ABSTRACT

Title of Dissertation: RESPONSE OF HYDROLOGIC CALIBRATION TO REPLACING GAUGE-BASED WITH NEXRAD-BASED PRECIPITATION DATA IN THE USEPA CHESAPEAKE BAY WATERSHED MODEL

Sunghee Kim, Doctor of Philosophy, 2012

Dissertation directed by: Associate Professor, Dr. Kaye L. Brubaker, Department of Civil and Environmental Engineering

This study investigated the response of hydrologic calibration to replacing gauge-based with radar-based precipitation data in the USEPA Chesapeake Bay Program (CBP) Watershed (CBW) model over the Potomac River Basin. Specific objectives were to (1) compare gauge-based and NEXRAD radar-based (Multisensor Precipitation Estimator, MPE) data at the (a) point-pixel and (b) spatially aggregated level; (2) evaluate the model's calibration accuracy using the different precipitation data sets; and (3) examine the response of model hydrology.

Hourly gauge-point and MPE-pixel data were compared at 80 locations. The CBP's interpolated and aggregated precipitation data at the model unit (county) level were compared with MPE data aggregated to the same 114 county-based spatial segments. The model calibration followed the CBP's automated approach, using observed streamflow at 37 gauge stations. Model performance was evaluated using

calibration and hydrologic statistics, and GIS-aided spatial information. Calibrated parameters and model hydrologic fluxes were compared.

The average annual gauge-point and MPE-pixel values (excluding hours when either was missing) agreed well. Differences in average annual values between the spatially aggregated data sets were, however, significant in parts of the study area.

When parameter constraints were relaxed to allow calibration to adjust to the smaller volume of precipitation, the model using MPE outperformed the model calibrated to CBP precipitation data at 65% of the 37calibration sites. The model response was controlled largely by the seasonal difference in precipitation inputs: (1) calibration process could not compensate for large differences in seasonal flow bias caused by the seasonal volume of precipitation; (2) seasonal flow bias affected the lower zone nominal soil moisture storage parameter (LZSN), mainly affecting interflow and groundwater flow. The surface flow component was generally the same for the different precipitation inputs. The two precipitation data types can be used interchangeably to simulate surface-flow dominated processes, but care must be taken in simulations where subsurface pathways and residence times are important.

MPE is a strong alternative to gauge-based precipitation data because of its spatiotemporal coverage and rare missing records. Using MPE in hydrologic modeling is appealing because of the improved calibration accuracy of the CBW model demonstrated in this study.

RESPONSE OF HYDROLOGIC CALIBRATION TO REPLACING GAUGE-BASED WITH NEXRAD-BASED PRECIPITATION DATA IN THE USEPA CHESAPEAKE BAY WATERSHED MODEL

by

Sunghee Kim

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:

Associate Professor Kaye Brubaker, Chair
Professor Allen Davis
Dr. Ross Mandel
Professor Richard H. McCuen
Professor Adel Shirmohammadi

In memory of my parents-in-law

DEDICATION

To my parents for their constant support

ACKNOWLEDGEMENT

I would like to thank my advisor, Dr. Kaye Brubaker for her guidance and support. I am grateful to Dr. Richard McCuen for his encouragement and invaluable advice throughout my study. My gratitude goes to my committee members, Dr. Allen Davis, Dr. Ross Mandel, Dr. Richard McCuen, and Dr. Adel Shirmohammadi for their academic mentorship. I acknowledge the support of Mr. Joseph Ostrowski and Mr. Edwin Pryor from the Middle Atlantic River Forecast Center for providing NEXRAD XMRG datasets and information on the status of the telemetry device at COOP gauge stations. The support provided by Mr. Joseph Hoffman and Mr. Carlton Haywood from the Interstate Commission on the Potomac River Basin is gratefully acknowledged.

My special thanks go to my three children, Sangduk, Jungduk, and Shinduk. I would especially like to thank Shinduk for being my English grammar mentor. My deepest appreciation goes to my husband DJ.

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations frequently used in this study

AGRM	ASOS Gauge based Rainfall Measurement
AGWO	Groundwater flow
AGWR	Groundwater Recession Index
AOPA	Aircraft Owners and Pilots Association
ASOS	Automated Surface Observing System
AWPAG	All Weather Precipitation Accumulation Gauge
CBP	Chesapeake Bay Program
CBW	Chesapeake Bay Watershed
CGRM	COOP Gauge based Rainfall Measurement
COOP	Cooperative Observer Program
ECDF	Empirical Cumulative Distribution Function
F&P	Fischer-Porter
FPU	Fischer-Porter Upgrade
GIS	Geographical Information System
GOF	Goodness-Of-Fit
GRM	Gauge based Rainfall Measurement
H&WQ	Hydrology and Water Quality
HPD	Hourly Precipitation Data
HRAP	Hydrologic Rainfall Analysis Project
HSPF	Hydrological Simulation Program-Fortran
ICPF	Intensity-based Cumulative Precipitation Function
IFWO	Interflow or subsurface flow

INFILT	Infiltration Index
INTFW	Interflow Inflow Index
IRC	Interflow Recession Index
KS	Kolmogorov-Smirnov test
KVARY	Groundwater recession parameter
LAND_EVAP	Land evaporation parameter (multiplication factor)
LARC/DCP	Limited Automatic Remote Collector/Data Collection Platform
LZRAT	Ratio of LZS to LZSN
LZSN	Nominal Lower Zone soil moisture Storage
MARFC	Middle Atlantic River Forecast Center
mCBP	Mean areal time series of gauge-based hourly precipitation developed by USGS/CBP
mMPE	Mean areal time series of radar-based (MPE) hourly precipitation developed for this study
MPE	Multisensor Precipitation Estimator
NCDC	National Climatic Data Center
NEXRAD	NEXt Generation Weather RADar
NSE	Nash-Sutcliffe Efficiency
NWSMC	National Weather Service Modernization Committee
PET	Potential EvapoTranspiration
PRB	Potomac River Basin
R-B index	Richards-Baker Flashiness index
RMSD	Root Mean Square Difference
RRE	Radar-based Rainfall Estimates

SURO	Direct Surface Runoff
TMDL	Total Maximum Daily Load
USDC	US Department of Commerce
USEPA	US Environmental Protection Agency
USGS	US Geological Survey
UZSN	Nominal Upper Zone soil moisture Storage
WDM	Watershed Data Management

CHAPTER ONE

Introduction

1.1 Research Motivation

1.1.1 Uncertainty in simulation models

Hydrologic simulation models are widely used as decision-making tools in water resources management, watershed planning, and regulatory policy-making. Although field monitoring is a desirable method, simulation models offer time- and cost-effective methods to conduct and evaluate management alternatives (Munoz-Carpena et al., 2006; Shirmohammadi et al, 2006).

From a management or decision-making perspective, the main concern is the inaccurate interpretation of model output by failing to account for uncertainty in model results. The sources of uncertainty in the modeling process include inaccuracies in both parameter values and input data. Inaccuracies in parameter values arise from the fact that parameter values cannot be measured directly, but can only be inferred by calibrating model outputs to historical observed data (Vrugt et al., 2005). Vrugt et al. (2005) also reported that, because of errors in inputs and outputs, inaccuracies in parameter values increase. Inaccuracies in input data can be introduced by measurements and/or data preprocessing (Smith et al., 2004) or by the selection of inappropriate data for a given modeling purpose (Das et al., 2008). Variability in model parameters or input data due to heterogeneity in a natural system such as climate, soil, or land use increases uncertainty in model outputs (Shirmohammadi et al., 2008).

The common approach to minimizing model output uncertainty is through model calibration in which model parameters are adjusted so that model predictions match observed data (Shirmohammadi et al., 2006). When natural variability is not properly represented, however, other components of a model are forced to compensate for shortcomings of the model during the calibration process. As a result, the propagation of errors occurs through the entire model simulation, affecting outputs.

1.1.2 Precipitation data

Uncertainties in precipitation data greatly affect the performance of models because precipitation is the major driving force in hydrology modeling (Dingman, 2002). Studies have shown that the inability to represent precipitation space/time variability results in the reduced performance of hydrology and water quality (H&WQ) models (Shirmohammadi et al., 2006). An attempt to overcome this problem is to use spatiotemporally distributed precipitation data. Then, the spatial variability of precipitation may be translated into more accurate predictions (Choi and Olivera, 2008; Cashman and Potter, 2006).

A critical constraint for fine-resolution spatiotemporal continuous H&WQ simulations is data availability. Accurate simulation of watershed hydrology requires accurate data about the amount, location, and timing of precipitation. Continuous simulation requires precipitation data at an hourly time scale or finer. A major source of the long-term hourly precipitation data (HPD) in the United States is the hourly Cooperative Observer Program (COOP) network. It uses the Fischer-Porter (F&P) weighing gauges utilizing paper-punch mechanism (NWSMC, 1998). Automated Surface

Observing System (ASOS) data sets are also available as gauge-based rainfall measurements (GRM). In general, the rain gauge station networks are too sparse to capture spatial variability (Schuurmans and Bierken, 2007). Dealing with missing data in the gauge records is also a concern to model users (Dingman, 2002).

An alternative to gauge-based rainfall data is the radar-based rainfall estimates (RRE). The U.S. National Weather Service's Next Generation Weather Radar (NEXRAD) system consists of the Weather Surveillance Radar-1988 Doppler (WSR-1988D) radars. The NEXRAD system measures reflectivity and the intensity of radar echo returns from precipitation particles (e.g., Krajewski and Smith, 2002; Seo and Breidenbach, 2002; Seo et al., 2000; Fulton et al, 1998). Then, the Z-R (reflectivity-rainfall rate) relationships are used to estimate precipitation rates from the reflectivity (e.g., Seo and Breidenbach, 2002; Seo et al., 2000). The RRE is produced with its own theoretical and technical challenges from selecting coefficients in the Z-R relationship to instrumental reliability. A well-known issue in producing radar-based precipitation is that uncertainties of precipitation increase as the distance from the NEXRAD radar stations increases (Seo et al., 2010).

Given that different types of precipitation data are available to model users, two questions arise. Which data type is better for a given model? Is a modeler able to assess the implication of using one data type over another?

1.1.3 Precipitation data for Chesapeake Bay Watershed model

Over the years, incorporating RRE data sets into the Chesapeake Bay Watershed (CBW) model developed by the USEPA Chesapeake Bay Program (CBP) has

been recommended (USEPA, 2010; Band et al., 2008; Band et al. 2005). From the modeling perspective, incorporating RRE for an operational use is recommended only when model accuracy improves using RRE. The CBW model is built on the Hydrological Simulation Program-Fortran (HSPF), which is a comprehensive software package for building spatially lumped watershed hydrology and water quality models. The CBW model parameters represent spatially averaged characteristics in the model hydrology, and therefore, the model may not recognize the merits of RRE. This issue is addressed by investigating the CBW model behavior with different types of precipitation data.

1.2 Research Questions and Objectives

This research investigated the following questions: Can model users replace gauge-based precipitation data with radar-based ones in the Chesapeake Bay Watershed model? What are the implications of using the model with the replaced precipitation data? To answer the research questions, the following specific objectives were investigated:

(1) Compare the hourly gauge-based and radar-based precipitation data.

(2) Calibrate the CBW model using two different types of precipitation data.

(3) Investigate model hydrology when the CBW model is calibrated with the different forcing inputs.

The fulfillment of these research objectives will demonstrate the potential contributions to the modeling community. First, this study will provide descriptive statistics that will identify the similarities and differences between gauge-based and radar-based

precipitation data sets. The results of the data comparison will enable modelers to assess the implication of using one data type over the other in hydrologic modeling. Second, this study will compare the CBW model performance using the two types of precipitation data. At the end of the study, modelers will be able to make a decision whether or not implementing the radar-based precipitation into the CBW model is beneficial. Third, this study will provide the technical aspects of implementing NEXRAD radar-based precipitation in the Chesapeake Bay Watershed model at a large scale.

This dissertation consists of six chapters, including this first chapter. Chapter 2 introduces the introduction of the CBW model with a brief overview of HSPF. Chapters 3, 4, and 5 are about Objectives 1, 2, and 3, respectively. Chapter 6 summarizes the research, describes its contribution to the modeling community, and discusses recommendations for future work.

CHAPTER TWO

The Chesapeake Bay Watershed Model

2.1 Introduction

This chapter describes the CBW model. Because the CBW model is built on HSPF, the implementation of HSPF into the CBW model system is explained. Since HSPF is a widely used and well documented modeling environment, this chapter emphasizes aspects that are different from a typical operation of HSPF. This section also introduces the Chesapeake Bay Program's preparation of the precipitation and meteorological time series for the CBW model.

2.2 The Chesapeake Bay Watershed (CBW) Model

After excessive nutrient problems were identified from the 7-year USEPA Study in 1983 (Martucci et al., 2006), an agreement signed by Chesapeake Bay watershed states, the District of Columbia, the Chesapeake Bay Commission and the U.S. Environmental Protection Agency (USEPA) led to the creation of the USEPA Chesapeake Bay Program (CBP). It was necessary to develop a model to simulate watershed processes to meet both CBP's and states' goals. The former is to restore the Bay by reducing nutrients and the latter is related to Total Maximum Daily Loads (TMDLs), which are mandated under the Clean Water Act. The CBW model was developed with modifications of the HSPF software to accommodate Best Management Practices (BMPs) and land use changes with time, as well as various postprocessing and analytical tools (Martucci et al., 2006). The results from the CBW model simulations are

intended to provide scientific information to evaluate the Bay restoration goals as well as to develop TMDLs.

The CBW model consists of two parts: the HSPF core and the surrounding code that handles input, output, and calibration (Shenk, 2010). According to the manual (Shenk, 2010), running the model requires the execution of a series of programs as follows (Figure 2-1). In the figure, land segments and river segments are spatial model units for hydrologic process simulation based on county and watershed subdivision, respectively (discussed below in Section 2.4). Land segments are further subdivided by land use to create noncontiguous Hydrologic Response Units. The Watershed Data Management (WDM) file type used in this process is a digital storage format for binary time series data. Meteorology (MET) data, including precipitation data, are required for hydrologic simulation. Atmospheric deposition (ATDEP) and point source (PS) data are needed for water quality modeling.

First, the Land UCI Generator (LUG) produces a UCI (User Controlled Input file) for each land use within each land segment. Then, HSPF modules are run on the land UCIs for each land use and land segment, and the simulated flows are stored in the WDM format. At this stage, all hydrologic fluxes are simulated per acre. Actual flows are calculated by multiplying by the actual size of each land use at a given land segment when the External Transfer Module (ETM) runs. The ETM creates a file of transfer coefficients between land variables to river variables, including unit conversion factors. Then, it aggregates outputs from land simulation and stores them as input data for river simulation. The River UCI Generator (RUG) produces a UCI for each river segment. Then, HSPF modules are run on the river UCIs and the river simulation results are

written back to the river WDM, preserving the input data. Finally, a system of postprocessors generates text outputs from the land and river WDMs. A model user can configure the postprocessor to obtain output on various time and spatial scales.

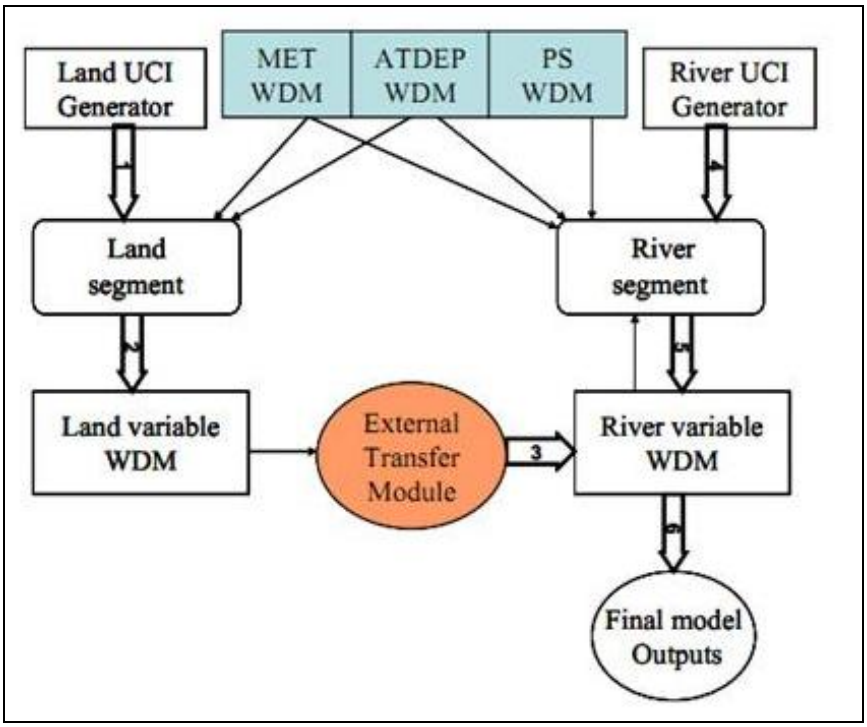


Figure 2-1. Chesapeake Bay Watershed model structure. Watershed Data Management (WDM) binary input files include: meteorology (MET), atmospheric deposition (ATDEP), and point source (PS) data. Adapted from CBP (2010).

2.3 HSPF in the CBW Model

The implementation of HSPF in the CBW model is different from the typical implementation of HSPF (Shenk, 2010). In the original HSPF implementation, hydrologic spatial segmentations (called model units) are spatial sub-areas of a watershed, divided by hydrological characteristics. The hydrological characteristics include rainfall and other meteorological information, watershed topography, stream network topology, stream reach characteristics, soil type, land use, and others. In the CBW model, however, the spatial segmentation is defined by two separate types of

layers: county-based land segments and watershed-based river segments. The original HSPF software uses a single UCI file to simulate each hydrologic segment, but in the CBW model, each land use in each land segment is simulated in a separate UCI and each watershed-based river segment is simulated in another separate UCI. The original HSPF contains SCHEMATIC and MASSLINKS modules to match land segments to river segments or one river segment to another, and hydrologic variables to hydraulic variables, respectively. In the CBW model, the ETM program is responsible for these functions. The ETM also manages point source, septic, and atmospheric deposition loads for water quality modeling (Shenk, 2010).

2.4 Model Units

The CBP uses counties to subdivide the Chesapeake Bay region because much land use and management information is available at the county level, and because land use policy decisions are made at this level. Digital county boundary data sets at a 1:100,000-scale were used to delineate land segments (Martucci et al., 2006). A single political county may be divided into two or three contiguous spatial units, depending on its meteorological homogeneity and other factors. The CBP and this study refer to these spatial units as "land segments". All inputs for the hydrologic simulation, including precipitation, were prepared at this county scale.

The watershed is also subdivided into contiguous units of land directly draining to the stream reaches. The CBP used 30-meter resolution digital elevation model (DEM) data sets to delineate a subwatershed for each river reach (Martucci et al., 2006). The

CBP and this study refer to these units as "river segments". All inputs for the simulation of stream reaches were prepared to represent these river segments.

The CBP's land segments and river segments are two different spatial subdivisions of the same region. The boundaries of land segments and river segments are not necessarily the same. The intersections between land segments and river segments are of interest during the model simulation; such intersections are referred to as land-river segments. After simulated values are calculated per acre for individual land use categories in each land segment, the value is multiplied by the total number of acres of each land use category in a given land-river segment (Figure 2-2).

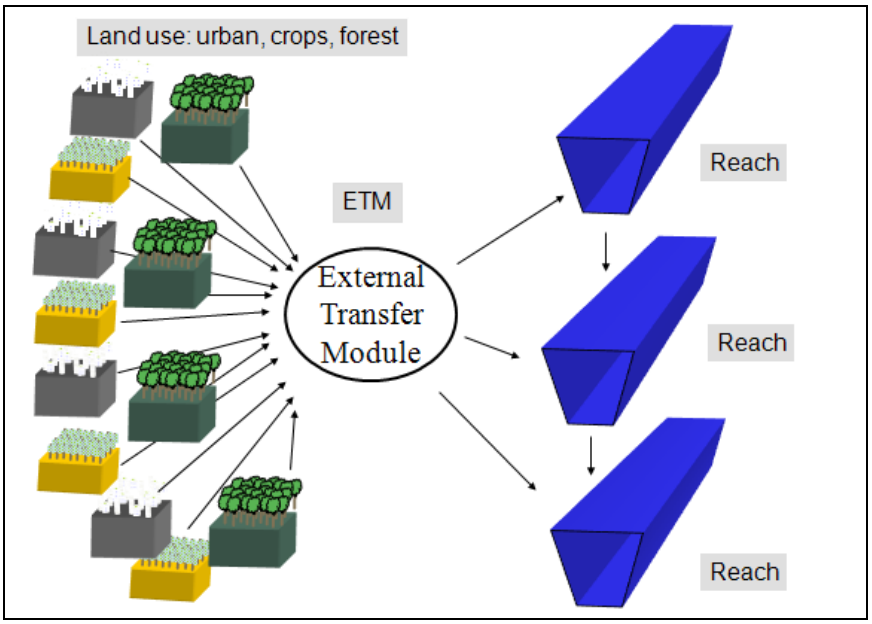


Figure 2-2. Land simulations (left side) and river simulations (right side) in the CBW model. Adapted from CBP (2010).

The results from all land-river segments contained in the river segment are then added and delivered to the associated stream reach for hydraulic and in-stream water quality simulations. The land simulation for each land use type is independent, whereas each reach simulation is dependent on both local land simulations and upstream reach

simulations. For further details on the model segmentation, the reader is referred to Martucci et al. (2006) and USEPA (2010).

2.5 Preparation of Precipitation Data

This section briefly reviews the development of the precipitation time series currently used for the CBW model. The USEPA (2010) documentation is the only publicly available source of this information. The following section is quoted directly from Subsection 2.2 in Section 2 "Meteorology and Precipitation" of USEPA (2010). Note that Phase 5.3 is the version of the CBW model developed at the time USEPA (2010) was written. In this study, Phase 5.2 was used. Both Phase 5.2 and 5.3 use the same precipitation data.

Development of Precipitation Time Series

Lauren Hay of the U.S. Geological Survey (USGS) National Research Program has developed a methodology for estimating the spatial distribution of precipitation and other meteorological variables (Hay et al. 1991, 2000a, 2000b, 2006). Observed meteorological data of precipitation and temperature are interpolated across the Phase 5.3 domain by fitting a multiple regression equation that relates the observed data to latitude, longitude, and elevation. In the case of the Phase 5.3 Model, the Chesapeake Bay basin was divided into six subregions, and a separate regression equation was fitted by month for each subregion. The fitted equations were then interpolated onto a 5-kilometer (km) grid and then averaged over land-segments. That procedure was used to estimate both precipitation and temperature inputs. As the Next Generation Weather Radar (NEXRAD) or the North American Reanalysis continues to evolve and improve,

future versions of the Watershed Model could use data from these sources for estimated precipitation inputs (Over et al. 2007, Mesinger et al. 2006).

Generation of Daily Rainfall Records

HSPF uses estimates of hourly precipitation and other meteorological variables for each model segment. To compute reliable estimates of these quantities, researchers at the USGS National Research Program in Denver have developed a method of interpolation of observed data across a basin to better represent basin climate variability. Significant physical factors affecting the spatial distribution of climate variables in a river basin are latitude (x), longitude (y), and elevation (z). In the method, multiple linear regression (MLR) equations are developed for each dependent climate variable (e.g., precipitation) using the independent variables of x, y, and z from the climate stations. The general form of the MLR equation for daily precipitation (p) is

$$p = b_0 + b_1 x + b_2 y + b_3 z$$

The resulting fit from the above equation describes a plane in three-dimensional space with slopes b1, b2, and b3 intersecting the p axis at b0. Similar equations are used for temperature. Using the station latitude and longitude coordinates in the MLR provides information on the local-scale influences on the climate variables that are not related to elevation, for example, the distance to a topographic barrier. To account for physiographic and seasonal climate variations, MLR equations are developed for each month using mean values from a set of selected stations in and around each subregion. The Chesapeake Bay watershed and southwestern Virginia have been divided into six physiographic subregions for analysis (Figure 2-2 in USEPA 2010). The monthly MLRs

are computed to determine the regression surface that describes the spatial relations between the monthly dependent variables and the independent variables (x, y, and z). Note that for each month, the best MLR relation will not always include all the independent variables (i.e., in some months, latitude, longitude, or elevation might be unimportant to the regression).

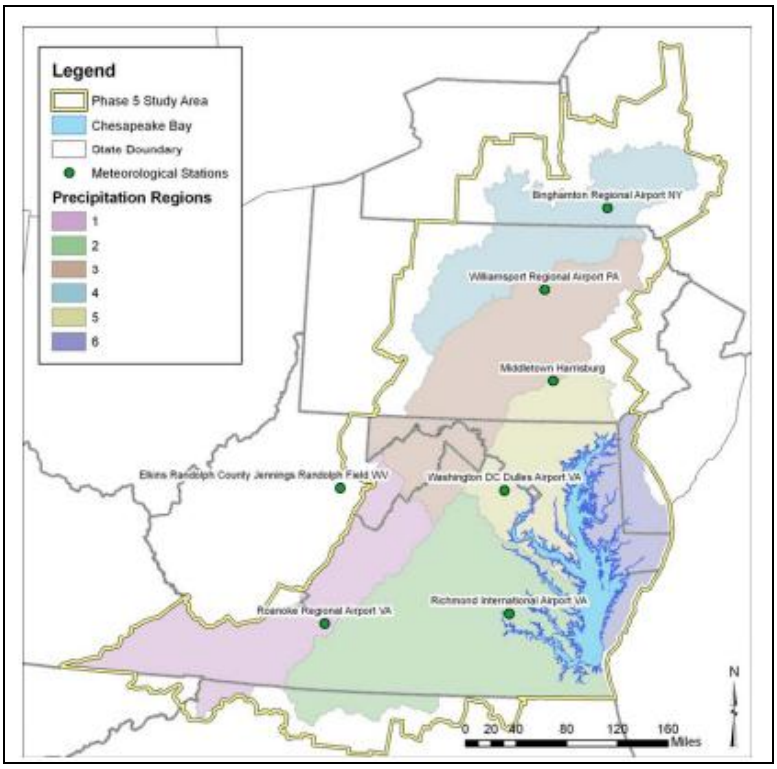


Figure 2-2 in USEPA(2010). The six precipitation regions used to develop the monthly MLRs used in the precipitation model.

To estimate daily precipitation for each land-segment, the following procedure was used: (1) mean daily precipitation (p) and corresponding mean latitude, longitude, and elevation (x, y, z) values from a selected station set (determined using an Exhaustive Search analysis) were used with the slopes (b1, b2, b3) of the monthly MLR to compute a unique b0 for that day; (2) the MLR equation was then solved using the x, y, z values of points on a 5-km grid; and (3) these gridded estimates were integrated over the land-

segment area. The process used for the precipitation model is graphically represented in Figure 2-3.

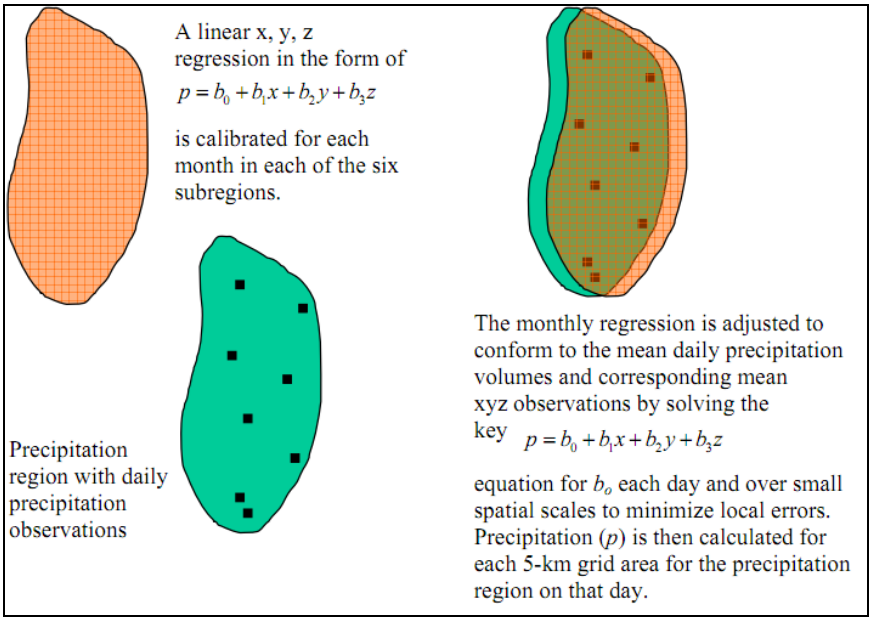


Figure 2-3 in USEPA (2010). Graphical representation of the precipitation model.

Generation of Hourly Rainfall Records

The daily rainfall records were used to derive the daily volume of precipitation. The volume was then disaggregated to hourly values for the land-segment (usually a county) using a nearest neighbor approach applied to about 200 hourly precipitation observed stations across the Phase 5.3 domain. Although there were about 200 hourly stations in the two decades of the data set, usually only about 10 hourly stations would be working on any one day. For that reason, the search pattern had a wide cast to capture hourly stations to disaggregate the daily rainfall data.

In the final precipitation data for the hourly disaggregation of the daily precipitation stations, 57 percent of the stations were disaggregated using an hourly station 100 km from a daily station with the precipitation volume within 100 percent of calculated daily

volume. Relaxing the distance constraint allowed an additional 26 percent of the daily stations to be disaggregated to hourly estimates. Relaxing both the distance and volume constraints allowed an additional 17 percent of the daily estimated precipitation estimates to be disaggregated. Finally, very few hourly stations (0.3 percent) were unresolved even with distance and volume constraints relaxed, and so disaggregation used daily values divided by 24. USEPA (2010)

In summary, the precipitation time series for the CBW model are gauge-based and interpolated to a 5x5 km grid using multilinear regression equations developed by USGS. Such 5 x 5 km grid values are spatially aggregated to the county level and temporally disaggregated to hourly time series.

2.6 Regional Meteorological Data

The hourly meteorological data inputs to the CBW Model include air temperature, wind speed, and solar radiation. According to USEPA (2000), air temperature was prepared based on observations from numerous meteorological stations (Figure 2-3) using a model similar to the precipitation model. Wind speed and solar radiation were compiled from observed meteorological data provided by National Oceanic and Atmospheric Administration's Climatic Data Center.

The entire Chesapeake Bay region is divided into seven primary meteorological regions. Regional meteorological stations are Binghamton, New York; Williamsport, Pennsylvania; Harrisburg, Pennsylvania; Elkton, West Virginia; Dulles Airport, Virginia; Richmond, Virginia; and Roanoke, Virginia (Figure 2-3). Data gaps in the primary

stations were filled by alternate stations or by an adjacent primary station if the data in the alternate station were also missing.

