DOWNWARD MODEL DEVELOPMENT OF THE SOIL MOISTURE ACCOUNTING LOSS METHOD IN HEC-HMS: REVELATIONS CONCERNING THE SOIL PROFILE

A Thesis

Submitted to the Faculty

of

Purdue University

by

Jessica A. Holberg

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Civil Engineering

May 2015

Purdue University

West Lafayette, Indiana

For my family, both new and old.

ACKNOWLEDGEMENTS

I would like to thank Dr. Venkatesh Merwade for his guidance and support as I discovered how to conduct research. His patience and direction have been invaluable to me during the course of this study. Thank you to Dr. Bowling for her advice on the appropriate application of statistical methods and for serving on my advisory committee. I would also like to thank Dr. Cary Troy for serving on my advisory committee.

I would also like to express appreciation to my family, friends, and colleagues who all provided me support when I needed it most.

Lastly, I would like to thank the Graduate School and the Lyles School of Civil Engineering for bringing me to Purdue. Boiler Up!

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Appendix Figure Page

ABSTRACT

Holberg, Jessica A. M.S.C.E., Purdue University, May 2015. Downward model development of the Soil Moisture Accounting method in HEC-HMS and its impact on peak flows. Civil Engineering Professor: Venkatesh Merwade.

Despite the fact that the soil profile is known to impact streamflow, most Curve Number (CN)-based models ignore subsurface processes. This study explores the influence of soil storage on peak flows. Two watersheds in flat, humid west-central Indiana were modeled using both the Natural Resources Conservation Service (NRCS) Curve Number and four versions of the Soil Moisture Accounting (SMA) loss methods in the United States Army Core of Engineers-developed (USACE) Hydrologic Engineering Center Hydrologic Modeling System (HEC-HMS). One watershed encompasses the Wabash and Tippecanoe Rivers' confluence; the other contains an ephemeral stream, Plum Creek. The CN-based model was developed using standard practices, but for the SMA-based model, four increasingly sophisticated SMA loss method arrangements of the two study areas were included and analyzed for summer and winter seasons. All four arrangements contain identical surface characteristics but vary in the soil profile parameters included. The first arrangement includes unlimited soil storage, the second includes limited tension zone storage, the third limits soil storage and includes groundwater parameters, and finally, the fourth includes baseflow characteristics. Results show that the streamflow from the four arrangements differs little for much of the year. However, significant differences in

model results are observed when the causative storm has relatively high maximum precipitation intensity. While these results do not necessarily coincide with the results of previous studies, the departure can be explained by the greater soil profile depth in the watersheds of interest. Comparison of streamflow from both the CN-based and SMA-based models with observed streamflow data show that these models do vary in their prediction of peak flow values.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Hydrologic models are used for a variety of purposes: streamflow forecasting, flood inundation mapping, infrastructure design, and water supply planning, among others. Many hydrologic models, such as the Natural Resources Conservation Service (NRCS) Curve Number-based (CN) model in the US Army Corps of Engineers-designed Hydrologic Engineering Center Hydrologic Modeling Software (HEC-HMS), in use today focus on surface processes but ignore or simplify the soil profile. Soil Moisture Accounting-based (SMA) models exist but are rarely employed, due to the challenges of parameter estimation and calibration (Tramblay et al. 2010). There is a need to clarify the soil profile's impact on streamflow so better planning practices can be used.

When the soil profile is not properly modeled, it can have a dire effect on both the economy and the public's trust in science. In 2013, the National Weather Service (NWS) predicted the Red River of the North in Fargo, North Dakota to crest between 11.6 to 12.8 meters. Citizens rallied, and the city spent approximately two million dollars and hundreds of volunteer hours building temporary sandbag dikes. When the river finally crested, it was at 2.4 meters below the prediction. NWS Hydrologist Steve Buan credits

NWS's inaccurate prediction to the model's failure to account for the dry soil condition that allowed much of the water to seep into the ground. Four years previously, NWS under-predicted the flood peak, and the city saw the highest flood stage in recorded history. This under-prediction was attributed to the model's inability to account for the extremely wet soil condition (Gunderson 2013). Additional conditions, such as frozen soil, may further contribute to the NWS's inaccurate predictions of peak flow in the Red River of the North.

Clearly, rainfall-runoff models need to better capture the antecedent soil moisture condition. SMA-based models continuously adjust the soil moisture based on recent hydrologic activity and soil-water processes; given a suitable spin-up period, the model itself determines the initial soil conditions. Conversely, initial soil conditions must be determined by the modeler before the event-based CN-based model can be run (Tramblay et al. 2010), adjusting the NRCS CN to reflect the Antecedent Moisture Condition (AMC) is a common method for defining initial soil conditions for CN-based models. Many recent studies have explored the deficiencies of this method, however. It results in poor prediction of runoff depth and peak flow—often an under-prediction of these parameters. This is attributed to the fact that the method is empirical, and therefore may not be suitable across a wide range of catchments (Huang et al. 2007, Brocca et al. 2008). This study explores the impact of various elements of the soil profile on peak flows via incorporating increasingly sophisticated soil moisture accounting through downward model development. The primary objectives are:

1. Determine whether a CN-based or SMA-based model more accurately predicts peak streamflow and streamflow recession behavior.
2. Define the specific role of soil profile elements in producing streamflow.
3. Identify characteristics of storm events and watersheds that necessitate modelling the soil profile for optimal hydrologic modelling efficiency.

1.2 Downward Model Development

Dissatisfied with the investigative methods of hydrologists, Klemeš (1983) suggests applying downward model development to solve hydrological problems. Klemeš notes shortcomings in hydrologists' understanding of the scales at which hydrological processes occur; this translates into poor model development practices that choose to ignore science in favor of ungrounded mathematical models with the sole goal of perfectly recreating observed hydrographs. Despite Klemeš' frustrations, many hydrologists have extensively explored streamflow generation mechanisms at various scales. For example, Thomas Dunne has investigated how a basin's spatial structure governs its flow processes (Beighley, Dunne, and Melack 2005) and how vegetation and microtopography affect surficial hydraulic conductivity and thus influence infiltration and surface runoff mechanisms (Dunne, Zhang, and Aubry 1991). Furthermore, many distributed hydrologic models, such as Topmodel (Beven et al. 1995) and MIKE SHE, attempt to properly represent subsurface flow along with a wide variety of flow generation mechanisms.

At its core, downward model development involves starting with an observed behavior at a certain scale and attempting to explain it via interactions at a slightly lower scale. To obtain the best understanding of hydrological processes, downward methods should be used in conjunction with upward methods, upward methods being essentially the inverse of downward (Klemeš 1983). But, this study focuses only on a downward investigation of the soil profile, because a pre-existing hydrologic modeling software, HEC-HMS, will be used to investigate the soil profile.

In practice, downward model development essentially involves creating a series of increasingly sophisticated models of the same process and using the results to pinpoint the influence of specific model processes. The analysis creates an understanding of the interactions between minor processes and their role within the greater context of watershed behavior. The focus is not on the input-output relationships, rather on the internal links of the system (Sivapalan et al. 2003).

Many hydrologists embrace downward model development. It has been used to explore subsurface flow at the catchment scale (Ewen and Birkinshaw 2007), the impact of hydrological parameters on the water balance (Farmer et al. 2003), the effect of storm patterns and soil profile composition on flood frequency (Kusumastuti et al. 2006), how time scales in relation to model complexity impact a model's ability to predict streamflow (Lan-Ahn and Willems 2011), and even to explain how geometric features influence runoff (Sivapalan et al. 2003). This study investigates the role of soil profile processes in shaping the streamflow hydrograph.

Studies by Farmer et al. (2003) and Kusumastuti et al. (2006) serve as the basis for the direction of this study. Farmer et al. (2003) developed an independent hydrologic model for a water balance study using the technique of downward model development. Due to issues of climatic and topographic variability, as well as routing, the study only examines watersheds that are similar in magnitude to 10 km^2. Here, catchments of 4,430 km^2 and 7 km^2 are investigated. The small watershed is selected for similarity with Farmer et al. The large watershed lies within a flood-prone area and serves as a basis for comparison with the smaller watershed. As a semi-distributed model, HEC-HMS does not have the same limitations of Farmer et al. Kusumastuti et al. (2006) also explores the role of the soil profile in generating streamflow but uses synthetic rainfall and hypothetical catchments with shallow soils. Both studies develop independent models based on existing hydrologic theory. This study models two real watersheds with deep soil profiles using historic precipitation as input to the USACE-developed HEC-HMS SMA.

1.3 Mechanisms of Streamflow Generation

Water can enter streams from three different sources: surface runoff, interflow, and baseflow. Surface runoff is precipitation that flows over the land surface to the stream channel instead of being infiltrated into the soil profile. Interflow is essentially subsurface runoff that enters the stream channel by travelling laterally through unsaturated soil in the upper region of the soil profile. When water percolates down into the saturated portion of

the soil profile and then flows underground to the stream channel, it is called baseflow (Gupta 2008).

There are three prevalent theories explaining how runoff is generated. The most common is Hortonian overland flow, in which precipitation in excess of the infiltration capacity of the soil fills surface depressions and runs downslope as overland flow. According to this theory, infiltration capacity is at its maximum in the initial stages of the precipitation event and then quickly decreases to reach a constant rate as the storm progresses. A second theory explains how water is able to flow just under the ground surface in densely vegetated, humid regions. The dense vegetation allows for almost all of the precipitation to be absorbed but floods can still occur due to the lateral transmission of absorbed water through the soil's unsaturated zone. This flow is essentially interflow. The final runoff generation theory addresses the idea of saturation excess flow. When a shallow soil profile is vertically restricted by a bottom bounding layer, such as bedrock, it can become saturated during precipitation events. The saturated soil profile cannot store any additional water, so precipitation immediately becomes overland flow. Generally only a small portion of a basin contributes to saturation overland flow, as such it is part of the variable source concept suggested by Hewlett and Hibbert in 1967 (Gupta 2008).

1.4 Soil Moisture Accounting

The Soil Moisture Accounting loss method is used to investigate soil profile behavior via downward model development. A continuously-simulated model with eight storage

components, the SMA method in HEC-HMS is the most flexible and extensive loss method available for the software (see Figure 1.1). To fully define these eight storage components, a total of 17 parameters are required. An in-depth discussion of parameter determination is included in Chapter 3. SMA is heavily based on Leavesley's Precipitation-Runoff Modeling System (PRMS); its basic operations are described below (HEC 2000). The SMA method in HEC-HMS is a one-dimensional, semi-distributed representation of soil processes. One-dimensional hydrologic models only allow water to flow in one direction during a time-step. This works well for many applications but has the potential to decrease model accuracy at larger spatial scales. Greater variability in topography and soil type is likely to occur when a large spatial scale is considered; a one-dimensional model may fail to capture the complex flow behavior that results from a varied landscape and anisotropic soils. HEC-HMS attempts to solve these issues by including semi-distributed modeling capabilities and multiple storage components in the soil profile. A more complete description of the mathematical models involved can be found in the model technical manual (HEC 2000) and in Bennett (1998).

SMA takes a precipitation hyetograph as its input and routes it through canopy, surface, and soil storages while taking into account groundwater, baseflow, and evapotranspiration processes before outputting a streamflow hydrograph. When precipitation occurs, the canopy storage is first filled; the surface storage is filled next. Once both of these storage components are filled, precipitation has a chance to infiltrate into the ground. If the precipitation intensity is greater than the maximum infiltration capacity of the soil profile, the excess precipitation will become surface runoff instead of

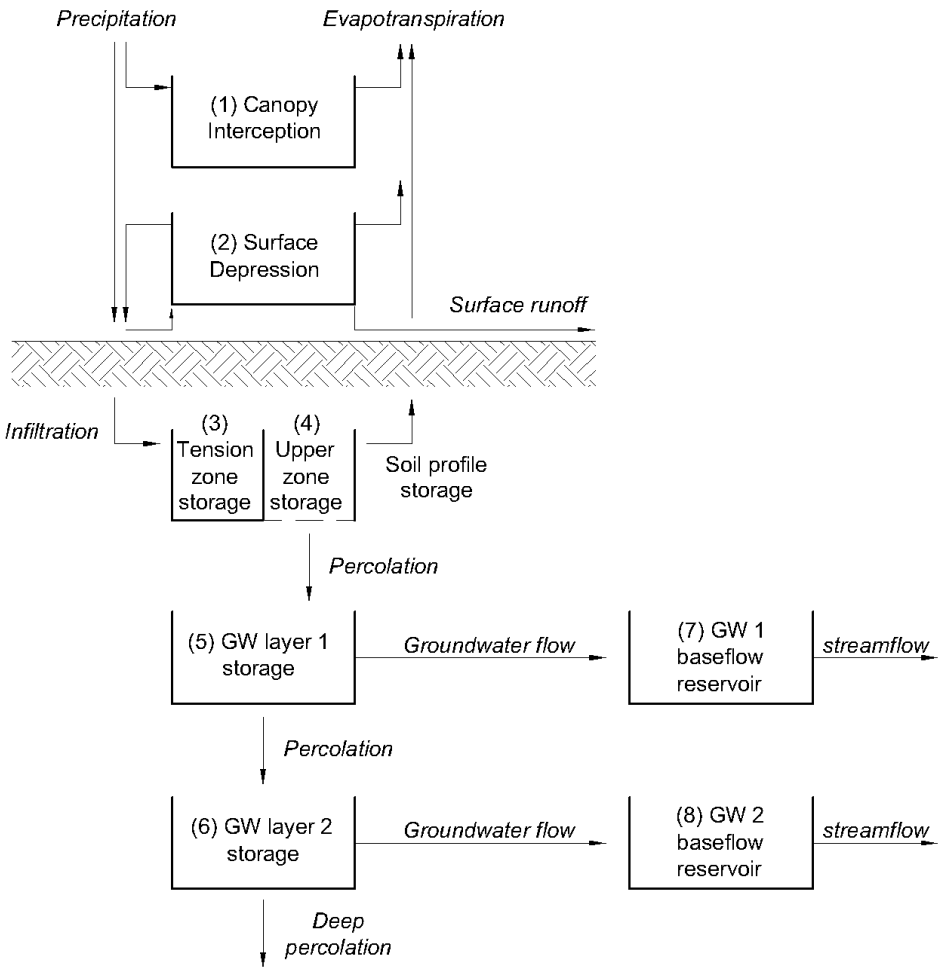


Figure 1.1: SMA Model Configuration, adapted from HEC (2000)

infiltrating. When precipitation infiltrates into the soil, it fills the tension zone first and then the upper zone. Precipitation can percolate from the upper zone, but not from the tension zone, into the groundwater layer one storage (GW1). Some water in GW1 will be routed to the first baseflow reservoir while the rest percolates down to groundwater layer two (GW2). From GW2, water can be transferred to the second baseflow reservoir, otherwise it percolates down to a deep aquifer and is considered lost from the system.

Water in the baseflow reservoirs is transformed to streamflow based on the characteristics of the reservoirs, such as quantity and the flow coefficient.

When precipitation does not occur, evapotranspiration occurs if water is present in the system. The rate of evapotranspiration is dependent upon the weather conditions of the region, but common values for temperate climates are about 170 mm per month during the summer season and 13 mm in the winter months (Fleming 2002). Evapotranspiration first occurs from the canopy storage, then the surface storage. If sufficient water is not present in the first two storage components to fulfill the evapotranspiration potential, water is first removed from the upper zone storage. When evapotranspiration occurs from one of these three storages, water is lost from the system at the potential evapotranspiration rate. If evapotranspiration is still not satisfied, water is then removed from the tension zone storage. Evapotranspiration from the tension zone storage occurs at a decreased rate based on the current soil storage depth and the maximum storage capacity of the tension zone.

1.5 Thesis Organization

This thesis is organized into five chapters. The first chapter contains introductory information and an overview of the Soil Moisture Accounting method in HEC-HMS. The second chapter includes a description of the study areas and an overview of the data used in the study. The third chapter provides an explanation of how the CN-based and SMA-based models are developed and an overview of the statistical methods used to analyze

the model results. The fourth chapter displays and discusses model results. The fifth chapter contains a summary of the study and the conclusions reached.

CHAPTER 2. STUDY AREA AND DATA

This chapter provides information regarding the study areas for which the CN-based and SMA-based models are developed. The data used to develop and validate these models and to perform a frequency storm analysis using the models is also explored.

2.1 Study Area

The study focuses on west-central Indiana. The region is primarily home to agricultural and industrial operations interspersed with small to medium-sized cities. Much of the agricultural land in the region is tile-drained, which greatly affects the generation of surface runoff and subsurface flow. The impact of tile drainage is expected to emerge during the course of this study. The climate of the region is temperate with no pronounced dry season. The area receives an average annual precipitation of 1040 mm, with the summer months producing slightly more precipitation. Soils in the area are primarily descended from limestone, dolomite, and shale. As a result of prior glacial activity, much of the soil is deep glacial till exhibiting little to no relief (USACE 2011). The primary waterway running through this region is the Wabash River, which flows for 820 km. This region was selected for the study, because it provides ample rainfall, flat terrain, and relatively deep soils, lending itself to exploration via downward model development.

2.1.1 Wabash/Tippecanoe Sub-watershed

The Wabash River at Lafayette Watershed (WRLW) covers much of northern Indiana and part of Ohio. The watershed outlet, located in Lafayette, Indiana, is approximately 98 km northwest of Indiana's largest city, Indianapolis. The watershed is 77% cultivated crops, half of which is likely planted on artificially drained soils (Zucker and Brown 1998). The second highest land use in the watershed is deciduous forests at 8.5%. For the purposes of this study, a portion of the WRLW is isolated using two gauges upstream of Lafayette as inflow gauges. This allows the hydrology of the flood-prone WRLW to be modelled despite the dearth of available data for managed reservoirs throughout the watershed; the reservoirs all lie upstream of the isolated study area. As the isolated watershed contains a high-order stream, it is expected that regional subsurface flow initiating outside the study area may be present in the stream network. This may present a challenge for accurately modeling the stream's recession characteristics.

The two United States Geological Survey (USGS) gauges selected for isolating the sub-watershed are the Wabash River at Logansport and the Tippecanoe River below Oakdale Dam. The Wabash River at Lafayette gauge, USGS gauge 03335500, sits at an elevation of 154 m. The Wabash River at Logansport gauge, USGS 03329000, sits at an elevation of 175 m. The Tippecanoe River below Oakdale Dam gauge, USGS 03332605, sits at an elevation of 171 m. By using the Logansport and Oakdale Dam gauges as inflow, the total modeled area is 4,430 km^2 (see Figure 2.1: Study Areas). The elevation within the isolated watershed ranges from 122 to 305 m, with an average elevation of 232 m. The

average slope of the watershed is 2.8%. This sub-watershed encompasses the confluence of the Wabash and Tippecanoe Rivers, and is hereafter called the Wabash/Tippecanoe sub-watershed.

2.1.2 Plum Creek Watershed

Since Farmer et al.'s (2003) theoretical model is only useful for watersheds similar in magnitude to 10 km^2, a significantly smaller watershed, Plum Creek near Bainbridge (Plum Creek), is also investigated. This is to determine if the downward model analysis of Plum Creek provides any indication that the Wabash/Tippecanoe sub-watershed is unsuitable for analysis via downward model development. Plum Creek covers a mere 7 km^2 and is located approximately 57 km west of Indianapolis (see Figure 2.1). As a first-order stream, Plum Creek is expected to generate less subsurface flow than the Wabash/Tippecanoe. This should be evident in the downward model analysis. The watershed is 63% cultivated crops with pasture/hay being the second most dominant land use, claiming 23.5% of the total land area. Plum Creek is monitored by USGS gauge 03357350. The gauge itself sits at an elevation of 252 m, with elevation within the watershed ranging from 252 m to 290 m and with an average elevation of 277 m. The average slope of the watershed is 2.6%. Plum Creek is an ephemeral stream that often runs dry in the summer, adding a unique aspect to this study.

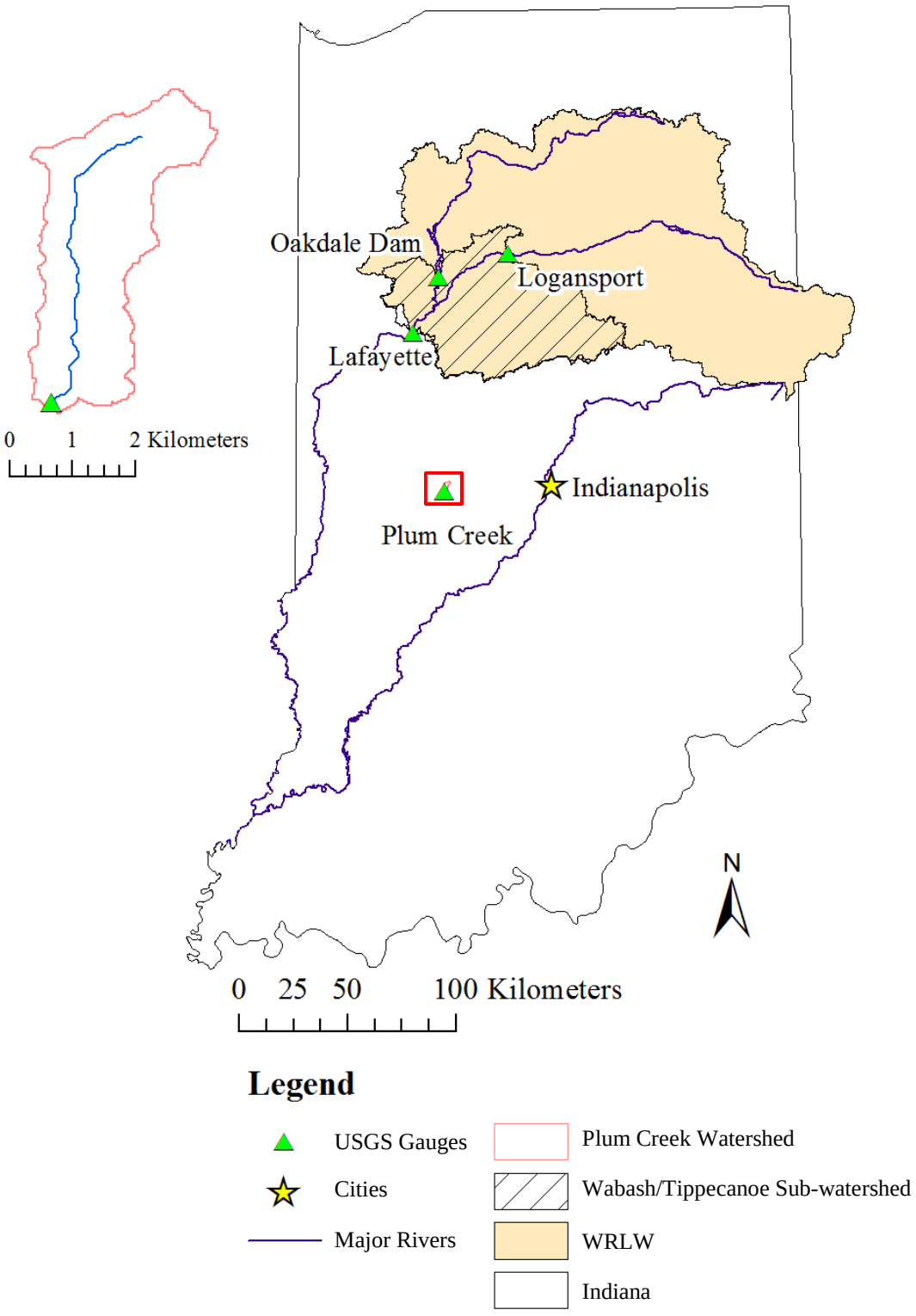


Figure 2.1: Study Areas

2.2 Data

2.2.1 SSURGO Database

The primary source of information used to develop the soil profile parameters required for the SMA-based model is the Soil Survey Geographic (SSURGO) Database. Soil data for areas of interest are available for download from the United States Department of Agriculture (USDA) website, http://websoilsurvey.sc.egov.usda.gov. Covering over 95% of counties, SSURGO stands as the sole authoritative source of soil data in the U.S. With a resolution of 30 m, SSURGO provides a vast supply of soil survey information. The specific SSURGO information used in this study is explored further in the methodology chapter.

