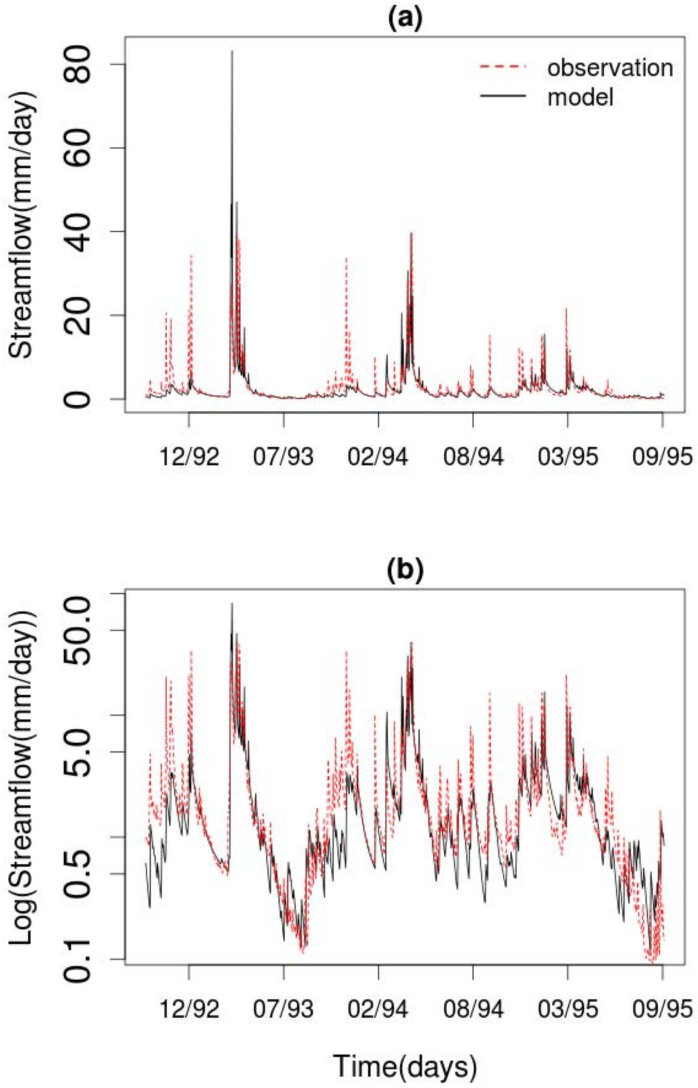


Figure 4.31. RHESSys streamflow predictions for Biscuit Brook: (a) measured daily streamflow vs modeled daily streamflow, and (b) measured log daily streamflow vs modeled log daily streamflow.

Following calibration for streamflow, the model was used to predict DOC concentration in the stream. For 1993, RHESSys overpredicts stream DOC in winter, but underpredicts in fall (Figure 4.32). High predicted winter DOC is likely related to underestimated streamflow (Figure 4.32b). Underestimated DOC in fall 1992 may be related to lower groundwater level because the temporal patterns of modeled DOC corresponds to the timing of rising groundwater level (Figure 4.32c). These results suggest that DOC export occurs when the groundwater table is close to the surface.

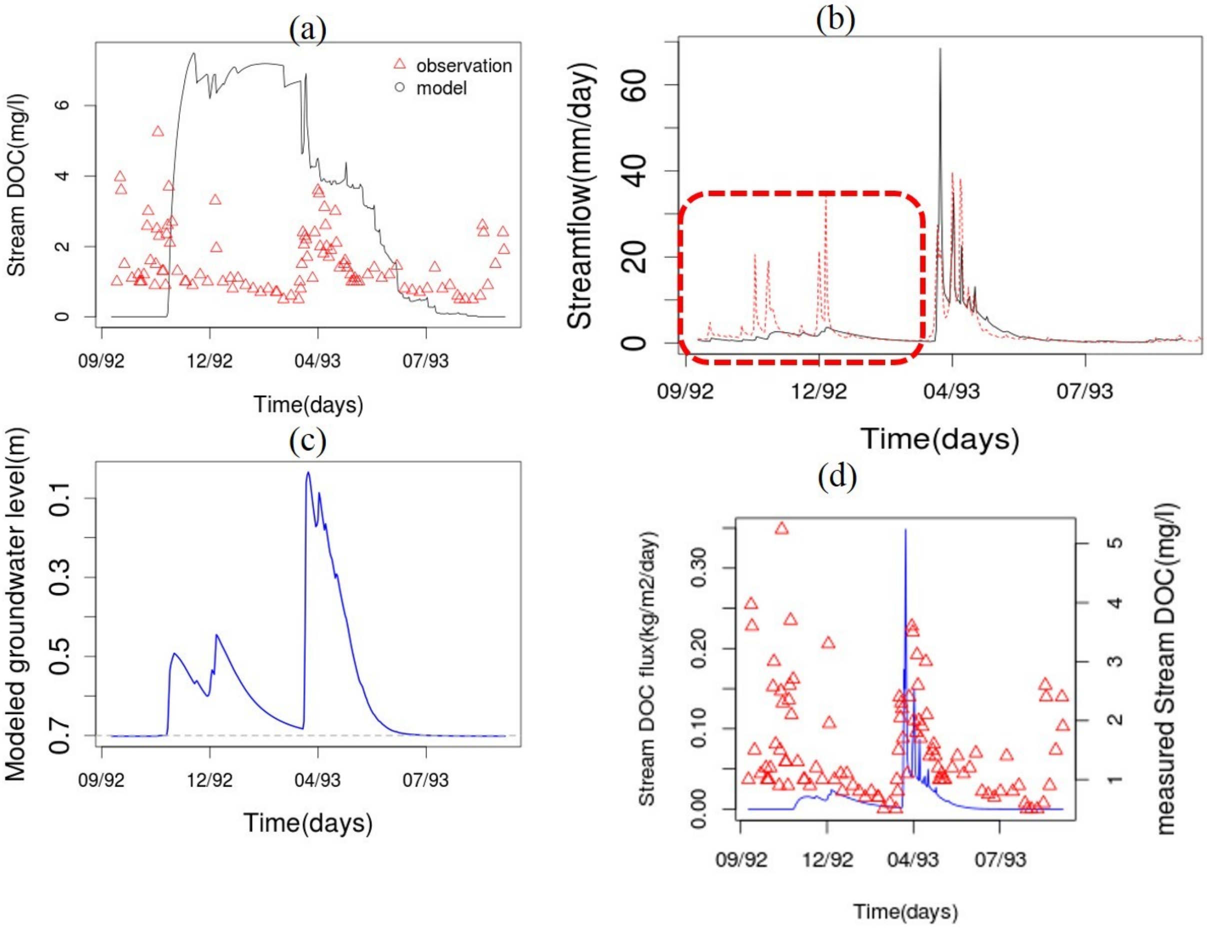


Figure 4.32. Model estimates of the calibrated RHESSys for Biscuit Brook in WY1993: (a) measured DOC concentration samples vs RHESSys estimates of stream DOC, (b) streamflow estimate (black solid line) vs measured streamflow (red dotted line), (c) watershed averaged groundwater estimates of RHESSys and (d) measured DOC concentration samples (red triangle) vs modelled DOC fluxes (blue solid line).

To improve RHESSys predictions of fall streamflow and stream DOC, we evaluated model sensitivity to soil depth. We preserved the spatial variation in soil depth, but reduced these by percentages ranging from 10% to 90%, with other soil parameters unchanged. We hypothesize that shallower soil depth will increase the peak flow estimates in fall and early winter, and that improved hydrologic predictions lead to better stream DOC predictions. In general, reducing soil depth increased peak flow as well as stream DOC prediction. Figure 4.33 shows the streamflow and DOC prediction with 10% of calibrated soil depth with fixed other calibrated soil parameters values. The results showed that peak flow is improved with shallower soil depth, and improved peak flow simulation lead to better capturing the measured DOC concentration and fluxes. Model predictions of DOC fluxes (R^2=0.30) are better than those of DOC concentration (R^2=0.04). These result suggest that estimate of DOC concentration is very sensitive to flow variability.

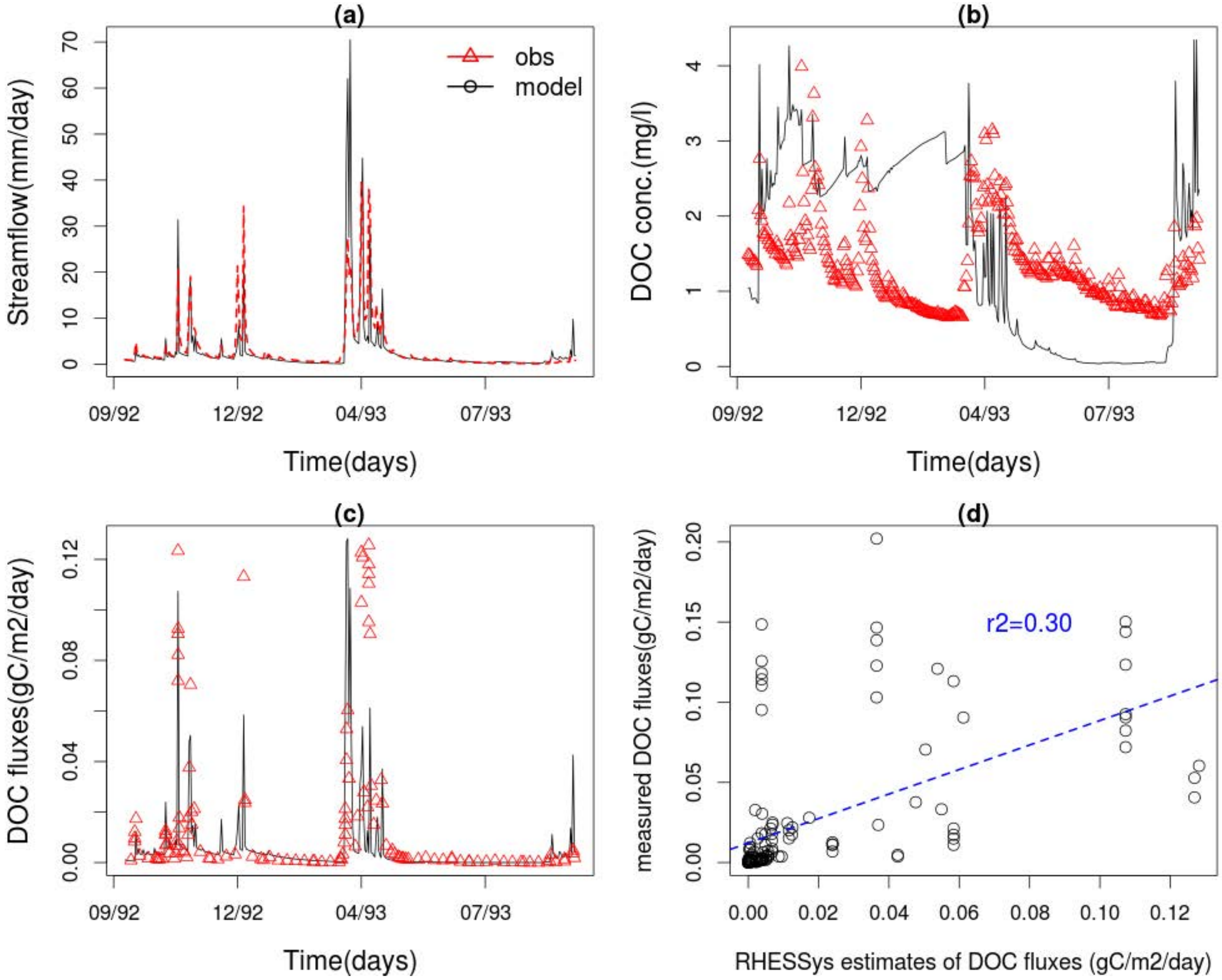


Figure 4.33. RHESSys streamflow predictions and DOC fluxes with shallower soil depths for Biscuit Brook: (a) daily streamflow, (b) model estimates of DOC concentration (black solid line) vs measured stream DOC concentration (red triangle), (c) modeled DOC fluxes (black solid line) vs measured stream DOC fluxes (red triangle), and (d) statistical relationship (R2) between measured stream DOC fluxes and modeled stream DOC fluxes.

The calibrated model was also used to compare the model estimates of LAI with Landsat TM-based LAI estimates (Figure 4.34). Average model estimates of LAI are lower than average TM-LAI values, and the two estimates have different spatial patterns. The RHESSys estimate of LAI showed that high flow accumulation areas tend to have higher LAI. However, the spatial patterns of TM-LAI values tend to reflect difference of aspect within the watershed. The sensitivity of LAI estimates to soil depth showed that lower LAI values of RHESSys estimates are due to shallower soil depth (not shown here). To correct the magnitude of LAI and its distribution in RHESSys simulation, we may consider improving soil depth parameterization. A suggestion is to parameterize the soil depth according to vegetation index (LAI, EVI and NDVI, etc.), and tree biomass. We hypothesize that deeper soil supports taller, larger trees. Turner *et al.* (2003) used the calibration of soil depth in Biome-BGC modeling to enable simulated LAI values to match with remotely sensed-based LAI values. In this study, LAI values were estimated using an empirical calibrated equation of NDVI and LAI developed in other studies (Hwang, unpublished data). Therefore, in order to improve the estimate of LAI, field measurement of LAI will be necessary. Recent studies (Ganguly *et al.*, 2008, 2012) show the potential utility of using radiative transfer theory of canopy spectral invariant to estimate the LAI at the Landsat TM scale (30 m). In addition, the MODIS LAI 1km product can be used to generate the variation of LAI over time, and these data will be very useful to diagnose RHESSys performance.

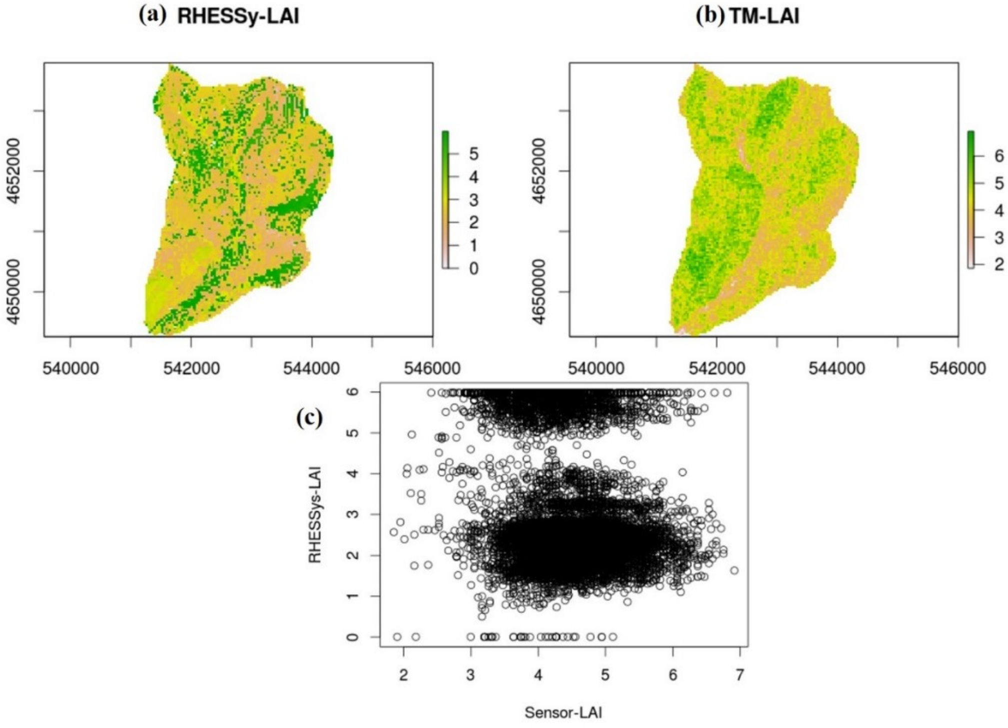


Figure 4.34. Comparison of RHESSys estimates of LAI and Landsat TM based LAI: (a) RHESSys estimates of LAI, (b) Landsat TM-based LAI estimates, and (c) the scatter plots of the two estimates.

4.3.5. Summary and conclusion

RHESSys was tested to predict streamflow, stream nitrate, and DOC concentration and fluxes for Shelter Creek and Biscuit Brook in the Neversink River Basin. The model results for Shelter Creek showed that the current model was only able to accurately reproduce streamflow at a monthly or longer time scale, and tended to overestimate the summer nitrate. The overestimation of summer nitrate may be due to the underestimation of denitrification in the groundwater storage. The model also failed to produce the seasonal variability of DOC concentration, even though the magnitude of DOC estimates is within the ranges of measured values. The RHESSys estimate shows a low temporal variability of soil DOC, suggesting that stream DOC processes is dominated by hydrologic transport mechanism. For Biscuit Brook, RHESSys-predicted streamflow showed a larger error in fall and early winter, and its error also leads to poor prediction of DOC concentration in the stream. Through lowering soil depth for the calibrated soil parameters, streamflow estimates in fall and early winter were improved, and its improvement also leads to better prediction of stream DOC concentration and fluxes. However, the model simulations with the shallow soil depth cause the underestimation of LAI values, and different spatial patterns, compared with remotely sensed based LAI values. The LAI estimated by RHESSys is sensitive to soil depth. We therefore suggest that vegetation properties (leaf area and biomass, etc.) can be used to guide the soil depth parameterization, and perhaps improve LAI estimates.

4.4. Impact of forest harvesting on streamflow in Neversink Reservoir streams

4.4.1. Introduction

Forest harvesting is a common forest management practice, but long-term consequences of forest harvesting to water yield and water quality are not consistent since harvesting may impact hydrology and forest ecosystem differently depending on the extent and intensity of forest harvesting, climate variability and natural forest succession in the recovery period. In the 1990s the USGS implemented an ambitious project in the Neversink Reservoir watershed to assess the long-term consequences of forest harvesting. A set of experimental catchments along Shelter Creek in the upper Neversink basin were established. These catchments were gaged, differing intensities of forest harvesting were applied, and the subsequent effects on stream chemistry were evaluated (Siemion *et al.*, 2011). Here, we revisit the Shelter Creek experimental watersheds to examine the effects of forest harvesting on streamflow, and to assess:

1. the recovery of the forest canopy following varying harvest intensities using remotely sensed data to derive vegetation indices,
2. the effect of forest harvest on streamflow using a paired (harvested/control) watershed approach.

4.4.2. Study sites

The study site is the Frost Valley YMCA Model Forest, located in the Neversink River Basin. Three areas within the site experienced different forest harvesting practices: light selective cut, heavy selective cut, and clear-cut. Figure 4.35 and Table 4.7 show the area of forest harvesting in Shelter Creek watersheds and the different harvesting history for each sub-watersheds, including CL25, DC57, NS25 and SS20. Figure 4.35b shows the forested areas harvested by USGS (McHale *et al*., 2008). This map is used as a reference when we analyze the forest harvesting areas using remotely sensor-based vegetation indices. SC20 is the entire Shelter Creek watershed. NS25 is ‘light selective harvest’ watershed with 2% basal area decrease after the harvest. SC20 is ‘heavy selective harvest’ watershed with 8% basal area decrease after the harvest. DC57 is ‘clear cut harvest’ watershed with 80% decrease in the basal area. CL25 is used as a control watershed. A detailed description of the forest harvesting history is provided in McHale *et al*., 2008.

Table 4.7. Shelter Creek sub-watersheds with different forest harvesting regimes and their harvesting periods.

Watershed	**Size (km^2)**	**Harvesting Period**	**Intensity**
CL25	0.52	---	Control (0%)
NS25	0.34	Winter, 1995-96	Light harvest (2% decrease in basal area)
SS20	0.43	Fall, 1996	Heavy harvest (8% decrease of basal area)
DC57	0.24	Winter, 1996-97	Clear Cut (97% cut, 90% decrease of basal area)
SC40	1.0	Winter, 1995-96 and Fall, 1996	NS2 and SS20 are the sub-watersheds of SC40
SC20	1.6	Winter, 1995-96 and Fall, 1996, Winter, 1996-97	DC57, NS2 and SS20 are sub-watersheds of SC20

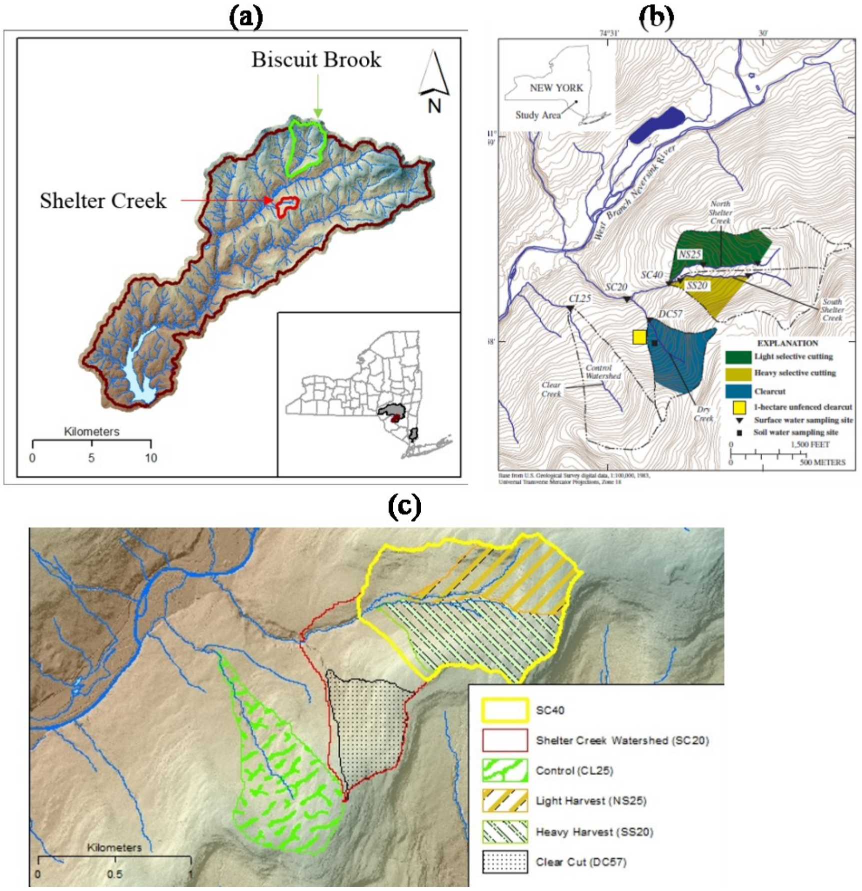


Figure 4.35. Shelter Creek watersheds including control watersheds, and forest harvested watersheds: (a) the location of Shelter Creek and Biscuit Brook, (b) forest harvesting maps (McHale *et al*., 2008), and (c) the map of sub-watersheds with different forest harvesting history.

4.4.3. Recovery of the forest canopy following varying harvest intensities based on remotely-sensed data derived vegetation indices

The recovery of the forest canopy following varying harvest intensities is assessed using remotely-sensed data derived vegetation indices. Landsat-TM images were used to derive three vegetation indices: normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized different moisture index (NDMI). Image analysis was conducted using GRASS 7.0 GIS imagery tools. The calculation of the vegetative indices is given by Equations (4.15), (4.16), and (4.17):

$$NDVI = \frac{NIR - VIS}{NIR + VIS} \tag{4.15}$$

$$EVI = 2.5 \times \frac{(NIR - RED)}{(NIR + 6 \times RED - 7.7 \times BLUE + 1.0)} \tag{4.16}$$

$$NDMI = \frac{NIR - MIR}{NIR + MIR} \tag{4.17}$$

where NIR is near infrared light (0.7 to 1.1µm wavelength), VIS is visible light (0.4 to 0.7µm), RED is red light (0.63-0.69 µm), BLUE is blue-green light (0.45-0.52 µm), and MIR is mid-infrared (1.55-1.75 µm).

Figure 4.36 shows the calculated NDVI, EVI and NDMI values in the period of 08/09/1995. NDVI values tend to have higher values than the other two indices, and the spatial patterns of each index are different. NDVI is sensitive to the density of leaf greenness. EVI is similar to NDVI but is preferred in that NDVI is easily saturated in dense forest environments. EVI also corrects for some distortion of light by air particles and the ground cover below the vegetation. Figure 4.36(b) shows several long strips along slopes and the strips related to exposed boulders. Figure 4.36(c) show the NDMI values. NDMI is sensitive to leaf water content. NDMI has the lowest values compared with the two indices (Figure 4.36d). EVI has the highest spatial variance, but NDVI has lowest spatial variance. Since each vegetation index shows different features of vegetation, we consider which vegetation index better reflects the hydrologic changes (evapotranspiration and streamflow) both in the forest harvesting periods and forest recovery periods.

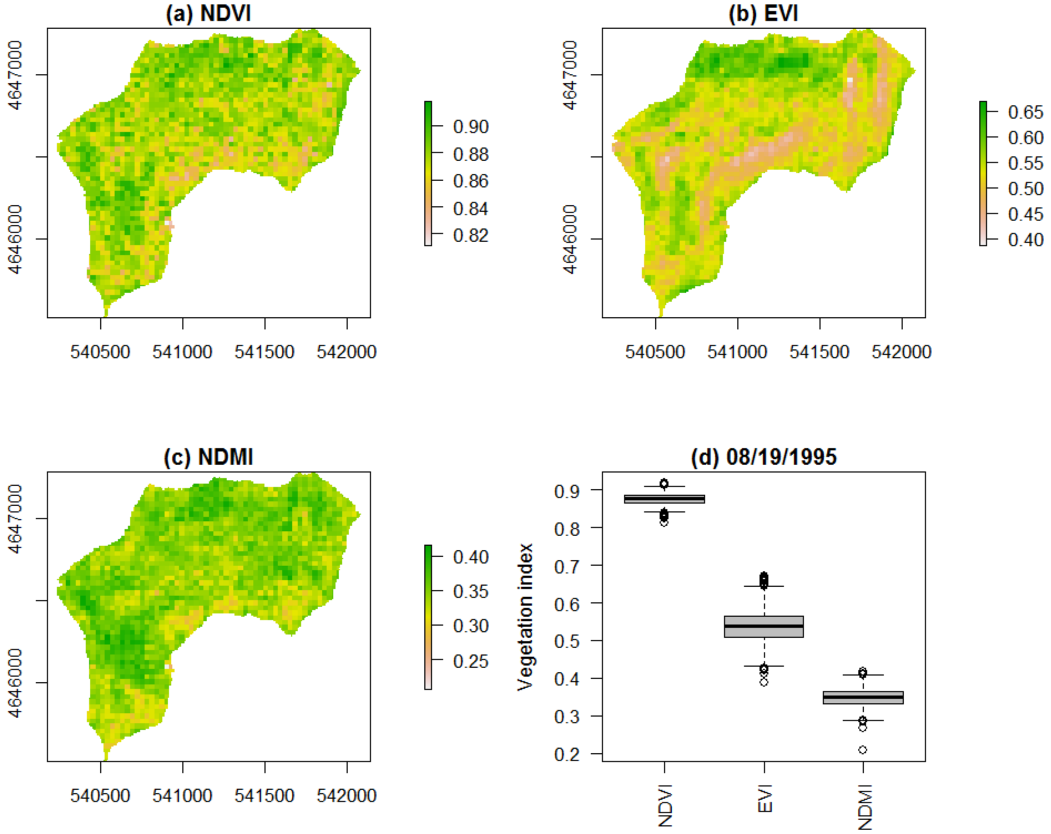


Figure 4.36. Calculated vegetation index: (a) NDVI, (b) EVI,(c) NDMI, and (d) summary of each vegetation index

4.4.4. Recovery after light forest harvesting

Three vegetation indices, NDVI, NDMI and EVI were used to detect the light forest harvested areas in NS25 and the pattern of recovery. Light harvesting was conducted in the winter of year 1995. Since the light harvested area also included the downstream area of NS25 as well as the areas within NS25, the indices values for SC40 was calculated instead. To detect the light forest harvested area and its recovery pattern, we used the two TM images collected on June 12, 1995 and June 2, 1996 to calculate the three indices per each date. Figure 4.37 (a), (d) and (g) show the vegetation indices values in early summer of year 1995 prior to forest harvesting. Figure 4.37(b), (e) and (h) show the vegetation indices in early summer of year 1996 after the light forest harvesting. NDVI ranges from 0.85 to 0.91 for 1995, and from 0.72 to 0.86 for 1996. The reduction of the NDVI values was only shown in the upper part of NS25, where forest harvesting was not conducted. However, the NDVI values of the harvested areas between the two dates is very similar. These patterns were observed in the other two indices. Therefore, we can conclude that the forest recovered in the light forest harvested areas within a year.

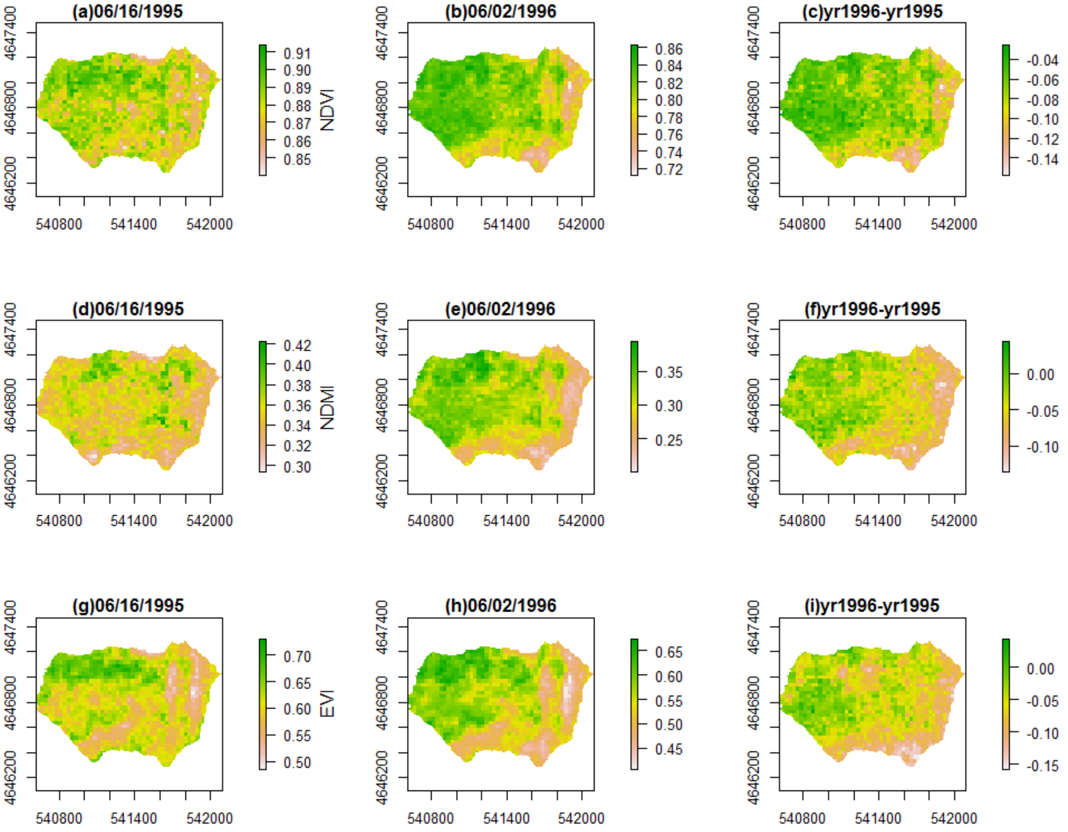


Figure 4.37. Calculated NDVI, NDMI and EVI for SC40 before and after light forest harvesting: (a) 06/16/1995, (b) 06/02/1996 , (c) the difference of the two images , (d) 06/16/1995, (e) 06/02/1996, (f) the difference of the two images,(g) 06/16/1995, (h) 06/02/1996 and (i) the difference of the two images.

4.4.5. Recovery after heavy partial forest harvesting

We explored the heavy forest harvested areas in SS20 using the three indices. Since SS20 had heavy partial forest harvesting in the fall of year 1996, we used three TM images collected at on October 3, 1994, October 24, 1996 and October 11, 1997. Figure 4.38 shows the calculated NDVI values at the three dates. The NDVI values of the lower parts of SS20 distinctively reduced, and the areas with the low values matched with existing forest harvesting map (Figure 4.35b). Figure 4.38c shows the difference of the NDVI values at the two dates (October 3,1994 and October 24, 1996) has the negative values (<-0.3) in the lower part of SS20. NDVI values at the October 11, 1997 shows positive values (>0.2) compared with the NDVI values at the October 24, 1996. These results are supported by the report of McHale *et al.* (2008). However, the NDVI values on October 11, 1997 still showed lower values than those at the October 3th, 1994. It is not clear that forest leaf fully recovered from the harvesting, or that the lower values of NDVI on October 11, 1997 can be attributed to climate difference. The analysis of using EVI values for the three dates are similar (Figure 4.39). The EVI values on October 24, 1996 had lower and negative (<-0.3).values, compared with the EVI values at the October 3, 1994 (Figure 4.39c), and the distinctive increase of EVI values was shown in the lower part of SS20 on October 11, 1997. However, the analysis of NDMI showed different results (Figure 4.40). Even though NDMI values on October 24, 1996 were lower than those on October 4, 1994, the lower part of SS20 had still very similar values between the two dates. These results suggest that NDVI and EVI may show better performance to detect the heavy forest harvesting area, and recovery patterns.

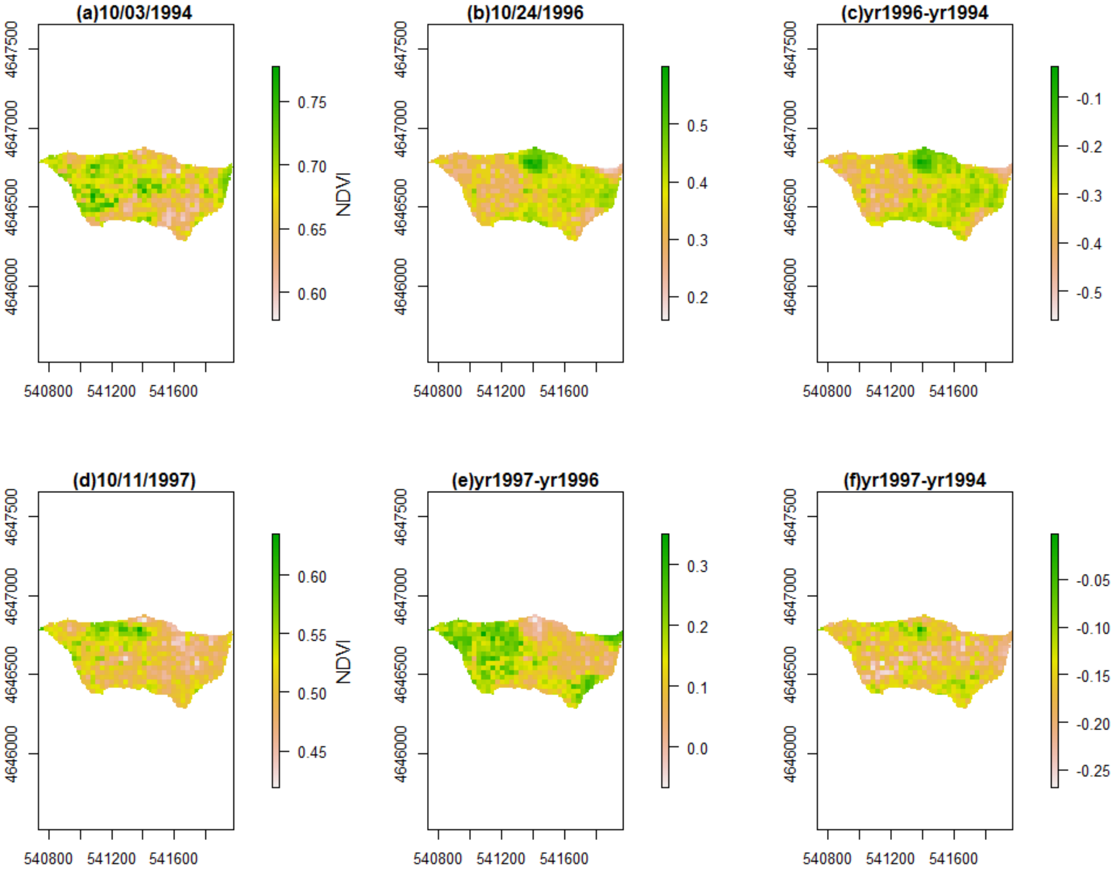


Figure 4.38. Calculated NDVI values for SS20: (a) 10/03/1994, (b) 10/24/1996 and (c) the difference of the indices at the two dates (10/3/1994 and 10/24/1996), (d) 10/11/1997, (e) the difference of the indices at the two dates (10/11/1997 and 10/24/1996), and (f) the difference of the indices at the two dates (10/11/1997 and 10/003/1994)

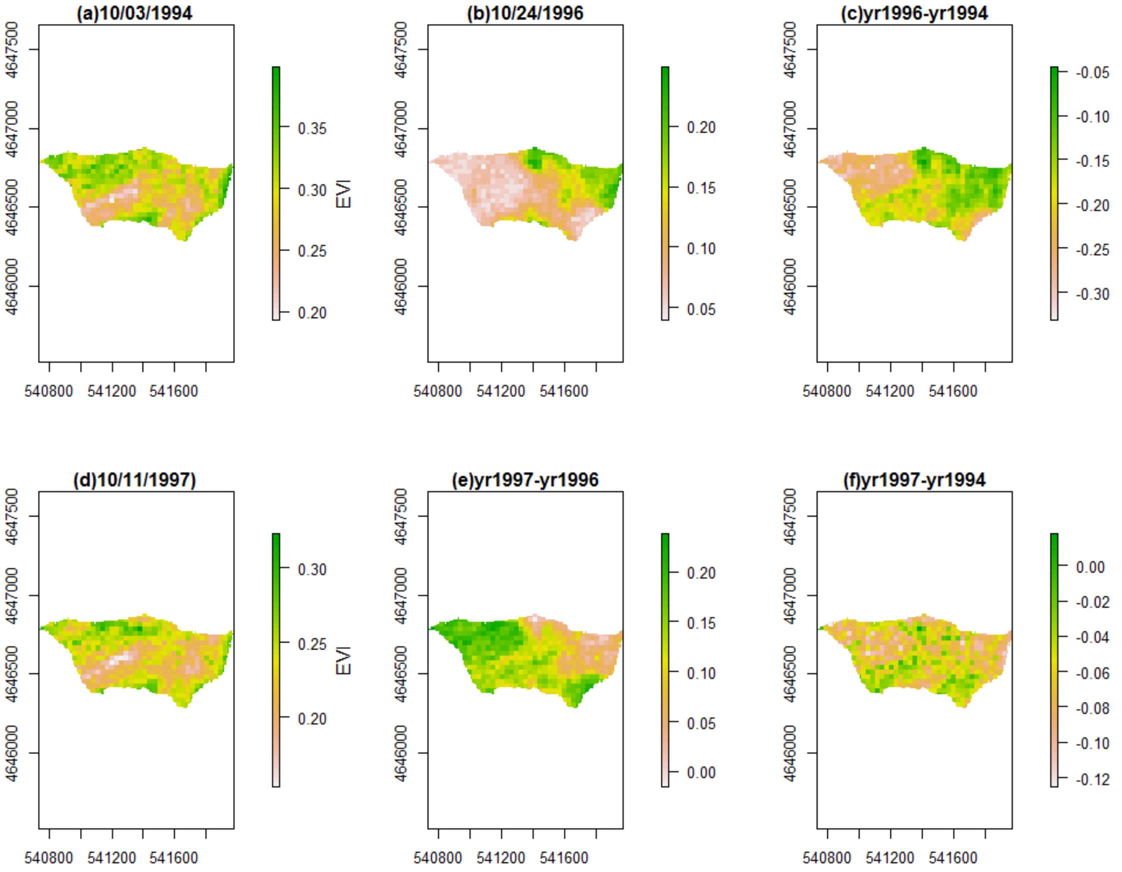


Figure 4.39. Calculated EVI values for SS20: (a) 10/03/1994, (b) 10/24/1996 and (c) the difference of the indices at the two dates (10/3/1994 and 10/24/1996), (d) 10/11/1997, (e) the difference of the indices at the two dates (10/11/1997 and 10/24/1996), and (f) the difference of the indices at the two dates (10/11/1997 and 10/003/1994).

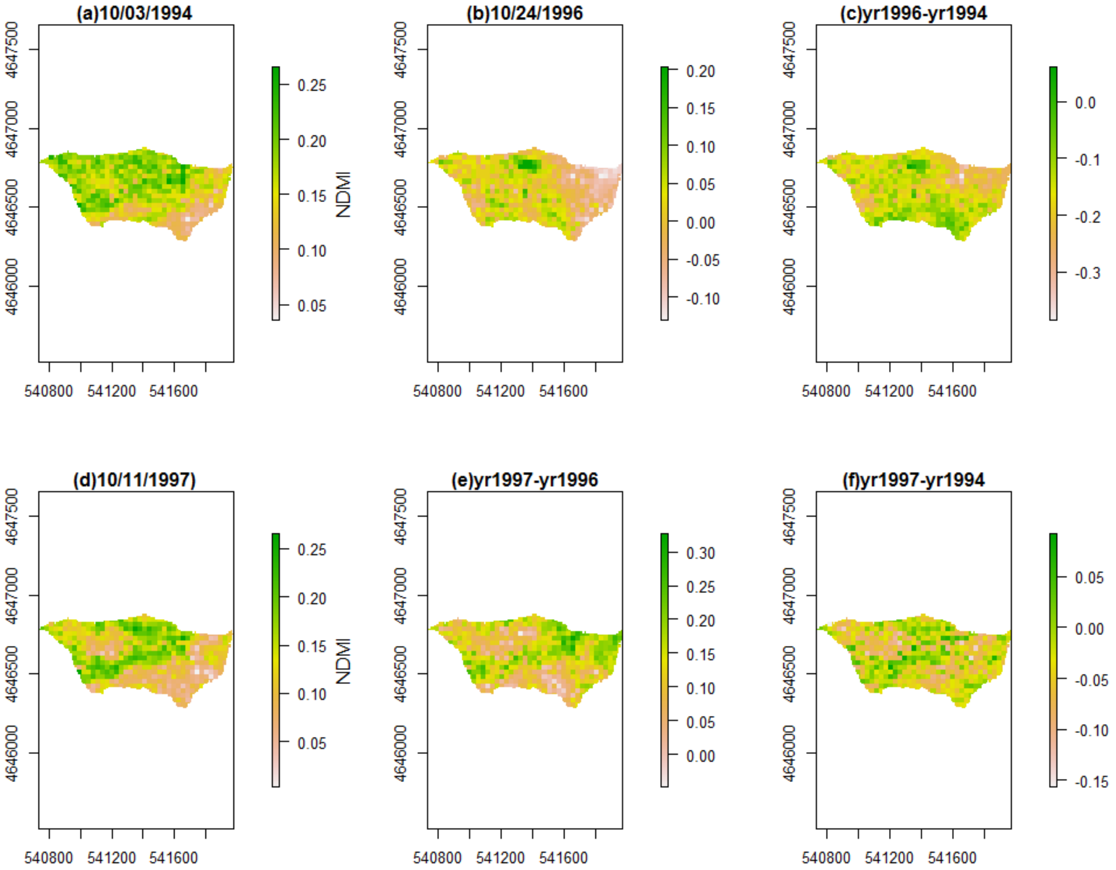


Figure 4.40. Calculated NDMI values for SS20: (a) 10/03/1994, (b) 10/24/1996 and (c) the difference of the indices at the two dates (10/3/1994 and 10/24/1996), (d) 10/11/1997, (e) the difference of the indices at the two dates (10/11/1997 and 10/24/1996), and (f) the difference of the indices at the two dates (10/11/1997 and 10/003/1994).