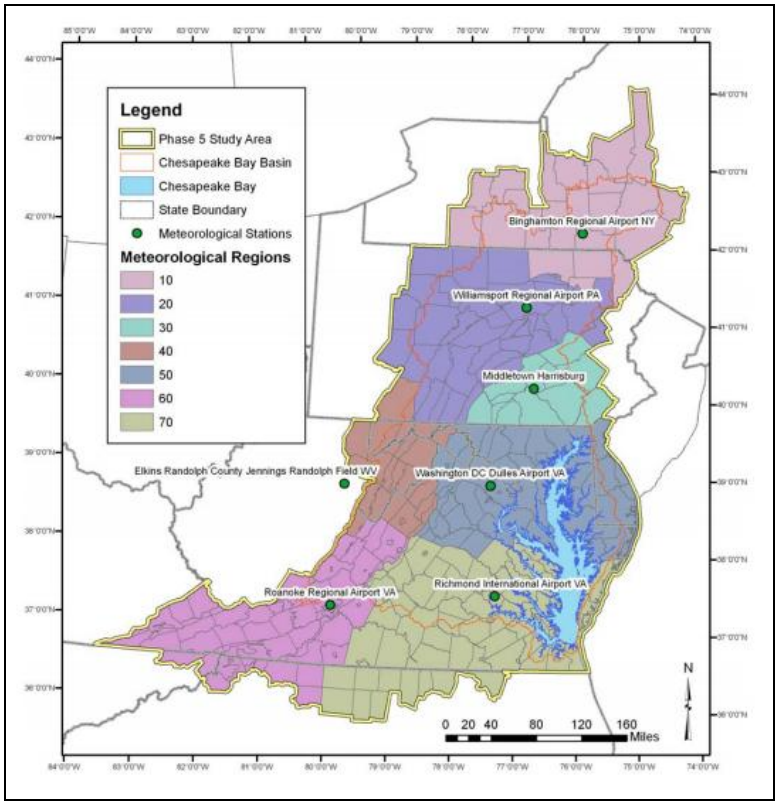


Figure 2-3. Regional meteorological stations and associated land segments. Adapted from USEPA (2010).

For details on the CBW model, readers are referred to the USEPA documentation (2010), and model manual (Shenk, 2010). For details on the HSPF modeling system, readers are referred to the HSPF documentation (Bicknell et al., 2001), and training materials such as lectures, exercises, and technical notes from the BASINS website (USEPA, 2010).

CHAPTER THREE

Comparison of Precipitation Data Sets

3.1 Introduction

Hydrologic simulation and forecasting models often require precipitation input data at a fine temporal resolution, and often fine spatial resolution as well. *In situ* point (gauge) precipitation measurements are generally held to be the most accurate, but may not be an appropriate representation of a large area. Remote (radar) estimates provide excellent spatial and temporal coverage, but with variable accuracy. Given that different types of precipitation data are available to model users, understanding the implications of using one type of input and interpreting model results correctly are critical steps in modeling. Knowing the characteristics of each type of precipitation data should help model users address these steps.

The goal of this chapter is to understand the statistical characteristics of the two types of precipitation data in space and time. This study compared hourly precipitation data from different observing/reporting systems, both *in situ* and remote. Specifically, the Cooperative Observer Program (COOP) gauge-based rainfall measurements (CGRM) and the Automated Surface Observing System (ASOS) gauge-based rainfall measurements (AGRM) were compared with one of the U.S. National Weather Service's NEXt generation RADar (NEXRAD)-based rainfall estimates products, Multisensor Precipitation Estimator (MPE). In addition, the mean areal time series of hourly precipitation previously developed from interpolated gauge data by the Chesapeake Bay Program (CBP) for their lumped model (mCBP) were compared with the mean areal

precipitation time series developed from MPE pixel data by the author for this study (mMPE).

Gauge- and radar-based precipitation data were compared to better understand their statistical characteristics from the model user's perspective. The study addressed the following questions:

(1) How well do hourly COOP/ASOS gauge (point) and MPE (pixel) time series agree?

(2) What information can be gained from MPE data during periods of invalid records in COOP/ASOS, and vice versa?

(3) How well do the CBP's interpolated and areally aggregated gauge precipitation time series (mCBP) agree with areally aggregated MPE time series (mMPE) at a county scale?

Each data set has its strengths and sources of errors; therefore, none can be treated as "truth" to evaluate or correct the other without thorough examination. A comparison of data will shed light on issues that should be addressed when using either or both data sources in hydrologic modeling.

3.2 Background

Precipitation data with a fine temporal resolution are critical not only for forecasting floods, but also for predicting patterns of heavy rains in response to climate changes (Kim, 2008). Precipitation data with a high temporal resolution are vital for studies on the 'contaminant first flush' effect during storm events (Ribarova et al., 2008; Debele et al., 2007). Hourly precipitation data (HPD) are required to simulate diurnal

variations in hydrodynamics and water quality processes (Ji, 2008). Although academic research demonstrates the benefits of using HPD, operational applications are constrained by scarcity of the data and data quality issues (Gyasi-Agyei and Mahbub, 2007).

3.2.1 Major sources of precipitation data

A major source of a long-term HPD in the United States is the COOP network, which has been in operation since 1890 (NWSMC, 1998). The COOP network uses Fischer-Porter (F&P) weighing gauges with a paper-punch recording mechanism. Because the F&P mechanism is prone to recording errors and frequent failure (NWSMC, 1998), F&P systems are being replaced with the Fischer-Porter Upgrade (FPU). A modification of the existing F&P, the FPU consists of modern load cell sensors and commercial electronic data loggers (USDC, 2007).

The Automated Surface Observing System (ASOS), maintained by a joint effort of the NWS, the Federal Aviation Administration (FAA), and the Department of Defense (DOD), has been installed at the NWS forecasting offices at airports. Precipitation information at ASOS stations is automatically updated on an hourly basis using heated tipping bucket gauges (NWSMC, 1998). To resolve problems with a heated tipping bucket in measuring frozen precipitation, the All Weather Precipitation Accumulation Gauge (AWPAG) has been implemented (USDC, 2003). The AWPAG measures both liquid and frozen precipitation by weighing total accumulation and converting the weight to an electrical signal (USDC, 2003). Depending on the level of services, ASOS data undergo augmentation processes conducted by human observers (AOPA, 1999). The services provided at ASOS stations are categorized by service level: levels A, B, and C

include human-augmentation processes and level D has a stand-alone ASOS which operates without corrections by human observers (AOPA, 1999).

The ASOS and COOP gauges have measurement resolutions of 0.01 inch (0.254 mm) and 0.1inch (2.54 mm), respectively. In the ASOS system, when the amount of rain collected in a given hour is less than 0.01 inch, "trace of water" is recorded as zero and flagged with 'T'. The COOP devices do not report "trace of water", recording as plain zeros until the accumulated precipitation amounts to 0.1 inch or greater in a given hour.

The Hydrometeorological Automated Data System (HADS) program is operated by the NWS Office of Hydrologic Development in cooperation with the Water Resources Division of the U.S. Geologic Survey, the U.S. Army Corps of Engineers, the U.S. Forest Service, and other agencies. The HADS acquires raw hydrologic and meteorologic observation messages from the Geostationary Operational Environmental Satellites Data Collection Platforms. The HADS processor program converts the raw information into Standard Hydrometeorological Exchange Format products. These products are distributed to river forecast centers and weather forecast offices via the internal NWS communication system, the Advanced Weather Interactive Processing System Satellite Broadcast Network. They are very useful products for the real-time calibration of radar-based precipitation products.

The spatial variability of precipitation is a major driver of watershed and river responses to precipitation events. Yet, except for in a few densely instrumented watersheds, adequate tools for assessing spatial characteristics of rainfall were not available until the advent of radar-based rainfall estimates (RRE). It is, therefore, natural that interest in using RRE has increased immensely in the water resources field.

Academic and operational research subjects in recent years include RRE algorithm development (Seo et al., 2009; Silvestro et al., 2009; Kuzmin et al., 2008; Giangrande and Ryzhkov, 2008; Tadesse and Anagnostou, 2005; Habib et al, 2004; Krajewski and Smith, 2002; Seo and Breidenbach, 2002; Seo et al., 2000 and 1999), RRE accuracy (Wang et al., 2008; Rogalus III and Ogden, 2007; Yilmaz et al., 2005), uncertainty/error analysis (Villarini et al., 2009; Habib et al., 2008; Zhang et al., 2007; Ciach et al., 2007; Borga et al, 2006), and various applications (Ramkellawan et al., 2009; Bedient et al., 2007; Jayawickreme and Hyndman, 2007; Vivoni et al., 2006; Jacobs and Srinivasan, 2005; Knebl et al., 2005). A substantial number of comparison studies have revealed significant differences between RRE and gauge-based rainfall measurements (GRM) (Skinner et al., 2009; Wang et al., 2008; Westcott et al., 2007; Kalinga and Gan, 2007; Xie et al., 2006; Jayakrishnan et al, 2004). Various approaches of obtaining optimal precipitation data have been suggested, including a concurrent multiplicative-additive approach (Garcia-Pintado et al., 2009), mean field bias correction method (Seo et al., 1999), or merging methods using various schemes (Goudenhoofdt and Delobbe, 2009; Skinner et al, 2009; Kalinga and Gan, 2007; Sokol, 2003; Sun et al., 2000).

In response to the necessity of reducing discrepancies between RRE and GRM, while capitalizing on the strengths of each, the NWS Office of Hydrologic Development (OHD) developed and implemented an algorithm to produce a new RRE product called the Multisensor Precipitation Estimator (MPE). The product is generated by merging ground-based radar rainfall estimates (NEXRAD), rainfall data from rain gauges, and satellite-based rainfall estimates (Kondragunta and Seo, 2004). The satellite-based rainfall estimates in this product refer to Satellite Precipitation Estimates produced from

the NESDIS-HydroEstimator algorithm (Kondragunta et al., 2005). The reporting precision for MPE is 0.01mm (Young et al., 2000).

Based on the algorithm, preprocessing, calibration, and quality control, RRE data can be categorized into several product levels (Wang et al., 2008; Young and Brunsell, 2008; Xie et al., 2005; Young et al., 2000; and many more). Before MPE became operationally available in the early 2000s, the most commonly used RRE product was the NEXRAD Stage III. A distinct difference between the NEXRAD Stage III and MPE is that the latter was developed by merging NEXRAD, other satellite-based rainfall estimates, and GRM. Types of GRM involved in the MPE processing include the ASOS network and the HADS precipitation data.

3.2.2 Uncertainty

All rain gauge data are subject to both systematic and random errors, including instrumental, wind-induced, sampling, evaporation-loss, and dynamic calibration-mechanistic errors (Wang et al., 2008; Molini et al., 2005; Chang and Harrison, 2005; Ciach, 2003; Habib et al., 2001; Humphrey et al., 1997). The error due to wind effects is known to be the largest, and methods to reduce wind-related errors have been well documented (Habib et al., 2008; Sieck et al, 2007; Sevruk and Nespor, 1998; Legates and Deliberty, 1993).

Besides these common types of errors, the hourly COOP data are subject to inconsistencies and inaccuracies due to partial reliance on manual intervention. The F&P gauge devices automatically record precipitation on paper tapes. Once a month, the paper tapes are changed manually and mailed to the National Weather Service (NWS)

local offices (NWSMC, 1998). In a FPU gauge, the amount of precipitation is automatically recorded onto a digital tape, instead of a paper tape. Each month, observers need to download data from a digital logger to the Data Key and mail it to the agency (USDC, 2007). By soliciting volunteer help, agencies including the NWS risk inaccuracies because volunteers may lack accountability and knowledge about information gathering and device maintenance (Daly et al., 2007).

Table 3-1. Sources of errors for estimating precipitation by weather radars. Quoted from Legates (2000).

Error type	Problem	Bias
Reflectivity measurement	Ground clutter contamination	overestimates
	Anomalous propagation	overestimates
	Partial beam filing	underestimates
	Wet radome attenuation	underestimates
	Incorrect hardware calibration	unsystematic
Reflectivity and rainfall relationship (Z-R)	Variations in dropsize distributions	underestimates
	Hail, mixed precipitation, snow events	overestimates
Below beam effects	Strong horizontal winds	unsystematic
	Evaporation of falling precipitation	overestimates
	Coalescence below the radar beam	underestimates

A major source of errors in the RRE products is systematic bias stemming from significant limitations inherent with the use of radar in estimating precipitation (Hunter, 1996). For example, the overshooting of precipitation by the radar beam often causes the underestimation of precipitation. Legates (2000) summarized the possible sources of errors in three categories (Table 3-1). Depending on these biases, the effective coverage of a radar station, which is defined as the spatial range where radar can detect precipitation consistently, varies from one station to another. To reduce such errors, the

procedure of producing MPE includes the delineation of the effective coverage of radar, mean field bias correction or local bias correction, and mosaicking precipitation estimates from multiple radars (Seo et al., 2010).

3.2.3 Lack of reference data

Although biases may exist in the data reported from the COOP Network stations, they have rarely been characterized. The major barrier to characterizing them is a lack of reference data for the comparison. The best available reference data might come, for example, from neighboring stations. However, using data from nearby stations may introduce another source of bias from ignored spatial variability.

It is also difficult to quantify the uncertainties of MPE. Since the MPE processing integrates various sources of rainfall data, limited independent measurements are available for evaluating the accuracy of the MPE product. The COOP daily precipitation measurements are widely perceived as independent from the development of MPE. Published comparison studies have used the COOP daily data as a reference tool to validate MPE (e.g., Young and Brunsell, 2008). The MPE hourly data are usually summed and compared with the COOP daily data. However, the two data sets are not necessarily independent: the COOP hourly precipitation data can be utilized in the MPE processing if a gauge is equipped with telephone (Limited Automatic Remote Collector, LARC) or satellite (Data Collection Platform, DCP) telemetry (Rohli and Vega, 2007; personal communication with MARFC). The LARC/DCP reporting system allows the field data to be available in real time to meteorologists in river forecast centers. In such cases, the COOP measurements affect the MPE values and no longer constitute

independent measurements for statistical comparison. The COOP HPD from gauge stations without telemetry systems are reported to the NWS once a month by observers. Therefore, these records are not available in real time; they are independent from the MPE processing and can be used to validate the MPE product. The COOP HPD archived at NCDC include data from gauge stations both with and without LARC/DCP telemetry systems. Since the information on the status of such devices is not available to the public through the NCDC database, thorough examinations are required to identify the stations that are appropriate for direct comparison with MPE.

A lack of information on the ground truth complicates an assessment of uncertainties in precipitation data. This study did not treat any data source as "truth", but compared the characteristics of each type, with and without spatial aggregation.

3.3 Study Site

The Potomac River Basin (PRB), the area of interest for this study, is a tributary of the Chesapeake Bay, the largest estuary in the United States. The PRB comprises 22% of the Chesapeake Bay watershed and provides the second largest source of fresh water to the Bay. It encompasses an area of approximately 37,800 km^2 (14,600 mi^2), lying in five physiographic regions: Appalachian Plateau, Valley and Ridge, Blue Ridge, Piedmont Plateau, and Coastal Plain. The basin lies across parts of the District of Columbia and four states: Maryland, Pennsylvania, Virginia, and West Virginia (Figure 3-1). The average annual precipitation ranges from 36 to 43 inches over the majority of the study area (Tiruneh, 2007). The population of the basin is approximately 5.8 million (as of 2005) and has increased by 8% since 2000. The population density varies greatly across

the region, with three-quarters of the population residing in the Washington, D.C., metropolitan area. As population grows at a spatially varying rate, it is a challenge for water resources planners and managers to implement uniform regulatory policies at a regional scale. Local policies are needed to maintain the balance between meeting local water-use demands by increased population and avoiding the over-withdrawal of water. To satisfy various levels of stakeholders' interest in carrying out local and regional policies, this region needs scientific information and support. Hydrologic simulation models are important decision-making tools to meet such needs.

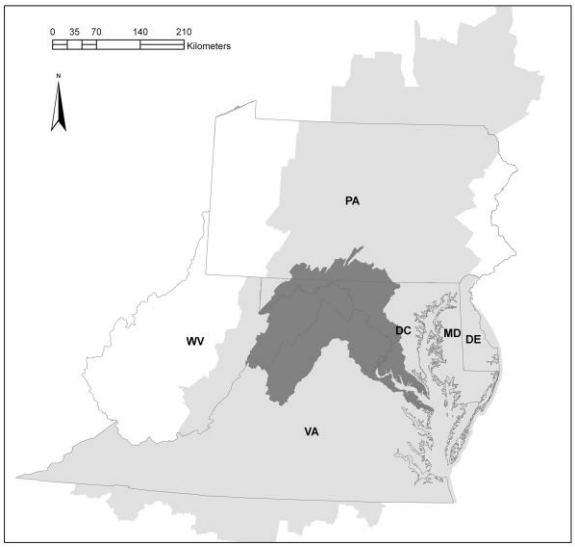


Figure 3-1. Potomac River Basin (dark gray). State boundaries are shown with the Chesapeake Bay region.

3.4 Data

This section describes the data types and processing procedures used in the comparison study (Table 3-2). At the time of the data preparation for this study, the MPE product was available as complete yearly sets from 2001 to 2007. Therefore, the study time period for the comparison of gauge-point and MPE-pixel pairs is 2001-2007. Since mCBP was prepared using gauge-based measurements during 1985-2005 by the CBP, the

study time period for the comparison of the mCBP-mMPE pairs is the overlap, 2001-2005.

Table 3-2. Summary of the data types evaluated in this study.

DATA Acronym in this study	Type	Source (agency)	Description	Study period
CGRM	COOP	NCDC	Gauge-point	2001-2007
AGRM	ASOS	NCDC	Gauge-point	2001-2007
MPE	MPE	MARFC[1]	NEXRAD radar-4x4 km^2 pixel	2001-2007
mCBP	COOP and others	CBP	Gauge interpolated to 5x5 km^2, then aggregated at a county scale	2001-2005
mMPE	MPE	N/A	MPE aggregated at a county scale	2001-2005

1. MARFC: Middle Atlantic River Forecast Center

3.4.1 Gauge rainfall measurement (GRM): COOP/ASOS

For the analysis of GRM, seventy-eight COOP hourly gauge stations were selected within the PRB and its 110-km bordering region. These stations were screened, and those stations with records for the study period, from 2001 to 2007 are selected (70 sites). The area on the south side of the PRB was extended due to the sparse density of the COOP network (Figure 3-2). Precipitation records for the selected stations were obtained from NCDC (2009). ASOS data sets are available under COOP Network ID numbers; they were obtained as are the COOP data sets. Table 3-3 summarizes the gauge stations employed in this study. Twenty of the seventy COOP hourly gauge stations have either LARC or DCP telemetry (Table 3-4). CGRM from these stations are available in real time and could be utilized in the MPE processing. Because the LARC/DCP telemetry systems are reporting systems, not measurement systems, they are not expected to affect measurement characteristics or the accuracy of CGRM. Therefore, the

telemetry-equipped COOP stations were analyzed in the same way as the other COOP stations.

Table 3-3. Hourly rain gauge stations used in this study by state and type.

State	ASOS	COOP	Total
DE	1	1	2
MD	1	5(1)	6(1)
PA	1	31(4)	32(4)
VA	6	17(3)	23(3)
WV	1	17	18
Total	10	70(8)	80(8)

Note: The number of parentheses indicates the number of stations excluded because of incomplete records over the study time period

Table 3-4. COOP gauge stations with the LARC/DCP telemetry system.

State	Station ID
DE	076410
MD	188065
PA	360785, 362569, 362838, 363321, 364214, 364397, 364896, 365825, 366066, 366852, 368846, 369022
VA	444128, 445880, 446712, 449159
WV	461323, 466163
Total	20

Note: This information was obtained through personal communication with MARFC (Middle Atlantic River Forecast Center)

The COOP/ASOS network HPD were processed from NCDC's alphanumeric online format into numerical single column ASCII precipitation time series for each site. During the format conversion process, data values with text flags indicating "deleted", "missing periods", or "trace of water" in the NCDC data were replaced with specific numerical values which are outside the range of possible data values.

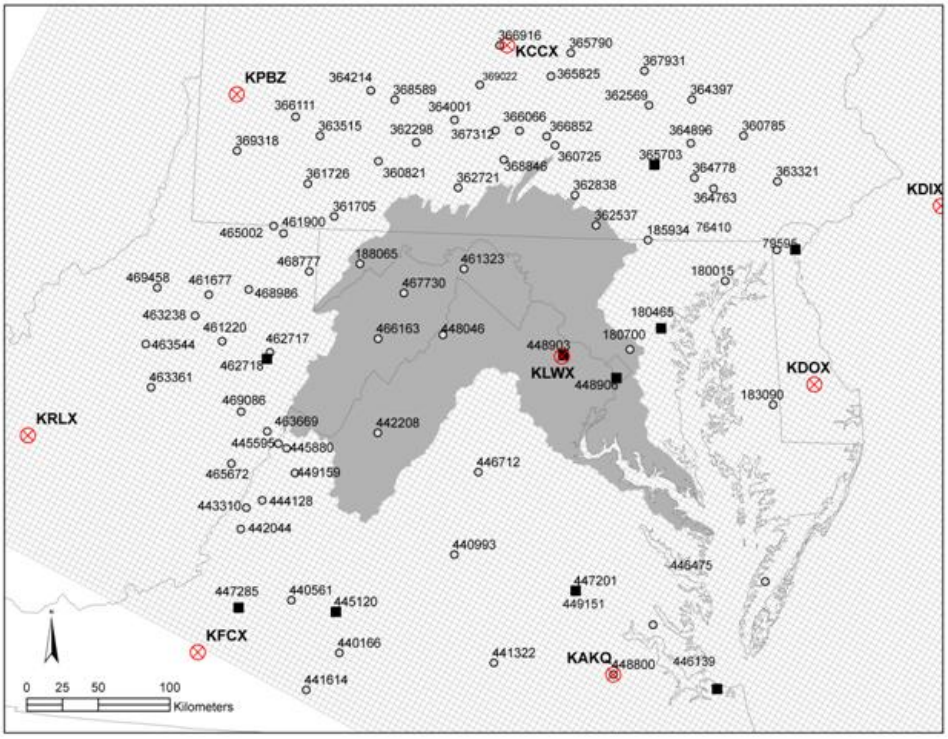


Figure 3-2. Distribution of rain gauge stations in the UTM-reprojected HRAP grid system: ASOS (■) and COOP (○). MPE (one value per mesh grid cell) created by the NWS using data from NEXRAD radar stations (⊗).

3.4.2 Radar-based rainfall estimates (RRE): MPE

The data were obtained in the binary format, XMRG, from MARFC. The data were delivered as a single tape archive (tar) file that contained gridded hourly values for seven years, from 2001 to 2007. They were processed in batch mode using the Korn shell-scripting language, C programming language, and Visual Basic for Application (VBA) on Linux and PC-Windows platforms. The process largely followed steps described in Xie (2005), except for a few technical modifications that resulted from employing different hardware and software.

3.4.3 Georeferencing and synchronizing time series

The coordinate system defining the geographic location of MPE is the Hydrologic Rainfall Analysis Project (HRAP) grid system (Reed and Maidment, 1999). Each MPE value represents a spatially averaged rainfall estimate over a given 4 km-by-4 km grid cell. The COOP data time series are referenced to point locations using the geographical coordinate system (latitude, longitude). The two data sets were mutually georeferenced using the Universal Transverse Mercator (UTM) system, which is commonly used for GIS-based analysis and modeling at a regional scale. Using a series of GIS tools, the HRAP grid was reprojected to the UTM Zone 18 North coordinate system (Figure 3-2). Figure 3-2 also shows COOP/ASOS stations plotted in the UTM coordinates projected from their latitudes and longitudes.

The spatial distribution of COOP and ASOS hourly rain gauge stations is not systematic: the shortest distance between any stations is 5.03 km. The shortest distance between any COOP-ASOS station pair is 5.10 km. Two rain gauge stations separated by such distance could be within the same HRAP grid cell, considering the 5.66 km diagonal length of a 4x4 km^2 grid cell. A visual screening, however, confirmed that each rain gauge station corresponds to a unique HRAP grid cell.

Each MPE value was extracted from each grid cell corresponding to the site of a COOP/ASOS station at every hour. Thus a 7-year MPE time series was generated at each site. Comparing MPE time series with their corresponding COOP/ASOS time series requires a time shift. The MPE time series is recorded in Coordinated Universal Time (UTC), while the COOP/ASOS time series is reported in local time, Eastern Standard Time (EST). To make the two rainfall time series concur, UTC was converted to local

time, consistent with hydrologic modeling and decision-making practices. By pairing MPE and GRM at an hourly scale in EST on the UTM spatial reference system, eighty paired time series were generated.

3.4.4 Mean areal precipitation: mCBP and mMPE

After comparing gauge-point values with MPE-pixel values, this study compared the spatially aggregated data sets. The mCBP data set from 1985 to 2005 was developed by the USGS for the CBW model (Section 2.5); this data set is currently used in the CBP model runs. This study refers to it as "mCBP". For the comparison in this study, the MPE pixel values were aggregated to the same spatial scale as mCPB. The GIS software was used to aggregate the pixel data to the county-scale land segments used by the CBW model. The spatially aggregated MPE data set is referred to as "mMPE". This study chose 114 land segments within/near the PRB for the comparison of mCBP-mMPE pairs (Figure 3-3).

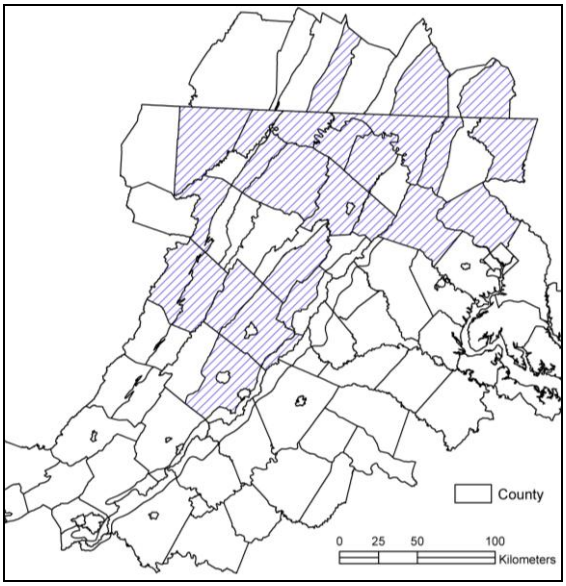


Figure 3-3. Spatial aggregation of MPE-pixel precipitation conducted to create mMPE at a county level. The mCBP data set is available at the CBP website. 26 land segments (hatched polygons) were selected for the comparison of storm-event pairs (Section 3.5.7).

3.5 Methods

3.5.1 Missing data

In this study, missing data in the gauge time series were identified and examined. They were not interpolated, estimated, or otherwise filled: they were excluded from quantitative statistical analyses of the station time series. Missing data in MPE-pixel were filled in by experts in the hydrology and meteorology fields before the data sets are made available to the public. Before the mMPE time series were created in this study, any missing values, if any, were filled by repeating preceding values. The mCBP time series were filled in by the data developers in the interpolation process (Chapter 2) and did not have missing data.

3.5.2 Comparison of collocated time series

This chapter first compared collated gauge-based and NEXRAD-based precipitation data at a point/pixel scale (COOP, ASOS/MPE). For the comparison, statistical measures for distributions and differences were used. Then, gauge-based and NEXRAD-based precipitation data were compared at a county scale (mCBP and mMPE) using statistical measures for moments and distributions over the study period. This study used descriptive statistics to identify similarities and differences among the precipitation data sets before they were used to drive hydrologic models.

3.5.3 Comparison tools for gauge-point and MPE-pixel data

Average annual precipitation values were calculated for each study location from GRM and MPE time series. The accumulated precipitation depths based on GRM and

MPE were plotted against each other over the seven years at a given site, *k*. The plot creates a double mass curve, commonly used to identify the consistency of two datasets over time (Searcy and Hardison, 1966).

The frequency distributions of hourly rainfall intensity were calculated and plotted. The relative frequency of rainfall intensity was calculated as follows:

$$\text{Relative frequency}\,(B) = \frac{n(B)}{T} \text{ x } 100 \tag{3-1}$$

where

B = bin for each rainfall intensity in increments of 0.01 inch (0.254 mm). The endpoint convention is to exclude the lower boundary and include the upper boundary.
T= number of study hours

It is useful to quantify the amount of precipitation that is contributed by individual intensities. Intensity-based cumulative precipitation (ICP) is defined as the total depth of precipitation contributed by hours with intensity less than or equal to a specified value:

$$ICP(x) = \sum_{i=1}^{n(x)} x_i \tag{3-2}$$

where

x= the precipitation intensity [mm/hr]
i =the index (from 1 to *n*) when the observations are ranked from smallest to largest
n(x)=the number of hours with precipitation intensity less than or equal to *x*

Mathematically, Eq. 3-2 is known as a partial sum (Weisstein, 2011).

Root mean square errors (Steiner et al., 1999) focus on the differences between GRM and MPE at an hourly scale. In this study, RMSE is renamed root mean square difference (RMSD), since both data types represent estimates and neither data type is the "true" value

$$\mathrm{RMSD}_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{g}_{i,k} - \mathrm{r}_{i,k})^2} \quad (3\text{-}3)$$

where

i = hour index
n = number of hours
$g_{i,k}$= CGRM or AGRM [mm] during hour i at gauge location, k
$r_{i,k}$ = MPE [mm] during hour i in HRAP grid cell corresponding to gauge location, k

For each station, the RMSD was calculated for conditional as well as unconditional pair selection methods. The conditional RMSD assesses the agreement for all of the hours in which both systems report nonzero precipitation, while the unconditional RMSD assesses when at least one system reports nonzero.

3.5.4 MPE as a reference tool

The GRM missing data or data recorded as zero were examined under several scenarios, including cases where a value of zero is recorded with a "trace-of-water" flag. The MPE were extracted and cumulated over 7 years (2001-2007) for all hours in which GRM contains invalid (missing or deleted), zeros of water, or trace of water (Equation 3-4). Since each station has different percentage of the record completion, the percentage of the total hours was also calculated.

$$\mathbf{R}_{k,m} = \sum_{i=1}^{n} r_{i,k} \quad if\ (g_{i,k} = m) \quad (3\text{-}4)$$

where

i = hour index
k =index for either COOP or ASOS gauge station
n = number of hours
$g_{i,k}$= CGRM or AGRM [mm] during hour i at gauge location, k

$r_{i,k}$ = MPE [mm] during hour i in HRAP grid cell corresponding to gauge location, k
m = invalid (missing, deleted), trace of water, or zero of water

3.5.5 Comparison of spatially aggregated data sets

For the comparison of mCBP and mMPE, average annual values for the study period (2001-2007) at each land segment were calculated. The relative difference between the two types of data was also calculated. Since mCBP is currently used in the CBW modeling community, it was treated as a standard and used as the denominator. At a given land segment,

$$\text{Relative difference} = \left[\frac{\sum_{i=1}^{n} x_i^{mMPE} - \sum_{i}^{n} x_i^{mCBP}}{\sum_{i=1}^{n} x_i^{mCBP}} \right] \quad (3\text{-}5)$$

where

i = hour index for all paired values
x =the precipitation depth [mm] during hour i from mMPE or mCBP

3.5.6 Comparison of storm events

Precipitation events from mMPE and mCBP time series were extracted to examine the characteristics of storm events (peak time, duration, peak depth, and event total) using both unpaired and paired storm events. To reduce the computational demand in this analysis, precipitation events were extracted from 26 county-based model segments (hatched polygons in Figure 3-3), a subset of the 114 land segments over the study area. Those 26 segments were used to simulate the most downstream river segment for any given calibration site in the model calibration process described in Chapter 4.

Following Habib (2009), a precipitation event was defined as "a continuous raining period interrupted by no longer than 6 hours of no rain".