2.2.2 DEM

Digital Elevation Models (DEMs) are primarily used to delineate the watersheds and stream networks in the aforementioned study areas. They are also used for topographic calculations such as watershed slope and longest flow path. DEMs contain elevation data for the entire country at a resolution of 30 m. This data is primarily collected via radar and is maintained by the USGS National Elevation Dataset. The DEMs used in this study were published on the National Map Viewer in 2013 and boast a vertical accuracy of 1.55 m (Gesch et al. 2014).

2.2.3 Land Cover

The National Land Cover Dataset (NLCD) 2006 is used to determine the NRCS curve number required for the CN model, the canopy storage grid required for the SMA-based model, and the impervious surface percentage needed for both the CN-based and SMA-based models. The NLCD datasets used in this study are the 30 m resolution land use data and 30 m resolution impervious surface percentage data. This data is maintained by the USGS and is available through the National Map Viewer.

2.2.4 Evapotranspiration Rates

The National Oceanic and Atmospheric Administration (NOAA) publishes state-wide monthly average pan evapotranspiration rates. This data is used as part of the meteorological model within the SMA-based model. For the purposes of this study, a monthly average of the pan evaporation rates from 2008-2012 is used for all years. A correction factor of 0.7 is used to convert the pan evaporation rate to potential evapotranspiration. This is acceptable, because evapotranspiration rates do not vary greatly year to year.

2.2.5 Precipitation

Fifteen minute precipitation data is used for the calibration, validation, and simulation periods of both the CN-based and SMA-based models. The precipitation data covers the years 1993-2003 and 2008-2012. It is obtained for either the most centrally located precipitation gauge or the nearest gauge with data available during the time periods of

interest. In the Wabash/Tippecanoe Sub-watershed, the precipitation gauge at Chalmers, Indiana is used; the Crawfordsville, Indiana gauge serves for the Plum Creek Watershed. The data is available through the National Climatic Data Center (NCDC).

Frequency precipitation data from NOAA's Precipitation Frequency Data Server, http://hdsc.nws.noaa.gov/hdsc/pfds/, is used to create frequency-based storms in HEC-HMS. Frequency precipitation provides an estimated precipitation depth for a specific storm duration and a return period; it includes a 90% confidence interval. HEC-HMS takes frequency precipitation as an input and generates a hyetograph, or frequency-based storm, for the specified storm duration and return period. For this study, a duration of 24 hours is used for return periods of 2, 5, 10, 25, 50, 100, 250, and 500 years. Frequency precipitation is generated using a Gumbel distribution with L-moment estimators to analyze a partial duration rainfall series that has been adjusted using the annual maximum series for the study area (Bonnin et al. 2006).

2.2.6 Streamflow Data

The highest resolution streamflow data, generally either 15-minute or 1-hour, available from the USGS is used in the calibration and validation periods of the models. Daily streamflow data are used to calculate the groundwater layer parameters required in the SMA-based model. Data for the years 2008-2012 are used from the gauges mentioned in the study area descriptions for the two watersheds. The data is available on the USGS website, http://nwis.waterdata.usgs.gov.

CHAPTER 3. METHODOLOGY

This chapter provides information regarding the tools and processes used to develop and analyze the CN-based and SMA-based HEC-HMS models.

3.1 Methodology Overview

The basis of this study is the knowledge that the soil profile has the potential to impact streamflow. Due to their simplicity, many models developed are CN-based, which significantly simplify soil profile parameters. Conversely, SMA-based models include a very developed soil profile with multiple storage components and processes. In order to investigate the effect of a fully-developed soil profile, a methodology is developed in which CN-based and SMA-based HEC-HMS models are created for two watersheds and the results compared. The level at which the soil profile begins to impact streamflow is explored by deconstructing the SMA-based model into four models of increasing complexity. In summary, the methodology consists of the following steps: (1) create CN-based models using standard methods; (2) create four SMA-based models using different soil profile representation or configurations based on downward scaling; (3) compare the results of the CN-based and SMA-based models; (4) perform statistical analysis to investigate precipitation intensity threshold levels, flow persistence, and flow generation mechanisms in the downward-developed SMA-based models.

3.2 HEC-HMS Overview

HEC-HMS contains options for mathematically simulating precipitation, evapotranspiration, infiltration, excess precipitation and transformation, baseflow, and open channel routing (HEC 2000). While primarily an event and CN-based, lumped model, HEC-HMS includes an option for SMA, which is semi-distributed and continuously-simulated, and for distributed runoff using the ModClark transformation method. Within the model framework, HEC-HMS includes basin models, meteorological models, control specifications, and time series data. HEC has also developed an ArcGIS add-in, the Geospatial Hydrologic Modeling Extension (HEC-GeoHMS). With a function allowing direct exportation to the HEC-HMS software, this tool significantly increases the ability to accurately develop a hydrologic model (Abushandi and Merkel 2013). More is explained about the model development process in the following section. HEC-HMS is selected as the modeling tool for this study, because it is flexible, provides reasonable results, and there is extensive literature available concerning its functions and abilities.

3.3 Model Development

3.3.1 CN-based Model Development, Calibration, and Validation

The CN-based HEC-HMS model is primarily developed using ArcGIS tools. ArcHydro, an extension in ArcGIS, is used to process terrain data, define streams, and delineate the watershed of interest. Once this is complete, the HEC-GeoHMS extension is used to

create HEC-HMS project files and assign model parameters. Table 3.1 provides the mathematical models selected for each model process.

Table 3.1 CN-based parameter models[1]

Component	Model
Loss	SCS Curve Number
Transform	SCS Unit Hydrograph
Baseflow	Recession
Routing	Muskingum

To begin the calibration process, precipitation and streamflow data for the years 2009-2011 are added to the time-series component of the model. Three storms each in the summer and winter seasons are selected as calibration storms for a total of six calibration storms. The grounds for storm selection are that the three storms must be: hydrologically isolated (Fleming 2002), occur throughout the season, and result in different magnitudes of peak streamflow. Lag time, percent impervious, and baseflow parameters (recession constant and ratio to peak) have the greatest impact on hydrograph shape and peak flow, so these are the primary parameters adjusted during the calibration process. CN is not calibrated as a means of preserving the physical characteristics of the watersheds as captured by the surface and soil data collected and maintained by the USGS and USDA. During model calibration, five objective functions (see Table 3.2) are used to determine the final model parameters: coefficient of determination (R^2), Nash-Sutcliffe efficiency

[1] The loss method defines what happens to precipitation that does not immediately become runoff. The transform method defines how water is transferred over the ground surface to the stream channel. The baseflow method defines how water is routed from subsurface flow to streamflow. The routing method defines how streamflow is carried down the stream channel.

(NSE), a normalized objective function (NOF), the sum of squared errors (SSE), and the model bias (MB).

Table 3.2 Objective Functions for Calibration

Objective Function	Equation
R^2	$\left(\frac{n\sum_{i=1}^{n}(Q_{obs,i}Q_{sim,i}) - \sum_{i=1}^{n}(Q_{obs,i})\sum_{i=1}^{n}(Q_{sim,i})}{\sqrt{n(\sum_{i=1}^{n}Q_{obs,i}{}^2) - (\sum_{i=1}^{n}\sum Q_{obs,i})^2}\sqrt{n(\sum_{i=1}^{n}Q_{sim,i}{}^2) - (\sum_{i=1}^{n}Q_{sim,i})^2}}\right)^2$
NSE	$1 - \frac{\sum_{i=1}^{n}(Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n}(Q_{obs,i} - \bar{Q}_{obs})^2}$
NOF	$\frac{1}{\bar{Q}_{obs}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{obs,i} - Q_{sim,i})^2}$
SSE	$\sum_{i=1}^{n}(Q_{obs,i} - Q_{sim,i})^2$
MB	$\left(\frac{\overline{Q_{sim}} - \overline{Q_{obs}}}{\overline{Q_{obs}}}\right) \times 100$

N.B. n is number of observations, Q_{obs} is observed streamflow, Q_{sim} is modeled streamflow

After calibration is complete, precipitation and streamflow data for 2012 are added to the time series data. From this year, a summer storm and a winter storm are selected to validate the model. As with the storm selection guidelines mentioned above, the validation storms are also hydrologically isolated and have a different peak flow from the calibration storms. The suitability of the model is determined via the objective functions listed in Table 3.2.

3.3.2 SMA-based Model Development, Calibration, and Validation

As with the CN-based model, most of the model parameters are determined via ArcGIS. The advanced development of the soil profile in the loss method used for SMA requires more extensive processing of land use and soil data than the CN-based model development. Not only does SMA provide a developed soil profile, it also includes surface and canopy storages. Table 3.3 provides the parameter models used in the SMA-based model.

Table 3.3 SMA-based parameter models[2]

Component	Model
Surface	Simple Surface
Canopy	Simple Canopy
Loss	SMA
Transform	SCS Unit Hydrograph
Baseflow	Linear Reservoir
Routing	Muskingum

The surface, canopy, loss, and baseflow methods for the SMA-based model utilize a total of 17 parameters; eight are estimated from soil and land use data (canopy storage, surface storage, infiltration rate, percent impervious, soil percolation rate, soil storage, tension zone storage, groundwater layer 1 percolation rate), four from streamflow recession analysis (groundwater layers 1 and 2 storage depth and coefficient), and five are calibrated (groundwater layer 2 percolation rate, groundwater layers 1 and 2 baseflow

[2] See footnote 1 for description of loss, transform, baseflow, and routing definitions. The surface method defines the amount of surface depression storage available in the watershed. The canopy method defines how much water can be stored on the leaves, branches, etc. of vegetation within the watershed.

coefficient and baseflow reservoir count). Groundwater layer 1 represents interflow, and groundwater layer 2 represents groundwater flow.

3.3.2.1 Parameters Estimated from Land Use Data

The maximum canopy storage and percent impervious grids are both estimated from land use data. The percent impervious grid provided by the USGS is used directly with HEC-GeoHMS, while the canopy storage grid must be calculated. The land cover grid contains NLCD classes whose descriptions can be found at http://www.mrlc.gov/nlcd06_leg.php. Using these descriptions and the values provided in Table 3.4 (Bennett 1998), canopy interception values are assigned to each NLCD class.

Table 3.4 Canopy Interception Values

Type of Vegetation	Canopy Interception (mm)
General Vegetation	1.270
Grasses and Deciduous Trees	2.032
Trees and Coniferous Trees	2.540

3.3.2.2 SSURGO Description

The SSURGO database contains extensive soil data for most of the country. The information is generally downloaded on a county-wide basis and then trimmed to the area of interest. One county download contains geographic information, generally in the form of a soil map compatible with ArcGIS, and a plethora of tables containing information

ranging from soil chemistry to erodibility to flood susceptibility. SSURGO data is organized on three levels (see Figure 3.1): map units, components, and horizons. A map unit is a geographic region that contains soils with properties that are different from neighboring soils. One map unit typically consists of a few different components. A component is a single type of soil, also known as a soil series. Each component has multiple horizontal soil layers, or horizons, all of the same soil type.

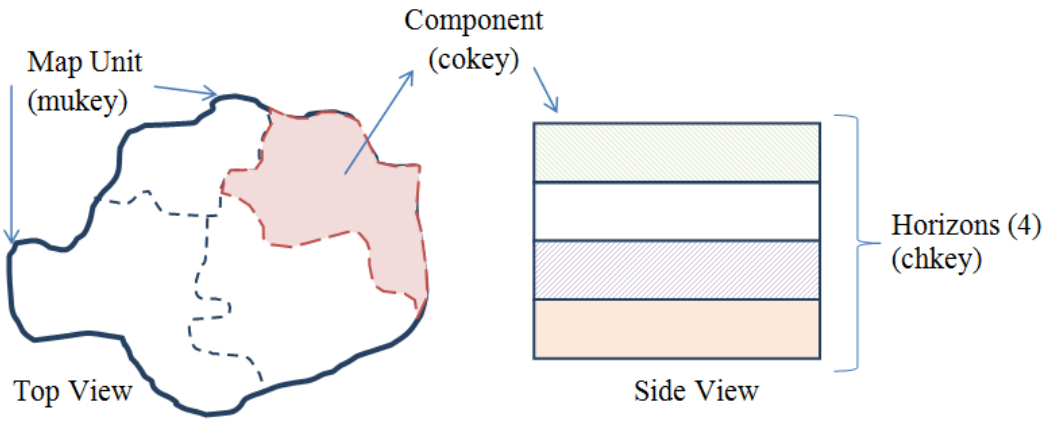


Figure 3.1: SSURGO Organization

Each map unit can be identified by a unique identifier: an mukey. This mukey is connected to each piece of information concerning that map unit throughout all of the tables provided in the SSURGO database. Similarly there are component keys (cokey) and horizon keys (chkey). When the same component is found in different map units, that component will always have the same cokey but a different mukey. See Figure 3.2 for an example of how mukeys, cokeys, and chkeys are used. Note that since Component 2 is found in both Map Unit 1 and 2, it has the same cokey but different mukeys.

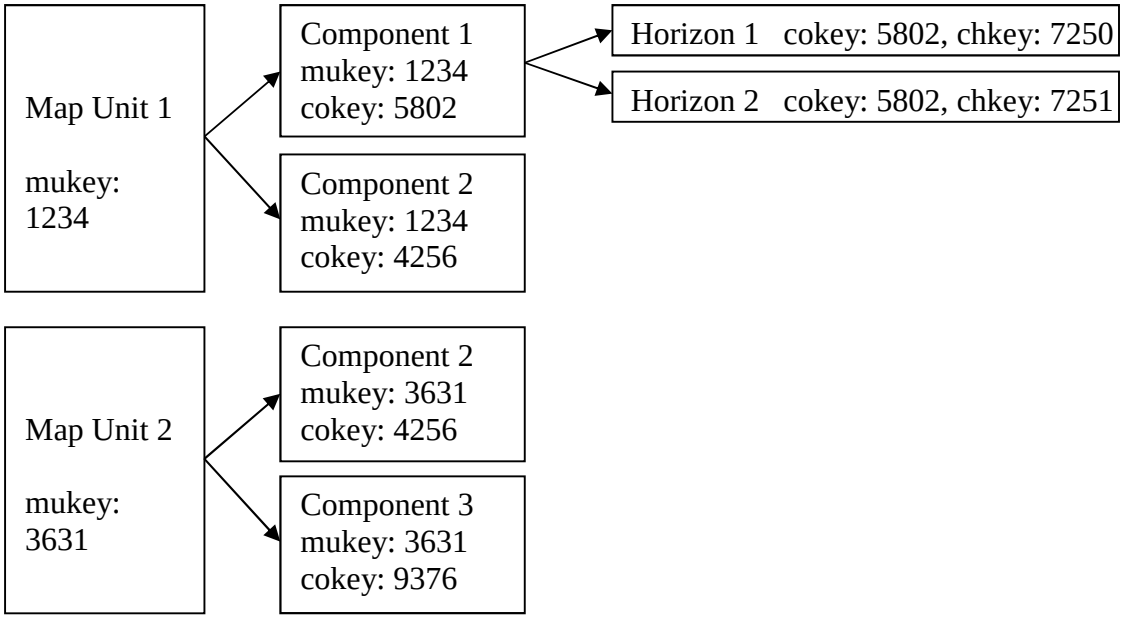


Figure 3.2: Map Unit, Component, and Horizon Identifiers

3.3.2.3 Parameters Estimated from SSURGO

Six SMA parameters are estimated from the SSURGO database: maximum surface storage, maximum infiltration rate, maximum soil percolation rate, soil storage, tension zone storage, and groundwater layer 1 maximum percolation rate. To estimate these parameters, only the map unit feature class and the chorizon and component tables in SSURGO are required. The chorizon table contains information about the soil horizons, while the component table includes information about the soil components. Table 3.5 contains the fields required for calculating the SMA parameters.

Table 3.5 SSURGO Table Field Definitions

Table	Field	Definition
Horizon	chkey	Horizon ID
	cokey	Component ID
	ksat_r	Representative saturated hydraulic conductivity
	hzdepb_r	Representative depth from soil surface to bottom of layer
	wsatiated_r	Representative soil porosity
	wthirdbar_r	Representative field capacity
Component	mukey	Map Unit ID
	cokey	Component ID
	comppct_r	Representative component percent
	slope_r	Representative ground slope

3.3.2.3.1 Soil Data Preprocessing

The chorizon and component tables are first prepared before calculating the required parameters. The chorizon table is exported to a spreadsheet and a running count of the number of horizons in each component is created. Next, the average saturated hydraulic conductivity, soil porosity, and field capacity values are calculated for each component by simply averaging the values for each horizon within the component. The saturated hydraulic conductivity for the topmost horizon in each component and the depth from the soil surface to the base of the bottommost horizon in each component are determined. Determination of these parameters is depicted in Figure 3.3. Note average soil porosity and field capacity have been excluded for the sake of brevity, but the calculations are identical to those of average saturated hydraulic conductivity.

chorizon Table

cokey	**chkey**	Number of Horizons	ksat_r	ksat_avg	ksat_layer1	hzdepb_r
9391673	26349605	1	21.88	-	21.88	23
9391673	26349604	2	21.88	-	-	48
9391673	26349606	3	0.92	14.89	-	152
9391674	26349607	1	23.29	-	23.29	23
9391674	26349608	2	23.29	-	-	76
9391674	26349609	3	0.92	15.83	-	203

Average saturated hydraulic conductivity: average of ksat_r for horizons 1, 2 and 3

Saturated hydraulic conductivity for topmost horizon

Depth from soil surface to bottommost horizon

Preprocessing Result

cokey	ksat_avg	ksat_layer1	hzdepb_r
9391673	14.89	21.88	152
9391674	15.83	23.29	203

Figure 3.3: SSURGO Preprocessing Example

All of the fields except those calculations mentioned here and the cokey field are deleted and the spreadsheet is imported back into ArcGIS. Using the joins and relates function in ArcGIS, the mukey, component percent, and ground slope from the component table are added to the edited chorizon table. The cokey is used as the common field. The edited chorizon table is then re-exported to a spreadsheet, and a weighted average parameter is calculated for each map unit based on the percent composition of each soil series. To achieve this, a running count of the number of components associated with each map unit is calculated, similar to the running count of horizons mentioned previously. At this point, the preprocessing is complete.

3.3.2.3.2 Parameter Calculations

SMA parameter calculations are performed in a spreadsheet. One value for each parameter is calculated for an entire map unit. ArcGIS is then used to create rasters from the spreadsheet. The rasters are directly used with the parameter estimation function in HEC-GeoHMS. A description of the calculations performed follows.

Maximum Surface Depression Storage. Surface depression storage is precipitation that is held at the ground surface in hollows or indentations. It can only escape through evaporation or infiltration into the soil. Previous studies indicate that the amount of water retained on the ground surface is related to the ground slope (see Table 3.6) (Bennett 1998). As such, the weighted average slope of each map unit is calculated by using the ground slope and component percent values. Using Table 3.6, surface storage values are assigned to each map unit (see Figure 3.4).

Table 3.6 Surface Depression Storage Values

Description	Slope (%)	Surface Storage (mm)
Paved Impervious Areas	NA	3.18-6.35
Flat, Furrowed Land	0-5	50.8
Moderate to Gentle Slopes	5-30	6.35-12.70
Steep, Smooth Slopes	>30	1.02

*taken from Fleming, 2002

Map Unit: 2387

Cokey: 4320 Component %: 45 Slope (%): 14	Cokey: 5625 Component %: 23 Slope: (%): 1	Cokey: 3467 Component %: 32 Slope (%): 3

Sample Calculation:

$$Weighted\ Avg.Slope = \left(\frac{45}{100} \times 14\right) + \left(\frac{23}{100} \times 1\right) + \left(\frac{32}{100} \times 3\right) = 7.49\%$$

From Table 3.6, surface depression storage is about 9.5 mm.