4.4.6. Recovery after clear-cutting

We explored forest clear-cut areas in DC57 using the TM images collected in summer from 1996 to 2003 (Figure 4.41). In the lower part of DC57, the three indices values had very low values in 1997 compared to 1996. The areas with the low values were consistent with the forest harvesting map (Figure 4.35b). In 2000, the three indices values are likely close to the values prior to the clear-cut harvesting. In the year 2003, the three indices showed that the lower part of DC57 (clear-cut sites) had higher values than the upper parts of DC57, which is similar to the pre-harvesting conditions (Figure 4.41a, Figure 4.41b and Figure 4.41c).

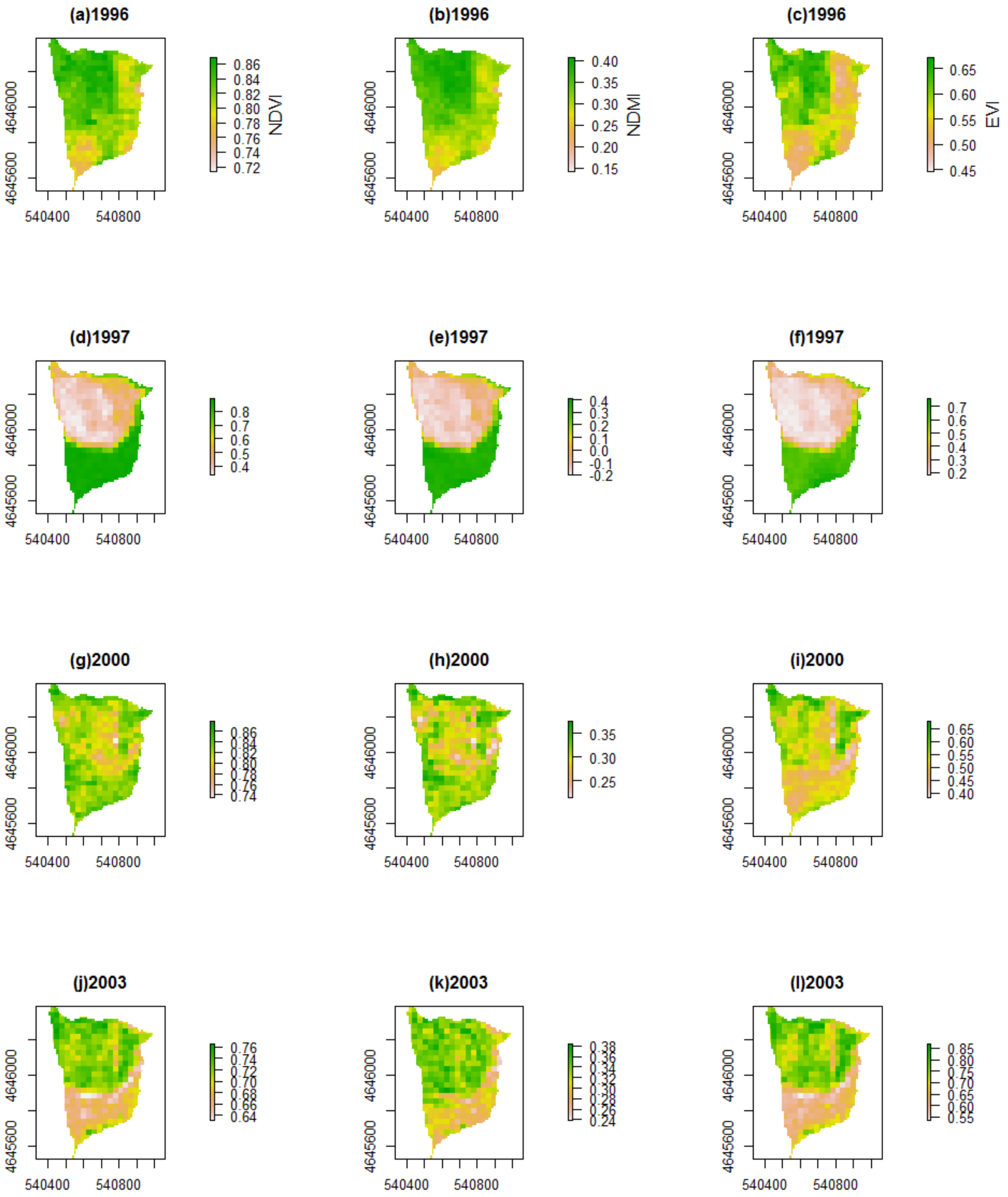


Figure 4.41. Calculated NDVI, NDMI and EVI for DC57 using summer TM images from 1996 to 2003.

4.4.7. Temporal changes in the three vegetation indices

We explored the temporal change of the three indices in CL25 (control site), DC57 (clear-cut site), and SC20. Figure 4.42 shows summer NDVI, NDMI and EVI values of CL25 from 1995 and 2003. Three indices show similar temporal patterns in most years, but in 2002 and 2003, NDVI and NDMI values became lower than year 2001, but the EVI values in the two years became larger than in 2001. In the clear-cut site (DC57) the three indices decreased in 1997, and thereafter increased (Figure 4.43). The NDMI values 2002, and the EVI values in year 2000, are close to the values in year 1995, respectively. After forest clear-cuts in DC57, the NDVI values increased until 2002, then decreased again in 2003. The NDVI values did not fully recover to the levels of 1995. These patterns are also found in CL25 (Figure 4.42) and SC20 (Figure 4.44). These results are similar to a study of Khomik *et al.* (2014). Their study showed the quick recovery of young trees in small study sites in the Harvard Forest Long-Term Ecological Research sites in Petersham, MA after forest clear cut; LAI value also quickly increases.

Forest regeneration following forest clear-cut also changed the forest stand structure as well as dominant forest species. Figure 4.45 shows the 2016 forest stand structure and species in the unharvested area, and clear-cutting site. The unharvested areas have red maples, yellow birch, black cherry, and beech but the regenerated forest compositions of the clear-cut sites are now dominated by pin cherry and yellow birch. The clear-cut sites have denser vegetation structures than the unharvested sites. We expect that the regenerated forest will alter the magnitude and timing of forest water use and streamflow as well as water qualities in the stream, even if the forest leaf area recovers to the conditions prior to forest harvesting.

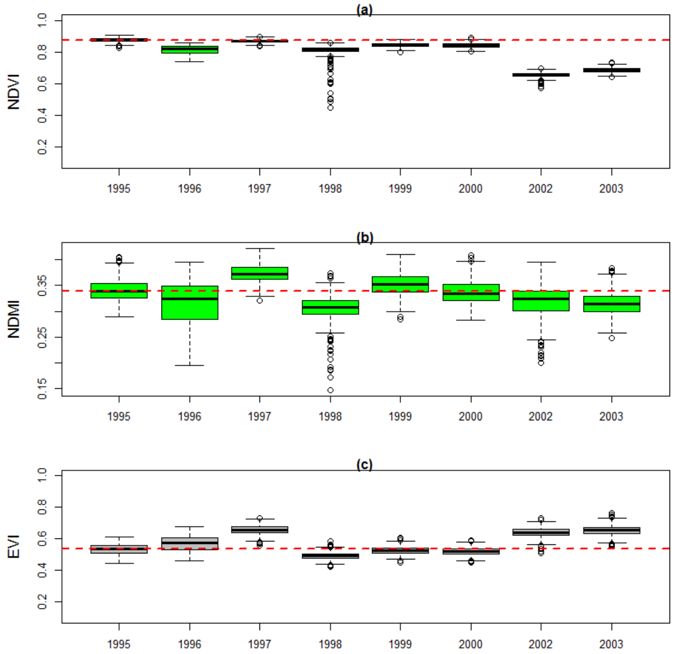


Figure 4.42. Calculated NDVI, NDVMI and EVI values of CL25 using Landsat TM images collected in summer of year 1995 to 2003: (a) NDVI, (b) NDMI and (c) EVI.

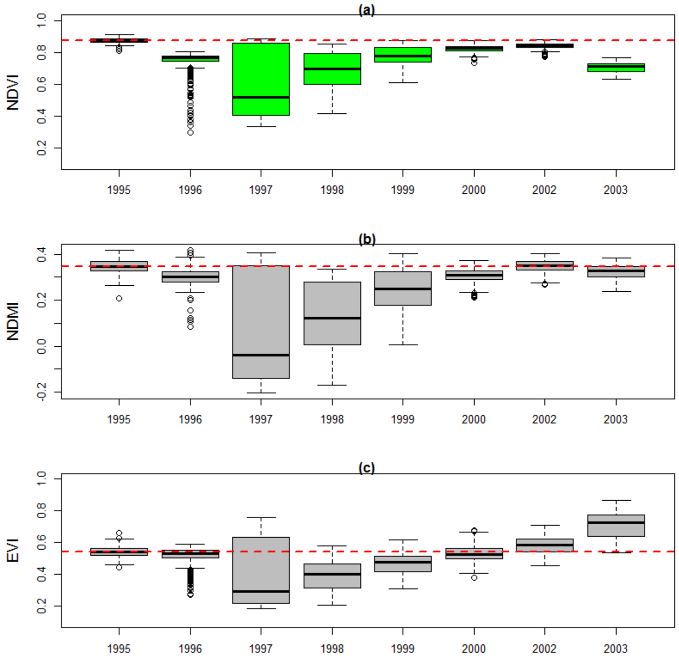


Figure 4.43. Calculated NDVI, NDMI and EVI for DC57 using Landsat TM images collected in summer of year 1995 to 2003: (a) NDVI, (b) NDMI and (c) EVI.

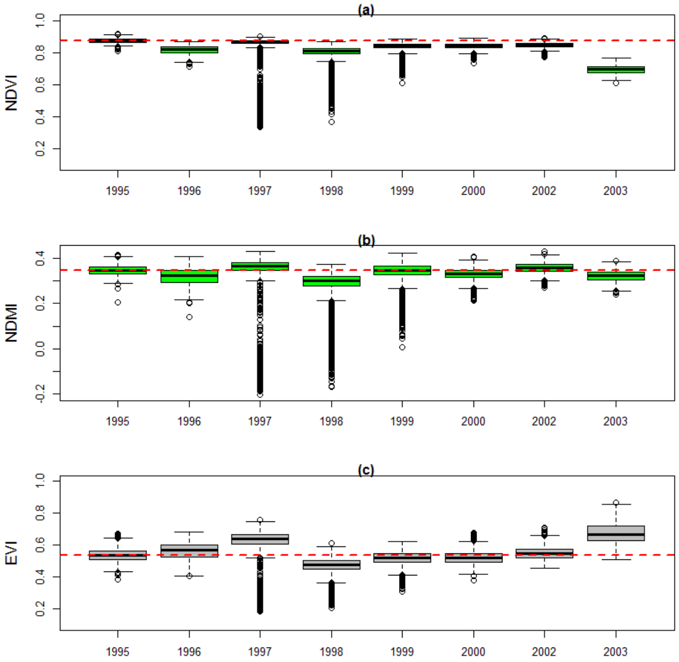
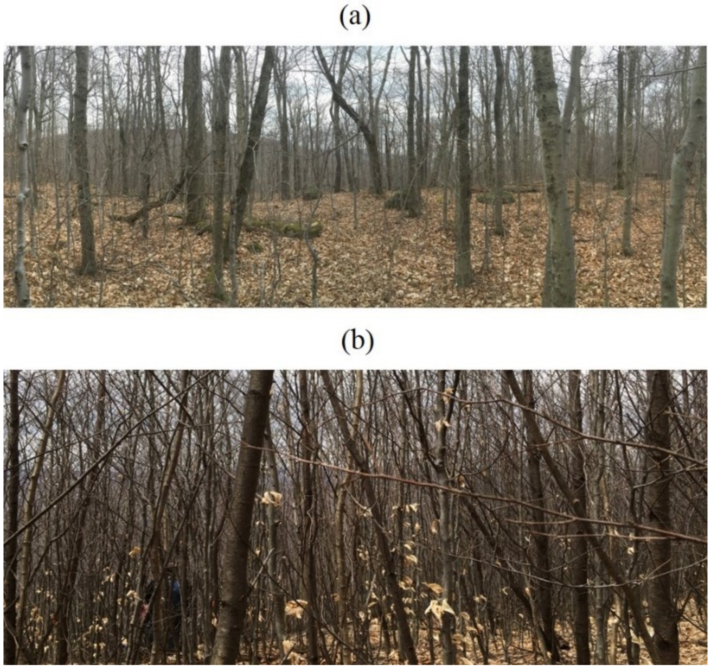


Figure 4.44. Calculated NDVI and EVI for SC20 using Landsat TM images collected in summer of year 1995 to 2003: (a) NDVI, (b) NDMI and (c) EVI.


(a)

(b)

Figure 4.45. Changes of forest structure and species after clear cut forest harvesting: (a) unharvested area in DC57, and (b) clear cut sites. This picture was taken in the year 2016.

4.4.8. Effect of forest harvest on streamflow using a paired (harvested/control) watershed approach

The paired watershed approach uses two hydrologically similar watersheds, a disturbed watershed and an undisturbed watershed. Biscuit Brook was selected as a control watershed because Biscuit Brook has not experienced forest harvesting practices and other forest disturbances, and has gaged streamflow data. We developed a linear regression model using measured streamflow data (10/1/1992 to 09/30/1995) at Biscuit Brook and SC20 prior to forest harvesting (Figure 4.46). This regression model is used to represent the streamflow variability of SC20 in the unharvested conditions from 10/1/1995 to 10/1/2003. The impact of forest harvesting on flow changes can be calculated by comparing the flow values of the developed regression model with the measured streamflow of SC20.

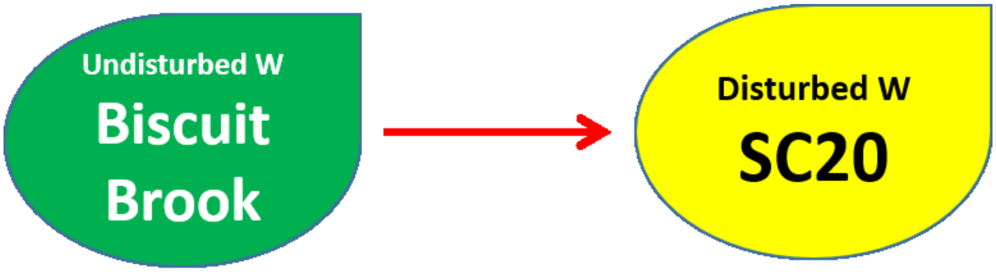


Figure 4.46. Paired watershed approach: streamflow of undistributed watershed (Biscuit Brook) is compared with streamflow of disturbed watershed (SC20).

Figure 4.47 shows the results of the developed regression model. The regression model shows high streamflow accuracies, with daily NSE of 0.92, and monthly NSE of 0.97. The regression model also captures the timing of the measured daily flow in SC20 (R^2=0.92). Observed low flows of SC20 tend to be lower than those of the regression model. Measured streamflow of SC20 in the calibration periods is about 140 mm larger than the measured streamflow of Biscuit Brook but the streamflow difference of the regression model and SC20 is less than 1%.

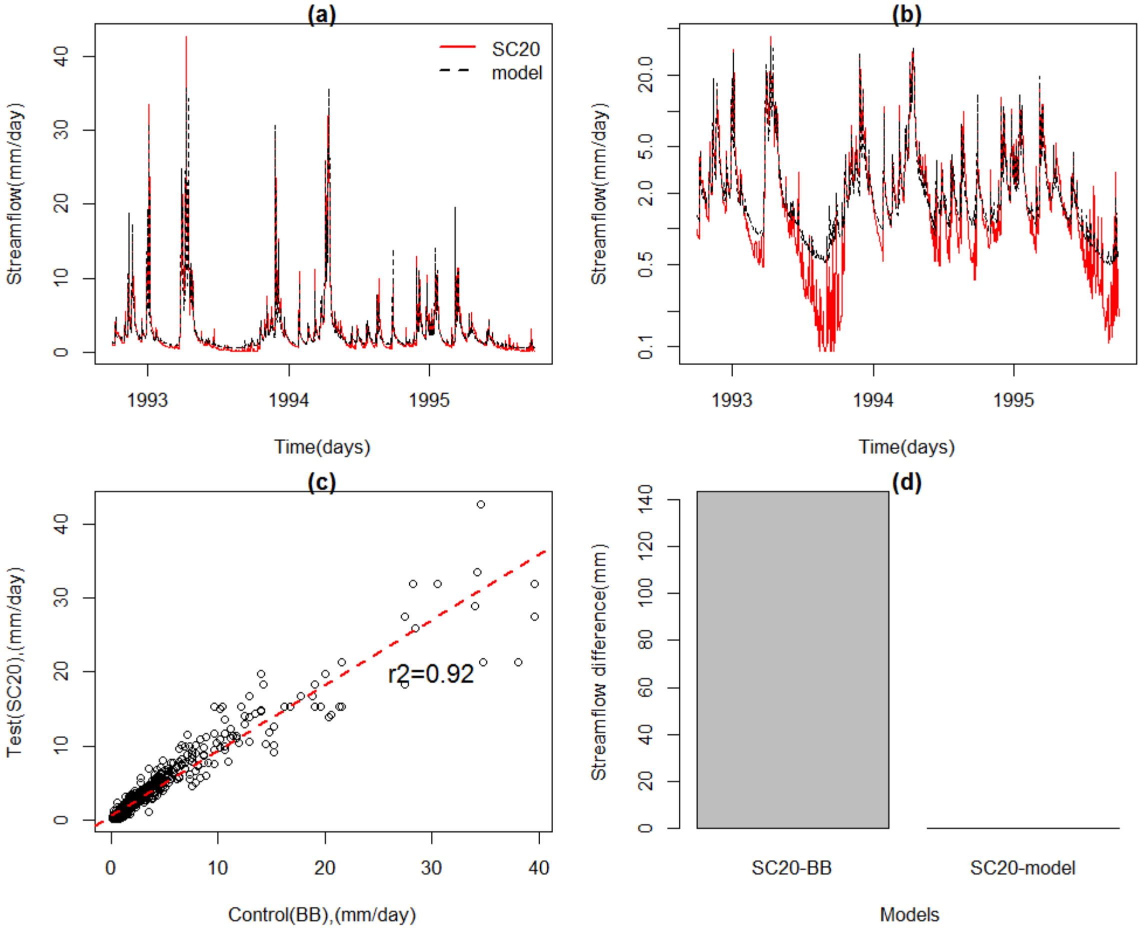


Figure 4.47. Comparison of predicted streamflow with measured streamflow of SC20 prior to harvesting: (a) comparison of modeled flow with measured streamflow of SC20, (b) comparison of modeled flow with measured streamflow of SC20 with log scale, (c) fitted streamflow of Biscuit Brook against the measured streamflow of SC20, and (d) the streamflow differences of Biscuit Book and SC20, and the difference of modeled streamflow of SC20 and measured streamflow.

This linear regression model was used to detect flow changes after the three different harvesting practices (Table 4.7). Here the regression model represents flow variability in the pre-harvesting conditions. Flow changes at SC20 were investigated at daily, monthly and annual time scales. Figure 4.48 shows the difference of daily streamflow between the regression model and SC20 in the light forest harvesting periods (winter 1995). SC20 tends to have higher peak streamflow, and lower baseflow, but it is difficult to conclude that these patterns are attributable to forest harvesting because the regression model showed mismatch of peak flows in some events and overestimation of baseflow in the calibration periods. However, we can observe the

R^2 value decrease from 0.92 to 0.81 and annual flow in the year 1996 increased to 118mm compared with the calibrated model. Figure 4.49 shows the difference of daily streamflow between the regression model and SC20 in the heavy forest harvesting periods (Fall of year 1996) and the clear-cut periods (Winter of year 1996-1997). The measured streamflow of SC20 shows higher peak flow and lower baseflow than those of the regression model. The flow timing of SC20 and the regression model is similar to pre-harvesting periods, with R^2 of 0.92, but annual flow of SC20 increases to 145mm, compared with the regression model.

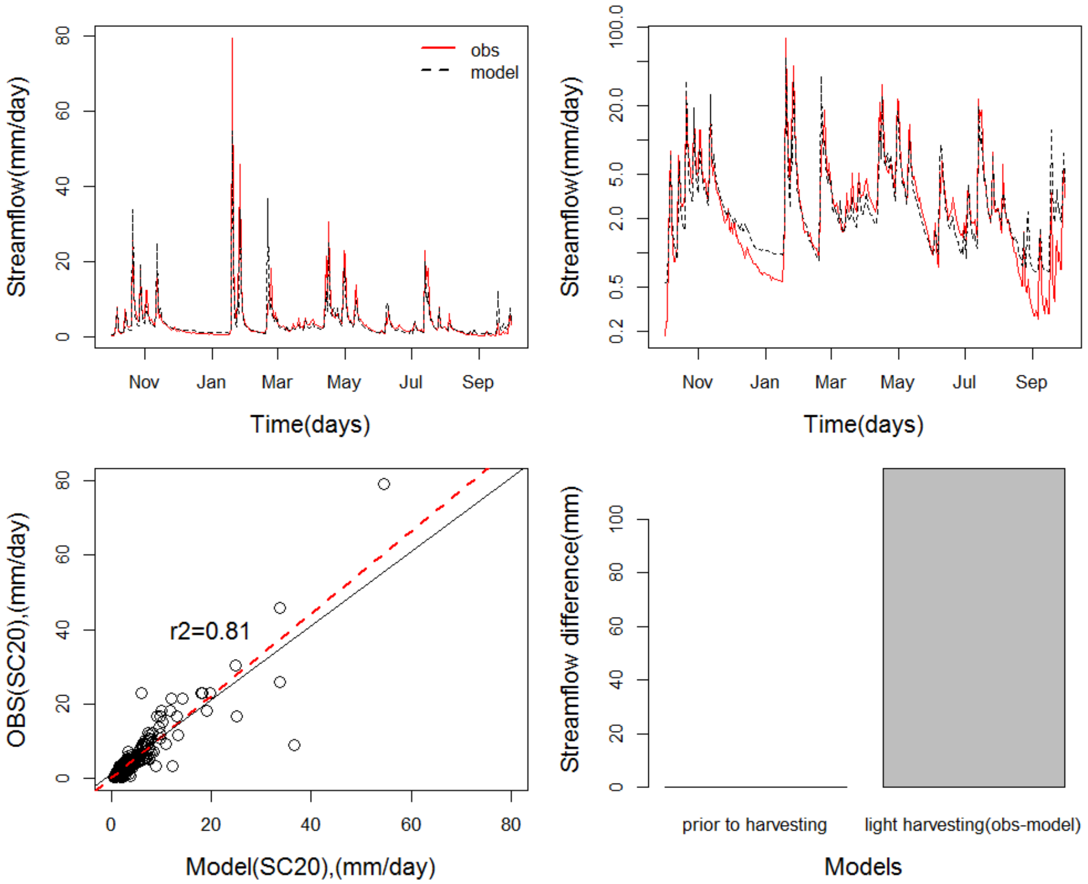


Figure 4.48. Flow change of SC20 following light selective harvesting: model is predicted streamflow based on measured streamflow of Biscuit Brook, and obs is measured streamflow at the outlet of SC20.

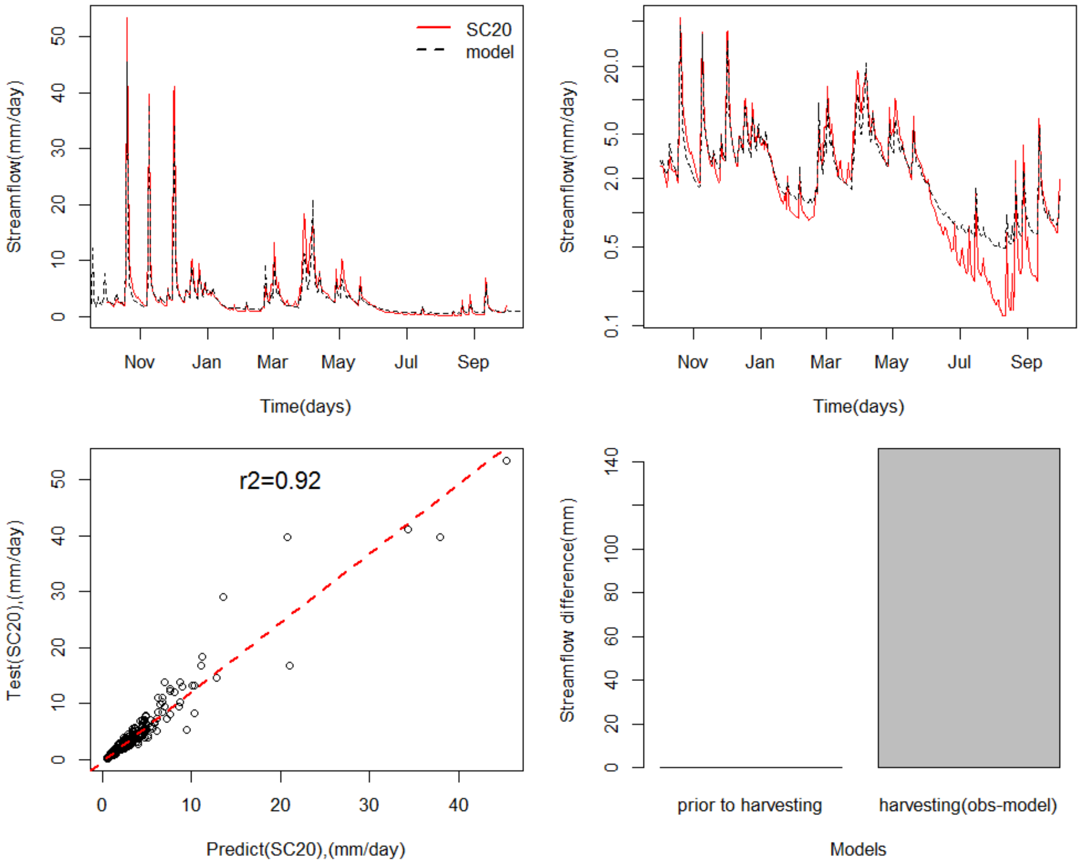


Figure 4.49. Flow change of SC20 following heavy selective harvesting in SS20 and clear- cut in DC57: model is predicted streamflow based on measured streamflow of Biscuit Brook, and obs is measured streamflow at the outlet of SC20.

We also explored the monthly flow changes following forest harvesting and forest recovery (Figure 4.50). According to the forest harvesting history and the analysis of vegetation indices, we divided the forest harvesting periods in the four periods, (1) the pre-harvesting periods (WY 1993 to 1995), (2) harvesting periods (WY 1996 to 1997), and (3) forest recovery periods (WY 1998 to 2000), and forest mature periods (WY 2001 to 2003). In the pre-harvesting periods, median of monthly flow difference between SC20 and the regression model is close to -4 mm. In the harvesting periods, median value of the streamflow difference is about 5 mm. In the forest recovery periods and mature periods, the median of the streamflow difference is 1 and -2 mm respectively. These results suggest that monthly flow in the forest mature periods is close to those in the pre-harvesting conditions. Figure 4.49b summarizes the annual flow changes before and after forest harvesting. In the pre-harvesting periods, the median of annual streamflow difference between SC20 and the regression model is close to -5 mm. In the harvesting periods, the median of the annual streamflow difference increases to 118mm. In the forest recovery periods, the median of the annual streamflow difference is about 178mm. However, in the forest mature periods, the median of annual streamflow difference is about 21mm. This result suggests that annual flow did not fully recover even if the forest leaf conditions (based on vegetation indices) recovered to the pre-harvesting conditions. This result can have two explanations. The first explanation is due to the limitation of the regression model. The annual flow estimates of the regression model are based on a simple assumption that the flow relationship between SC20 and Biscuit Brook in the calibration periods remains same in the un-calibrated periods. Therefore, the difference in annual flow between the pre-harvesting periods and the forest mature periods may attribute to the limitation of this model assumption. An alternative explanation is that remotely sensed vegetation indices may not fully capture the hydrological behaviors such as evapotranspiration.

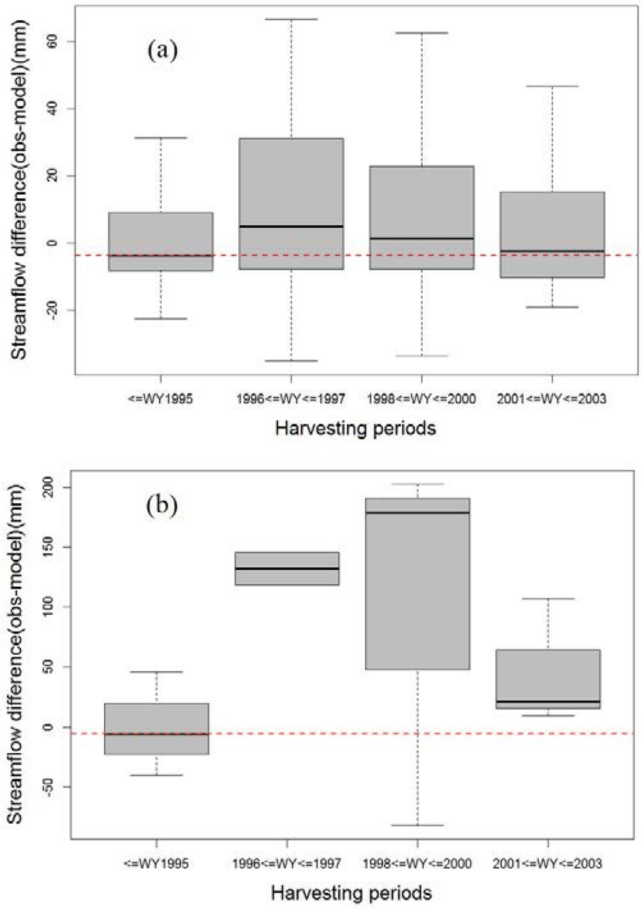


Figure 4.50. Changes in monthly flow and annual flow after forest harvesting completed: (a) monthly flow, and (b) annual flow. Model is the predicted streamflow of SC20 based on measured streamflow of Biscuit Brook, and obs is the measured streamflow at the outlet of SC20

4.4.9. Summary and conclusion

This analysis focused on detecting forest recovery patterns using remotely sensed vegetation indices, and improving our understanding of the relationship between forest changes and streamflow using paired watershed approach. Landsat TM images were used to calculate three different vegetation indices (NDVI, NDMI and EVI), and these indices were analyzed to detect forest recovery patterns after harvesting. Each index showed different sensitivity to forest harvesting and forest recovery pattern. NDMI is the least sensitive to forest changes, while EVI is most sensitive. For example, NDMI values cannot detect the heavy forest harvesting patterns in SS20. The derived vegetation indices can detect the forest recovery patterns following the three forest harvesting practices. The derived vegetation indices quickly recovered within a year after partial forest harvesting. For DC57 (clear cut site), recovery pattern of the forest varies among the three indices. The analysis using the NDMI and EVI values suggests that the forest fully recovered by 5 years and 3 years, respectively. However, the analysis of using NDVI values suggest that the forest did not fully recovered by year 2003.

The paired watershed approach using a linear regression model was used to examine the impact of forest harvesting on streamflow. This approach assumes that the difference of the modeled streamflow and measured streamflow includes the impact of forest harvesting and inherent model uncertainty. The regression model showed that forest harvesting increased the annual and monthly flow, which was generally expected due to reduction in evapotranspiration. In the forest recovery periods, the annual and monthly flow decreased. However, the regression model showed that annual flow did not fully recover to the conditions that existed prior to harvesting even when the forest leaf (based on vegetation indices) recovered to the pre-harvesting conditions. However, the monthly flow did recover to nearly pre-harvesting conditions. These results can be interpreted in two ways: (1) the assumption in the paired watershed approach that the flow relationship between SC20 and Biscuit Brook determined for the calibration periods remains same in the uncalibrated periods may not be valid, and (2) the remotely-sensed vegetation indices do not fully represent hydrological changes of the watersheds, such as evapotranspiration, that are associated with shifts in forest stand structure and species.

In addition to the impact of forest harvesting on streamflow changes, it was observed that forest harvesting changed the biogeochemistry of soils and stream (Siemion *et al.*, 2011). These changes will be critical to maintain high quality water in streams. Future research will use the paired-watershed approach to quantify the impact of forest harvesting on dissolved organic carbon and nitrate in streams.

4.5. Preliminary testing of a turbidity model for Neversink Reservoir

The two-dimensional model for temperature and turbidity for Neversink Reservoir is based on CE-QUAL-W2 (Cole and Wells 2013). The turbidity submodel has previously undergone development, testing, and application for Schoharie, Ashokan, Kensico and Rondout reservoirs during earlier Catskill Turbidity Control studies (Gelda and Effler 2007, Gelda *et al*. 2009, 2012, 2013). The Water Quality Modeling has plans to extend the testing and application of this model to the remaining West of Hudson reservoirs, these being Cannonsville, Neversink, and Pepacton. This section describes application and testing of the model for Neversink Reservoir. Model calibration for 2015 and validation for 2012–2014 is described here. This effort is preliminary; further testing of the model will be conducted as more data become available. The model may be integrated into OST in the future.

4.5.1. Input Data

The bathymetric information needed for the model was developed from a survey conducted by USGS. The reservoir was represented by a grid of segments and layers in the model, as summarized in Table 4.8.

Table 4.8. Neversink Reservoir: specification of model grid and hydraulic features.

Parameter	Value	Intake/Release	Elevation[*]
No. of branches	1	NR1	1314-1328 ft.
No. of active segments	19	NR2	1344-1358 ft.
Maximum No. of active layers	54	NR3	1374-1388 ft.
Average length of a segment	530 m	NR4	1404-1418 ft.
Height of a layer	1 m		
Spillway elevation, NAVD88	438.67 m (1439.21 ft.)	Dam Release	1314 ft.
Spillway length	182.88 m (600 ft.)		

[*] *BWS datum*

The primary tributary of the reservoir, Neversink River, is gauged at Claryville, NY, capturing runoff from ~ 74% of the watershed. Outflow from the reservoir occurs via an aqueduct (Neversink Tunnel) that discharges into Rondout Reservoir, through release works located in the Dam, and over the spillway. The ungauged inflow for the study period was estimated by performing a flow budget according to Owens *et al*. (1998). It represented ~ 26% of the total inflow. Other input data included meteorological data, inflow temperature, and inflow turbidity, which are described in Sections 5.1, 5.3, and 5.4, respectively. Data for model testing comprised observed profiles of temperature and turbidity at the single robotic monitoring site, biweekly observations at 3-4 depths at DEP’s routine monitoring sites, and observations (5 days/week) in Neversink Tunnel at the point of discharge into Rondout Reservoir. The coefficients and their values adopted in this preliminary effort are listed in Table 4.9.

Table 4.9. Coefficient values for two-dimensional hydrothermal/transport submodel for Neversink Reservoir.

Coefficient	**Value**
Longitudinal eddy viscosity	1 m^2 s^{-1}
Longitudinal eddy diffusivity	0.4 m^2 s^{-1}
Chezy coefficient	70 $m^{0.5}$ s^{-1}
Wind sheltering coefficient (dimensionless)	1
Fraction of incident solar radiation absorbed at the water surface	0.45

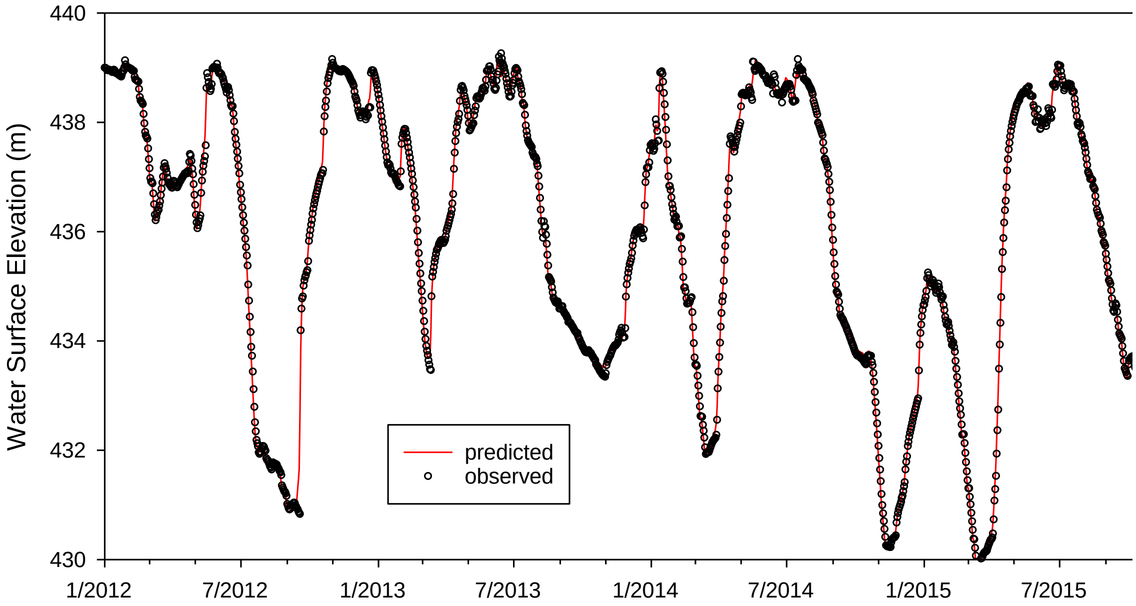


Figure 4.51. Evaluation of the hydrologic budget calculations as comparison of predicted and observed water surface elevation for 2012–2015. Mean absolute error was 0.084 m.

4.5.2. Model Performance

Comparison of the observed and model-predicted water surface elevation for 2012–2015 indicate good performance (Figure 4.51) with a mean absolute error of 0.084 m. For temperature and turbidity, model performance was evaluated both in terms of predictions of in-reservoir vertical profiles, and withdrawal conditions. The submodel performed well in simulating the stratification regime of the reservoir, including the timing of the onset of stratification in spring, the duration of stratification, the onset of fall turnover, and layer temperatures, with the exception of slight over prediction in the epilimnetic layers during July–September interval. Comparisons of detailed vertical profiles of observations and simulations are presented for the April to November interval of 2015 (Figure 4.52–Figure 4.54), which illustrate generally the high level of performance. Performance for 2013–2014 was similar (results not shown). Over prediction in the epilimnetic layers may be due to over mixing caused by wind stress. In future work, we plan to evaluate this effect by conducting sensitivity runs with dynamic wind sheltering coefficient.

Performance of the turbidity submodel for the detailed vertical profiles data for 2015 is shown in Figure 4.55–Figure 4.57, and in Figure 4.58–Figure 4.60 for 2012–2015 with limited laboratory-measured data. 2015 did not represent a rigorous test of the model as there were no major turbidity events in the reservoir in that year. However, the model performed satisfactorily in maintaining the generally uniform vertical structure in turbidity, with occasional signatures of interflow with somewhat higher turbidity. The most recent significant turbidity event in Neversink Reservoir occurred in September, 2012, when it peaked around 350 NTU at site 2NN (Figure 4.58; profile #25). The model predicted the timing and the longitudinal extent of the turbid interflow well, however the overall magnitude was under predicted and can be attributed to uncertainty in the estimates of loading.

In conclusion, time series of temperature and turbidity observations in Neversink Tunnel at NRR2CM are compared to continuous simulations for the entire 2012–2015 period of testing in Figure 4.61 and Figure 4.62, respectively. The model under predicted withdrawal temperatures during the summer stratification period, despite good performance in the water column. This is likely due to the uncertainty in the empirical selective withdrawal algorithm in the model (Cole and Wells 2013) and specification of the effective depth of withdrawal. This will be investigated in the future work. The average absolute error was 2.2 °C. Performance of the model with respect to withdrawal turbidity was good (average absolute error = 0.7 NTU; Figure 4.62). Performance, in this format, during September, 2012 for high turbidity event could not be assessed because the tunnel was shut down and no observations of turbidity. However, limited observations from elevation tap in the reservoir (NR2) that draws water from the corresponding diversion elevation indicate good agreement with the predictions (Figure 4.62). Further work may include sensitivity runs, improving the estimates of turbidity loading from Neversink River, additional testing, and hindcasting.

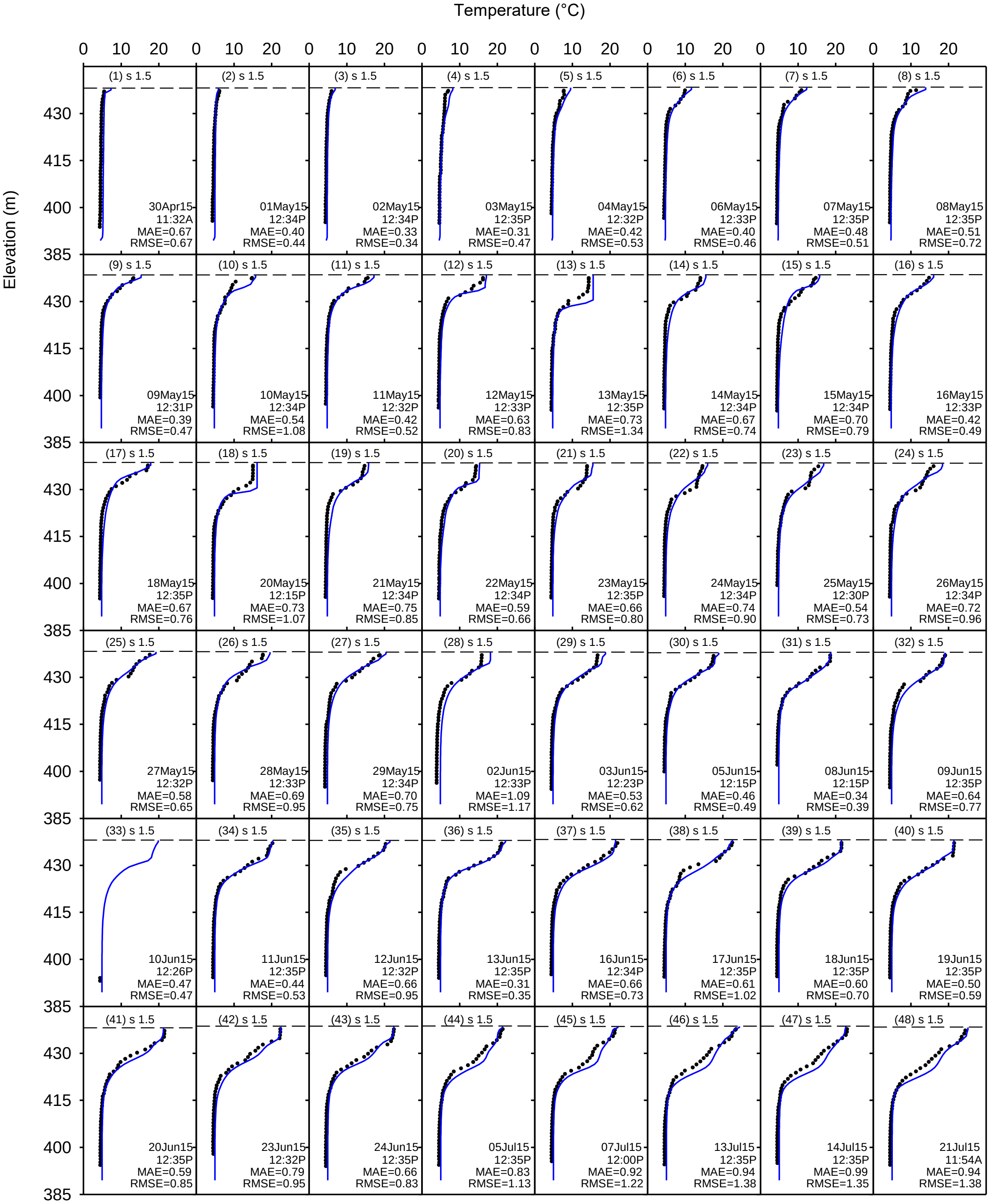


Figure 4.52. Performance of model for Neversink Reservoir for April–July, 2015 as comparisons of predicted and observed vertical profiles of temperature at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (°C) and root mean square error (°C), respectively.