3.5.6.1 Unpaired mMPE and mCBP precipitation events

The characteristics of unpaired storm events were compared using the empirical cumulative distribution function (ECDF). The ECDF is a non-parametric tool for quantitatively describing the distribution of a data set or comparing the distributions of multiple data sets. The ECDF, *F(x)* is defined as follows:

$$F(x) = \frac{n(x)}{N} \tag{3-6}$$

where

n(x)= number of precipitation events with duration (or peak depth) less than or equal to *x*
N = number of total precipitation events

3.5.6.2 Paired mMPE and mCBP precipitation events

For the paired precipitation event comparison, precipitation events were selected only if the total depth per event is at least 5 mm or 0.19685 inches (Habib, 2009). Then, the mCBP and mMPE events were paired using a 24-hour time window. Time windows are the time frames used for pairing mMPE- and mCBP-precipitation events. A paired event with a 6-hour window, for example, refers to a pair of mCBP and mMPE events that start within 6 hours of each other. Except for the case where a single mMPE event corresponded with a single mCBP event within the 24-hour time window, all of the cases were programmatically re-screened to make one-to-one pairs using a user-defined rule, checking the difference in the starting time of an event, selecting the two events that had the smallest difference in their starting time. Once one-to-one relationships between

mCBP and mMPE events were established, paired mCBP and mMPE were categorized based on different sizes of time window: 12-, 6-, 3-, and 0-hour. The effect of various time windows on the paired storm characteristics was also examined using descriptive statistics.

3.6 Results

3.6.1 Missing data

Invalid (unknown or missing) CGRM values occur randomly, with the COOP stations reporting a wide range of percentages of invalid CGRM at different times during the study period. Of the 10 ASOS sites, only Site 180465 shows invalid values. It has 744 hours of invalid AGRM from August 1, 2006, to August 31, 2006. Based on the pattern of invalid values, the cause is most likely related to mechanical problems. In the MPE records, missing values are rarely found because the majority of missing data was filled in by experts in the hydrology and meteorology fields before being made available to the public. There are, however, a few periods of invalid values in the MPE data set. For one hour beginning at 3 am on March 8, 2001, MPE records are concurrently invalid across six separate study locations. Similarly, for two hours from noon to 2 pm on March 3, 2004, seven study locations show invalid values.

To quantify the magnitude and effects of invalid CGRM, cumulative MPE records were extracted. The cumulative MPE occurring during periods of invalid CGRM amounts to as much as 8000 mm over the seven-year period at COOP station 461900 (Figure 3-4), corresponding to 96% of total MPE (8300 mm) at the site. In addition, the closest two sites in the PRB show a distinctive difference in missing values, indicating

great spatial variability. For example, COOP station 465002 is invalid for 14% of the total hours during which about 1000 mm of MPE are accumulated, whereas COOP station 461900 which has the shortest distance from the site 465002 (approximately 5km away) shows invalid values for 94% of the total study hours, with a corresponding MPE accumulation of 8000 mm (Figure 3-4).

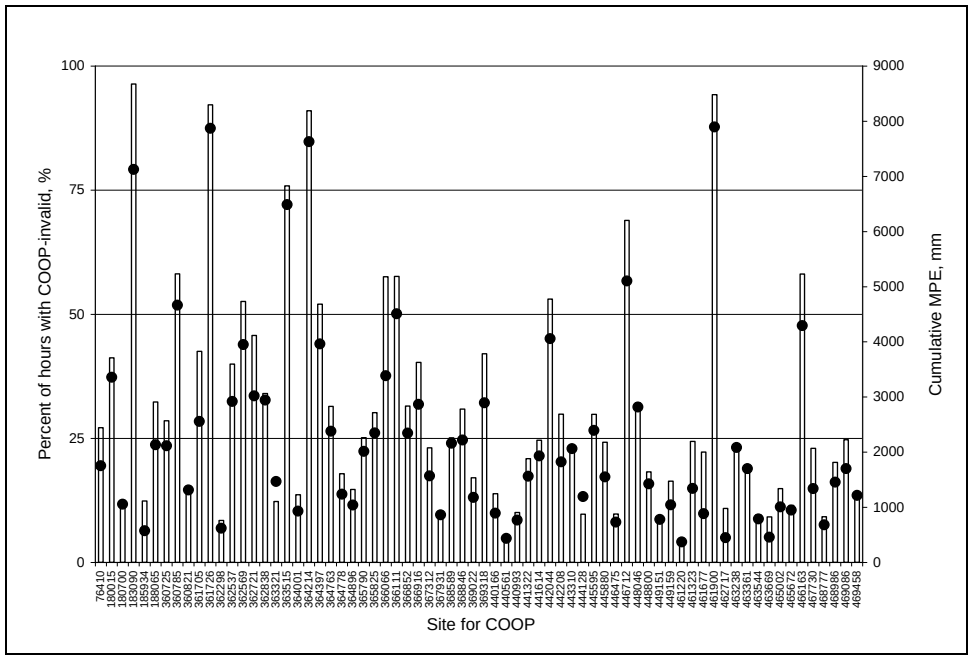


Figure 3-4. Percent of hours with invalid values at COOP sites (▯) and cumulative MPE (●) when COOP has invalid values at each gauge site.

The implementation of the FPU in the PRB area started in May 2005 (Table 3-5). This change in recording devices may have contributed to a change in the frequency of invalid data. One site demonstrates an improvement (COOP site 462717). Three other locations, however, show a higher percentage of invalid values after the FPU implementation, perhaps due to issues during the transition period between the two systems. Two locations show no change in the fraction of hours of invalid values before

and after the upgrade, indicating that the invalid values may have been caused by data management factors during the manual-intervention process, not by the devices.

Table 3-5. Status of FPU implementation (USDC, 2007) at COOP study sites and comparison of hours for invalid values before and after the upgrade.

COOP ID	Start of FPU operations	Before			After		
		Total hours	Hours with invalid records	% invalid	Total hours	Hours with invalid records	% invalid
180015	04/07/06	46128	19036	41.2	15211	6247	41.1
360785	05/22/06	47208	27826	58.9	14131	7823	55.4
446475	08/02/05	41424	3461	8.4	19915	2524	12.7
446712	11/18/05	42768	23689	55.4	18571	18571	100.0
462717	05/20/05	38400	6311	16.4	22939	350	1.5
467730	07/28/06	48816	9672	19.8	12523	4427	35.4

3.6.2 Comparison of gauge-point (GRM) and MPE-pixel time series

3.6.2.1 Average annual recorded precipitation

Figure 3-5 compared average annual recorded precipitation depth [mm/yr] using the different data sources across the region. The average annual reported precipitation ranges widely, from 40mm/yr to 1150mm/yr. As explained in Section 3.5.1, data gaps are not filled: GRM-MPE pairs are discarded if either GRM or MPE is missing or unknown. As a result, the average annual recorded rainfall can be small if the majority of values at a given site are invalid and therefore discarded. For example, the percentages of unknown values at COOP stations 183090 and 461900 during this study time period are 96.4% and 94.2%, respectively. As one might expect, these COOP stations have low average annual values: 38mm/yr and 64mm/yr, respectively (Figure 3-5). The average annual values at ASOS stations demonstrate higher recorded rainfall per year. The

average annual reported MPE pixel and GRM-point precipitation values (when both are valid) match well, lying close to the 1-to-1 line +/- 10% (Figure 3-5).

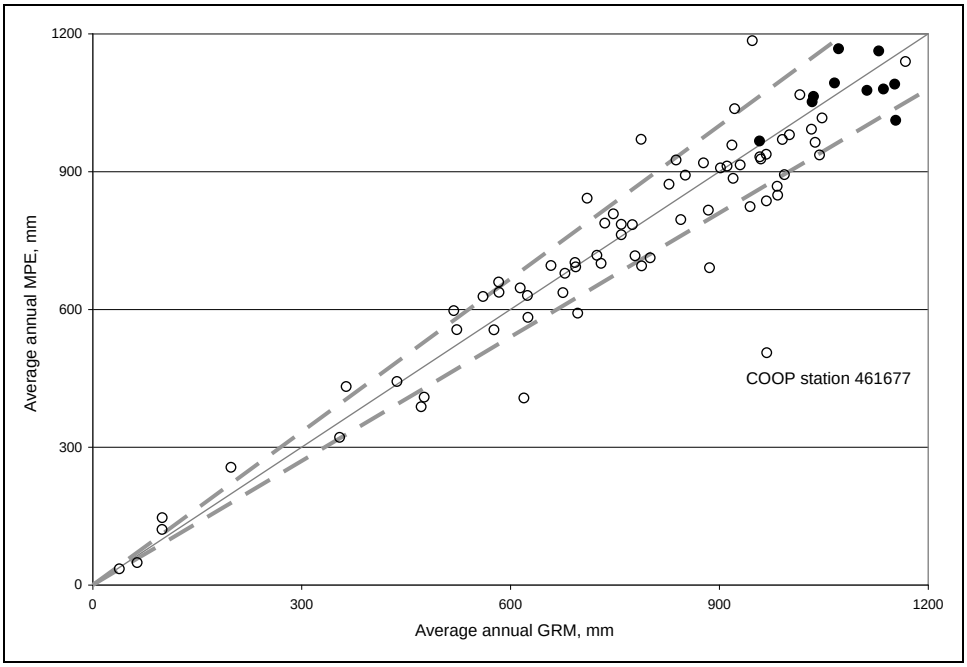


Figure 3-5. Comparison of average annual GRM and average annual MPE at COOP sites (○) and at ASOS sites (●), excluding hours in which either report was invalid or missing. The two dashed lines fall 10% above and below the 1-to-1 line.

The largest discrepancy in average annual precipitation is found at COOP station 461677 (CGRM of 968 mm versus MPE of 506 mm). The neighboring stations of COOP 461677 are COOP 463238, 468986, 461220, 469458. They are 17 km, 28 km, 34 km, and 36 km away from COOP 461677, respectively. The average annual for the (CGRM, MPE) pairs at their stations are (827, 872), (1000, 980), (1176, 1139), and (959, 927). The closer agreement between CGRM and MPE at the neighboring stations suggests that the discrepancy at COOP 461677 is not related to geographical factors such as topography. Considering that the range of the average annual MPE at the neighboring stations is between 870 and 1140 mm, there may be issues at the MPE pixel corresponding to COOP 461677 in estimating precipitation from radar reflectivity.

3.6.2.2 Double mass curve

A double mass curve reveals the status of the pair relationship over time at a given location. For instance, Figure 3-6 (a) illustrates the relationship between cumulative MPE and CGRM at COOP station 440166 over the years of the study time period. The ratio of MPE to CGRM changes when MPE accumulates faster than CGRM, resulting in a slope change from 1 to a value greater than 1. The slope change occurs around November 6, 2003, when the cumulative CGRM is 2331mm and the cumulative MPE is 2455mm. A plot of the hourly paired data set shows that the range of CGRM (0 to 35 mm on x-axis) remains the same before [Figure 3-7(a)] and after 11/06/2003 [Figure 3-7(b)], whereas the range of MPE after 11/06/2003 doubles (y-axis). Considering that the slope of the curve remains constant after the change occurred in 2003, the slope change may be related to the change in algorithms for generating the MPE product, rather than radar operational issues. Accessing and analyzing the algorithms are beyond the scope of this study.

Figure 3-6 (b) illustrates that the slope of the double mass curve is constant with a value less than 1 at COOP 461677, indicating that CGRM exceeds MPE by a consistent ratio over the study time period. This finding is consistent with the discrepancy between COOP-based and MPE-based average annual rainfall shown in Figure 3-5.

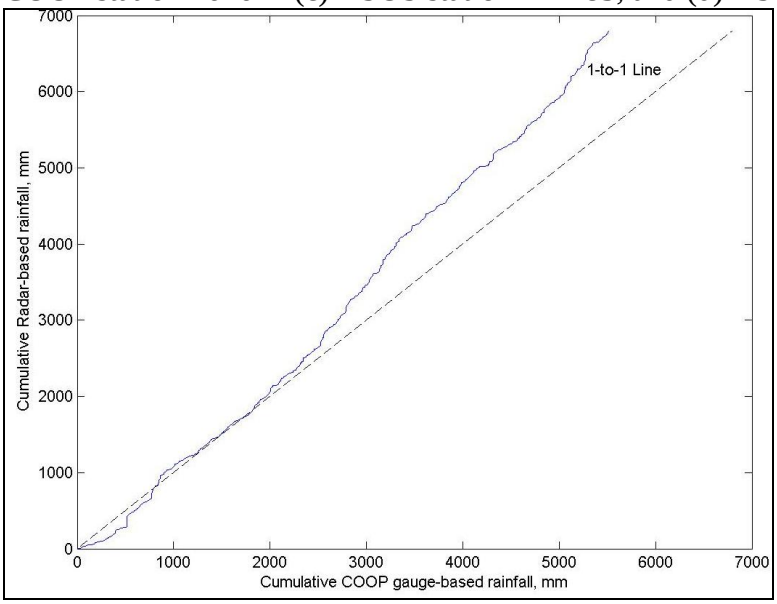
On the other hand, the slope of the double mass curve for ASOS station 447285 in Figure 3-6 (c) is approximately 1, meaning MPE is approximately equal to AGRM throughout the study time period. The slope remains constant before and after AWPAG was installed on August 18, 2004 (Table 3-6), indicating that its installation did not affect the proportionality of the two data sets. Because ASOS data are used to create MPE, one

might expect that ASOS data match well with MPE data, but that does not happen at all of the ASOS stations: at ASOS station 462718, for example, MPE is consistently greater than AGRM [Figure 3-6 (d)]. According to AOPA (1999), ASOS 462718 belongs to Level D (Table 3-6), which means no corrections are made by humans, whereas 447285 belongs to Level B. At ASOS stations with Level B, human intervention is expected to improve their accuracy (Section 3.2.1).

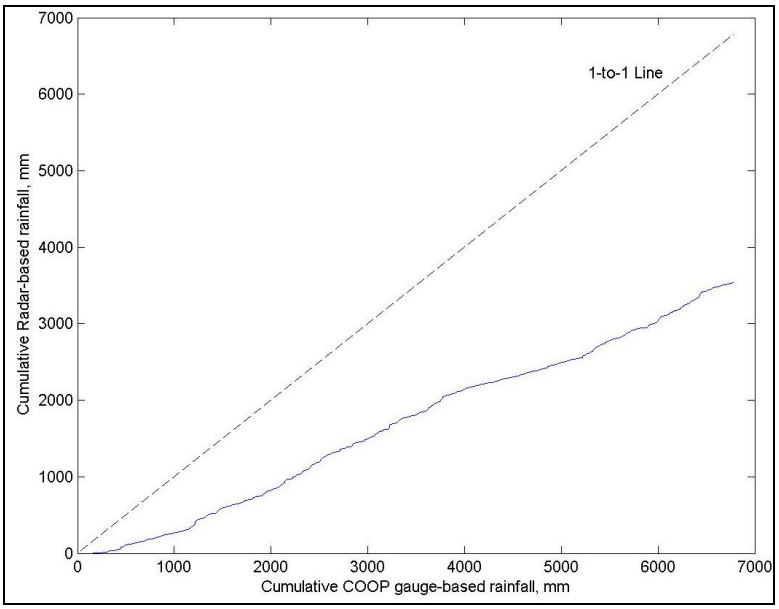
Table 3-6. Status of AWPAG implementation (USDC, 2003) at ASOS study sites.

Station ID	Site Name	AWPAG installed	Level (AOPA, 1999)
079595	WILMINGTON NEW CASTLE CO	8/11/2004	C
180465	BALTIMORE WASHINGTON INT'L AP	12/9/2004	A
365703	HARRISBURG INT'L AP	2/3/2006	B
445120	LYNCHBURG REGINAL AP	7/29/2004	C
446139	NORFOLK INTERNATIONAL AP	7/29/2005	B
447201	RICHMOND INTERNATIONAL AP	7/29/2005	A
447285	ROANOKE REGIONAL AP	8/18/2004	B
448903	WASHINGTON DULLES INT'L AP	7/20/2004	A
448906	RONALD REAGAN WASHINGTON AP	8/17/2004	A
462718	JENNINGS RANDOLPH FIELD	1/10/2005	D

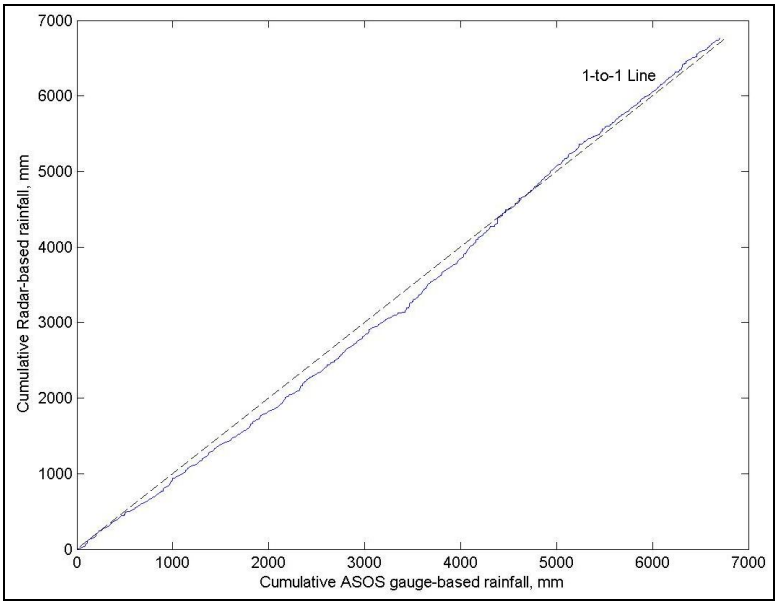
Figure 3-6. Double mass curve based on GRM and MPE at (a) COOP station 440166 (b) COOP station 461677 (c) ASOS station 447285, and (d) ASOS station 462718.



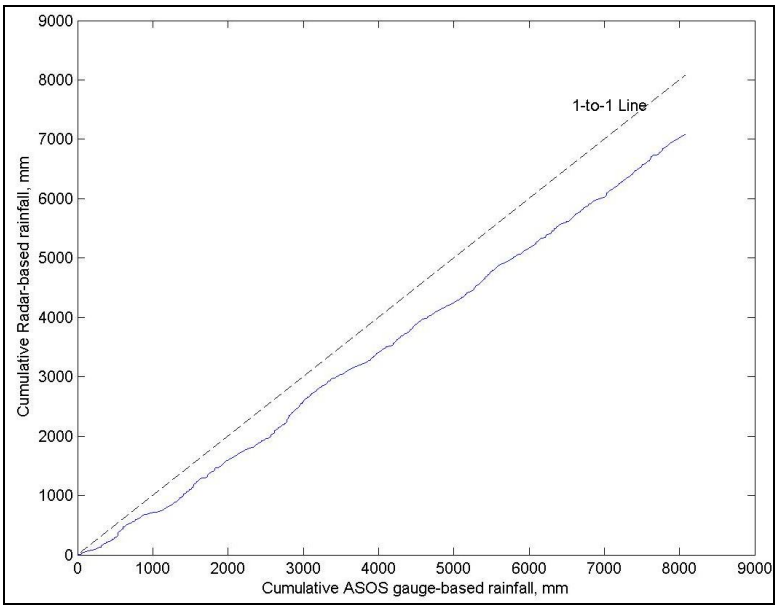
(a)


1-to-1 Line
Cumulative Radar-based rainfall, mm
Cumulative COOP gauge-based rainfall, mm
0
1000
2000
3000
4000
5000
6000
7000

(b)

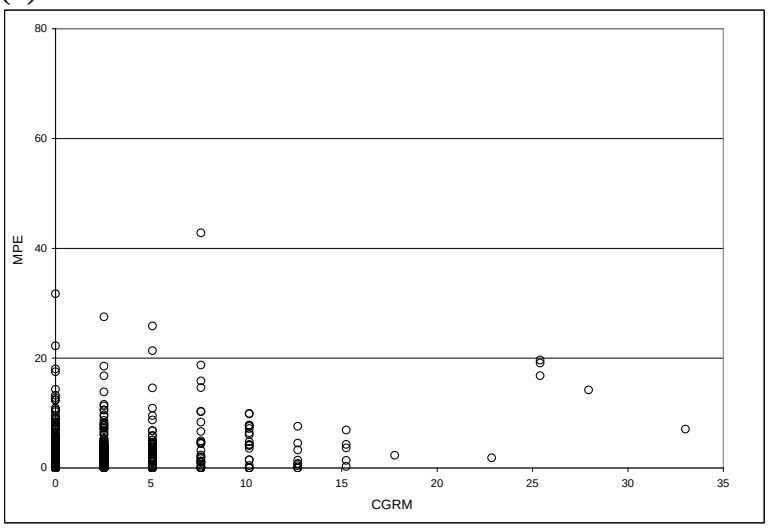

1-to-1 Line
Cumulative Radar-based rainfall, mm
Cumulative ASOS gauge-based rainfall, mm
0
1000
2000
3000
4000
5000
6000
7000

(c)

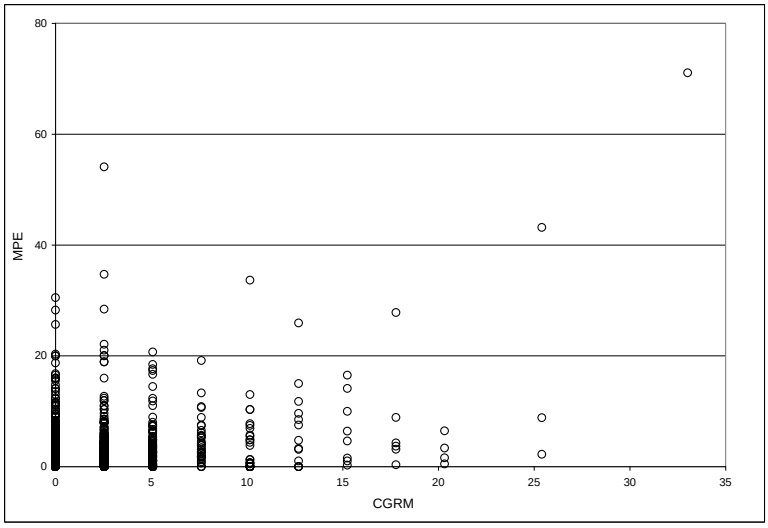


(d)

Figure 3-7. Hourly paired CGRM-MPE at COOP 440166 (a) from 1/1/2001 to 11/6/2003 (b) after 11/6/2003. Units are millimeters.



(a)

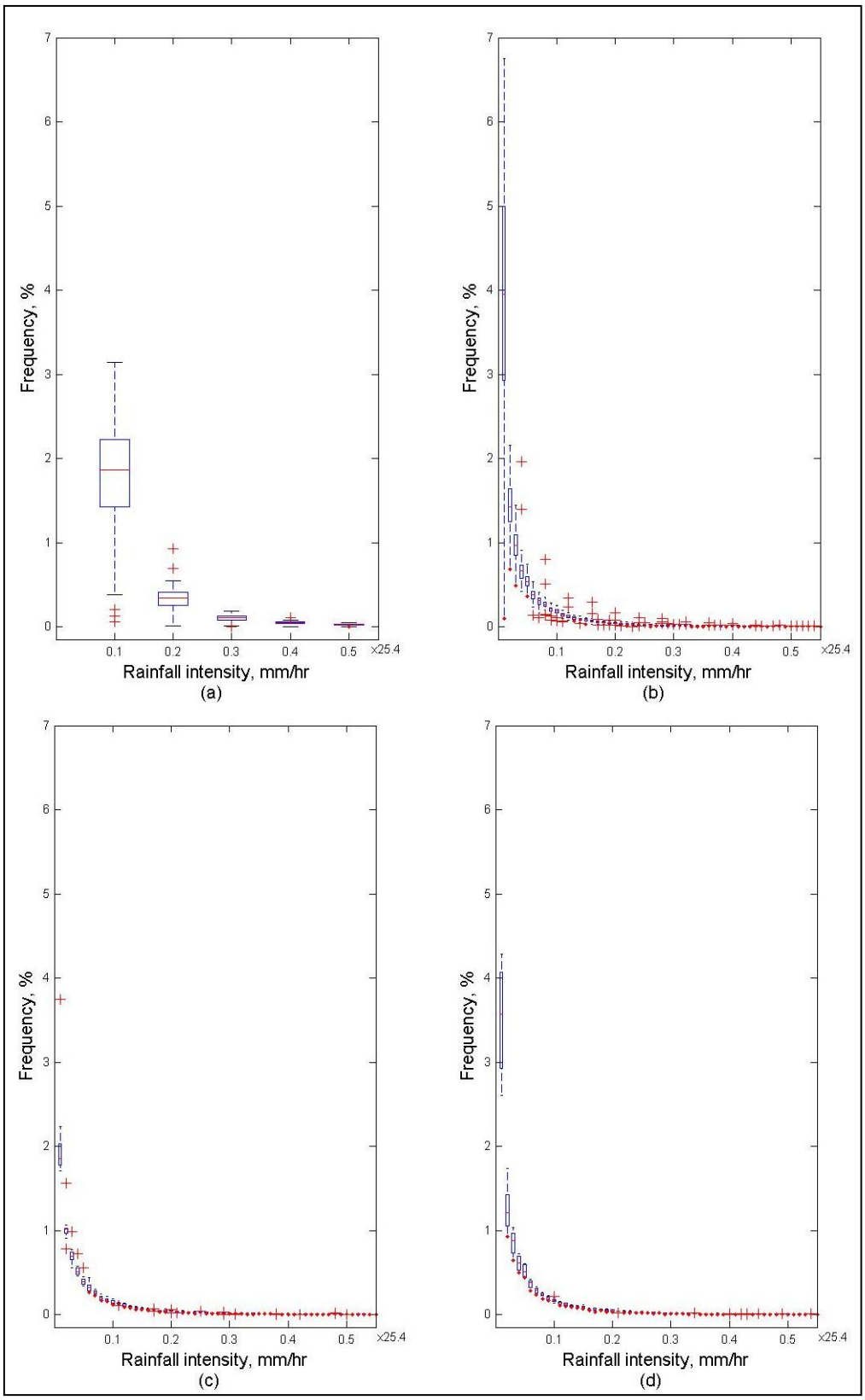


(b)

3.6.2.3 Frequency distribution of rainfall intensity

The effect of the different measurement systems and their resolutions on frequencies of rainfall intensity was examined in Figure 3-8. The frequencies of zero and invalid values, particularly for CGRM, are high, with 85% ~ 90% of total hours being recorded as such. To better display the frequency distribution of nonzero rainfall intensities, the frequency distributions are plotted with the bins for zero and invalid values omitted, and frequencies of rainfall intensity greater than 14 mm/hr (0.55 in/hr) are not shown in the plots.

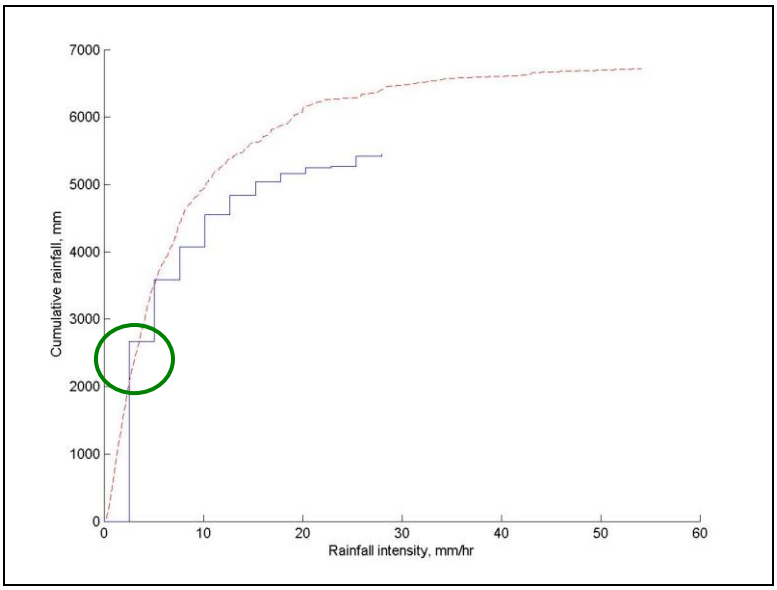
Figure 3-8. Frequency distribution of rainfall intensity. (a) COOP sites (70 samples); (b) MPE from HRAP grid cells corresponding to COOP sites; (c) ASOS sites (10 samples) (d) MPE from HRAP grid cells corresponding to ASOS sites. The lower and upper lines of the each box show the 25th (Q1) and 75th (Q3) percentiles. The limits of whiskers are 5th and 95th percentiles. The "plus" symbols indicate outliers which are defined as values outside of Q3+1.5(Q3-Q1).


Frequency, %
Rainfall intensity, mm/hr
×25.4
(a)
(b)
(c)
(d)

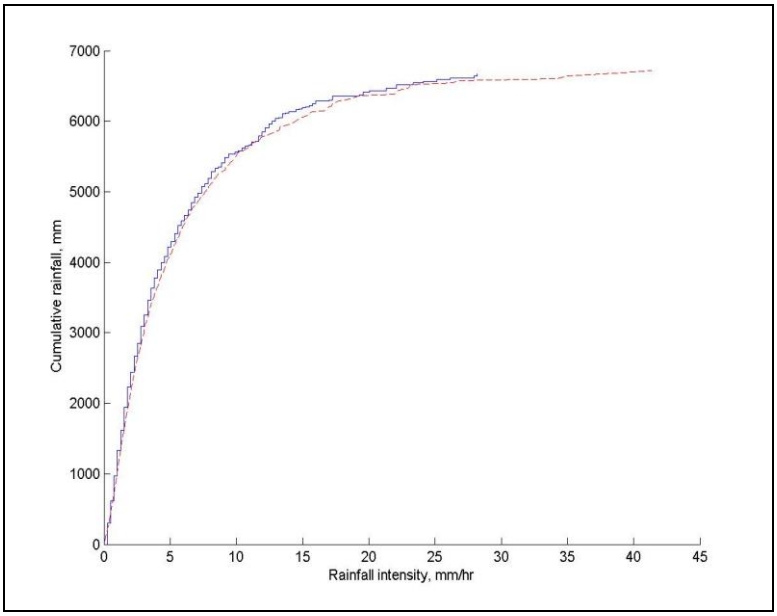
The COOP system does not record rainfall until 2.54 mm (0.1 in) have accumulated. That depth is assigned to the time period in which this measurement threshold is reached, so any hours in which precipitation is actually between 0 and 2.54 mm are included in the 2.54-mm (0.1 in) bin [Figure 3-8 (a)]. This accumulation process appears only in the COOP system. The ASOS system does not have a threshold depth, causing a continuous decrease in the relative frequency, starting at 0.254 mm/hr [Figure 3-8 (c)]. The frequency distribution of AGRM is similar to that of MPE at grid cells containing individual ASOS gauge stations [Figures 3-8 (c) and (d)]. It is noticeable that Figure 3-8 (d) depicts a longer box at a given interval for low intensities than (c); this difference indicates that the variability in frequencies of MPE-rainfall intensity is higher than the variability in corresponding ASOS measurements for the same locations. The boxes and whiskers in the boxplots of Figure 3-8 (b) are longer than those in Figure 3-8 (d) for most rainfall intensity intervals. This suggests that the MPE frequencies corresponding to COOP stations vary from one site to another at a larger degree than those corresponding to ASOS sites.

3.6.2.4 Intensity-based cumulative precipitation function (ICPF)

An intensity-based cumulative precipitation function (ICPF) indicates the total depth of rainfall contributed by hours of intensity less than or equal to a specified value (Figure 3-9). Half of the total depth at COOP station 440166 and ASOS station 447285 was produced during hours with intensities less than or equal to 5mm/hr, and approximately 25% of the total depth in both cases was contributed by intensities less than 2.54 mm/hr, which is the threshold reporting depth for the COOP system.



(a)



(b)

Figure 3-9. Cumulative depth based on rainfall intensity, gauge (—) and MPE (····) at (a) COOP gauge station 440166; (b) ASOS gauge station 447285.