Figure 3.4: Surface Depression Storage Sample Calculation

Maximum Infiltration Rate. The maximum infiltration rate or infiltration capacity is the fastest rate at which precipitation can seep from the ground surface into the soil profile. The hydraulic conductivity of a soil is greatest when the soil is saturated; it decreases significantly as the water content of the soil decreases. SMA mimics this relationship by relating the infiltration rate to soil storage availability (HEC 2000). Since the maximum hydraulic conductivity is the saturated hydraulic conductivity, the maximum infiltration rate of each map unit is taken as the weighted average of the saturated hydraulic conductivity for the topmost horizon of each component (see Figure 3.5). This is achieved using the component percent and the saturated hydraulic conductivity of the first horizon (Fleming 2002).

Map Unit: 2387

Cokey: 4320 Component %: 45 Layer 1 saturated hydraulic conductivity (µm/s): 9.17	Cokey: 5625 Component %: 23 Layer 1 saturated hydraulic conductivity (µm/s): 28.23	Cokey: 3467 Component %: 32 Layer 1 saturated hydraulic conductivity (µm/s): 91.74

Sample Calculation:

$$Max.Infiltration\ Rate = \left(\frac{45}{100} \times 9.17\right) + \left(\frac{23}{100} \times 28.23\right) + \left(\frac{32}{100} \times 91.74\right)$$

$$= 39.98\ \mu m/s$$

Figure 3.5: Maximum Infiltration Rate Sample Calculation

Maximum Percolation Rate. Percolation is the process by which water is transferred through the soil profile and groundwater layer(s). This generally occurs due to gravity, but can also occur due to capillary forces (Chow 1964). The percolation rate is limited by the lowest hydraulic conductivity of the soil layers through which the water is travelling (Zaslavsky and Rogowski 1969). In this study, the average saturated hydraulic conductivity of all horizons in a component is used to calculate the maximum percolation rate, as described in Bennett (1998) and Fleming (2002). The maximum percolation rate is taken as the weighted average of the horizon-average saturated hydraulic conductivity for all components in a map unit (see Figure 3.6). Refer back to Section 3.3.2.3.1 for clarification, as the approach is similar to what is described there. This percolation rate calculated here is used for both the soil profile and groundwater layer 1 percolation rates.

Map Unit: 2387

Cokey: 4320 Component %: 45 Average saturated hydraulic conductivity (μm/s): 4.65	Cokey: 5625 Component %: 23 Average saturated hydraulic conductivity (μm/s): 12.70	Cokey: 3467 Component %: 32 Average saturated hydraulic conductivity (μm/s): 37.76

Sample Calculation:

$$Max.Percolation\ Rate = \left(\frac{45}{100} \times 4.65\right) + \left(\frac{23}{100} \times 12.70\right) + \left(\frac{32}{100} \times 37.76\right)$$

$$= 17.1\ \mu m/s$$

Figure 3.6: Maximum Percolation Rate Sample Calculation

Maximum Soil Profile Storage. The maximum soil profile storage is the storage depth available in voids and soil pores when the soil is dry. Soil voids can be drained by gravity or evaporation (HEC 2000). The soil profile storage is calculated by multiplying the component percent, average porosity, and the depth from the soil surface to the deepest horizon together for each component and then summing these values to reach a total for each map unit (see Figure 3.7). Porosity is the fraction of total soil volume that is not occupied by the soil medium; it includes voids and pore space (Chow 1964).

Map Unit: 2387

Cokey: 4320 Component %: 45 Porosity (%): 33 Depth from soil surface (cm): 203	Cokey: 5625 Component %: 23 Porosity (%): 37 Depth from soil surface (cm): 152	Cokey: 3467 Component %: 32 Porosity (%): 42 Depth from soil surface (cm): 191

Sample Calculation:

$$Max.Soil\ Storage$$

$$= \left(\frac{45}{100} \times \frac{33}{100} \times 203\right) + \left(\frac{23}{100} \times \frac{37}{100} \times 152\right) + \left(\frac{32}{100} \times \frac{42}{100} \times 191\right)$$

$$= 68.75\ cm$$

Figure 3.7: Maximum Soil Storage Sample Calculation

Maximum Tension Zone Storage. The maximum tension zone storage is the storage depth available in the form of water attached to soil particles (HEC 2000). This water can only be removed via evaporation, suction, or contact with a dry, porous material (Jury and Horton 2004). Field capacity is the amount of water left in the soil profile after water has stopped draining from the soil; it is analogous to the tension zone (Veihmeyer and Hendrickson 1931). The tension zone storage is calculated by multiplying the component percent, average field capacity, and the depth from the soil surface to the deepest horizon together for each component and then summing these values to reach a total for each map unit (see Figure 3.8).

In Figure 1.1, presented in the introduction chapter, the soil profile is shown to have two parts: the tension zone and the upper zone. SMA does not require a value for the upper

zone directly; rather it calculates the storage depth of the upper zone as the maximum soil profile storage minus the maximum tension zone storage (HEC 2000).

Map Unit: 2387

Cokey: 4320 Component %: 45 Field capacity (%): 27 Depth from soil surface (cm): 203	Cokey: 5625 Component %: 23 Field capacity (%): 10 Depth from soil surface (cm): 152	Cokey: 3467 Component %: 32 Field capacity (%): 39 Depth from soil surface (cm): 191

Sample Calculation:

$$Max.Tension\ Zone\ Storage$$

$$= \left(\frac{45}{100} \times \frac{27}{100} \times 203\right) + \left(\frac{23}{100} \times \frac{10}{100} \times 152\right) + \left(\frac{32}{100} \times \frac{39}{100} \times 191\right)$$

$$= 52.00\ cm$$

Figure 3.8: Maximum Tension Zone Storage Sample Calculation

3.3.2.4 Parameters Estimated from Streamflow

This section explains the calculation of the four parameters estimated from streamflow recession analysis: groundwater layers 1 and 2 storage depth and coefficient. Groundwater layer 1 (GW1) represents interflow, and groundwater layer 2 (GW2) represents groundwater flow (Fleming 2002). Interflow is water that flows laterally through the soil profile when the water content falls between field capacity and saturation (Steenhuis and Muck 1988).

Streams carry water from three different sources: surface runoff, surface soil (interflow), and groundwater. A streamflow hydrograph can be deconstructed into its various components to calculate the aforementioned parameters (Linsley et al. 1958). For this process, six hydrologically isolated storms from different months are used. The storms used are independent of the calibration storms. Daily streamflow values are plotted on a semi-logarithmic plot for this analysis. Excel is used to perform the recession analysis.

Streamflow hydrographs contain three regions: a rising limb, a peak, and a receding limb. The tail-end of the receding limb represents the time when groundwater is the only source contributing to streamflow, as surface runoff and interflow have stopped (Linsley et al. 1958). At this point, an inflection point is visible, indicating the end of surface runoff (see Figure 3.9). To begin the deconstruction process, the groundwater is separated from the baseflow by projecting a line backwards from the tail-end of the receding limb to the time of peak flow while maintaining the slope of the tail-end portion. This line is then connected to the point at which the hydrograph begins to rise as a result of runoff. This is the groundwater contribution to streamflow, or GW2 (Linsley et al. 1958, Fleming 2002). It is the dashed line in Figure 3.9.

To determine the portion of the hydrograph that is made up of surface runoff and interflow (SR-I), the groundwater is subtracted from the total streamflow hydrograph. This is depicted as the dash dot line in Figure 3.9. To separate interflow from the SR-I portion, a line is projected backwards from the area of lowest slope in the receding limb

of the SR-I to the time of peak flow. As with the groundwater separation, this line is then connected to the point at which the SR-I hydrograph begins to rise. This is the interflow contribution to streamflow, or GW1 (Linsley et al. 1958, Fleming 2002). It is the dotted line in Figure 3.9. The SR-I and Interflow lines are truncated, because they drop to zero after the final point shown, and zero values cannot be plotted on a logarithmic axis.

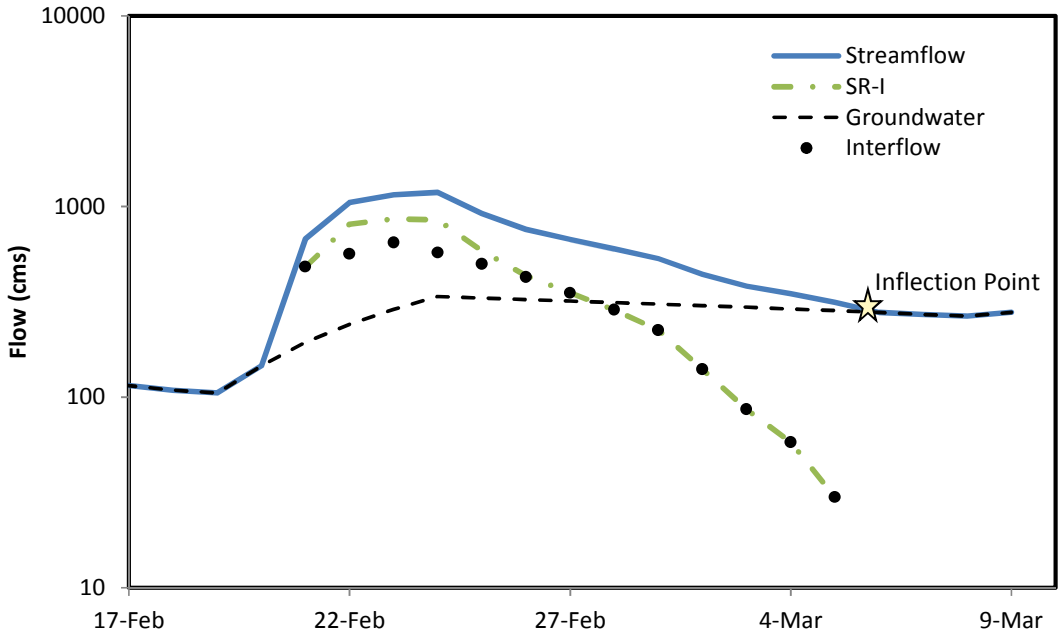


Figure 3.9: Streamflow Hydrograph Deconstruction

Using the data from the groundwater and interflow lines calculated above, the SMA parameters are calculated. The recession curve, or receding limb of a hydrograph, can be described by Equation 3.1, below.

$$q_t = q_0 K_r = q_0 * \exp(-\alpha t) \tag{3.1}$$

Where q_0 is the initial streamflow, q_t is the streamflow at a later time, t, K_r is a recession constant less than one, and $\alpha = -\ln K_r$. The recommended time step for streamflow regression analysis is one day, but a shorter time step can be used for a smaller basin (Linsley et al. 1958). Using the area of shallowest slope of the receding limb of the groundwater hydrograph and Equation 3.1, the GW2 α-value for each step is calculated. After averaging the α-values for the current hydrograph, the GW2 Recession Coefficient is calculated using Equation 3.2, below.

$$Recession\ Coefficient = 1/\alpha \tag{3.2}$$

Using the same section of the groundwater hydrograph and Equation 3.3, the GW2 Storage Depth is calculated for each step. The maximum value produced by this calculation is taken as the GW2 Storage Depth, or storage capacity. The maximum instantaneous storage is used for the storage depth, because it is the most accurate estimate of storage capacity that can be obtained using streamflow recession analysis.

$$S_t = \frac{q_t}{\alpha \times A} \tag{3.3}$$

Where S_t is the storage in the basin at time, t and A is the area of the watershed. The same calculations are repeated using the interflow hydrograph to determine the GW1 Recession Coefficient and GW1 Storage Depth.

Once complete, the values are summarized in one spreadsheet and examined to see how they changed over different months and seasons. Since there is a fairly drastic difference between the parameters calculated for summer and winter storms, it is evident a bi-annual hydrologic model is necessary to accurately capture watershed behavior (see Table 3.7).

Using the parameters as guidelines, July to November is set as the summer season and December to June is set as the winter season. Once the bi-annual model has been determined, the recession coefficients and storage capacities are averaged across the relevant months to provide one parameter value of each type for each season.

Table 3.7 Streamflow Recession Analysis

Event Month	GW2 Recession Coefficient (hr)	GW2 Storage (mm)	GW1 Recession Coefficient (hr)	GW1 Storage (mm)
Oct	400	31	70	3
Sept	414	27	35	2
July	324	26	76	5
May	547	207	34	10
Apr	324	236	57	36
March	439	168	85	13

3.3.2.5 Model Preparation, Calibration, and Validation

Once the aforementioned parameters are calculated, HEC-GeoHMS is used to assign subbasin parameters and export the project files to HEC-HMS. In HEC-HMS, the monthly pan evapotranspiration data is added to meteorological models; precipitation and streamflow data from 2009-2011 are added to the time series data. At this point, the model is copied and one designated the winter model and the other for summer. The season-specific GW1 and GW2 parameters are assigned for the SMA-based model. Initial values of the calibration-determined parameters, GW2 percolation rate, GW1 and GW2 baseflow coefficient and baseflow reservoir count, are set. A sensitivity analysis indicates that percent impervious is the most sensitive parameter in the model. It shows that the

GW2 percolation rate has little to no influence on storm event streamflow; this parameter is not altered during calibration.

The calibration session begins with running the model and examining the baseflow output. GW1 and GW2 baseflow coefficients and number of baseflow reservoirs are adjusted to permit the groundwater to travel through the baseflow model with little to no attenuation (Fleming 2002). The linear reservoir baseflow method in HEC-HMS is based on the Clark Unit Hydrograph (UH) method for transferring flow through reservoirs. The GW1 and GW2 baseflow coefficients are analogous to the attenuation, or storage, coefficient in the Clark UH method (HEC 2000) and similar to the GW1 and GW2 storage coefficients calculated in Section 3.3.2.4. Interflow (GW1) travels faster than groundwater flow (GW2), but slower than surface runoff (Kirkby 1978). The GW1 coefficient should be smaller than the GW2 coefficient. A high baseflow coefficient means that less of the inflow to the reservoir is immediately transferred through the reservoir; rather it will have a higher residence time in the reservoir. Once these values are set, the calibration continued by testing various percent impervious values and determining model performance with the objective functions listed in the previously presented Table 3.2. This process did not yield satisfactory results, so another influential parameter is also considered: surface depression storage.

As mentioned in the study area descriptions, both watersheds of interest likely have extensive artificial drainage. Artificial drainage captures and conveys soil water to the edge of a cultivated field, where it is then transferred to a local stream or surface ditch

(Skaggs et al. 1994). With a clay fraction of about 0.21 for both watersheds, the presence of artificial drainage is expected to increase peak streamflow (Rahman et al. 2014). HEC-HMS does not have a built-in function to express this behavior; reducing the surface depression storage of the watershed best mimics artificial drainage. Decreasing surface storage results in more precipitation becoming surface runoff rather than infiltration, which produces the same result as artificial drainage: quicker conveyance of water to the stream. Therefore, the maximum surface storage is reduced to 12.7 mm from the 50.8 mm as recommended by Fleming (2002). This value agrees with Chow (1964). Final surface depression storage values are determined via calibration.

After satisfactory model calibration, the validation process begins. Precipitation and streamflow data from 2012 are added to the time series data. The models are run; the aforementioned objective functions (see Table 3.2) serve to indicate model suitability. Refer to Section 3.3.1 for an explanation of the CN-based model calibration technique.

3.3.2.6 Downward Model Development

The completed SMA-based model is split into four models of increasing sophistication per Farmer et al. (2003) and Kusumastuti et al. (2006). The first model includes unlimited soil storage, the second includes limited tension zone storage, the third limits soil storage and includes groundwater parameters, and finally, the fourth includes baseflow characteristics (see Table 3.8 and Figure 1.1, included below). This configuration permits

inferences to be drawn concerning the impact of specific soil parameters on streamflow (Klemeš 1983).

Table 3.8 Model Elements

Model	Elements
M1	1, 2, 4 (unlimited)
M2	1, 2, 3, 4 (unlimited)
M3	1, 2, 3, 4 (limited), 5, 6
M4	1, 2, 3, 4 (limited), 5, 6, 7, 8

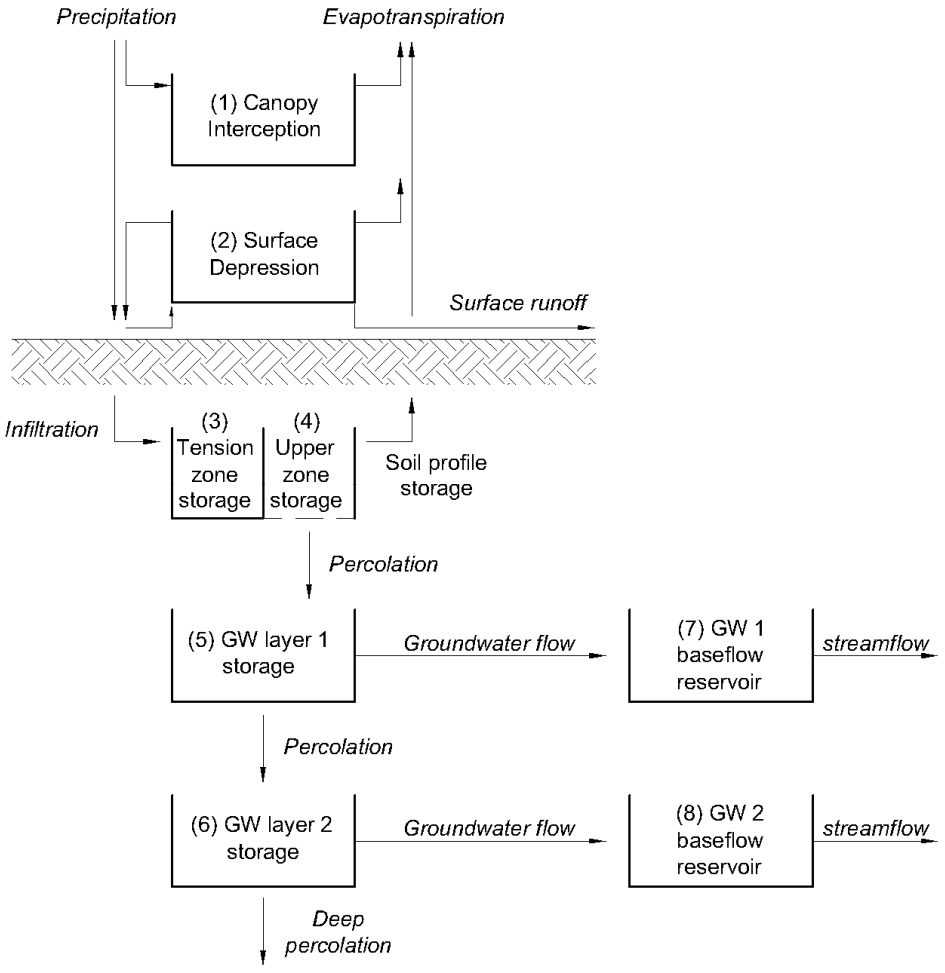


Figure 1.1: SMA Model Configuration, adapted from HEC (2000)

3.4 Analyzing Model Results

The downward models are run continuously at a 10-minute time step for a 10-year simulation period from 1994 to 2003 using historic precipitation data, and the results are analyzed via the methods described in this section. Ten minutes is used for model computations, because it allows a high resolution investigation of the differences in modelled streamflow. When running the Wabash/Tippecanoe models, the inflow hydrographs at Logansport and Oakdale Dam are excluded, allowing for the isolation of the watershed. Thus, discrepancies in streamflow among the four downward-developed models are easily identified. The models are also run for frequency-based storms. Refer to Section 2.2.5 for a discussion of frequency-based storms. The sections following explain the methods used for analyzing the model results.

3.4.1 CN-based and SMA-based Model Comparison

To determine whether the CN-based or SMA-based model captures the hydrologic behavior of the watershed better, the model streamflows are compared with the observed streamflow. Berthet et al. (2009) states that an objective function, time to peak error, and a visual comparison of the observed and modeled hydrographs can serve as a basis for determining which model performs better. As such, these three methods are used to compare CN-based and SMA-based model performance.

3.4.2 SMA-based Downward-developed Models

The impact of the soil profile is examined via an analysis of the variation seen in the four downward-developed model results. A sign test, flow duration curves, and flood frequency analysis serve as the basis for this examination.

3.4.2.1 Sign Test

The sign test can be used to determine whether data pairs are typically different from each other. The null hypothesis of the test is that there is not a statistically significant difference in the data pairs. It is a fully nonparametric test, as it does not require assumptions of normality or symmetry (Helsel and Hirsch 2002). Examination of the downward model development results indicates that the soil profile only influences streamflow after a storm event with a high precipitation intensity. So, the sign test is applied at an alpha value of 0.05 to peak flows associated with specific maximum precipitation intensities. Only peak flows above the 90^{th} or 95^{th} flow percentile for the Wabash/Tippecanoe and the Plum Creek watersheds, respectively, are tested. A total of 149 peak flows are tested for the Wabash/Tippecanoe, and 504 peak flows are tested for Plum Creek. These peak flows represent every peak flow above the aforementioned flow percentiles that occurs during the ten year simulation period. Local peak flows due to first flush runoff were omitted when detected. Plum Creek exhibits significantly more peak flows than the Wabash/Tippecanoe, because it is a much flashier watershed with a time to peak of approximately two hours. The data pairs used are M1 vs. M2, M2 vs. M3, and M3 vs. M4. These are selected, because M4 streamflow is always greater than or equal to

M3 streamflow, M3 streamflow is always greater than or equal to M2 streamflow, etc. As such, it is inexpedient to test M1 vs. M4, because the result can be reasonably inferred from the result of M3 vs. M4. Also, unnecessarily testing additional data pairs simply reduces the power of the test (Kutner et al. 2005).

Since the sign test is performed multiple times with the same set of data, the issue of multiple comparisons is considered. When the same set of data is used for simultaneous hypothesis testing, it increases the probability that the test will return an incorrect conclusion. To protect against this error, a correction is made to the alpha value. In this case, a Bonferroni Correction is the most appropriate since it does not assume anything regarding the distribution of the data (Kutner et al. 2005). The corrected alpha value for the sign test is 0.0083 (0.05/(2*3)).

3.4.2.2 Flow Duration Curve

Flow duration curves are developed for M1-4 to show the distribution of flow values. A flow duration curve is simply a quantile plot of the flow data (Helsel and Hirsch 2002). It can be useful in determining general flow characteristics, such as the impact of baseflow or how quickly a watershed transitions from high to low flows (Farmer et al. 2003). Flow duration curves are developed using the plotting position formula shown in Equation 3.4.

$$p = \frac{i}{n+1} \tag{3.4}$$

Where p is the exceedance probability, i is the rank of the data, and n is the total number of data points. For the Wabash/Tippecanoe sub-watershed, the flow relative to the median flow is plotted against the exceedance probability. For the Plum Creek watershed, flow is simply plotted against the exceedance probability, because the median flow in the watershed is zero.