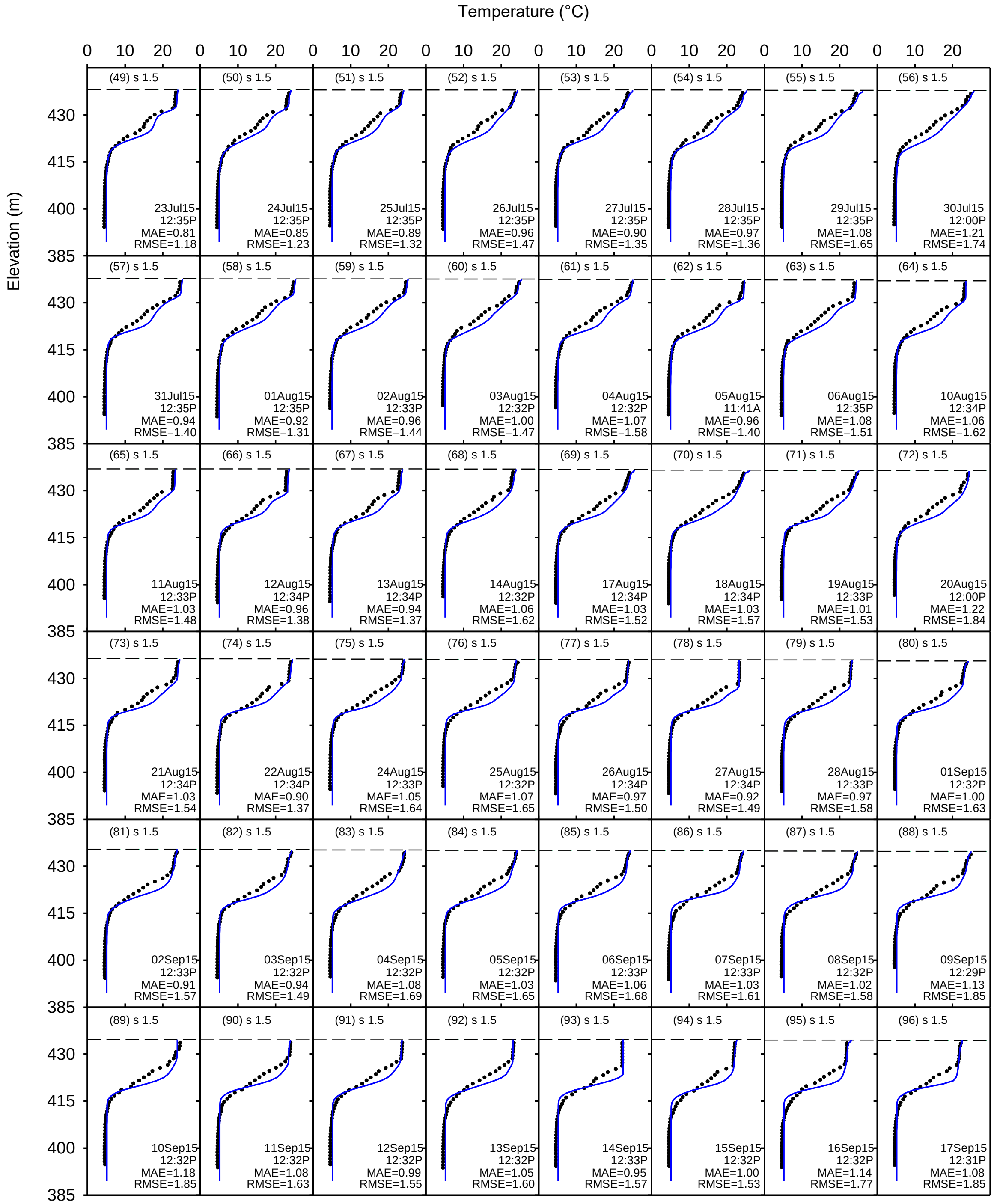


Figure 4.53. Performance of model for Neversink Reservoir for July–September, 2015 as comparisons of predicted and observed vertical profiles of temperature at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (°C) and root mean square error (°C), respectively.

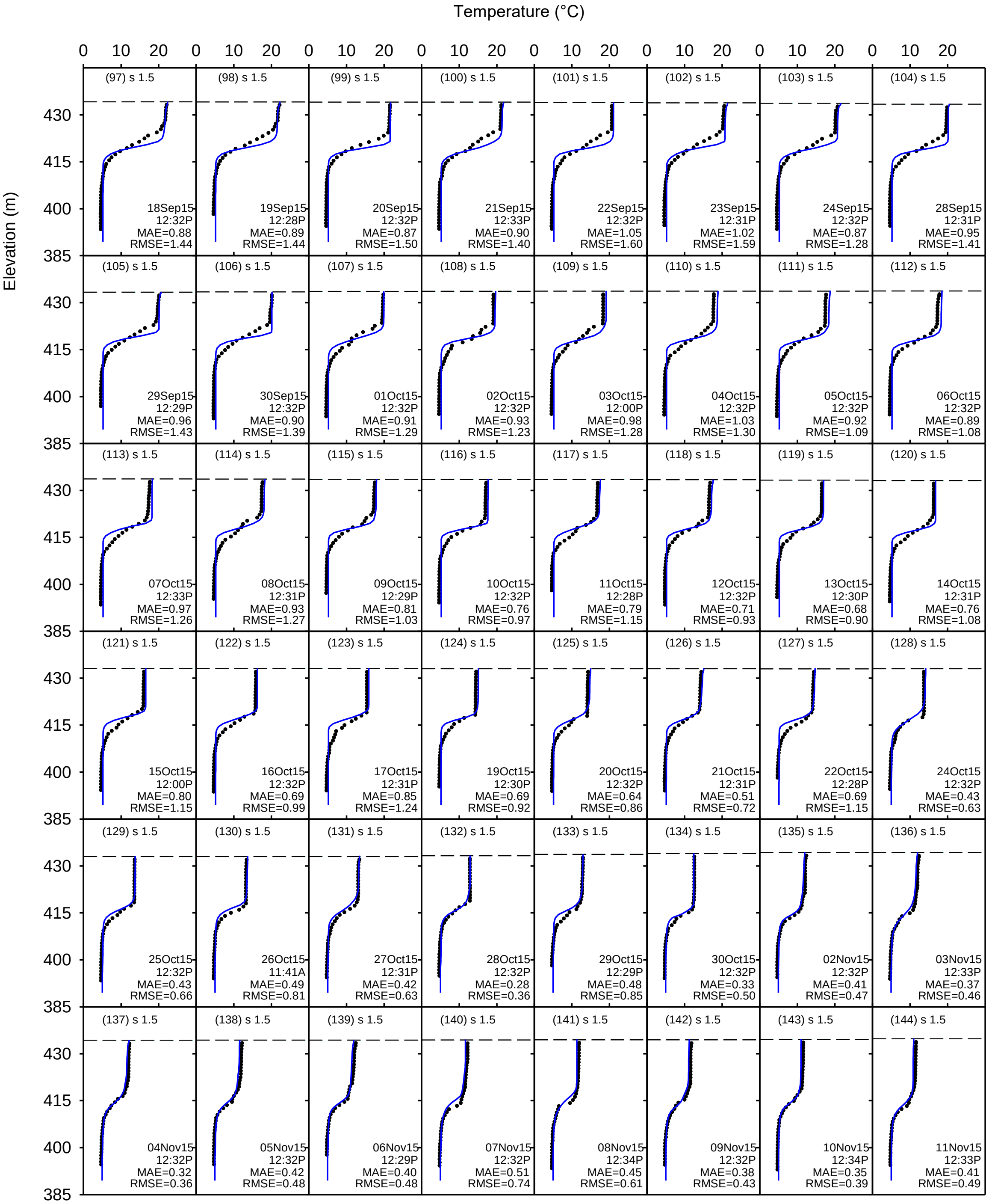


Figure 4.54. Performance of model for Neversink Reservoir for September–November, 2015 as comparisons of predicted and observed vertical profiles of temperature at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (°C) and root mean square error (°C), respectively.

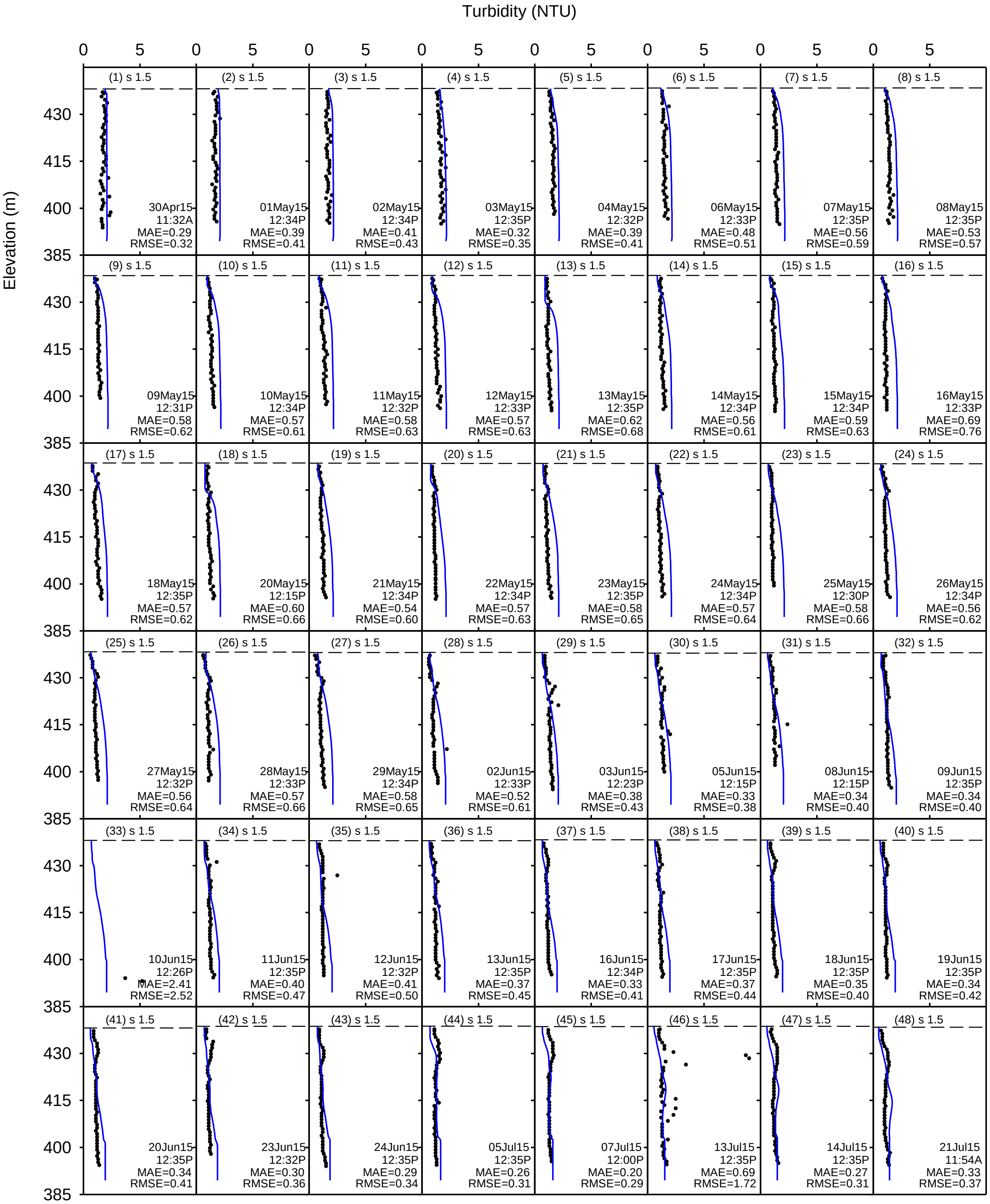


Figure 4.55. Performance of model for Neversink Reservoir for April–July, 2015 as comparisons of predicted and observed vertical profiles of turbidity at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

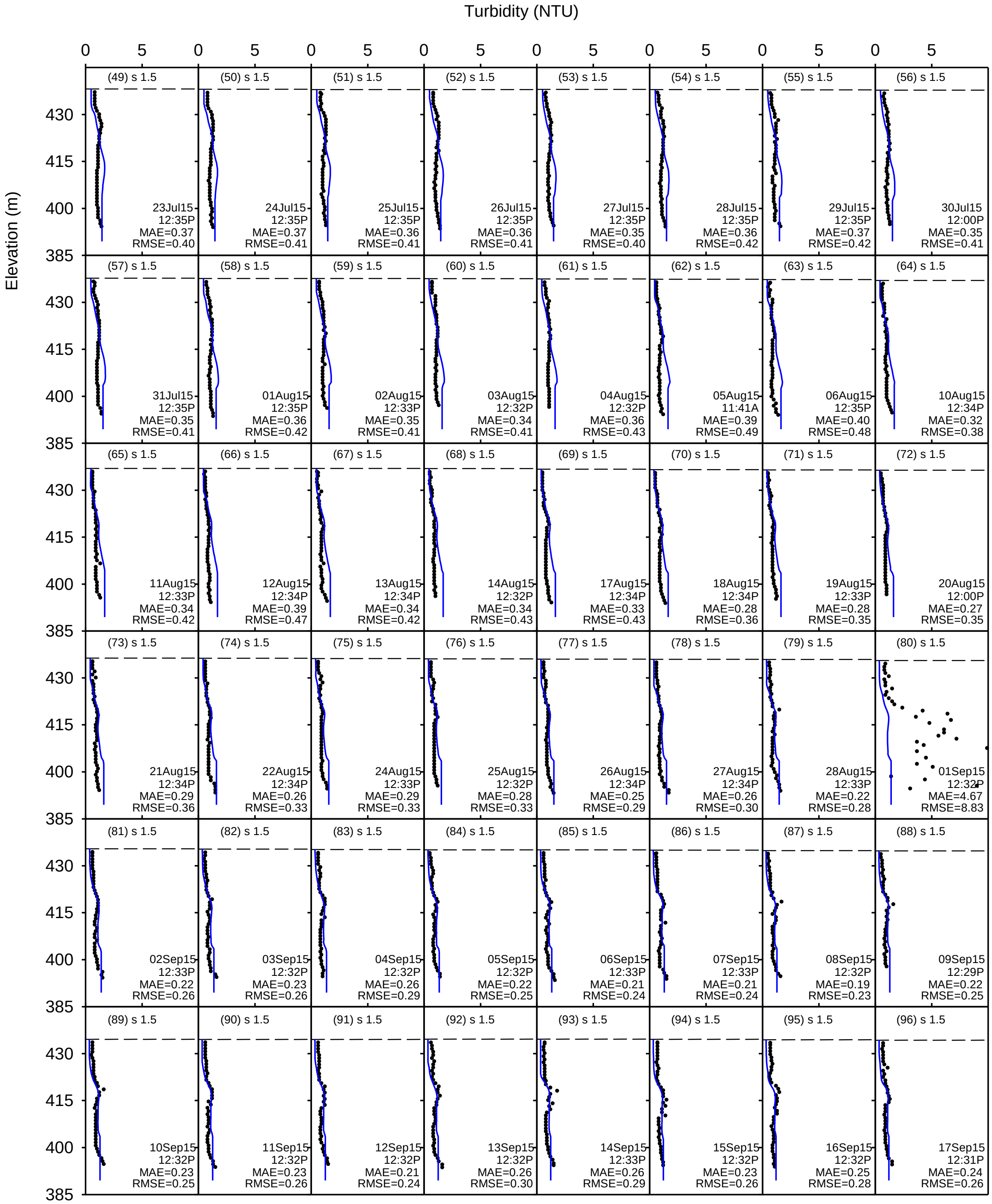


Figure 4.56. Performance of model for Neversink Reservoir for July–September, 2015 as comparisons of predicted and observed vertical profiles of turbidity at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

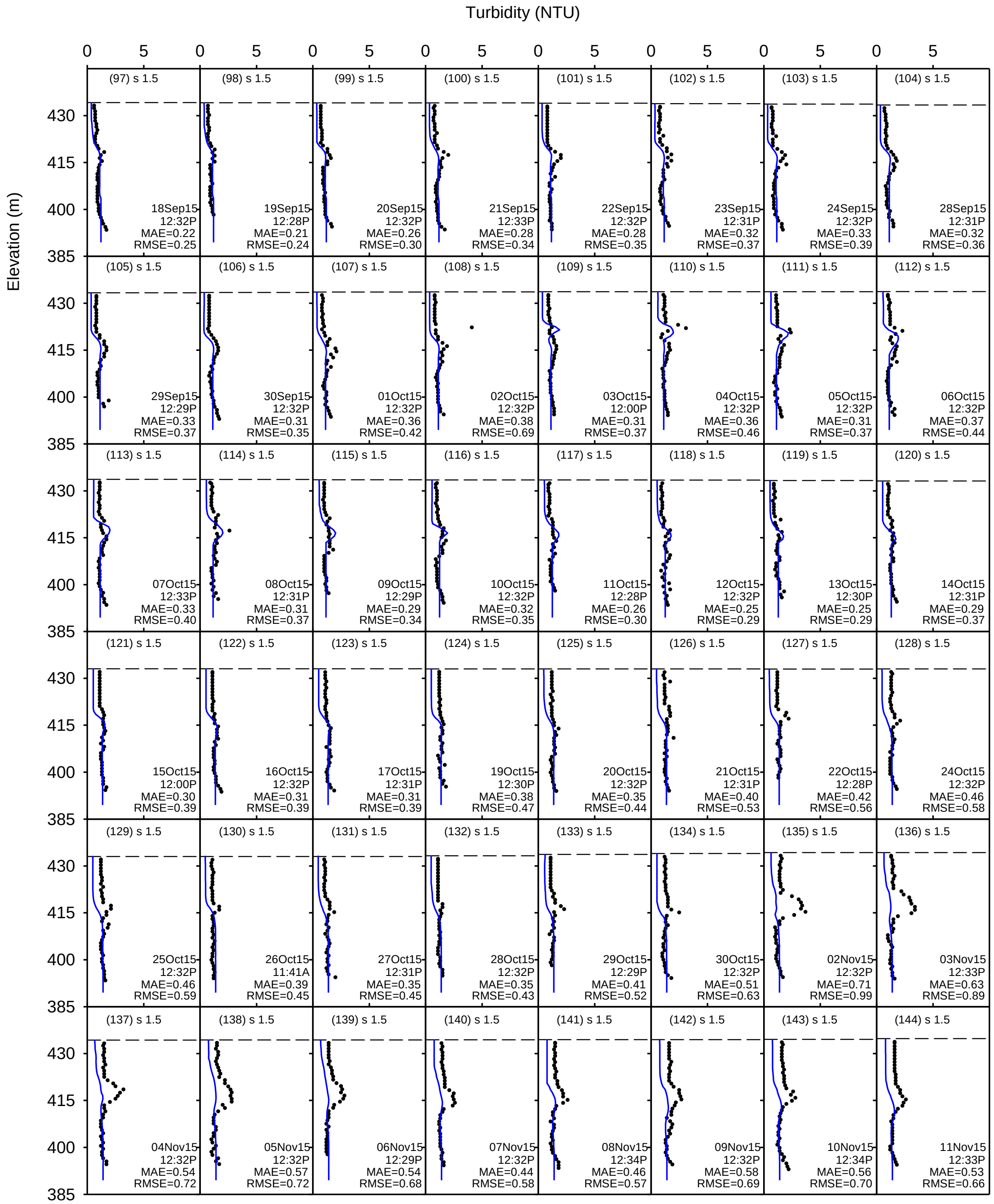


Figure 4.57. Performance of model for Neversink Reservoir for September–November, 2015 as comparisons of predicted and observed vertical profiles of turbidity at robotic monitoring site 1.5NN. MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

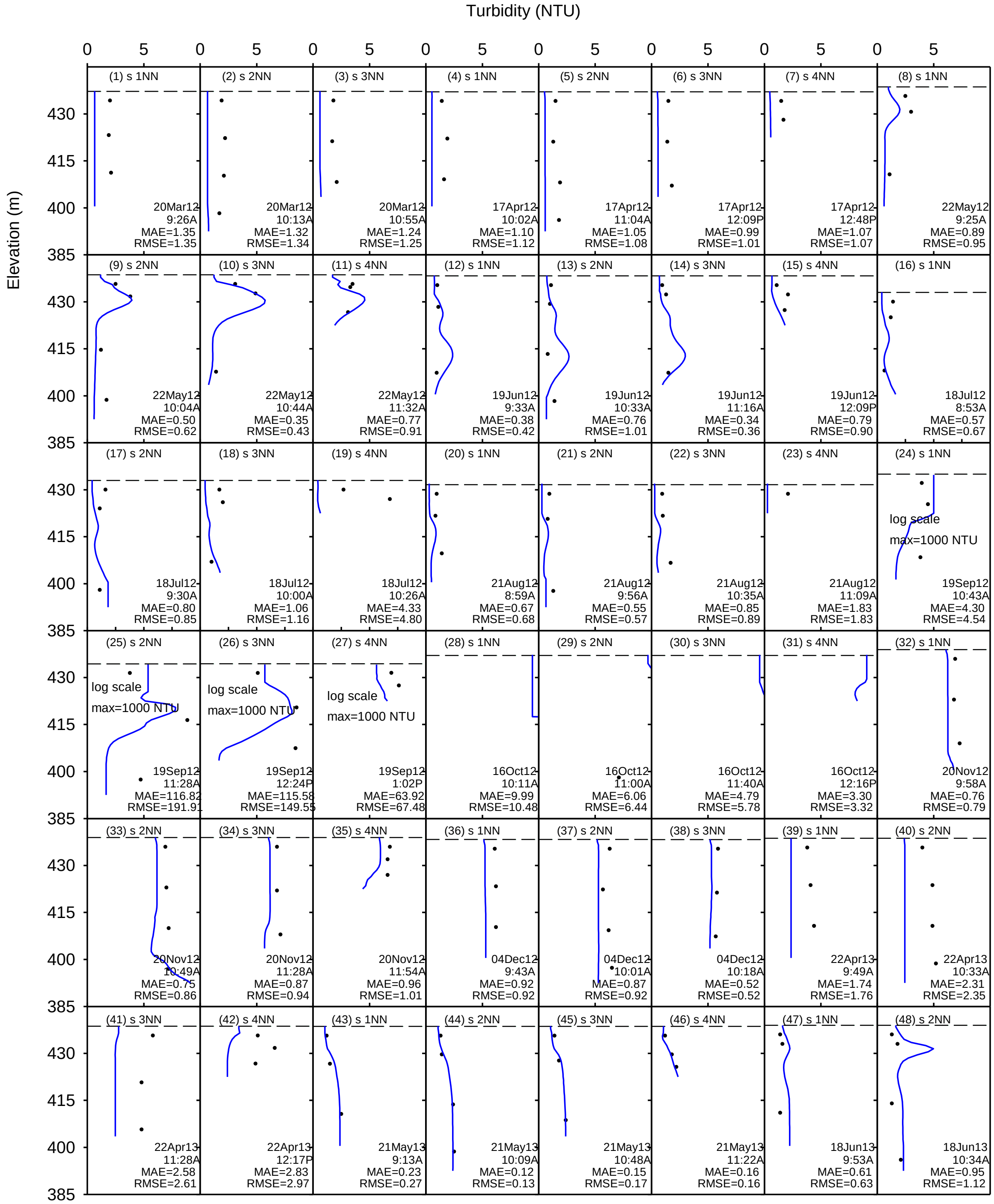


Figure 4.58. Performance of model for Neversink Reservoir for March, 2012–June, 2013 as comparisons of predicted and observed vertical profiles of turbidity at routine monitoring sites (1NN, 2NN, 3NN, 4NN). MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

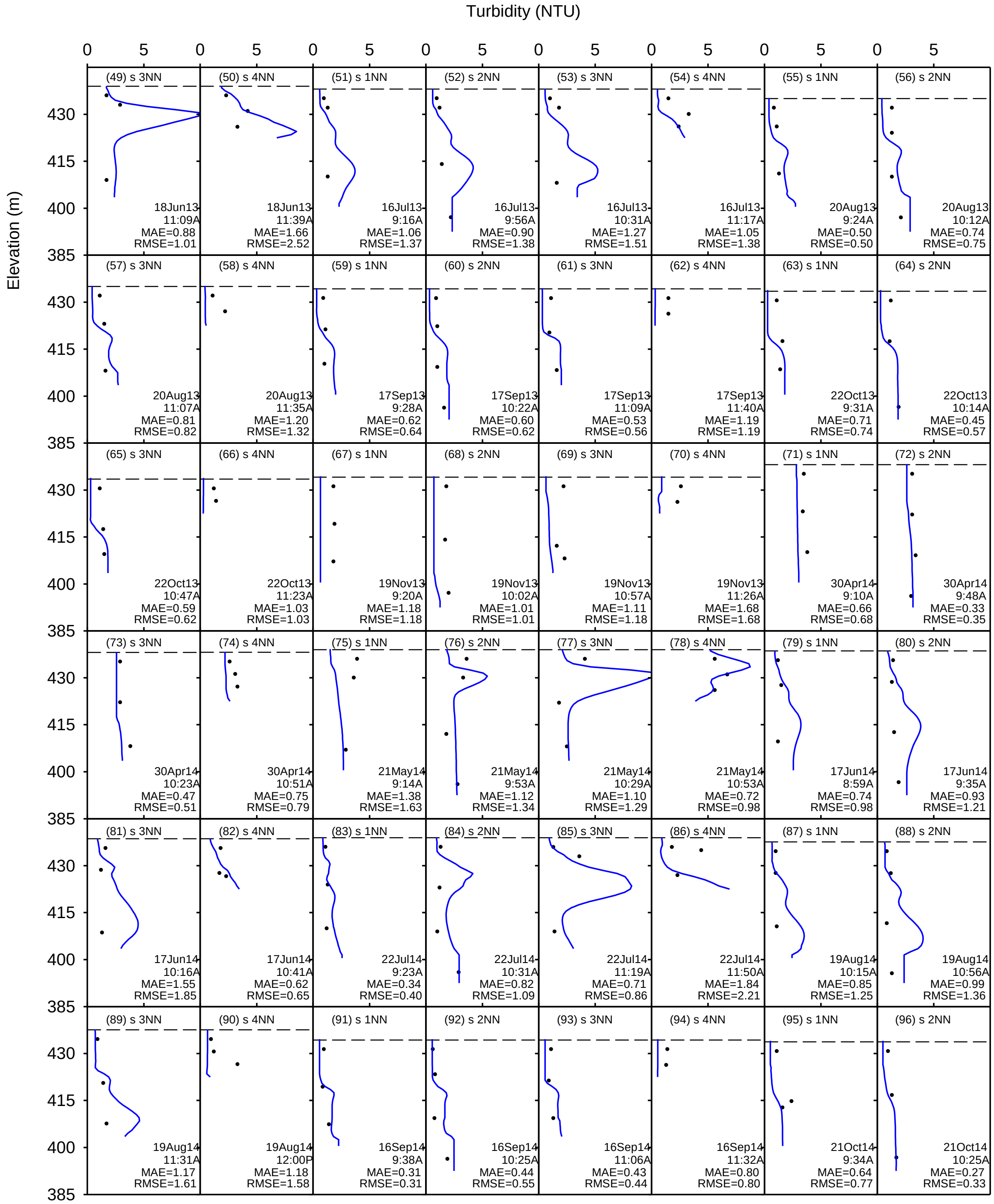


Figure 4.59. Performance of model for Neversink Reservoir for June, 2013–October, 2014 as comparisons of predicted and observed vertical profiles of turbidity at routine monitoring sites (1NN, 2NN, 3NN, 4NN). MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

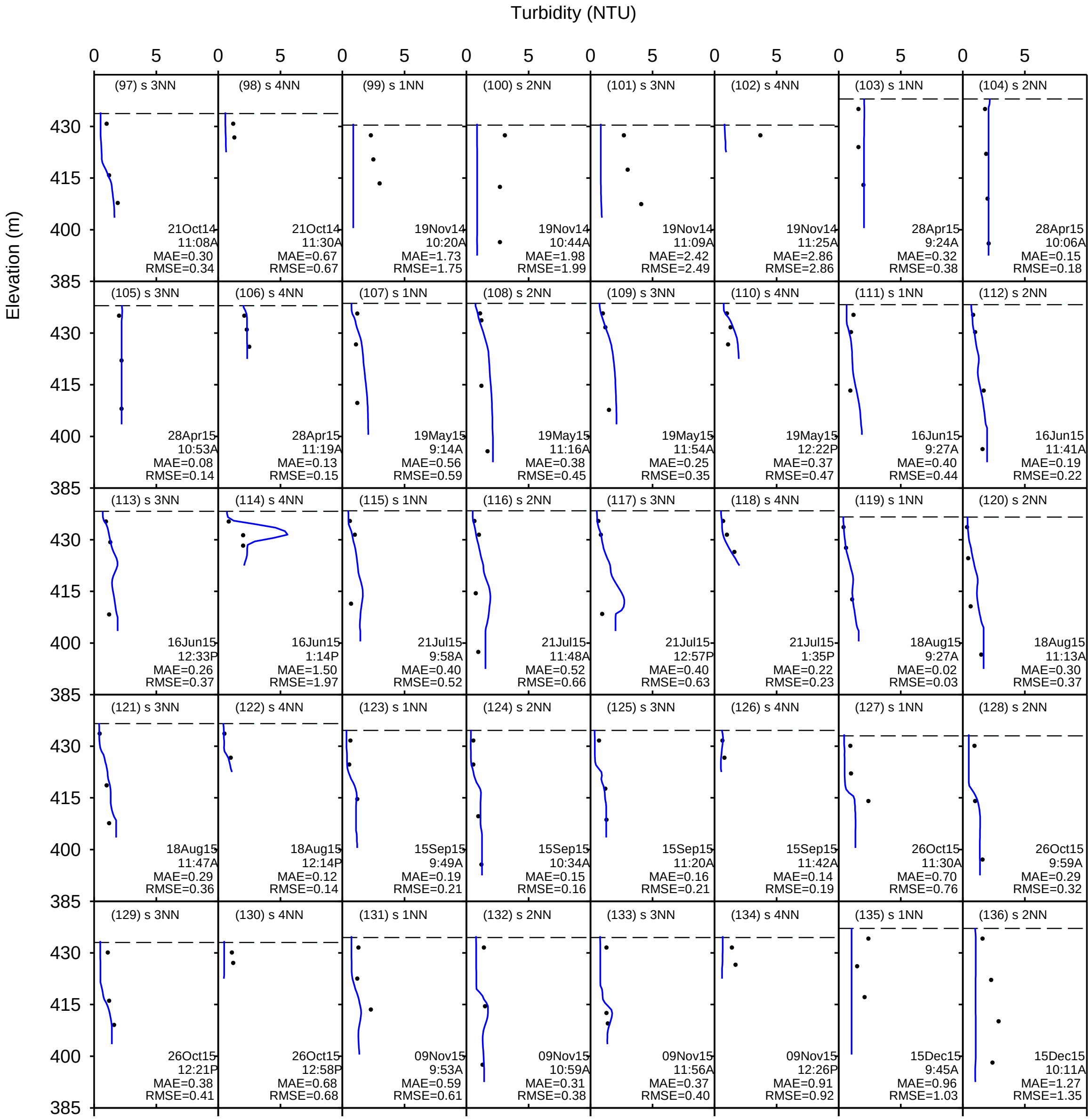


Figure 4.60. Performance of model for Neversink Reservoir for October, 2014–December, 2015 as comparisons of predicted and observed vertical profiles of turbidity at routine monitoring sites (1NN, 2NN, 3NN, 4NN). MAE and RMSE indicate mean absolute error (NTU) and root mean square error (NTU), respectively.

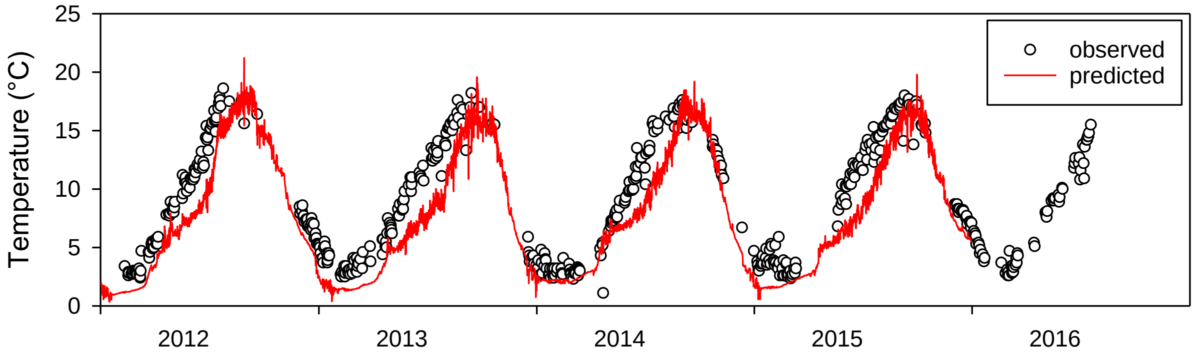


Figure 4.61. Performance of model for Neversink Reservoir for 2012–2015 as comparison of predicted and observed temperatures in Neversink Tunnel (NRR2CM; diversion).

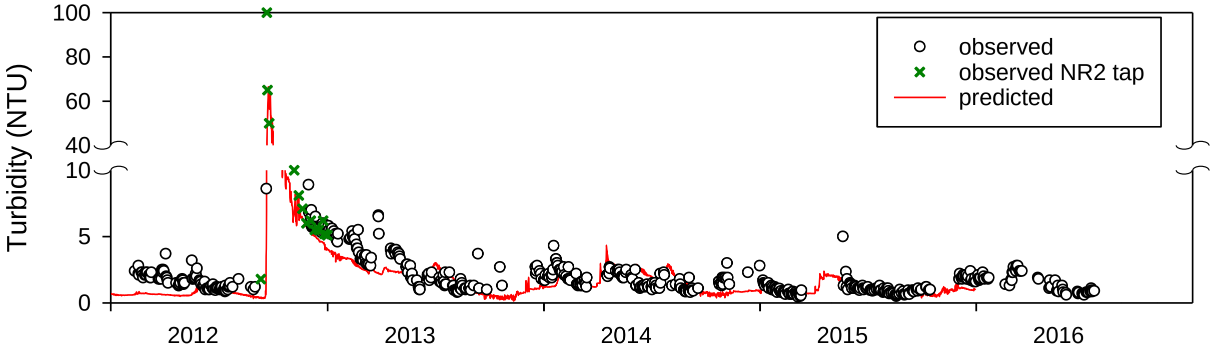


Figure 4.62. Performance of model for Neversink Reservoir for 2012–2015 as comparison of predicted and observed turbidities in Neversink Tunnel (NRR2CM; diversion). Observed turbidities from NR2 (an in-reservoir location at the point of withdrawal at level 2) are also shown for the storm event of September 2012.

4.5.3. Routine Water Quality Forecasting

DEP's Operations Support Tool (OST) is used to guide reservoir operating decisions involving both the quantity and quality of water. In the position analysis (PA) mode, OST can be used to generate near-term as well as long-term forecasts of storage and turbidity levels in the NYC's water supply system. Water quality model runs with OST have been conducted on demand, generally in response to forecast of a storm in the watershed. In the future, it is envisioned that in addition to on-demand operation, OST forecasts will be generated on a regular schedule, such as once per week. Such routine water quality forecasting is intended to provide

Bureau staff with the opportunity to conduct OST runs and interpret output more frequently. This will allow any shortcomings in software or procedures to be identified and corrected quickly. We will then be better prepared to rapidly respond to watershed storm events and adjust water supply operations effectively and efficiently. The Water Quality Modeling group is developing software tools based on OST to greatly reduce the time and effort required to setup and complete water quality runs with OST.

The scope of the work will include predictions of turbidity as the water quality parameter at three important keypoints in the system – the water supply diversion from Rondout Reservoir (RDRR), the diversion from Ashokan Reservoir (EARCM), and the diversion from Kensico (DEL18DT). The predictions will be generated in a probabilistic format indicating the probabilities of exceedance of selected levels of turbidities that are appropriate for a keypoint at the time of forecast. The forecasts would be generated for a period of two weeks and updated once a week. This software to generate these forecasts is currently under development.

4.6. Streamflow simulation for East of Hudson watersheds

4.6.1. Background

Streamflow contribution from watersheds draining to NYC DEP's East of Hudson (EOH) reservoirs are required to make simulations with Operations Support Tool (OST) and to study the impact of climate change on NYC water supply. Currently regression-based mass balance methods are used to estimate historical inflows to EOH reservoirs (DEP 2016). One approach to estimate inflows is based on observed water surface elevation, but this is limited by the availability of daily records of reservoir water surface elevation, and diversions and other reservoir outflows. An alternative approach to estimate inflows to a reservoir is the use of a watershed scale hydrologic model. Watershed models have the advantage of applicability to ungauged watersheds using calibrated model parameters from nearby watersheds with similar landscape characteristics. In 2016, staff from DEP's Water Quality Modeling Section and Operations Directorate jointly completed a modeling project using GWLF model to estimate historical streamflow from EOH watersheds. A brief summary of the progress on this project follows.

4.6.2. Methods

The GWLF-VSA model was set up for five natural inflow gauged watersheds, seven flow regulated watersheds and two controlled lake watersheds that are part of the Croton system of NYC water supply (Figure 4.63). A 2009 land use map was overlaid on a 10-class topographic wetness index map to create land use-wetness index combinations and was used as input to the model. The Parameter-elevation Relationships on Independent Slopes Model (PRISM) climate data (daily precipitation and daily maximum and minimum air temperature (T_{MAX} and T_{MIN})) available on a 4-km grid were used as input to drive the model. For each watershed the PRISM grid near the centroid of the watershed was used and this dataset was accessed through the applied climate information system (http://www.rccacis.org/). Other required climate data including solar radiation and relative humidity were computed using a climate interpolation model that uses air temperature and precipitation data (MT-CLIM, (Running *et al*. 1987)).

Figure 4.63. Location of NYC East of Hudson Watershed.

.

The GWLF model was calibrated to observed streamflow for the five natural inflow reservoir watersheds using available streamflow observations from USGS stream gauges for the period from 10/1/2011 to 09/30/2015. The calibrated models were validated by comparing model simulated streamflows to observed streamflows for an independent time period. Model performance was evaluated using R^2 and Nash Sutcliffe efficiency (NSE) coefficient. In addition, inflow volumes simulated by the model was compared to measured volumes for the period of available USGS data record for each watershed. Calibrated parameters from the five gauged watersheds were then applied to estimate the total inflows into the same five reservoirs including inflows from ungauged areas. The same approach was used to estimate inflows from the seven flow regulated watersheds and two controlled lake watersheds by using calibrated parameters from the natural inflow watersheds. Finally, long-term simulations were made for all 14 EOH reservoir watersheds for the period from 01/01/1981 to 12/31/2015 to generate historical inflows.

4.6.3. Results

Simulated streamflows were comparable to observed streamflows for all five calibrated watersheds with R^2 ranging from 0.60 to 0.78 and NSE ranging from 0.60 to 0.76 during the calibration period. During the validation period R^2 values ranged from 0.69 to 0.75 and NSE values ranged from 0.66 and 0.75 (Table 4.10). In all cases the model underestimated the high flow events (Figure 4.64). Difference in inflow volumes simulated by the model for the period of available USGS streamflow record were as follows: (Amawalk at USGS station #1374930 +6.7%; Boyd Corners at USGS station# 1374559, +2.9%; Cross River at USGS station# 1374890, -7.8%; East Branch at USGS station# 137449480, +2.6%; Titicus at USGS station# 1374781, -5.9%).

A comparison of simulated monthly streamflow with observations for the period of available USGS streamflow record shows close agreement (Figure 4.65). This corresponds to the calibration/validation period of this study. Use of average model parameters did not cause major change in model performance at the daily time step except for Titicus at USGS station# 1374781 where R^2 and NSE dropped from 0.60 to 0.55 and 0.54 for the calibration period. For the validation period this change was from 0.73 to 0.66 for R^2 and 0.70 to 0.60 for NSE.

Overall, this analysis shows that using a simple lumped parameter model like GWLF it is possible to use parameters from a calibrated model to simulate streamflow responses in nearby small watersheds. This also provides confidence in simulated estimates of long-term (1981-2015) inflows from the seven regulated and two controlled lake watersheds. This preliminary work is part of a study investigating multiple complementary approaches to obtain best estimates of inflows from watersheds draining to NYC EOH reservoirs.

Table 4.10. GWLF model performance for selected (natural inflow) east of Hudson reservoir watersheds (R^2 and NSE are for daily streamflow).

		R^2	**NSE**	**R^2**	**NSE**	
Watershed	**Drainage area (km^2)**	Calibration period (10/01/2011 to 09/30/2015)		Validation period		**Validation period**
Amawalk at USGS station# 1374930	34.16	0.70	0.69	0.75	0.75	10/01/1995 to 09/30/2011
Boyd Corners at USGS station# 1374559	28.42	0.70	0.68	0.69	0.66	10/01/1995 to 09/30/2011
Cross River at USGS station# 1374890	44.01	0.78	0.76	0.72	0.70	12/08/1995 to 09/30/2011
East Branch at USGS station# 137449480	157.18	0.71	0.70	0.69	0.69	10/01/1995 to 09/30/2011
Titicus at USGS station# 1374781	33.4	0.60	0.60	0.73	0.70	12/08/2006 to 09/30/2011

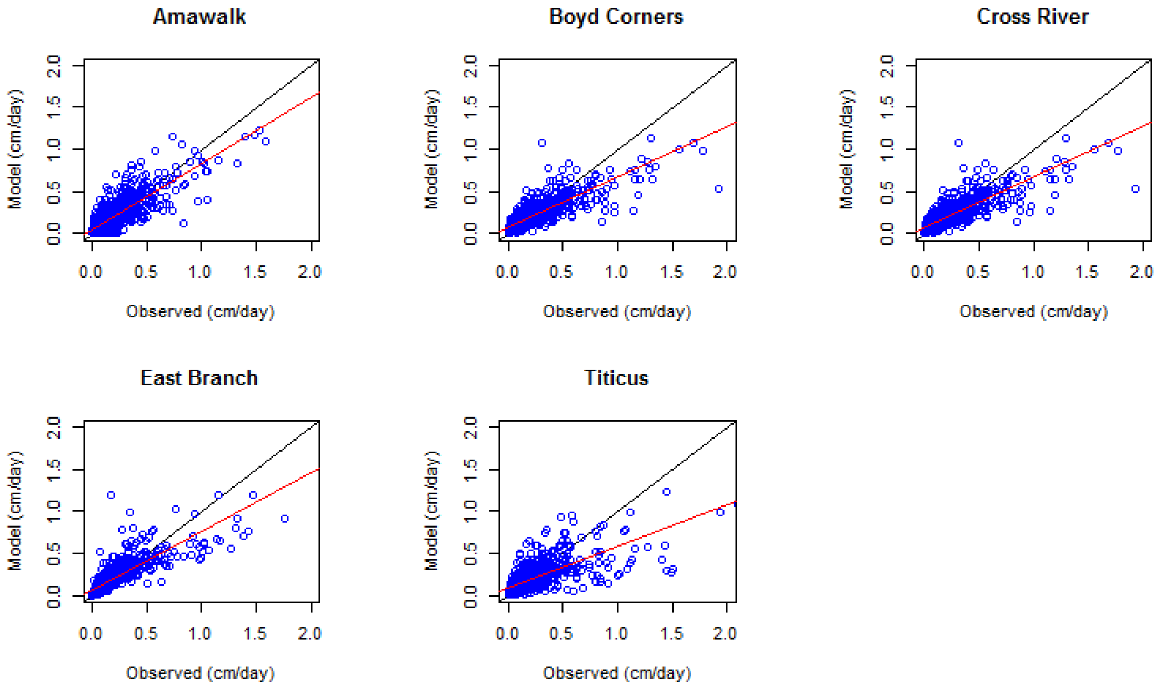


Figure 4.64. Comparison of GWLF model simulated vs. observed daily streamflow (cm/day) for five natural inflow EOH watersheds. Linear regression lines of model vs. simulated values are shown in red and 1:1 regression line in black.