Due to the reporting precision of 2.54 mm, the COOP-ICPF shows a coarser resolution than the ASOS-ICPF does, creating a step function. The COOP-ICP(2.54) is the cumulative GRM at the threshold reporting depth in the COOP system [circle in

Figure 3-9 (a)]. A sudden step does not appear in the ASOS-ICPF [Figure 3-9 (b)] due to its reporting resolution of 0.254 mm (0.01in). At 62 COOP gauge stations, COOP-ICP (2.54) exceeds MPE-ICP (2.54) regardless of the shapes of the ICPFs afterwards. By visual inspection, the ASOS-ICPF agrees MPE-ICPF over most of the range of rainfall intensity at ASOS station 447285 in Figure 3-9 (b). Also, the horizontal asymptotes of the two ICPFs are approximately the same, because the total precipitation by the two methods is the same at this station. The overall analysis of ICPF for the ASOS system (10 sites) did not find any systematic features between ASOS- and MPE-ICPFs.

3.6.2.5 Root mean square difference between hourly GRM and MPE

Based on the frequency distribution and the ICPF analysis, a "threshold reporting depth bias" may be introduced because the COOP system reports precipitation only when its depth reaches a threshold of 2.54mm (0.1in) and then assigns that precipitation to the time period in which this threshold is reached. To evaluate a potential threshold reporting depth bias, RMSD (Equation 3-3) was calculated in two ways. First, GRM-MPE pairs when both systems report nonzero values were selected and analyzed ["conditional" in Figure 3-10(a)]. This analysis does not consider hours during which only one system reported rain. The other analysis contains all GRM-MPE pairs when either reports nonzero precipitation, excluding only no-rain conditions, when GRM = 0.0 and MPE = 0.0 ["unconditional" in Figure 3-10(b)]. This method allows all reported precipitation to be accounted for, but perhaps not agreeing in their temporal context.

The $RMSD_{unconditional}$ tends to be smaller than $RMSD_{conditional}$. This finding indicates that the better agreement between CGRM and MPE occurs when MPE nonzero

values are taken into account while CGRM accumulates to the threshold reporting depth. The variation of $RMSD_{unconditional}$ among sites is also smaller than that of $RMSD_{conditional}$. For the unconditional analysis, $RMSD_{CGRM\text{-}MPE}$ ranges from 2.04 to 5.78. For the conditional analysis, $RMSD_{CGRM\text{-}MPE}$ ranges from 2.97 at COOP station 369318 to 10.66 at COOP station 183090. The unconditional analysis includes more hours (larger sample) and the differences are small during hours when CGRM is accumulating and not recording: both these factors reduce RMSD. It is possible that rainfall reported by CGRM actually fell during time periods recorded as zero by CGRM but was captured as nonzero by MPE.

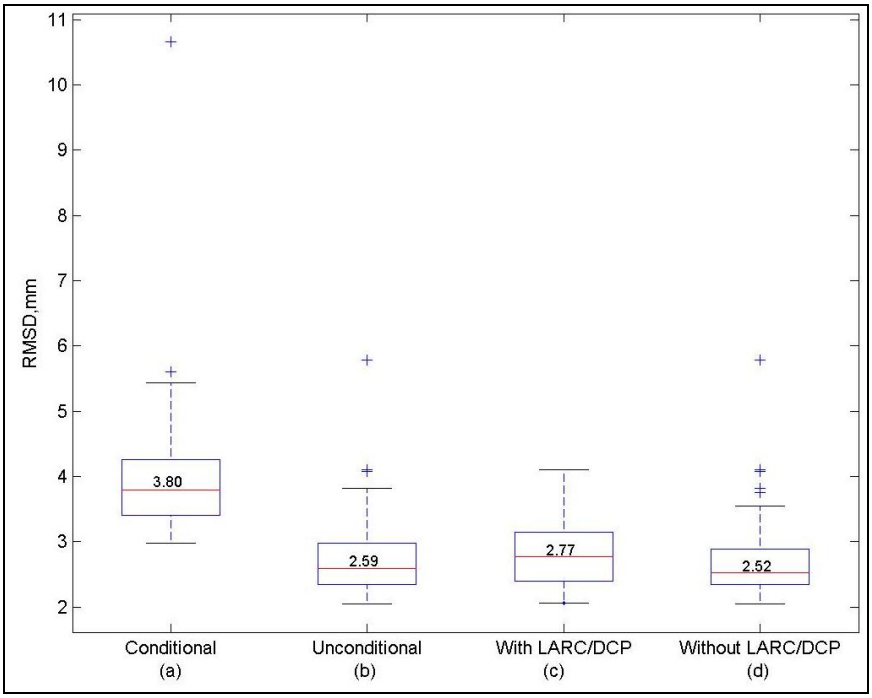


Figure 3-10. RMSD based on (a) conditional analysis (70 sites); (b) unconditional analysis (70 sites). $RMSD_{unconditional}$ based on CGRM reported by (c) telemeters (20 sites); (d) human observers (50 sites).

It might be expected that the RMSD for sites using CGRM reported by telemetry systems would be lower than for sites using CGRM reported by human observers, due to

the real-time involvement of the telemetry system during the process of producing MPE. However, no significant difference between two reporting systems is found based on the median values of 2.77 and 2.52 [Figures 3-10 (c) and (d)].

The RMSD is also used to examine one of the well-known issues in creating the MPE product: uncertainties in MPE values increase as the distance from NEXRAD radar sites increase. Plotting $RMSD_{unconditional}$ as a function of the distance shows the quality of MPE by testing it against CGRM data (Figure 3-11). In the plot, two COOP gauge stations are located within 5500m of the radar stations; STATE COLLEGE (KCCX) – COOP366916 and NORFOLK (KAKQ) – COOP448800. Their unconditional RMSDs are near the median values, 2.095 and 2.976, respectively. The highest RMSD, 5.78, is reported at COOP station 183090, within 32000m of NEXRAD radar station KDOX. The two-sided t-test on nonzero R was conducted at $\alpha=0.05$. The null hypothesis is accepted ($p = 0.713$), meaning there is no linear relationship between the two variables. The level of agreement of the MPE product with gauge data does not change with the distance from NEXRAD radar stations.

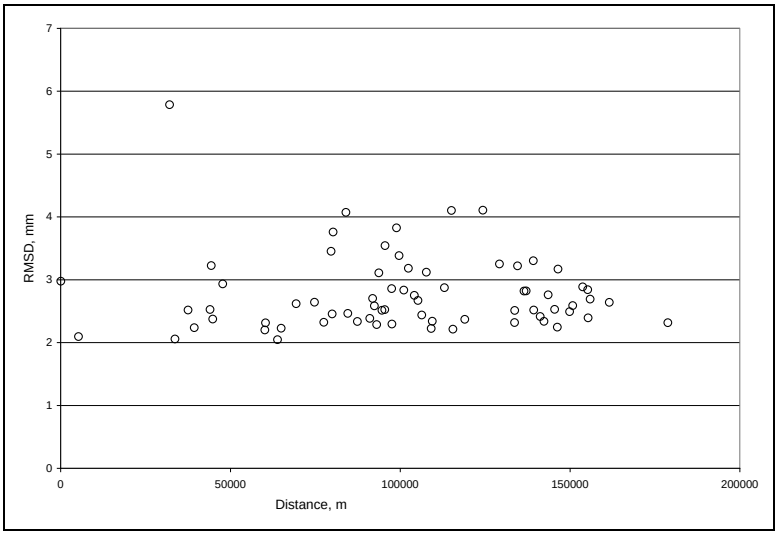


Figure 3-11. Plot of RMSD and the distance from the nearest NEXRAD radar station to each COOP gauge site. R=0.002

3.6.2.6 Analysis of gauge data recorded as trace and zero

Modelers who use gauge data often enter zero precipitation into their model for time periods when records are "trace" or zero. As discussed earlier, no data source is considered absolute "truth", however, gauge data recorded as trace/zero can at least be checked by examining another data set at the location and time.

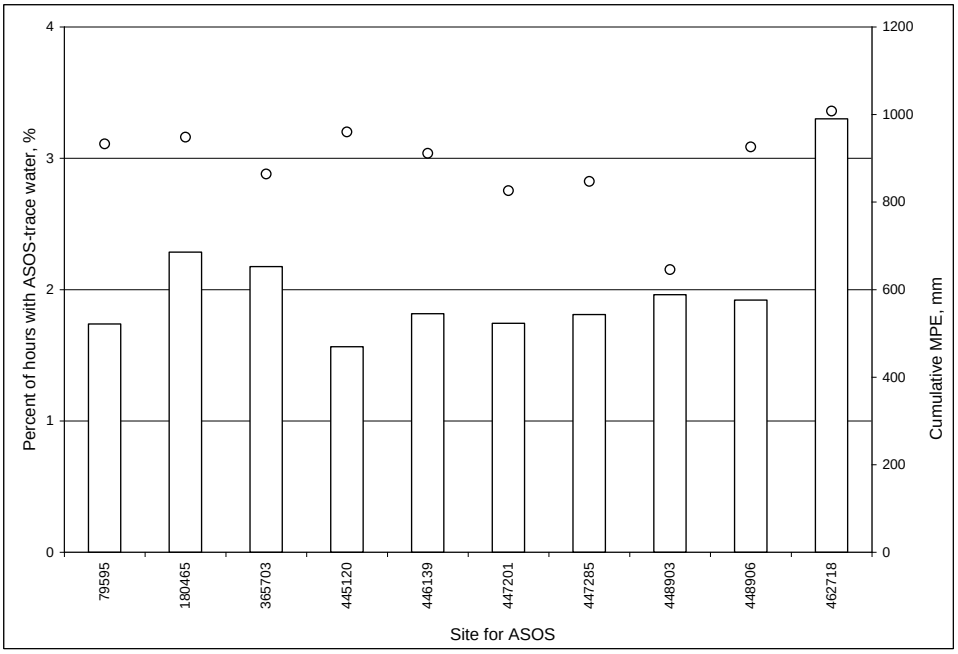


Figure 3-12. Percent of hours with ASOS-trace of water (▯) and cumulative MPE (○) during the corresponding hours in the grid cell containing the ASOS site.

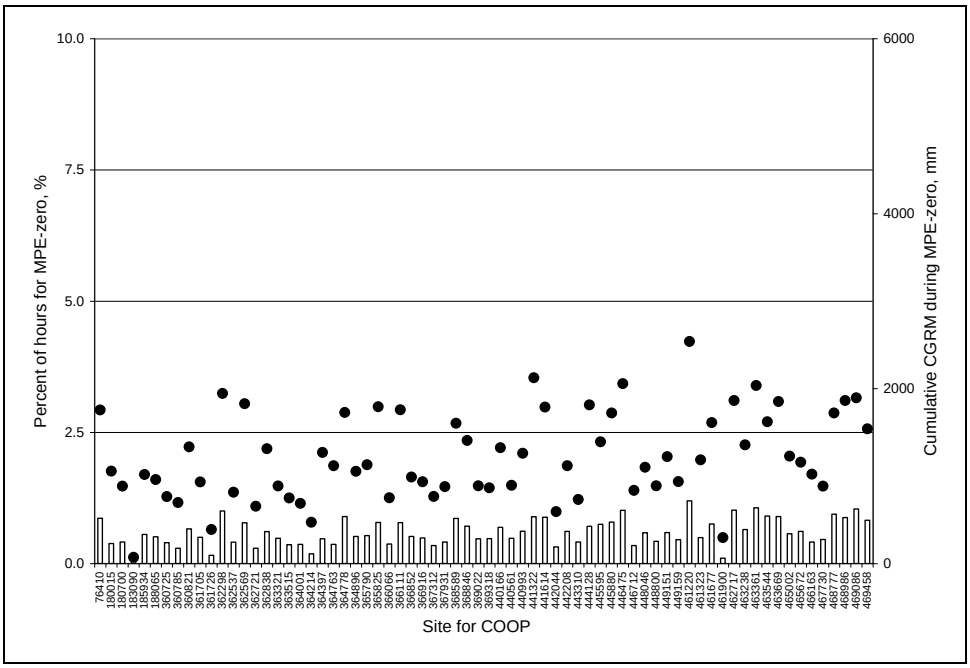
Using the MPE-AGRM pairs, this study examined what MPE reports during AGRM-trace (Figure 3-12). ASOS gauge sites report a "trace of water" for approximately 2% of the 61340 total study hours. During the corresponding hours, MPE accumulates an average of 886 mm. This implies that, on average, MPE records 0.681 mm (0.0268 in) for every hour that the ASOS system reports a "trace of water". This value is inconsistent with the definition of "trace" as less than 0.254 mm (0.01 in). The

overestimation of MPE may be caused by either the ASOS point to MPE pixel comparison, or bias in estimating light precipitation from radar reflectivity.

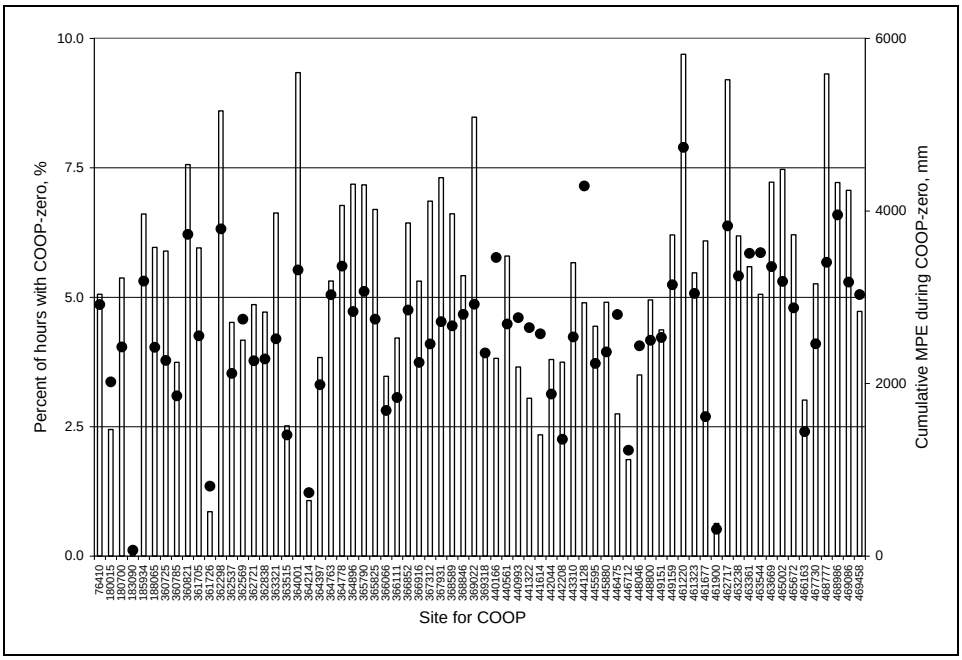
Three distinct scenarios involve zero values: (1) When both GRM and MPE show zeros at a given time and location, the pair is eliminated from the analysis as no rainfall is assumed; (2) Gauge reports nonzero precipitation and MPE reports zero in the corresponding grid cell. Even though a gauge site may report rain, rain elsewhere within its corresponding 4 km-by-4 km grid cell may be so light that the grid average rainfall is less than 0.01mm, which is MPE -measurement precision. In other words, the spatial variability of rainfall and different measuring/reporting systems for each data source may contribute to the discrepancies in rainfall records at a given time and location; (3) GRM reports zero at a site when MPE reports nonzero values in the corresponding grid cell. It may be that precipitation falls elsewhere in the grid cell and either misses the gauge site or is too light to be recorded. At COOP gauge sites, CGRM reports zero until the hour in which the COOP system accumulates 2.54mm of rain. At ASOS gauge sites, AGRM may be simply zero or zero with a "trace-of-water" flag. To investigate the frequency and characteristics of such events, this study extracted pairs of GRM-MPE which satisfy the condition of having both a nonzero MPE report and a zero gauge report at a given time and location.

Among the three cases explained above, Figure 3-13 (a) illustrates the fraction of hours in which MPE is zero while CGRM is nonzero (Case 2). Case 2 occurs in about 1% of the total study hours (0.05% ~ 1.2%) at all COOP sites across the region.

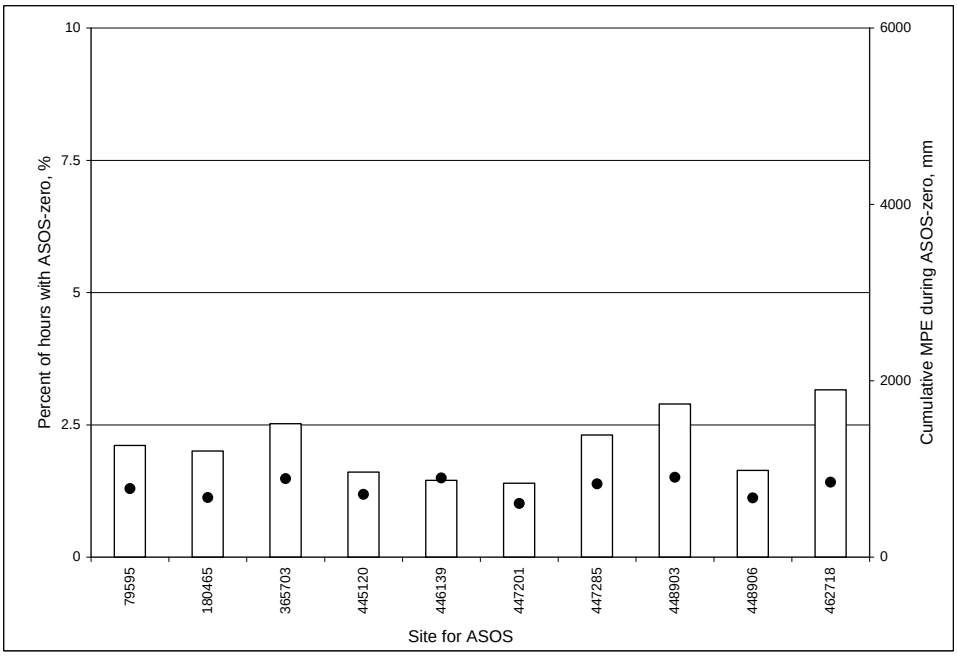
Figure 3-13. (a) Percent of hours with MPE-zero (▯) and cumulative CGRM during MPE-zero (●) at each COOP site for Case 2. (b) Percent of hours with CGRM-zero (▯) and cumulative MPE during CGRM-zero (●) at each COOP site for Case 3. (c) Percent of hours with AGRM-zero (▯) and cumulative MPE during AGRM-zero (●) period at each ASOS site.



(a)



(b)



(c)

The analysis shows that Case 3 occurs for more than 5% of the total hours at more than 60% of the COOP sites [Figure 3-13 (b)]. The two COOP sites closest to each other have a significant discrepancy: 0.6% of the total hours at COOP Site 461900 versus 7.5% of total hours at COOP Station 465002, approximately 5km away. At these sites over the corresponding time periods, MPE accumulates 310 mm and 3200 mm, respectively. This demonstrates that the relationship between the two measurement methods has a large spatial variation.

When the same analysis was conducted for ASOS stations [Figure 3-13 (c)], cumulative MPE during AGRM-zero consistently amounts to 900 mm per site across the region, even though the ASOS stations are located in diverse geographic settings (Figure 3-2). The cumulative MPE during GRM-zero is compared between COOP and ASOS stations located 5.1 km apart, the closest distance between COOP-ASOS sites in the PRB area. At ASOS Station 462718, the 7-year cumulative MPE amounts to 850 mm. At

nearby COOP Station 462717, the cumulative MPE amounts to 3800 mm, which corresponds to 50% of the average annual precipitation of 1000 mm (40 in) in the PRB region (Section 3.3). This also demonstrates that the spatial variability of precipitation measurement methods is significant.

All the COOP gauge stations report Case 3 more than Case 2. MPE accumulates as much as 4730mm during CGRM-zero (10% of the total hours) at COOP station 461220 [Figure 3-13 (b)]. At the same station, CGRM records nonzero values during MPE-zero for only 1.2% of the total hours. Figure 3-14 demonstrates that GRM reports a higher percentage of zeros than MPE across the region.

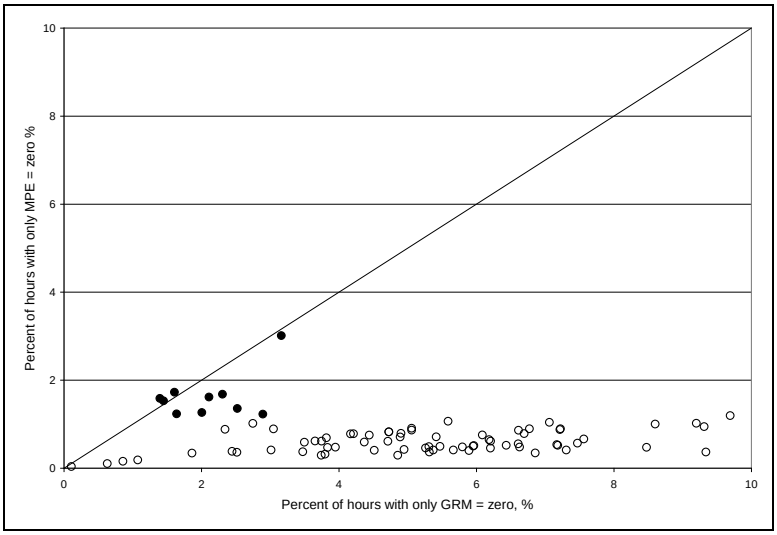


Figure 3-14. Plots of percent hours of only GRM-zero and percent hours of only MPE-zero at a given COOP gauge location (○) and ASOS gauge location (●).

The likelihood of detecting precipitation is higher in MPE 4x4 km^2 pixel than at the gauge point. A finer network of gauges would be needed to decide whether this finding is due to the physical likelihood of detecting precipitation in a 4x4 km^2 pixel (as opposed to a point in that pixel) or due to properties of the radar system. The COOP delayed response due to the threshold reporting depth also contributes. ASOS gauge stations

show a pattern similar to the COOP gauge stations', but with less difference in the percentages and with slightly higher percentages of the MPE-zero occurrence. Because of instrumental precision and their threshold reporting depth, COOP stations record zero more often than ASOS stations do.

By investigating what MPE reports for the time periods of uncertainty reported in the gauge records, this study shows that MPE can be a guide when analysts and modelers handle missing data. The study demonstrates that one cannot simply interpret zero values as no precipitation. It also demonstrates that interpolating from neighboring gauges may introduce errors.

3.6.3 Comparison of mean areal precipitation: mCBP and mMPE

The county-based mean areal MPE (mMPE) is plotted against the county-based mean areal CBP (mCBP) (Figure 3-15). Average annual mCBP ranges from 40 to 60 inches (1016-1524 mm) with a median value of 44.5 inches (1130 mm). Average annual mMPE ranges from 22 to 53 inches (560-1346mm) with a median value of 41.8 inches (1061 mm). During the growing season, from May to October, the average annual mCBP and mMPE ranges are 23-32 and 15-31 inches, respectively. The average annual mCBP exceeds the average annual mMPE in nearly 85% of the land segments.

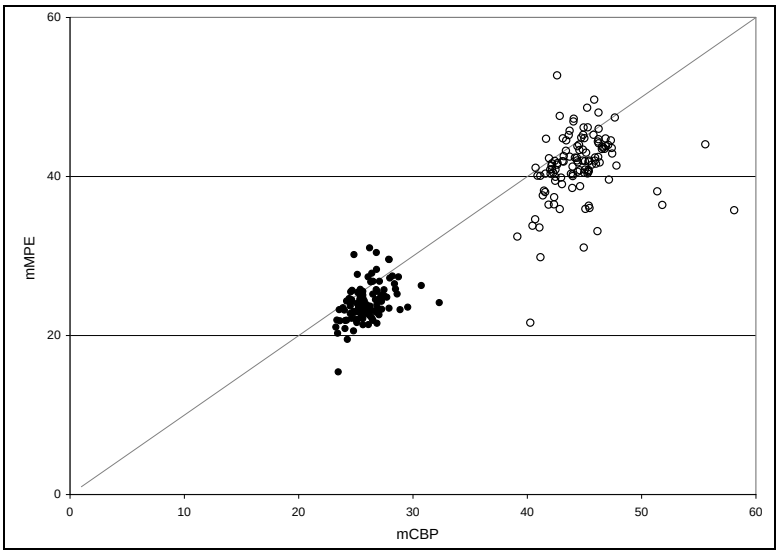


Figure 3-15. Scatter plot of average annual mCBP and mMPE based on all months (○) and growing season months (●). Of the 114 counties, 96 are below the 1-to-1 line when all months are taken into account and 97 are below the 1-to-1 line for growing season months. Units are inches.

The seasonal difference between mCBP and mMPE varies from one county to another to the largest degree in winter (Figure 3-16). This explains the feature found in Figure 3-15: the variance among sites for growing months is smaller than that for all seasons. The median difference is negative for every season, meaning that each seasonal total mCBP exceeds the corresponding mMPE in more than half of the 114 county land segments. Among the counties, the median difference is the largest in summer, and the spatial variability is the greatest in winter.

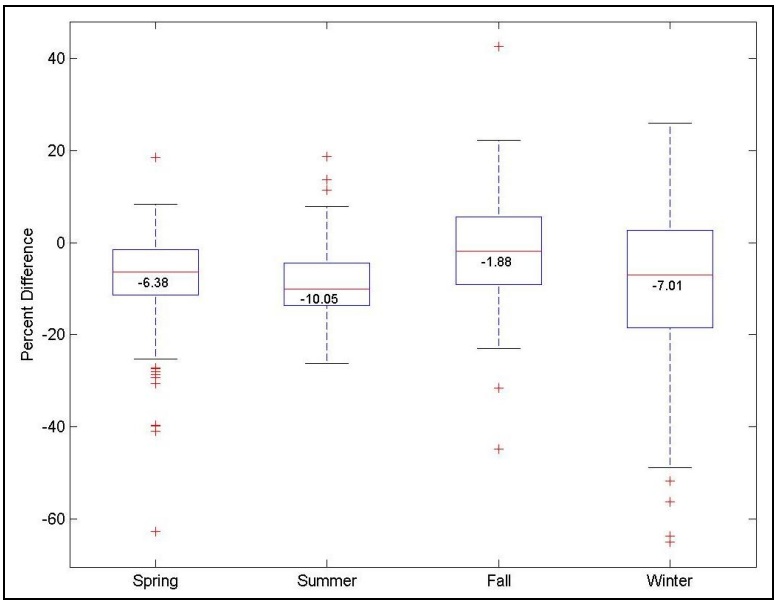


Figure 3-16. Boxplot of the seasonal relative percent difference [100 x (mMPE-mCBP)/mCBP] in precipitation. The values inside the boxes are the medians of the 114 county-based differences. In this study, the spring months are March, April, and May. The summer months are June, July, and August. The fall months are September, October, and November. The remaining months are for winter.

The spatial distribution of the average annual difference shows the exceedance of mCBP across the basin, except for a narrow band region as well as a few other counties in the Valley and Ridge physiographic province (Figure 3-17). On an annual basis, mMPE underestimates mCBP the most in the western part of the basin. Based on the gauge point-MPE pixel analysis using hours for both valid values (Section 3.6.2), the average annual MPE-pixel values exceed the average annual gauge point values across the basin without any spatial patterns (circles in Figure 3-17). In Somerset and Alleghany Counties, MPE exceeds the COOP annual precipitation at both gauge stations within each county, but the spatially-aggregated mCBP exceeds mMPE (Table 3-7). In Highland and Preston Counties, the average annual difference between mMPE and mCBP has a far more negative magnitude than the difference between the MPE pixel and

the gauge point values. At all ASOS gauge stations, MPE pixel precipitation exceeds gauge point precipitation, whereas counties containing those ASOS stations show that mCBP exceeds mMPE. These observations consistently demonstrate that mMPE underestimates mCBP, regardless of the difference between each pair of gauge point and MPE pixel. This indicates that there exist biases between point-pixel precipitation pairs and the mean areal precipitation pairs. Such biases could come from radar-related uncertainties or uncertainties that arise in handling invalid data or other processes during the creation of mCBP time series. This finding might be explained by the mathematics of interpolation, although this study does not have information on the exact procedure for the creation of the mCBP data set. Mathematically speaking, most interpolation schemes would not assume zeros between two nonzero values, resulting in overestimation. Another explanation for these biases could be that the interpolation scheme overestimates the spatial correlation of the precipitation field.

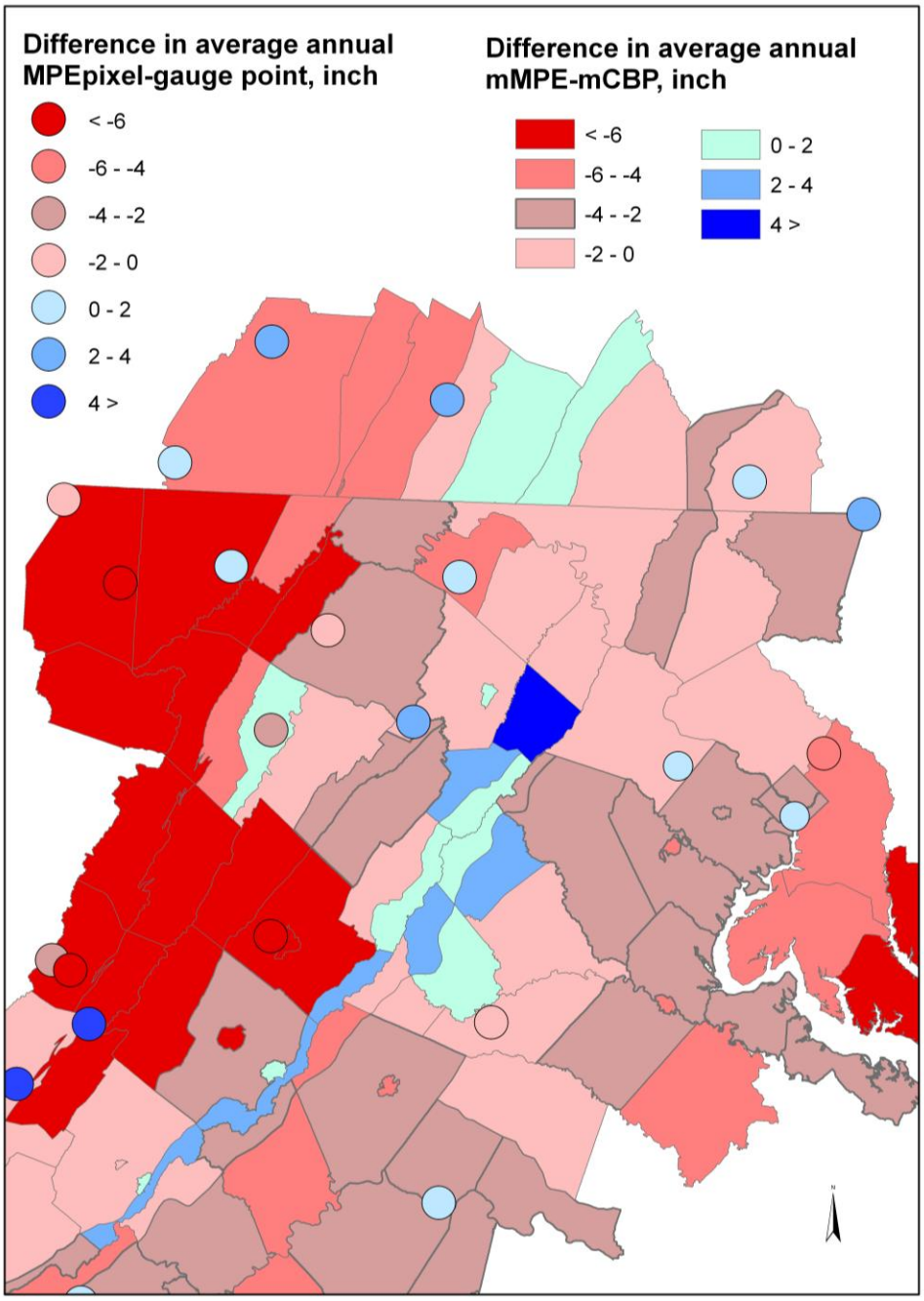


Figure 3-17. Spatial distribution of the average annual difference between (a) mMPE and mCBP during 2001-2005 (polygon) and (b) MPE pixel and COOP gauge point during 2001-2007 (circle)

Table 3-7. Average annual difference with and without spatial aggregation. The point-pixel difference is based on a 7-year time period, 2001-2007 (Section 3.6.1). The difference in mean areal precipitation data sets is based on a 5-year time period, 2001-2005. Stations in this table are located within the PRB. Units are inches.