3.4.2.3 Flood Frequency Analysis

Flood frequency analysis is carried out using the built-in frequency storm function in HEC-HMS. Precipitation data from the NOAA's Precipitation Frequency Data Server is used to create frequency storms with the following return periods: 2, 5, 10, 25, 50, 100, 250, and 500 years. The simulated peak flows from each frequency storm are plotted for M1-4. In downward model development, frequency storms can be useful in identifying shifts in the dominant flow mechanism (Kusumastuti et al. 2006).

CHAPTER 4. RESULTS AND DISCUSSION

The results of the CN-based and SMA-based model calibration and validation are presented herein along with findings from statistical analyses. A discussion of their meaning and significance is also included in this chapter.

4.1 Model Parameter Values

The results of the SMA-based model calibration are shown in Table 4.1 for both watersheds of interest. The Wabash/Tippecanoe sub-watershed is modelled using 17 sub-basins; Table 4.1 shows parameters for only one of the 17 sub-basins. The results shown are typical. Note the similarity in values between the two watersheds for the parameters calculated via the land use, SSURGO, and streamflow recession analyses. This is expected, as the watersheds lie in a region of geographic similarity (Gray 2000). The most striking differences seen in the values are the maximum surface storage and GW2 coefficient. For the maximum surface storage, Wabash/Tippecanoe requires a value of 7.3 mm, while Plum Creek requires a value of 2.5 mm. Generally speaking, furrowed agricultural land captures and retains significantly more water (see Table 3.6) than a natural landscape. The Wabash/Tippecanoe sub-watershed has more land under cultivation than Plum Creek, resulting in a higher capacity for surface depression storage. Despite the higher fraction of the Wabash/Tippecanoe with artificial drainage (see

Section 2.1), the overall effect of the non-artificially drained agricultural land is to allow more surface storage in the watershed than in Plum Creek. For the GW2 coefficient, Wabash/Tippecanoe requires a value of 416.8 hours, while Plum Creek requires a value of 167.7 hours. GW2 represents groundwater flow. The subbasins in the reduced Wabash/Tippecanoe sub-watershed range from 9.6 to 1096.3 km^2, whereas the Plum Creek watershed is a mere 7 km^2. That the variable derived from groundwater persistence is so much higher for the larger watershed is understandable, as a watershed's time of concentration is proportional to its area (Chow 1964).

Table 4.1 Calibrated SMA Parameters, Summer

Summer Model Parameters	Wabash/Tippecanoe Subbasin W520	Plum Creek
Max. Canopy Storage (mm)	1.3	1.5
Max. Surface Storage (mm)	7.3	2.5
Max. Infiltration Rate (mm/hr)	33.7	31.7
% Impervious	4.3	5.0
Soil Storage (mm)	557.5	567.0
Tension Zone Storage (mm)	440.1	433.4
Soil Percolation Rate (mm/hr)	27.4	25.1
GW1 Storage (mm)	19.7	25.2
GW1 Percolation Rate (mm/hr)	27.4	25.1
GW1 Coefficient (hr)	43.9	42.3
GW2 Storage (mm)	203.5	115.3
GW2 Percolation Rate (mm/hr)	1.3	1.3
GW2 Coefficient (hr)	416.8	167.7
GW1 Baseflow Coefficient (hr)	8	100
GW1 Baseflow Reservoirs	5	4
GW2 Baseflow Coefficient (hr)	450	120
GW2 Baseflow Reservoirs	5	2

A big discrepancy is also seen between the Wabash/Tippecanoe and Plum Creek baseflow parameters. The GW1 baseflow reservoirs convey more water than the GW2 reservoirs. Most of the soil water is laterally transferred to the GW1 baseflow reservoirs before it has time to percolate through GW1 storage and into the GW2 storage. As such, the shape of the receding limb produced by the SMA-based model is much more sensitive to the GW1 baseflow parameters than the GW2 baseflow parameters.

The Wabash/Tippecanoe requires a smaller GW1 baseflow coefficient than Plum Creek. A smaller baseflow coefficient results in quicker recession and less attenuation of baseflow, i.e. more baseflow is transferred to streamflow at a quicker rate. The shape of Wabash/Tippecanoe's receding limb is heavily influenced by streamflow upriver of the watershed, whereas the entire length of Plum Creek is contained within the watershed boundary. As such, the Wabash/Tippecanoe GW1 baseflow coefficient primarily serves to generate the appropriate quantity of baseflow. Conversely, the Plum Creek GW1 baseflow coefficient primarily serves to define the shape of the receding limb. During the summer, very little precipitation reaches the baseflow reservoirs due to the high rate of evapotranspiration and high intensity of precipitation; most water is lost before percolating through to the baseflow reservoirs. This increases the difficulty of appropriately calibrating the baseflow parameters, since there are very few baseflow occurrences to use for direction. In the winter, both precipitation intensity and evapotranspiration are much lower, allowing water to reach the baseflow reservoirs and direct the calibration process.

4.2 CN-based and SMA-based Model Performance Comparison

For the Wabash/Tippecanoe sub-watershed, model comparisons for individual validation storms show that the SMA-based model is at least as good as, if not better than, the CN-based model (see Figure 4.1 and Table 4.1). Calibration storm hydrographs are shown in the appendix. For the summer season, the SMA-based model correctly simulates the general shape and magnitude of the hydrograph, but it does not model the specific idiosyncrasies of the flow as well as the CN-based model, despite the fact that both models are run with the same time step. The SMA-based model simulates a fairly smooth hydrograph, whereas the CN-based model produces the same bumps and crevices seen in the observed streamflow. This is true for both the summer and winter seasons. Given that the SMA-based model passes water through multiple storage components before it is transformed into streamflow, it is reasonable to expect the resulting hydrograph to appear more processed. Despite this inability, it is clear from a visual comparison and the values presented in Table 4.1 that the SMA-based model performs better than the CN-based model during the summer. The SMA-based model exhibits better measures for every aspect except the time to peak (t_{peak}) error, where it posts a 4% greater error. However, at a significance level of 0.05, the null hypothesis that the t_{peak} for both models is the same cannot be rejected.

Examining the winter model hydrographs and the objective function results, it cannot be concluded that the SMA-based model performs better than the CN-based model. The CN-based model provides an equally good or better fit for every measure except t_{peak} error,

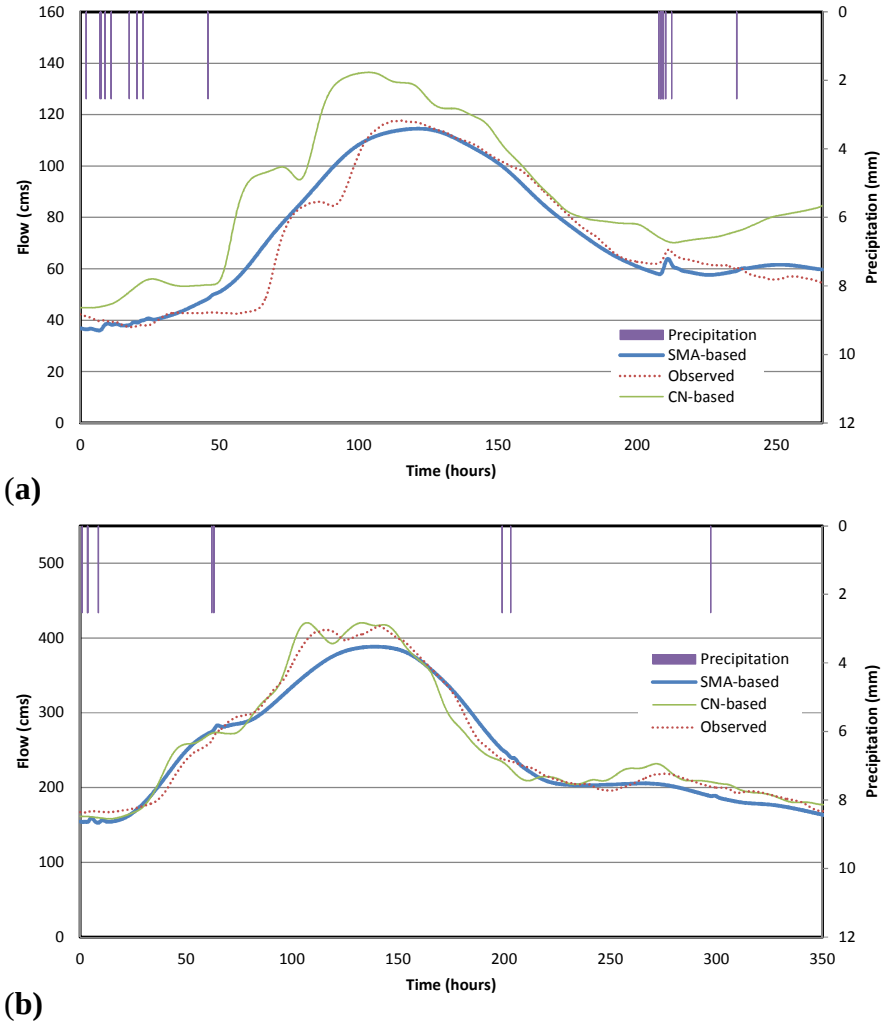


Figure 4.1: Wabash/Tippecanoe Model Comparison, **(a)** summer **(b)** winter

where it exhibits a 23% greater error than the SMA-based model. But, this seemingly glaring error can be essentially ignored due to the fact that the winter CN-based model and observed hydrographs both peak twice. The CN-based model's highest flow occurs on the first peak, and the observed hydrograph's highest flow occurs on the second peak; the difference in the magnitude of the two peaks is negligible. As such, the CN-based t_{peak} error is an artifact of the double peak. Overall, the performance difference of the SMA-

based and CN-based models is not great enough for the winter season to soundly conclude that one performs better.

Table 4.2 Wabash/Tippecanoe Goodness of Fit Parameters

	Summer		Winter	
	CN-based	**SMA-based**	**CN-based**	**SMA-based**
SSE	655,335	62,250	373,331	534,767
NSE	0.38	0.94	0.99	0.99
R^2	0.80	0.94	0.97	0.97
MB	22.37	1.69	0.58	-2.72
NOF	0.28	0.09	0.05	0.06
t_{peak} Error	6%	10%	24%	1%

For the Plum Creek watershed, model comparisons for the summer and winter seasons show that the SMA-based model performs significantly better than the CN-based model (see Figure 4.2 and Table 4.). Despite this, it still does not perform satisfactorily as it significantly underestimates peak flows. As this occurs with both the CN-based and SMA-based models and a suitable solution could not be achieved via calibration, it is most likely caused by shortcomings in the data used to construct and run the models. The discrepancy between the model results and observed data can be attributed to the precipitation data, as the nearest precipitation gauge to the watershed is about 29 km away. One indicator of this is that the highest observed peak flow that occurred during a storm used to calibrate the summer CN-based model occurred after the smallest precipitation event with relatively low precipitation intensity. Another factor could be that Plum Creek is an ephemeral stream. Gan et al. (1997) note that dry catchments are much more difficult to model due to model structure, the use of objective functions

during calibration, and data quality. Plum Creek certainly falls within this category, as streamflow only occurs for a few storms during the summer months. Plum Creek runs dry for much of the summer and portions of the winter months. The issues noted by Gan et al. (1997) may be less of an issue for the CN-based model, particularly during the winter season.

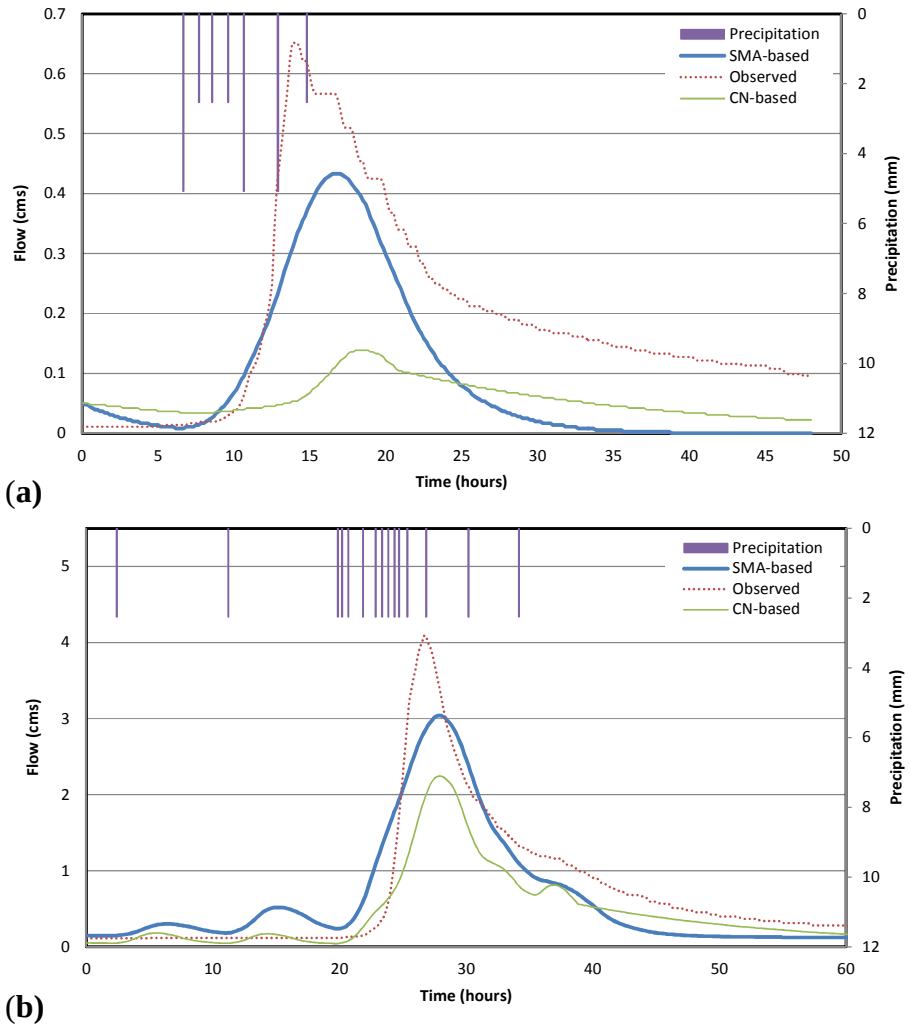


Figure 4.2: Plum Creek Model Comparison, **(a)** summer, **(b)** winter

Table 4. shows that the SMA-based model performs better than or the same as the CN-based model in every instance except the coefficient of determination for the winter season. The discrepancy between the two R^2 values is barely significant at only 0.05. The coefficient of determination is indicative of how well the model explains variance in the observed dataset. The CN-based model has a higher R^2 value because it is better able to model the shape of the receding limb of the winter hydrograph. This is primarily due to the manner in which the CN-based model simulates baseflow in this study—via the commonly used recession method. With this method, the shape of the receding limb is extremely sensitive to baseflow parameters in the CN-based model; this makes calibrating for recession behavior fairly easy. Conversely, the SMA method requires the use of the linear reservoir method. With this method, water in the SMA-based model must percolate through groundwater storage and baseflow reservoirs before appearing in the stream. The complicated nature of this process reduces the ability to define the shape of the recession curve via calibration. While the CN-based model captures some aspects of the Plum Creek watershed's behavior, the SMA-based model performs much better.

Table 4.3 Plum Creek Goodness of Fit Parameters

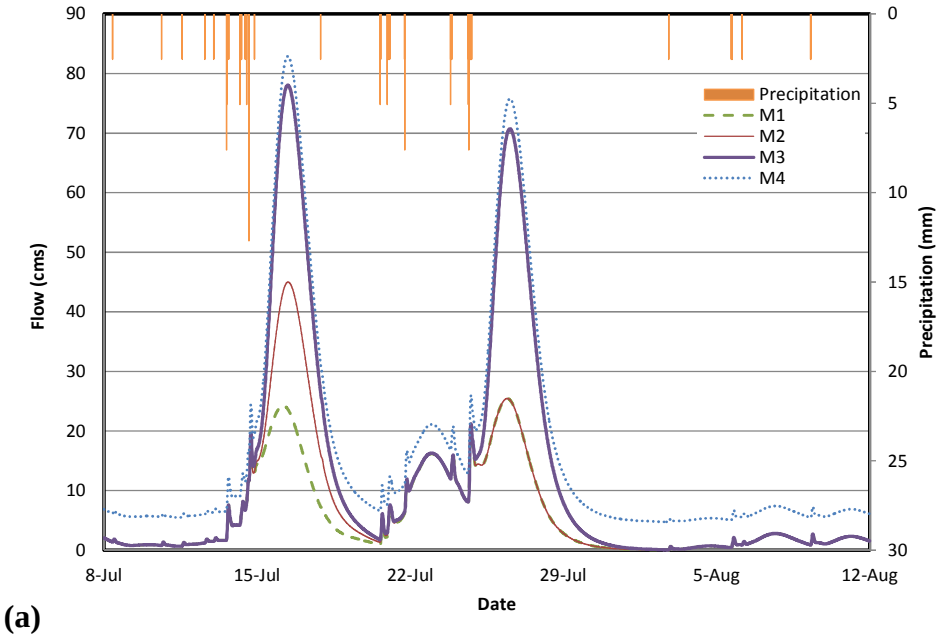
	Summer		Winter	
	CN-based	**SMA-based**	**CN-based**	**SMA-based**
SSE	4,108	5,069	119,659	67,791
NSE	-0.42	0.49	0.73	0.86
R^2	0.49	0.80	0.92	0.87
MB	-69.81	-47.97	-43.14	-15.23
NOF	1.04	0.62	0.80	0.58
t_{peak} Error	31%	19%	4%	4%

4.3 Downward Model Development Results

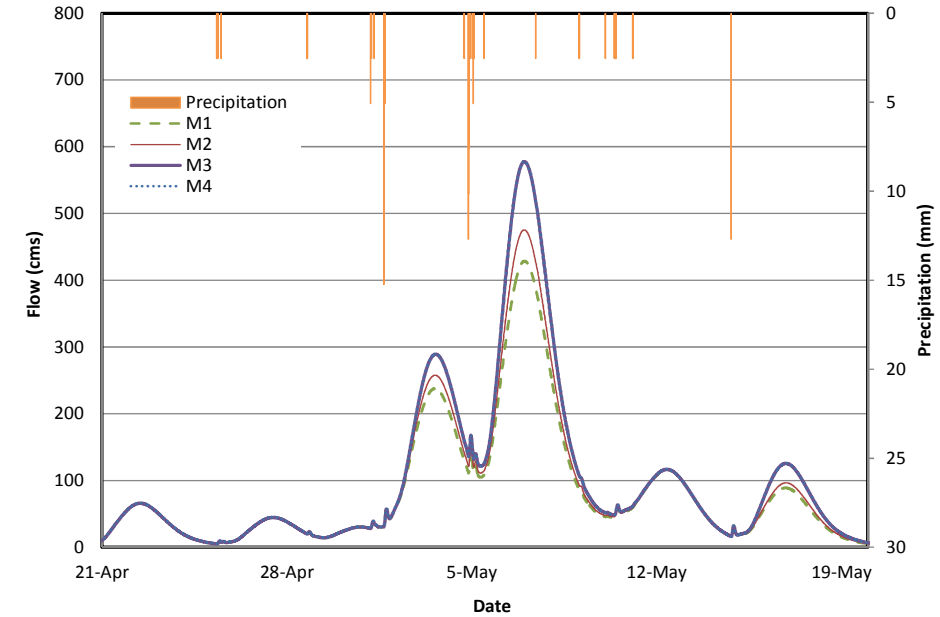
A comparison of the streamflow hydrographs from the downward developed models provides some insight into the influence of specific model parameters. Figure 4.3 and Figure 4.4 give a representative comparison of streamflow from the four SMA-based models for both the Wabash/Tippecanoe sub-watershed and Plum Creek watershed. M1-M3 essentially collapse to the same streamflow when the precipitation intensity is low. M4, which includes baseflow, results in greater streamflow than the other three models. At times, this is difficult to determine visually, but it is verifiable via an examination of the model outputs. Observed streamflow is omitted from Figure 4.3 and Figure 4.4, because streamflow is not available from the USGS website for the simulation time period, 1994-2003. Also, Figure 4.3 depicts flow from the isolated sub-watershed; USGS data includes streamflow from regions outside of the study area.

When the precipitation intensity is low, most rainfall immediately infiltrates into the soil; the surface characteristics then become the most influential factor in determining streamflow, which is why little to no difference in M1-3 is evident at low precipitation intensities. M1 and M2 contain infinite soil storage; whereas M3 and M4 limit soil storage (see Table 3.8 and Figure 1.1). The signature of this is evident in the significant increase in streamflow between M2 and M3. The precipitation infiltration rate is proportional to the amount of available storage in the soil profile (Chow 1964). With infinite storage, the infiltration rate is maximized, considerably reducing runoff potential.

This is why M1 and M2 produce less streamflow than the soil-storage-limiting M3 and M4.