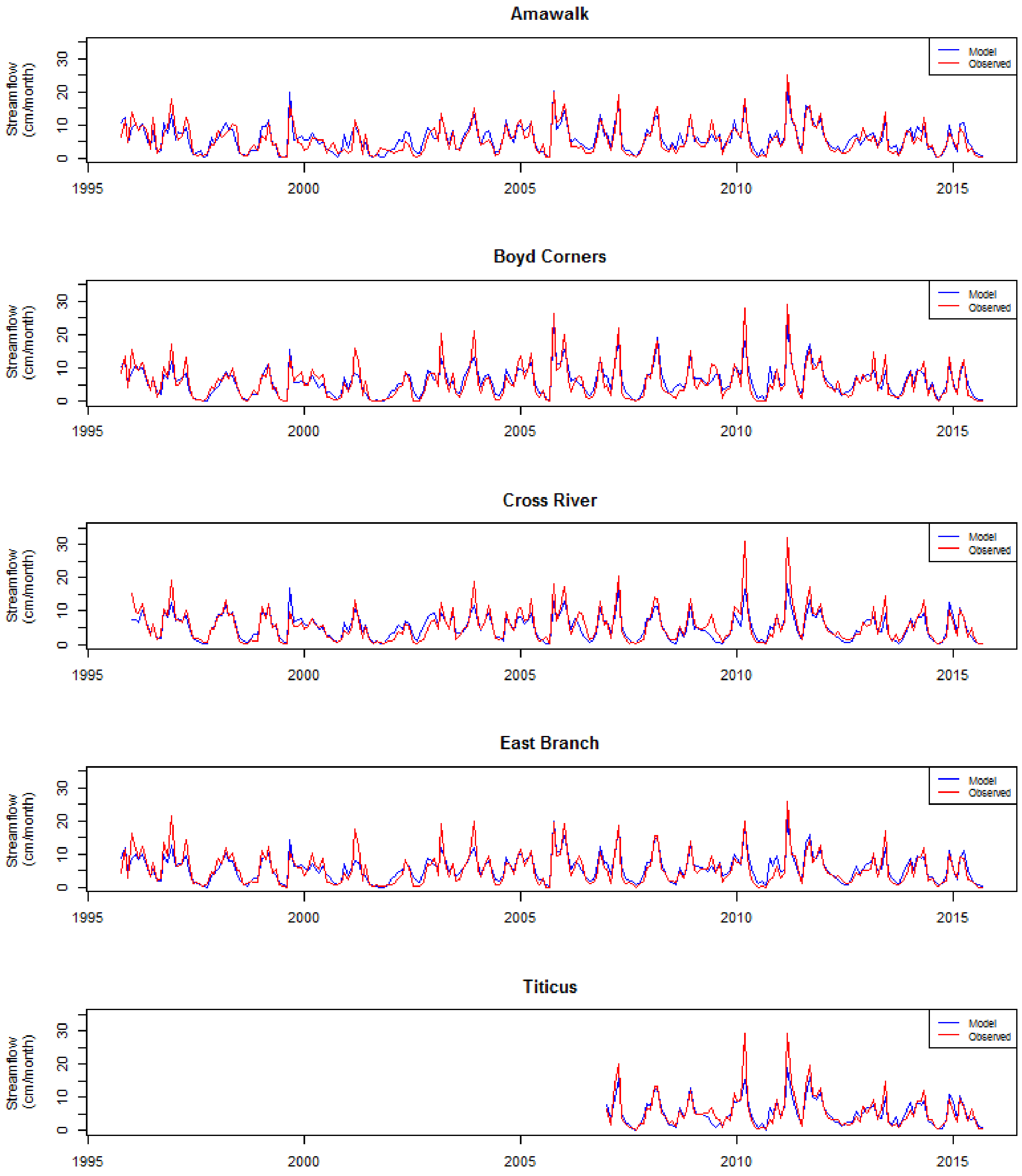


Figure 4.65. Comparison of GWLF model simulated and observed monthly streamflow during the calibration and validation periods for five natural inflow EOH watersheds.

4.7. Effect of PRISM precipitation input on streamflow simulation in NYC WOH watersheds

4.7.1. Introduction

Weather is one of the most important drivers in watershed scale hydrologic models (Beven 2001) and the accuracy of any model simulation is directly linked to the accuracy in weather inputs. In recent years there has been a steady decline in the number of precipitation monitoring stations in NYC West of Hudson (WOH) watersheds from 18 stations in the 1990s to 8 stations in the late 2000s. These are mostly cooperator stations recognized by the National Climate Data Center (NCDC). Currently there are only two active precipitation stations in the region. This has necessitated testing and use of alternative climate data sources, including radar based products in watershed models.

One source of climate data is the climate forecast system reanalysis (CFSR) data (http://globalweather.tamu.edu/). This is an interpolated dataset on a 38-km grid based on the National Weather Service Global Forecast system. The CFSR is based on hourly forecasts generated using information from the global weather station network and satellite-derived products, covering any land location in the world. A review of recent articles on studies using CFSR data for hydrological modeling indicates satisfactory or comparable to poor performance when compared to conventional rain gauge data (Dile and Srinivasan 2014, Fuka *et al*. 2014, Worqlul *et al*. 2015, Yang *et al*. 2014). Another source of interpolated daily weather data that has not been commonly used or tested in watershed models is the Parameter-elevation Relationships on Independent Slopes Model (PRISM) data (Daly *et al*. 2008, Di Luzio *et al*. 2008) available on a 4-km grid covering the conterminous U.S. for the period from 1981 to present (www.prism.oregonstate.edu/). The daily time series of temperature and precipitation starting 1981 was developed using a process called climatologically aided interpolation (CAI). For precipitation, CAI, using 1981-2010 monthly climatology as the predictor grids, is applied in the western U.S. (Rockies westward), and, 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 are included in the PRISM dataset. A recent study by Radcliffe and Mukundan (2017) compared PRISM climate data with other data sources including CFSR and NCDC observer station data in simulating streamflow using the SWAT model for two small watersheds in Georgia and Louisiana. PRISM models outperformed models using CFSR and NCDC rain gauge data, but the CFSR grid/rain gauge in this study was located outside the watershed boundary in each of the watersheds. While intra-basin variability in rainfall was minimal among various PRISM stations in the small watersheds tested (44 and 132 km^2), agreement among NCDC, PRISM, and CFSR rainfall data were dependent on the proximity of the stations compared. They concluded that using PRISM data is a promising approach when rain gauge data are not available as PRISM data are available at near real time. However additional testing is needed at different scales and locations including snowmelt dominated watersheds. The objective of this study was to evaluate the performance of PRISM climate data in simulating streamflow in NYC WOH watersheds.

4.7.2. Methods

The Generalized Watershed Loading Function-Variable Source Area (GWLF-VSA) model was set up to simulate streamflow at 10 USGS gauging stations within WOH watersheds with drainage areas ranging from 65 to 859 km^2 (Table 4.11). Two sets of models were calibrated for streamflow at the gauging stations for a five-year period from 10/1/2004 to 9/30/2009; one using observer station climate data as input and the other using PRISM climate data as input. A single time series of daily precipitation and air temperature (T_{MAX} and T_{MIN}) from 10 observer stations in and around WOH watersheds and 324 PRISM grid points were used as model input (Figure 4.66). For each watershed daily station/grid data were averaged for the entire basin using Thiessen polygon method interpolation for precipitation and inverse distance squared interpolation for air temperature. Over the duration of the simulation period the number of observer stations reporting precipitation declined to seven. The PRISM dataset was accessed through the applied climate information system (http://www.rcc-acis.org/). Streamflow simulated by the two sets of models for each watershed was compared to observations using standard hydrologic model evaluation statistics. In addition, final calibrated values of model parameters were compared to evaluate the effect of climate input on streamflow simulation.

Table 4.11. Location of USGS gauging stations and corresponding drainage areas.

USGS Gauge No.	Watershed Description	Reservoir	Drainage Area (km^2)
1414500	Millbrook near Dunraven	Pepacton	63.8
1350080	Manor Kill near Gilboa	Schoharie	84.0
1415000	Tremper Kill near Andes	Pepacton	85.5
1365000	Rondout Cr. near Lowes Corners	Rondout	99.5
1362200	Esopus Creek at Allaben	Ashokan	164.8
1435000	Neversink River near Claryville	Neversink	172.5
1413500	EBDR at Margaretville	Pepacton	421.7
1362500	Esopus Creek at Coldbrook	Ashokan	493.2
1350000	Schoharie Creek at Prattsville	Schoharie	612.5
1423000	WBDR at Walton	Cannonsville	859.3

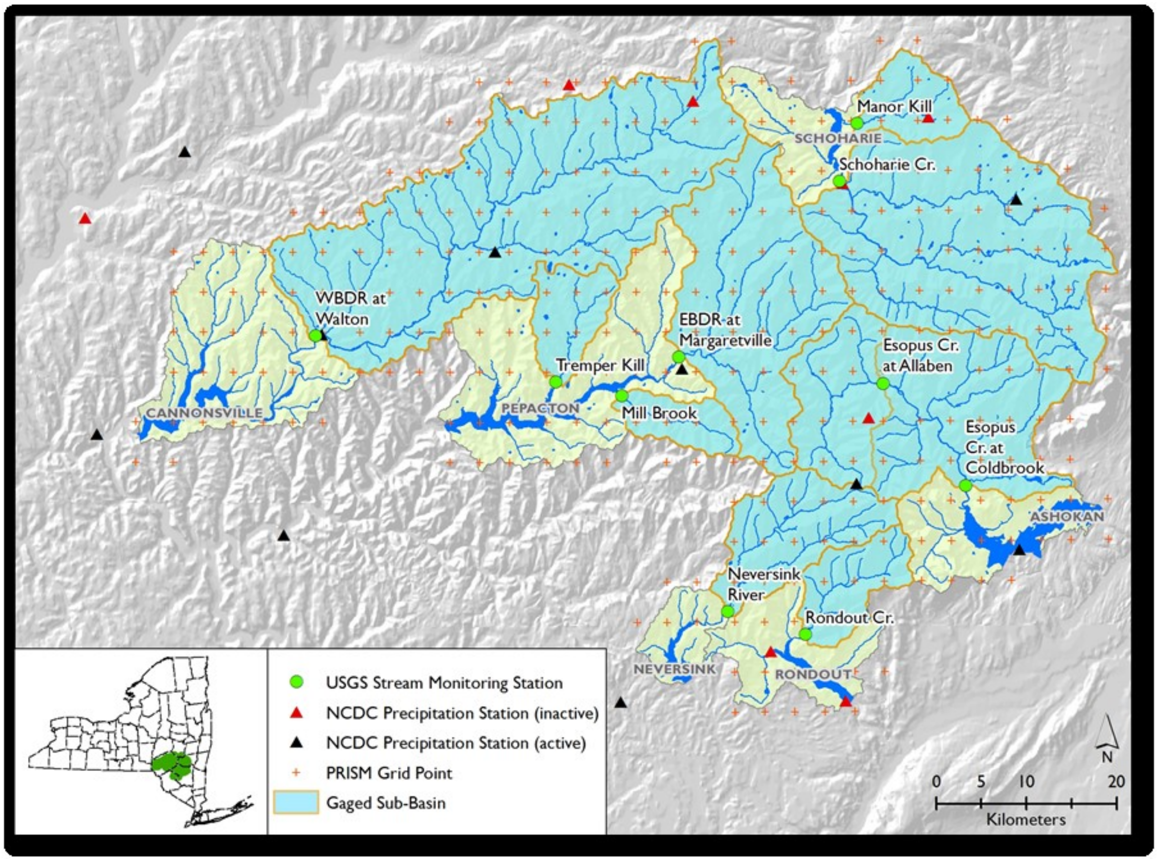


Figure 4.66. Gauged watersheds that provide major inflows to NYC West of Hudson Reservoirs, and locations of NCDC stations and 4-km PRISM grid points.

4.7.3. Results and Discussion

Streamflow simulations using PRISM data were comparable or better than simulations using available observer station data for nine out of ten streams based on R^2 and Nash-Sutcliffe efficiency (NSE) values (Table 4.12). Nevertheless all simulations qualified as “good simulations” based on hydrologic model evaluation criteria (monthly NSE > 0.65) proposed by Moriasi *et al*. (2007). In all cases simulations showed GWLF model underestimating the peak flow values in all watersheds simulated when compared to observed values (Figure 4.67). Model vs. observed regression lines were all below the 1:1 line. For Esopus Creek at Allaben and Coldbrook, Schoharie Creek at Prattsville, and Rondout Creek near Lowes Corners improvement in streamflow simulation when using PRISM was minimal. No apparent differences in model performance were observed for Neversink River near Claryville and Millbrook near Dunraven. The only stream where observer station clearly outperformed PRISM was in Manor Kill at Conesville near Gilboa. This was found to be due to the fact that PRISM gave a low value for the largest precipitation event during this period when observer station reported a relatively high value (Figure 4.67).

PRISM based models clearly outperformed observer station based models for three streams. These include West Branch Delaware River (WBDR) at Walton, East Branch Delaware River (EBDR) at Margaretville, and Tremper Kill near Andes. While there were two observer stations within the Cannonsville watershed during the period of analysis, their locations were at the middle or near the outlet of the watershed. Therefore, improved simulation of streamflow at WBDR when using PRISM input could be due to better estimation of total rainfall within the watershed. Other hydrologic modeling studies have reported that accurate estimates of total rainfall within a watershed are more important than spatial distribution or timing. (Andréassian *et al*. 2001, Obled *et al*. 1994). In the case of EBDR at Margaretville improved streamflow simulations could be due the denser network of PRISM grid points producing more accurate estimates of total rainfall within the watershed compared to a single observer station below the outlet of this relatively large (421 km^2) watershed. Moreover, unlike PRISM data, observer station records were discontinuous due to inconsistent reporting from this observer station. For Tremper Kill near Andes, the observer station was outside the watershed boundary and near the outlet resulting in PRISM providing better estimates of total rainfall in this relatively small (85 km^2) watershed.

Table 4.12. Model performance evaluation and comparison for daily streamflow.

	Observer station	***PRISM***	***Observer station***	***PRISM***
	Esopus at Allaben		Esopus at Coldbrook	
NSE	0.66	**0.67**	0.64	**0.66**
R^2	0.66	**0.67**	0.65	**0.65**
	WBDR at Walton		Neversink near Claryville	
NSE	0.65	**0.76**	**0.58**	0.56
R^2	0.65	**0.77**	**0.61**	0.59
	EBDR at Margaretville		Mill Brook near Dunraven	
NSE	0.62	**0.71**	**0.59**	0.56
R^2	0.64	**0.73**	**0.62**	0.60
	Tremper Kill near Andes		Rondout Creek at Lowes	
NSE	0.59	**0.66**	0.62	**0.66**
R^2	0.60	**0.68**	0.65	**0.68**
	Schoharie Creek at Prattsville		Manor Kill near Gilboa	
NSE	0.62	**0.64**	**0.65**	0.55
R^2	0.65	**0.65**	**0.68**	0.58

*better fit values are highlighted

Differences in GWLF predictions of streamflow when using different climate input were reflected in four calibrated model parameters in all watersheds. These parameters are summarized in Table 4.13 and include:

1. *Precipitation correction factor*: Factor to adjust precipitation input to basin. Default of 1 yields no adjustment. Calibrated to minimize cumulative error in simulated vs. measured streamflow, where long-term measured streamflow data is available.

2. *Runoff recession coefficient*: Surface water recession constant. The fraction of surface water stored in stream channels and waterbodies within the watershed which takes 1 day or less to reach the outlet.
3. *Recession coefficient*: Coefficient to determine rate at which baseflow is drawn from the saturated zone.
4. *Bypass coefficient*: Fraction of infiltration that bypasses the unsaturated zone and directly enters the saturated zone.

Minimal adjustments were made to precipitation correction factor when using PRISM as input to GWLF compared to observer station data. This implies that PRISM provides a better estimate of total rainfall within the watershed compared to estimates made from limited number of available weather stations in and around the watersheds studied. Consistently greater calibrated values were used for runoff coefficient, recession coefficient, and bypass coefficient when using PRISM data reflecting model response to input climate data.

Table 4.13. GWLF parameters and values reflecting predicted streamflow differences.

	Parameter	Obs. Station range	PRISM range
1	Precipitation factor	0.84-1.12 (1.04)	0.92-1.06 (1.01)
2	Runoff coefficient	0.36-0.58 (0.46)	0.41-0.61 (0.52)
3	Recession coefficient	0.05-0.06 (0.05)	0.06-0.07 (0.06)
4	Bypass coefficient	0.0-0.08 (0.04)	0.03-0.14 (0.07)

*average values are shown in parenthesis

4.7.4. Conclusions

Streamflows simulated using PRISM based precipitation were comparable or better than simulations using precipitation values interpolated from a coarse network of ten available weather stations used in this analysis. Improvements in streamflow simulation using PRISM depend on the availability and proximity of weather station in the watershed simulated. In the absence of observer stations PRISM can be a useful source of climate data for watershed modeling in WOH watersheds. Unlike observer stations, PRISM offers a continuous time series of climate record from 1981 to present with no geographic data gaps or inconsistent reporting periods. Model simulations using PRISM were in most cases classified as “good” as per model evaluation guidelines at the daily time step. Although PRISM climate data are useful for watershed modeling they are not a substitute for observer station data. Simulated streamflows in this analysis underestimated high flows in all cases, possibly due to inaccurate estimates of total rainfall during large events. PRISM data incorporates station data, so an active network of local stations will better inform PRISM, and likely result in better estimation of high flows. For best model estimates of streamflow from watersheds that provide major inflows to NYC WOH reservoirs, an active network of observer stations would be desirable.

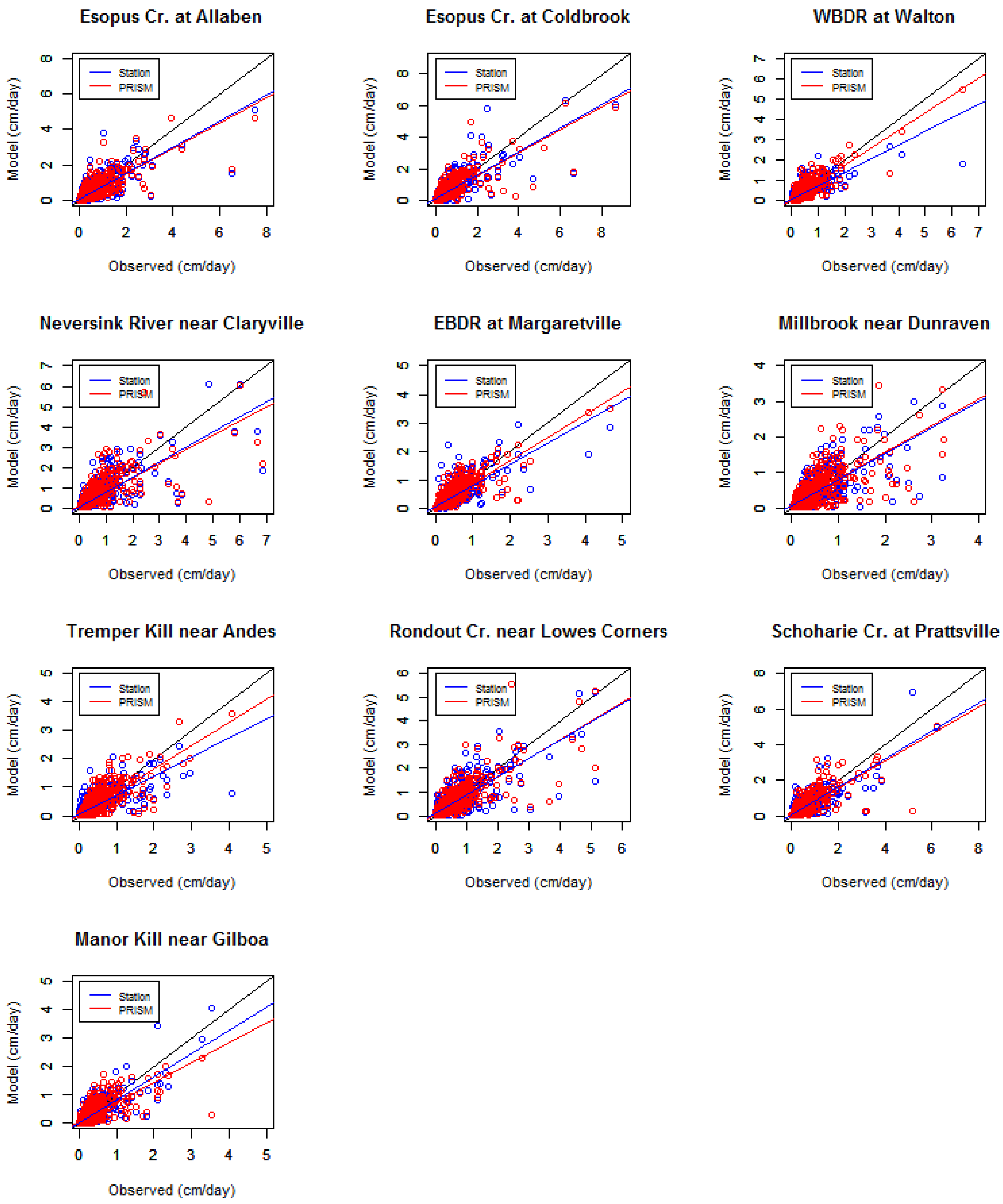


Figure 4.67. Comparison of GWLF model simulated vs. observed daily streamflow (cm/day) for watersheds that provide major inflows to NYC West of Hudson Reservoirs. Linear regression lines are shown in color and 1:1 line in black.

4.8. Estimating predictive uncertainty in turbidity-flow relations

4.8.1. Background

Like most utilities managing surface water supplies, turbidity control is a critical function of New York City (NYC) water supply, currently operating under a renewable filtration avoidance determination granted by the New York State Department of Health and the US Environmental Protection Agency. While water quality is usually pristine, high magnitude runoff events can cause significant increases in stream and reservoir turbidity, which at times limits the use of this unfiltered drinking water supply (Effler *et al*. 1998, Gelda *et al*. 2009). A logical strategy for turbidity control is to reduce the diversions or temporarily remove the turbid reservoirs from service. It is thus important to understand how long a particular turbidity event will affect the system. In order to understand the duration, intensity and total load of a turbidity event, predictions of future in-stream turbidity values are important. Traditionally, turbidity predictions have been carried out by applying turbidity-flow rating curves. However, predictions from rating curves are often inaccurate due to inter and intra event variability in turbidity-flow relations. Variability in sediment-flow relations has been reported in several studies (Asselman 1999, Lenzi and Marchi 2000, Seeger *et al*. 2004, Walling and Webb 1988, Zabaleta *et al*. 2007). In the Catskill streams this variability can be due to turbidity-flow hysteresis, geologic factors, and antecedent conditions. Our previous analysis has shown that for a given value of streamflow the rising limb of streamflow hydrograph contributed higher turbidity compared to the falling limb due to sediment source depletion as an event progresses (Mukundan *et al*. 2013). While it is often difficult to make accurate predictions of stream turbidity given the variability in erosion rates and sediment transport in space and time within a watershed, quantifying uncertainty in stream turbidity predictions will be valuable for predicting reservoir turbidity and for operational decisions for maintaining high water quality in the NYC water supply.

4.8.2. Method

A turbidity rating curve is an empirical relationship between measured stream turbidity and streamflow. In NYC Catskill streams a second order polynomial (non-linear least square regression) is commonly used to model this relationship which is expressed as

$$LogT = a\ LogQ^2 + b\ LogQ + c \qquad 4.18$$

where T is the turbidity in NTU; Q is the streamflow;a,b, and c are regression coefficients.

The quantile regression approach was used to estimate regression coefficients at various quantiles (e.g. 0.5) of this general regression model in addition to regression on the mean. This approach can capture the broad range of variability in measured data particularly at high streamflows. Thus it is possible to derive a probabilistic distribution of turbidity value for a given streamflow value by using multiple quantile-dependent turbidity rating curves instead of a single one that provides a mean estimate. Shiau and Chen (2015) applied this approach to a power law function sediment rating curve for a river in southern Taiwan. Theoretical details of quantile regression can be found in Koenker (2005). The quantile regression based turbidity rating curve for equation 1.1 for various quantiles can be expressed as

$$LogT_q = a_q\ LogQ^2 + b_q\ LogQ + c_q \qquad 4.19$$

where q is the quantile which can range between 0 and 1; T_q is a conditional turbidity value at quantile q; a_q, b_q, and c_q are quantile-dependent regression coefficients determined for a selected quantile by minimizing the sum of asymmetrically weighted absolute values of residuals as

$$min\left[(q)\sum_{LogT_i \geq \widehat{LogT_i}}\left|LogT_i - \widehat{LogT_i}\right| + (1-q)\sum_{LogT_i < \widehat{LogT_i}}\left|LogT_i - \widehat{LogT_i}\right|\right] \qquad 4.20$$

where $\widehat{LogT_i}$ is the predicted value of $LogT_i$. The first component of the sum in the above expression is the total of vertical distances of data points that are above the fitted line at a selected quantile (e.g. 0.75). The second component of the sum is for all data points that are below the fitted line for the selected quantile. For a selected quantile, observations above the regression line are given a weight of q and those below the line are given a weight of 1-q. This approach is different from estimates of least squares that involves the minimization of sum of squares of the residuals. Quantile-dependent regression coefficients were estimated numerically by solving an algorithm developed by Koenker and d'Orey (1987) using the R-package *quantreg* (Koenker 2014).

4.8.3. Test Application

The methodology outlined above was applied to a turbidity-flow rating curve used to predict average daily stream turbidity for the Esopus Creek at Coldbrook, using average daily streamflow and flow weighted average daily turbidity (calculated from turbidity and flow measurements made at 15-minute intervals). A total of 736 pairs of average daily flow and daily turbidity values between 6/13/2003 to 5/17/2016 were included in the analysis. Data collected during the period immediately after Hurricane Irene and Tropical Storm Lee (Aug-Sept 2011) and up to 12/31/2012 were excluded from the analysis to avoid biases resulting from the effect of extreme events and post-event restoration efforts carried out in the streams.

A comparison of turbidity estimates using non-linear least square (NLS) regression (Equation 1.1) and quantile regression at the 0.5 quantile (median) with observed turbidity is shown in Figure 4.68. A regression line of simulated vs. observed turbidity shows that NLS regression over-predicted the observed values whereas quantile regression method under-predicted the observations. However, a vast majority of the high turbidity values that may create water quality concerns to NYC water supply, when predicted using quantile regression were closer to the 1:1 regression line compared to NLS estimates. Overall predictive capability of the two regression models were compared using model evaluation statistics and is shown in Table 4.14.

Table 4.14. Model performance evaluation statistics

Statistic	**Non-linear least square (NLS) regression**	**Quantile regression (0.5 quantile)**
Coefficient of determination, R^2	0.83	0.83
Nash-Sutcliffe coefficient, NSE	0.77	0.81
Mean Absolute Error (NTU)	13.65	13.07
Root Mean Square Error (NTU)	42.21	38.36

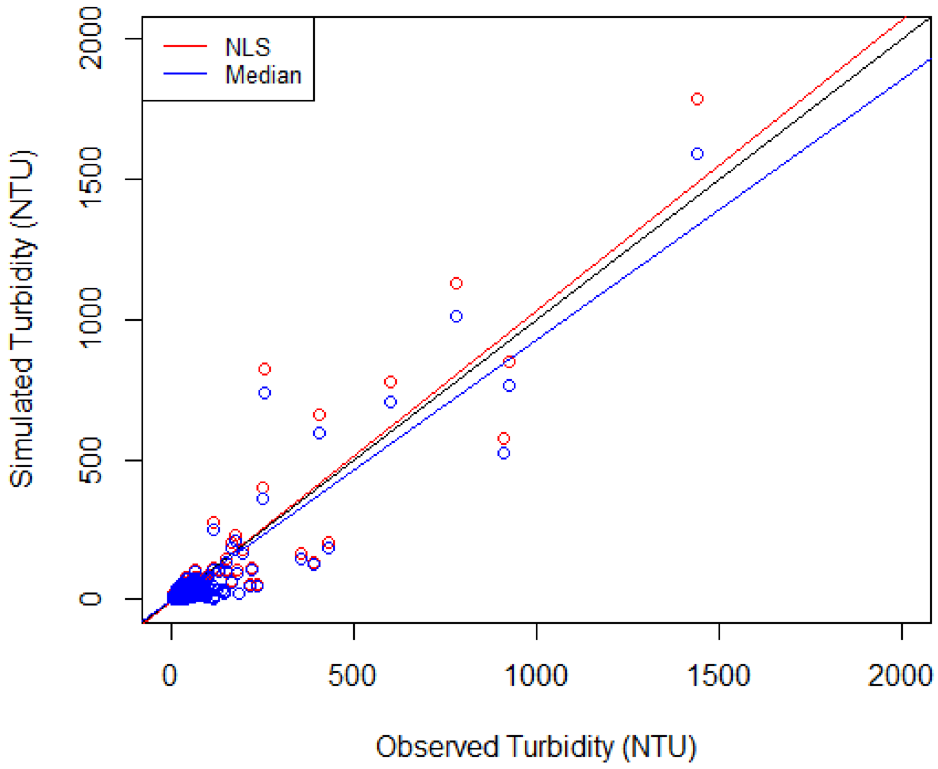


Figure 4.68. Comparison of simulated vs. observed turbidity for non-linear least square (NLS) regression and 0.5 quantile (median) regression estimates. The 1:1 regression line is shown in black.

In addition to the 0.5 quantile (median), turbidity estimates were made at various quantiles (0.1, 0.25, 0.75, and 0.9) to obtain the uncertainty in turbidity predictions for the range of observed flows (Figure 4.69). This information was used to generate the 50% and 80% prediction intervals of estimated turbidity as illustrated in Figure 4.70.

Performance of quantile regression based turbidity-flow rating curve (0.5 quantile) for the Esopus Creek at Coldbrook were comparable or better when compared to non-linear least square regression estimates. Additional testing is needed at multiple sites to further evaluate the method. Quantile regression based method provides the ability to estimate and quantify the wide range in uncertainty observed in turbidity-flow relations.

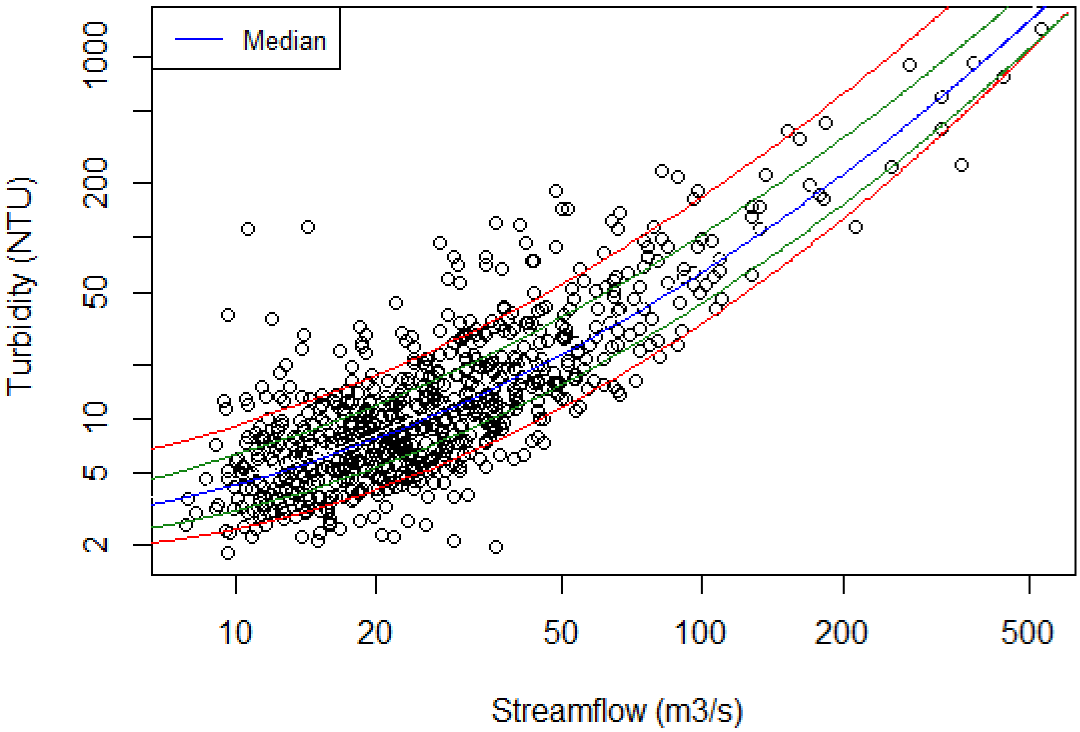


Figure 4.69. Quantile regression based turbidity rating curves for Esopus Creek at Coldbrook. Quantiles shown are 0.9, 0.75, 0.5, 0.25, and 0.1 from top to bottom.

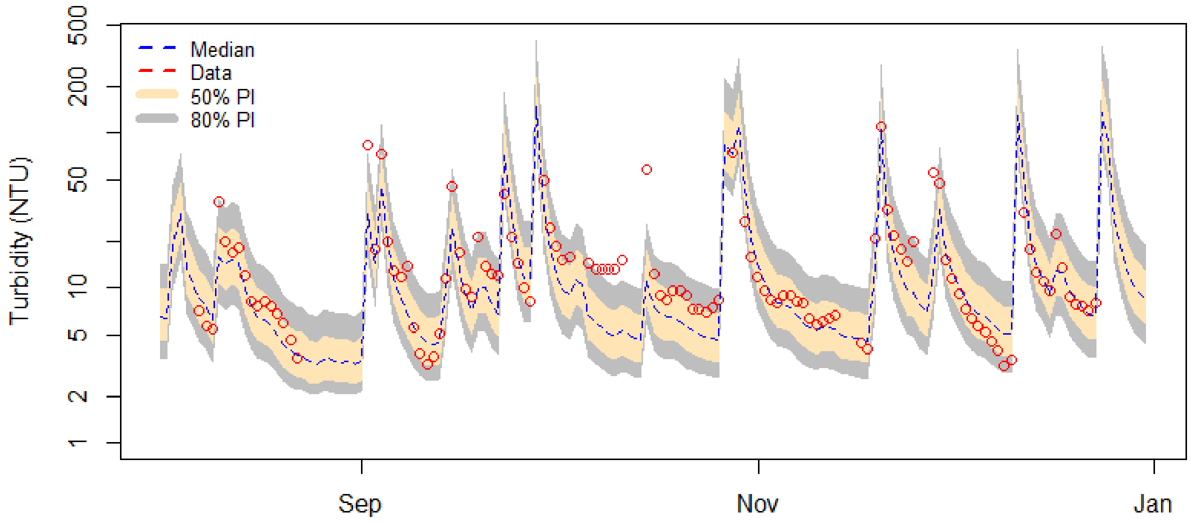


Figure 4.70. Illustration of uncertainty in turbidity predictions estimated using quantile regression method. Values shown are turbidity for the Esopus Creek at Coldbrook for the period 08/01/2003 to 12/31/2003.

4.8.4. Comparison of methods used to estimate turbidity rating curve coefficients

- Coefficients based on quantile regression are robust and are estimated by minimizing the sum of absolute residuals compared to minimizing sum of squared residuals in least square estimates.

- Quantile regression based estimates of regression coefficients are more efficient than those based on least squares when the error terms are not normally distributed or when there are outliers in the dataset as often seen in turbidity-flow relations.

- Coefficients are derived iteratively in quantile regression using least square estimates as the initial value. Coefficient are derived analytically in least square estimates.

- Least square estimates provide one best fitting regression line whereas quantile regression may find more than one best fitting line. This will enable users of the quantile regression approach to select or modify the best fitting line (e.g. 0.52 or 0.55 quantile) for the range of flow and turbidity conditions that are of interest.

5. Data Analysis to Support Modeling

5.1. Meteorological data analyses to support development of turbidity models for Cannonsville, Pepacton, and Neversink Reservoirs

5.1.1. Introduction

Turbidity models for Cannonsville, Pepacton, and Neversink reservoirs are planned for development, which will be based on the hydrodynamic and transport framework of CE-QUAL-W2 (W2; Cole and Wells 2013). Similar turbidity models, based on W2, for Rondout, Schoharie, Ashokan, and Kensico reservoirs had been developed in earlier studies (Gelda and Effler 2007, Gelda *et al*. 2009, 2012, 2013). W2 requires specifications of surface boundary conditions (surface heat exchange, solar radiation absorption, wind stress, and gas exchange) calculated from meteorological parameters. These parameters are air temperature, dew point temperature (or relative humidity), wind velocity, and solar radiation (or cloud cover), all specified at an hourly time step.

DEP began installing meteorological stations throughout the NYC watershed since mid-1990s, including some stations on the reservoir monitoring buoys (RoboMon). Data with varying duration of records are available at these sites. Common problems with these data are (i) missing observations, (ii) systematic errors (i.e., bias) due to long-term drift of sensors, (iii) errors caused by malfunctioning of instruments, and (iv) errors made during data processing. In addition, data from certain sites are not representative of the local reservoir conditions or simply, the duration of available data is not long-enough for model calibration and validation purposes.

This study was conducted to develop representative and continuous, long-term (e.g., greater than 60 years) records of meteorological data for input to the reservoir models, and eventually input to Operations Support Tool (OST).

5.1.2. Methods

The general approach consisted of using data from DEP's local watershed stations as well as NOAA's regional weather stations and establishing empirical inter-site relationships for various parameters. Such relationships are then used to fill-in missing observations, check consistency and quality of the observations, and extrapolate to time intervals outside the records of observations.

The stations considered in this study are Binghamton Airport site operated by NWS of NOAA, and sites DCM074, DPM110, DPM114, DNM146, DNM152 operated by DEP (Table 5.1).

Table 5.1. Summary of meteorological data.

Location	Site	Years	Agency
Cannonsville Reservoir Dam	DCM074	1995-2015	DEP
Pepacton Reservoir Dam	DPM110	1997-2015	DEP
Pepacton Watershed	DPM114	2013-2015	DEP
Neversink Reservoir Dam	DNM146	1997-2015	DEP
Neversink Watershed	DNM152	2012-2015	DEP
Greater Binghamton Airport	4725	1948-2015	NOAA

For DEP sites, dew point temperature was computed from relative humidity and air temperature (Lawrence 2005). For NOAA site, solar radiation was estimated from cloud cover data (Cole and Wells 2013). Hourly observations were averaged to daily values. Hourly observations of wind velocity were vector averaged and then the X-Y components were used to establish inter-site relationships. Linear regression models were fitted to the paired data from different sites for air temperature, dew point, X-Y wind components, and solar radiation, while forcing the intercept value to zero. The slope value of the regressions thus can be applied to hourly observations without altering the dynamics of diurnal variations that may be present in the data.

5.1.3. Results and Discussion

Evaluation of the inter-site regression relationships from the point of view of establishing meteorological inputs for Cannonsville, Pepacton, and Neversink reservoirs is presented in Figure 5.1–Figure 5.11. Generally, excellent relationship was obtained for air temperature and dew point (i.e., slope value close to 1) in all pairings of the data. Relationships for solar radiation exhibited relatively greater scatter even though the slope value remained approximately equal to 0.9 or greater indicating random nature of the variability among different locations for this variable. Relationships for the components of wind velocity were quite variable perhaps due to strong effects of local topography. It is proposed that the following general hierarchy be used to build long-term input data files for hydrodynamic, water quality models and OST: (i) use off-shore data from reservoir buoys when available, (ii) use on-shore (adjusted to reservoir surface if paired measurements available) data, (iii) use adjusted data from local sites in the watershed, (iv) use adjusted data from the regional NOAA sites. When no buoy data are available, as in the current study, all other data will be adjusted to the on-shore site data. Finally, wind sheltering coefficient (WSC) in the models will be calibrated to achieve a representative wind field over the reservoir surface.