Type	COOPID	MPE pixel – gauge point	County name	mMPE-mCBP
COOP	180700	-4.520	PRINCE GEORGES	-4.590
COOP	185934	2.033	CARROLL	-3.573
COOP	188065	0.222	GARRETT	-13.259
COOP	362537	0.022	ADAMS	-0.192
COOP	362721	2.649	BEDFORD (PA)	-4.975
COOP	440561	0.248	BEDFORD (VA)	-3.745
COOP	440993	1.571	BUCKINGHAM	-3.947
COOP	442208	-8.357	ROCKINGHAM	-11.335
COOP	444128	9.349	BATH	-0.531
COOP	446712	-1.301	ORANGE	-1.816
COOP	448046	2.121	FREDERICK	-0.524
COOP	466163	-2.638	HARDY	0.361
COOP	467730	-0.051	HAMPSHIRE	-3.761
COOP	449159	5.199	BATH	-9.434
COOP	461323	1.296	MORGAN	-5.870
COOP	360821	3.439	SOMERSET	-5.865
COOP	361705	0.370		
COOP	442044	2.674	ALLEGHANY	-1.329
COOP	443310	2.059		
COOP	445595	-3.483	HIGHLAND	-9.062
COOP	445880	-7.693		
COOP	461900	-0.591	PRESTON	-11.551
COOP	468777	-8.008		
ASOS	445120	0.737	CAMPBELL	-2.360
ASOS	448903	1.083	LOUDOUN	-0.257
ASOS	448906	1.112	ARLINGTON	-3.298
ASOS	447285	0.346	ROANOKE	-1.257

3.6.4 Comparison of Precipitation events using spatially aggregated time series

The number of precipitation events differs depending on the analysis settings described in Section 3.5.6. Overall, mCBP events outnumber mMPE events by 33%. After eliminating events according to the criterion of having a total depth per event of at least 5mm, the number of mCBP events exceeds mMPE events by only 18% (from data set #1 to data set #2 in Table 3-8). Events with a total depth of less than 5 mm occurred

more frequently in mCBP than in mMPE: there are 28,139 mCBP events and 17,724 mMPE events. As the time windows applied to select mCBP and mMPE pairs become narrower, the number of matching pairs decreases (Table 3-8). In the five years of record, mCBP and mMPE events coincide exactly in time (0 hour window) for only 197 precipitation events.

Table 3-8. Type of analyses and number of precipitation events during 2001-2005 from 26 counties.

<table>
<tr><td colspan="3">Precipitation event</td><td>Analytical setting (Data set)</td><td>mMPE</td><td>mCBP</td></tr>
<tr><td colspan="3">All events</td><td>#1</td><td>24,143</td><td>35,964</td></tr>
<tr><td rowspan="5">Total depth per event >= 5mm (0.196 in)</td><td colspan="2">unpaired</td><td>#2</td><td>6,419</td><td>7,825</td></tr>
<tr><td rowspan="4">Paired</td><td>12-hour window</td><td rowspan="4">#3</td><td colspan="2">3,453</td></tr>
<tr><td>6-hour window</td><td colspan="2">2,524</td></tr>
<tr><td>3-hour window</td><td colspan="2">1,541</td></tr>
<tr><td>0-hour window</td><td colspan="2">197</td></tr>
</table>

3.6.4.1 Comparison of unpaired precipitation events

Using the data set #1(all events), a comparison of the monthly precipitation events demonstrates that mCBP events occur more often than mMPE events every month (Figure 3-18). In January, mCBP events exceed mMPE events by as much as 44%. The smallest difference between mMPE and mCBP events occurs in December at 23%. The monthly pattern found in mCBP is similar to that in mMPE: May has the most events, followed by July and August. September has the fewest events.

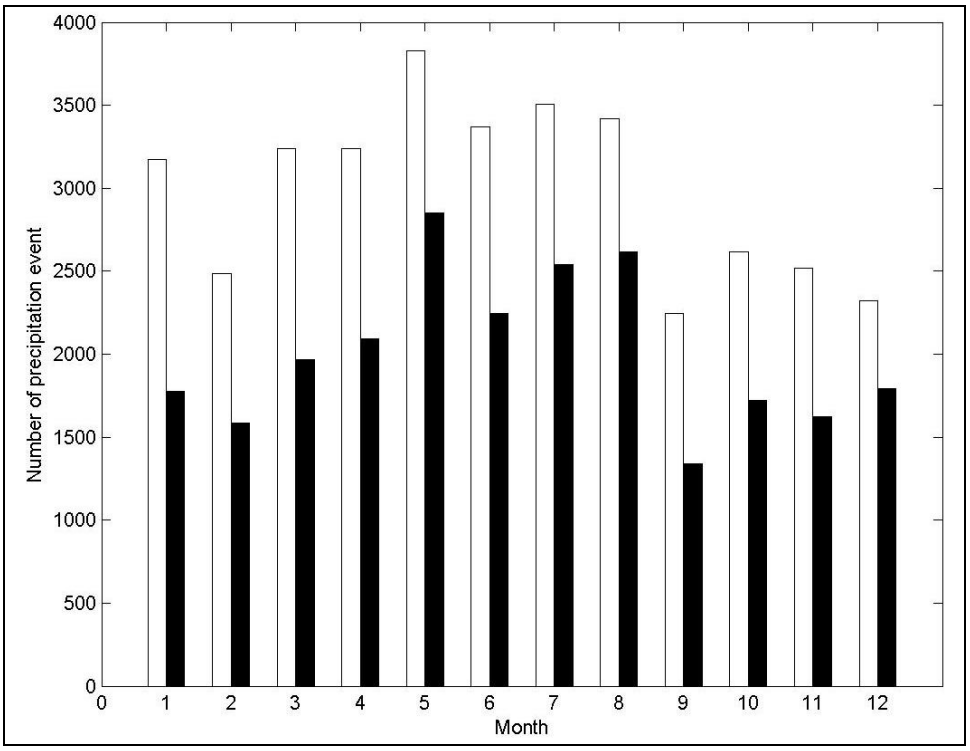
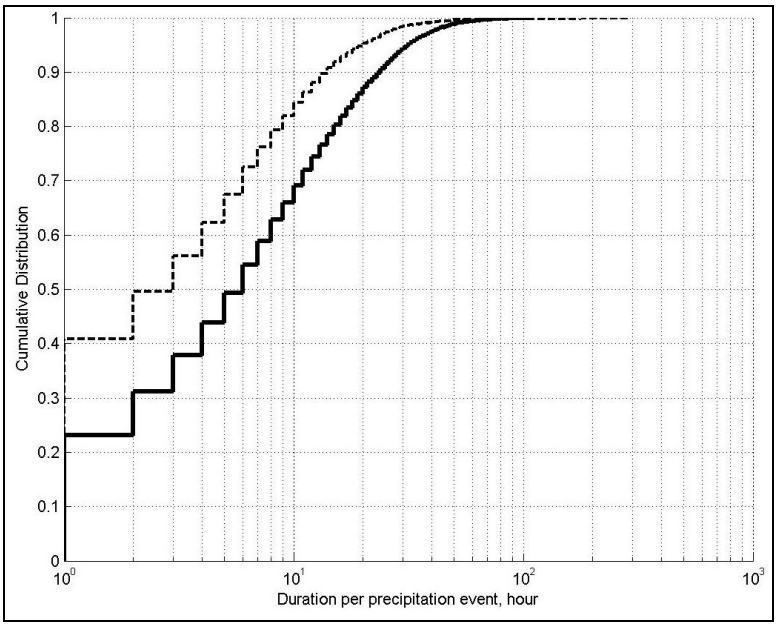
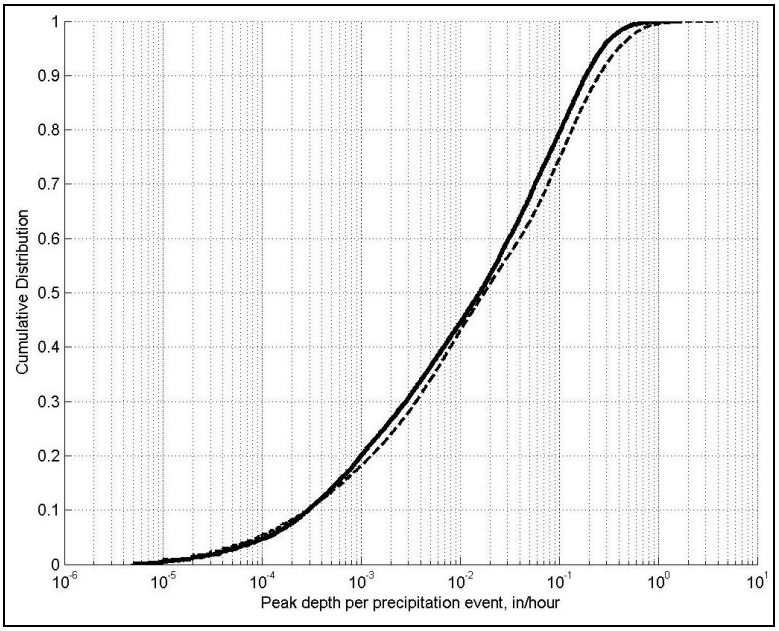


Figure 3-18. Number of precipitation events, mCBP (▯) and mMPE (▮), based on the data set #1 in Table 3-8 on a monthly basis.

Since the actual number of mMPE-events is different from the number of mCBP-events (data set #2 in Table 3-8), their characteristics (duration and peak depth) were compared using ECDFs [Figures 3-19 (a) and (b)]. The two ECDFs indicate that mCBP events generally have shorter durations than mMPE events [Figure 3-19 (a)]. However, the longest event is found in the mCBP event which lasted 264 hours. The longest duration among mMPE events is 110 hours. Approximately 40% of the mCBP events have a one-hour duration, whereas only 23% of the mMPE events have a one-hour duration. Heavy peak intensities are slightly more frequent for mCBP events than for mMPE events [Figure 3-19 (b)].



(a)



(b)

Figure 3-19. ECDFs for the charateristics (a) duration; (b) peak depth, of mMPE (—) and mCBP (···) from 26 counties.

3.6.4.2 Comparison of paired precipitation events

Using data set #3, mCBP and mMPE events paired using different time windows were compared to examine several characteristics of precipitation events, including

duration, peak depth, peak time, and total depth (Table 3-9). Applying different time windows results in different numbers of events for comparison. As the time window grows, the number of mMPE-mCBP paired events increase and the differences between the two become more noticeable. In general, mMPE events have longer durations and lighter peak depths than their paired mCBP events. The total depth per mMPE event slightly exceeds the mCBP's. Peaks in mMPE events occur earlier than those in mCBP events. Using the median values shown in Table 3-9, the ratio of peak time to event duration was calculated and converted into a peak time in the Soil Conservation Service (SCS) 24-hour rainfall distribution (last two rows in Table 3-9). Peaks in mCBP events are similar to Type II storms in the SCS 24-hour storm distributions: the peak occurs at the center of a storm (Figure 3-20). Peaks in mMPE events are similar to Type I storms, in which the peak occurs at a storm time of about 8 hour. According to the "Approximate geographic boundaries for SCS rainfall distributions" (McCuen, 1998), the Type II storm is applied for most of the U.S., including the Potomac River Basin. Standard deviations show that characteristics such as duration, peak depth, and peak time vary more in mCBP events than in mMPE events.

The characteristics of paired/unpaired precipitation events are different from those based on the entire precipitation time series. For example, the average annual mMPE underestimates the average annual mCBP in the majority of the land segments (Column 5 in Table 3-7), whereas the total depth per mMPE event slightly exceeds the mCBP's, as seen in Table 3-9. This discrepancy is due to the number of mCBP precipitation events excluded from the statistical comparison by Habib (2009)'s definition of "precipitation event".

Table 3-9. Statistics of precipitation event pairs with various time windows. Peak time refers to the number of hours from the time an event starts to the time a peak depth is reached. Ratio = median (Peak time)/ median (Duration)

Time Window		+/-12		+/-6		+/-3		+/-0	
Num events		3453		2524		1541		197	
Characteristics	stats	mMPE	mCBP	mMPE	mCBP	mMPE	mCBP	mMPE	mCBP
Duration (hour)	Mean	21.539	16.580	19.888	17.179	18.589	17.561	20.157	20.437
	Median	19	14	17	15	16	15	18	18
	25^{th}	12	8	11	9	11	9	12	12
	75^{th}	27	22	26	23	24	23	26	25
	90^{th}	38	30	34	30	32	29.4	36.8	36.8
	Std.dev	13.105	13.273	12.120	13.745	11.335	15.044	10.948	13.814
Peak Depth (inch)	Mean	0.204	0.269	0.201	0.260	0.196	0.245	0.200	0.230
	Median	0.163	0.199	0.162	0.194	0.158	0.186	0.166	0.178
	25^{th}	0.104	0.124	0.104	0.120	0.105	0.116	0.109	0.115
	75^{th}	0.258	0.343	0.258	0.328	0.244	0.315	0.264	0.308
	90^{th}	0.390	0.546	0.381	0.528	0.366	0.501	0.379	0.465
	Std.dev	0.149	0.223	0.146	0.213	0.141	0.191	0.128	0.160
Event-total (inch)	Mean	0.818	0.714	0.778	0.702	0.715	0.674	0.805	0.731
	Median	0.607	0.555	0.589	0.548	0.554	0.531	0.660	0.575
	25^{th}	0.370	0.358	0.358	0.362	0.347	0.357	0.379	0.345
	75^{th}	1.006	0.875	0.943	0.858	0.858	0.828	1.008	0.895
	90^{th}	1.672	1.324	1.552	1.305	1.402	1.268	1.649	1.440
	Std.dev	0.678	0.559	0.641	0.538	0.570	0.497	0.598	0.550
Peak time (hour)	Mean	9.193	9.338	7.982	9.718	7.260	9.750	8.432	11.715
	Median	6.961	6.961	6.000	7.919	5.760	7.919	6.960	10.078
	25^{th}	3.120	2.880	2.880	3.120	2.161	3.838	2.880	5.042
	75^{th}	12.000	12.960	10.440	12.961	10.078	12.960	11.040	15.901
	90^{th}	18.958	20.880	16.081	20.880	15.840	20.880	17.041	24.000
	Std.dev	8.332	9.420	7.511	9.383	6.741	9.449	7.227	9.192
Ratio[1]		0.366	0.497	0.353	0.528	0.360	0.528	0.387	0.560
Peak time converted to 24-hour rainfall distribution		8.793	11.933	8.471	12.670	8.640	12.670	9.280	13.438

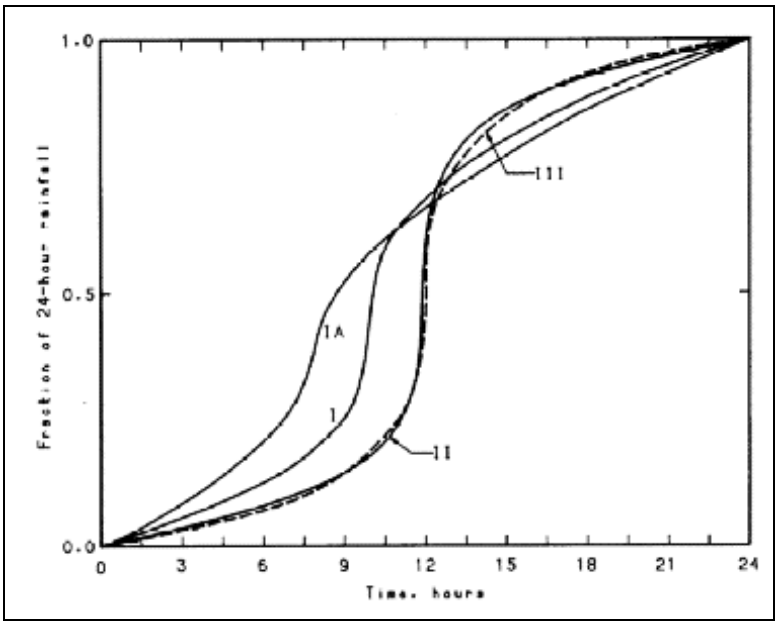


Figure 3-20. SCS 24-hour rainfall distributions (McCuen, 1998).

3.7 Discussion and Conclusion

This study examined three types of hourly precipitation data sets: COOP gauge-based rainfall measurement (CGRM), ASOS gauge-based rainfall measurement (AGRM), and Next Generation Weather Radar (NEXRAD)-based Multisensor Precipitation Estimator (MPE). Gauge records were compared to the corresponding 4x4 km^2 pixel MPE data. Then, spatially aggregated data sets at a county scale, gauge-based (mCBP) and NEXRAD-based mean areal precipitation data (mMPE) were compared (Table 3-1). From the mCBP and mMPE time series, both unpaired and paired precipitation events were extracted, and compared. Questions proposed at the beginning of the study are discussed in this section.

3.7.1 Comparison of collocated gauge point and MPE pixel time series

The annual reported point-GRM and pixel-MPE data sets, averaged over hours when both reports are valid, agree well at a given space and time at the majority of the 70 COOP and 10 ASOS stations. Statistical measures indicate a cumulative depth bias between COOP and MPE at the COOP threshold reporting depth of 2.54mm (0.1 in). Because of the high rate of missing values, ranging from 3% to 96% of the total study hours, the COOP gauge data used in this study are not directly usable in hydrologic modeling: an additional step of filling the missing time periods is required.

In the ASOS system (10 stations in the Potomac Basin), recording failure is rare. Because AGRM is used to create MPE, AGRM would be expected to agree well with MPE. However, the double mass curve reveals that AGRM from the ASOS station with Level D does not support this expectation.

The MPE data over the 7-year time period has only a few hours of invalid or missing values over the entire Basin. Although radar-related uncertainties in the MPE data sets are found at some locations, the average annual MPE values match well with both the AGRM- and CGRM values across the region.

This study demonstrates that an accumulation of small or invalid precipitation values can add up to a significant amount, although a single missing amount may be negligibly small. This study's analysis of cumulative MPE corresponding to CGRM-invalid values indicates that missing or invalid CGRM values should not simply be replaced with values from nearest stations because of the high spatial variability between stations. The analysis of cumulative MPE corresponding to CGRM-zero values suggests that CGRM-zeros should be verified before being accepted as valid zeros.

3.7.2 Assessment of GRM and MPE as a reference tool

This study explored whether uncertainties in one type of precipitation (gauge, radar) can be quantified using the other type (radar, gauge) as a reference. The GRM recorded as invalid (missing or deleted), trace of water, or zeros were quantified by using MPE. The analysis demonstrates that MPE can provide modelers with guidance to deal with missing values and bias due to the threshold reporting depth when using historical hourly gauge-based data.

The difference between the MPE and CGRM pair was examined as a function of the distance from the nearest NEXRAD radar station to each gauge station. This study reveals that the distances are irrelevant to the difference between the two. Sometimes the quality of CGRM is inadequate to serve as a reference tool because of the high rate of invalid values. At a given location, however, CGRM can serve as a water-budget guidance tool to which MPE might be adjusted or compared.

Each data set has its strengths and weaknesses. None of the individual data sets can be considered to be the ultimate "truth" of precipitation measurement, but a better understanding of the difference among precipitation data sets would improve the reliability, quality, and applicability of the individual datasets.

3.7.3 Comparison of collocated areally aggregated gauge and MPE time series (mCBP and mMPE)

The average annual mean areal Chesapeake Bay Program precipitation (mCBP) exceeds the average annual aggregated MPE precipitation (mMPE) in the majority of the 114 county-based land segments across the basin. However, the gauge point annual

precipitation value exceeds the MPE pixel value at only one-third of the 27 gauge stations, which may indicate that an excessive amount of precipitation is accounted for during the process of generating the mean areal mCBP data set. This process involves filling invalid values, spatially interpolating values using multi-linear regression equations, and disaggregating daily values into hourly values.

Based on the analysis of unpaired precipitation events, when "event" is defined by an interevent period of 6 hours, mCBP events outnumber mMPE events by 33%. However, approximately 40% of the mCBP events have a one-hour duration with the threshold reporting depth of 0.1 inch, whereas only 23% of the mMPE events have a one-hour duration with various reporting values. In its analysis of paired precipitation events, this study finds that the characteristics of events are different in the two data sets: mMPE events have longer durations, lower peak depths, and earlier peaks than mCBP events. The total depth per mMPE event slightly exceeds that of mCBP's. The characteristics of paired and unpaired precipitation events are not necessarily the same as the characteristics of the complete precipitation time series. Overall, mCBP delivers more precipitation, but much of that precipitation occurs in events that are not reflected in mMPE events.

3.7.4 Use of MPE in hydrologic and water quality models

Using MPE in hydrologic and water quality modeling is appealing because of its spatial coverage and temporal resolution. It does not suffer from the temporal data gaps of the gauge record. Unlike gauge-based precipitation data, spatial interpolation is not required. This study, however, finds significant differences between the mean areal time

series of hourly precipitation developed based on MPE and those previously developed from interpolated gauge data by the Chesapeake Bay Program for their HSPF-based lumped model. In addition, continuous hydrologic simulations are driven by the characteristics of all events and interevent periods in the entire time series, rather than by the characteristics of individual storm events. Because precipitation is the primary driver of hydrologic models, differences in precipitation data sets are expected to affect model performance. Chapter 4 explores the impact of precipitation input changes on model calibration accuracy.

CHAPTER FOUR

Calibration of Chesapeake Bay Watershed Model with Gauge- and Radar-Based Mean Areal Precipitation Data

4.1 Introduction

Calibration is the process of identifying model parameters that cannot be determined by direct measurement or physical laws and adjusting them to make modeled output time series as similar to the observed ones as possible. Because the modeling process simplifies and generalizes hydrological processes in the natural system, the hydrologic calibration is necessary to make a model reflect area-specific characteristics of a watershed. Studies show that calibrations also need to be input-specific: the calibration and prediction/application should use consistent types of inputs (Stisen and Sandholt, 2010; Troutmann, 1982). According to those studies, when a hydrologic model uses a different type of precipitation data, the model should be recalibrated using a new input to obtain adjusted parameters. From the model user's perspective, it is important to know whether or not recalibration is a necessary step when implementing radar-based precipitation into hydrologic models already calibrated with gauge-based precipitation.

The goal of this chapter is to examine the impact on model calibration accuracy of using different types of precipitation data. The Chesapeake Bay watershed (CBW) model was calibrated using both spatially aggregated gauge-based and radar-based precipitation inputs, mCBP and mMPE, respectively. The model calibration accuracy was compared by using both calibration and hydrologic statistics, where "calibration statistics" are

defined as those used as objective functions in the calibration process, and "hydrologic statistics" are the postprocessing analysis of the model calibration.

The study followed the autocalibration approach that the CBP currently uses. The purpose was to examine the feasibility and implications of implementing mMPE into the CBW model without any modification of the model system. This study addressed the following questions:

(1) How does model calibration accuracy change based on calibration statistics when mCBP is replaced with mMPE?

(2) How does model calibration accuracy change based on hydrologic statistics when mCBP is replaced with mMPE?

(3) What are the implications of using mMPE for the CBW model?

4.2 Background

4.2.1 Calibration

Model calibration is a critical process in developing models as decision-making tools in management, planning, risk assessment and regulatory policy-making (Shirmohammadi et al., 2008). It is conducted through the process of estimating model parameters by comparing model outputs at a certain condition with observed data for the same condition (Moriasi et al., 2007; Shirmohammadi et al., 2006). A complex hydrologic model with many parameters requires sophisticated calibration to ensure that the model can represent the real physical meaning of individual parameters.

Since the 1960s and 1970s, researchers have explored automated methods that require objective functions, a search algorithm, and a criterion for the termination of the

automated procedure (Moriasi et al., 2007; Gupta et al., 1999). As more advanced technologies have become available, the complexity of objective functions has evolved with the use of more comprehensive techniques for search algorithms, including a single multi-component objective function and stepwise multiple objective functions (Gutierrez-Magness and McCuen, 2005; Hay et al., 2006). Depending on how the objective functions are constructed, they are optimized using different criteria. This study used the same autocalibration approach that the CBP implemented for the CBW model. The CBP developed a series of procedures for optimization consisting of calibration statistics, objective functions, targeted parameters, and updating factors.

If a model is given a different type of precipitation data input, model performance will most likely be affected because precipitation is a major driving force in the model (Dingman, 2002). Finnerty et al. (1997) found that the model parameters were strongly linked to the temporal and spatial scales of the precipitation forcing inputs (cited in J. Guo et al., 2004). During calibration, model parameters adjust to compensate for the biases and sampling characteristics of a type of precipitation data (Bradley and Kruger, 1998; Bradley, 1997). Therefore, the hydrologic model parameters calibrated based on rain gauge networks may not be suitable for use with the NEXRAD precipitation data: recalibration may be needed.

Literature on model recalibration shows inconsistent views. Some comparison studies showed similar or improved model performance using NEXRAD-based precipitation data with recalibration compared to model results forced by gauge-based precipitation (Sexton et al., 2010; Kalin and Hantush, 2006; Guo et al., 2004).

Some studies using NEXRAD precipitation for their hydrologic model applications included recalibration process. For example, a framework for flood modeling using the HEC-HMS/RAS integrated the NEXRAD-based Stage III product with calibration (Knebl et al., 2005). A study evaluated the existing storm-water system to understand dynamic storm water response and ultimate solutions over the Texas Medical Center (Bedient et al., 2007). In their study, the Storm water management model (SWMM) with input hydrographs from HEC-HMS was calibrated using NERAD-based precipitation.

Comparison studies without recalibration process were also found. A few studies compared the SWAT model performance using gauge-based and NEXRAD-based Stage III data sets (Tuppad et al., 2010; Tobin and Bennett, 2009; Jayakrishnan et al., 2005). They reported that using Stage III produced better simulation results, even without recalibration. Using a physically based distributed parameter hydrologic model (GSSHA), Sharif et al. found that on a 30-meter square grid, the model simulation forced by NEXRAD-based MPE product outperformed the model using gauge-based precipitation (2010). Without recalibration, a study found that the SWAT model using NEXRAD-based data overpredicted high flow events and underpredicted low-flow events (Moon et al., 2004). The Large Basin Runoff Model (LBRM) forced by NEXRAD-based Stage II/III without recalibration showed poor performance compared to model results using gauge-based precipitation (Watkins et al., 2007). The USEPA BASINS-HSPF for 7 watersheds in the Chesapeake Bay performed better when using the NASA-modified precipitation (NLDAS) compared to when using NEXRAD-based data (Nigro et al., 2010).

Gauge based precipitation and MPE represent two different estimates of real-world precipitation. As shown in Chapter 3, the two data sets have different statistical properties when they are aggregated to the same spatial scale. It is worthwhile to examine the calibration statistics for the CBW model using mCBP and mMPE.

4.2.2 Statistics on the automated model calibration

The CBP defines statistics specifically for the CBW model calibration. These statistics are calculated at specified locations over the entire time period of a model run. The CBP developed a series of programs that automatically calculate these statistics after each model run (USEPA, 2010). All statistical calculations are based on the days when both simulated and observed values are available. The automated model calculation includes an algorithm for separating peak flows and a method for estimating baseflow before calculating the statistics for the individual components of a hydrograph. To calculate the statistics of peak flows, the program requires user-specified numbers of peak flows (this study used 50, following the CBP practice). The peak flows are ranked in descending order and included in the statistical calculation, only if the peaks from simulated flow and observed flow occur on the same day (USEPA, 2010). The PART software developed by the USGS (Rutledge, 1998) is used to determine baseflow from observed or model-generated streamflow records.

In this section, the calibration statistics are briefly defined; further details can be found in USEPA (2010). Capital letters are used to express all statistical terms adapted from the CBP documentation. In this study, the term "BIAS" is used when the model simulation is compared to measured river flow, assuming that the measurements reflect

the true state of the flow variables. Various versions of "BIAS" are taken from the CBP's statistical definitions and used to compare model results for the calibration and validation periods. "Relative difference" is used to compare precipitation, because all precipitation sources are treated as estimates of an unknown true state.

BIAS is a statistical measure of the agreement between simulated and observed streamflow. Total bias (TBIAS) is a dimensionless quantity defined as follows:

$$\text{TBIAS} = \left[\frac{\sum_{i=1}^{n} y_i^{sim} - \sum_{i=1}^{n} y_i^{obs}}{\sum_{i=1}^{n} y^{obs}} \right] \quad (4\text{-}1)$$

Where

i = daily index for all paired values
$y_i^{obs} = i^{th}$ observed stream flow
$y_i^{sim} = i^{th}$ simulated stream flow

The CBP also defines the following BIAS statistics calculated as in Equation 4-1, but for different subsets of the flow time series:

BBIAS = (baseflow bias) using daily index, *i* for all paired baseflows

WBIAS = (winter bias) using daily index, *i* for all paired values from December, January, and February

SBIAS = (summer bias) using daily index, *i* for all paired values from June, July, and August

PBIAS = (peak flow bias) using daily index, *i* for all paired peak flows as described above.

VPBIAS = (volume peak bias) using daily index, *i* for volumes of storms related to paired peak flows from flow preceding to another peak (USEPA, 2010).

LOW10bias = (low flow bias) using daily index, *i* for all paired flows below the 10^{th} percentile.

A BIAS statistic equal to zero is desirable: this indicates that the model simulation produces the same volume of flow as the observed in the specified range or time period. It does not, however, indicate a "perfect" model because it gives no information on simultaneous agreement.

Bias indices normalize the different BIAS statistics by the total bias (TBIAS). The CBP defines three types of bias indices as follows:

$$\text{BSTAT} = \left[\frac{BBIAS + 1}{TBIAS + 1}\right] \tag{4-2}$$

$$\text{WSTAT} = \left[\frac{WBIAS + 1}{TBIAS + 1}\right] \tag{4-3}$$

$$\text{SSTAT} = \left[\frac{SBIAS + 1}{TBIAS + 1}\right] \tag{4-4}$$

Given that a bias statistic equal to zero is desirable, adding 1s to the numerator and denominator to avoid the division by zero.

Recession indices measure the ratio of flow at day *n+1* to flow at day *n* during the recession period. The groundwater recession index measures the rate at which groundwater outflows to the stream. The interflow recession index measures the rate at which interflow is discharged from storage. The CBP developed two recession index terms, BaveRI and QaveRI to capture how well the observed base flow or interflow recession is represented in the model. They are defined as follows:

Let *Q* be a vector for flows during the *n+1*days of a recession period,

[q(t_1), q(t_2), q(t_3),..... q(t_n), q(t_{n+1})]. Periods of baseflow and stormflow (interflow) recession are identified using the USGS method described in this section, and an average recession index (avgRI) is calculated as follows:

$$avgRI = \left[\frac{avgRI^{sim}}{avgRI^{obs}} \right] \tag{4-5}$$

where

$$avgRI^{sim} = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{q(t_{i+1})^{sim}}{q(t_i)^{sim}} \right] \tag{4-6-1}$$

$$avgRI^{obs} = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{q(t_{i+1})^{obs}}{q(t_i)^{obs}} \right] \tag{4-6-2}$$

where n represents the number of observations in the respective period. The index is calculated separately for baseflow and stormflow (interflow), that is,

BavgRI = avgRI for baseflow

QavgRI = avgRI for stormflow

4.2.3 Automated model calibration

The CBW model calibration is an automated iterative process. Key hydrologic parameters are updated automatically by a set of rules based on calibration statistics. The process continues until a maximum number of iteration is reached.