(a)

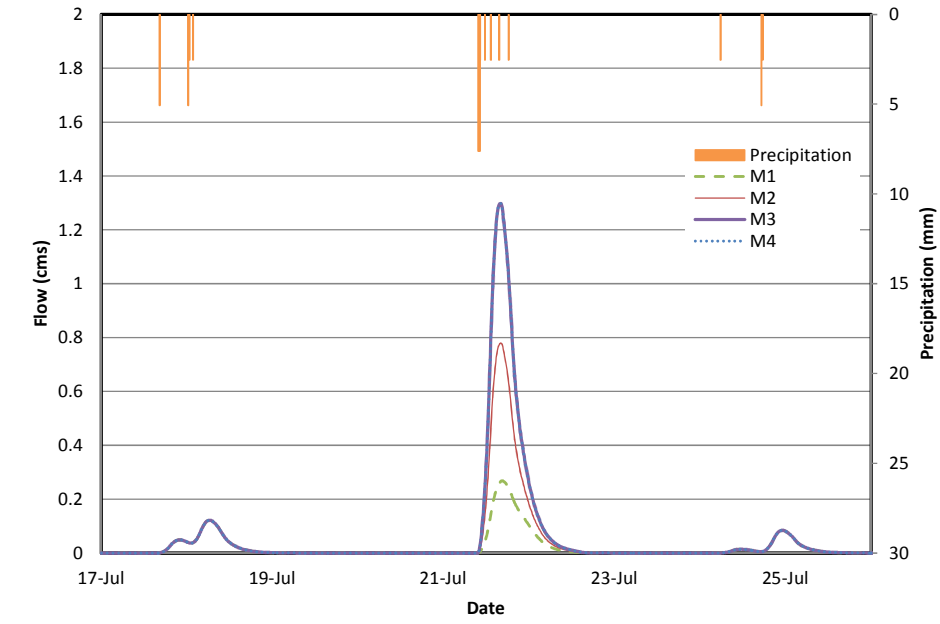


(b)

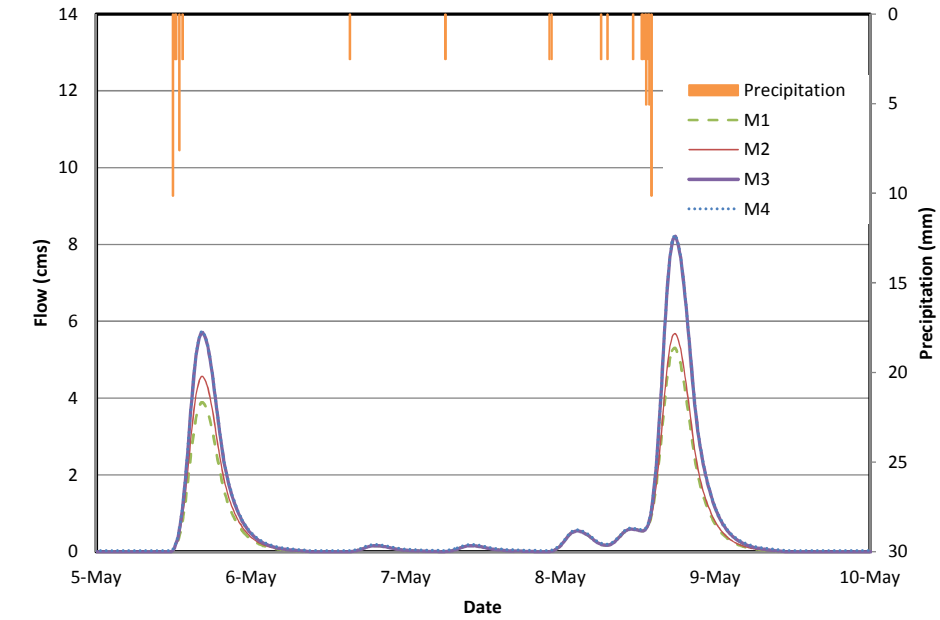
Figure 4.3: Streamflow Hydrograph for Four Wabash/Tippecanoe Models, **(a)** summer: July 8-August 12, 2000, **(b)** winter: April 21-May 19, 2003

The summer season in the Wabash/Tippecanoe displays a significant difference between M3 and M4 not seen in the winter season (see Figure 4.3). This is due to the slow decay of baseflow during the summer season, not because the summer season produces more baseflow than the winter. In reality, the summer model rarely produces baseflow, as the majority of soil water is evaporated before having time to percolate down through the groundwater layers to the baseflow reservoirs. With an evapotranspiration rate nearly a tenth of the summer value, the winter produces baseflow quite frequently. It is not evident in Figure 4.3b, because the magnitude of baseflow is small compared to the peak streamflow.

As with the Wabash/Tippecanoe sub-watershed, the Plum Creek watershed also produces more baseflow in the winter, even though it is not evident in Figure 4.4. In Plum Creek, the relative difference between M1 and M2 is much greater during the summer season than the winter season. To understand this cause, it is first important to note that M1 and M2 contain the same amount of soil storage; the only difference is the manner of storage. In M1, both evapotranspiration and percolation occur from the entire soil profile since it is all modeled as upper zone storage in the downward development configuration. In M2, evapotranspiration occurs throughout, but percolation only occurs from the upper zone storage (see Table 3.8 and Figure 1.1). So, the soil profile is likely to maintain a higher degree of saturation in M2 than in M1.



(a)



(b)

Figure 4.4: Streamflow Hydrograph for Four Plum Creek Models, **(a)** summer: July 17-July 26, 1996, **(b)** winter: May 5-May 10, 1996

Evapotranspiration first occurs from the upper zone, but precipitation fills the tension zone first (HEC 2000). In the summer, the evapotranspiration rate is high. When evapotranspiration occurs from the tension zone, the rate of evapotranspiration is reduced relative to the ratio of current soil storage to tension zone capacity (HEC 2000). In M1, all soil storage is upper zone storage, so the evapotranspiration rate is always at its maximum. As such, there is less water stored in the soil to then influence the infiltration rate and consequently streamflow. In summary, the tension zone serves to reduce the rate of evapotranspiration, and therefore it increases the ability of the soil profile to retain water and increases the potential for surface runoff due to lowered rates of infiltration.

4.3.1 Precipitation Intensity and the Soil Profile

The results of the sign test for soil profile significance at different maximum precipitation intensities in the Wabash/Tippecanoe sub-watershed are shown in Table 4.. The p-values are displayed for the 10-year simulation period. The application of a Bonferroni correction results in significance at p-values less than 0.0083. Note that M2 generates peak flows that are significantly different from M1 at precipitation intensities of 1.5 and 3.0 cm/hour, but then are insignificant until a precipitation intensity of 7.6 cm/hour is reached. At low precipitation intensities, the infiltration rate has a high probability of being equal to the precipitation intensity. As such, most of the precipitation enters the soil profile, and this allows the soil profile to play a significant role in determining streamflow. This also suggests the point at which tension zone storage begins to impact streamflow is at a maximum precipitation intensity of 7.6 cm/hour when there is a high

probability that much of the precipitation becomes runoff rather than infiltrating into the soil. Among the 17 sub-basins in the Wabash/Tippecanoe, the average maximum infiltration rate is 4.4 cm/hr. M3 peak streamflow only becomes significantly different from M2 when a maximum precipitation intensity of 6.1 cm/hour is reached. Since M2 contains unlimited soil storage and M3 contains limited soil storage, this indicates the availability of soil profile storage begins to have a sizable impact on streamflow at a precipitation intensity of 6.1 cm/hour. Furthermore, M4 peak streamflow is always significantly different from peak streamflow in M3. This can be explained as the influence of baseflow.

Table 4.4 Peak Streamflow Significance for Wabash/Tippecanoe Model Comparison

Max. Rainfall Intensity	M1 vs. M2	M2 vs. M3	M3 vs. M4
1.5 cm/hour	S (0.0003)	NS (0.1435)	S (<0.0001)
3.0 cm/hour	S (0.001)	NS (0.2668)	S (<0.0001)
4.6 cm/hour	NS (0.0654)	NS (0.0215)	S (<0.0001)
6.1 cm/hour	NS (0.0215)	S (0.0001)	S (0.0001)
7.6 cm/hour	S (0.001)	S (0.001)	S (0.001)

Key: S- significant, NS- not significant; p-value shown is for 10-year simulation period (1994-2003).

The results of the sign test for soil profile significance in the Plum Creek watershed are shown in Table 4.. The p-values are again displayed for the 10-year simulation period, showing significance at p-values less than 0.0083. Note that M2 becomes significantly different from M1 at a precipitation intensity of 4.6 cm/hour; this suggests the point at which tension zone storage begins to impact streamflow in the Plum Creek watershed is 4.6 cm/hour. M3 peak streamflow is significantly different from M2 at the minimum

precipitation intensity, 1.5 cm/hour, and the same is true for the difference between M3 and M4. Since M2 contains unlimited soil storage and M3 contains limited soil storage, this indicates that the availability of soil profile storage always has a sizable impact on streamflow for this watershed. The difference in peak streamflow for M3 and M4 can again be explained as the influence of baseflow.

Table 4.5 Peak Streamflow Significance for Plum Creek Model Comparison

Max. Rainfall Intensity	M1 vs. M2	M2 vs. M3	M3 vs. M4
1.5 cm/hour	NS (0.125)	S (0.0066)	S (<0.0001)
3.0 cm/hour	NS (0.125)	S (<0.0001)	S (<0.0001)
4.6 cm/hour	S (<0.0001)	S (<0.0001)	S (<0.0001)
6.1 cm/hour	S (<0.0001)	S (<0.0001)	S (<0.0001)
7.6 cm/hour	S (0.0020)	S (0.0020)	S (0.0039)

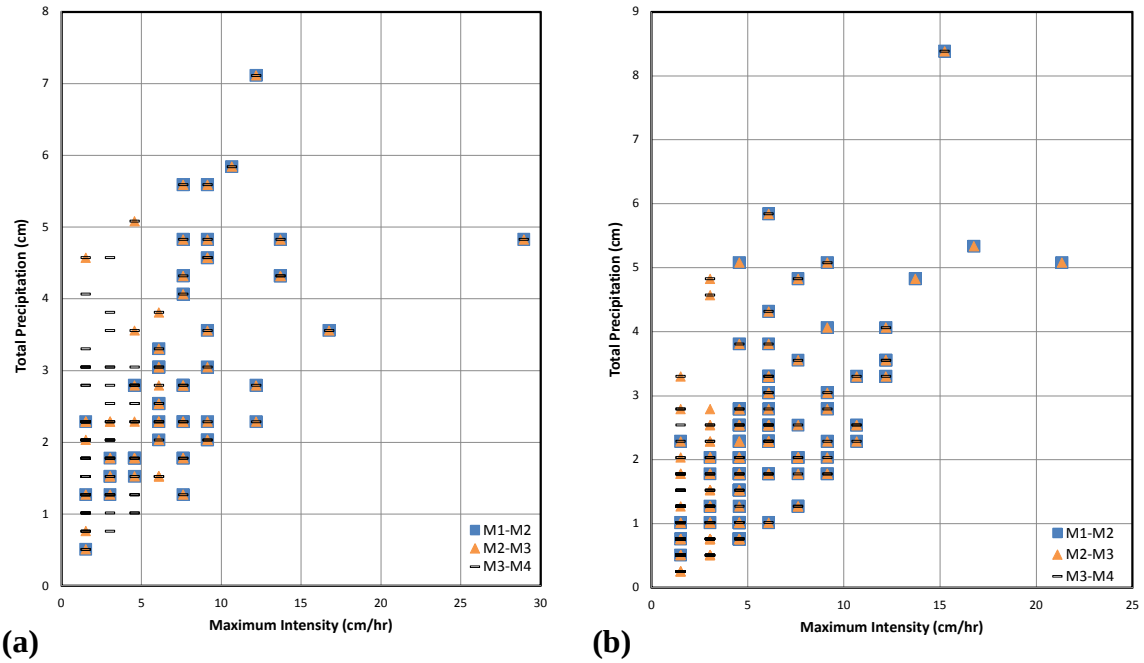
Key: S- significant, NS- not significant; p-value shown is for 10-year simulation period (1994-2003).

A visualization of the occurrence of differences between the models in relation to total storm precipitation and maximum precipitation intensity provides additional insight into the behavior of the soil profile (see Figure 4.5). The dataset shown in the figure is the aggregate of the peak flows tested with the sign test. Remember that these are all of the peak flows above the 90^{th} and 95^{th} flow percentile for the Wabash/Tippecanoe and Plum Creek, respectively, which occurred during the ten-year simulation period. The figure shows total storm precipitation on the y-axis and maximum precipitation intensity on the x-axis. For every combination of total storm precipitation and maximum precipitation intensity that occurred during the ten-year simulation period, a marker indicates that two models produced different peak flows. There are three types of markers displayed: a blue

square for a difference in peak flow between M1 and M2, denoted M1-M2, an orange triangle for a difference between M2 and M3, denoted M2-M3, and a black rectangle for a difference between M3 and M4, denoted M3-M4. A marker appears regardless of the number of times a difference was detected or of its statistical significance. The total storm precipitation and maximum precipitation intensity appear as discrete values, because precipitation data is reported by NCDC at 0.254 cm intervals.

Although the difference between M1 and M2 in the Wabash/Tippecanoe sub-watershed displayed significance at precipitation intensities of 1.5 and 3.0 cm/hour (see Table 4.), Figure 4.5a shows that this difference only occurs for six out of 26 total precipitation/maximum intensity combinations for both intensity levels. This indicates that the differences seen can be attributed more to antecedent soil moisture conditions than to the particular influence of tension zone storage. An interesting phenomenon is seen in the Plum Creek watershed but not in the Wabash/Tippecanoe. At extremely high maximum precipitation intensities, Plum Creek ceases to produce a difference in peak flows between M3 and M4, while still producing differences in M1, M2, and M3 (see Figure 4.5b). For high precipitation intensities, this indicates that while some precipitation is infiltrated, the majority of precipitation is transferred to the river via Hortonian overland flow. As a result, baseflow influence of streamflow is small compared to the stormflow response. This effect is seen in Plum Creek but not in the larger Wabash/Tippecanoe due to the great differences in time of concentration between the two watersheds. The time of concentration in the Wabash/Tippecanoe is so long that some of the water reaching the baseflow reservoirs has enough time to be transferred to

the river network before the peak flow is achieved. Plum Creek's significantly shorter time of concentration disallows this effect.



Key: M1-M2: difference seen between M1 and M2 peak flows, etc.

Figure 4.5: Occurence of Peak Flow Differences between Four Models, **(a)** Wabash/ Tippecanoe, **(b)** Plum Creek

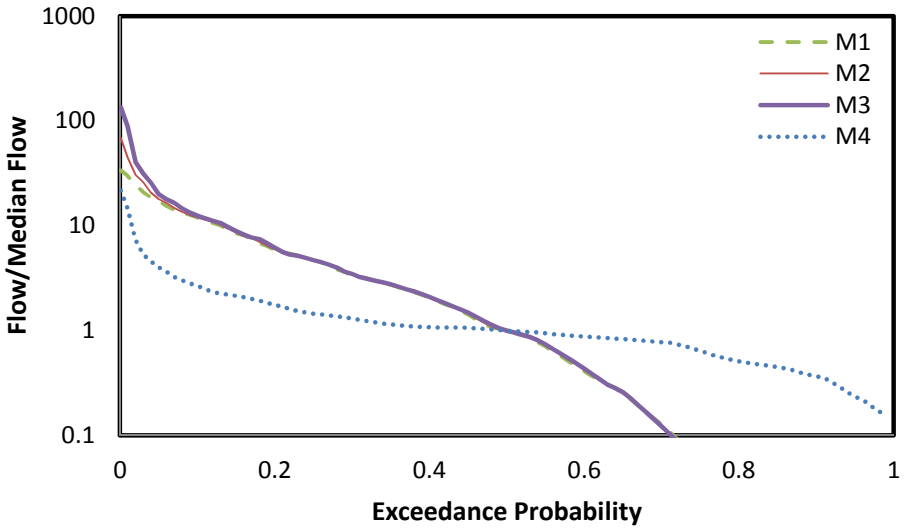
The application of downward model development to the Wabash/Tippecanoe and Plum Creek watersheds provides the ability to identify the specific impact of various soil parameters on streamflow. The tension zone, or field capacity, of the soil profile only truly influences streamflow in the Wabash/Tippecanoe sub-watershed at unusually high precipitation intensities (7.6 cm/hour) while for Plum Creek, this value is 4.6 cm/hour. Percolation to lower soil storage levels cannot occur from tension zone storage. Thus, the tension zone only affects streamflow when it is full, or full enough to significantly reduce

the maximum infiltration rate. The depth of the soil profile impacts streamflow much more than tension zone storage. The impact begins at a lower precipitation intensity, 6.1 cm/hour for Wabash/Tippecanoe and 1.5 cm/hour for Plum Creek, and results in the greatest magnitude of change among all the soil profile parameters explored. While baseflow impacts streamflow for all precipitation intensities examined, the impact on the magnitude of peak flow is relatively minor. Its impact is so small that it cannot often be detected on the streamflow hydrographs (see Figure 4.3b-Figure 4.4).

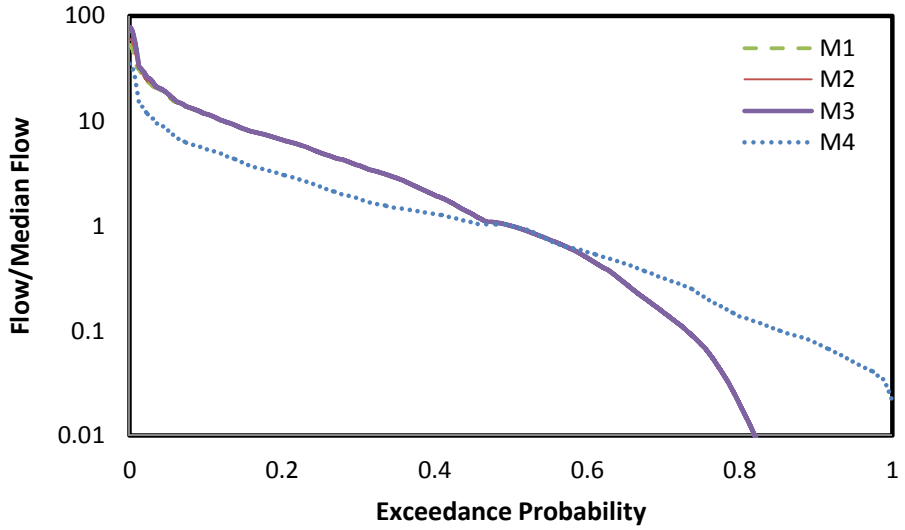
4.3.2 Flow Duration Curves

Flow duration curves for the four SMA-based Wabash/Tippecanoe models are presented in Figure 4.6. There is little difference in the shape of the curves for M1-3, but M4 produces less extreme flow values when compared to the median flow. This indicates that a fully developed soil profile dampens the effect of both high-intensity precipitation and low streamflow. The flow duration curves for M1-3 also exhibit a steeper slope than that of M4, underscoring the inability of these models to fully capture the streamflow recession behavior of the watershed (Farmer et al. 2003). The absence of baseflow reservoirs in M1-3 limits the ability of the models to convert infiltrated precipitation into streamflow. The flat slope of the M4 flow duration curves also suggests that M4 has a greater capacity to store water than the other three models (Gupta 2008). In effect, the baseflow parameters included in the complete model, M4, are necessary to appropriately transform precipitation and to fully capture the storage capabilities of the watershed. An observed flow duration curve is not shown due streamflow data limitations for the

simulation period. However, the flow duration curve from M4 is expected to closely mimic the observed flow duration curve; the full model (M4) is calibrated to historic streamflow data.



(a)



(b)

Figure 4.6: Flow Duration Curve for Wabash/Tippecanoe Model **(a)** summer, **(b)** winter

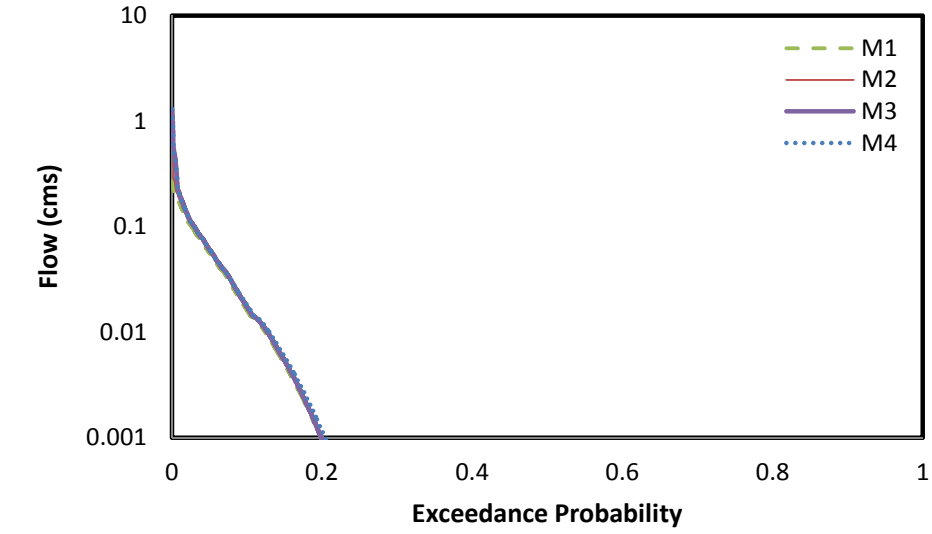
The flow duration curve for the summer season M4 is much flatter than the winter season M4. As explored earlier in this section, this can be attributed to the fact that baseflow is transferred to streamflow much quicker during the winter months, resulting in a flashier stream system.

Flow duration curves for the four SMA-based Plum Creek models are presented in Figure 4.7. As is also seen in Farmer et al. (2003), the flow duration curves for the less sophisticated models, M1-3, fail to capture the flow persistence of the river system. This, coupled with the fact that Plum Creek flows intermittently, explains why the curves do not cover the spectrum of exceedance probability. Compared to the Wabash/Tippecanoe curves, Plum Creek produces extremely steep flow duration curves; Plum Creek watershed is much flashier and has less storage capacity than the larger Wabash/Tippecanoe sub-watershed. This is expected, as Plum Creek contains a first order stream and Wabash/Tippecanoe contains a high-order stream. First order streams have a much smaller area contributing to streamflow, which results in a much shorter time to peak and a lower ability to generate baseflow. For low-order watersheds, surface and interflow processes play a dominant role in generating streamflow. For high-order watersheds, the large contributing area results in baseflow processes generating a greater portion of streamflow than in low-order watersheds.