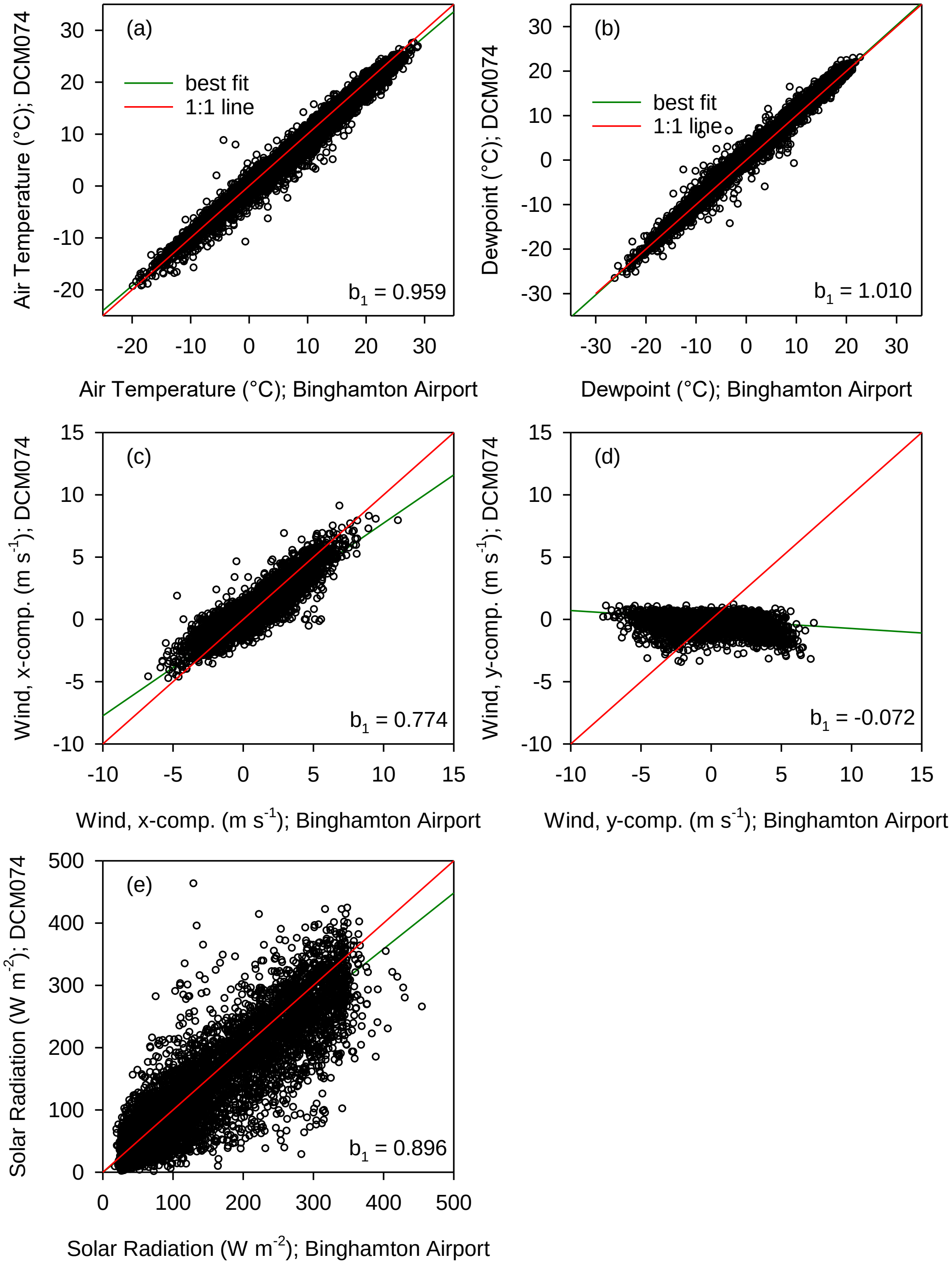


Figure 5.1. Evaluation of relationship between daily average meteorological data observed at Binghamton (New York) Airport and Cannonsville Reservoir site DCM074, for 1994–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

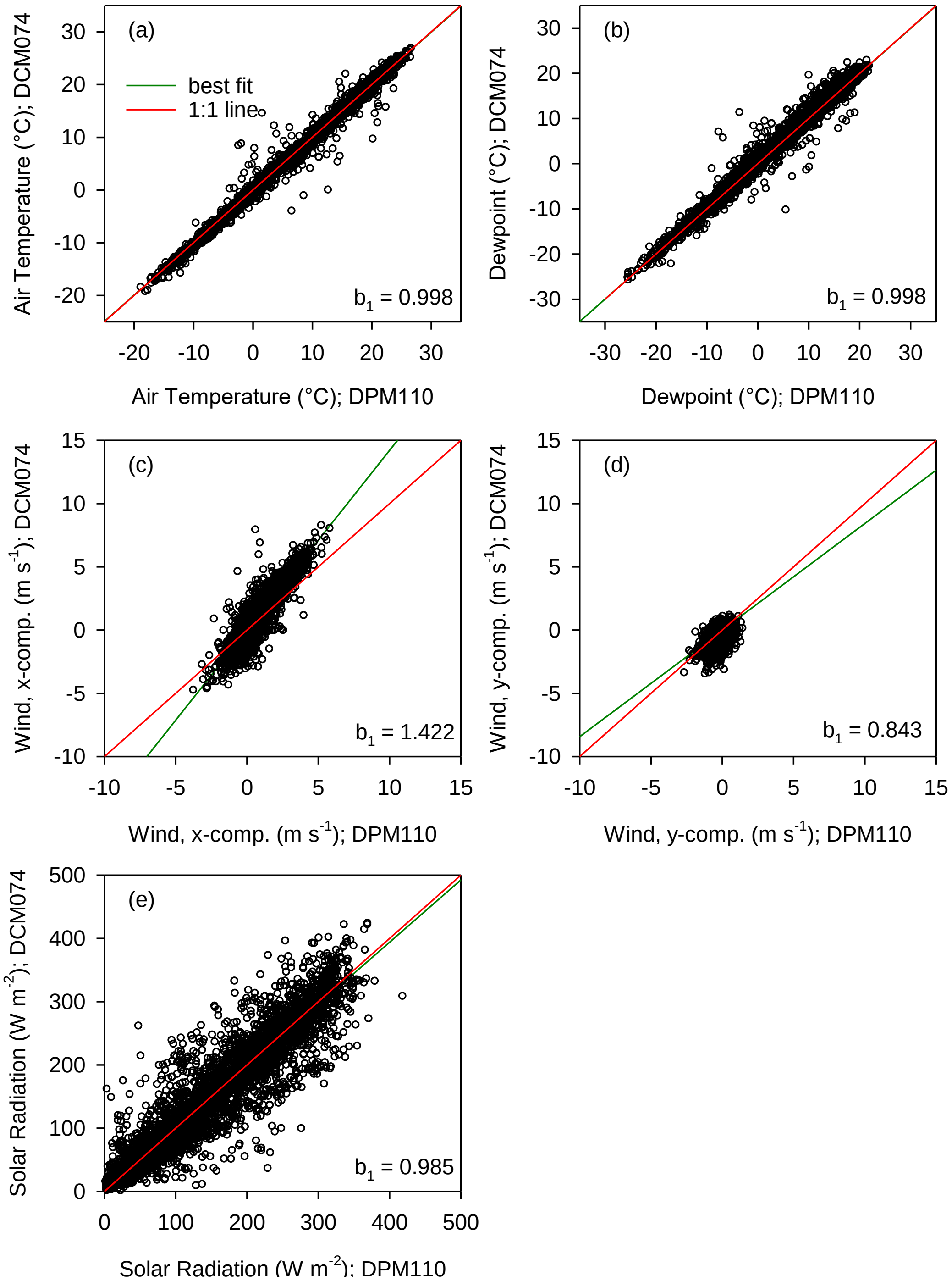


Figure 5.2. Evaluation of relationship between daily average meteorological data observed at Pepacton Reservoir site DPM110 and Cannonsville Reservoir site DCM074, for 1996–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

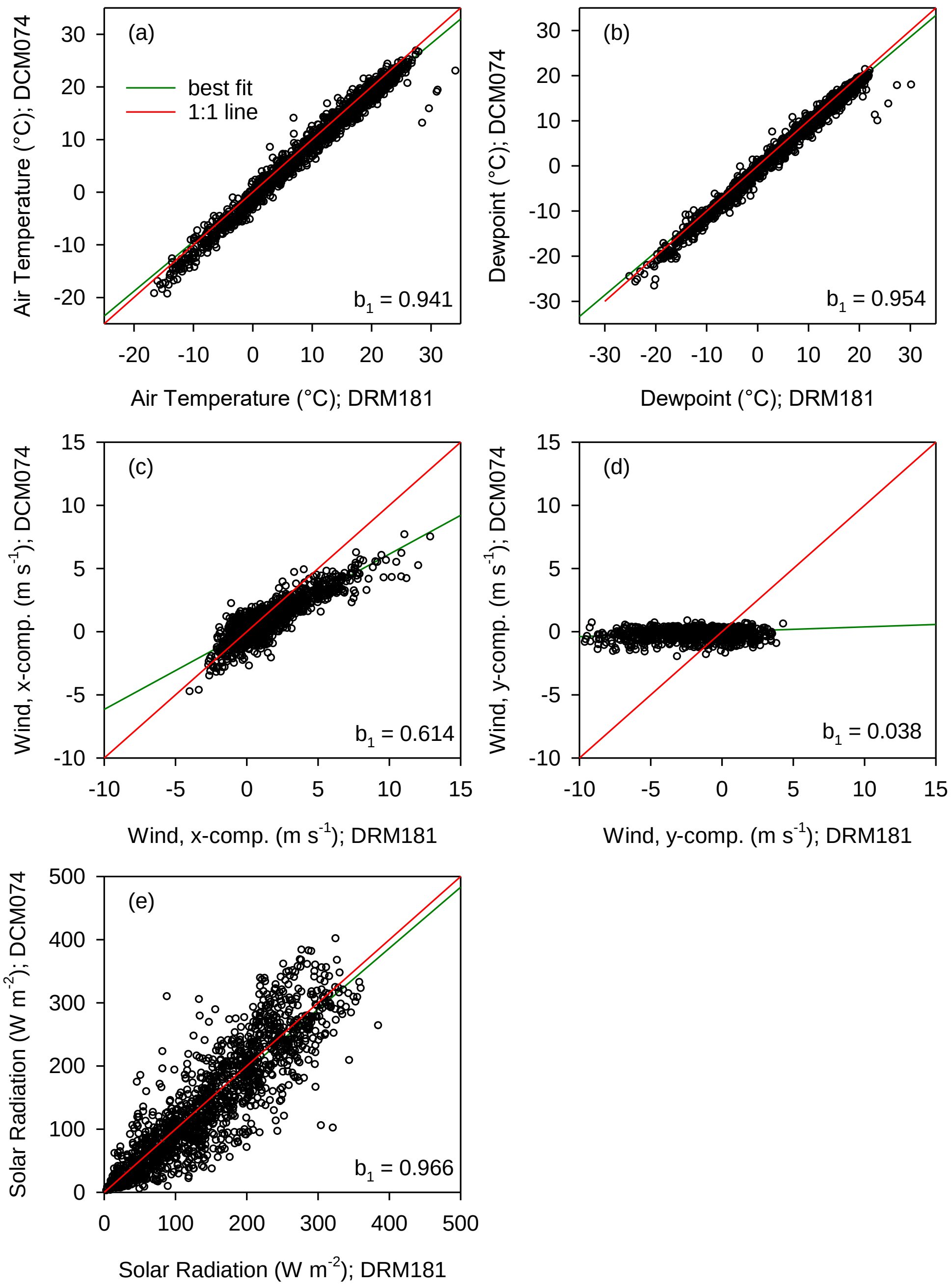


Figure 5.3. Evaluation of relationship between daily average meteorological data observed at Rondout Reservoir site DRM181 and Cannonsville Reservoir site DCM074, for 2011–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

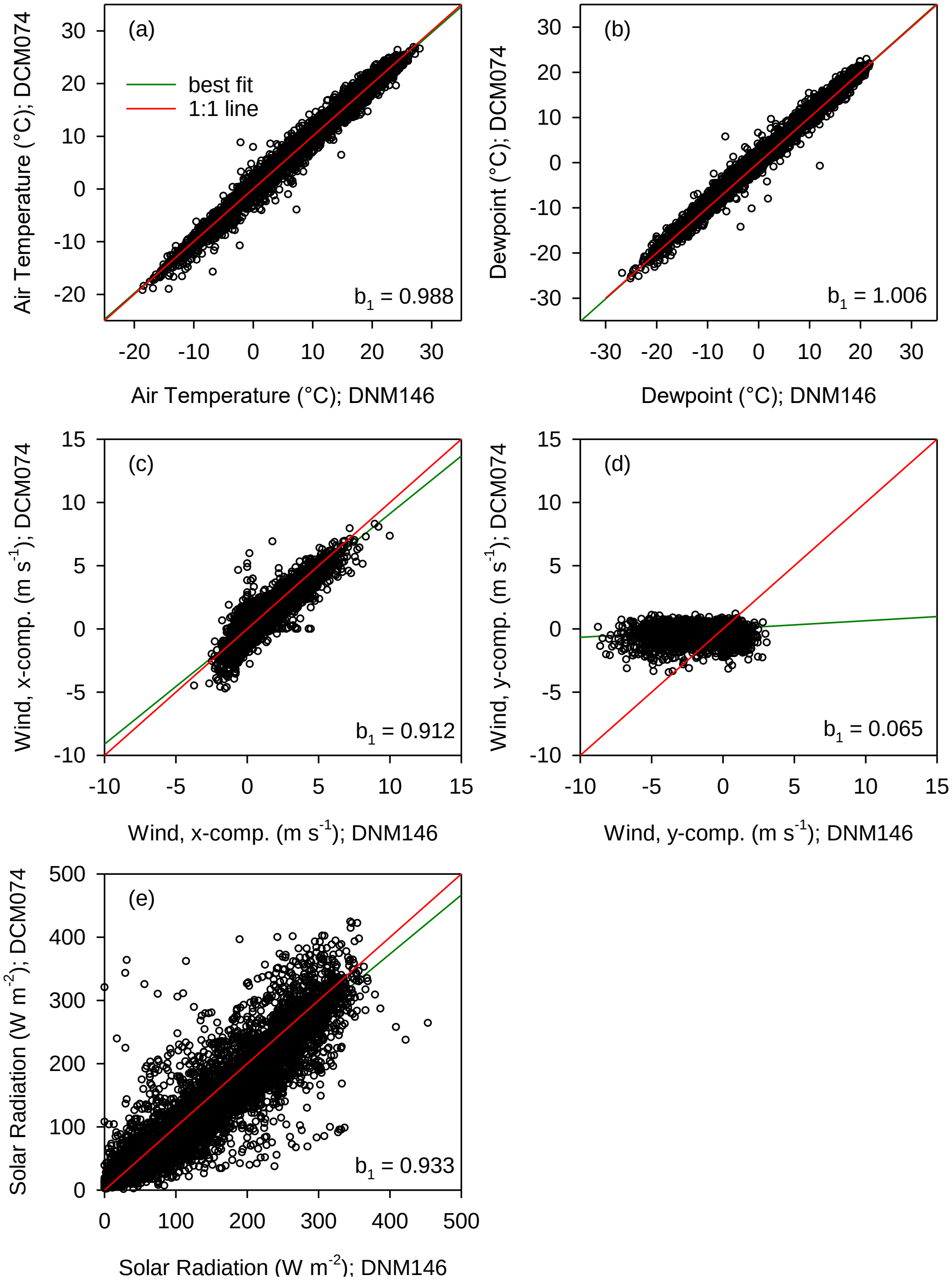


Figure 5.4. Evaluation of relationship between daily average meteorological data observed at Neversink Reservoir site DNM146 and Cannonsville Reservoir site DCM074, for 1996–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

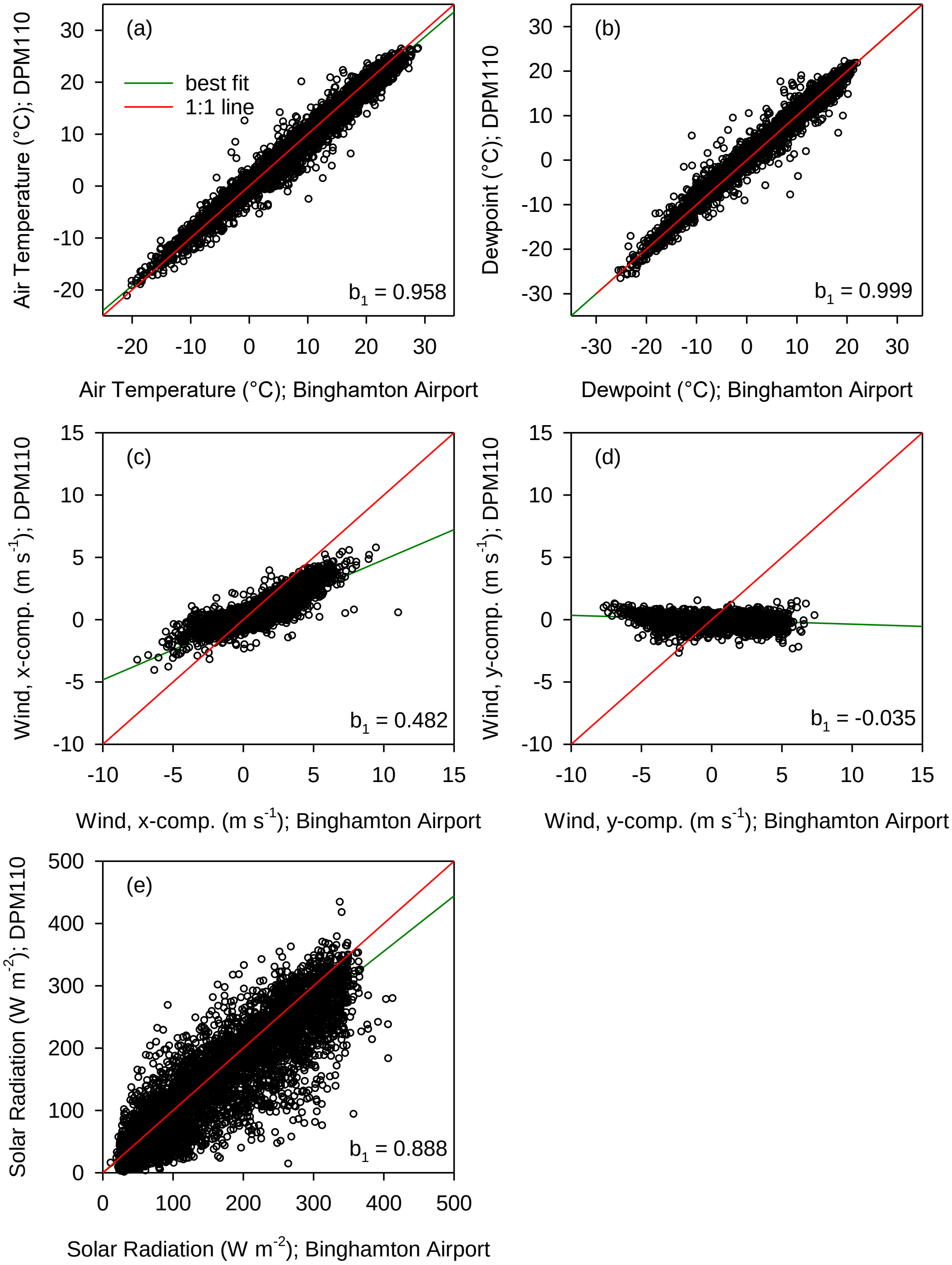


Figure 5.5. Evaluation of relationship between daily average meteorological data observed at Binghamton (New York) Airport and Pepacton Reservoir site DPM110, for 1996–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

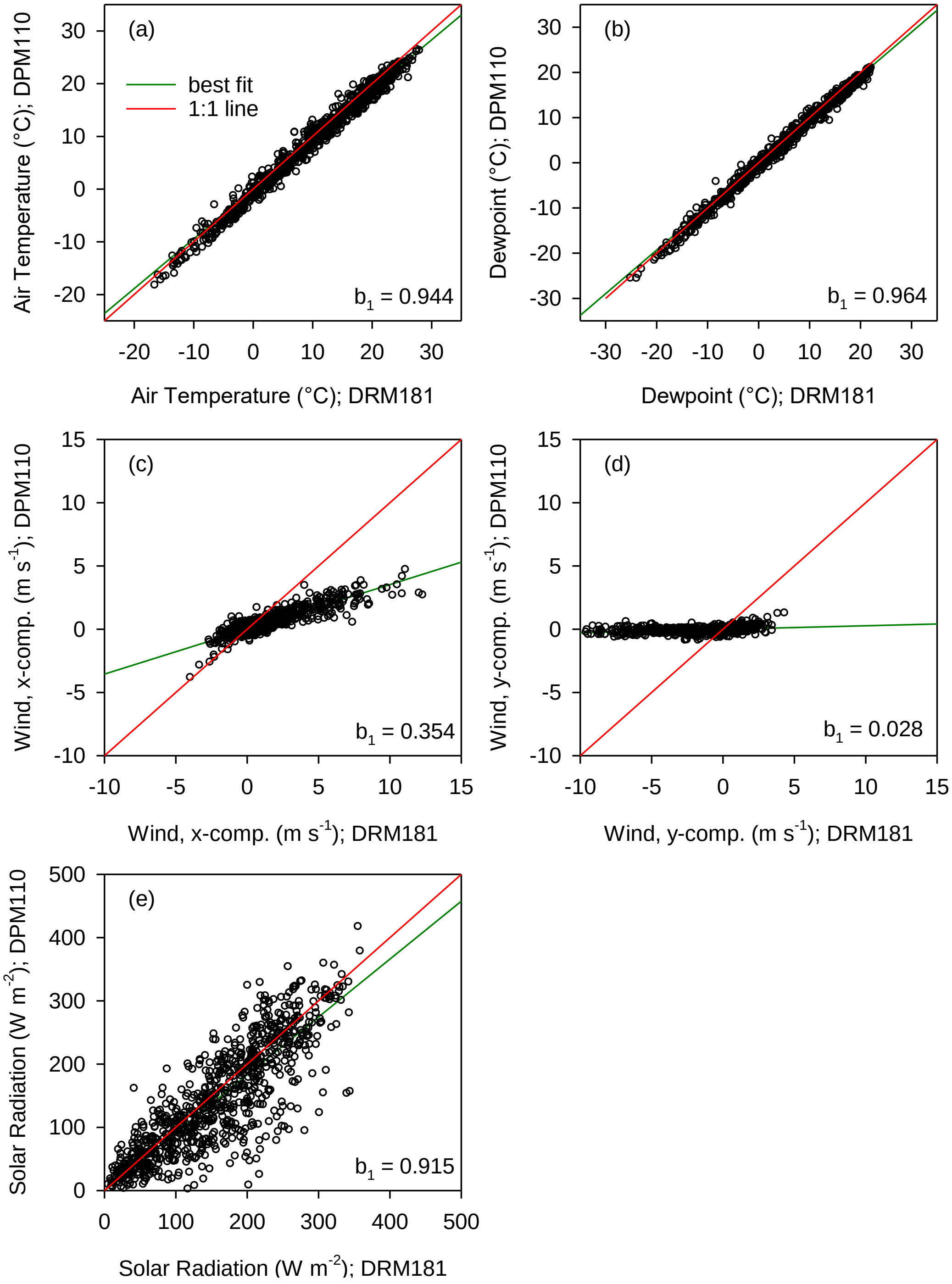


Figure 5.6. Evaluation of relationship between daily average meteorological data observed at Rondout Reservoir site DRM181 and Pepacton Reservoir site DPM110, for 2011–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

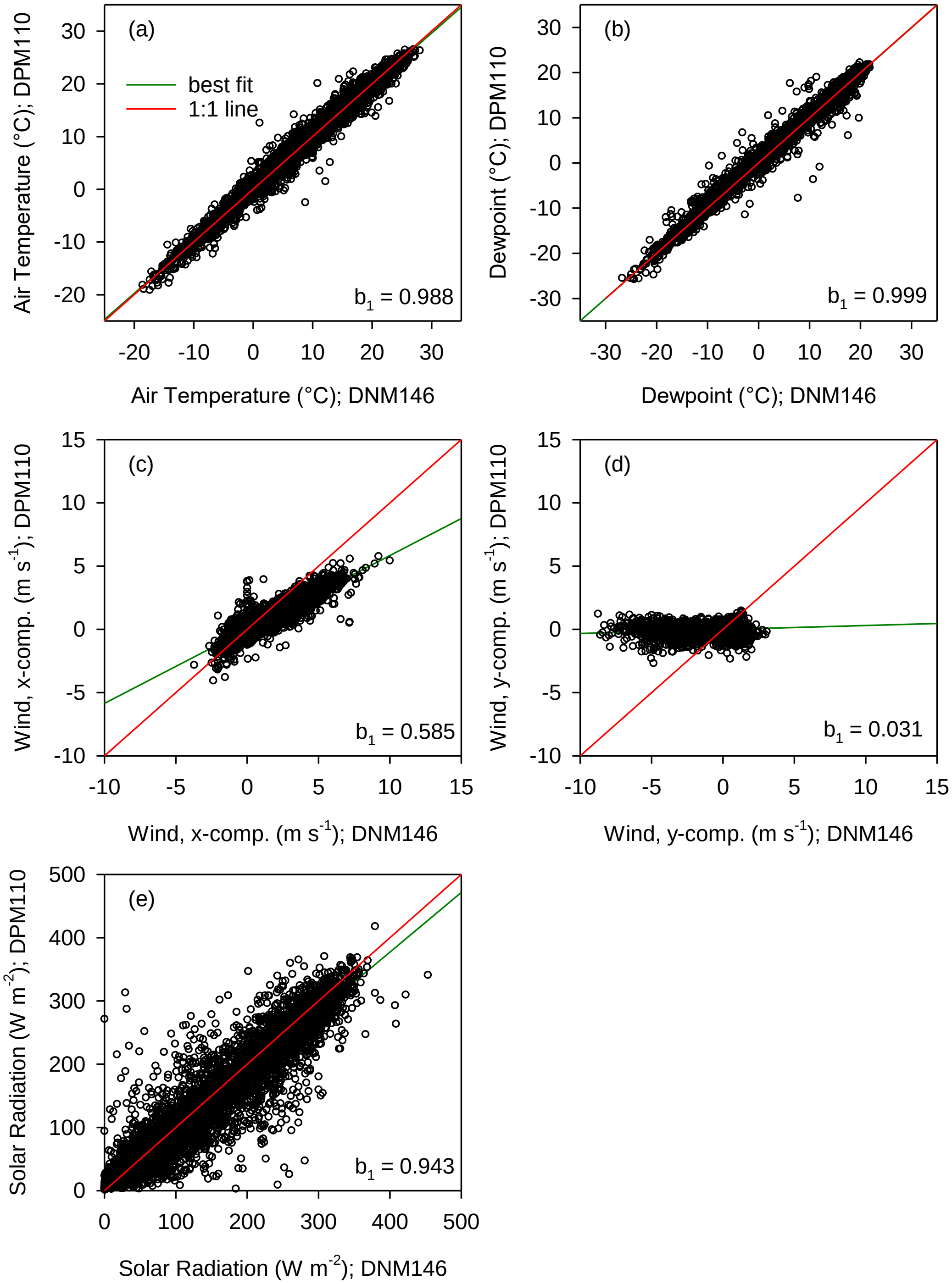


Figure 5.7. Evaluation of relationship between daily average meteorological data observed at Neversink Reservoir site DNM146 and Pepacton Reservoir site DPM110, for 1996–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

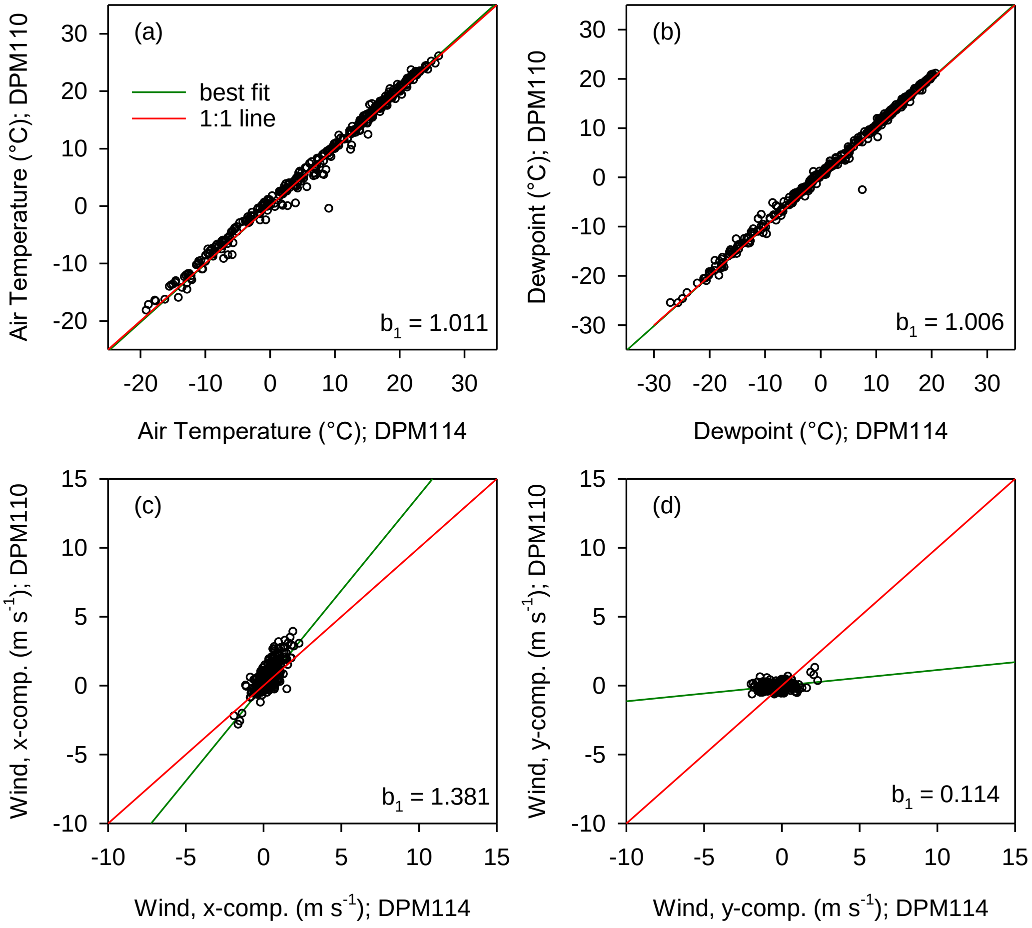


Figure 5.8. Evaluation of relationship between daily average meteorological data observed at Pepacton Reservoir sites DPM114 and DPM110, for 2013–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, and (d) y-component of wind speed. Solar radiation data at these two sites were not available for comparison.

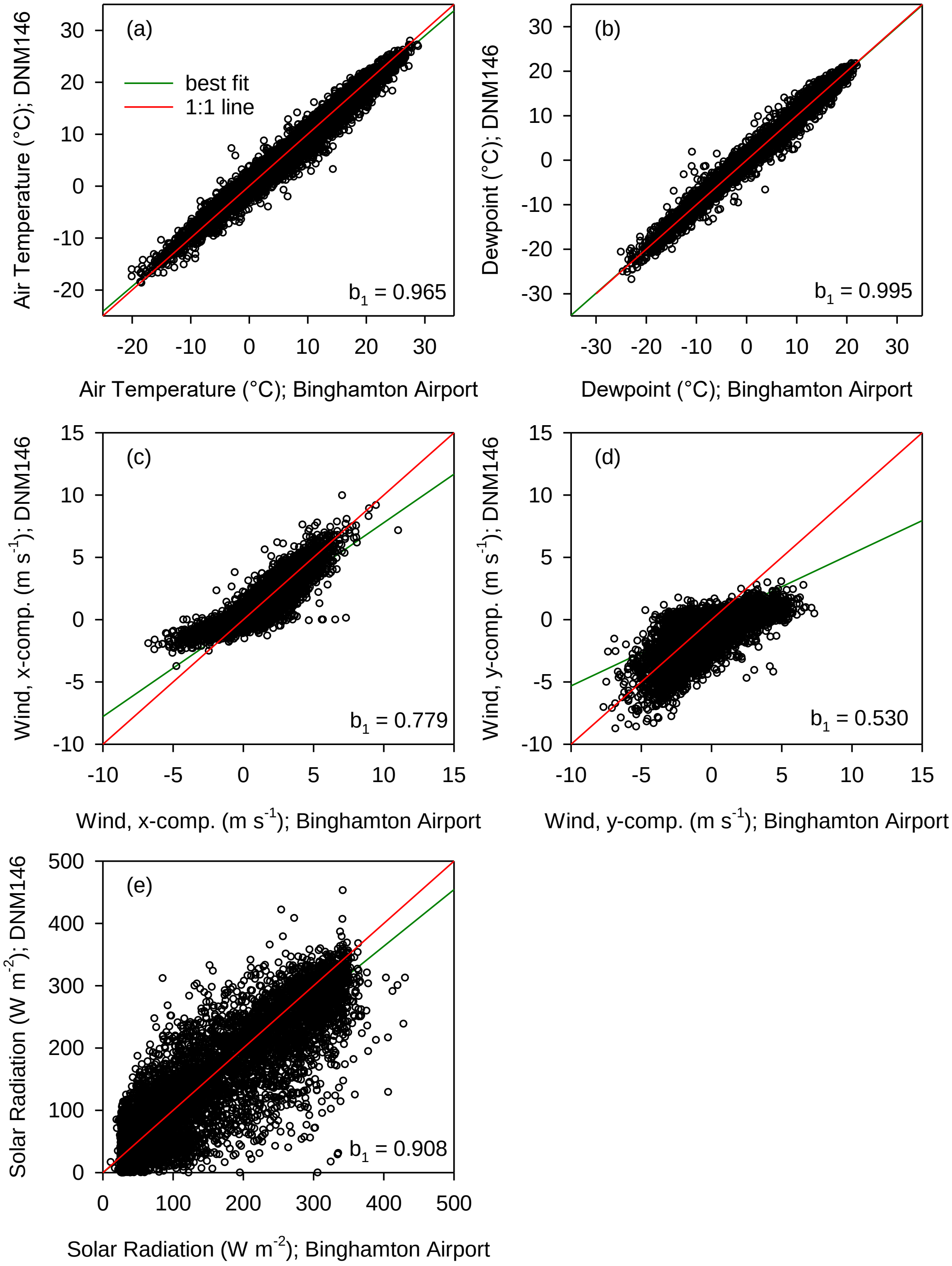


Figure 5.9. Evaluation of relationship between daily average meteorological data observed at Binghamton (New York) Airport and Neversink Reservoir site DNM146, for 1996–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

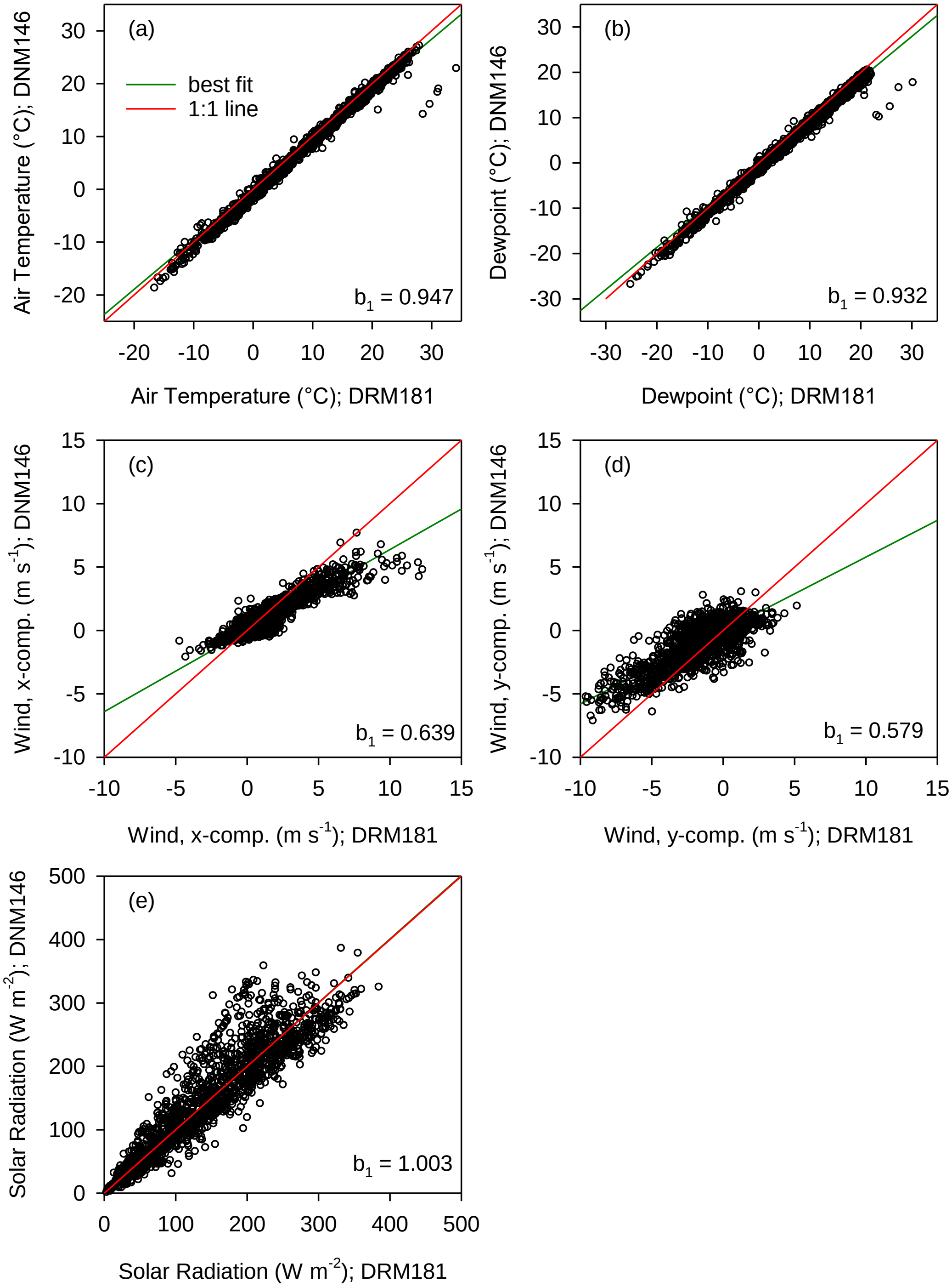


Figure 5.10. Evaluation of relationship between daily average meteorological data observed at Rondout Reservoir sites DRM181 and Neversink Reservoir site DNM146, for 2012–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

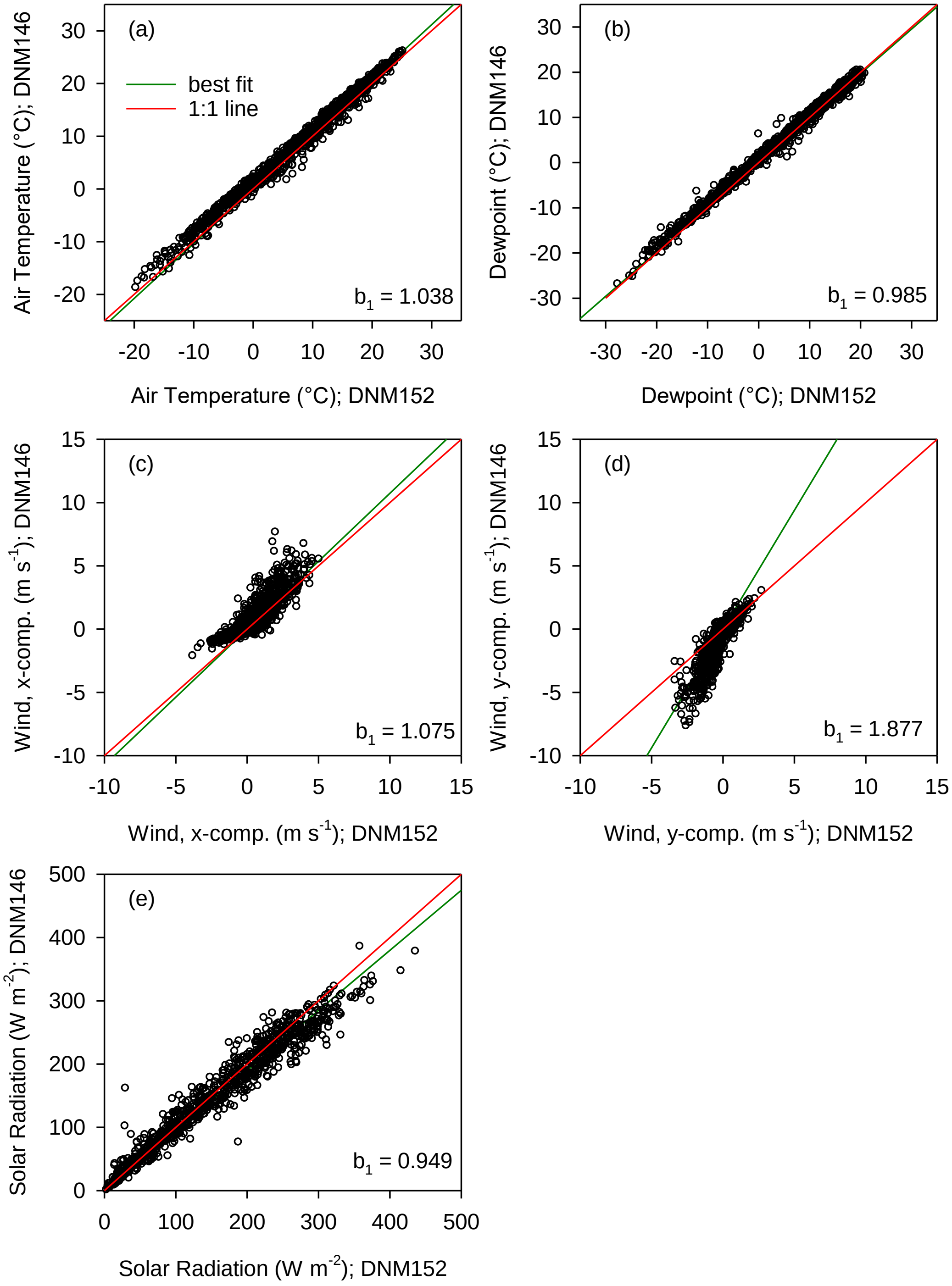


Figure 5.11. Evaluation of relationship between daily average meteorological data observed at Neversink Reservoir sites DNM152 and DNM146, for 2012–2015: (a) air temperature, (b) dewpoint temperature, (c) x-component of wind speed, (d) y-component of wind speed, and (e) solar radiation.

5.2. Development of polynomials for release temperatures below Cannonsville and Pepacton Reservoirs

The analysis presented here supports the development and application of a stream temperature model for the West and East Branch Delaware River below Cannonsville (site CNB) and Pepacton (site PDB) reservoirs. Polynomials were fitted through the temperatures observed at sites CNB and PDB located below Cannonsville and Pepacton dams, respectively. For planning and evaluation purposes, these polynomials may be used in OST to predict release temperatures and force the empirical stream temperature model (to be developed). Additionally, uncertainty in the predictions of release temperatures may be included in OST by conducting Monte-Carlo simulation within a range of ± 1 standard deviation or ± 10th and 90th percentiles of observed values.

This analysis will be integrated in an empirical model of daily average stream temperature for the lower Delaware reaches for the months of June through September. The model is expected to help develop short-term forecasts to estimate the volume of reservoir releases needed to keep water temperatures below a threshold to protect cold water fisheries from potentially harmful warm water.

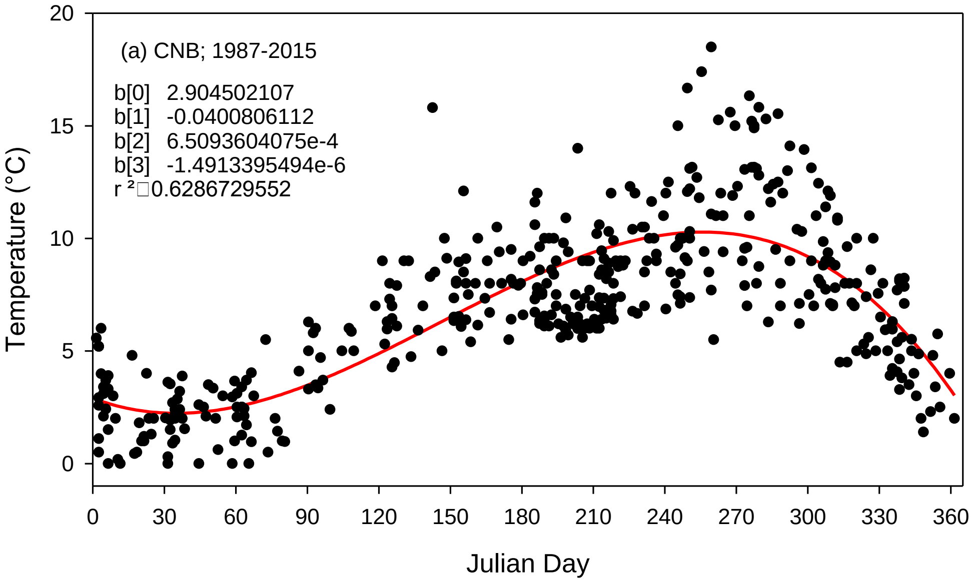


Figure 5.12. Temperatures observed at site CNB below Cannonsville Reservoir for 1987-2015. A best-fit polynomial is shown. Data points with spill flow > 0 are excluded.

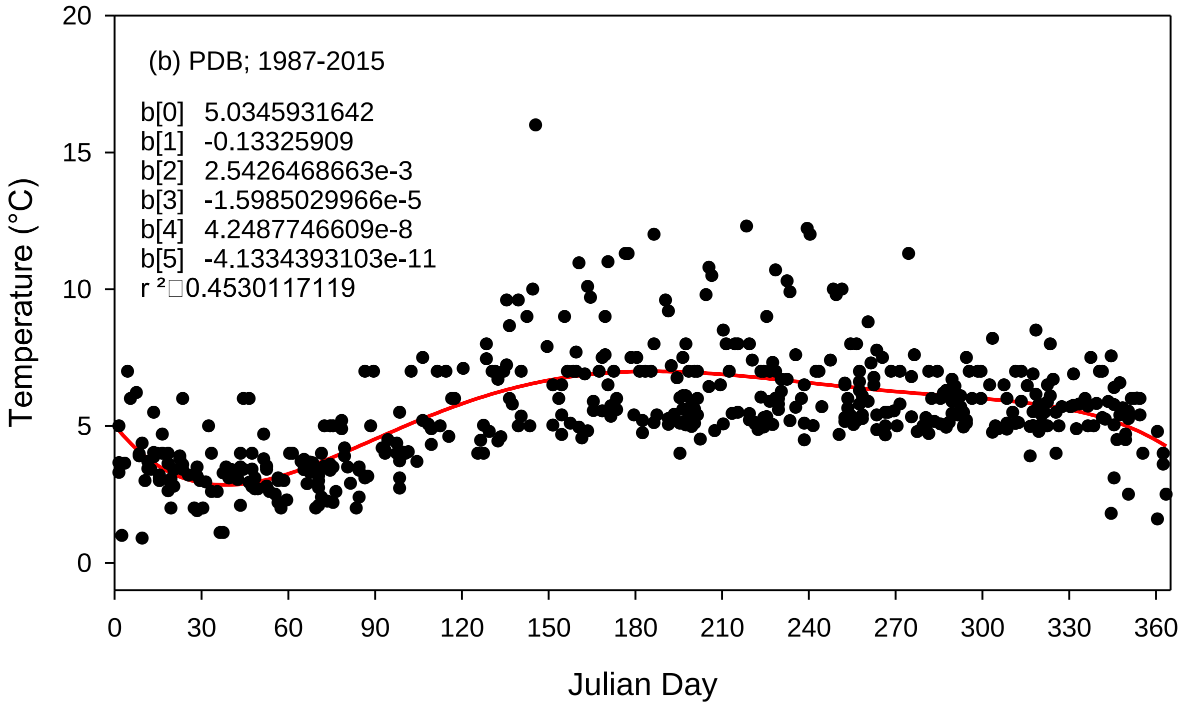


Figure 5.13. Temperatures observed at site PDB below Pepacton Reservoir for 1987-2015. A best-fit polynomial is shown. Data points with spill flow > 0 are excluded.