The CBP selected six hydrologic parameters for the model calibration. These parameters were identified by a series of hand calibrations as the most sensitive in the hydrologic process simulation (USEPA, 2010). The six parameters are land evaporation (LAND_EVAP), lower zone nominal storage (LZSN), infiltration index (INFILT),

interflow inflow index (INTFW), interflow recession index (IRC), and groundwater recession index (AGWR). These parameters are required when running the model for land segments; parameter values are assigned by county and by land use (USEPA, 2010). For example, INFILT has at most 21 different values within a county, excluding three types of land use, belonging to impervious areas (impervious land use requires a different set of parameters which are not affected by the hydrology calibration in this study). However, not all 21 parameter values are estimated independently: the CBP established a series of rules for setting up parameters. The rules include fixed ratios of parameters for all other land uses to cropland parameter values, and of upper zone nominal storage (USZN) to LSZN by land use. Thus the parameters that are free to be estimated in calibration are the cropland values of these six parameters for each land segment. Other land use parameters change in the course of calibration, but always in proportion to the cropland parameters for that land segment.

Besides these six parameters for every county-scale land segment, all of the other hydrologic parameters must be established manually prior to automatic calibration runs, and remain constant. This study used values determined by the CBP (USEPA, 2010). Details on parameters are to be examined in Chapter 5.

Objective functions required for the automated calibration process are constructed such that each parameter is associated with a single objective function and calibrated by minimizing that objective function independently from other parameters (Table 4-1). In the CBW model calibration, the value of the individual objective function at the end of a model run is used to calculate an updating factor that multiplies the current parameter value to generate a new value for the next model run. All of the objective functions and

updating parameters are dimensionless. The relationships between the particular statistics used in objective functions and their targeted parameters are explained in USEPA (2010).

Ideally, each updating factor (φ in Table 4-1) would converge to 1.0, indicating that a parameter value has been identified that causes the simulation to match the observations by that particular measure. This is not always possible, as the user specifies a maximum number of iterations, after which the calibration is stopped.

Table 4-1. Objective functions, their targeted parameters, and updating factors for the autocalibration approach implemented in the CBW model.

Objective function	Targeted parameter (units)	Updating factor
TBIAS	LAND_EVAP	$\varphi_1 = \frac{2}{2 - TBIAS}$
SSTAT, WSTAT	LZSN (inch)	$\varphi_2 = \frac{2.5 - \frac{SSTAT}{WSTAT}}{1.5}$
BSTAT	INFILT (in/hr)	$\varphi_3 = \frac{1}{BSTAT}$
QavgRI	IRC	$\varphi_4 = \frac{2}{(1 + QaveRI)}$
BavgRI	AGWR	$\varphi_5 = \frac{2}{(1 + BaveRI)}$
Max (PBIAS,VBIAS)	INTFW	if PBIAS*VPBIAS >=0 $\varphi_6 = 1 + \frac{\max(PBIAS, VPBIAS)}{2}$ Otherwise, $\varphi_6 = 1$

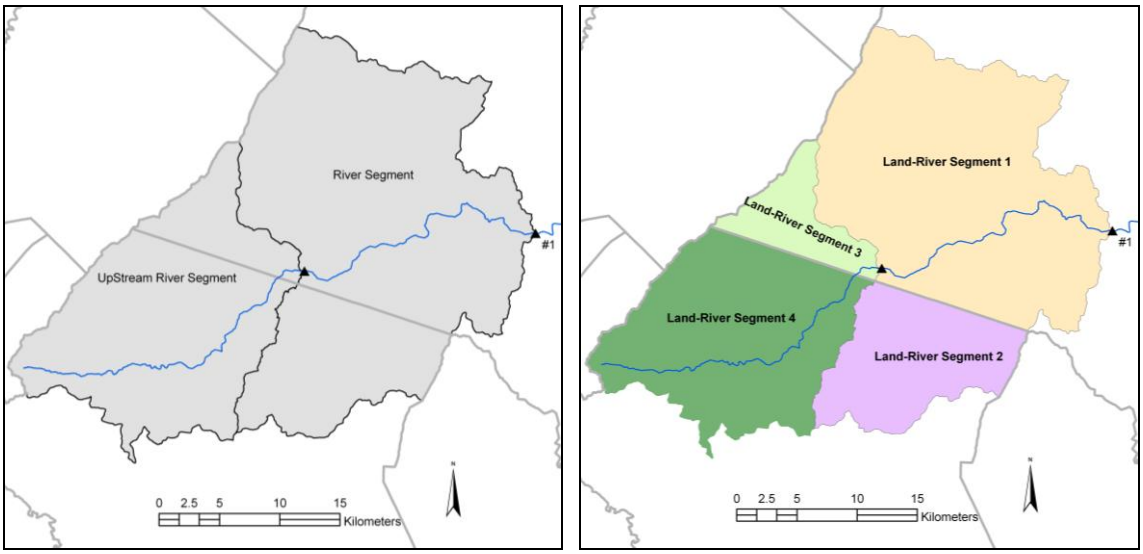
The optimization approach the CBP developed is similar to the classic multi-objective optimization. However, the CBP method has its own distinctive features: (1) the CBP approach treats different objectives as functions of individual target parameters, whereas the classic approach treats each objective as a function of all parameters; (2) The CBP method seeks a value of each objective function as close as possible to a specified

value, whereas the classic method seeks a maximum or minimum of an objective function; (3) objective functions are used to calculate updating factors to update parameters in the CBP method, whereas a gradient search is used to update parameters in the classic method; (4) the CBP method uses a user-specified maximum number of iterations as termination criteria, whereas the classic method uses numerical convergence criteria.

4.2.4 Link between land simulation and river simulation

The parameters involved in the autocalibration procedure correspond to county-based land segments. The land hydrologic fluxes are reallocated to watershed-based river segment. The calibration statistics are calculated by comparing flow simulated at the end of a reach segment with the observed streamflow at the corresponding stream gauge. As described in Section 2.4, the land and river segments do not necessarily agree with each other: a given county-based land segment may provide hydrologic fluxes to more than one river segment. Therefore, the parameter updating factors calculated after a model run in calibration do not correspond simply to specific land segments.

To explain the CBP updating procedure, a simplified case is illustrated in Figure 4-1. The CBW model programmatically identifies the upstream river segments of a given stream gauge station (USEPA, 2010), for example, the two river segments upstream of Station #1 in Figure 4-1(a). Then, the model creates a list of intermediate model units, called land-watershed segments, by intersecting the land and river segments (Martucci et al., 2006) as illustrated in Fig 4-1 (b). The model system has a look-up table that assigns a county-based precipitation data set to each land-watershed segment.



(a) (b)

Figure 4-1 Boundary disagreement between land segment and river segment. (a) A river segment and its upstream river segment for a calibration station #1 are presented with the land segments which are county boundaries; (b) the CBW model uses land-watershed segments to link land simulations to river simulations. In this case, four land-river segments contribute to the simulated streamflow at Station #1.

At the end of each calibration iteration run, the model produces the statistical matrix by comparing the simulated and observed flows at each calibration site and calculates the φ values, as listed in Table 4-1. The φ values serve as multipliers used to update each land segment-based parameter for the following iteration. As explained above, however, a single φ may reflect multiple land segments, and a single land segment may contribute to more than one φ. To apply the updating multipliers to the land parameters, the CBP established a weight scheme for cases when the relationship between the land and river segments does not completely match. For example, when multiple land segments exist in multiple adjacent river segments, one land segment can relate to multiple nested river segments. The CBP applies this weight scheme over the entire Chesapeake Bay.

Based on the CBP's documentation (USEPA, 2010), the weight scheme is illustrated as follows. Let land segment **A** drain to gauges X and Y. First, the importance of land segment **A** to the stream gauges is decided by the percentage of drainage area occupied by **A**. For example, let **A** be the only land segment for gauge station X and one of two (50%) of the land segments for gauge station Y. The relative contributions are then calculated (Equations 4-7-1 and 4-7-2). The parameters for **A** will be updated using the recommended multiplier by the weighted factor calculated (Equation 4-8).

$$\text{Relative importance of } \mathbf{A} \text{ to the gauge X} = \left[\frac{100}{100+50}\right] \text{x } 100 = 66.7 \qquad (4\text{-}7\text{-}1)$$

$$\text{Relative importance of } \mathbf{A} \text{ to the gauge Y} = \left[\frac{50}{100+50}\right] \text{x } 100 = 33.3 \qquad (4\text{-}7\text{-}2)$$

$$PRM^{new} = PRM^{old} \text{ x}(\varphi^{X} \text{ x } 66.7\% + \varphi^{Y} \text{ x } 33.3\%) \qquad (4\text{-}8)$$

where

PRM^{new} = new land parameter for the following calibration run
PRM^{old} = land parameter used in the current calibration run
φ^{X} = updating factor calculated based on calibration statistics at gauge X
φ^{Y} = updating factor calculated based on calibration statistics at gauge Y

4.3 Study Site

This study selected the Potomac River Basin (PRB) to examine the effect of using different types of precipitation inputs on the CBW model calibration. The Potomac River is the second largest tributary to the Bay, so the PRB is an important region in modeling the Bay. Details on the PRB can be found at the website of the Interstate Commission on the Potomac River Basin (ICPRB). Thirty-seven USGS stream gauge stations were selected as calibration sites, depending on the availability of observed data sets during the

study time period within the PRB (Figure 4-2). Eighteen stations were selected from the Upper Potomac, seven stations were from the Middle Potomac, and twelve stations were from the Shenandoah. This study used watersheds delineated by the CBP as model river units for the watershed simulation (Section 2.4).

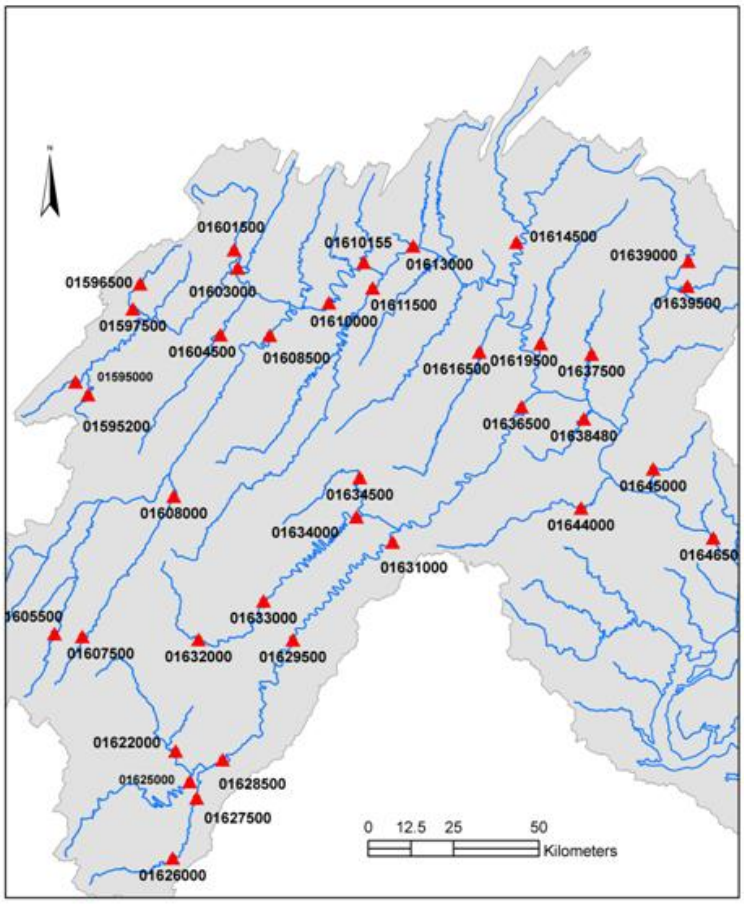


Figure 4-2. Calibration stations used in the model calibration.

4.4 Data

The data collected for this study were spatially aggregated on a model land segment (county-based) basis at an hourly scale, regardless of spatiotemporal resolution of the data sources. The data sets were divided into three calendar-year periods: (1) 2001 for spin-up; (2) 2002-2004 for calibration; (3) 2005 for validation.

For gauge-based precipitation inputs, this study used the mCBP prepared by the CBP (Chapter 2). According to EPA/CBP documentation (2010), the USGS National Research Program developed a methodology for estimating the spatial distribution of precipitation, as explained below. Physical characteristics such as latitude, longitude and elevation were employed as critical factors in developing a mathematical model (Equation 2-1). The regression was calibrated by month and adjusted to the mean observed daily precipitation volumes. The daily precipitation was calculated per 5x5 km^2 grid and then aggregated over the county-wide model unit. Daily values were then disaggregated into hourly data using a nearest neighbor approach based on the observed hourly data. If no observed hourly data were available, the daily values were divided by 24 (USEPA, 2010). The mCBP time series for this study were obtained from the CBP website (CBP, 2009).

The MPE time series obtained from the Middle Atlantic River Forecast Center was provided on a 4x4 km^2 grid basis at an hourly scale. The datasets were reprocessed into county-averaged MPE using the geographical information system (GIS) software and stored in a Watershed Data Management (WDM) format. They are called mMPE time series.

Observed daily streamflow time series were obtained from the CBP website (2009) and the USGS water data website (USGS, 2009) to evaluate the CBW model performance during the hydrologic calibration process. The land use data set was also obtained from the CBP website (2009). The land use categories (Table 3) were derived by the CBP from various sources such as the 2000 land cover data developed by the University of Maryland's Regional Earth Science Applications Center, the 2000 NLCD

map, and the agricultural census data. Details on land use are found in USEPA (2010). Meteorological data sets from 2001 to 2005, including air temperature, wind speed, solar radiation, and cloud cover were obtained in a WDM format from the CBP website (2009). The CBP compiled these input data sets from observed meteorological data obtained from the NCDC/NOAA (USEPA, 2010). Potential evapotranspiration (PET) time series are also available at the CBP website. PET was calculated using the Hamon method, which was recommended for the Chesapeake region (Lu et al., 2005 cited in USEPA 2010). Function tables required for the hydraulic routing were also obtained from the CBP website (CBP, 2009).

4.5 Methods

4.5.1 Calibration settings

This study explored how the model performance statistics change and model parameters adjust when the model is recalibrated using different types of precipitation. This chapter focuses on calibration and hydrologic statistics. Details on parameter values are discussed in Chapter 5. Several autocalibration runs were conducted to answer different questions (Table 4-2), including concerns about parameter constraints. Parameter constraints refer to the lower and upper limits on allowed parameter values. Establishing parameter constraints forces every parameter to stay within the user-defined range while the model is autocalibrated based on statistics. Without constraints, the autocalibration might satisfy statistical objectives by assigning non-physical or irrational parameter values.

If the termination criterion were convergence, as opposed to the user-defined number of iterations, the value of LZSN for example, would continue to decrease by an updating factor of less than 1. That way, statistical indices eventually converge at the end of the iteration process. On the other hand, if a new LZSN generated by multiplying the updating factor to the old value is beyond the lower and upper parameter boundaries, it would be reset to the violated boundary. In this case, no improvement in the statistical indices would be found during further iterations. In summary, there are two possibly influential factors for the final value of LZSN: the finite number of iterations and parameter constraints.

A series of 4 calibration runs (Runs 1 to 4) were set for 2002-2004. Six parameters listed in Table 4-1 were adjusted during the calibration process in each Run. All other model settings are the original CBW inputs, established based on data from 1985-2005 by the CBP. Run 1 uses mCBP with the parameter constraints set by the CBP (first row in Table 4-2). Run 2 replaces mCBP with mMPE and maintains the same parameter constraints. Run 3 uses mMPE, but with different parameter constraints to allow a broader range of the parameters. Run 4 uses the new parameter ranges, but with mCBP precipitation. To avoid any confounding factors, only one parameter (LZSN) was involved in changing the constraints (Table 4-2).

Table 4-2. Calibration settings used in this study.

Calibration	Precipitation input	Parameter constraints
Run 1	mCBP	CBP (LZSN: 8~12)
Run 2	mMPE	CBP (LZSN: 8~12)
Run 3	mMPE	NEW (LZSN: 2~12)
Run 4	mCBP	NEW (LZSN: 2~12)

4.5.2 Model calibration accuracy statistics

The model calibration accuracy was compared using calibration statistics that are objective functions in the calibration process: TBIAS, SSTAT, WSTAT, BSTAT, QaveRI, BaveRI, and PBIAS (VBIAS). Calibrated values of the objective functions are compared for four Runs (Table 4-2).

A two-sample Kolmogorov-Smirnov (2-sided) hypothesis test was applied to determine whether the simulated and observed streamflow are likely to be sampled from identical populations. The test is based on the statistic D, the maximum discrepancy between the two empirical cumulative distribution functions (ECDFs), $F(x)_{mMPE}$ and $F(x)_{mCBP}$ (Section 3.5.6), and evaluates the hypothesis defined by:

$\mathbf{H_o}$: The two data sets are from the same probability distribution.

$\mathbf{H_a}$: The two data sets are not from the same probability distribution.

If D is larger than the critical value for a given level of significance, the null hypothesis $\mathbf{H}_o$ is rejected, and the two data sets are not from the same probability distribution. The critical value for the D-statistic is determined by the effective number of samples, taking autocorrelation in streamflow into account:

$$D_\alpha = K(\alpha)\sqrt{\frac{n_{e1}+n_{e2}}{n_{e1}n_{e2}}} \tag{4-9}$$

where

K is a function of the selected level of significance (K=1.36 for α=0.05)

n_e is effective record length, as defined by the following expression:

$$n_e = \frac{n}{\left(\frac{1+\rho}{1-\rho}\right) - 2\rho\left(\frac{1-\rho^n}{n(1-\rho^2)}\right)} \tag{4-10}$$

where

n is actual record length and ρ is a lag-one auto correlation coefficient

Nash-Sutcliffe efficiency (NSE) is one of the most widely used statistics that assess the performance of a hydrologic model. It measures the correspondence between simultaneous observed and simulated flows. It is defined as follows:

$$\text{NSE} = 1 - \left[\frac{\sum_{i=1}^{n} (y_i^{obs} - y_i^{sim})^2}{\sum_{i=1}^{n} (y_i^{obs} - y^{obs,mean})^2} \right] \quad (4\text{-}11)$$

where

$y_i^{obs} = i^{th}$ observed stream flow

$y_i^{sim} = i^{th}$ simulated stream flow

$y^{obs,mean} = \frac{1}{n} \sum_{i=1}^{n} y_i^{obs}$

A value of 1 indicates a perfect model performance. The NSE can take negative values if the sum of squared errors introduced by the model (Numerator in Equation 4-11) exceeds the variance of the measured variable (denominator in Equation 4-11); in such cases, the average of the observed variable is a better predictor than the model. If the prediction of a linear model is biased, NSE values can be negative (McCuen et al., 2006). Studies show that NSE index should accompany other measures to describe model performance, because a low value of NSE may not necessarily mean a poor model performance and vice versa (Jain and Sudheer, 2008; McCuen et al., 2006).

Relative standard error of estimate (Se/Sy) measures the model accuracy relative to predictability when using the mean value of the dependent variable, y. It quantifies the reduced error by using the model instead of the mean value of the quantity being

predicted. When Se/Sy is near 0, the model significantly improves the accuracy of prediction over the prediction made with the mean (McCuen, 1993). Se/Sy can be applied to both linear and non-linear models.

$$\text{Se/Sy} = \frac{\left[\frac{1}{\nu}\sum_{i=1}^{n}(y_i^{obs} - y_i^{sim})^2\right]^{0.5}}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(y_i^{obs} - y^{obs,mean})^2\right]^{0.5}} \quad (4\text{-}12)$$

where ν is the degree of freedom (*n-1* in this study).

4.5.3 Statistics on flow characteristics

Richards-Baker Flashiness(R-B) index measures flow flashiness, indicating the frequency and rapidity of short term changes in streamflow (Baker et al., 2004). Baker and others (2004) derived this index by dividing the sum of the absolute value of daily change in *y* by total *y* for the corresponding time period (2002-2004 in this study):

$$\frac{\sum_{i=1}^{n}|y_i - y_{i-1}|}{\sum_{i=1}^{n} y_i} \quad (4\text{-}13)$$

where

y_i =mean daily flow in cfs for a given day, *i*
n = number of days of recorded data at a given gauging station

$R_{Q10\text{-}90}$ is the ratio of the discharge equaled or exceeded 10% of the time to the discharge equaled or exceeded 90% of the time (Baker et al., 2004). The percentiles are based on the entire calibration time period. The higher the value of $R_{Q10\text{-}90}$, the greater the flow variability.

$$R_{Q10\text{-}90} = \frac{y_{10}}{y_{90}} \tag{4-14}$$

This study also calculated $R_{Q25\text{-}75}$ to examine the flow variability in moderate streamflows.

Lag- one autocorrelation (ρ) measures the degree of correlation between values separated by one time interval, meaning adjacent values (McCuen, 2003).

$$\rho = \frac{\Sigma_{i=1}^{n-1} y_i y_{i+1} - \frac{1}{(n-1)}\left(\Sigma_{i=1}^{n-1} y_i\right)\left(\Sigma_{i=2}^{n} y_i\right)}{\left[\Sigma_{i=1}^{n-1} y_i^2 - \frac{1}{(n-1)}\left(\Sigma_{i=1}^{n-1} y_i\right)^2\right]^{0.5}\left[\Sigma_{i=2}^{n} y_i^2 - \frac{1}{(n-1)}\left(\Sigma_{i=2}^{n} y_i\right)^2\right]^{0.5}} \tag{4-15}$$

where

y_i = mean daily flow in cfs for a given day, i
n = number of recorded days at a given calibration station

4.6 Results

In the CBP's autocalibration approach, several single-component objective functions are used to calibrate individual target parameters independently, using parameter updating rules based on the intermediate values of those objective functions (Table 4-1). This section examines the values of the objective functions obtained in calibration. The hydrologic statistics on model output after autocalibration runs are then examined to evaluate model calibration accuracy. Performance statistics are also calculated for a one-year period not used in the calibration. Due to the limited sample size, statistics for the validation period are presented for reference only. The statistical comparison is conducted for all four runs (Table 4-2).

4.6.1 Automated model calibration statistics

The first objective function, TBIAS, is an index for water balance: the model must adjust evaporation and streamflow to balance the volume of precipitation. The land segment parameter LAND_EVAP is adjusted to move TBIAS toward its optimal value of zero. As TBIAS approaches zero, the updating factor φ_1 converges to 1 and the value of parameter LAND_EVAP is accepted as the optimal value (Table 4-1). At the majority of the 37 calibration stations, Runs 2 and 3 show smaller TBIAS than Runs 1 and 4 (Table 4-3).

Table 4-3. Final values of objective functions for four calibration runs.

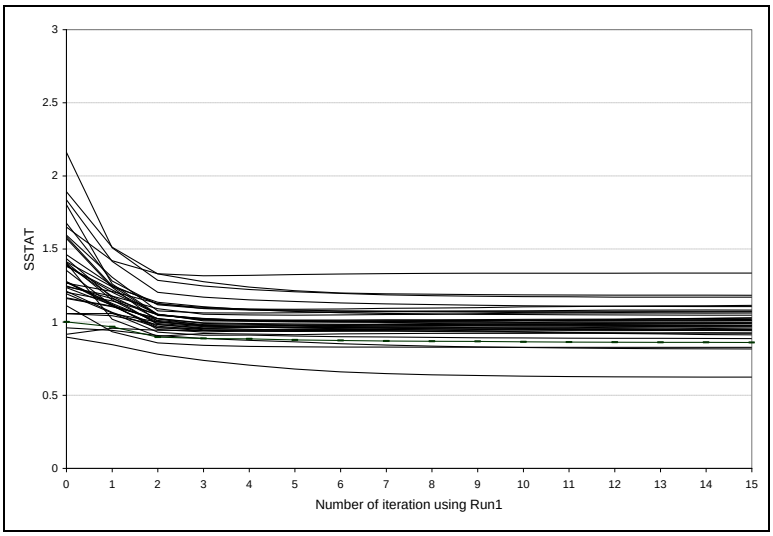
	Objective function	Calibration setting	Run1	Run2	Run3	Run4
		Precipitation	mCBP	mMPE	mMPE	mCBP
		Parameter constraints	LZSN: 8~12	LZSN: 8~12	LZSN: 2~12	LZSN: 2~12
1	*TBIAS*	Q1	-0.009	-0.053	-0.025	-0.003
		Median	0.002	0.001	0.003	0.101
		Q3	0.071	0.019	0.045	0.186
2	*SSTAT / WSTAT*	Q1	0.999	0.947	0.924	0.972
		Median	0.998	1.277	1.040	0.977
		Q3	0.992	1.194	1.110	0.948
3	*BSTAT*	Q1	0.983	0.986	0.989	0.991
		Median	1.013	1.003	1.008	1.016
		Q3	1.046	1.034	1.034	1.051
4	*QaveRI*	Q1	0.977	0.984	0.985	0.986
		Median	1.014	1.016	1.020	1.027
		Q3	1.066	1.046	1.060	1.060
5	*BaveRI*	Q1	0.999	0.998	0.999	0.999
		Median	1.001	1.001	1.001	1.001
		Q3	1.004	1.004	1.005	1.003
6	*PBIAS*	Q1	-0.042	-0.170	-0.159	-0.044
		Median	0.043	-0.094	-0.085	0.070
		Q3	0.115	-0.019	-0.170	-0.042
	VPBIAS	Q1	-0.042	-0.173	-0.118	-0.05
		Median	0.045	-0.052	-0.008	0.102
		Q3	0.113	0.144	-0.173	-0.042

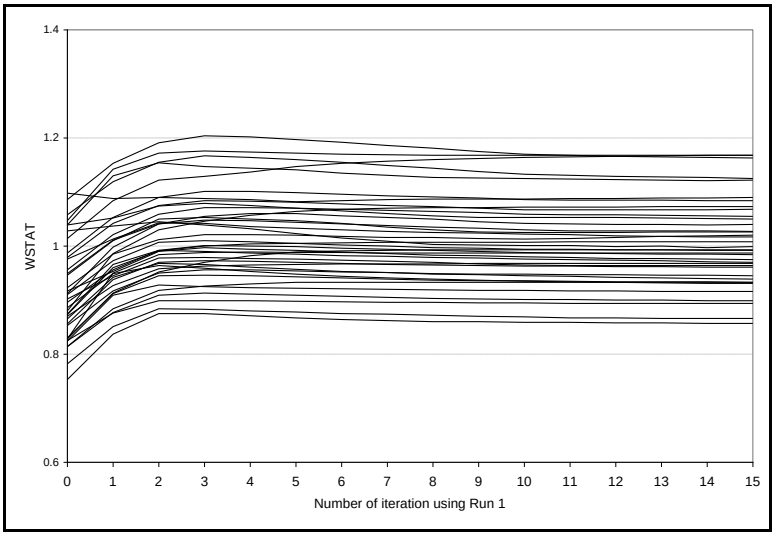
Note: This table shows the lower and upper quartiles (25^{th} and 75^{th} percentile) and medians of 37 values from all calibration sites.

The autocalibration attempts to adjust LZSN to move the objective function 2 (the ratio of SSTAT to WSTAT) to a value of 1. In Run 1, the ratio approaches 1, but Run 2 shows SSTAT to be higher than WSTAT, making the ratio greater than 1 (Table 4-3). In Run 4, the objective function 2 becomes less than 1, making the updating factor φ_2 less than 1 as well, and therefore, LZSN still decreases even at the end of the calibration, not converging to an optimum value.

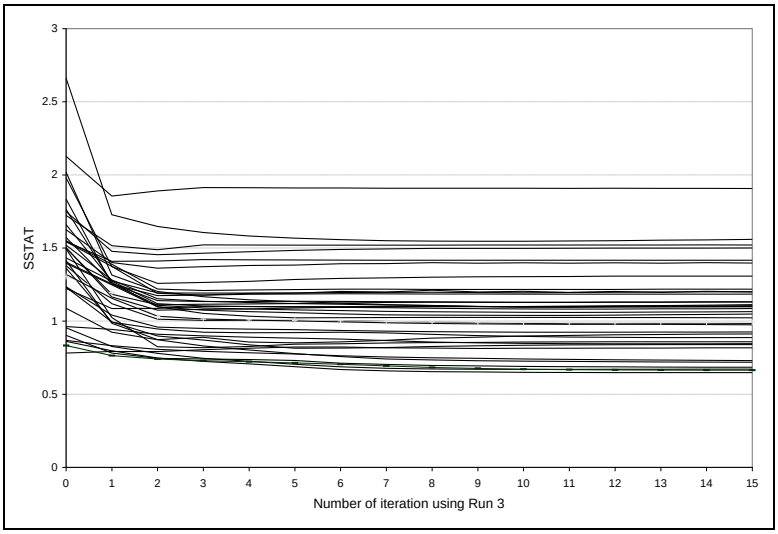
To determine whether the final values in objective function 2 and LZSN were controlled by convergence or maximum number of iterations at a given calibration site (Section 4.5.1), the trajectories of changes in SSTAT and WSTAT during the iteration processes of Runs 1 and 3 are examined (Figure 4-3). Regardless of the magnitude of total bias, SSTAT and WSTAT in Run 1 approach a certain value by the third iteration and remain the same for the rest of iterations at the majority of the calibration sites. Whether CBP parameter constraints (Run 1) or NEW parameter constraints (Run 3) are used, the pattern of the trajectory is similar. This visual analysis indicates that the number of iterations performed, 15 in this study, is not a factor in deciding the final value of LZSN. The early convergence of the statistical values, SSTAT and WSTAT will be further discussed in the following chapter in its examination of LZSN-value changes during iterations.

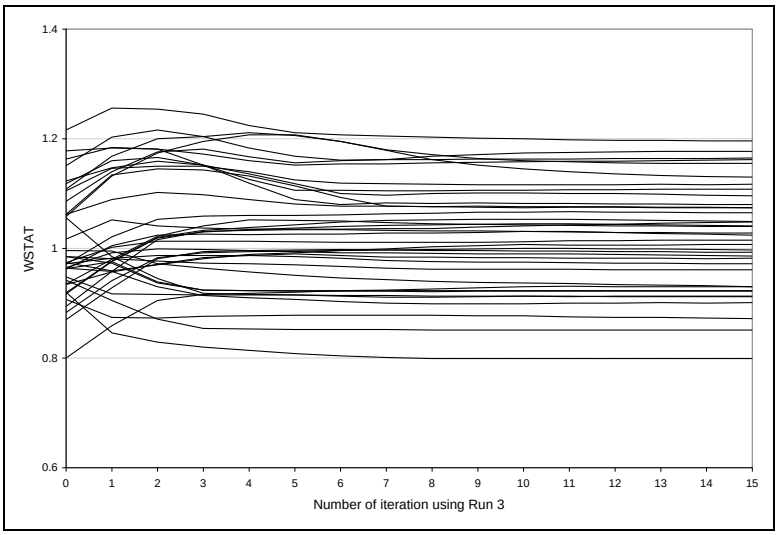
Figure 4-3. Values of SSTAT (upper panel) and WSTAT (lower panel) during the autocalibration process using (a) Run 1 and (b) Run 3.