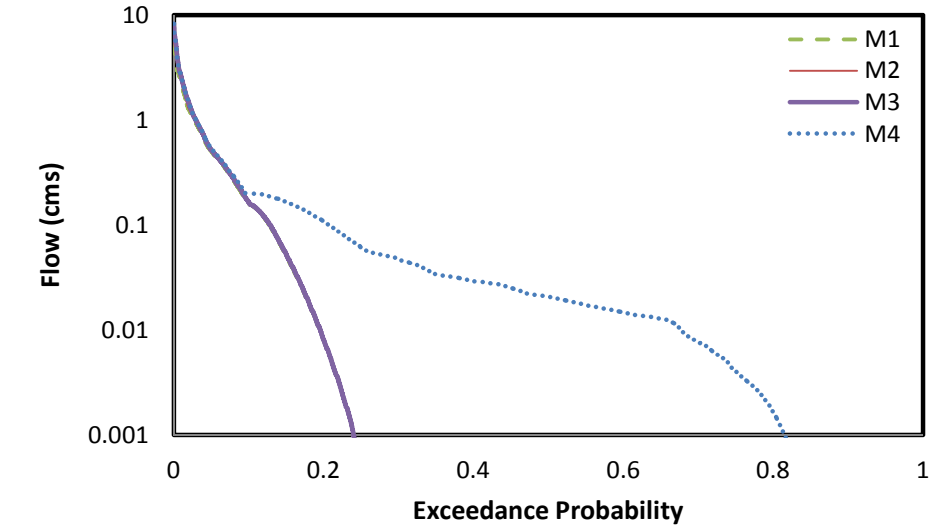
The Plum Creek watershed is so flashy, that the median streamflow for M1-3 for the summer and winter seasons is zero, as evidenced by the fact that the curves disappear at an exceedance probability of about 0.2. In fact, M4 shows that streamflow is less than

0.001 m^3/s for approximately 80% of summer and 20% of the winter. For the summer season (see Figure 4.7a), there is not a visible difference between the flow duration curves for M1-4. This indicates that baseflow does not play a substantial role in Plum Creek, as it does in the Wabash/Tippecanoe sub-watershed. This is due in part to the smaller size of the watershed and dominance of overland flow and interflow processes and in part to the higher rate rater of evapotranspiration during the summer season.

For the winter season, M4 does have a median flow greater than zero, and it also exhibits the significant influence of baseflow. Including baseflow in the Plum Creek model for the winter season shows that baseflow helps to reduce the simulated flashiness of the watershed system. In an ephemeral stream, such as Plum Creek, baseflow is vital to the maintenance of aquatic habitats in the streambed, because, as suggested by Figure 4.7b, it provides water to the stream long after a precipitation event has passed.



(a)



(b)

Figure 4.7: Flow Duration Curve for Plum Creek Model **(a)** summer, **(b)** winter

4.3.3 Flood Frequency Analysis

Figure 4.8 and Figure 4.9 display the peak flows and total runoff depths for various return period storm events for the summer and winter seasons for both watersheds. The difference between M1 and M2 peak flows and again between M3 and M4 peak flows is

negligible. This indicates that for flood events, neither tension zone storage nor baseflow significantly impacts peak flows. Rather, the most important factor in flow magnitude is soil profile storage capacity, as is also seen in Figure 4.3 and Figure 4.4.

For both watersheds, there is a negligible difference between the M1 and M2 runoff depths. Not only do M1 and M2 have essentially the same peak flows, they also have identical cumulative runoff depth over the length of the storm event. This suggests that the tension zone does not significantly impact recession behavior. Conversely, despite having the same peak flows, M3 and M4 display a substantial difference in runoff depths, except for the summer Plum Creek model. This implies that baseflow significantly influences recession behavior in a watershed. This finding agrees with the basic definition of baseflow (Gupta 2008). The difference between runoff depths for M2 and M3 is expected, as the peaks flows are also different. The reason for this difference is again, the result of limiting soil storage capacity.

During the summer season, the relative magnitude of peak flows and runoff depths is significantly smaller at short return periods than during the winter season. This is evidenced by the steep curve extending far towards the origin of each summer season plot (see Figure 4.8a,c and Figure 4.9a,c). Two explanations for this behavior are probable. First, the higher rate of evapotranspiration in the summer means the soil profile is emptier and can thus store more water, reducing runoff. Second, during the summer season, the impervious surface percentage is lower than the winter, because the ground is not frozen. As a result, the soil profile can capture and store more of the precipitation in the summer

than in the winter. This effect is reduced when the precipitation intensity is higher than the maximum infiltration rate, as it often is with long-return period storms, causing excess rainfall to immediately become Hortonian overland flow.

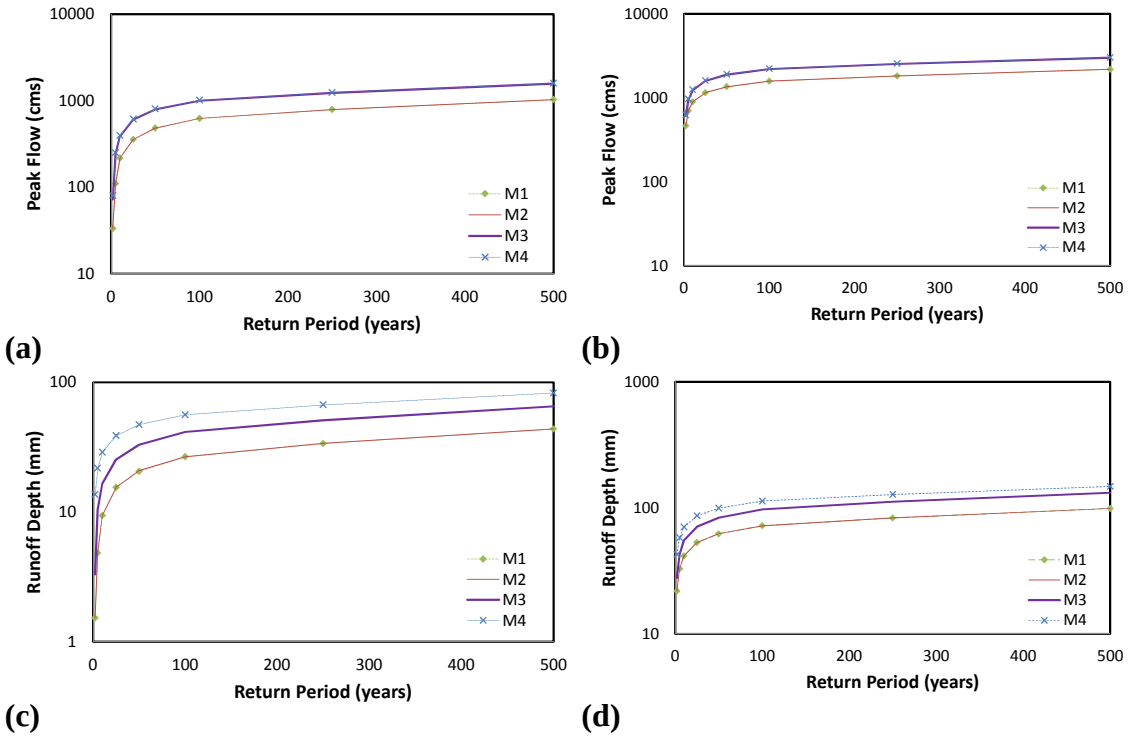


Figure 4.8: Frequency Analysis for Wabash/Tippecanoe Models for **(a)** summer model peak flows, **(b)** winter model peak flows, **(c)** summer model runoff depth, **(d)** winter model runoff depth

The smooth transitions between peak flow values on the frequency curves, as seen in Figure 4.8 and Figure 4.9, indicate that the tension zone in the watersheds contains enough storage that only one flow mechanism dominates: surface runoff. Kusumastuti et al. (2006) notes that a jagged jump in peak flow values between return periods marks a change in the flow mechanism from subsurface flow to saturation-excess surface flow.

However, the opposite would be true in this study. In Kusumastuti et al.'s study of hypothetical watersheds in Australia, the tension zone storage depth explored is 45 mm and ranges from 11 to 45% of the total soil profile storage. In this study, the tension zone storage depth ranges from 275 to 470 mm or 59 to 79% of the total soil profile storage. The tension zone is so much greater in this study, because the soils in Indiana are significantly deeper than those found in Australia. Also, the field capacity of Indiana soils is much greater than soils in Australia. The deep tension zone in this study provides enough storage that little, if any, precipitation reaches the upper zone storage from which percolation to the groundwater layers occur, and subsurface flow is generated. As such, flow is primarily generated through surface runoff either due to impervious surface cover or precipitation intensity in excess of the maximum infiltration rate. In summary, a deep tension zone provides enough storage that it is unlikely that subsurface flow will influence peak flows during high precipitation intensity flood events.

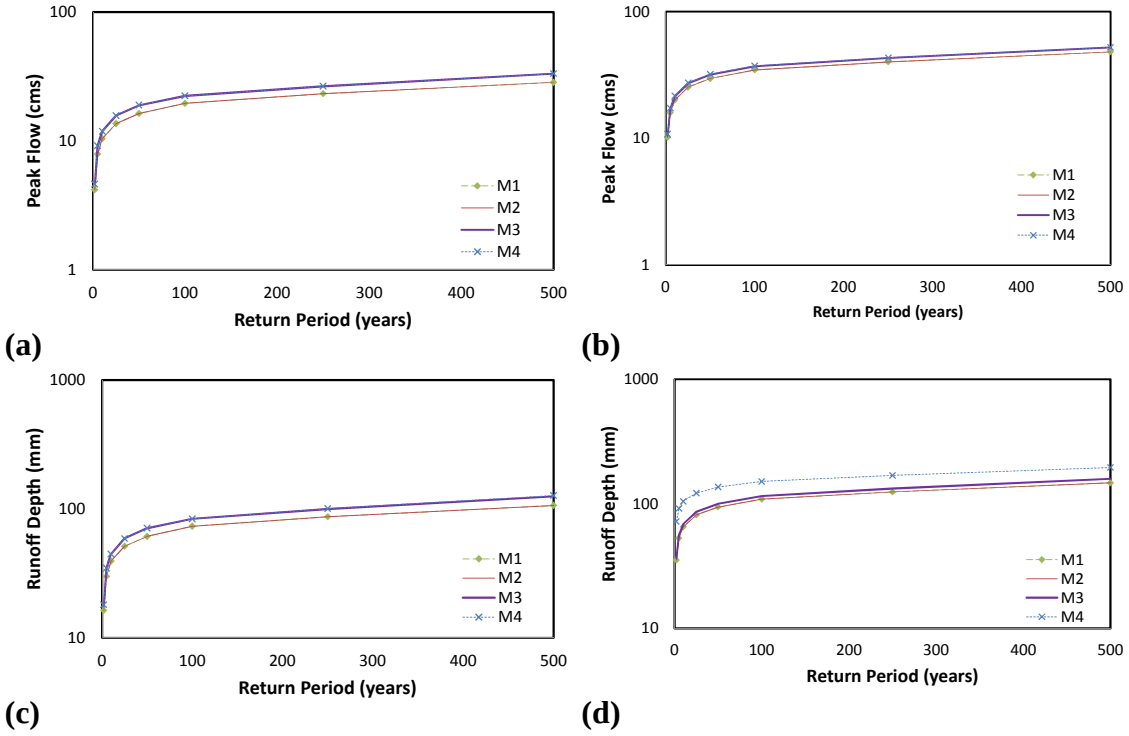


Figure 4.9: Frequency Analysis for Plum Creek Models for **(a)** summer model peak flows, **(b)** winter model peak flows, **(c)** summer model runoff depth, **(d)** winter model runoff depth

4.4 Study Limitations

While every attempt is been made to accurately represent the hydrology of the Wabash/Tippecanoe sub-watershed and Plum Creek watershed, this study includes limitations. For future studies, the SMA-based model parameters should be derived with additional attention to the limitations presented here.

For example, vertical flow of water through soil is generally determined by the minimum saturated hydraulic conductivity (k_{sat}) value of the soil horizons through which it passes,

not the average used in Figure 3.3. Since this study uses an average k_{sat}, the rate of percolation through the upper zone storage and GW1 is higher than it should be. As such, water is lost from these storage components much quicker than is expected to occur in nature. The primary impact of this is that water may be lost from the interflow component (GW1) before it has the chance to be transferred to the baseflow reservoirs and be transformed into streamflow.

While vertical flow through soil is determined by the minimum k_{sat}, the infiltration rate is limited by the k_{sat} of the first soil horizon. The actual infiltration rate has the potential to be much greater than the k_{sat} of the first soil horizon. Since the k_{sat} in this study, is used as the maximum, not the minimum, infiltration rate, more precipitation is likely to become surface runoff. This decreases the ability of the soil profile to impact streamflow.

In this study, the porosity of a soil series is simply taken as the average porosity of the soil horizons. Realistically, a weighted average porosity based on the depth of each soil horizon should be used. The impact of this depends on the relative soil porosity and depths of the soil horizons. In this study, it has the potential to either increase or decreases the upper zone storage. An increase in upper zone storage would decrease the generation of interflow, since more water could be stored in the soil profile, and vice versa.

This study is greatly limited by the challenges presented by modelling artificial drainage in a substitutive manner. In this study, artificial drainage is modeled by decreasing

surface depression storage, which results in more surface runoff. In reality, artificial drainage removes water from the soil profile storage and transfers it underground to the stream network. This behavior most closely mimics interflow. As such, artificial drainage should be modeled using the GW1 storage and GW1 baseflow reservoirs in the SMA-based model in HEC-HMS. Since artificial drainage occurs via a perforated pipe network, the GW1 storage and baseflow parameters would need to be adjusted such that the transfer of water to streamflow is relatively quick compared to traditional baseflow. In this study, the manner in which artificial drainage is modeled could have a significant impact on the study results. Primarily, it may prevent the tension zone from reaching saturation, which also inhibits the production of subsurface flow. This may explain why the fully-developed SMA-based model fails to fully capture the recession characteristics of the watersheds. The particular influence of decreasing surface storage is likely dependent upon the rate of evapotranspiration, so further investigation would be required to clarify the impacts of modelling artificial drainage in this manner.

CHAPTER 5. SUMMARY AND CONCLUSIONS

The application of downward model development to the SMA-based loss method in HEC-HMS is explored in this study. As a starting point, CN-based and SMA-based models are developed and their performance compared. Overall, the SMA-based models performed as well as or better than the CN-based models for specific storm events. However, the performance of the SMA-based model may vary when compared to that of a CN-based model during continuous simulation of dry periods. It is expected that the SMA-based model would perform significantly better since it continuously adjusts soil moisture conditions. Interactions among specific soil profile processes can be related to model outputs in the SMA-based model because of its fully developed soil profile. This is not possible with the CN-based model, since the soil profile is greatly simplified. The downward analysis shows that individual soil profile processes do significantly impact streamflow.

Streamflow hydrographs from the four downwardly developed models showed significant differences in prediction of peak flows among the four models for storms with high precipitation intensities. This indicates that various components of the soil profile only begin to play a role in generating streamflow after a threshold precipitation intensity is reached. This threshold exists, because the primary method for simulating streamflow after a storm event is via surface runoff in this study. An

alternative study that models artificial drainage using interflow mechanisms may have different findings. The volume of surface runoff is directly dependent upon the rate at which precipitation infiltrates into the soil and on the soil's ability to retain water. Characteristics of the soil profile, such as storage depth, percolation rate, evapotranspiration rate, and groundwater storage, play a role in defining infiltration rate and water retention capabilities.

Flow duration curves from the four downwardly developed models showed that a complete soil profile is required to properly define the flow persistence characteristics of streamflow. Baseflow mechanisms allow precipitation to maintain streamflow long after the storm has ceased. The presence of baseflow also increases the storage capacity of a watershed. The flow duration curves confirm that watersheds with ephemeral streams are more variable, especially in the summer when monthly evapotranspiration is close to or exceeds monthly precipitation.

Flood frequency curves of peak flow and runoff depth from the four downwardly developed models showed that a deep tension zone prevents a significant portion of the precipitation from reaching the upper zone storage and generating subsurface flow. As such, the dominant flow mechanism is surface runoff for return periods up to 500 years. Also, tension zone storage does not influence the recession behavior of a watershed, but the total soil profile storage does.

In general, soil profile storage does impact streamflow, but it only becomes consistently significant after a threshold precipitation intensity is reached. This threshold value will vary based on the characteristics of individual watersheds, such as size, land cover and climate. For a large agricultural watershed in central Indiana, this value is about 6.1 cm/hour. For a small agricultural watershed, there is not a defined threshold; soil profile storage is shown to influence streamflow at every level of precipitation intensity. But, baseflow stops influencing streamflow in the small watershed when storms involve extremely high precipitation intensities; this is due to a short time of concentration. The total storage capacity of the soil profile is the most important factor in accurately determining the magnitude of peak streamflow. Limiting the storage capacity of the soil profile results in a sizable increase in streamflow. In shallower soils, the soil profile has a greater ability to influence streamflow, because the actual infiltration rate is inversely proportional to the soil profile depth. When possible, hydrologic models should always include soil profile parameters, as they are known to affect streamflow.

The rate of evapotranspiration is almost as important as soil profile storage for determining streamflow. During the summer months, significantly less streamflow occurs, despite higher precipitation intensity and nearly equivalent total precipitation in central Indiana. This is largely attributed to the higher rate of evapotranspiration in the summer. Since evapotranspiration first occurs from the tension zone, it cannot retain water in the summer. This results in higher rates of infiltration and therefore

less surface runoff. While it is important to understand the soil profile, the effects of evapotranspiration cannot be ignored either.

While SMA-based models perform better than CN-based models, they are not always the best choice for hydrologic modeling. SMA-based models require a lot more time and data to develop than CN-based models. The benefit from these additional efforts is not always warranted by the project, especially since the CN-based model performs equally well at times. The CN-based model certainly captures the recession behavior of the stream network better than the SMA-based models. If the project requires great accuracy in this aspect, CN-based modeling is undeniably the better choice. For the large watershed, the CN-based, event model performed almost as well as the SMA-based, continuous model. This is largely attributed to the fact that large watersheds are slower to respond to precipitation events and are therefore less flashy. As such, the initial conditions set for the watershed do not influence the model results as much, because the model has some time to equilibrate before the streamflow peaks. Large watersheds with long memory may be sufficiently modeled using CN-based models for peak flow prediction. However, SMA-based models should be used if the objective is to accurately predict streamflow long after the storm occurs, because the fully-developed soil profile allows the model to accurately simulate the memory of the watershed. With smaller, flashier watersheds, the initial conditions greatly influence peak streamflow. By using a respectable spin-up period, the SMA-based model essentially determines the watershed's initial conditions itself. As such, SMA-based models can be a useful tool when modeling flashy, small watersheds.

Based on the results of this study, the following recommendations are made regarding the implementation of a fully developed soil profile:

1. The SMA-based model development methodology discussed in this study is effective, but special attention should be paid to soil profile parameters in regions with artificial drainage.
2. Regions experiencing frequent high-intensity precipitation events should always opt to create hydrologic models with a fully developed soil profile, as it significantly influences streamflow under these conditions. The watershed used in this study has a precipitation intensity threshold of 6.1 cm/hour, but more studies need to be done to see if this threshold can be generalized for large watersheds.
3. Generally, small, flashy watersheds should be modeled with an SMA-based method, since the antecedent moisture condition greatly affects streamflow generation

Further work should be performed to explore the relationships between the soil profile and streamflow in watersheds of various compositions and sizes. While hydrologists have developed a reasonably accurate description of soil profile processes, downward model development can help improve this understanding at the watershed scale or even greater. Given the perceived climatic shifts, it is also important to begin considering how changes to the soil profile may impact streamflow in future precipitation events. In addition, deeper studies should be

undertaken concerning the impact of artificial drainage on the hydrology of a watershed. This could help identify the most accurate method for incorporating the influence of artificial drainage into hydrologic models, as it is a major component of many agricultural watersheds.

LIST OF REFERENCES

LIST OF REFERENCES

Abushandi, E. and Merkel, B. (2013). “Modelling rainfall runoff relations using HEC-HMS and IHACRES for a single rain event in and arid region of Jordan.” *Water Resour. Manage.*, 27, 2391-2409.

Beighley, R.E., Dunne, T. and Melack, J.M. (2005). “Understanding and modeling basin hydrology: Interpreting the hydrogeological signature.” *Hydrol. Processes*, 19, 1333–1353.

Bennett, T. (1998). "Development and Application of a Continuous Soil Moisture Accounting Algorithm for the Hydrologic Engineering Center Hydrologic Modeling System (HEC-HMS)." MS thesis, Dept. of Civil and Environmental Engineering, University of California, Davis, Davis, CA.

Berthet, L., Andréassian, V., Perrin, C., and Javelle, P. (2009). “How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments.” *Hydrol. Earth Syst. Sci.*, 13, 819-831.

Beven K.J., Lamb, R., Quinn, P.F., Romanowicz, R., and Freer, J. (1995). TOPMODEL. in *Computer Models of Watershed Hydrology*, SinghVP *(ed.). Water Resources Publications:* 627–668.

Brocca, L., Melone, F., Moramarco, T. (2008). “On the estimation of antecedent wetness conditions in rainfall–runoff modelling.” *Hydrol. Processes*, 22, 629–642.

Bonnin, G.M. et al. (2006). *Precipitation Frequency Atlas of the United States*. NOAA, 2(3). <http://www.nws.noaa.gov/oh/hdsc/PF_documents/Atlas14_ Volume2.pdf>. (Nov. 24, 2014).

Chow, V.T. *Handbook of Applied Hydrology*. New York: McGraw-Hill, 1964.

Danish Hydraulic Institute (DHI). (2007). *MIKE SHE User Manual: Volume 1*. DHI Water & Environment.

Dunne, T., Zhang, W, and Aubry, B.F. (1991). “Effects of rainfall intensity, vegetation, and microtopography on infiltration and runoff.” *Water Resources Res*., 27, 2271-2285. Ewen, J. and Birkinshaw, S.J. (2007). "Lumped hysteretic model for subsurface stormflow developed using downward approach." *Hydrol. Process*., 21, 1496-1505.

Farmer, D., Sivapalan, M., and Jothityangkoon, C (2003). "Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis." *Water Resour. Res.*, AGU, 39(2), doi:10.1029/2001WR000328.

Fleming, M. (2002). "Continuous hydrologic modeling with HMS: parameter estimation and model calibration and validation." MS thesis, Dept. of Civil and Environmental Engineering, Tennessee Technological Univ., Cookeville, Tenn.

Gan, T.Y., Dlamini, E.M., and Biftu, G.F. (1997). “Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling.” *J. Hydrol.*, 192, 81-103.