5.3. Development of an empirical stream temperature model for Neversink River

This section describes the development of an empirical regression model to predict historic temperatures in the Neversink River at Claryville, NY where it enters Neversink Reservoir. The predicted temperatures will be used as input to the long-term probabilistic water quality model for Neversink Reservoir that is currently under development. The data used to build the model were hourly average air temperature recorded at the Binghamton Airport, hourly average flow measured at Claryville, NY in Neversink River, and hourly average stream temperature at Claryville, NY in Neversink River. These data were for the period of 2012-2015.

Multiple nonlinear regressions were developed for individual months to specify water temperature T_w at an hourly frequency as a function of stream flow, Q and air temperature, T_{air}, according to the relationship:

$$T_{w,i} = a_0 + a_1 T_{air,i-2} + a_2 \log Q_i \tag{5.1}$$

where $T_{w,i}$ = Neversink River temperature (°C) for hour i; $T_{air,i-2}$ = air temperature (°C) for 2 hours before hour i; Q_i = stream flow ($m^3 \cdot s^{-1}$) for Neversink River for hour i; and a_0, a_1, and a_2 are regression coefficients. The analysis was based on paired values from 2012-2015 interval for which measurements of $T_{w,i}$ were available.

Model coefficients a_0, a_1, and a_2 were computed for each month used least squares regression (Table 5.2). Plots of observed and predicted water temperature T_w for individual months are shown in Figure 5.14 through Figure 5.25. The model(s) generally performed well at within-a-day and day-to-day time scales for the various months, as summarized by the performance statistics in the table below. Adding other independent variable(s), e.g., ice-conditions in the River may improve model performance for the winter months, however, such data are scarce and would severely limit model's application.

Table 5.2. Coefficients and performance statistics for the Neversink River temperature model.

	a_0	a_1	a_2	r^2	RMSE (°C)
January	1.544598	0.060403	-0.05482	0.404423	0.658601
February	1.34221	0.069384	0.183285	0.350289	0.604665
March	1.76543	0.212519	1.102734	0.591571	1.436336
April	6.560054	0.286242	-3.2646	0.59465	1.659607
May	7.292366	0.312197	-1.01666	0.60824	1.519336
June	8.841445	0.344087	-1.45693	0.703056	1.118783
July	11.17401	0.317921	-2.36552	0.667484	1.105493
August	9.808569	0.332045	0.39693	0.624923	1.006289
September	9.510929	0.335726	-0.36038	0.681606	1.256662
October	7.946925	0.303309	-0.61619	0.633223	1.223196
November	4.261754	0.284066	0.52697	0.591717	1.469512
December	2.921218	0.24825	0.228926	0.574917	1.232532

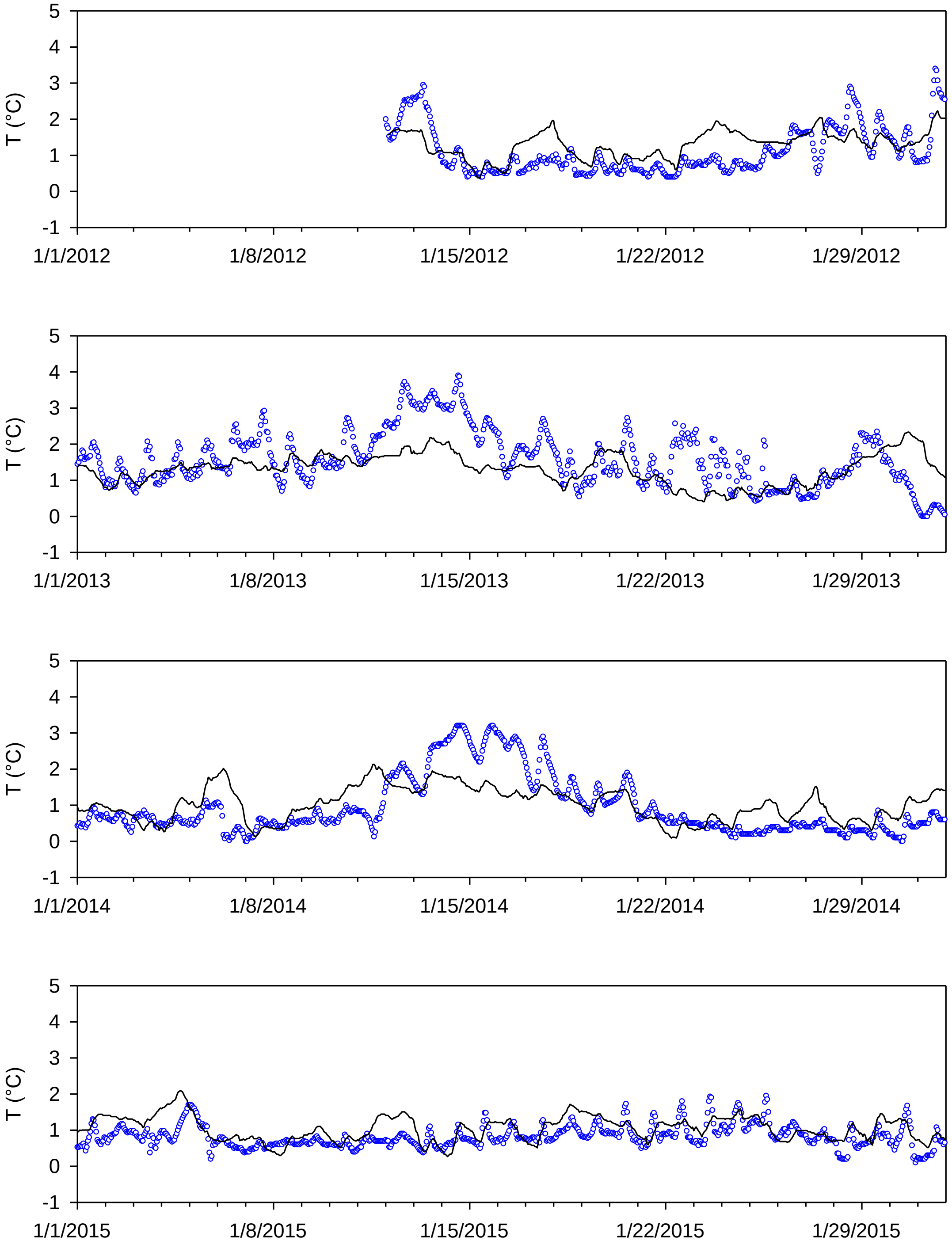


Figure 5.14. Performance of empirical temperature model for Neversink River for hourly observations in January 2012 – January 2015.

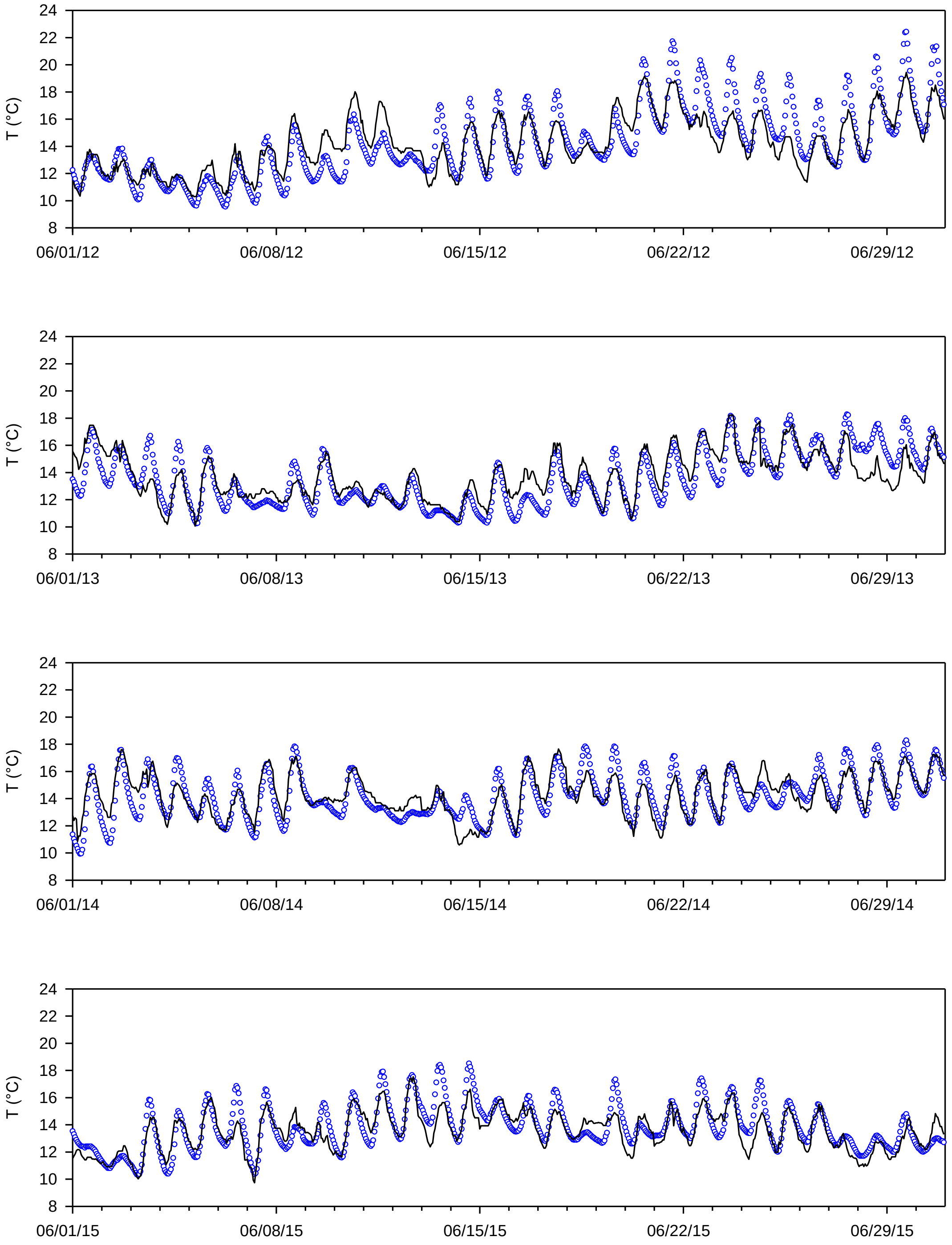


Figure 5.15. Performance of empirical temperature model for Neversink River for hourly observations in June 2012 – June 2015.

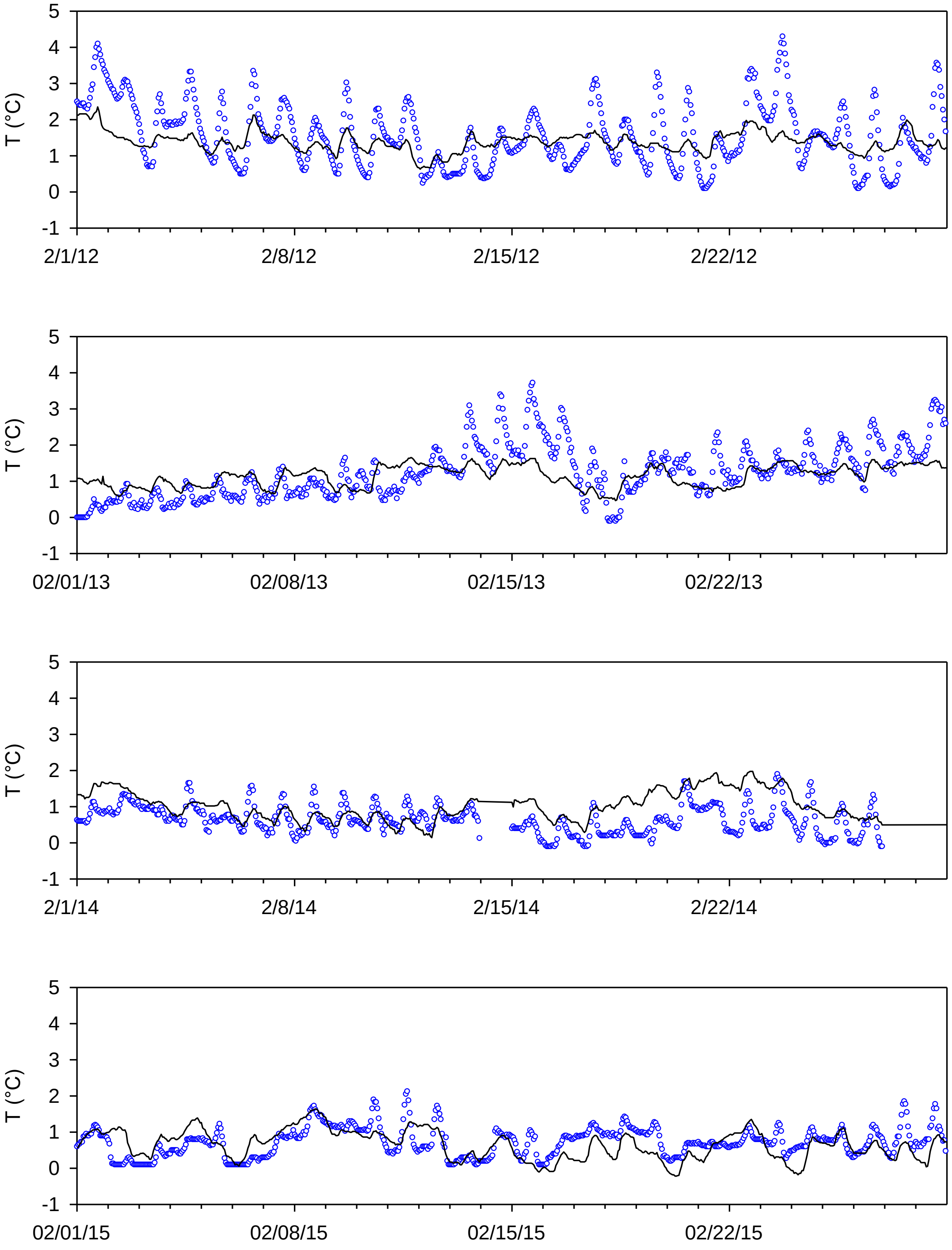


Figure 5.16. Performance of empirical temperature model for Neversink River for hourly observations in February 2012 – February 2015.

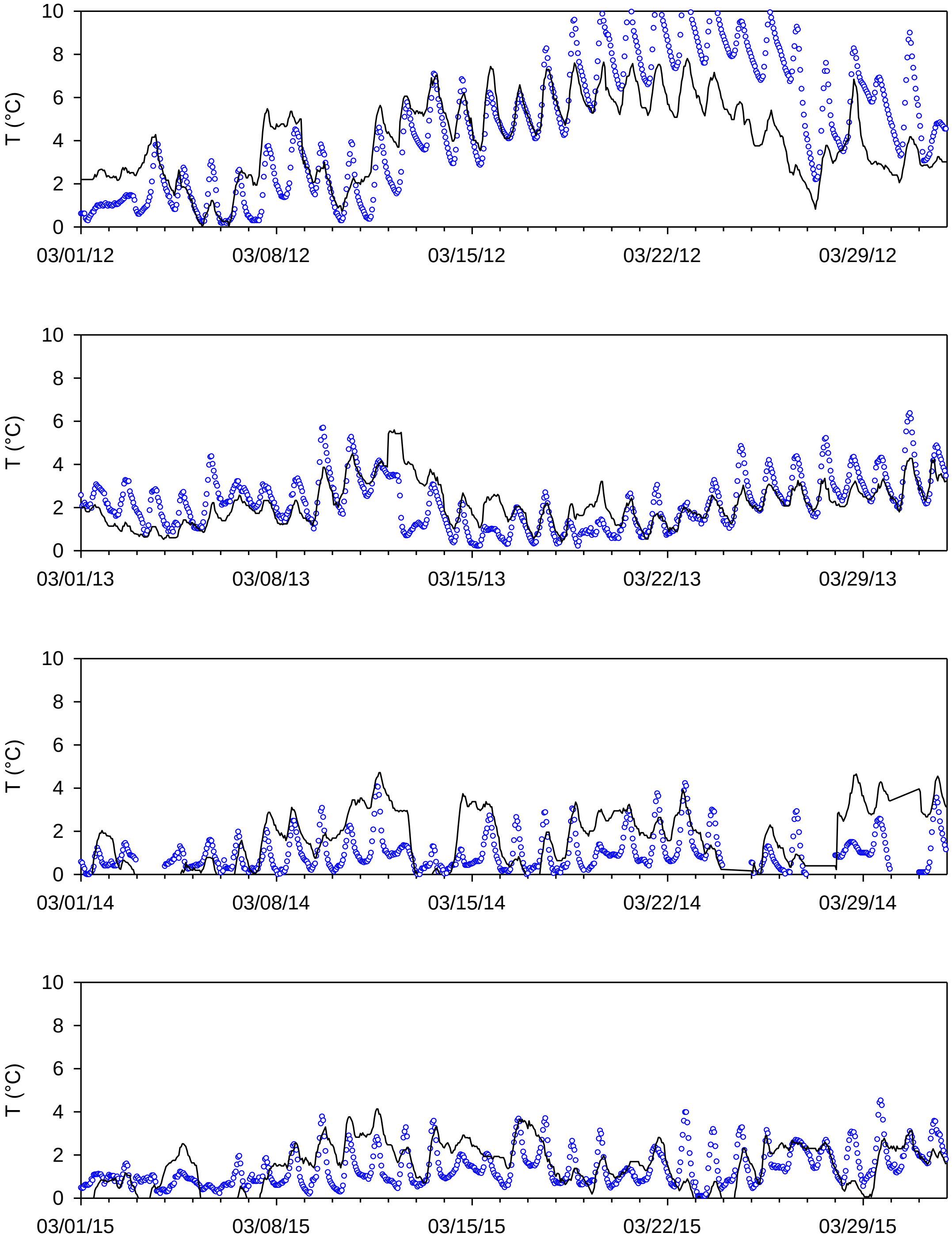


Figure 5.17. Performance of empirical temperature model for Neversink River for hourly observations in March 2012 – March 2015.

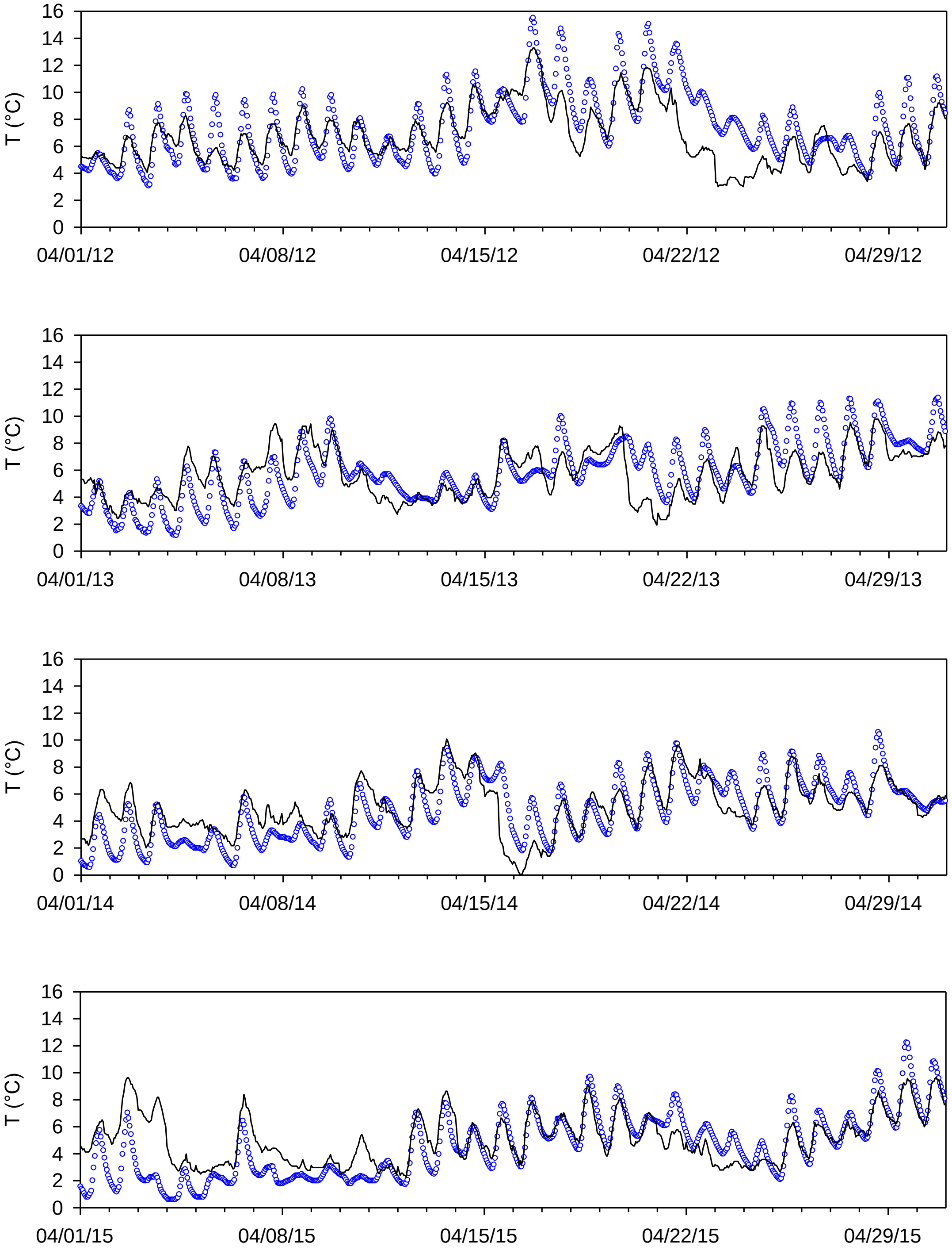


Figure 5.18. Performance of empirical temperature model for Neversink River for hourly observations in April 2012 – April 2015.

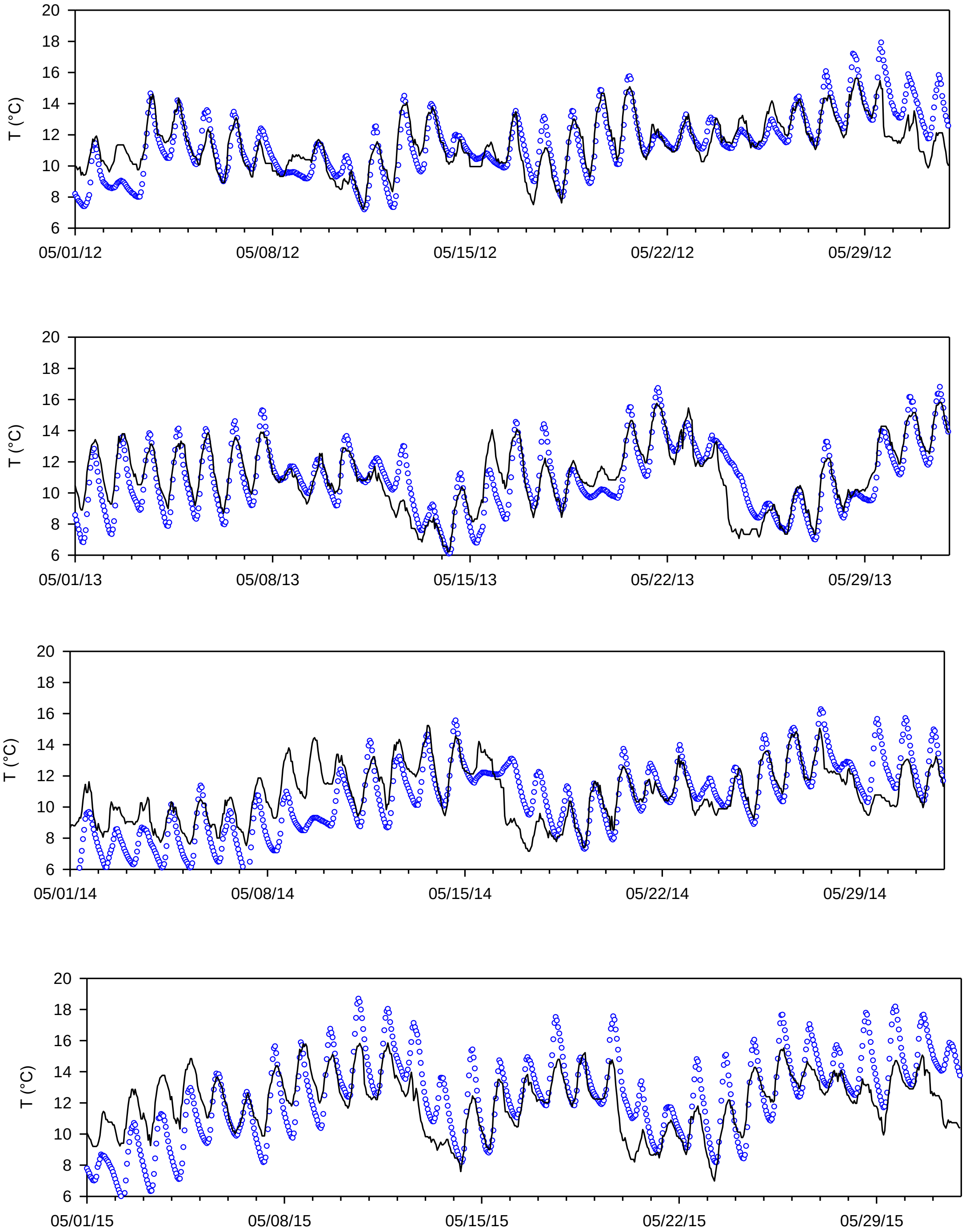


Figure 5.19. Performance of empirical temperature model for Neversink River for hourly observations in May 2012 – May 2015.

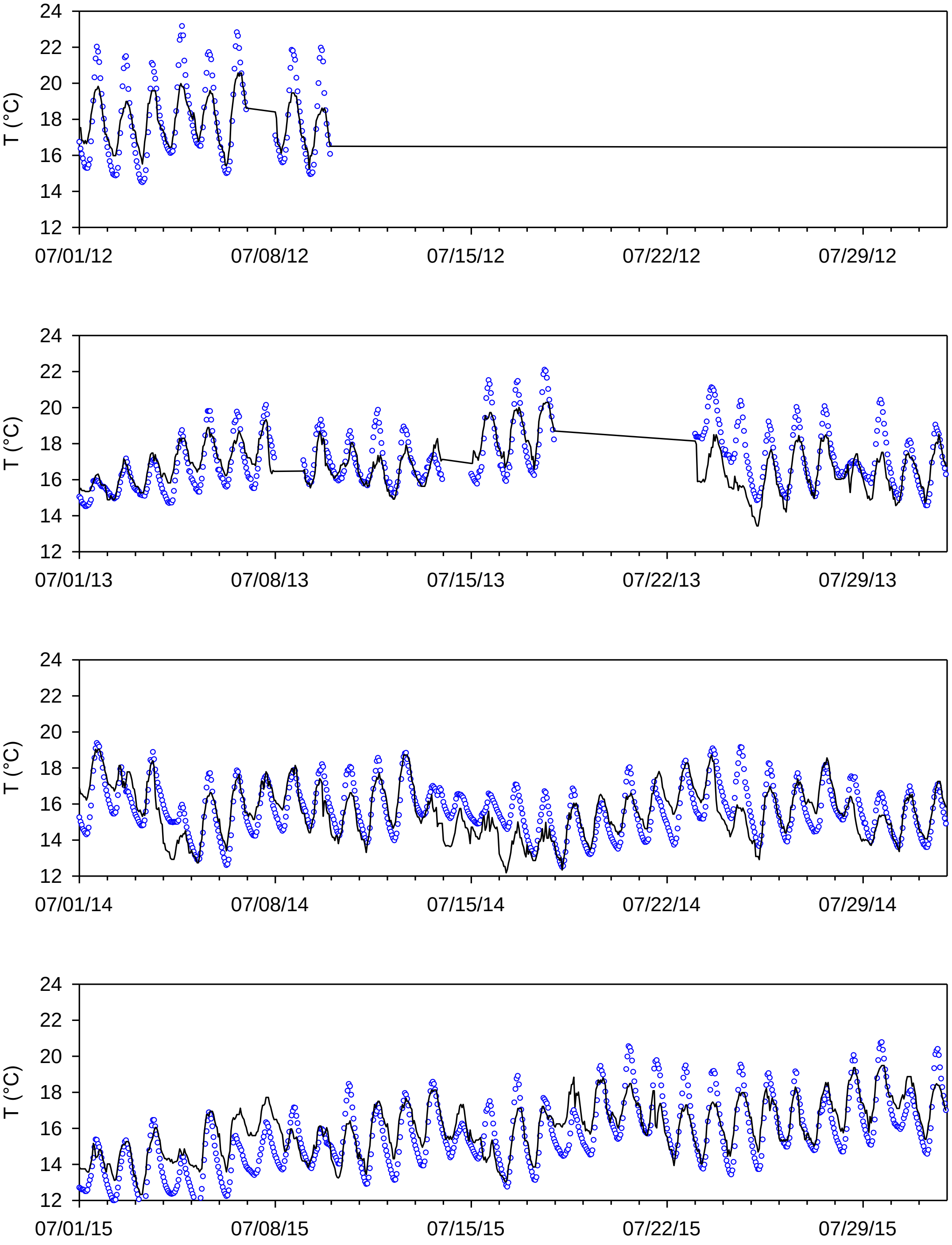


Figure 5.20. Performance of empirical temperature model for Neversink River for hourly observations in July 2012 – July 2015.

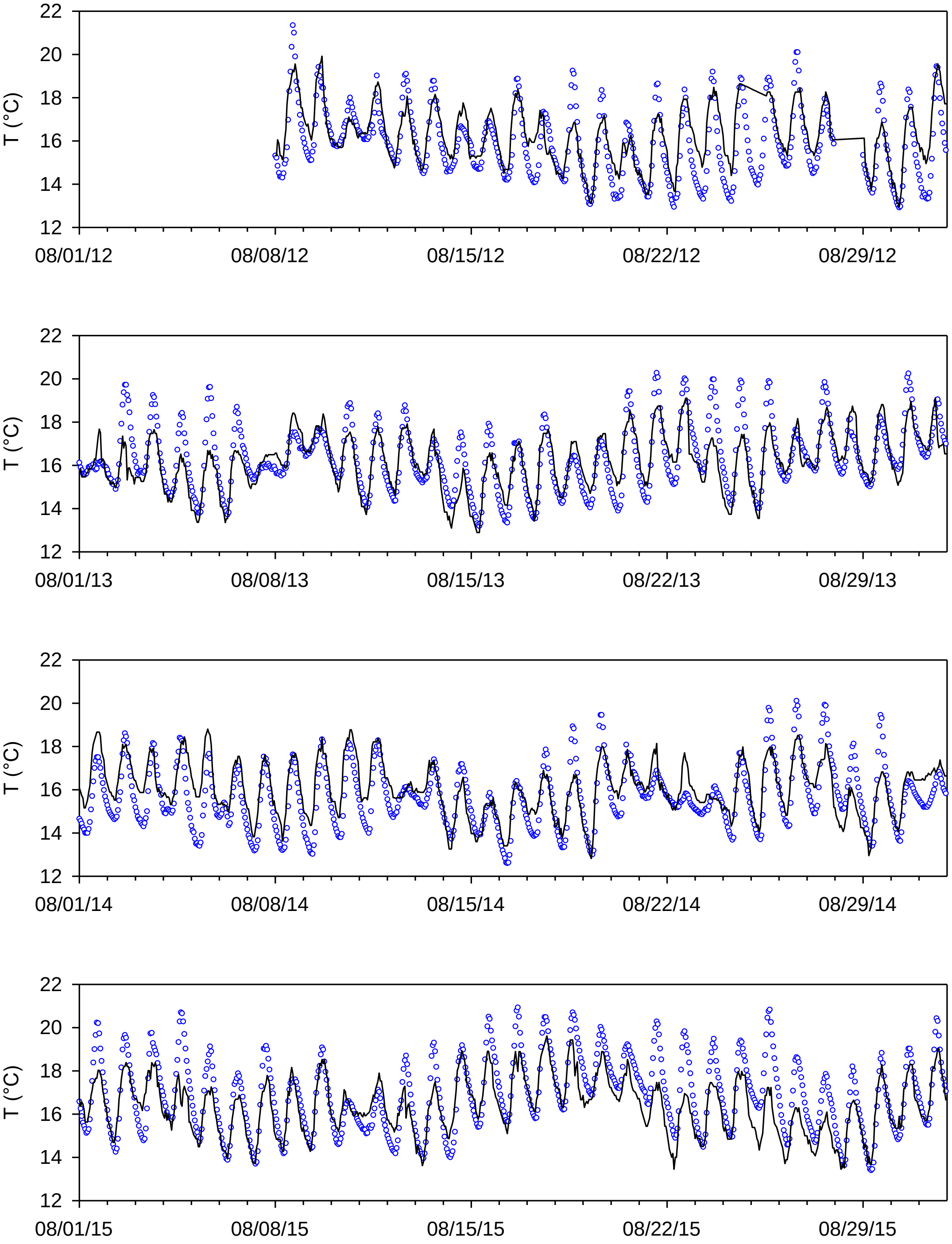


Figure 5.21. Performance of empirical temperature model for Neversink River for hourly observations in August 2012 – August 2015.

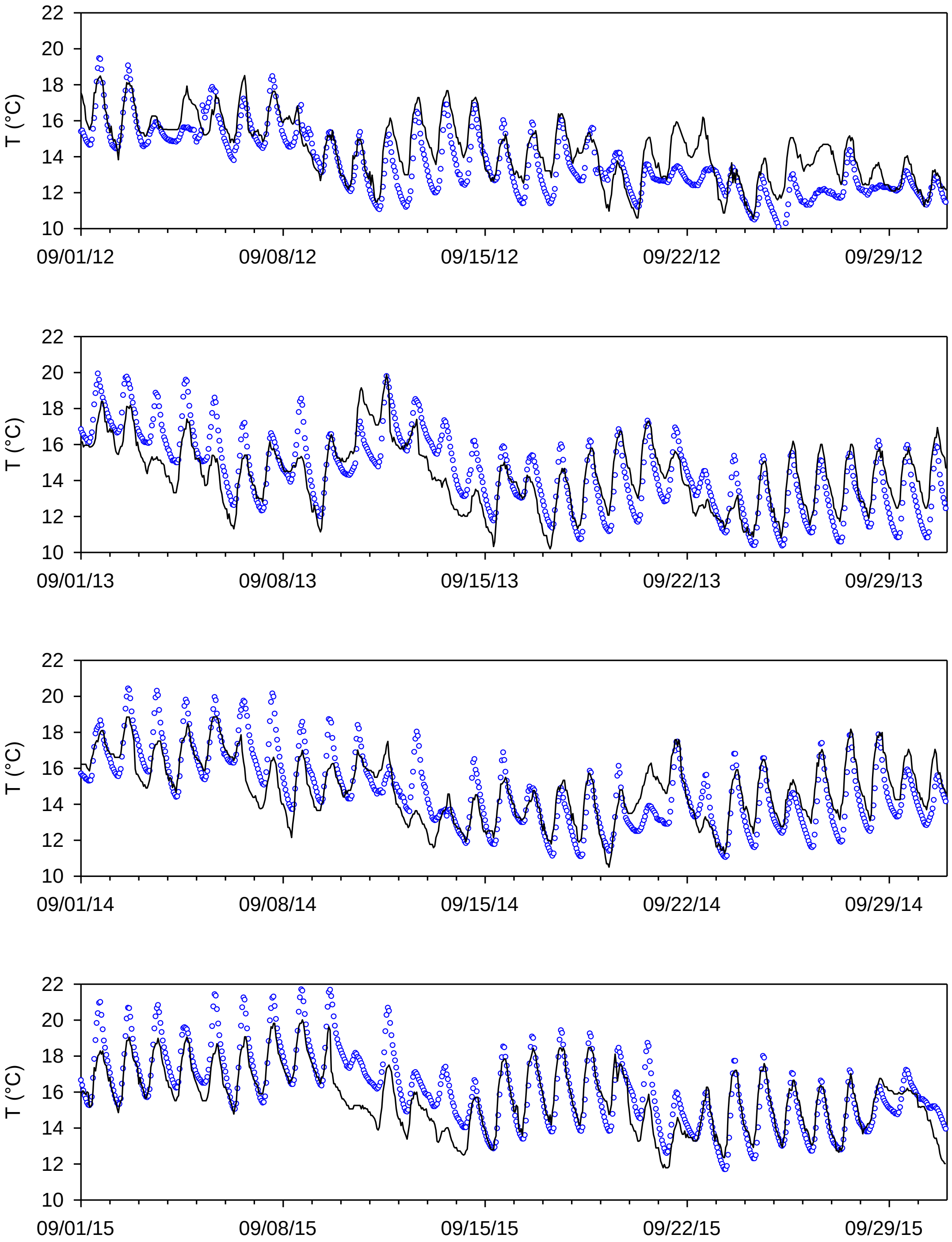


Figure 5.22. Performance of empirical temperature model for Neversink River for hourly observations in September 2012 – September 2015.

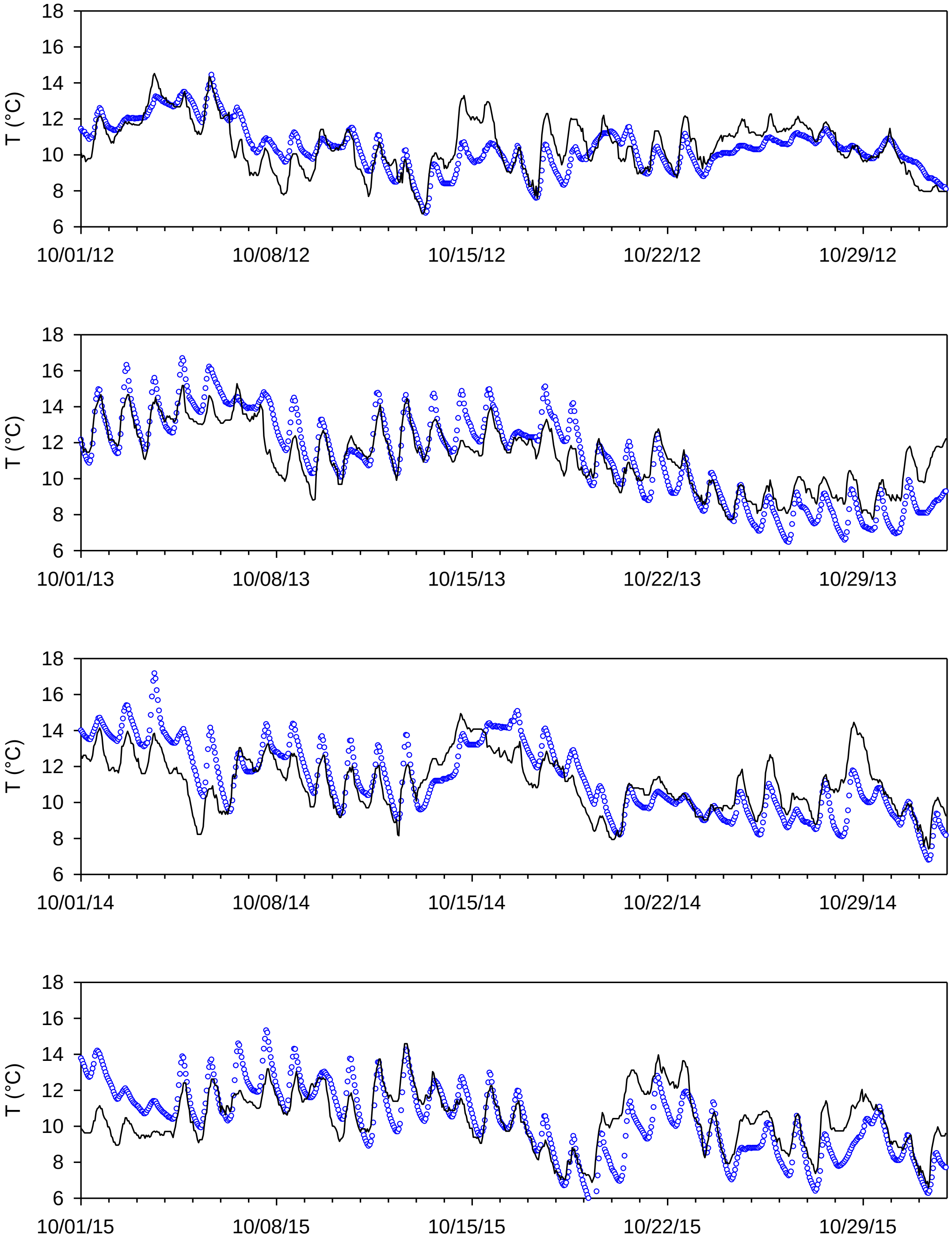


Figure 5.23. Performance of empirical temperature model for Neversink River for hourly observations in October 2012 – October 2015.

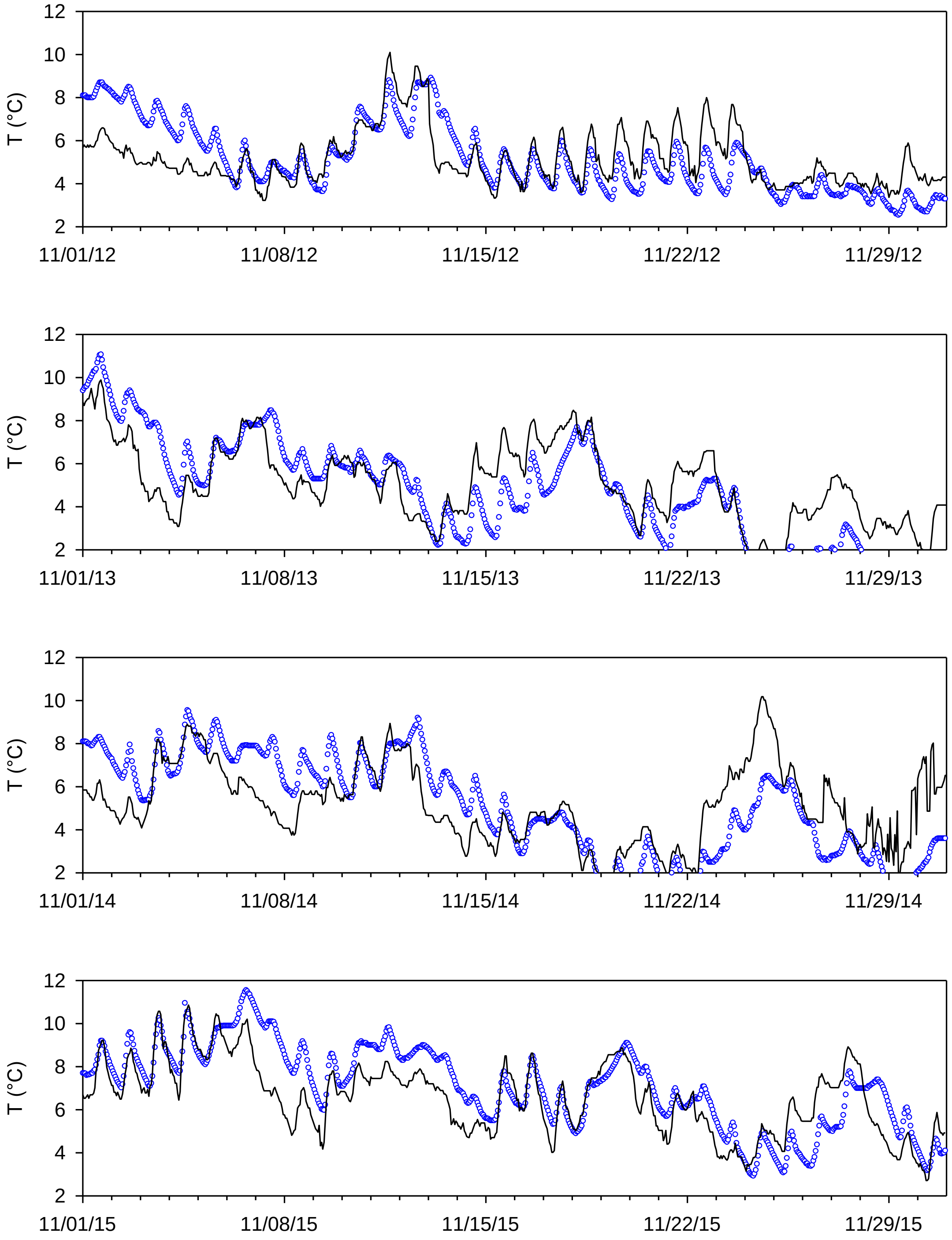


Figure 5.24. Performance of empirical temperature model for Neversink River for hourly observations in November 2012 – November 2015.