SSTAT
Number of iteration using Run1


WSTAT
Number of iteration using Run 1

(a)


SSTAT
Number of iteration using Run 3



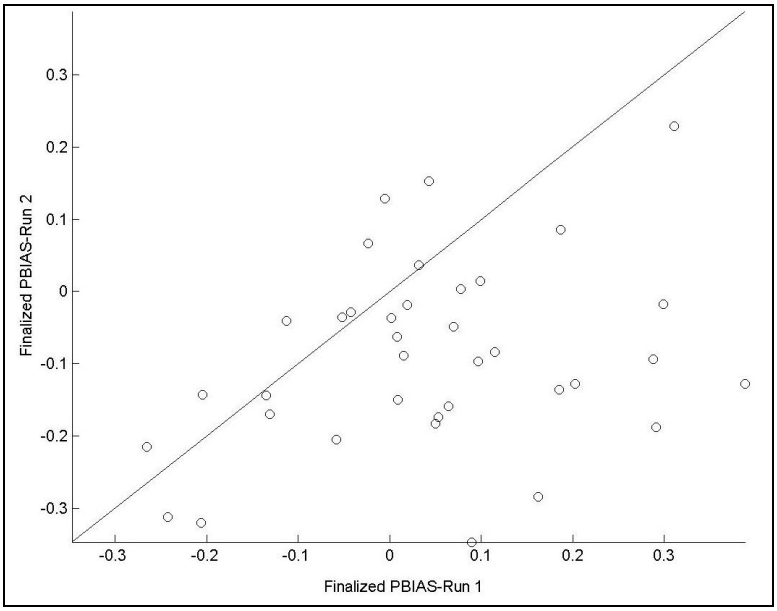
(b)

The objective function BSTAT should approach a value of one. This value is sought by changing the INFILT parameter. Runs 2 and 3 show a slightly smaller final value of BSTAT when mMPE is used, regardless of parameter constraints (Table 4-3).

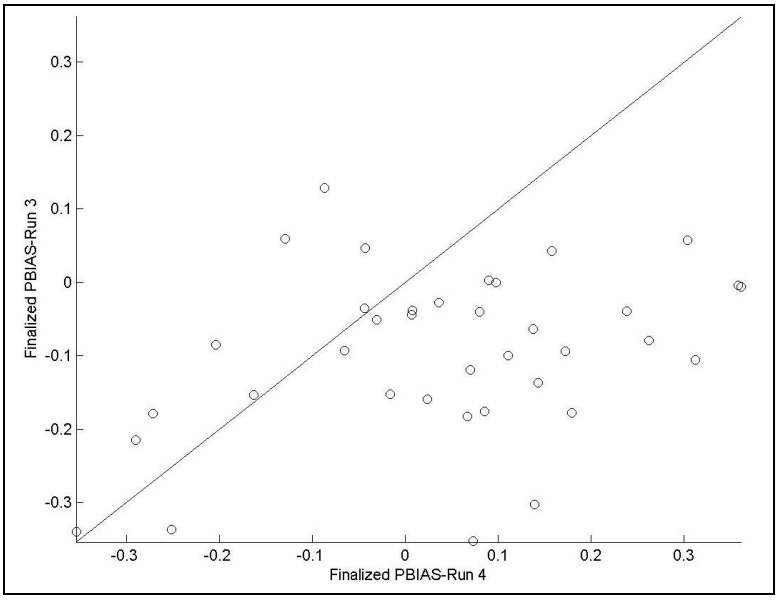
The objective functions 4 and 5 are related to recession indices. These functions are used to estimate IRC and AGWR, controlling interflow and groundwater flow, respectively. The final values of these recession indices among the Runs are similar (Table 4-3). The statistics on baseflow and recession indices seem to be insensitive to both the type of precipitation and parameter constraints.

The final PBIAS is greater in Runs 1 and 4 than Runs 2 and 3, respectively (Figure 4-4). Runs 2 and 3 underestimate the observed peak flow at the majority of the calibration sites, showing that most of points stay on the negative side on the y-axis (Figure 4-4). Run 3, with NEW parameter constraints, performs better than Run2 does by this measure. Considering that the mMPE time series has lighter peak depths than the

mCBP does (Section 3.6.3), model calibration accuracy by this measure is affected mostly by the type of precipitation.



(a)



(b)

Figure 4-4. Plot of the final Peak bias statistics for calibration runs: (a) Run 1 and Run 2; (b) Run 4 and Run 3.

The autocalibration procedure does not include a constraint to minimize the overall bias from each hydrologic component: the model system handles several objective functions independently, as opposed to having a single multi-term objective function. Therefore, individual biases are reflected in TBIAS, but not compensated for biases coming from other hydrologic components. The bias from Peak flow (or peak volume) is rather high among biases from other hydrologic components, which is not unusual, given that the variation of peak flows is great from event to event.

Relaxing the LZSN parameter constraints with mMPE precipitation slightly improves the calibration accuracy (Run 3 improvement over Run 2). On the other hand, the model calibration accuracy, as measured by calibration objective functions such as TBIAS and SSTAT/WSTAT, is mainly affected by the type of precipitation, regardless of parameter constraints. TBIAS in Runs 2 and 3 (using mMPE) is smaller than that in Runs 1 and 4 (using mCBP). As measured by the ratio of SSTAT to WSTAT, Runs 1 and 4 perform better than Runs 2 and 3.

4.6.2 Hydrologic statistics

An essential step of post-hydrologic calibration is to examine the model accuracy using statistics of the output function (streamflow, in this case) that are not used as calibration objectives. In this section, the overall statistical distributions of simulated and observed streamflow are compared. Then, Goodness-of-fit (GOF) statistics are evaluated for the efficiency of model calibration. Statistics based on selected flow characteristics also provide information on the calibration accuracy.

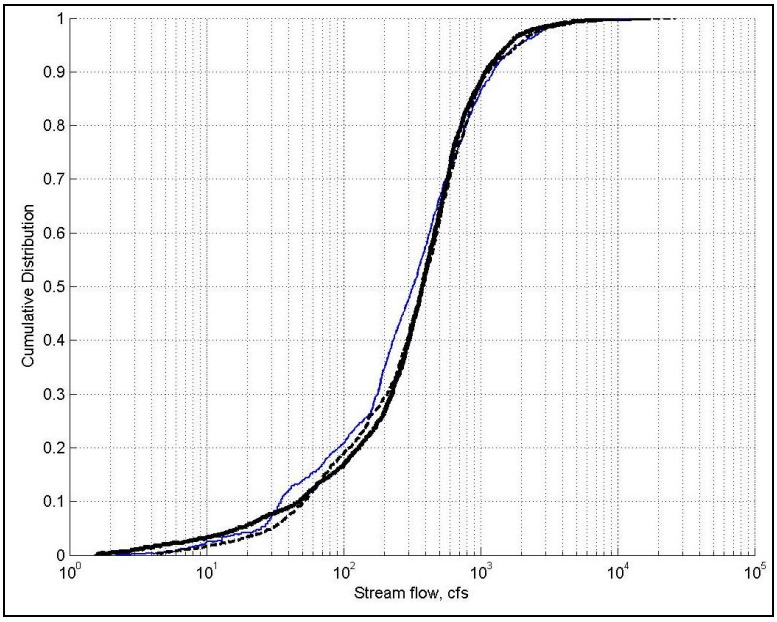
To examine the agreement of the overall distribution between the simulated and observed streamflow for the calibration period, the KS-test was performed at each calibration station (Table 4-4). When the D value is calculated based on the actual record length, the null hypothesis is rejected for nearly all of the calibration sites, indicating that the simulated and observed streamflows are not from the same probability distribution. When autocorrelation is taken into account, the effective sample size is reduced. Using the effective sample size, 54 to 70% of the calibration sites show that the simulated (using either mCBP or mMPE) and observed streamflow are from the same probability distribution. Using mCBP, Run 1 has a better performance than Run 4 does: more calibration sites show that simulated and observed flows are from the same distribution. Using mMPE, both Runs 2 and 3 have higher calibration accuracies by this measure than Runs 1 and 4.

Table 4-4. Number of calibration sites where the hypothesis of equality in the KS-test is accepted (rejected) at a level of significance of 0.05.

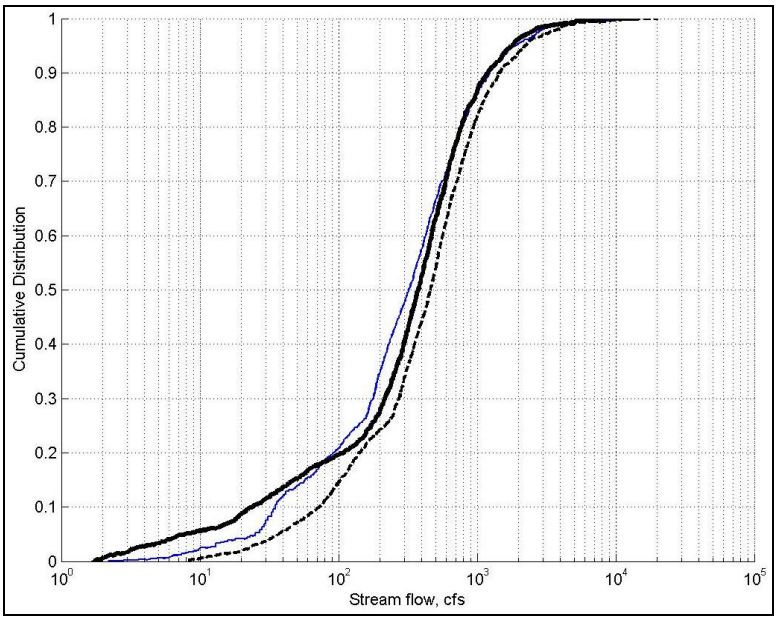
Record length used	Null hypothesis accepted (rejected)			
	Run1	Run2	Run3	Run4
Actual *n* (1096 days)	3 (34)	3 (34)	3 (34)	2 (35)
Effective *n* (36 to 662 days, depends on site-specific autocorrelation)	25 (12)	27 (10)	26 (11)	20 (17)

Figure 4-5 illustrates ECDFs, $F(x)_{mMPE}$ and $F(x)_{mCBP}$ at USGS gauge 1633000 as an example in which the null hypothesis is accepted in Runs 1 and 2. The hypothesis is also accepted in Run 3, but not in Run 4. Throughout the entire range of flow, Run 4 consistently overestimates the observed flow after the recalibration with NEW parameter constraints [dotted curve in Figure 4-5 (b)]. This suggests that the parameter constraints

prepared for the model using mMPE are not appropriate for the recalibration of the model using mCBP at this calibration site.



(a)



(b)

Figure 4-5. ECDFs of simulated streamflow at USGS 1633000 using the mMPE (—) and mCBP (···) with (a) CPB parameter constraints; (b) NEW parameter constraints. The thin solid line is for the observed stream flow.

To evaluate model calibration accuracy, the Nash-Sutcliffe coefficient (Table 4-5) and relative standard errors of estimate (Table 4-6) were examined. The NSEs for the three hydrologic components, all flow, quick flow, and baseflow, are summarized (Table 4-5).

Table 4-5. Nash-Sutcliffe coefficient (NSE) on daily simulated and observed flows at 37 calibration sites for all four runs.

NSE		Statistics of 37 calibration sites	Run1	Run2	Run3	Run4
calibration	All flow	Min	0.109	-0.020	0.054	0.016
		Q1	0.477	0.414	0.525	0.493
		Median	0.566	0.557	0.614	0.567
		Q3	0.681	0.637	0.688	0.695
		Max	0.775	0.782	0.839	0.774
	Baseflow	Min	0.264	-0.008	-0.218	-0.008
		Q1	0.505	0.528	0.424	0.438
		Median	0.641	0.611	0.583	0.581
		Q3	0.742	0.700	0.674	0.697
		Max	0.857	0.810	0.857	0.813
	Quickflow	Min	-0.611	-0.636	-0.474	-0.250
		Q1	0.390	0.367	0.497	0.351
		Median	0.483	0.509	0.599	0.513
		Q3	0.633	0.614	0.668	0.627
		Max	0.758	0.790	0.859	0.733
validation	All flow	Min	0.175	-0.936	-0.911	-0.377
		Q1	0.417	0.259	0.344	0.390
		Median	0.623	0.544	0.557	0.543
		Q3	0.824	0.644	0.645	0.748
		Max	0.922	0.836	0.828	0.873
	Baseflow	Min	-0.206	-1.823	-1.888	-0.373
		Q1	0.593	0.278	0.230	0.584
		Median	0.757	0.418	0.427	0.696
		Q3	0.838	0.576	0.603	0.790
		Max	0.928	0.817	0.835	0.915
	Quickflow	Min	-0.082	-1.762	-1.720	-1.001
		Q1	0.305	0.144	0.195	0.259
		Median	0.443	0.475	0.468	0.404
		Q3	0.754	0.584	0.560	0.686
		Max	0.901	0.819	0.796	0.814

Note: Based on a rule of thumb accepted by many professionals, the criterion for a good model performance is a NSE greater than 0.5.

Table 4-6. Relative standard error of estimates (Se/Sy) by calibration stations.

Se/Sy	CBP parameter constraints		NEW parameter constraints	
USGS ID	Run 1	Run 2	Run3	Run 4
1595000	0.722	0.771	0.725	0.705
1595200	0.666	0.766	0.758	0.677
1596500	0.695	0.796	0.779	0.671
1597500	0.847	1.009	0.959	0.755
1601500	0.666	0.608	0.595	0.609
1603000	0.59	0.715	0.689	0.54
1604500	0.687	0.625	0.57	0.812
1605500	0.732	0.791	0.678	0.799
1607500	0.755	0.936	0.706	0.658
1608000	0.654	0.765	0.614	0.644
1608500	0.66	0.69	0.587	0.712
1610000	0.588	0.649	0.559	0.6
1610155	0.58	0.745	0.747	0.576
1611500	0.867	0.681	0.677	0.939
1613000	0.66	0.601	0.536	0.706
1614500	0.486	0.56	0.557	0.484
1616500	0.648	0.778	0.797	0.671
1619500	0.495	0.665	0.661	0.49
1622000	0.845	0.884	0.658	0.682
1625000	0.683	0.602	0.569	0.726
1626000	0.565	0.491	0.463	0.573
1627500	0.526	0.468	0.448	0.54
1628500	0.502	0.511	0.42	0.544
1629500	0.493	0.475	0.401	0.528
1631000	0.474	0.494	0.42	0.517
1632000	0.767	0.76	0.64	0.766
1633000	0.789	0.707	0.575	0.81
1634000	0.932	0.75	0.622	0.991
1634500	0.857	0.638	0.65	0.88
1636500	0.515	0.564	0.484	0.552
1637500	0.498	0.974	0.972	0.475
1638480	0.708	0.602	0.601	0.7
1639000	0.631	0.694	0.694	0.628
1639500	0.666	0.664	0.664	0.668
1644000	0.619	0.633	0.633	0.604
1645000	0.595	0.623	0.616	0.594
1646500	0.502	0.504	0.487	0.516

Note: Based on a rule of thumb accepted by many professionals, the criteria for a good model performance is Se/Sy less than 0.3. The criteria for moderate performance can be between 0.3 and 0.7.

For both all flow and quick flow, Run 3 outperforms the other runs at the majority of the calibration sites. For baseflow, NSE in Run 1 is higher than in Run 3, although all runs show similar performance by the calibration statistics, BSTAT (Table 4-3). The reason is, as mentioned in Section 4.2.2, that BSTAT does not consider information on simultaneous agreement. For the validation period, Run 1 outperforms the other runs. Run 4 using the same precipitation as Run 1 (mCBP) performs better than Run 2 or Run 3.

The relative standard error of estimates Se/Sy on all flow was also calculated and summarized by stations (Table 4-6). All four runs fail to show a Se/Sy less than 0.3 at any stations, indicating that no calibration station exists with high calibration accuracy by this measure. However, the model using mMPE performs better than when using mCBP: Run 3 outperforms Run 1 at 24 of the 37calibration sites (65%). With the mMPE input, Run 3 performs better than Run 2, showing Se/Sy decreases at 31 of the 37 calibration sites (84%) when using NEW. With the mCBP input, Run 1 (using CBP parameter constraints) performs better than Run 4 (using NEW parameter constraints) at 20 out of the 37 calibration sites.

To quantify the relative changes, differences in Se/Sy between Runs were calculated and plotted (Figure 4-6). Se/Sy increases by as much as 18% when the CBW model using mCBP uses the NEW parameter constraints. This suggests that the CBW model performs better with CBP parameter constraints when the input forcing is mCBP precipitation. On the other hand, Se/Sy decreases by as much as 25% when the CBW model using mMPE uses NEW. This indicates that the CBW model is better calibrated with NEW parameter constraints when the input forcing is mMPE precipitation. When

the parameter settings are set appropriately for different forcing inputs, the model is better calibrated: Run 1 outperforms Run 2, and Run 3 outperforms Run 4.

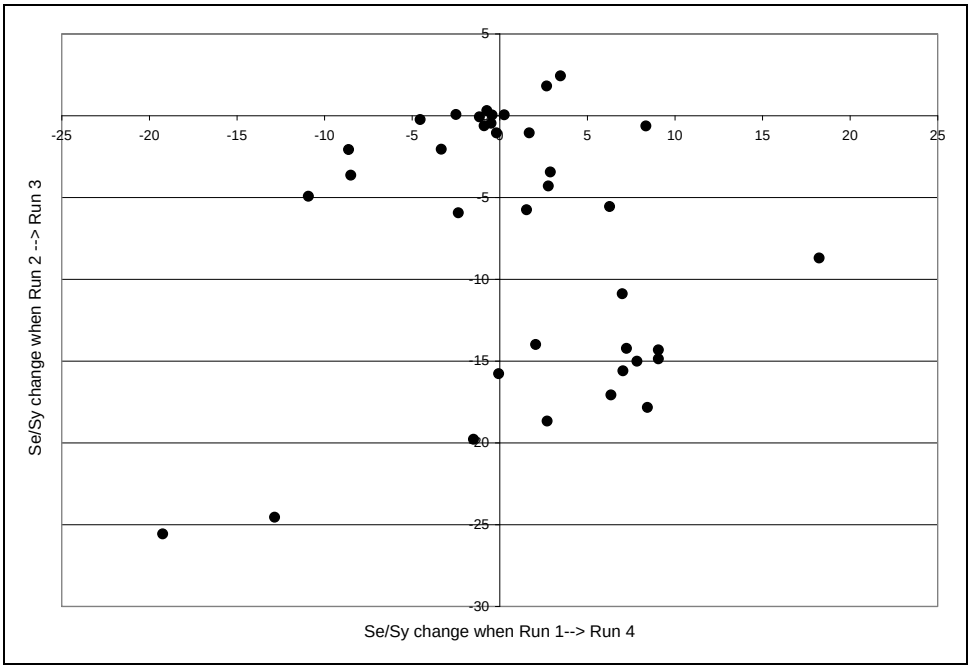


Figure 4-6. Se/Sy change from Run 1 to Run 4 and Se/Sy change from Run 2 to Run 3. The quantity on the x-axis is the relative difference in Se/Sy between Run1 and Run4. The quantity on the y-axis is the relative difference in Se/Sy between Run2 and Run3.

To examine the spatial distribution of the model efficiency, the Se/Sy for Run1 and Run 3 are mapped (Figures 4-7 and 4-8). The map uses the average annual difference between mMPE and mCBP from Chapter 3 to identify the relationship, if any, between model efficiency and precipitation. Overall, Run 1 performs better in the Middle Potomac River, whereas Run 3 performs better around the Shenandoah River, which is a principal tributary of the Potomac. Both Runs 1 and 3 show a high Se/Sy in the western headwaters of the Potomac River. This area also shows significant differences between mMPE and mCBP. Neither one of the precipitation data sets might be accurate for this region, causing poor performances by both Runs 1 and 3. Because of complex error sources in estimating precipitation using the radar system in the

mountainous region (Meischner, 2004), the mMPE time series may not be accurate. The mCBP values may also not be accurate in generating the mean areal product from gauge-point data in this mountainous region. Overall, any spatial patterns of model performance over the PRB are not found. It is also difficult to identify a relationship between calibration accuracy and differences in mMPE and mCBP.

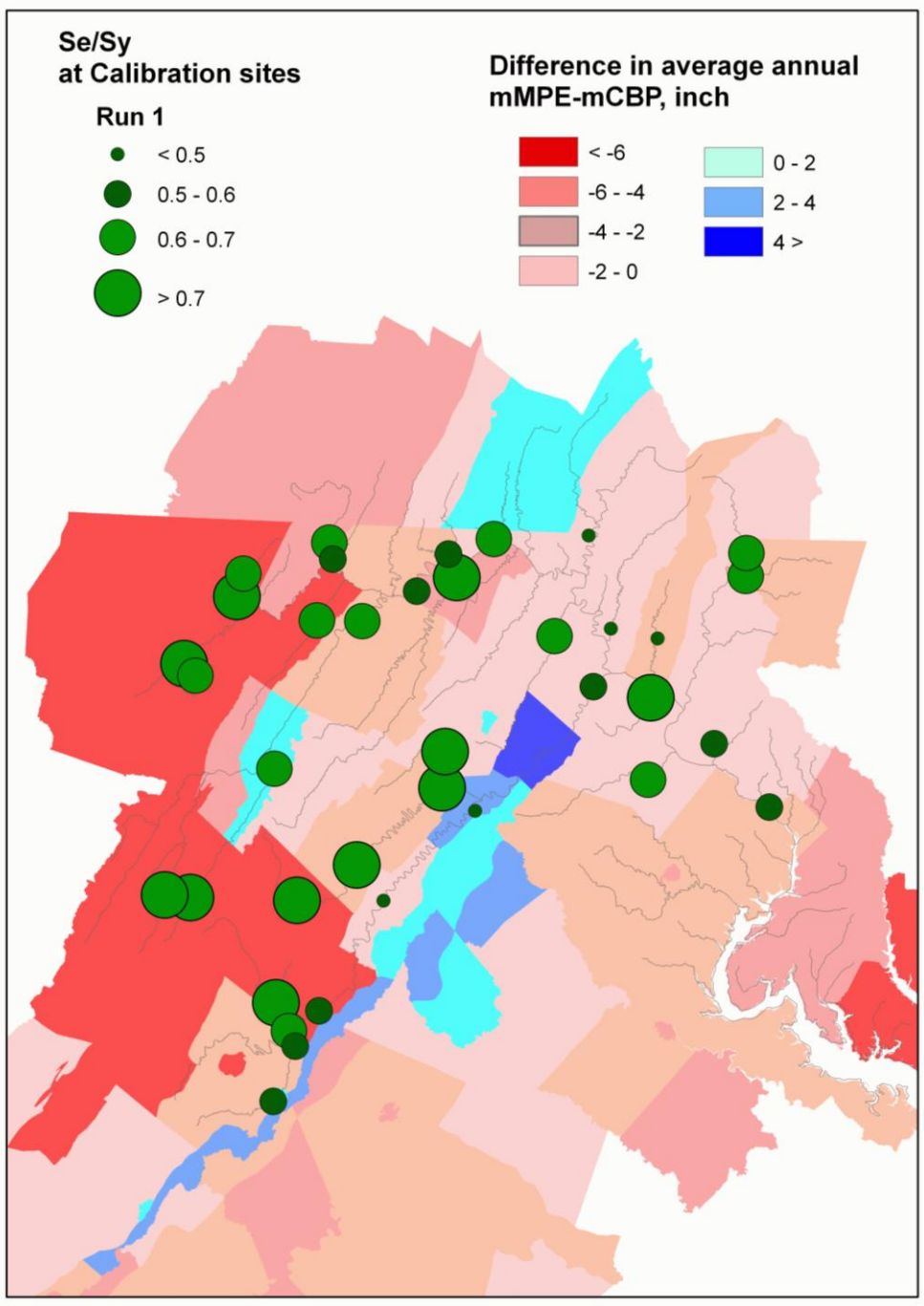


Figure 4-7. Spatial distribution of Se/Sy for Run 1 using mCBP. The difference in annual precipitation is based on Figure 3-17.

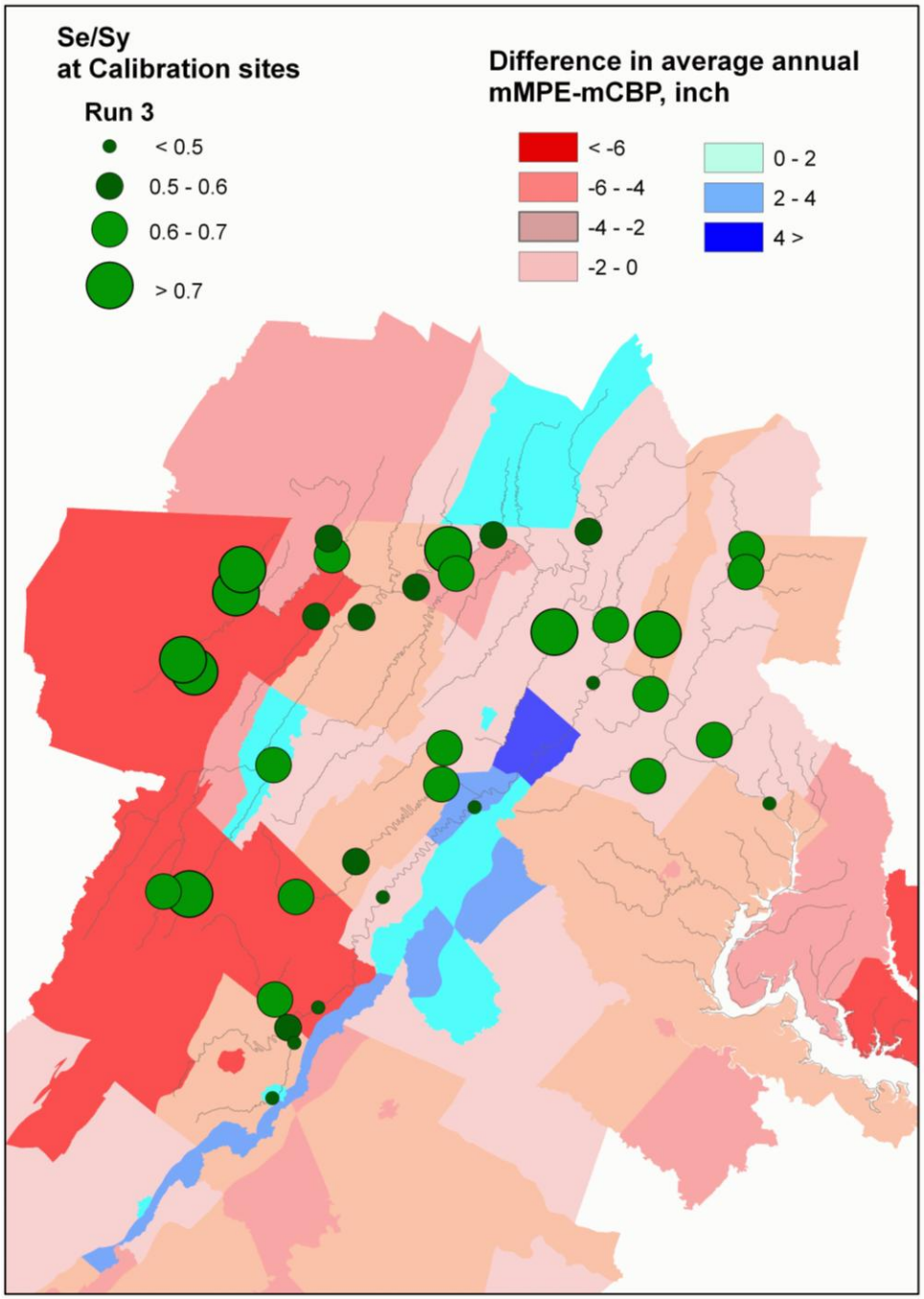


Figure 4-8. Spatial distribution of Se/Sy for Run 3 using mMPE. The difference in annual precipitation is based on Figure 3-17.

In addition to GOF statistics, flow characteristics were also used to evaluate model calibration accuracy (Table 4-7). For flow variability based on $\mathbf{R}_{10\text{-}90}$, the indices in Runs 1 and 4 show better agreement with those in the observed flow than Runs 2 and3.

However, Run3 agrees better with the observed flow variability using moderate flow, $\mathbf{R}_{25\text{-}75}$. It seems that characteristics of extremely low or high flow are well represented by flows simulated by the model using mCBP.

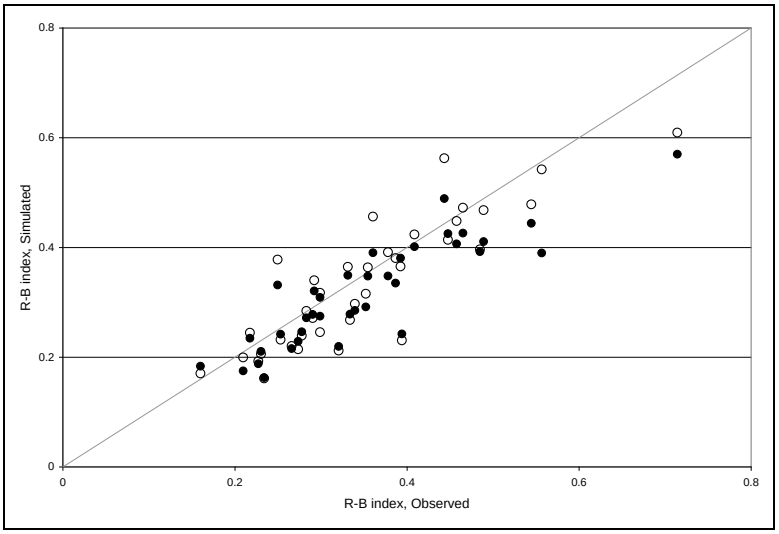
Table 4-7. Characteristics of observed and simulated flows.

characteristics	Statistics of 37 calibration sites	Observed	Run1	Run2	Run3	Run4
$\mathbf{R}_{10\text{-}90}$	Q1	0.031	0.048	0.027	0.021	0.049
	median	0.056	0.057	0.041	0.032	0.057
	Q3	0.073	0.079	0.063	0.062	0.075
$\mathbf{R}_{25\text{-}75}$	Q1	0.170	0.232	0.226	0.198	0.238
	median	0.228	0.250	0.271	0.248	0.275
	Q3	0.255	0.299	0.290	0.276	0.294
R-B index	Q1	0.274	0.232	0.253	0.242	0.215
	median	0.334	0.317	0.311	0.309	0.301
	Q3	0.409	0.414	0.399	0.390	0.388
Lag-one Autocorrelation (ρ)	Q1	0.561	0.477	0.517	0.606	0.582
	median	0.688	0.618	0.642	0.652	0.721
	Q3	0.776	0.763	0.745	0.801	0.827

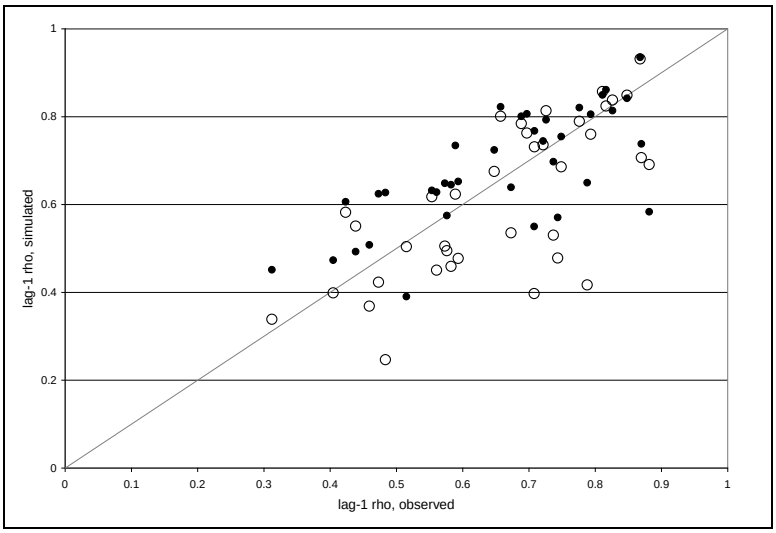
The R-B index measures flashiness in flow (Baker et al., 2004). When the flow has low/moderate flashiness, the indices for both Runs 1 and 3 agree well with observed ones at the majority of the calibration site [Figure 4-9 (a)]. However, the model using mCBP better represents high flashiness than the model using mMPE.