Gesch, D.B., Oimoen, M.J., and Evans, G.A., (2014). “Accuracy assessment of the U.S. Geological Survey National Elevation Dataset, and comparison with other large-area elevation datasets—SRTM and ASTER.” *U.S. Geological Survey Open-File Report 2014–1008*, 10 p.

Gray, H.H. (2000). “Physiographic divisions of Indiana.” *Indiana Geological Survey Special Report 61*, 15 p.

Gunderson, D. (Narrator). (2013, May 1). Why was the Fargo flood forecast off by so much? [Radio broadcast episode]. In A. Silverman and G. Smith (Producers), *All Things Considered*. Washington, DC: National Public Radio.

Gupta, R.S. *Hydrology and Hydraulic Structures*. Long Grove, IL: Waveland Press, 2008.

Helsel, D.R. and R. M. Hirsch. (2002). “Statistical Methods in Water Resources.” *Techniques of Water Resources Investigations*, Book 4, Chapter A3. U.S. Geological Survey. 522 pages.

Huang, M., Gallichand, J., Dong, C., Wang, Z., Shao, M. (2007). “Use of soil moisture data and curve number method for estimating runoff in the Loess Plateau of China.” *Hydrol. Processes*, 21 (11), 1471–1481.

Hydrologic Engineering Center (HEC). (2000). *Hydrologic modeling system HEC-HMS: Technical reference manual*, U.S. Army Corps of Engineers, Hydrologic Engineering Center, Davis, Calif.

Jury, W.A. and Horton, R. *Soil Physics*. Hoboken, NJ: Wiley, 2004.

Kirkby, M.J. *Hillslope Hydrology*. Great Britain: Wiley, 1978.

Klemeš, V. (1983). "Conceptualization and scale in hydrology." *J. Hydrol*., 65, 1-23.

Kusumastuti, D.I. et al. (2006). "Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency." *Hydrol. Earth Syst. Sci. Discuss.* 3, 3239-3277.

Kutner, M.H., Nachtscheim, C.J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*. Boston: McGraw-Hill.

Lan-Anh, N.T., and Willems, P. (2011). "Adopting the downward approach in hydrological model development: the Bradford catchment case study." *Hydrol. Process.*, 25, 1681-1693.

Linsley, R., Kohler, M., and Paulhus, J. (1982). *Hydrology for Engineers.* New York: McGraw-Hill.

Rahman, M.M, Lin, Z., Jia, X., Steele, D.D., and DeSutter, T.M. (2014). “Impact of subsurface drainage on streamflows in the Red River of the North basin.” *J. Hydrol*., 511, 474-483.

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R. (2003). "Downward approach to hydrological prediction." *Hydrol. Process.*, 17, 2101-2111.

Skaggs, R.W., Breve, M.A., and Gilliam, J.W. (1994). "Hydrologic and water quality impacts of agricultural drainage." *Crit. Rev. Env. Sci. Technol.*, 24(1), 1-32.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. <http://websoilsurvey.nrcs.usda.gov/>. (Mar. 17, 2014).

Steenhuis, T.S. and Muck, R.E. (1988). "Preferred movement of nonadsorbed chemicals on wet, shallow, sloping soils." *J. Environ. Qual*., 17, 376-384.

Tramblay, Y., Bouvier, C., Martin, C., Didon-Lescot, J.-F., Todorovik, D., and Domergue, J.-M. (2010). “Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling.” *J. Hydrol.*, 387, 176-187.

United States Army Corps of Engineers (USACE). (2011). "Wabash River Watershed: Section 729 Initial Watershed Assessment." <http://www.lrl.usace.army.mil/Portals/64/docs/CWProjects/WabashStudy.pdf>. (Dec. 31, 2014).

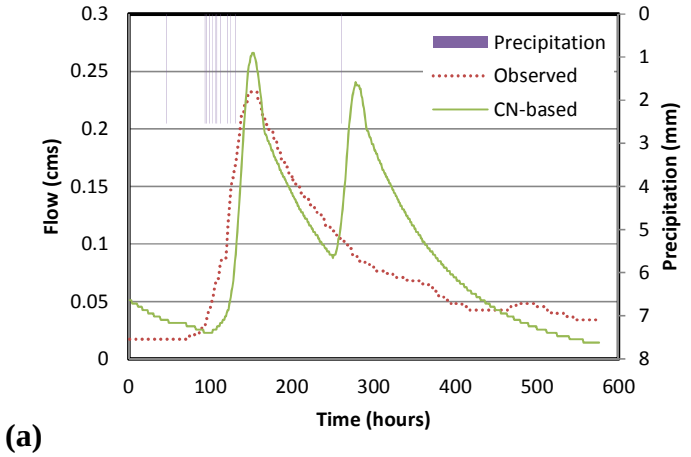
Veihmeyer F.J., and Hendrickson A. (1931). "The moisture equivalent as a measure of the field capacity of soils." *Soil Sci*., 32, 181–193.

Zaslavsky, D. and Rogowski, A.S. (1969). "Hydrologic and morphologic implications of anisotropy and infiltration in soil profile development." *Soil Sci. Soc. Proc*., 33, 594-599.

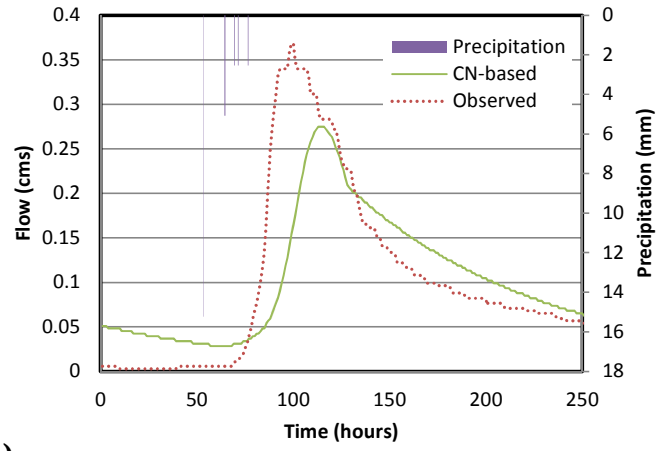
Zucker L.A., and Brown L.C. (1998). "Agricultural drainage: water quality impacts and subsurface drainage studies in the Midwest." *Ohio State University Extension Bulletin 871*. The Ohio State University.

APPENDIX

APPENDIX

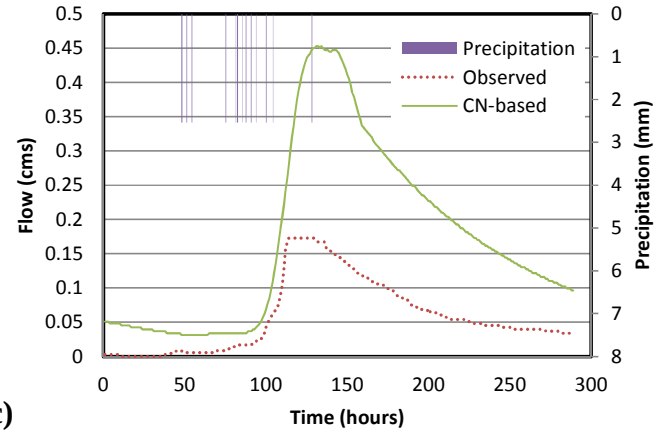


(a)



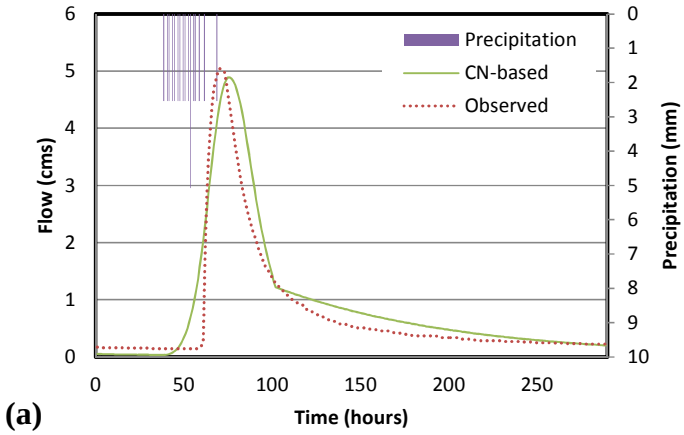
(b)

Figure A.1: Plum Creek CN-based summer calibration storms **(a)** July 2009, **(b)** August 2009, **(c)** November 2011

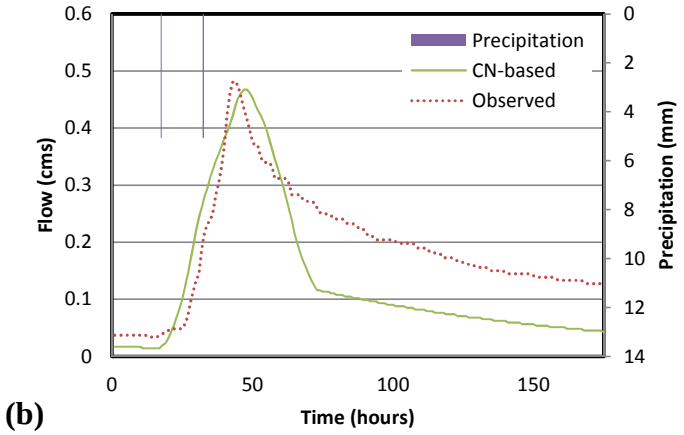


(c)

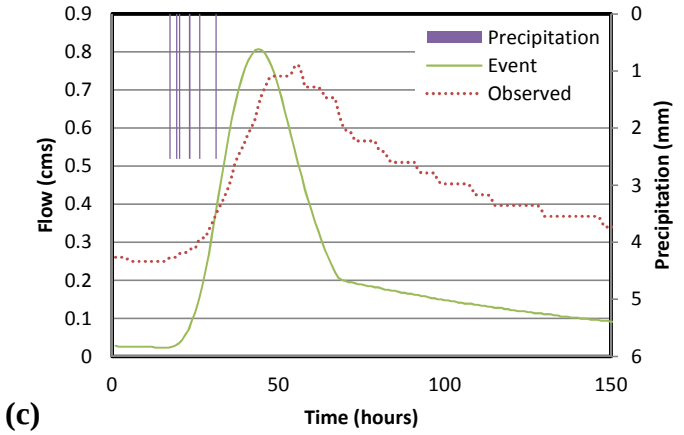
Figure A.1 (cont.): Plum Creek CN-based summer calibration storms **(a)** July 2009, **(b)** August 2009, **(c)** November 2011



(a)



(b)



(c)

Figure A.2: Plum Creek CN-based winter calibration storms (a) December 2008, (b) June 2009, (c) April 2011

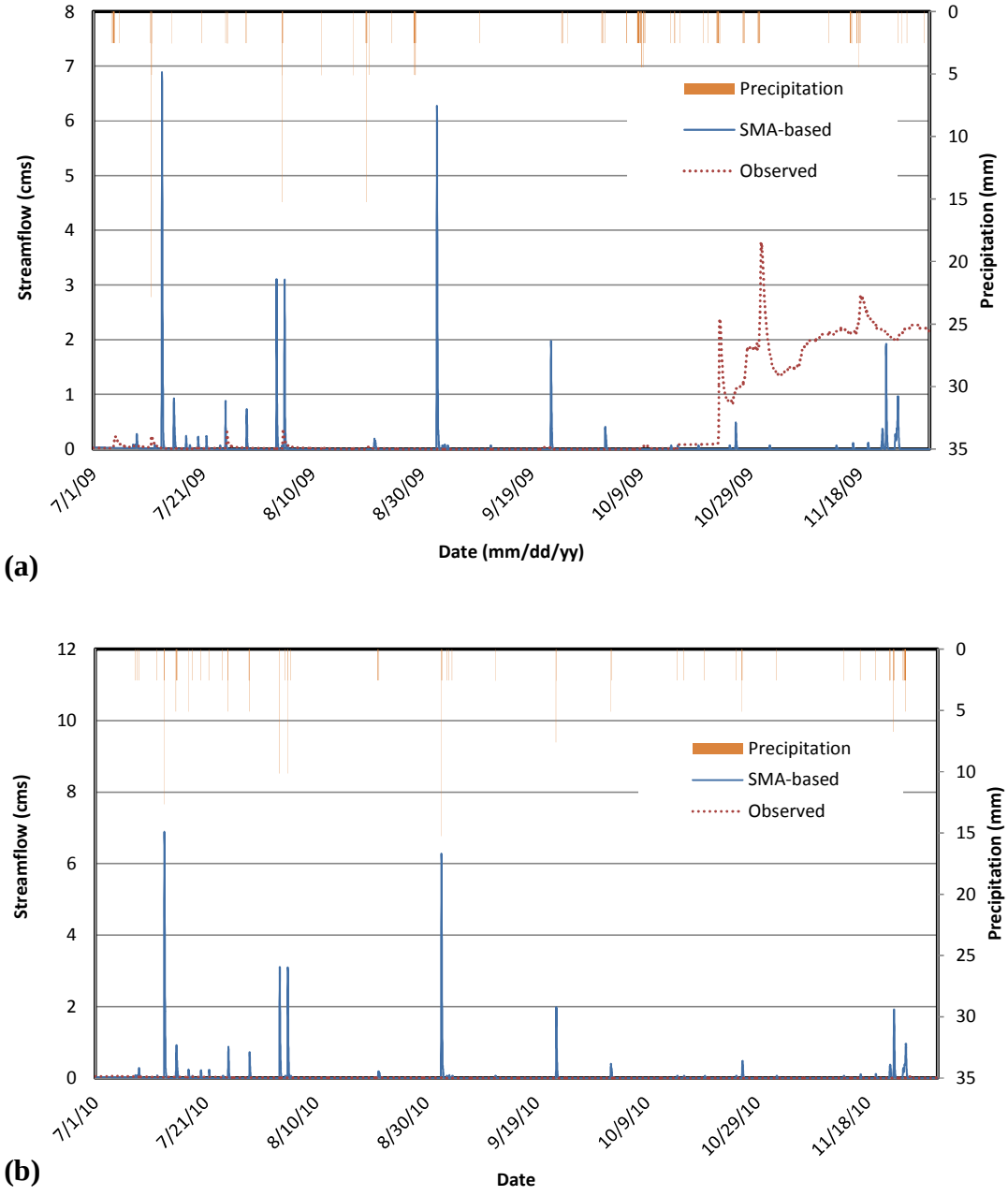
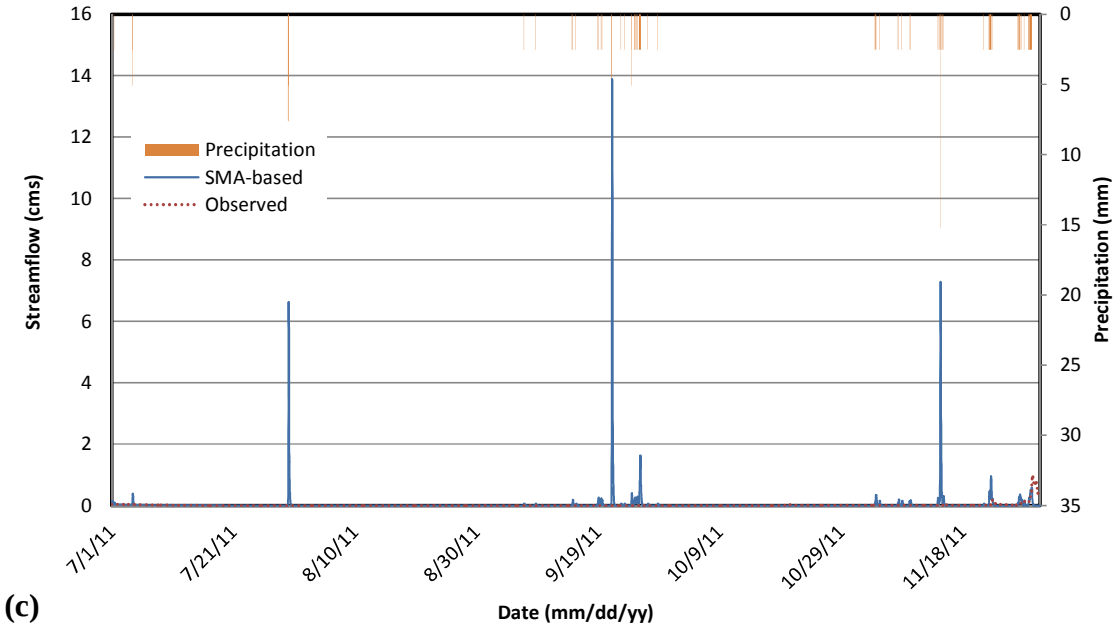
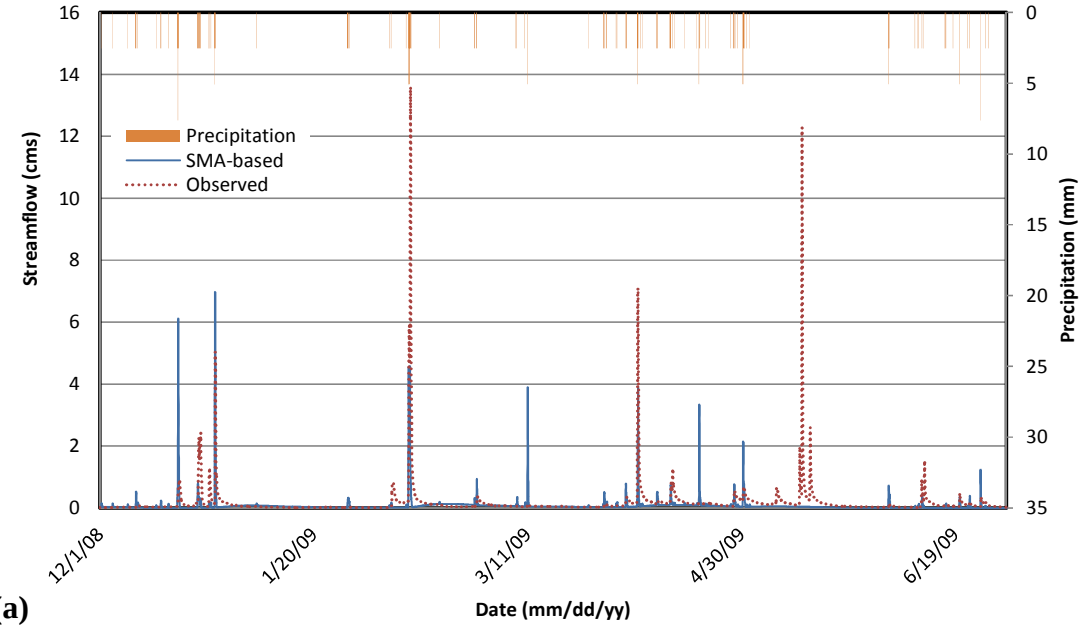


Figure A.3: Plum Creek SMA-based summer calibration **(a)** summer 2009, **(b)** summer 2010, **(c)** summer 2011



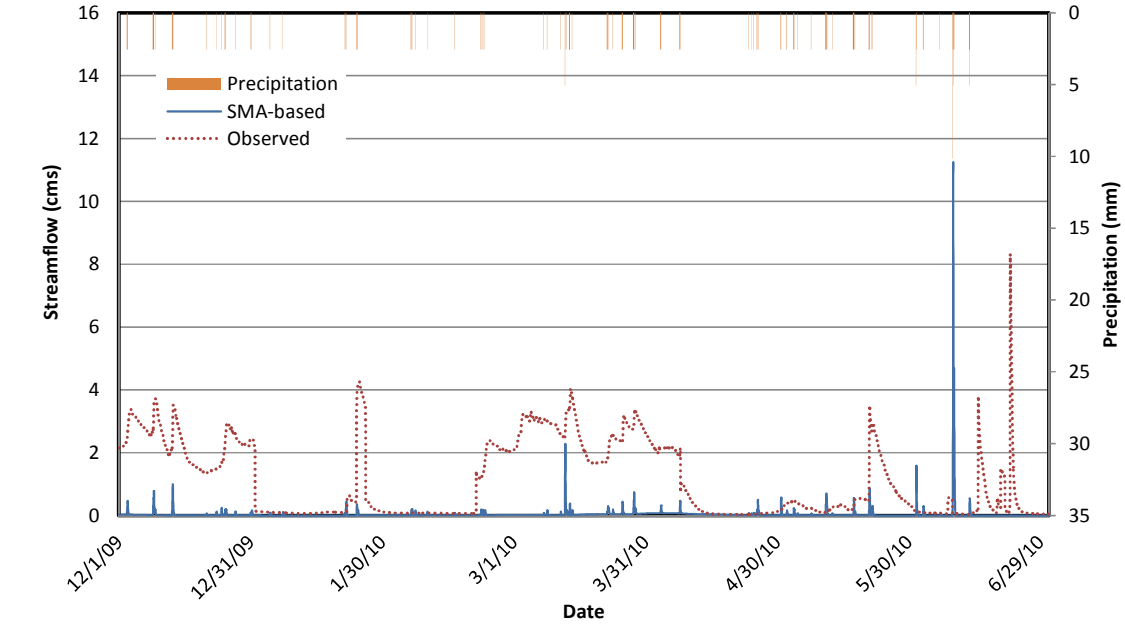
(c)

Figure A.3 (cont.): Plum Creek SMA-based summer calibration **(a)** summer 2009, **(b)** summer 2010, **(c)** summer 2011