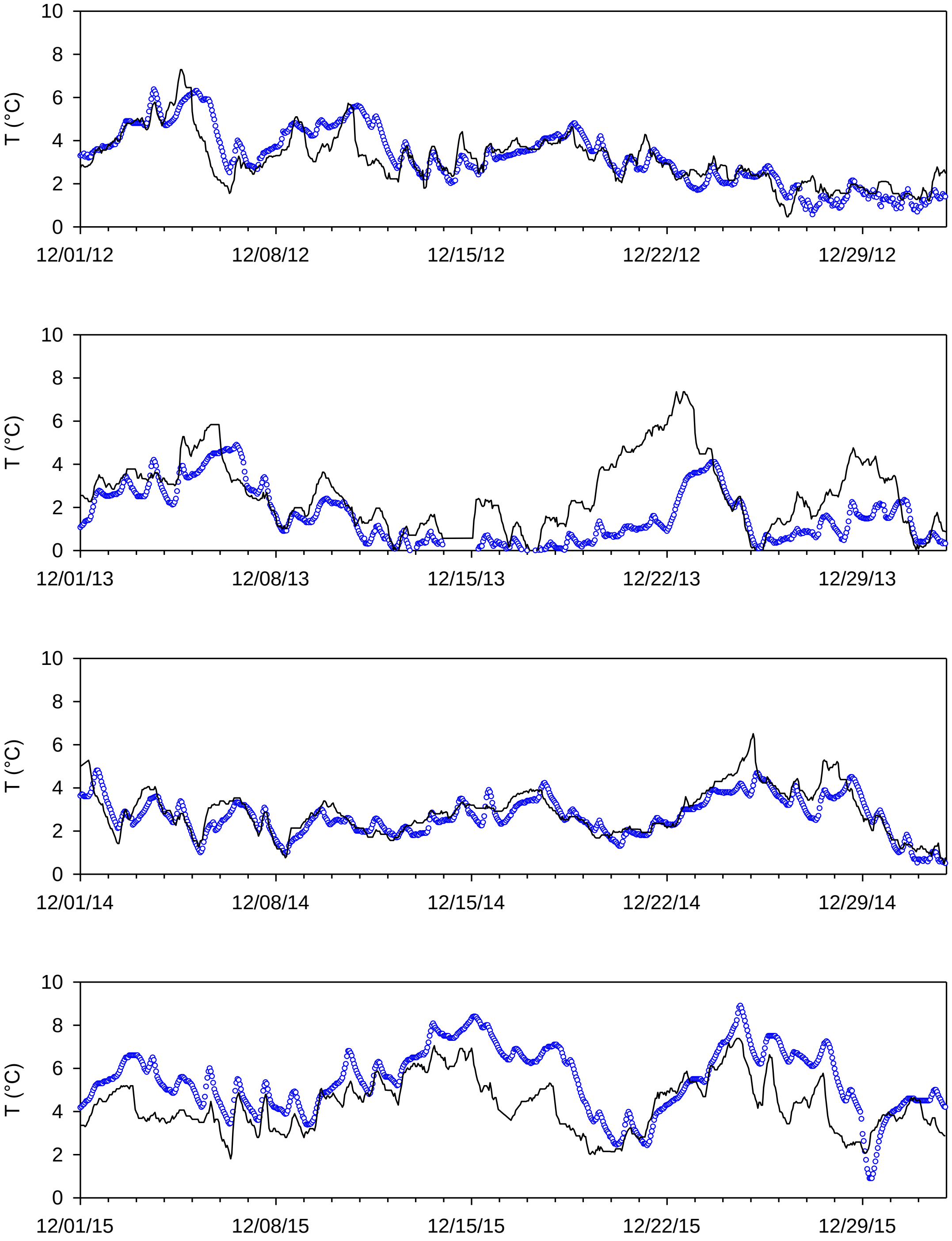


Figure 5.25. Performance of empirical temperature model for Neversink River for hourly observations in December 2012 – December 2015.

5.4. Development of an empirical stream turbidity model for Neversink River

To develop loading of turbidity for a turbidity model for Neversink Reservoir, an analysis of the flow-turbidity relationship for Neversink River was conducted. Data from 2013–2016 based on frequent (once every 15 min) sonde measurements were used. From the high-frequency flow and turbidity data, the time series of daily average streamflow (Q) and flow-weighted daily average turbidity (Tn) were computed. These were then the basic data which were analyzed as described below. A regression equation of the general form

$$\log Tn = a_2 \left(\log Q\right)^2 + a_1 \log Q + a_0 \tag{5.2}$$

was used, where the logs are base 10. A least-squares regression was used to determine the following coefficient values based on the entire dataset: a_2= 0.8133, a_1= - 0.0785, and a_0= - 0.2391, yielding a correlation coefficient r^2=0.363.

The predicted stream turbidity using Equation (5.2) is shown together with the daily average Q, Tn data and correlation coefficient in Figure 5.26. The period of data did not include sufficient runoff events to conduct additional event-specific empirical analyses. A quantile regression approach is under consideration to address the uncertainty in this simple regression (see Section 4.8 of this report).

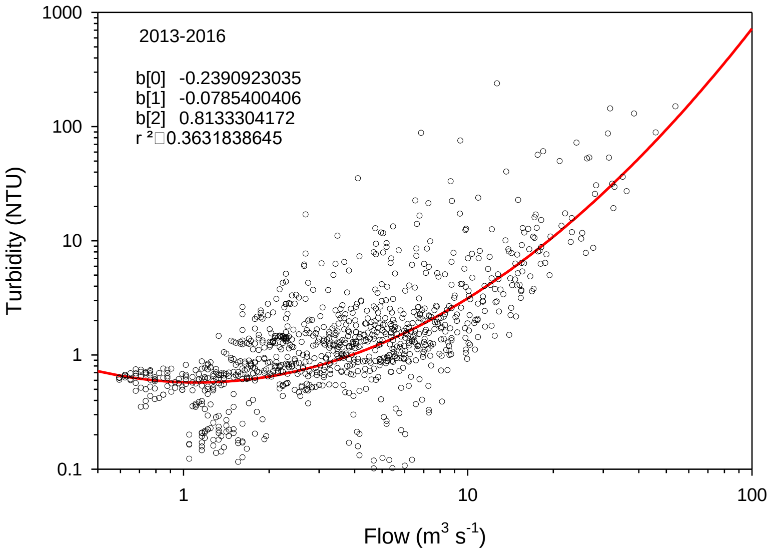


Figure 5.26. Neversink River turbidity model. Flow-weighted daily average turbidity versus daily average flow rate.

5.5. Relative contribution of turbidity loads from the Esopus Creek versus Shandaken Tunnel

The Esopus Creek is the primary tributary of Ashokan Reservoir that carries suspended material from its watershed and from the Schoharie Reservoir discharge (Shandaken Tunnel). Elevated turbidity levels are known to occur in the Esopus Creek during runoff events, even during periods when the tunnel is not discharging. The turbidity of the water discharged from Schoharie Reservoir can vary seasonally and year-to-year, depending upon the operation of the reservoir and runoff events in its own watershed. Thus, the turbidity of the water entering Ashokan Reservoir depends upon the flow rate and turbidity of both the tunnel and the Esopus Creek (e.g., stream banks) sources. Here the relative importance of each of these sources of turbidity is quantified.

Site SRR2 is located at the end of Shandaken Tunnel just before it meets the Esopus Creek. Site E16i is proximate to the mouth of the Esopus Creek at Coldbrook where it enters Ashokan Reservoir. The flow and turbidity monitored at E16i reflect contributions from both the tunnel and Esopus watershed. Turbidity, while an optical property, behaves as an intensive property, just like concentration of any other water quality constituent due to the additive character of its sources and components. This behavior allows us to perform mass balance type calculations for turbidity and to estimate turbidity (quasi-) loads (Davies-Colley *et al*. 1993).

Estimates of daily values of turbidity at SRR2 were obtained by linearly interpolating observations made 5 days/week by DEP. To obtain daily values of turbidity at E16i, an empirical model (called best-fit deterministic) developed earlier (Gannett Fleming & Hazen and Sawyer, 2008, Catskill turbidity control studies: Phase III Implementation Plan) was used. Then, daily turbidity loading was computed by simply multiplying turbidity by the flow rate and then summing up for the whole year to yield annual estimates. As shown in Figure 5.27, the contributions from the Shandaken Tunnel have remained very low for the entire period of the analysis.

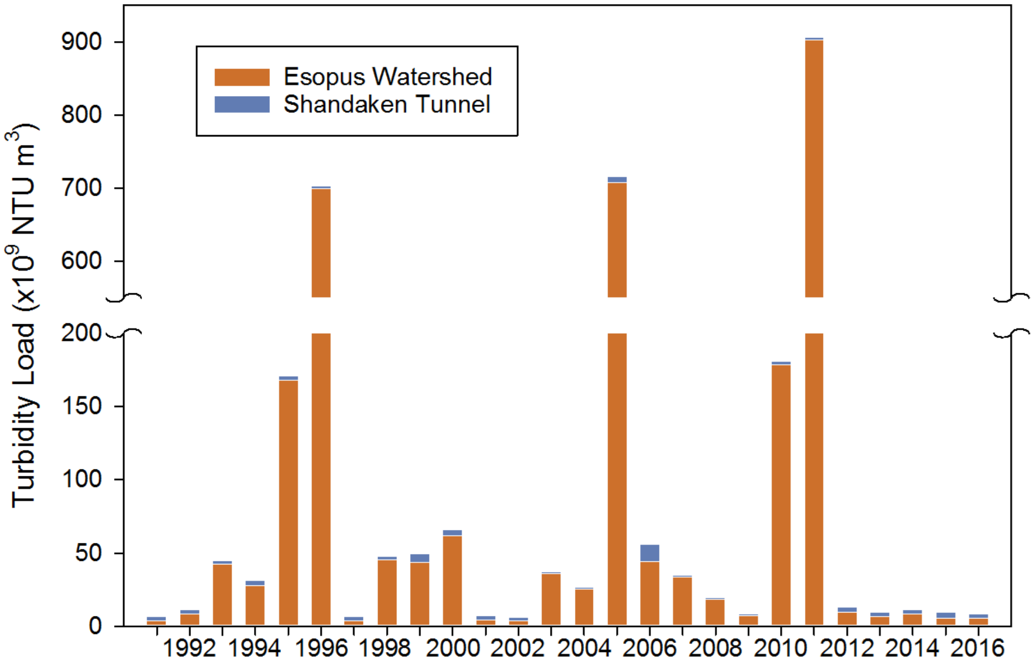


Figure 5.27. Esopus Creek watershed turbidity loading from BFD regression curves (UFI July 2008 report), Shandaken turbidity loading from observations (interpolated, DEP, unpublished). 2016 data through 12/12/2016.

6. Model Data Analysis and Acquisition

6.1. GIS Data Development for Modeling

6.1.1. WQ Monitoring Sites

In 2016, the GIS data library was updated with additions to water quality monitoring sites and biomonitoring sites. These locations were georeferenced from site descriptions provided by data samplers as part of the LIMS database entry. DEP snow monitoring and snow pillow locations data were updated with revised information provided by the Operations Directorate.

6.1.2. Support for modeling projects

GIS support was provided to members of the DEP Water Quality Modeling Section, the Division of Water Quality Science and Research, as well as CUNY Postdoctoral fellows. This support included the production and revision of maps for inclusion in annual reports, conference presentations and posters.

Table 6.1 lists the types of GIS data used by the Modeling Section. SWAT and RHESSys are watershed models used by DEP to assess the effects of terrestrial processes on inflows to WOH reservoirs. The CE-QUAL-W2 (W2) reservoir water quality model requires many input datasets, including bathymetry. The W2 model uses as inputs bathymetry data divided into segments across the reservoir. The GIS data delivered for the WOH reservoir bathymetry survey project were resampled from 1-meter resolution source rasters to 2.5, 10m 100 and 250 meter versions of the dataset. These rasters have been converted to W2 input files, and will be tested to assess the sensitivity of the model to the various resolutions. The results of this sensitivity analysis will determine the appropriate version of the bathymetry which will be used to update the W2 model.

Table 6.1. Inventory of GIS data used in modeling.

Data Type	Data Source	Data Description	Dates	Modeling Needs
LiDAR	DEP	Tree height, canopy fraction and LAI	2009	RHESSys Model Input
DEM (1 meter)	DEP	Watershed delineation, topographic parameters	2009	Watershed delineation
Land Use/ Land Cover	DEP	Spatial extent of types of land use and land cover	2009	SWAT Model Input
SSURGO2 Soil Survey	USDA	Soil Characteristics, Percent Clay, Ksat, Ksattop, Available Water Content, Depth to Water Table,	2012	SWAT Model Input
Aerial Imagery	USGS	Landsat TM multispectral imagery	1990-current	Leaf area index (LAI) values and tree phenology
	NASA	MODIS multispectral imagery	2002-current	LAI, and tree phenology
Forest Stands	DEP, USFS	Forest spatial extent	2009	RHESSys Model Input
Vegetation	Cary Institute of Ecosystem Studies	Spatial extent of 11 vegetation species	1986-1993	RHESSys Model Input
Reservoir Bathymetry	DEP	1 meter resolution bathymetric surfaces	2015	CE-QUAL-W2 Model Input

6.2. Ongoing Modeling/GIS Projects

6.2.1. West of Hudson Reservoir Bathymetry Surveys

Through an Intergovernmental Agreement signed with the USGS in 2013, the six West of Hudson reservoirs were surveyed from 2013-15 to generate updated bathymetric surfaces. The USGS employed a single-beam echosounder to survey transects across each reservoir. The transects were designed to measure depths below the survey boat with a maximum spacing between transects of less than 1% of the reservoir length. Higher density transects were completed near the intakes and spillways for each of the reservoirs to increase the accuracy of the surveys in those areas.

To accurately convert the measured depths to elevation above sea level, all reservoir spillways were surveyed in the NAVD88 vertical datum. Offsets between the original spillway elevation values and the survey elevations were calculated to enable comparison between as-built volume estimates and the updated survey volumes. The resulting echosounder elevation data were combined with secondary datasets and converted to a triangular irregular network (TIN), producing a bathymetric surface model. The raw TIN surfaces were then edited to enforce linear features such as channels and slopes. The final edited TINs were used to generate raster versions of the surfaces for inclusion in DEP GIS databases, 2-foot increment contours, represented in depth below spillway and elevation, and elevation-volume capacity tables calculated at 0.01 foot increments.

The final bathymetry data were delivered to DEP in August 2016, and draft report was delivered to DEP in December, 2016. DEP is in the process of reviewing the recent bathymetry data for incorporation into our existing models.

6.2.2. East of Hudson Reservoir Bathymetry Surveys

Following the surveys done for the West of Hudson reservoirs, DEP and the USGS are working to complete a new Intergovernmental Agreement to survey the 13 reservoirs and 3 controlled lakes in the East of Hudson System. Bathymetry data for the EOH reservoirs are incomplete, or of poor quality, having been developed from generalized atlas pages or fishing maps. The scope of work proposed for the Intergovernmental Agreement will produce results comparable to those of the WOH surveys. The project is scheduled to run from 2017-2020, and will produce surface models, 2-foot elevation contours and elevation-area-capacity tables.

6.2.3. Modeling Database

A modeling database is currently under development to store, format and document the datasets used by the Modeling Section. The database is being constructed using SQL Server database management software, along with a series of scripts written in python. The scripts automate the process of importing data into the database from source locations, and reformat the data into input files for use in DEP's models.

In 2016, the initial database was created on a DEP server. The database currently includes tables that define data sources external to DEP which are required by the Modeling Section. While internal DEP datasets are not intended to be duplicated, external datasets are copied locally to facilitate use of the data. Scripts have been written to collect data from the USGS and NOAA. The USGS-maintained stream gages collect hourly data on stream flow volume, water surface elevation, and other variables. The data are available in as a JSON formatted web service, which is imported into the database, and checked against previously downloaded data for any changes in values or data status, such as provisional data being finalized or corrected. NOAA publishes hourly meteorology data through the Integrated Surface Database (ISD). The ISD include data from locations such as airports and other meteorological stations. NOAA publishes these data to an FTP site as text files, which are added to the database with python scripts.

6.3. Time Series Data Development

6.3.1. Model Input Datasets

An inventory of the necessary raw time series data for watershed and reservoir model input and calibration is presented in Table 6.2 and Table 6.3, respectively. The time series data includes meteorology, streamflow, water quality, and point source loads for watershed models. For reservoir models the data includes meteorology, streamflow, stream, reservoir and key point water quality and reservoir operations. Data sets are updated as new data become available. Lag times between the current date and the dataset end dates are the result of QA/QC processes at the data source and/or procurement timelines driving the acquisition of any purchased data.

Table 6.2. Inventory of time-series data used for watershed modeling.

Data Type	**Data Source**	**Data Description**	**Dates***	**Modeling Needs**
Meteorology	Northeast Regional Climate Center	Daily Precipitation and Max/Min Temperature	Pre 1960-2013	Model Input
	PRISM	Climate Data	1981-present	Model Input
Wastewater Treatment Plants	DEP	Monthly WWTP Nutrient Loads	1990-2009	Model Input
Streamflow	USGS	Daily and Instantaneous Streamflow	Period of record available online via USGS	Hydrology Module Calibration / Nutrient and Sediment Loads
Water Quality	DEP	Routine and Storm Stream Monitoring	Period of record avail. via LIMS	Nutrient and Sediment Loads for Water Quality Calibration
	NYSDEC**	Stream Monitoring at West Branch Delaware River	1992-2010 w/ recent years avail. via LIMS	Nutrient and Sediment Loads for Water Quality Calibration

*Dates represent total span for all data sets combined. Individual station records vary.

**Now part of the DEP Water Quality dataset.

Table 6.3. Inventory of time-series data used for reservoir modeling.

Data Type	Data Source	Data Description	Dates*	Modeling Needs
Meteorology	DEP	Air Temperature, Relative Humidity, Solar Radiation, PAR, Wind Speed, Wind Direction, and Precipitation	1994-June, 2010 Period of record avail. Operations	Model Input
Keypoint and Reservoir Operations	DEP	Tunnel Water Quality, Flow and Temperature; Reservoir Storage, Spill, Withdrawal, and Elevation	Period of record avail. via LIMS	Model Input
Streamflow	USGS	Daily and Instantaneous Streamflow	Period of record available online via USGS	Model Input
Stream Hydrology	DEP	Stream Water Quality, Flow and Temperature	Period of record avail. via LIMS	Model Input
Limnology	DEP	Reservoir Water Quality, and Temperature Profiles	Period of record avail. via LIMS	Model Input

*Dates represent total span for all data sets combined. Individual station records vary.

6.3.2. General Circulation Models Time Series

The CMIP5 (Coupled Model Intercomparison Project Phase 5) program provides a framework to generate modeled time series climate data hindcasts and long term forecasts under a variety of carbon emissions scenarios. Through the CMIP5 project, 20 General Circulation Models (GCMs) have been developed to simulate the global effects of greenhouse gas emissions on climatic variables, including temperature and precipitation. These GCMs can be adapted for use as the meteorological drivers in DEP’s watershed and reservoir models. The GCMs were produced with daily values for minimum and maximum temperature and precipitation at a global scale with resolutions ranging from ½ degree to 5 degrees.

To standardize the datasets for future downscaling analyses, python scripts were developed to resample the GCM outputs to a standard 1.5-degree grid cell. Each time series was resampled using the natural neighbor interpolation method (Sibson, 1981) in ArcGIS and sampled at the centroid of the WOH watershed. The natural neighbor algorithm calculates a weighted average for each value based on the nearby input point values. Table 6.4 lists the CMIP5 climate models that were resampled. Note that not all climate models included all variables or future emissions scenarios, but all model outputs were resampled for future evaluation of their utility to the Modeling Section.

Table 6.4 CMIP5 GCM climate models used in resampling analysis. Variables used in this analysis included precipitation (Pr), minimum daily temperature (Tmin) and maximum daily temperature (Tmax). The representative concentration pathways (RCP's) available and time period of each model are listed.

Climate Model	Variable			RCP Scenarios				Time Period
	Pr	Tmin	Tmax	2.6	4.5	6	8.5	
BCC-CSM 1.1	●	●	●	●	●	●	●	1850-2301
BCC-CSM1.1(m)	●			●	●	●	●	1850-2101
CanCM4	●				●			1961-2036
CanESM2	●			●	●		●	1850-2301
CESM1(BGC)	●				●		●	1850-2101
CESM1(CAM5)	●			●	●	●	●	1850-2101
CMCC-CESM	●	●	●				●	1850-2101
CMCC-CM	●	●	●			●	●	1850-2101
CMCC-CMS	●	●	●			●	●	1850-2101
CNRM-CM5	●			●	●		●	1850-2101
GFDL-CM3	●			●	●	●	●	1860-2101
GFDL-ESM2G	●			●	●	●	●	1861-2101
GFDL-ESM2M	●			●	●	●	●	1861-2101
GISS-E2-H	●	●	●				●	1850-2101
GISS-E2-R	●				●		●	1850-2101
HadCM3	●				●			1934-2035
HadGEM2-CC	●	●	●		●		●	1859-2101
INM-CM4	●				●		●	1850-2101
IPSL-CM5A-LR	●			●	●	●	●	1850-2101
IPSL-CM5A-MR	●			●	●	●	●	1850-2101
IPSL-CM5B-LR	●				●		●	1850-2101
MIROC-ESM-CHEM	●			●	●	●	●	1850-2101
MIROC-ESM	●			●	●	●	●	1850-2101
MIROC4h	●				●			1950-2036
MIROC5	●			●	●	●	●	1850-2101
MPI-ESM-MR	●			●	●		●	1850-2101
MRI-CGCM3	●			●	●	●	●	1850-2101
MRI-ESM1	●						●	1851-2101
NorESM1-M	●			●	●	●	●	1850-2101

7. Modeling Program Collaboration

7.1. Water Utility Climate Alliance (WUCA)

New York City DEP is one of ten large water utilities in the United States that form the Water Utility Climate Alliance. This group was formed in order to identify, understand, assess the impact of climate change, and to plan and implement programs to meet climate challenges. WUCA members are involved in enhancing climate change research and improving water management decision-making to ensure that water utilities will be positioned to respond to climate change and protect our water supplies. Two white papers recently released by the Water Utility Climate Alliance feature case studies of water utilities, including DEP, addressing the threat of climate changes. These white papers advance understanding of how the relatively new enterprise of climate change assessment and adaptation is developing.

DEP is also one of four WUCA utilities (New York, Tampa Bay, Seattle, and Portland) participating in the Piloting Utility Modeling Applications (PUMA) project. In this program, the four PUMA utilities have formed partnerships with scientific institutions to explore how to integrate climate considerations into their specific programs for water quality and quantity management. PUMA has convened workshops where water utility representatives and researchers meet to discuss and compare approaches for addressing the impact of climate change on water utilities. The four utilities pursued customized approaches based on specific utility needs and learned important lessons in conducting assessments that may be of interest to the wider adaptation community. In addition, these projects attempted to create a “climate services” environment in which utility managers worked collaboratively and iteratively with climate scientists to understand both utility concerns and the ability or limitations of today’s climate science to respond to those concerns.

7.2. Global Lake Ecological Observatory Network (GLEON)

GLEON is a 10-year old organization that has been built around issues associated with the setup and deployment of robotic buoys for observing physical and water quality conditions in lakes and reservoirs, storage, processing, and analysis of the high-frequency data gathered by such buoys, and use of the data in modeling. DEP staff have attended recent annual GLEON meetings in Quebec and South Korea. DEP staff are also collaborating with other GLEON members in the intervals between meetings, including sharing of selected data. This participation has helped to ensure that DEP is getting the most out of its sizable investment in robotic monitoring in the reservoirs and tributary streams. DEP has made use of GLEON software tools in the analysis of robotic buoy data.

DEP is also applying the reservoir hydrothermal model GLM (General Lake Model) and associated Aquatic Ecodynamics (AED) water quality model. These models are “open source” software, and are thus open to use and revision by other researchers and professionals. These models are currently being applied to Cannonsville and Neversink Reservoirs by one of the CUNY postdoctoral researchers working in DEP's Water Quality Modeling Section.

Several collaborations have developed from DEP's participation in GLEON annual meetings, where scientists meet yearly to develop ideas and tools to analyze data from an array of lake and reservoir sensors deployed around the globe to address local issues for individual aquatic ecosystems. Additionally, this network of collaborators works to document changes in lake and reservoir ecosystems that occur in response to different environmental conditions and stressors. This is done in part by sharing and interpreting high-frequency sensor data and other water quality and environmental data. DEP contributed data to two collaborative GLEON research projects in 2016.

7.3. NYCDEP – City University of New York (CUNY) Modeling Program

Through the reporting period, DEP has maintained a contract with the Research Foundation of the City University of New York (RF-CUNY) that provides support for model and data development by providing postdoctoral scientists who work with DEP water quality modeling staff, and supporting research advisors. The results described in 3.2, 4.1,4.2, 4.3, and 4.4 of this report has been completed largely by CUNY postdoctoral research staff, working together with their research advisors and staff of the Water Quality Modeling Section.

In August, 2014, a new 4-year contract was initiated between DEP and RF-CUNY. Under this contract, RF-CUNY has hired four fulltime postdoctoral researchers who work in DEP's Water Quality Modeling office in Kingston, NY. Each of the researchers has an associated research advisor who receive part-time support under this contract. The research that has been initiated by these researchers continues to be a significant and critical component of DEP's modeling work. The postdoctoral program provides support in the form of providing model development and application expertise, modeling software, and data sets and in three project areas: (1) Evaluation of the effects of climate change on watershed processes and reservoir water quality as a part of CCIMP; (2) evaluation of FAD programs and land use changes on watershed processes and stream and reservoir water quality; and (3) development of the modeling capability to simulate watershed loading of dissolved organic carbon (DOC), and reservoir and water supply concentrations of DOC and disinfection byproduct formation potential (DBPFP).

The post-doctoral position in reservoir modeling of DOC and DBPFP was open from July 2016 to the end of the year. A candidate was selected to fill the position in August and has started work in January 2017. The climate science post-doctoral position was open from September 2016 to the end of the year. A candidate was selected in November and started work in February 2017. As a result, at the time of preparation of this report (March 2017), all four post-doctoral positions were filled. Planning is underway for a new 4-year CUNY-DEP contract to begin in August 2018.

8. Modeling Program Scientific Papers and Presentations

8.1. Scientific Papers

Acharya, N., A. Frei, J. Chen, L. DeCristofaro, and E. M. Owens, 2017. Evaluating Stochastic Precipitation Generators for Climate Change Impact Studies of New York City's Primary Water Supply. *Journal of Hydrometeorology* 18(3):879-896. DOI: 10.1175/JHM-D-16-0169.1.

Hoang, L, E. M. Schneiderman, R. Mukundan, K. E. B. Moore, E. M. Owens, and T. S. Steenhuis, 2017. Predicting saturation-excess runoff distribution with a lumped hillslope model: SWAT-HS. *Hydrological Processes* (accepted).

8.2. Conference Presentations

Water Utility Climate Alliance (WUCA) Meeting
May 2-3, 2016
Boulder CO

New York City's Climate Change Integrated Modeling Project
E. M. Owens, A. Cohn, NYCDEP, A. Frei, Hunter College, CUNY

Abstract: This presentation described the NYC water supply system, and the efforts of the NYC Department of Environmental Protection to forecast the effects of climate change on this system using predictive models.

American Water Resources Assoc. (AWRA) Summer Specialty Conference, GIS and Water Resources
July 11-13, 2016
Sacramento, CA

Creation of a Data Warehouse to Support New York City Water Quality Modeling Program
J. Gass and E. M. Owens, NYCDEP

Abstract: The NYC Department of Environmental Protection (DEP) is the agency responsible for supplying drinking water to the 8.5 million residents of New York City. Comprised on 19 reservoirs and 3 controlled lakes with a combined watershed area of more than 1.25 million acres, the drinking water system must provide more than 1 billion gallons of clean drinking water each day. With up to 90% of this demand provided from the 6 reservoirs that are part of the unfiltered Catskill/Delaware System, special attention must be paid to the water quality of that system.

Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI) Biennial Symposium
July 24 - 27, 2016
National Conservation Training Center, Shepherdstown, WV

The relationship of dissolved organic carbon to catchment characteristics in the Neversink River Basin, New York

K. Son and K. E. B. Moore, NYCDEP

Abstract: Recent studies of long-term trends have shown an increase in dissolved organic carbon (DOC) concentrations in surface waters (streams and lakes) across northern Europe and North America. This increase in DOC may be problematic for drinking water supplies due to the potential for a corresponding increase in disinfection by products (DBPs). In this study, a variety of alternative hypotheses are proposed and tested to explore the mechanisms for observed DOC increases in order to understand sources of variability and change in this important water quality parameter. This testing was based on DOC data from 11 USGS stream gauging stations within the Neversink River Basin, located in the Catskill Mountains, New York and included an analysis of the changes in DOC concentration at the event, seasonal, and annual scales. We examined the relationships between DOC concentrations observed at these 11 sites and catchment characteristics. The catchment characteristics include topography (slope, elevation, and aspect), soil properties (soil hydraulic features and soil carbon content), vegetation (density and dominant species), and climate influences (precipitation, temperature, and atmospheric deposition). This study will increase our understanding of several factors influencing DOC concentrations in stream water in this locale. This is a first step in developing a model to accurately estimate DOC concentrations in an area that is an important drinking water source, beginning with headwater streams and scaling up to the whole watershed in an effort to inform future watershed management strategies.

New York City Watershed Science and Technical Conference
September 16, 2016
Saugerties, New York

Mass Balance Modeling and Laboratory Experiments to Evaluate Production/Loss of Trihalomethane Precursors

E. M. Owens, NYCDEP

Abstract: EPA's Stage 2 Disinfectants/Disinfection Byproducts rule places restrictions on trihalomethanes (THMs) and haloacetic acids (HAAs) in drinking water. In order to comply with these federal regulations, and to more generally understand the behavior of the precursors to these disinfection byproducts (DBPs), NYCDEP is developing models that are capable of predicting the sources, transport, and fate of DBP precursors in the watersheds and reservoirs that make up the water supply. A monitoring program for THM precursors and dissolved organic carbon (DOC) was conducted on six West of Hudson reservoirs in 1997-98. This program involved weekly sampling and analysis over the full reservoir depth at a single site, at the mouth of the largest stream inflow to each reservoir, and in aqueducts discharging to downstream reservoirs. This monitoring data is presented and analyzed, leading to estimation of external precursor (allochthonous) loading and the net internal (autochthonous) production within each reservoir over the spring to fall period for an individual year. A series of laboratory experiments were conducted to directly measure the rate of various processes that produce and deplete organic carbon and THM precursors in the reservoir water column, including carbon fixation and excretion in laboratory cultures, THM precursor production, biodegradation, and photolysis. The results of these experiments are summarized. A predictive model for the internal production of THM precursors is presented and evaluated using the field and laboratory data.

Development of High-Resolution Bathymetry for the 6 West of Hudson Reservoirs: Implications for Water Quality Management

J. Gass, NYCDEP

Abstract: The New York City Department of Environmental Protection (DEP) utilizes simulation models to inform management decisions for reservoir operations and water quality management. These models require a variety of accurate input data, including reservoir bathymetry, to generate reliable predictions. Comprehensive bathymetric surveys of the 6 West of Hudson (WOH) reservoirs were last completed in 1997. Since that time, processes such as sedimentation, erosion and watershed management may have changed the overall morphology of the reservoirs. These processes necessitate updated bathymetry to support current modeling activities, produced at a higher resolution available with current technology. In 2013, DEP contracted with the United States Geologic Service to survey the 6 WOH reservoirs to generate a high resolution bathymetry dataset for use in water quality and operations modeling. This project will improve DEP's knowledge of the reservoirs by using more accurate methods and technology than were available during the previous surveys. This presentation will focus on the newly collected reservoir data, including comparisons with the dataset currently in use. The bathymetric data will be reviewed for any spatial patterns that may help to explain the change, and thereby inform future management scenarios. It will describe how the storage capacity estimates for each reservoir have changed, and discuss some of the possible reasons for these changes.

Development and Testing of a Probabilistic Turbidity Model for Rondout Reservoir.

R. K. Gelda, NYCDEP

Abstract: NYCDEP now routinely uses the Operations Support Tool (OST) software to help guide reservoir operating decisions involving both the quantity and quality of water. An important water quality parameter of concern for the City’s water supply is turbidity. Here we develop a tool to forecast more realistic simulations of turbidity in Rondout Reservoir withdrawal which can then be specified as input to Kensico Reservoir turbidity model in OST thus making the forecasts of turbidity in the Kensico Reservoir withdrawal more accurate. The proposed modeling tool is based on a separate turbidity model for Rondout Reservoir, which was developed and validated earlier. With the added capability of using short-term ensemble forecasts of hydrological inputs, turbidity, and climatology as the model drivers, it can generate probabilistic forecasts of turbidity. The tool is expected to help guide operations of Rondout Reservoir during storm events in the watershed as well as provide realistic estimates of turbidity inputs for the Kensico Reservoir model in OST.

Realistically predicting saturation excess runoff with SWAT-Hillslope.

L. Hoang, E. M. Schneiderman, T. S. Steenhuis, K. E. Moore, E. M. Owens

Abstract: Saturation excess runoff is without doubt the major runoff mechanism in the humid well vegetated areas where infiltration rates often exceed the medium rainfall intensity. Despite its preponderance, incorporating this runoff in distributed models has been slow and fraught with difficulties. Soil and Water Assessment Tool (SWAT) uses information of soil plant characteristics and hydrologic condition to predict runoff and thus is implicitly based on infiltration-excess runoff. Previous attempts to incorporate saturation excess runoff mechanism in SWAT fell short because the inability to distribute water from a Hydrological Response Unit (HRU) to another. This paper introduces a modified version of SWAT, referred as SWAT-Hillslope. It improves the simulation of saturation excess runoff by redefining HRUs to include landscape position, grouping the newly defined HRUs into wetness classes and by introducing a perched water table with the ability to route interflow from “dryer” to “wetter” HRUs. Mathematically, the perched aquifer is a non-linear reservoir that generates rapid subsurface stormflow as the perched water table rises. The SWAT-Hillslope model was tested in the Town Brook watershed in the upper reaches of the West Branch of the Delaware in the Catskill Mountains. The results show that SWAT-Hillslope could predict discharge well with Nash-Sutcliffe Efficiency of 0.68 and 0.87 for daily and monthly time steps. Compared to the original SWAT model, SWAT-Hillslope predicted less surface runoff and groundwater flow and more lateral flow. The saturated areas in SWAT-Hillslope were concentrated in locations with high topographic index and was in agreement with field observations. With the incorporation of topography characteristics and the addition of the perched aquifer, SWAT-Hillslope gives a realistic representation of hydrological

processes and will lead to better water quality models where the source of the surface runoff matters.

The predictability of forest recovery using a mechanistic hydro-ecological model.

K. Son, NYCDEP

Abstract: Forests plays a key role in regulating water quality in streams, and forest disturbances and harvesting can affect the water quality through major changes in hydrologic function and ecosystem. However, the prediction of forest recovery following forest disturbance through modeling is still difficult given the complex interactions among climate, hydrology and biogeochemistry. As well, the forest recovery data over large areas and long time periods are limited. This study tested the predictability of a mechanistic hydro-ecological model, Regional Hydro-Ecologic Simulation System (RHESSys) for capturing forest recovery after forest harvesting. The study site is Frost Valley YMCA Model Forest, located in the Neversink River Basin that is also a part of New York City Water supply watersheds. Three areas within the site experienced different forest harvesting practices: light selective cut, heavy selective cut and clear-cut. First, we used vegetation indices derived from Landsat imagery in order to develop the forest recover data for three sites. Second, we calibrated the RHESSys model with measured streamflow, nitrate and dissolved organic carbon (DOC) fluxes. Third, the predictions of leaf area index using the calibrated model was compared with the estimated forestry recovery data. This study provides a framework to improve the model predictability of forest recovery after forest harvesting.

Deciphering the Climate Signal and Other Sources of Change in Stream Water Quality, NYC Watershed Science and Technical Conference

K. E. B. Moore, NYCDEP, and K. Son, NYCDEP

Abstract: Climate projections for New York City’s water supply watersheds generally indicate that increases in temperature, precipitation, and potential evapotranspiration, with consequent reductions in snowpack and changes in the timing of runoff, are anticipated. These changes have important implications for impacts on the nitrogen cycle, with potential increases in the release of nitrate to surface waters. Murdoch *et al*. (1998) used a novel approach to examine the long-term effects of temperature on nitrate leaching from Biscuit Brook, a headwater stream in the Catskill region, and showed that mineralization and nitrification, rather than deposition or vegetation uptake, were the primary factors controlling nitrogen leaching from the watershed for the period of 1983-1995. We revisited their approach for later years at the same site and expand our inquiry to other sites in the region to investigate trends in nitrate concentration over a 20+ year period of record. We used a variety of approaches to look at trends in annual, seasonal, and event-based nitrate concentration, and the relationship between trends in nitrate concentration and different flow conditions. We also explored the patterns in dissolved

organic carbon and look at the interrelationships between nitrate and other water quality indicators. We expect that using long-term records of water quality in this study will allow us to have a better understanding of changes in stream water quality with respect to climate change.

Catskill Environmental Research & Monitoring (CERM) Conference
October 27-28, 2016
Belleayre Ski Center, Highmount, NY

Estimating carbon budgets of the Biscuit Brook watershed using RHESSys

K. Son, E.S. Schneiderman, E.M. Owens, L. Lin, and L. Band

Abstract: Carbon inputs to New York City (NYC) reservoirs are potential precursors to disinfection by-products (DBPs) which are subject to regulatory limits in the NYC drinking water distribution system. As a first step towards increasing understanding of the potential sources of watershed carbon, a forest ecosystem model (Regional Hydrologic Ecologic System Simulation, RHESSys) was applied to estimate the carbon pools and fluxes associated with soil, litter, coarse woody debris and live trees in selected NYC watersheds under historical climate conditions. As a case study, RHESSys was applied to Biscuit Brook, a headwater stream in the watershed of Neversink Reservoir that has a long record of streamflow, dissolved organic carbon (DOC) and other water chemistry data. Using measured streamflow, RHESSys was calibrated to estimate soil parameters. To parameterize the soil carbon pools in RHESSys, we estimated forest stand ages for 30m grid cells within the NYC watersheds using LIDAR-based stand heights and an empirical age-height relationship. Vegetation type for each grid is characterized from the Landsat TM based- vegetation map for the Catskill Mountains. RHESSys was first run in spin-up mode to bring soil carbon pools in approximate equilibrium with local climate conditions. The spun-up model was then run for as many years as the oldest stand. For each model output variable of interest (carbon pools, including leaves, fine roots, coarse roots, coarse woody debris, litter, and soil), a composite map was constructed by extracting, for each grid cell, the value of the variable at the age in the model run that corresponds to the stand age previously derived from LIDAR. The resultant spatial data are summarized to provide a carbon snapshot of the watershed and compared with existing national carbon database and the SSURGO carbon database. In addition to soil carbon, simulated LAI is compared with LAI based on Landsat TM imagery. Related parameters were calibrated to match the predicted LAI with the remote sensing-based LAI. The simulated DOC values at the outlet of a watershed were compared to measured values. Following application to Biscuit Brook and other watersheds of similar size, the model application will be scaled up to the Neversink and other NYC reservoir watersheds in an effort to inform future water supply management strategies.

Integrated tools to improve the understanding of linkages among hydrology, sediment and carbon in Catskills headwater streams

M. Vian, K. Son and R. Mukundan, NYCDEP

Abstract: New York City (NYC) daily consumes one billion gallons of water, provided by an unfiltered water supply delivered largely from headwater streams in the Catskill Mountains. Numerous anthropogenic influences present challenges to NYC's efforts to maintain high raw quality for this water supply. Notable influences are hydrologic alteration due to climate change, and impacts on riparian forest cover due to introduced forest pests, including Hemlock Woody Adelgid and Emerald Ash Borer. Through a series of complex functional linkages, these changes increase sediment and carbon exports to streams that potentially diminish raw water quality. In order to maintain high raw water quality, resource managers need to understand the factors contributing to entrainment of organic carbon and sediment. To date, no single model adequately represents the complexity of the linkages among hillslope hydrology, riparian forest conditions, and channel processes. Here we present a framework for integrating an ecohydrologic model, a landslide mapping model, and several channel morphologic and hydraulic models. Integration of these tools can spatially quantify the probability of wood and sediment loading from the stream corridor and adjoining hillslopes, predict the progressive cascading downstream of related disturbances, and provide estimates of their cumulative impact on sediment and carbon fluxes. Model analyses can estimate impacts on these fluxes resulting from climate change and forest pests. These tools can then provide key information to target management treatments. Application of the suite of tools is demonstrated in a case study of Biscuit Brook, a tributary to the Neversink River.

Modeling of DOC and Disinfection By-Product Precursors in the NYC Water Supply

E. M. Owens, NYCDEP

Abstract: EPA's Stage 2 Disinfectants/Disinfection Byproducts (DBP) rule places restrictions on trihalomethanes (THMs) and haloacetic acids (HAAs) in drinking water. In order to comply with these federal regulations, and to more generally understand the behavior of the precursors to these disinfection byproducts (DBPs), NYCDEP is developing models that are capable of predicting the sources, transport, and fate of DBP precursors in the watersheds and reservoirs that make up the water supply. These models will also be used to evaluate the effect of watershed protection programs, and of climate change and extreme weather events, on the water supply. The models include two existing distributed parameter terrestrial (watershed) models. In addition, NYCDEP is investigating two alternative approaches for modeling the production and loss processes for DBP precursors in the City's water supply reservoirs. Preliminary testing and analysis of these two alternative approaches is presented here. This analysis focuses on evaluating net production and loss of precursors occurring in Cannonsville and Neversink

Reservoirs, and on the spatial (vertical) and seasonal patterns in the net production/loss in the reservoirs.

2016 North American Lake Management Society (NALMS) International Symposium
November 2–4, 2016
Banff Springs Hotel, Banff, Alberta, Canada

Futuristic Simulation of Turbidity in Rondout Reservoir, New York
R. K. Gelda, NYCDEP

Abstract: Rondout Reservoir is one of New York City's (NYC) 19 water supply reservoirs, located approximately 130 km northwest of the city. This narrow dimictic reservoir receives withdrawals from three upstream reservoirs (Cannonsville, Pepacton, and Neversink) within the NYC Delaware water supply system. Water withdrawn from Rondout Reservoir enters an aqueduct for conveyance to a further downstream reservoir where it mixes with water from other parts of the system in Kensico Reservoir before disinfection and supply to NYC (no filtration). An important water quality parameter of concern for the City's water supply is turbidity. The turbidity of the water withdrawn from Rondout Reservoir is less than 2.8 NTU 95% of the time (1987-2014), however, higher turbidities have been observed during extreme runoff events. Mathematical models of turbidity are desired for such situations to predict transport and fate of turbidity-causing particles and to guide reservoir operations. Here we adopt a previously validated two-dimensional multi-size class dynamic turbidity model for this reservoir and develop a tool to generate probabilistic forecasts of withdrawal turbidity. The tool utilizes short-term ensemble forecasts of hydrological inputs, turbidity inputs, and climatology as the model drivers, and generates probabilistic forecasts of in-reservoir and withdrawal turbidity. The tool is expected to help guide operations of Rondout Reservoir during storm events in the watershed as well as provide realistic estimates of turbidity inputs for the downstream Kensico Reservoir turbidity model.