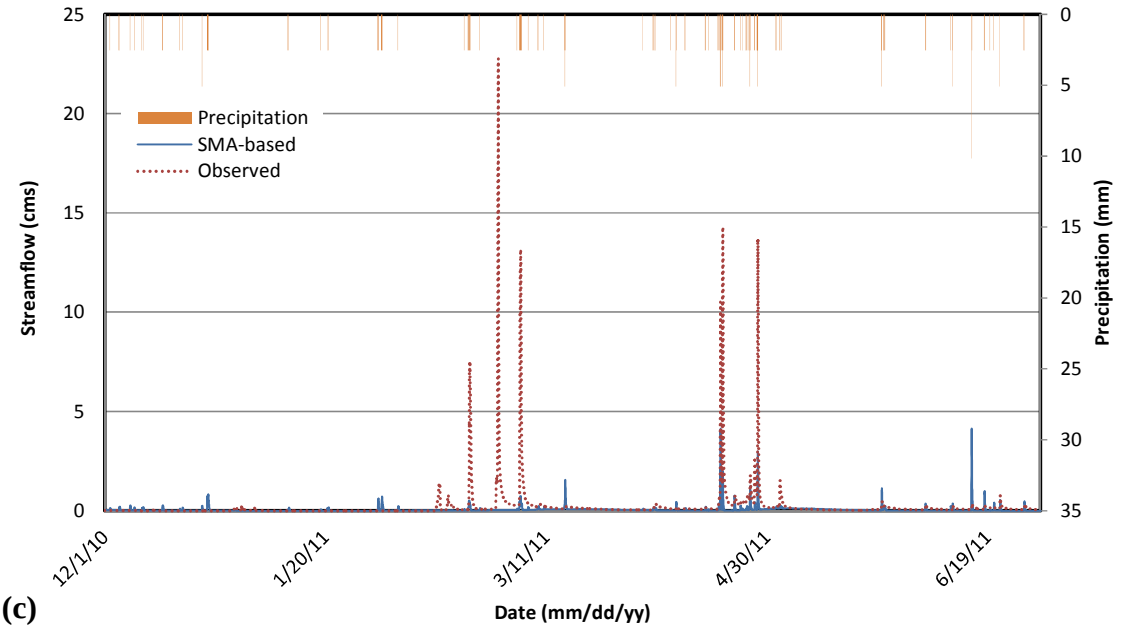


(a)

Figure A.4: Plum Creek SMA-based winter calibration **(a)** winter 2009, **(b)** winter 2010, **(c)** winter 2011

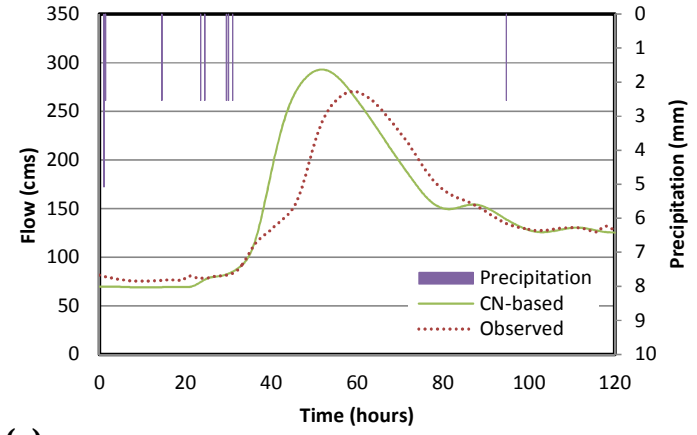


(b)

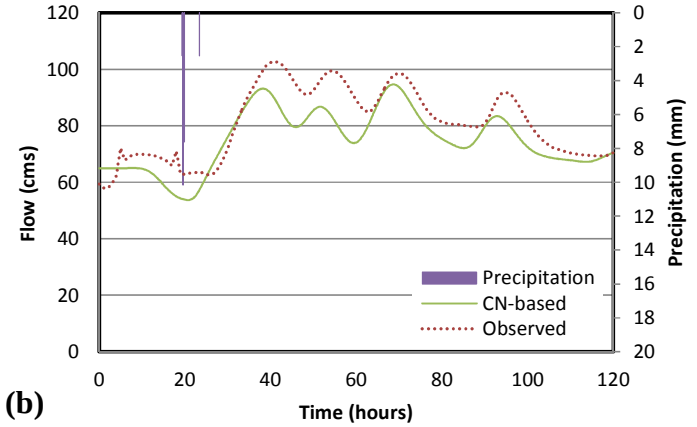


(c)

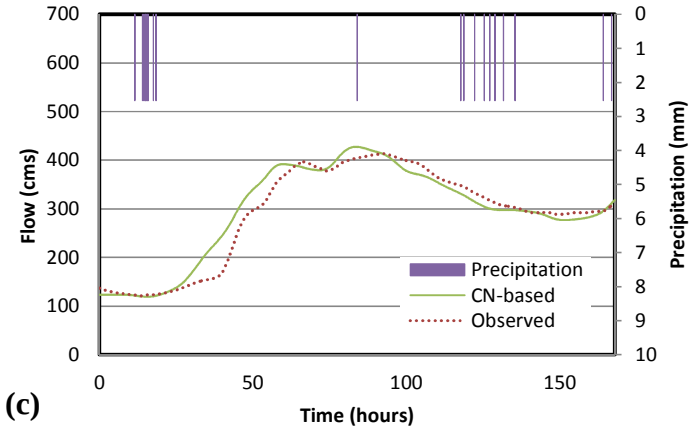
Figure A.4 (cont.): Plum Creek SMA-based winter calibration **(a)** winter 2009, **(b)** winter 2010, **(c)** winter 2011



(a)

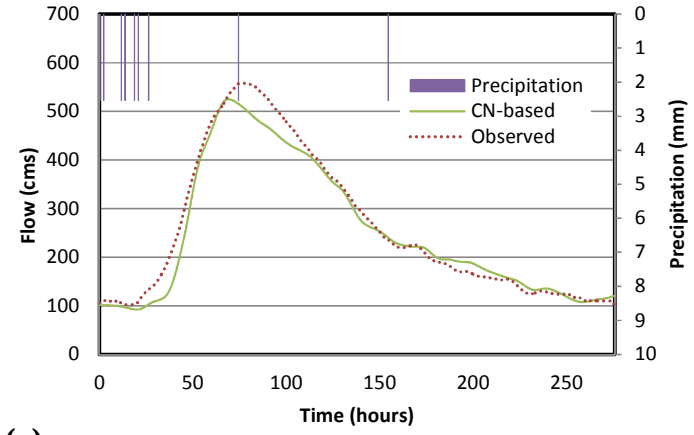


(b)

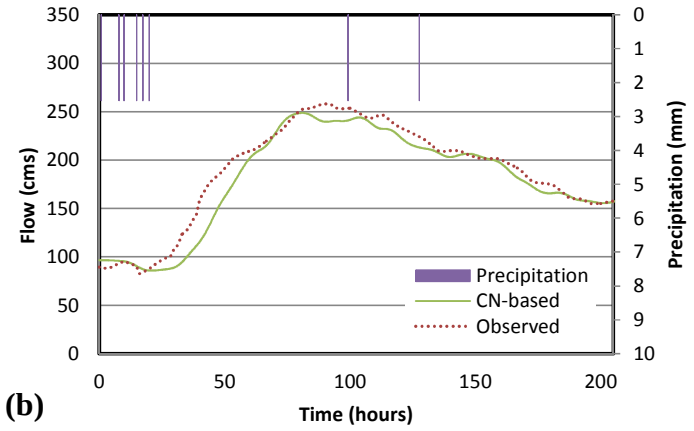


(c)

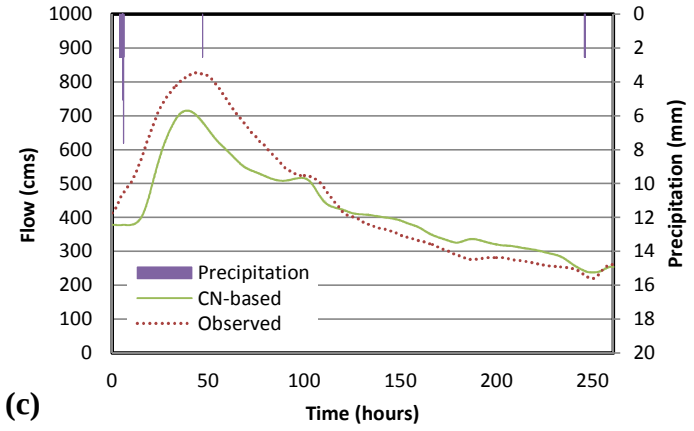
Figure A.5: Wabash/Tippecanoe CN-based summer calibration storms **(a)** July 2009, **(b)** August 2011, **(c)** November 2011



(a)

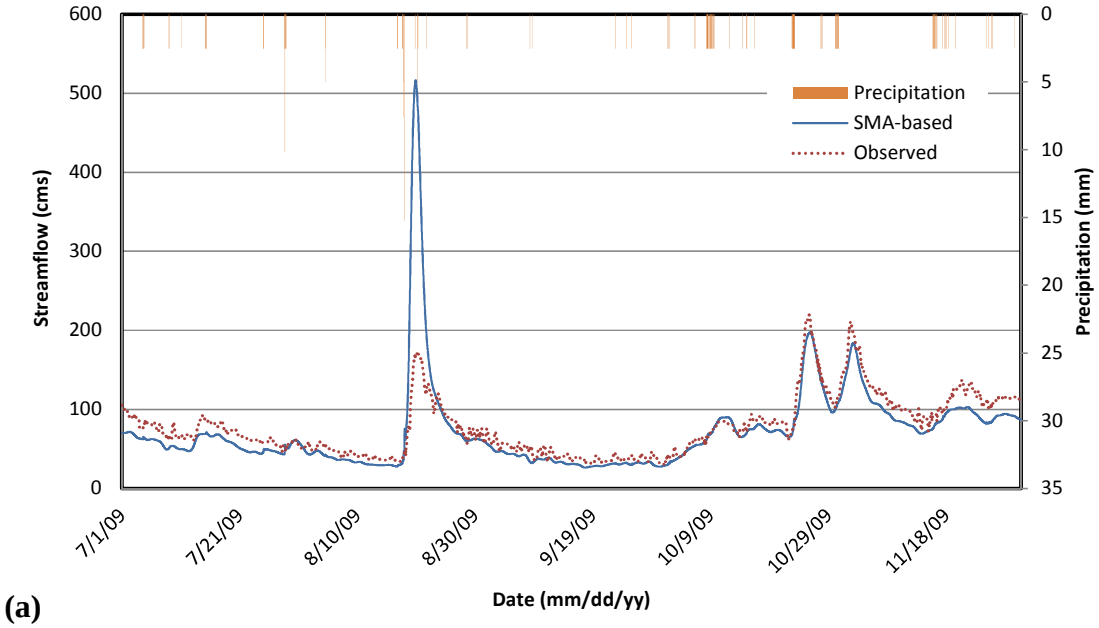


(b)

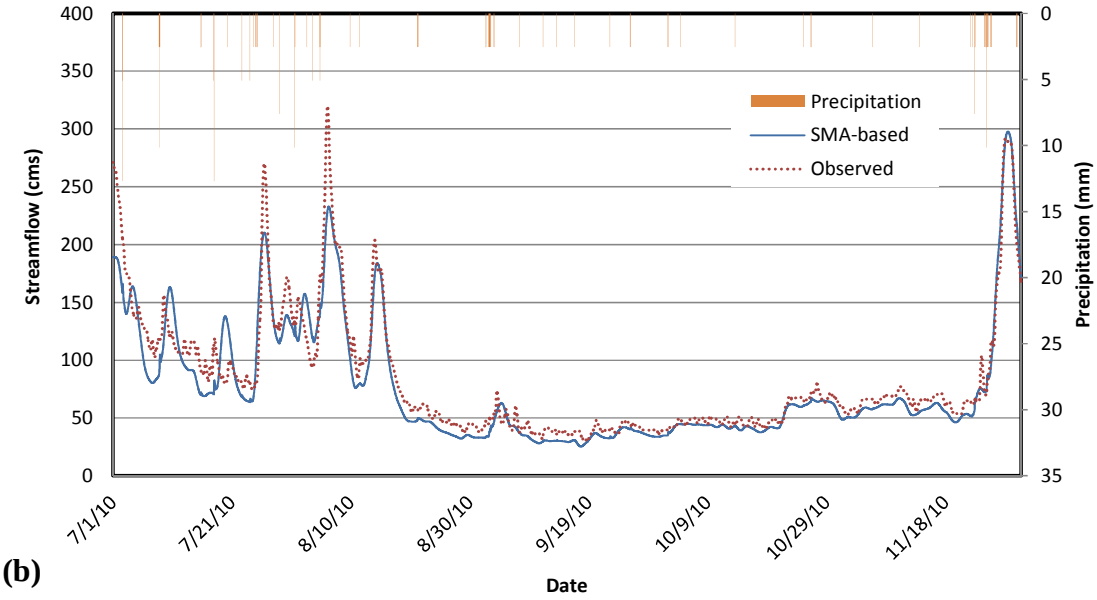


(c)

Figure A.6: Wabash/Tippecanoe CN-based winter calibration storms **(a)** December 2009, **(b)** February 2010, **(c)** May 2010

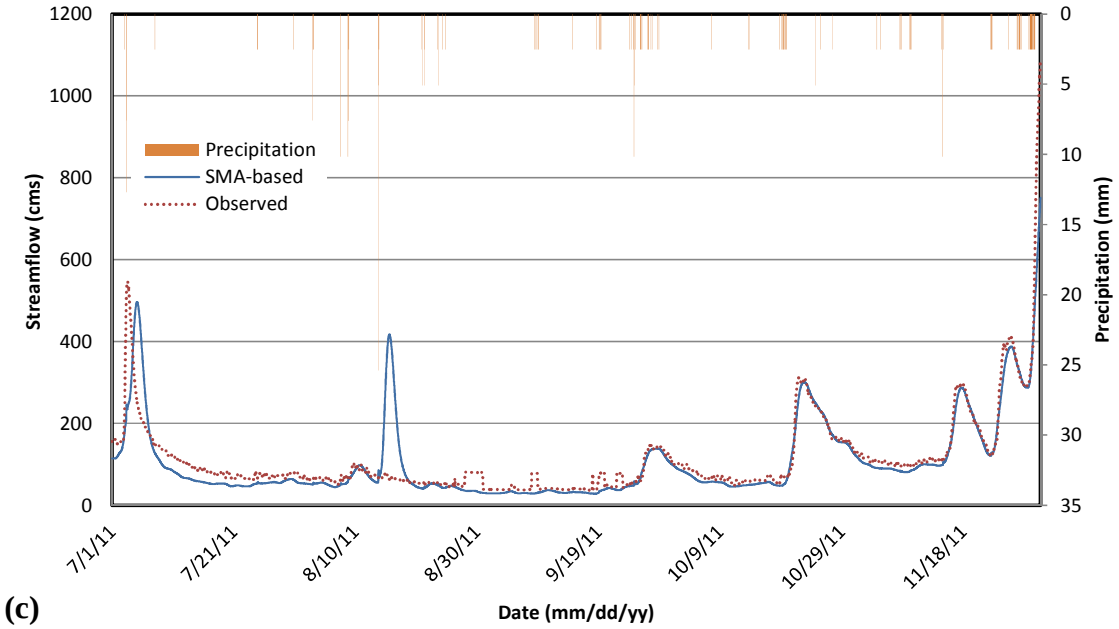


(a)



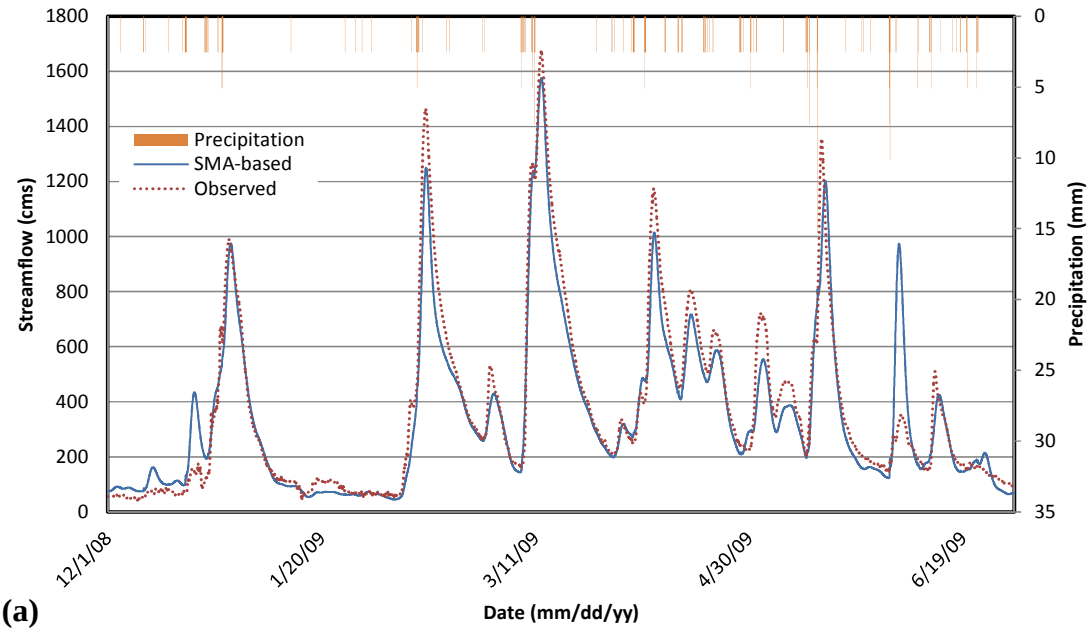
(b)

Figure A.7: Wabash/Tippecanoe SMA-based summer calibration **(a)** summer 2009, **(b)** summer 2010, **(c)** summer 2011



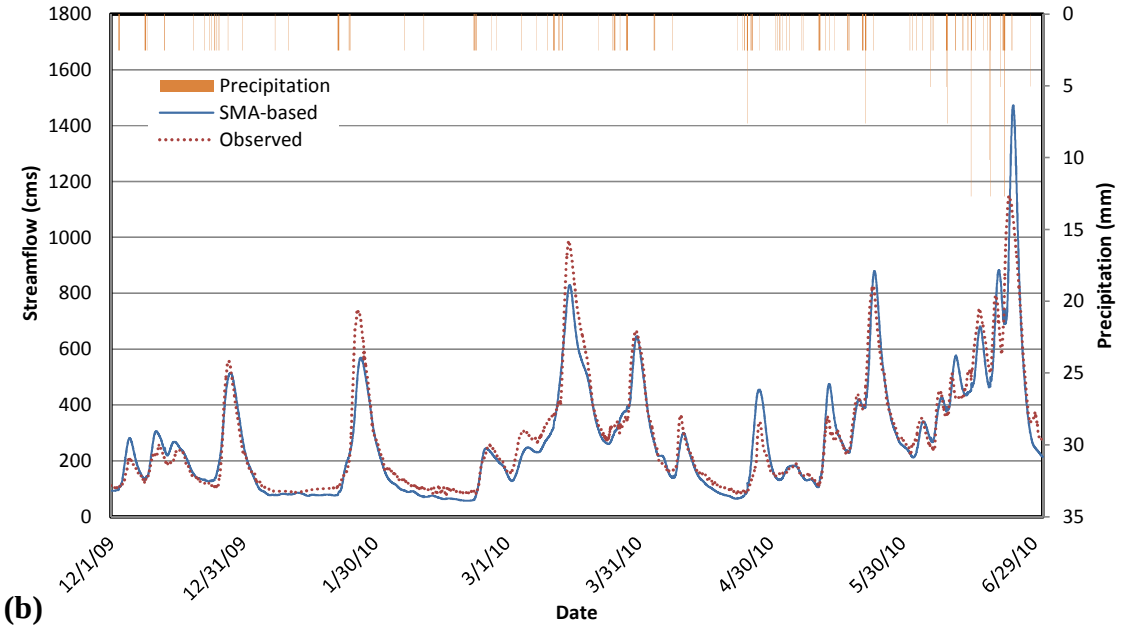
(c)

Figure A.7 (cont.): Wabash/Tippecanoe SMA-based summer calibration **(a)** summer 2009, **(b)** summer 2010, **(c)** summer 2011

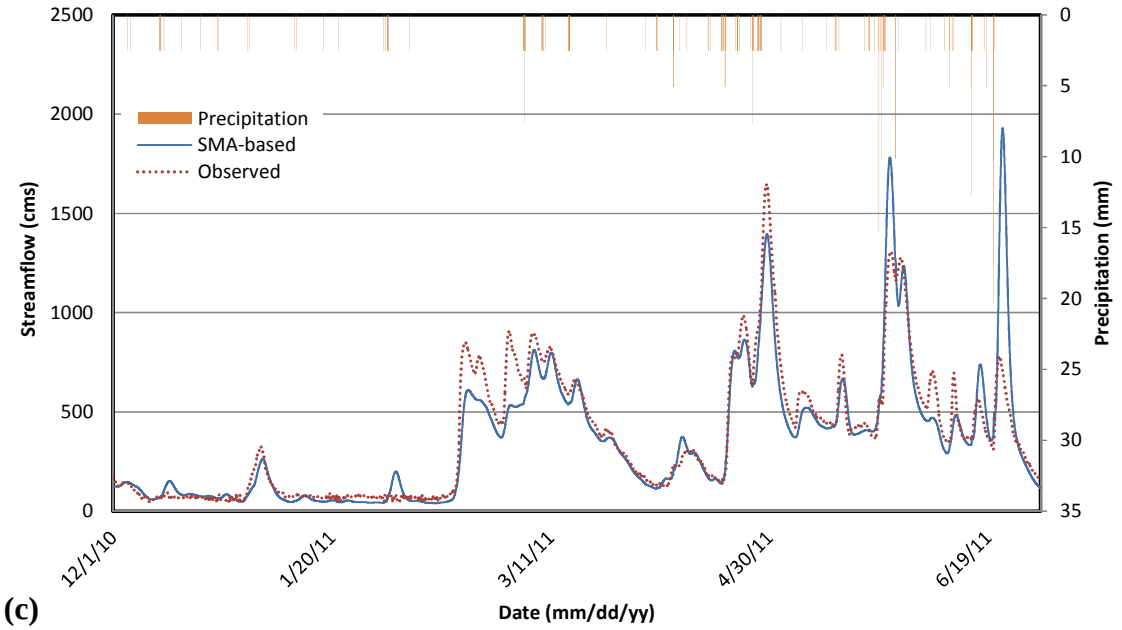


(a)

Figure A.8: Wabash/Tippecanoe SMA-based winter calibration **(a)** winter 2009, **(b)** winter 2010, **(c)** winter 2011



(b)



(c)

Figure A.8 (cont.): Wabash/Tippecanoe SMA-based winter calibration **(a)** winter 2009, **(b)** winter 2010, **(c)** winter 2011