American Geophysical Union Fall Meeting
December 12-16, 2016
San Francisco, CA

Streamflow simulation by a watershed model using stochastically generated weather in New York City watersheds
R. Mukundan, N. Acharya, R. K. Gelda, E. M. Owens, A. Frei, and E. M. Schneiderman

Abstract: Recent studies have reported increasing trends in total precipitation, and in the frequency and magnitude of extreme precipitation events in the West of Hudson (WOH) watersheds of the New York City (NYC) water supply. The potential effects of these changes may pose challenges for both water quality (such as increased sediment and nutrient loading) and quantity (such as reservoir storage and management). The NYC Dept. of Environmental Protection Climate Change Integrated Modeling Project (CCIMP) is using “bottom-up” or vulnerability based methods to explore climate impacts on water resources. Stochastic weather generators (SWGs) are an integral component of the bottom-up approach. Previous work has identified and evaluated the skill of alternative stochastic weather generators of varying complexity for simulating the statistical characteristics of observed minimum and maximum daily air temperature and occurrence and amount of precipitation. This evaluation focused on the skill in representing extreme streamflow event probabilities across NYC West of Hudson (WOH) watersheds. Synthetic weather time series from the selected (skewed normal) SWG were used to drive the Generalized Watershed Loading Function (GWLF) watershed model for a 600-year period to simulate daily streamflows for WOH watersheds under a wide range of hydrologic conditions. Long-term average daily streamflows generated using the synthetic weather time series were comparable to values generated using observed long-term (1950-2009) weather time series. This study demonstrates the ability of the selected weather generator to adequately represent the hydrologic response in WOH watersheds with respect to the total, peak, and seasonality in streamflows. Future application of SWGs in NYC watersheds will include generating multiple scenarios of changing climate to evaluate water supply system vulnerability and selection of appropriate adaptation measures.

The effect of input data complexity on the uncertainty of streamflow simulation using SWAT-HS

L. Hoang, E.M. Schneiderman, T. S. Steenhuis, K. E. B. Moore, E. M. Owens

Abstract: SWAT-Hillslope (SWAT-HS) is a modified version of the Soil and Water Assessment Tool (SWAT) that incorporates the simulation of saturation-excess runoff in SWAT and can be used in humid vegetated areas where this form of runoff is common. Similar to other distributed hydrological models, model uncertainty is an issue of concern due to its effects on prediction and subsequent decision makihng. Previous research has shown how SWAT predictions are affected by uncertainty in parameter estimation and input data resolution. Nevertheless, little information is available on how parameter uncertainty is affected by input data of varying complexity. SWAT-HS was applied to study the effects of input data with varying degrees of complexity on parameter uncertainty and output uncertainty. Four digital elevation model (DEM) resolutions (30, 10, 3 and 1 m) were tested for their ability to predict streamflow and saturated areas. In a second analysis, three soil maps and three land use maps were used to build nine SWAT-HS setups from simple (fewer soil types/ land use classes) to complex (more types/ classes), which were then compared to study the effect of input data complexity on model uncertainty. The case study was the Town Brook watershed in the upper reaches of the West Branch Delaware River in the Catskill Region, New York, USA. The results showed that grid size did not affect the simulation of streamflow but significantly affected the spatial pattern of saturated areas, with the 10m being the most appropriate grid size to use. The comparison of nine model setups revealed that input data complexity did not affect parameter uncertainty. The most simple soil/ land use specification resulted in a very low probability of good model performance, while the most complex setup did not give the highest probability of good performance either. It was found that increasing model complexity may not necessarily improve model performance or reduce parameter and output uncertainty, but using multiple temporal and spatial observations can aid in finding the most realistic parameter set and in reducing uncertainty.

A study of dissolved organic carbon and nitrate export in Catskill Mountain watersheds

K. Son, K. E. B. Moore, L. Lin, E. M. Schneiderman and L. Band

Abstract: Watersheds in the Catskill Mountain region of New York State have historically experienced soil and stream acidification due to atmospheric deposition of SO2 and NOx. Recent studies in this region, and elsewhere in North America and Europe, have shown increases in dissolved organic carbon (DOC) in streams and lakes. Catskill watersheds are the major source of drinking water for New York City and other communities in the region. Due to use of chlorine for disinfection, there is potential for the increase in DOC to lead to increased levels of disinfection byproducts in treated drinking water. Therefore, developing an improved understanding of the sources, fate and transport mechanisms, and export patterns for nitrate and DOC is important for informing watershed and water supply management. In this study, we analyzed the relationships between watershed characteristics, nitrate, and DOC for 12 gauged streams in the Neversink River watershed. Watershed characteristics included topography (elevation, slope, topographic wetness index), vegetation (leaf area index, species composition), soil (soil hydraulic parameters, soil carbon, wetland soil), atmospheric deposition (SO2, NOx), and climate (precipitation, temperature). Our preliminary analysis showed that both watershed slope and baseflow ratio are negatively correlated with annual median DOC concentration. At Biscuit Brook in the Neversink watershed, annual precipitation explained about 25% of annual DOC median concentration. DOC concentration was highly correlated with storm runoff in spring, summer, and fall, but stream nitrate concentration was weakly correlated with storm runoff in most seasons except summer when it was highly correlated with baseflow. We also applied a process-based ecohydrologic model (Regional Hydrologic Ecologic System Simulation, RHESSys) to the Biscuit Brook watershed to explore sources of nitrate and DOC and their movement within the watershed. We expect that this study will increase our understanding of how, when, and where DOC and nitrate are stored and transported to streams, as well as give insights into the key controls on nitrate and DOC processes in Catskill Mountain watersheds.

9. References

Acharya, N., A. Frei, J. Chen, L. DeCristofaro, and E.M. Owens. 2017. Evaluating Stochastic Precipitation Generators for Climate Change Impact Studies of New York City's Primary Water Supply. *Journal of Hydrometeorology*. DOI: http://dx.doi.org/10.1175/JHM-D-16-0169.1

Agnew. L.J., S. Lyon, P. Gérard-Marchant, V.B. Collins, A.J. Lembo, T.S. Steenhuis, and M.T. Walter. 2006. Identifying hydrologically sensitive areas: Bridging the gap between science and application. *Journal of Environmental Management*, 78: 63-76. DOI: http://dx.doi.org/10.1016/j.jenvman.2005.04.021.

Ailliot, P., D. Allard, V. Monbet, and P. Naveau, P. 2015. Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistique*, 156(1), 101-113.

Ameli, A. A., J.J. McDonnell, and K. Bishop 2016. The exponential decline in saturated hydraulic conductivity with depth: a novel method for exploring its effect on water flow paths and transit time distribution. *Hydrological Processes*, 30: 2438-2450. DOI: 10.1002/hyp.10777.

Anandhi, A., A. Frei, D. C. Pierson, E. M. Schneiderman, M. S. Zion, D. Lounsbury and A. H. Matonse. 2011a. Examination of change factor methodologies for climate change impact assessment. *Water Resources Research* 47, W03501 10p.

Anandhi A., A. Frei, S.M. Pradhanang, M.S. Zion, D.C. Pierson, and E.M. Schneiderman. 2011b. AR4 climate model performance in simulating snow water equivalent over Catskill Mountain watersheds, New York, USA. *Hydrological Processes* 25: 3302–3311.

Anandhi, A., M.S. Zion, P.H. Gowda, D.C. Pierson, D. Lounsbury, and A. Frei. 2013. Past and future changes in frost day indices in Catskill Mountain region of New York. *Hydrologic Processes* 27: 3094–3104.

Andréassian, V., C. Perrin, C. Michel, I. Usart-Sanchez, and J. Lavabre. 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Journal of Hydrology* 250(1), 206-223.

Apipattanavis, S., G. Podesta, B. Rajagopalan, and R.W. Katz. 2007. A semiparametric multivariate and multisite weather generator. *Water Resources Research* 43, W11401.

Arnold J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams, 1998. Large area hydrologic modeling and assessment part 1: Model development. *J. American Water Resources Assoc.* 34: 73-89.

Arnold, J.G., P.M. Allen, M. Volk, J.R. Williams, and D.D. Bosch, 2010. Assessment of different representations of spatial variability on SWAT model performance. *Transaction of ASABE*, 53: 1433-1443. DOI: 10.13031/2013.34913.

Asselman, N.E. 1999. Suspended sediment dynamics in a large drainage basin: The River Rhine. *Hydrological Processes* 13(10), 1437-1450.

Bárdossy, A., and S.K. Singh, 2008. Robust estimation of hydrological model parameters. *Hydrology and Earth System Science*, 12: 1273-1283. DOI: 10.5194/hess-12-1273-2008.

Bayabil, H.K., S.A. Tilahun, A.S. Collick, B. Yitaferu B, and T.S. Steenhuis, 2010. Are runoff processes ecologically or topographically driven in the (sub) humid Ethiopian highlands? The case of the Maybar watershed. *Ecohydrology*, 3: 457-466. DOI: 10.1002/eco.170.

Beven, K., and J. Freer. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrology* 249: 11–29.

Beven, K., 2001. Rainfall runoff modeling: The Primer. J. Wiley & Sons, Ltd, Chichester, London.

Beven K., 1986. Runoff Production and Flood Frequency in Catchments of Order n: An Alternative Approach. In: Scale Problems in Hydrology, V.K. Gupta, I. Rodríguez-Iturbe, E.F. Wood (eds.) Springer Netherlands, pp: 107-131.

Beven, K., and A. Binley, 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6: 279-298. DOI: 10.1002/hyp.3360060305.

Beven, K., and P. Young. 2013. A guide to good practice in modeling semantics for authors and referees. *Water Resources Research*, 49: 5092-5098.

Beven, K.J., and M.J. Kirkby, 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24: 43-69. DOI: 10.1080/02626667909491834.

Bingner, R.L., and F.D. Theurer, 2001. AnnAGNPS Technical Processes: Documentation Version 2. Available at www.sedlab.olemiss.edu/AGNPS.html.

Borah, D.K., and M. Bera, 2003. Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Transaction of ASAE*, 46: 1553-1566.

Bosch, D.D., J.G. Arnold, M. Volk, and P.M. Allen. 2010. Simulation of a low-gradient coastal plain watershed using the SWAT landscape model. *Transaction of ASABE*, 53: 1445-1456.

Boulanger J.P., F. Martinez F, O. Penalba, and E.C. Segura. 2007. Neural Network based daily precipitation generator (NNGEN-P). *Climate Dynamics* 28: 307–324.

Boyer E.W., G.M. Hornberger, K.E. Bencala, and D. McKnight. 1996. Overview of a simple model describing variation of dissolved organic carbon in an upland catchment. *Ecol. Modelling* 86: 183–188.

Boyer, E.W., G.M. Hornberger, K.E. Bencala, and D.M. McKnight. 1997. Response Characteristics of DOC Flushing in an Alpine Catchment. *Hydrologic Processes* 11:1635–1647.

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

Boughton, W. 2004. The Australian water balance model. *Environ. Modeling & Software*, 19: 943-956.

Brooks, E.S., J. Boll, and P.A. McDaniel. 2004. A hillslope-scale experiment to measure lateral saturated hydraulic conductivity. *Water Resources Research*, 40: DOI: 10.1029/2003WR002858.

Bryant, R., T. Veith, P. Kleinman, and W. Gburek. 2008. Cannonsville Reservoir and Town Brook watersheds: Documenting conservation efforts to protect New York City's drinking water. *Journal of Soil and Water Conservation*, 63: 339-344.

Brown, C., W. Werick, W. Leger, and D. Fay. 2011. A decision-analytic approach to managing climate risks: Application to the Upper Great Lakes. *J. American Water Resources Assoc*., 47(3): 524–534.

Buishand, T.A. 1978. Some remarks on the use of daily rainfall models. *J. of Hydrology* 36:295–308.

Buishand, T.A. and T. Brandsma. 2001. Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling, *Water Resources Research* 37(11): 2761–2776.

Burns, D.A., J. Klaus, and M.R. McHale. 2007. Recent climate trends and implications for water resources in the Catskill Mountain region, New York, USA. *J. Hydrology* 336:155–170.

Burrough, P.A. 1986. Principles of geographical information systems for land resources assessment. Oxford University Press.

Chen J, F.P. Brissette and R. Leconte. 2010. A daily stochastic weather generator for preserving low-frequency of climate variability. *J. Hydrology* 388:480–490.

Chen J. and F. Brissette and P. Zielinski. 2015. Constraining frequency distributions with the probable maximum precipitation for the stochastic generation of realistic extreme events, *Journal of Extreme Events*, DOI: 10.1142/S2345737615500098

Chen J., and F. Brissette. 2014a. Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the Loess Plateau of China. *Intl. J. Climatology*, 34: 3089–3105.

Chen J., and F. Brissette. 2014b. Stochastic generation of daily precipitation amounts: review and evaluation of different models. *Climate Research*, 59: 189-206.

Chen, J., F. Brissette, R. Leconte, and A. Caron. 2012. A versatile weather generator for daily precipitation and temperature. *Trans. Amer. Soc. Agric. and Biological Engineers*. 55(3), 895-906.

Cole, T.M., and S.A. Wells. 2013. CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.71, Department of Civil and Environmental Engineering, Portland State University, Portland, Oregon.

Dahlke, H.E., Z.M. Easton, D.R. Fuka, S.W. Lyon, and T.S. Steenhuis. 2009. Modelling variable source area dynamics in a CEAP watershed. *Ecohydrology*, 2: 337-349. DOI: 10.1002/eco.58.

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

Daniel, E.B., J.V. Camp, E.J. LeBoeuf, J.R. Penrod, J.P Dobbins, and M.D. Abkowitz. 2011. Watershed modeling and its applications: A state-of-the-art review. *Open Hydrology Journal*, 5: 26-50.

Davies-Colley R.J., W.N. Vant, and D.G. Smith. 1993. Colour and Clarity of Natural Waters. The Blackburn Press, 210 pages.

DeGaetano, A. T. and C. M. Castellano. 2013. Recent and future changes in extreme rainfall in the Catskills region of New York. *Annals of the New York Academy of Sciences*, 1298: 43–51.

Di Luzio, M., G.L. Johnson, C. Daly, J.K. Eischeid, and J.G. Arnold. 2008. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. *Journal of Applied Meteorology and Climatology* 47(2), 475-497.

Dile, Y.T., and R. Srinivasan. 2014. Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. *J. Amer. Water Res.Assoc.*50(5), 1226-1241.

Dooge, J.C.I., 2005. Bringing it all together. *Hydrology & Earth System Sciences Discussions*, 9: 3-14.

Dooge, J.C.I., 1986. Looking for hydrologic laws. *Water Resources Research*, 22(9S):46-58.

Douglas-Mankin, K.R., R. Srinivansan, and J.G. Arnold. 2010. Soil and Water Assessment Tool (SWAT) model: Current developments and applications. *Transactions of the ASABE,* 53: 1423-1431.

Driese K.L., W.A. Reiners, G.M. Lovett, and S.M. Simkin. 2004. A Vegetation Map for the Catskill Park, NY, Derived from Multi-temporal Landsat Imagery and GIS Data. *Northeastern Naturalist* 11 (4): 421–442.

Dunne, T., and R.D. Black. 1970. Partial area contributions to storm runoff in a small New England watershed. *Water Resource Research* 6: 1296-1311.

Dunne, T., and L Leopold. 1978. Water in Environmental Planning. W.H. Freeman: New York.

Easton, Z.M., D.R. Fuka, M.T. Walter, D.M. Cowan, E.M. Schneiderman, and T.S. Steenhuis. 2008. Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. *Journal of Hydrology*, 348: 279-291.

Effler, S.W., M.G. Perkins, N. Ohrazda, C.M. Brooks, B.A. Wagner, D.L. Johnson, F. Peng and A. Bennett. 1998. Turbidity and particle signatures imparted by runoff events in Ashokan Reservoir, NY. *Lake and Reservoir Management* 14(2-3), 254-265.

Evans C.D., D.T. Monteith, and D.M. Cooper. 2005. Long-term increases in surface water dissolved organic carbon: Observations, possible causes and environmental impacts. *Environmental Pollution* 137 (1): 55–71 DOI: 10.1016/j.envpol.2004.12.031.

Feldman, A.D. 2000. Hydrologic Modeling System HEC-HMS: Technical Reference Manual. US Army Corps of Engineers, Hydrologic Engineering Center.

Findlay, S.E.G. 2005. Increased carbon transport in the Hudson River: unexpected consequences of nitrogen deposition? *Frontiers in Ecology and Environment* 3(3):133-137.

Frankenberger, J.R., E.S. Brooks, M.T. Walter, M.F. Walter, and T.S. Steenhuis. 1999. A GIS-based variable source area hydrology model. *Hydrological Processes* 13: 805-822.

Fuka, D.R., M.T. Walter, C. MacAlister, A.T. Degaetano, T.S. Steenhuis and Z.M. Easton. 2014. Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes* 28(22), 5613-5623.

Gabriel, K.R. and J. Neumann. 1962. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society*, 88: 90–95.

Ganguly S., R.R. Nemani, G. Zhang, H. Hashimoto, C. Milesi, A. Michaelis, W. Wang, P. Votava, A. Samanta, and F. Melton. 2012. Generating global Leaf Area Index from Landsat: Algorithm formulation and demonstration. *Remote Sensing of Environment* 122: 185–202.

Ganguly S., M.A. Schull, A. Samanta, N.V. Shabanov, C. Milesi, R.R. Nemani, Y. Knyazikhin, and R.B. Myneni. 2008. Generating vegetation leaf area index earth system data record from multiple sensors. Part 1: Theory. *Remote Sensing of Environment* 112 (12): 4333–4343.

Gannett Fleming & Hazen and Sawyer. 2008. Catskill turbidity control studies: Phase III Implementation Plan, Prepared for New York City Department of Environmental Protection, Bureau of Engineering Design and Construction.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold. 2007. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. *Trans. ASABE* 50: 1211-1250.

Gelda, R.K., S.W. Effler and F. Peng. 2012. Modeling turbidity and the effects of alum application for a water supply reservoir. *Journal of Environmental Engineering*. 138(1):38-47.

Gelda, R.K., S.W. Effler, A.R. Prestigiacomo, F. Peng, A.J.P. Effler, B.A. Wagner, M.G. Perkins, D.M. O'Donnell, S.M. O'Donnell and D.C. Pierson. 2013. Characterizations and modeling of turbidity in a water supply reservoir following an extreme runoff event. *Inland Waters*. 3(3):377-390.

Gelda, R.K. and S.W. Effler. 2007. Modeling Turbidity in a Water Supply Reservoir: Advancements and Issues. *Journal of Environmental Engineering*. 133(2):139-148.

Gelda, R.K., S.W. Effler, F. Peng, E.M. Owens, and D.C. Pierson. 2009. Turbidity model for Ashokan Reservoir, New York: Case Study. *Journal of Environmental Engineering* 135(9), 885-895.

Geng, S., J.S. Auburn, E. Brandstetter, and B. Li. 1988. A Program to Simulate Meteorological Variables: Documentation for SIMMETEO, Agronomy Progress Rep 204. Department of Agronomy and Range Science, University of California, Davis, California.

Grayson, R., and G. Blöschl. 2001. Spatial patterns in catchment hydrology: observations and modelling. Cambridge University Press.

Gupta, H.V., H. Kling, K.K Yilmaz, and G.F. Martinez. 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377: 80-91.

Gupta, H.V., S. Sorooshian, and P.O. Yapo. 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering* 4: 135-143.

Haith, D.A., and C.L. Shoemaker. 1987. Generalized watershed loading functions for stream flow nutrients. *J. American Water Resources Association* 23: 471-478.

Hanson, C.L. and G.L. Johnson, 1998: GEM (Generation of Weather Elements for Multiple Applications): Its Application in Areas of Complex Terrain. *Hydrology, Water Resources and Ecology in Headwaters IAHS* 248:27-32.

Harpold A.A., D.A. Burns, T. Walter, S.B. Shaw and T.S. Steenhuis. 2010. Relating hydrogeomorphic properties to stream buffering chemistry in the Neversink River watershed, New York State, USA. *Hydrological Processes* 24 (26): 3759–3771 DOI: 10.1002/hyp.7802

Harpold, A.A., S.W. Lyon, P.A. Troch, and T.S. Steenhuis. 2010. The hydrological effects of lateral preferential flow paths in a glaciated watershed in the Northeastern USA. *Vadose Zone Journal* 9: 397-414.

Hoang, L, E.M. Schneiderman, R. Mukundan, K.E.B. Moore, E.M. Owens, and T.S. Steenhuis, 2017. Predicting saturation-excess runoff distribution with a lumped hillslope model: SWAT-HS. *Hydrological Processes* (accepted).

Horton, R.E. 1933. The role of infiltration in the hydrologic cycle. *Transactions of American Geophysical Union* 14: 446-460.

Horton, R.E. 1940. An approach toward a physical interpretation of infiltration capacity. *Soil Science Society of America Proceedings* 4: 339-417.

Ison N.T., A.M. Feyerherm and L.D. Bark. 1971. Wet period precipitation and the Gamma distribution. *J. Applied Meteorology* 10(4): 658–665.

Katz, R.W. 1977. Precipitation as a chain-dependent process. *J. Applied Meteorology*, 16: 671–76.

Khazaei, M.R., S. Ahmadi, B. Saghafian, and B. Zahabiyoun. 2013. A new daily weather generator to preserve extremes and low-frequency variability. *Climate Change*, 119:631–645.

Khomik M., C.A. Williams, M.K. Vanderhoof, R.G. Maclean and S.Y. Dillen. 2014. On the causes of rising gross ecosystem productivity in a regenerating clearcut environment: Leaf area vs. species composition. *Tree Physiology* 34 (7): 686–700 DOI: 10.1093/treephys/tpu049

King, L. M., I.A. McLeod, and S.P. Simonovic. 2015. Improved Weather Generator Algorithm for Multisite Simulation of Precipitation and Temperature. *J. Amer. Water Res.Assoc.* 51(5):1305-1320.

Koenker, R. 2005 Quantile regression, Cambridge University Press.

Koenker, R. 2014. quantreg: Quantile regression, 2011. R package version 5.http://CRAN.R-project.org/package=quantreg.

Koenker, R.W. and V. d'Orey. 1987. Algorithm AS 229: Computing regression quantiles. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 36(3), 383-393.

Krysanova, V., D.-I. Müller-Wohlfeil, and A. Becker. 1998. Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecol. Modelling* 106: 261-289.

Lawrence, M.G. 2005. The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin Amer. Meteorological Society*. 86(2):225-233.

Lenzi, M.A. and L. Marchi. 2000. Suspended sediment load during floods in a small stream of the Dolomites (northeastern Italy). *Catena* 39(4), 267-282.

Lewis, C.D. 1982. Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting. Butterworth Scientific, London.

Li, Z., Brissette, F. and Chen, J. 2013. Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in Nordic watersheds. *Hydrological Processes* 27(25), 3718-3729.

Li, C., V.P. Singh and A.K. Mishra. 2012. Simulation of the entire range of daily precipitation using a hybrid probability distribution. *Water Resources Res.* 48, W03521. DOI: 10.1029/2011WR011446.

Liang, X., D.P. Lettenmaier, E.F. Wood, and S.J. Burges. 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* 99(D7):14415-14428.

Lyon, S.W., M.T. Walter, P. Gérard-Marchant, and T.S. Steenhuis. 2004. Using a topographic index to distribute variable source area runoff predicted with the SCS curve-number equation. *Hydrological Processes* 18: 2757-2771. DOI: 10.1002/hyp.1494.

Matonse, A.H. and A. Frei. 2013. A seasonal shift in the frequency of extreme hydrological events in southern New York State. *Journal of Climate* 26(23), 9577-9593.

Matonse. A.H, D.C. Pierson, A. Frei, M.S. Zion, A. Anandhi, E.M. Schneiderman, and B. Wright. 2013. Investigating the impact of climate change on New York City’s primary water supply. *Climatic Change,* 116:437-456.

Matonse, A.H., D.C. Pierson, A. Frei, M.S. Zion, E.M. Schneiderman, A. Anandhi, R. Mukundan, and S.M. Pradhanang. 2011. Effects of changes in snow pattern and the timing of runoff on NYC water supply system. *Hydrological Processes* 25(21), 3278-3288.

McDonnell, J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, and M. Roderick. 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, 43.

McHale, M.R., P.S. Murdoch and D.A. Burns. 2008. Effects of Forest Harvesting on Ecosystem Health in the Headwaters of the New York City Water Supply, Catskill Mountains, New York. U.S. Geological Survey Scientific Investigations Report 2008-5057, 22 p.,

Mehrotra R, and A. Sharma. 2007. A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability. *J Hydrol (Amst)*, 335: 180−193.

Montanari, A., and D. Koutsoyiannis. 2012. A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48, W09555, doi:10.1029/2011WR11412.

Moore, R. 2007. The PDM rainfall-runoff model. *Hydrol. Earth System Sci. Discussions*, 11: 483-499.

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

Mukundan, R., D. Pierson, E. Schneiderman, S. Pradhanang, M. Zion and A. Matonse. 2013. Factors affecting storm event turbidity in a New York City water supply stream. *Catena* 107, 80-88.

Mukundan, R., D.C. Pierson, L. Wang, A.H. Matonse, N.R. Samal, M.S. Zion, and E.M. Schneiderman. 2013. Effect of projected changes in winter streamflow on stream turbidity, Esopus Creek watershed in New York, USA. *Hydrological Processes* 27(21), 3014-3023.

Murdoch, P.S. and J.L. Stoddard. 1992. The role of nitrate in the acidification of streams in the Catskill Mountains of New York. *Water Resources Research* 28 (10): 2707–2720 DOI: 10.1029/92WR00953

Nash, J.E., and J.V. Sutcliffe. 1970. River flow forecasting through conceptual models part I -- A discussion of principles. *Journal of Hydrology* 10: 282-290. DOI: 10.1016/0022-1694(70)90255-6.

Nathan R.J. and T.A. McMahon. 1990. Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research* 26 (7): 1465–1473 DOI: 10.1029/WR026i007p01465.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2011. Soil and Water Assessment Tool Theoretical Documentation, Version 2009. Texas Water Resources Institute.

New York City Department of Environmental Protection (NYCDEP). 2016. Watershed protection program summary and assessment. March 2016. http://www.nyc.gov/html/dep/pdf/reports/fad_5.1_watershed_protection_program_summary_and_assessment_report_03-16.pdf

Nicks A.D. and G.A. Gander. 1994. CLIGEN: a weather generator for climate inputs to water resource and other model. *In:* Proceedings of the 5th International Conference on Computers in Agriculture. American Society of Agricultural Engineers, St. Joseph, MI, 3–94.

NYSDOH [New York State Department of Health]. 2014. New York City Filtration Avoidance Determination. Final Revised 2007 FAD. 99 p. http://www.health.ny.gov/environmental/water/drinking/nycfad/docs/final_revised_2007_fad_may_2014.pdf.

Obled, C., J. Wendling and K. Beven. 1994. The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology* 159(1-4), 305-333.

Owens, E.M., R.K. Gelda, S.W. Effler and J.M. Hassett. 1998. Hydrologic analysis and model development for Cannonsville Reservoir. *Lake and Reservoir Management*. 14(2-3):140-151.

Pradhanang, S.M., A. Anandhi, R. Mukundan, M.S. Zion, D.C. Pierson, E.M. Schneiderman, A. Matonse and A. Frei. 2011. Application of SWAT model to assess snowpack development and streamflow in the Cannonsville watershed, New York, USA. *Hydrological Processes*, 25:3268-3277.

Pradhanang, S.M., R Mukundan, E.M. Schneiderman, M.S. Zion, A. Anandhi, D.C. Pierson, A. Frei, Z.M. Easton, D. Fuka and T.S. Steenhuis. 2013. Streamflow Responses to Climate Change: Analysis of Hydrologic Indicators in a New York City Water Supply Watershed. *Journal of the American Water Resources Association* 49(6):1308-1326.

Prodanovic, P., and S.P. Simonovic. 2008. Intensity Duration Frequency Analysis under Changing Climatic Conditions. In: 4th International Symposium on Flood Defense: Managing Flood Risk, Reliability and Vulnerability. S.P. Simonovic and P. Bourget (Editors), Toronto, Canada.

Prudhomme, C., E. Sauquet and G. Watts. 2015. Low Flow Response Surfaces for Drought Decision Support: A Case Study from the UK, *J. Extreme Events*, 2(2), 1550005, DOI: 10.1142/S2345737615500050

Prudhomme, C., R.L. Wilby, S. Crooks, A.L. Kay and N.S. Reynard. 2010. Scenario-neutral approach to climate change impact studies: Application to flood risk, *J. Hydrology*, 390(3-4):198-209.

Qian, B., S. Gameda, H. Hayhoe, R. De Jong, and A. Bootsma. 2004. Comparison of LARS-WG and AAFC-WG Stochastic Weather Generators for Diverse Canadian Climates. *Climate Research*, 26:175-191.

Quinn, P.F., and K.J. Beven. 1993. Spatial and temporal predictions of soil moisture dynamics, runoff, variable source areas and evapotranspiration for plynlimon, mid-wales. *Hydrol. Processes* 7: 425-448.

Radcliffe, D.E. and R. Mukundan. 2017. PRISM vs. CFSR Precipitation Data Effects on Calibration and Validation of SWAT Models. *Journal American Water Resources Association* 53(1), 89-100.

Rajagopalan, B., and U. Lall. 1999. A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resources Research* 35(10): 3089–3101.

Rathjens, H., N. Oppelt, D.D. Bosch, J.G. Arnold, and M. Volk. 2015. Development of a grid-based version of the SWAT landscape model. *Hydrological Processes* 29: 900-914.

Rawls, W., L. Ahuja, D. Brakensiek, A. Shirmohammadi, and D. Maidment. 1993. Handbook of Hydrology. Infiltration and soil water movement. McGraw-Hill, New York, 5.1-5.51.

Ray, P. A. and C.M. Brown. 2015. Confronting Climate Uncertainty in Water Resources Planning and Project Design: The Decision Tree Framework. Washington, DC: World Bank.

Raymond, P.A. and J.E. Saiers. 2010. Event controlled DOC export from forested watersheds. *Biogeochemistry* 100 (1): 197–209 DOI: 10.1007/s10533-010-9416-7.

Refsgaard, J.C., and B. Storm. 1995. MIKE SHE. In: *Computer Models of Watershed Hydrology*, V.P. Singh (ed.) Water Resources Publications, Highlands Ranch, CO, pp: 809 - 846.

Refsgaard, J.C., B. Storm, and T. Clausen. 2010. Système Hydrologique Europeén (SHE): review and perspectives after 30 years' development in distributed physically-based hydrological modelling. *Hydrology Research* 41: 355-377. DOI: 10.2166/nh.2010.009.

Richardson, C.W and D.A. Wright, 1984. WGEN: a model for generating daily weather variables. US Department of Agriculture, Agricultural Research Service, ARS-8, 83 p.

Richardson, C.W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, 17: 182–90.

Rode, M., G. Arhonditsis, D. Balin, T. Kedebe, V. Krysanova, A. Van Griensven, and S. van der Zee. 2010. New challenges in Integrated Water Quality Modeling. *Hydrological Processes* 24: 3447-3461.

Roldan, J., and D.A. Woolhiser. 1982. Stochastic daily precipitation models. 1. A comparison of occurrence processes. *Water Resources Research* 18:1451–1459.

Rossi, N., L. DeCristofaro, S. Steinschneider, C. Brown, and R. Palmer. 2015. Potential Impacts of Changes in Climate on Turbidity in New York City’s Ashokan Reservoir. *J. Water Resour. Plann. Manage*., 10.1061/(ASCE)WR.1943-5452.0000614, 04015066.

Running S.W., R.R. Nemani and R.D. Hungerford. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research* 17: 472–483 DOI: 10.1139/x87-081.

Savenije, H.H.G. 2010. Topography driven conceptual modelling (FLEX-TOPO). *Hydrology and Earth System Sciences* 14: 2681-2692. DOI: 10.5194/hess-14-2681-2010.

Sayama, T., J.J. McDonnell, A. Dhakal, and K. Sullivan. 2011. How much water can a watershed store? *Hydrological Processes* 25: 3899-3908.

Schneiderman, E.M., T.S. Steenhuis, D.J. Thongs, Z.M. Easton, M.S. Zion, A.L. Neal, G.F. Mendoza, and M.T. Walter. 2007. Incorporating variable source area hydrology into a curve-number-based watershed model. *Hydrological Processes* 21: 3420-3430. DOI: 10.1002/hyp.6556.

Seeger, M., M.P. Errea, S. Beguerıa, J. Arnáez, C. Martı J. and Garcıa-Ruiz. 2004. Catchment soil moisture and rainfall characteristics as determinant factors for discharge/suspended sediment hysteretic loops in a small headwater catchment in the Spanish Pyrenees. *Journal of Hydrology* 288(3), 299-311.

Semenov, M.A. and E.M. Barrow. 1997. Use of stochastic weather generator in the development of climate change scenarios. *Climate Change* 35:397–414.

Semenov, M.A. and E.M. Barrow. 2002. LARS-WG, A Stochastic Weather Generator for Use in Climate Impact Studies, User Manual. http://www.rothamsted.ac.uk/mas-models/download/LARS-WGManual.pdf.

Sharif, M. and D.H. Burn. 2007. Improved K-Nearest Neighbor Weather Generating Model. *Journal of Hydrologic Engineering* 12:42-51.

Sharma, A. and U. Lall. 1999. A nonparametric approach for daily rainfall simulation. *Mathematics and Computers in Simulation* 48: 361−371.

Sharma, A., D.G. Tarboton, and U. Lall. 1997. Streamflow simulation: A nonparametric approach. *Water Resources Research*, 33(2): 291-308.

Shiau, J.T. and T.J. Chen. 2015. Quantile regression-based probabilistic estimation scheme for daily and annual suspended sediment loads. *Water Resources Management* 29(8), 2805-2818.

Singh, J., H.V. Knapp, and M. Demissie. 2004. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. In: Illinois State Water Survey Contract Report 2004-08.

Siemion, J., D.A. Burns, P.S. Murdoch and R.H. Germain. 2011. The relation of harvesting intensity to changes in soil, soil water, and stream chemistry in a northern hardwood forest, Catskill Mountains, USA. *Forest Ecology and Management* 261 (9): 1510–1519 DOI: 10.1016/j.foreco.2011.01.036

Sillmann, J., V.V. Kharin, X. Zhang, F.W. Zwiers and D. Bronaugh. 2013. Climate extremes indices in the CMIP5 multi-model ensemble. Part 1: model evaluation in the present climate. *J. Geophysical Research* 118: 1716-1733.

Sivapalan, M., R.A. WoodsA, and J.D. Kalma. 1997. Variable bucket representation of TOPMODEL and investigation of the effects of rainfall heterogeneity. *Hydrological Processes* 11: 1307-1330.

Soltani, A. and G. Hoogenboom. 2003. A Statistical Comparison of the Stochastic Weather Generators WGEN and SIMMENTO. *Climate Research* 24:215-230.

Srikanthan, R., and T. McMahon. 2001. Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and Earth System Sciences Discussions*, 5(4): 653-670.

Steenhuis, T.S., M. Winchell, J. Rossing, J.A. Zollweg, and M.T. Walter. 1995. SCS runoff equation revisited for variable-source runoff areas. *J. Irrigation and Drainage Engineering* 121: 234-238.

Steinschneider, S. and C. Brown, 2013. A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, *Water Resour. Res*., 49: 7205–7220.

Stockle, C.O., G.S. Campbell and R. Nelson. 1999. ClimGen Manual. Pullman, WA: Biological Systems Engineering Department, Washington State University.

Tague, C., C. McMichael, A. Hope, J. Choate and R. Clark. 2004. Application of the RHESSys model to a California semiarid shrubland watershed. *J. American Water Resources Assoc.* 40 (3): 575–589.

Tague, C.L. and L.E. Band. 2004. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. *Earth Interactions* 8 (19): 1–42 DOI: 10.1175/1087-3562(2004)8<1:RRHSSO>2.0.CO;2

Tank, A.M.G., F.W. Zwiers and X. Zhang. 2009. Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation, Climate data and monitoring WCDMP-No. 72, WMO-TD No. 1500, 56 pp.

Tarboton, D.G. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research* 33: 309-319. DOI: 10.1029/96WR03137.

Tarboton, D.G., and I.N. Mohammed. 2010. TauDEM 5.1 Quick Start Guide to Using the TauDEM ArcGIS Toolbox. http://hydrology.usu.edu/taudem/taudem5/TauDEM51GettingStartedGuide.pdf

Thaler, J.S. 1996. Catskill Weather. Purple Mountain Press, Fleischmanns, NY.

Thiebeault, J. and A. Seth. 2014. Changing climate extremes in the Northeast United States: observations and projections from CMIP5, *Climatic Change* 127:273–287.

Thiessen A.H. 1911. Precipitation averages for large areas. *Monthly Weather Review*, 39: 1082–1089.

Todorovic, P. and D.A. Woolhiser. 1975. A stochastic model of n-day precipitation. *Journal of Applied Meteorology* 14: 17–24.

Tilahun, S.A., C.D. Guzman, A.D. Zegeye, E.K. Ayana, A.S. Collick, B. Yitaferu, and T.S. Steenhuis. 2014. Spatial and temporal patterns of soil erosion in the semi-humid Ethiopian highlands: a case study of Debre Mawi watershed. In: Nile River Basin, Springer, pp: 149-163.

Turner, B.D.P., M. Guzy, M.A. Lefsky, S.V.A.N. Tuyl, C. Daly, B.E. Law and F. Science. 2003. Effects of land use and fine-scale environmental heterogeneity on net ecosystem production over a temperate coniferous forest landscape: *Tellus B,* 55: 657–668

Upstate Freshwater Institute. 2007. Selected features of the character and origins of turbidity in the Catskill System, Upstate Freshwater Institute, Syracuse, NY.

USDA-NRCS. 2012. Soil Survey Geographic (SSURGO) Database. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed 08/08/2012.

USDA-SCS. 1972. National Engineering Handbook. Natural Resources Conservation Service, U.S. Department of Agriculture.

USDA. 2006. Soil Survey of Delaware County, New York.

van Griensven, A., and T. Meixner. 2006. Methods to quantify and identify the sources of uncertainty for river basin water quality models. *Water Science and Technology* 53: 51-59.

Walling, D. and B. Webb. 1988. The reliability of rating curve estimates of suspended sediment yield: some further comments. In: *Sediment Budgets,* Proc. Porto Allegro Symp., *IAHS Publ. No* 174.

Walter, M., V. Mehta, A. Marrone, J. Boll, P. Gérard-Marchant, T. Steenhuis, and M. Walter. 2003. Simple Estimation of Prevalence of Hortonian Flow in New York City Watersheds. *Journal of Hydrologic Engineering,* 8: 214-218. DOI: 10.1061/(ASCE)1084-0699(2003)8:4(214).

Walter, M.T., E.S. Brooks, M.F. Walter, T.S. Steenhuis, C.A. Scott, and J. Boll. 2001. Evaluation of soluble phosphorus loading from manure-applied fields under various spreading strategies. *Journal of Soil and Water Conservation* 56: 329-335.

Webster, K.L., I.F. Creed, F.D. Beall and R.A. Bourbonnire. 2011. A topographic template for estimating soil carbon pools in forested catchments. *Geoderma* 160 (3–4): 457–467 DOI: 10.1016/j.geoderma.2010.10.016.

White, E.D., Z.M. Easton, D.R. Fuka, A.S. Collick, E. Adgo, M. McCartney, S.B. Awulachew, Y.G. Selassie, and T.S. Steenhuis. 2011. Development and application of a physically based landscape water balance in the SWAT model. *Hydrological Processes* 25: 915-925. DOI: 10.1002/hyp.7876.

Wigmosta, M.S., and D.P. Lettenmaier. 1999. A comparison of simplified methods for routing topographically driven subsurface flow. *Water Resource Research* 35: 255-264.

Wigmosta, M.S., L.W. Vail, and D.P. Lettenmaier. 1994. A distributed hydrology-vegetation model for complex terrain. *Water Resource Research* 30: 1665-1679.

Wilby, R.L. and S. Dessai. 2010. Robust adaptation to climate change, *Weather*, 65(7): 180–185.

Wiley, M.W. and R.N. Palmer. 2008. Estimating the impacts and uncertainty of climate change on a municipal water supply system, *ASCE J. Water Resources Planning Management* 134(3): 239–246.

Wilks, D.S. and R.L. Wilby. 1999. The weather generation game: A review of stochastic weather models, *Prog. Phys. Geogr*., 23: 329–357.

Wilks, D.S. 1992. Adapting stochastic weather generation algorithms for climate change studies. *Climatic Change* 22: 67–84.

Wilks, D.S. 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology* 93: 153–69.

Wood, E.F., D.P. Lettenmaier, and V.G. Zartarian. 1992. A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research: Atmospheres* 97: 2717-2728. DOI: 10.1029/91JD01786.

Worqlul, A., A. Collick, S. Tilahun, S. Langan, T. Rientjes and T. Steenhuis. 2015. Comparing TRMM 3B42, CFSR and ground-based rainfall estimates as input for hydrological models, in data scarce regions: The Upper Blue Nile Basin, Ethiopia. *Hydrology and Earth System Sciences Discussions* 12(2), 2081-2112.

Yang, Y., G. Wang, L. Wang, J. Yu and Z. Xu. 2014. Evaluation of gridded precipitation data for driving SWAT model in area upstream of three gorges reservoir. PloS one 9(11), e112725, doi:10.1371/journal.pone.0112725.

Yang, Y., and L. Wang. 2010. A Review of Modelling Tools for Implementation of the EU Water Framework Directive in Handling Diffuse Water Pollution. *Water Resource Manage*. 24: 1819-1843.

Yates, D., S. Gangopadhyay, B. Rajagopalan and K. Strzepek, 2003. A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *Water Resour. Res.,* 39(7), 7-1 to 7-15.

Young, K.C. 1994. A multivariate chain model for simulating climatic parameters with daily data. *J. Applied Meteorology* 33(6): 661-671.

Young, R.A., C.A. Onstad, D.D. Bosch, and W.P Anderson. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. *J. Soil and Water Conservation* 44: 168-173.

Zabaleta, A., M. Martínez, J.A. Uriarte and I. Antigüedad. 2007. Factors controlling suspended sediment yield during runoff events in small headwater catchments of the Basque Country. *Catena* 71(1), 179-190.

Zhang, X., L. Alexander, G.C. Hegerl, P. Jones, A.K. Tank, T.C. Peterson, B. Trewin, and F.W. Zwiers. 2011. Indices for monitoring changes in extremes based on daily temperature and precipitation data, WIREs *Climate Change* 2: 851–870.

Zhao, R., and X. Liu. 1995. The Xinanjiang model. *In:* Computer models of watershed hydrology. (V.P. Singh, ed.), 215-232. Water Resources Publications, Littleton Colorado.215-232.

Zhao, R., Z. Yilin, F. Leren, L. Xinren, and Z. Quansheng. 1980. The Xinanjiang Model. In: *Hydrological Forecasting*, pg. 351-356, Wallingford IAHS Press.

Zierl, B., H. Bugmann and C.L. Tague. 2007. Water and carbon fluxes of European ecosystems: an evaluation of the ecohydrological model RHESSys. *Hydrologic Processes* 21 (24): 3328–3339.

Zion, M.S., S.M. Pradhanang, D.C. Pierson, A. Anandhi, D.G. Lounsbury, A.H. Matonse, and E.M. Schneiderman. 2011. Investigation and Modeling of winter streamflow timing and magnitude under changing climate conditions for the Catskill Mountain region, New York, USA. *Hydrological Processes* 25(21), 3289-3301.