With either CBP or NEW parameter constraints, the model using mMPE better captures the autocorrelation of observed flow at the majority of the calibration sites. Run3 shows near to or slightly higher indices than ρ in observed flow, whereas Run1 tends to show the opposite trend [Figure 4-9(b)]. In other words, the simulated flow by Run 3 has lower temporal variability than the simulated flow by Run 1, but similar to that in the observed flow. Generally speaking, the model using mMPE represents flow

characteristics more similarly to the observed flow than the model using mCBP, except for extreme flow.



(a)



(b)

Figure 4-9. Comparison of (a) R-B index; (b) Lag-one autocorrelations between observed flow and simulated flow by Run1 (○) and Run3 (●) at 37 calibration sites.

4.6.3 Referenced work

The model calibration accuracy in Runs 1 and 3 was compared to the results from the master's thesis *Calibrating Shenandoah Watershed Swat Model Using A Nonlinear Groundwater Algorithm* (Wang, 2011).

The goal of this chapter was to explore possibilities for improving model calibration accuracy using a modified nonlinear-reservoir groundwater algorithm. A modified version of SWAT (Soil Water Assessment Tool), iSWAT was calibrated for the Shenandoah River basin, which lies within the Potomac River Basin. The model was forced with daily station precipitation and meteorological data, using a simple nearest-neighbor interpolation. The Parameter Estimation Software Tool (PEST) was coupled with the model to evaluate its optimization method using different weighting schemes. The best results obtained from several calibration runs are reported in Table 4-8.

Table 4-8. NSE by calibration sites in the Shenandoah River Basin.

USGS ID	*i*SWAT model	Run1-CBWmodel	Run3-CBWmodel
1622000	0.596	0.534	0.566
1625000	0.387	0.473	0.676
1626000	0.445	0.672	0.786
1627500	0.500	0.708	0.799
1628500	0.706	0.704	0.824
1629500	0.689	0.722	0.839
1631000	0.565	0.733	0.823
1632000	0.368	0.413	0.590
1634000	0.482	0.016	0.614
1634500	0.314	0.225	0.578
1636500	0.482	0.695	0.765

Based on the NSE indices in the comparison matrix, the CBW model using mMPE with NEW parameter constraints outperforms the other two model calibrations at all of the calibration stations except for one (station 1622000). Whereas the indices from

iSWAT model range from 0.31 to 0.71 across the region, Run3 shows indices greater than 0.5, ranging from 0.57 to 0.84. This indicates a good performance, based on a rule of thumb accepted by professionals.

Because of the different simulation time periods (1996-2006 vs. 2002-2004), a direct comparison between these studies is not appropriate. However, this comparison provides a general idea of how well the hydrologic models perform in this region.

4.7 Discussion and Conclusion

This chapter examined the calibration accuracy of the model using two different precipitation input data sets, mCBP and mMPE. The model calibration accuracy was compared by both calibration and hydrology statistics. Questions proposed at the beginning of the study are discussed in this section.

4.7.1 Comparison of calibration statistics

Based on calibration statistics related to total bias and seasonal biases such as TBIAS, the ratio of SSTAT to WSTAT, the calibration accuracy is closely related to the type of precipitation data. TBIAS in Runs 2 and 3 (using mMPE) is smaller than that in Runs 1 and 4 (using mCBP). As measured by the ratio of SSTAT to WSTAT, Runs 1 and 4 perform better than Runs 2 and 3 (Table 4-3).

The updating factor for seasonal flow bias (φ_2) does not reach an optimum value, showing that the calibration process does not affect the calibration accuracy at an early stage of the iterative process. Regarding PBIAS or VBIAS, the model using mMPE underestimates the observed peak flow (or volume) at the majority of the calibration sites,

reflecting the fact that mMPE time series has lighter peak depths than the mCBP does in the previous study. The recalibration of the model using mMPE with NEW parameter constraints improves the calibration accuracy (Figure 4). Based on statistics on baseflow or recession indices, the calibration accuracy is affected by both the type of precipitation and the parameter settings. When the model is calibrated or recalibrated with its appropriate parameter constraints, the calibration accuracy increases.

4.7.2 Comparison of hydrologic statistics

Based on the KS-test on the overall distribution, Runs 2 and 3 show higher calibration accuracies than Runs 1 and 4, indicating that the overall distribution of simulated streamflow is affected by the type of precipitation data (Table 4-4). According to the NSE indices, the model using mMPE shows higher calibration accuracies on all flows and quick flows, with a slight improvement when recalibrated with NEW parameter constraints (Table 4-5). The statistics on baseflow are insensitive to the changes in calibration within the scope of this study where only the LZSN-parameter constraints are manipulated.

By the measure of Se/Sy, Run 3 outperforms Run 1 at 65% of the 37 calibration stations (Table 4-6). This study also shows that revising parameter constraints for each type of precipitation data makes calibration accuracy increase. At 31 of the 37 calibration sites, Se/Sy decreases from Runs 2 to 3 (Figure 4-6). The statistic increases from Runs 1 to 4 at 20 out of the 37 calibration sites. Based on the spatial analysis, this study shows that both Runs 1 and 3 perform poorly in the western part of the PRB where the difference between mMPE and mCPB is significant. Except for the middle Potomac

River, Run 3 performs slightly better than or comparable to Run 1 across the Basin, by this measure. The various indices for flow variability suggest that the physical characteristics of observed flow are better represented by the model using mMPE, except for extreme flow (Table 4-7 and Figure 4-9).

4.7.3 Implications of using mMPE for the CBW model

When one type of precipitation input is replaced with another, the model needs to be recalibrated with revised parameter constraints. This is especially true in an empirically and theoretically/physically based model like HSPF (Shirmohammadi et al., 2006). Shirmohammadi et al. (2006) pointed out that empirical algorithms are regression equations that are developed based on a set of observed data for defined climatic, soils, and land use scenarios. Thus, their application to conditions other than those for which these algorithms were developed may produce outputs with significant uncertainty. This study demonstrates that even recalibration with revised constraints cannot improve calibration accuracy by measures of, for example, SSTAT and WSTAT. Considering the two types of precipitation data sets show large differences in summer and winter (Figure 3-16), one can infer that the model calibration process cannot compensate for any shortcomings from the characteristics of precipitation inputs.

In this study, NSE and Se/Sy, and indices for flow variability generally suggest that the model using mMPE with NEW parameter constraints (Run 3) performs better than the model using mCBP with CBP parameter constraints (Run 1). The only difference between calibrations using CBP- and NEW parameter constraints is the range of allowed values for a single parameter. Findings in this study indicate that the model

interacts differently with different forcing inputs, requiring a different range of allowed parameter values to match simulated flows to the observed flow during calibration process. It is of interest to examine how the model hydrology, as reflected by the parameter values estimated in calibration, responds to the different types of precipitation inputs: the model hydrology including parameters and fluxes are the topic of the next chapter.

CHAPTER FIVE

Impact of Using Different Precipitation Sources on the CBW Model Hydrology

5.1 Introduction

In Chapter 4, multiple statistic measures demonstrated that when the parameter constraints were adjusted for different forcing inputs, the model performance changed: Run1 (CBP precipitation, CBP constraints) outperformed Run4 (CBP precipitation, NEW constraints); Run3 (MPE precipitation, NEW constraints) outperformed Run2 (MPE precipitation, CBP constraints). Chapter 4 concluded that the CBW model should be recalibrated with revised parameter constraints when mCBP is replaced with mMPE. This conclusion is consistent with Neary et al.'s finding that parameter constraints were reset to avoid unreasonable parameter values when the model interacts differently with different forcing inputs during autocalibration (2004). To set parameter constraints individually for each type of precipitation input, mMPE and mCBP in this study, it is important to understand how model parameters respond differently to those inputs.

In contrast to a transfer function that would generate an outflow time series by operating on a precipitation time series, a hydrologic model such as HSPF attempts to calculate fluxes, explicitly within the watershed. The calculated values recorded in the course of model execution can be examined to evaluate the physical realism of the modeled processes. To better understand the CBW model hydrology, it is important to examine various fluxes that deliver water to the model's streams and rivers.

The goal of this chapter is to examine how hydrologic parameters and simulated fluxes vary when the CBW model is calibrated with different precipitation data sets. This study follows the definition of "parameter" established by the American Meteorological Society (Geer, 1996): *a variable coefficient in some model, formula, or other relationship that can be adjusted to apply the general model to particular cases.* The use of this definition in this study is consistent with its usage in the CBW modeling community.

The analysis in this chapter was based on the model calibration described in Chapter 4. Run 1 is used as the best-calibrated CBW model using mCBP with CBP model constraints. Run 3 is used as the best-recalibrated CBW model using mMPE with NEW model constraints. Runs 2 and 4 are also included in this study for comparison purposes.

The scope of the study was limited to model hydrology processes for the land simulation. The CBP established the model environment for the river simulation, including segmentation, variables, and corresponding coefficients based on information from the years 1985-2005. The autocalibration procedure set by the CBP does not include river parameters and therefore the model settings for the river simulation remain the same throughout this study. The study addressed the following questions:

(1) How does the model hydrology, including parameters and fluxes, change when the model is calibrated with two different precipitation input data sets?

(2) What should model users consider when selecting precipitation inputs for the CBW model?

This chapter begins with further background on hydrologic processes in HSPF and the roles of the CBP calibration parameters, building on the overview given in Chapter 4.

The methods of investigation and comparison are described, followed by results and discussion.

5.2 Background

5.2.1 Modeling hydrological process in HSPF

To understand the role of parameters in hydrologic simulation, this section provides a brief review of modeling hydrological processes in HSPF, as implemented on the CBW model. The HSPF modeling environment employs a modular structure to simulate the watershed and in-stream processes. Pervious land, impervious land, and river reaches are given different conceptual and mathematical treatments by way of the computational modules PERLND, IMPLND, and RCHRES, respectively. The watershed process includes a flow simulation for pervious land using the PWATER section in the PERLND module and for impervious land using the IWATER section in the IMPLND module. The surface and subsurface flow contributions from the pervious and impervious land segments are routed in model stream and river reaches using the HYDR section in the RCHRES module. Because this study deals only with land hydrology and related parameters, only the pervious land hydrologic processes are described in this section. Impervious land use requires a different set of parameters which are not affected by the hydrology calibration in this study. More details on the HSPF modeling process can be found in the User's Manual for Hydrological Simulation Program-FORTRAN (Bicknell et al., 2001).

The PWATER section in the PERLND module is primarily used to generate the surface and subsurface components of runoff from a pervious area (Figure 5-1). HSPF

calculates flows based on simulated storage components: active groundwater, low soil, upper soil, interflow storage, and surface storage. For the land segment as a whole, the water balance is satisfied for any time period,

$$R = P - ET - IG - \Delta S \quad (5\text{-}1)$$

where R is total runoff, P is precipitation, ET is evapotranspiration, IG is loss to deep/inactive groundwater, and ΔS is total change in storage (both soil layers, active groundwater, surface, and interflow storage).

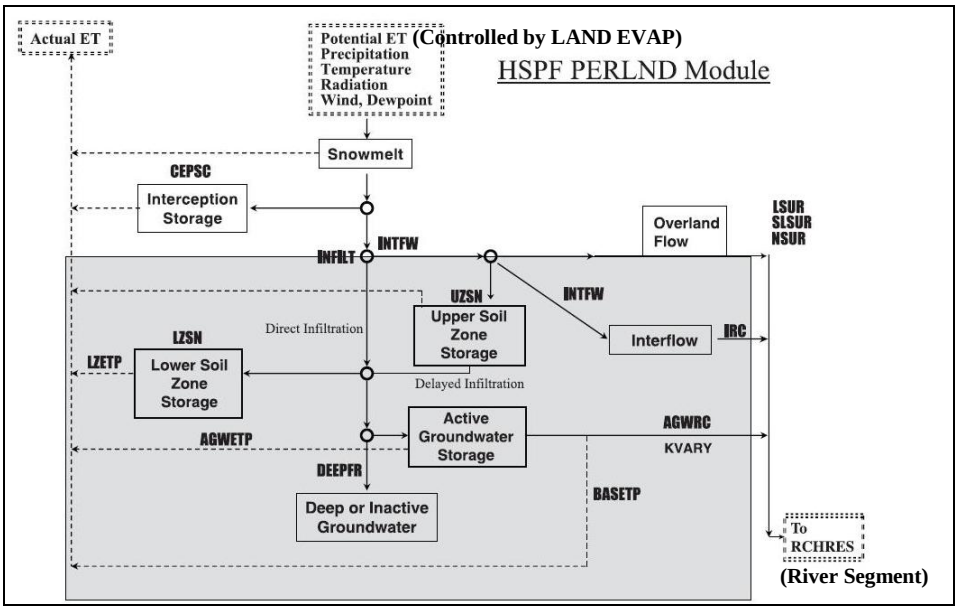


Figure 5-1. Model simulation of hydrologic process for pervious land segment. The shaded area indicates subsurface processes. Adapted and slightly modified for this study from Atkins et al. (2005).

Precipitation is an externally specified forcing, or independent variable. Actual evapotranspiration (ET) is calculated by drawing from a sequence of available storages to attempt to satisfy the externally specified independent variable potential evapotranspiration (PET). The PET is multiplied by the CBP's adjustment parameter, LAND_EVAP. The model equations that allocate precipitation among surface flow,

interflow, groundwater, base flow, interception, evaporation, and their associated parameters, are described below.

Precipitation intercepted by vegetation or other land covers is tracked by the interception storage variable. The interception storage cannot exceed the parameter CEPSC, which dictates interception storage capacity. Water exceeding the interception capacity becomes available for surface runoff and infiltration.

Several subsurface storage variables are updated by mass balance at every time step: upper and lower zone soil water, active groundwater, and inactive groundwater. The parameter INFILT controls the direct infiltration, which also depends on soil saturation as expressed by the ratios of upper and lower zone storage to their respective nominal values (UZSN and LZSN). The distribution of interflow/upper zone storage/overland flow is determined by parameters such as UZSN and INTFW. INTFW fixes the ratio of interflow to surface flow. Arriving water that is neither intercepted nor infiltrated becomes overland flow. The parameters LSUR, SLSUR, and NSUR describe the characteristics of the overland flow plane (average length, average slope, and average roughness, respectively) and control the rate at which water flows overland to streams. Water stored in the upper soil zone is slowly released, as controlled by the parameter IRC, a recession coefficient for interflow. Water can be infiltrated deeper to lower soil zone storage, depending on the saturation state of the lower zone. The capacity parameters UZSN and LZSN are nominal storage values of the surface soil layer and the lower soil layer, respectively; they are shape parameters of nonlinear response curves and do not represent the maximum water holding capacity. LZSN controls seasonal flow: water is deeply infiltrated into the lower zone in the winter and evaporates in the summer

(USEPA, 2010). The parameter DEEPER controls the portion of active groundwater moving into deep or inactive groundwater. Water stored in the inactive groundwater zone is considered 'lost' from the model system. The rate at which active groundwater is released into streams is controlled by the storage of groundwater in this compartment, with the recession parameter AGWR. The contribution to evaporation from active groundwater is controlled by the parameter AGWTP. Details on modeling the hydrologic process using the HSPF model can be found in numerous studies (Bicknell et al., 2001; USEPA, 2010;Shenk, 2010; Atkins et al., 2005) or websites of USEPA BASINS.

The parameter LZSN plays a significant role in the HSPF model hydrology. Its main contribution is to control the amount of infiltrated and percolated water that enters the lower zone according to the following equations (Bicknell et al., 2001):

$$LZfraction = 1.0 - \left(\frac{LZS}{LZSN}\right) \times \left(\frac{1.0}{(1+indx)}\right)^{indx} \qquad if\left(\frac{LZS}{LZSN}\right) < 1 \tag{5-2}$$

$$LZfraction = \left(\frac{1.0}{(1+indx)}\right)^{indx} \qquad if\left(\frac{LZS}{LZSN}\right) > 1 \tag{5-3}$$

where

$$indx = 1.5 \times abs\left(\frac{LZS}{LZSN} - 1.0\right) + 1.0 \tag{5-4}$$

LZfraction = the fraction of infiltration plus percolation plus lower zone lateral inflow that enters lower zone storage.

Percolation is defined as vertical movement of water from the upper soil zone to the lower zone. The percolation rate is also affected as follows by the LZSN parameter values (Bicknell et al., 2001):

$$PERC = 0.1 \times INFILT \times INFFAC \times UZSN \times \left(\frac{UZS}{UZSN} - \frac{LZS}{LZSN} \right)^3 \quad (5\text{-}5)$$

where

PERC = percolation from the upper zone
INFILT = infiltration parameter
INFFAC = factor to account for frozen ground, if any
= *MAX (1.0-FZG*PACKI, FZGL)*

where *FZG* = impact of icing on infiltration/percolation
PACKI = ice in snow pack
FZGL = minimum value of INFFAC

The parameter LZSN also affects the determination of infiltration and interflow. Infiltration is based on equations developed by Phillips (Bicknell et al., 2001). The distribution of the moisture available to the land surface is divided into potential direct runoff, potential interflow inflow, and infiltration. The distribution depends on the locations of two linear distribution functions (Bicknell et al., 2001), intended to represent variation of infiltration capacity over the land segment. The mean, maximum and minimum capacities are calculated to define Line I (Equations 5-6 to 5-8). Line II is calculated by multiplying the ordinates of Line I by RATIO (Equation 5-9). Details on the figures and equations are found in Bicknell et al. (2001).

$$IBAR = INFFAC \times \left(INFILT \times \left(\frac{LZSN}{LZS} \right) \right)^{INFEXP} \quad (5\text{-}6)$$

$$IMAX = INFILD \times IBAR \quad (5\text{-}7)$$

$$IMIN = IBAR - (IMAX - IBAR) \quad (5\text{-}8)$$

$$RATIO = INTFW \times 2.0^{\left(\frac{LZS}{LZSN} \right)} \quad (5\text{-}9)$$

where

INFFAC = a factor to account for frozen ground effects, if applicable.
INFILD = the ratio of IMAX to mean infiltration capacity, default value is 2
INFEXP = exponent parameter, default value is 2
INTFW = interflow inflow parameter

LZSN is also used in the actual evapotranspiration routine to calculate the index for the maximum ET opportunity using Equation 5-10.

$$RPARM = \left(\frac{0.25}{1.0 - LZETP}\right) \times \left(\frac{LZS}{LZSN}\right) \times \frac{DELT60}{24} \qquad (5\text{-}10)$$

where

RPARM = maximum ET opportunity
LZETP = lower zone ET parameter
DELT60 = time conversion factor

Three components of the flow from pervious land segments to streams are calculated: direct surface runoff (SURO), subsurface flow or interflow (IFWO), and groundwater flow (AGWO). The total flux for the pervious areas, PERO, is the sum of SURO, IFWO, and AGWO. The runoff from impervious land has a single component: direct surface runoff (SURO). The components of flow from land segments are generally routed in the model stream network to the outlet of a basin. The aggregated and routed flow can then be compared to the observed stream flow. An important concern for the model user is that, by the time the land outflows reach a basin outlet, the distinctive characteristics of the three components within a given land use or among different land uses may be smeared due to hydrologic connectivity.

5.2.2 Parameter values for the CBW model calibration

In calibrating the CBW model, its developers divided many HSPF parameters into two types. Type I parameters are key parameters for the hydrologic processes: they are updated automatically by the calibration algorithm within the user-defined ranges as constraints for the optimization process. The six Type I parameters are the potential evapotranspiration adjustment parameter (LAND_EVAP), lower zone nominal storage (LZSN), infiltration index (INFILT), interflow inflow index (INTFW), interflow recession index (IRC), and groundwater recession index (AGWR). These six parameters are estimated for each land use for any given land segment.

Another set of parameters (Type II) are all the other parameters which must be established before the automatic calibration procedure, and which remain constant. These parameter values are extracted from the site-specific field survey, GIS operation or literature reports. Type II parameters include all the parameters for the IWATER module, which are used to model hydrologic processes for the impervious land. Table 5-1 lists the Type I and II parameters for the HSPF pervious land hydrologic module in the CBW. Also included are reported values for these parameters from calibration studies of the HSPF-based hydrologic model using gauge- and radar- based precipitation.

Table 5-1. Parameters used to simulate the hydrologic process using PWATER and IWATER modules in HSPF. Reported values are based on studies using precipitation data prepared from gauge measurements and based on studies using NEXRAD-based precipitation.

Parameter(unit)		Description	Reported values	
			Gauge-based precipitation	Radar-based precipitation
Type I	LAND_EVAP	PET adjustment factor	Unique to the Chesapeake Bay program Watershed model	
	INFILT (inch/hr)	Index to the infiltration capacity of the soil	0.01 – 0.25 [a] 0.001 – 0.90 [b] 0.0394 [c]	0.0018 – 0.0042 [d]
	INTFW	Interflow inflow parameter	1.0 - 3.0 [a] 0.05 – 3.4 [b] 0.0197 [c]	
	IRC	Interflow recession parameter	0.50 -0.7 [a] 0.02 – 0.9 [b] 0.0197 [c]	0.586 – 0.85 [d]
	LZSN (inch)	Lower zone nominal storage	3.0 – 8.0 [a] 1.5 – 10.0 [b] 5.0 [c]	2.0 – 3.26 [d]
	AGWR	Basic groundwater recession rate (when KVARY is zero)	0.92 – 0.99 [a] 0.91 – 0.99 [b] 0.99 [c]	0.965 – 0.997 [d]
Type II (pervious area)	NSUR	Manning's N for the assumed overland flow plane	0.15 – 0.35 [a] 0.02 -0.77 [b] 0.106 -0.212 [c]	0.2 [d]
	LSUR (feet)	Length of the assumed overland flow plane	Estimates from GIS	-
	SLSUR	Slope of the assumed overland flow plane	Estimates from GIS	-
	KVARY (1/inch)	Variable groundwater recession parameter	1.0 – 3.0 [a] 0.0 – 4.7 [b] 0.000945 [c]	0 [d]
	AGWETP	Fraction of remaining PET which can be satisfied from active groundwater storage	0.0 – 0.05 [a] 0.0 – 0.0 [b] 0.15 [c]	0.001 – 0.022 [d]
	LZETP	Lower zone ET parameter: an index to the density of deep-rooted vegetation	0.2 – 0.7 [a] 0.1 – 0.9 [b] 0.42 [c]	0.2 [d]
	DEEPER	Fraction of groundwater inflow which will enter deep groundwater	0.0 – 0.2 [a] 0.0 – 0.4 [b] 0.18 [c]	0.017 – 0.138 [d]
	UZSN (inch)	Upper zone nominal storage	0.1 – 1.0 [a] 0.02 – 1.12 [b] 0.70079 [c]	0.81 – 1.972 [d]
	BASETP	Fraction of remaining PET which can be satisfied from baseflow	0.0 – 0.05 [a] 0.0 – 0.005 [b] 1.0 [c]	0.03 – 0.156 [d]
	CEPSC (inch)	Interception storage capacity	0.03 – 0.2 [a] 0.0 0– 0.1 [b] 0.0 [c]	0.0004 – 0.0156 [d]

Parameter(unit)		Description	Reported values	
			Gauge-based precipitation	Radar-based precipitation
Type II (Impervious area)	LSUR (feet)	Length of the assumed overland flow plane	Estimates from GIS	-
	SLSUR	Average slope of the assumed overland flow plane	Estimates from GIS	-
	NSUR	Manning's n	0.05 – 0.10 [a]	-
	RETSC (inch)	Retention storage of the impervious surface	0.1 – 0.5 [a]	-

a. USEPA, 2010 (AOI: N/A)
b. Atkins et al., 2005 (AOI: Mountainous basins, WV)
c. Engelmann et al., 2002 (AOI: Ohio watershed)
d. Ryu, 2009 (AOI: Basins for DMIP2, OK)

5.2.3 Rules for setting initial parameters for land simulation

A model simulation for land is conducted for 24 categories of land use (Table 5-2) within a given county-based model segment. The model requires individual parameter sets for all the land uses. For each pervious land use, nine parameters specify options for modeling hydrologic processes, such as whether a parameter can vary on a monthly basis. Based on these options, from 20 to 70 (including monthly varying values) hydrologic parameters need to be assigned for each land use in each land segment. The CBP established a set of rules for systematically assigning values to this large number of parameters. The rules include: (1) fixed ratios of all other land use parameters to crop land values; (2) fixed ratios of UZSN to LZSN by land use; and (3) monthly factors for UZSN [Tables 5-3 (a) to (c)]. In addition, LAND_EVAP is the same for all land uses within a given land segment. As a result, the number of parameters to specify is greatly reduced. For each model segment for land simulation, only six parameters corresponding to the land use of hightill-without-manure ("hom") are estimated in the autocalibration.

All other land use parameters in a land segment are scaled by the respective factors listed in Table 5-3 (c).

Table 5-2. Land use category for land simulation (USEPA, 2010).

General Land Use	Code	Description
FOREST (FOR)	for	wooded / other
	hvf	harvested forest
CROP (AG)	urs	nursery
	alf	alfalfa
	hyw	hay with nutrients
	hyo	hay without nutrients
	hom	hightill without manure
	hwm	hightill with manure
	lwm	lowtill with manure
	nhy	nutrient management hay
	nal	nutrient management alfalfa
	nho	nutrient management hitill without manure
	nhi	nutrient management hitill with manure
	nlo	nutrient management lotill
	Afo	animal feeding operations (the impervious land)
Impervious URBAN (IMPVURB)	iml	low intensity impervious urban
	imh	high intensity impervious urban
Pervious URBAN (PERVURB)	pul	low intensity pervious urban
	puh	high intensity pervious urban
PASTURE (PAS)	pas	pasture
	trp	degraded riparian pasture
	npa	nutrient management pasture
OTHERS	ext	extractive
	bar	construction

Table 5-3. Rules established for parameter settings by the CBP (USEPA, 2010). (a) ratio of UZSN to LZSN by land use; (b) Monthly factors for UZSN for crop land; (c) ratio of hydrology parameters to their corresponding values for "hom" by pervious land use.

(a)

Land Use	Forest	Crop	Grass, Pasture, Hay	Pervious Urban	Extractive
UZSN: LZSN	0.1-0.12	0.12-0.14	0.07-0.1	0.1	0.08

(b)

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Factor for UZSN	0.6	0.6	0.6	0.6	0.6	0.7	0.95	1.0	1.0	0.8	0.7	0.65

(c)

Land use by code		LZSN, AGWR, IRC	INTFW	INFILT
FORST	for	1.0	1.25	1.6
	hvf	1.0	1.0	1.3
CROP	urs	1.0	1.0	0.8
	alf	1.0	1.0	1.5
	hyw	1.0	1.0	1.5
	hyo	1.0	1.0	1.5
	hom	**1.0**	**1.0**	**1.0**
	hwm	1.0	1.0	1.0
	lwm	1.0	1.0	1.5
	nhy	1.0	1.0	1.5
	nal	1.0	1.0	1.5
	nho	1.0	1.0	1.0
	nhi	1.0	1.0	1.0
	nlo	1.0	1.0	1.5
Pervious URBAN	puh	1.0	1.0	0.8
	pul	1.0	1.0	0.8
PASTURE	pas	1.0	1.0	1.5
	trp	1.0	1.0	0.8
	npa	1.0	1.0	1.5
OTHERS	ext	1.0	0.7	0.8
	bar	1.0	0.5	0.7

For forests, for example, UZSN is set at 10 to 12 percent of the LZSN [Table 5-3 (a)]. UZSN takes its maximum value in August and September, and declines to 60 percent of that value in Jan. – May [Table 5-3 (b)]. Forest land is assumed to have 1.6 times the infiltration opportunity that hightill-without-manure land has [Table 5-3 (c)].

5.3 Study Site

This chapter examined the land parameters and fluxes for the model used in Chapter 4. The calibration accuracy was evaluated at the 37 calibration stations in the Potomac River Basin (PRB). As discussed in Section 2.4, the CBP watershed model

subdivides land in two ways: the county-based land segments, and the watershed-based river segments. For each calibration site, the model programmatically identifies the upstream river segments contributing to that site. The model then creates a list of land segments that need to be simulated to generate runoff to streams from each upstream river segment. Fifty-seven land segments are involved in the calibration of the 37 sites. This study focused on parameter changes in these land segments (Figure 5-2).

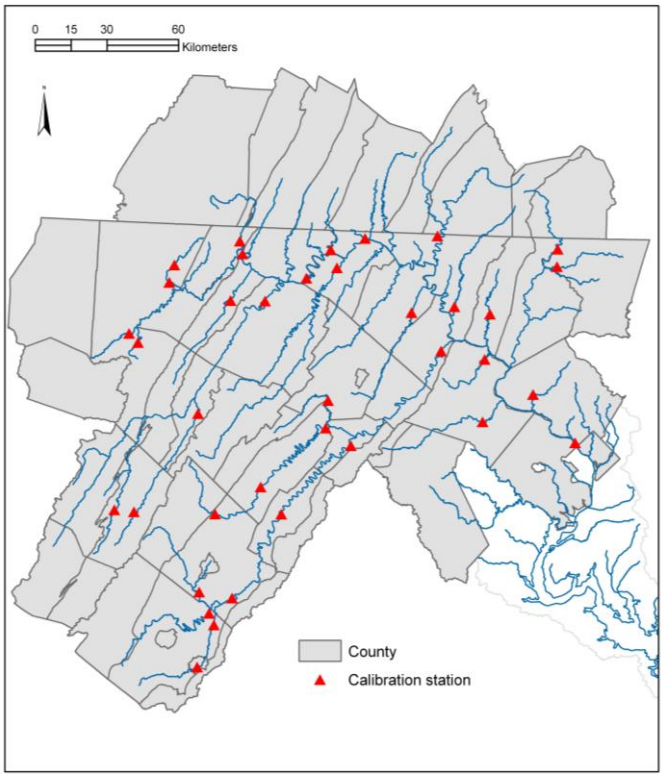


Figure 5-2. Land segments involved in the autocalibration procedure in Chapter 4.

5.4 Methods

All four Runs introduced in Chapter 4 were used to examine calibrated parameters. Fluxes were then examined using Run 1(the best-calibrated model using mCBP, as shown in Chapter 4) and Run 3 (the best-calibrated model using mMPE). This

study also estimated the errors that occur when the model is used to simulate streamflow using a different type of precipitation without the recalibration.

5.4.1 Trial scenarios for Run3

As demonstrated in Chapter 4, a recalibration of the model with the revised parameter constraints allowed for improved calibration statistics. The parameter constraints for the model using mCBP were set by the CBP. This study explored various parameter settings for the model using mMPE. This section explains the different scenarios explored by this study, aiming at finding the optimum setting for the model by measures of various calibration statistics.

This study experimented with different scenarios to find whether the model system can respond better when mCBP is replaced with mMPE. The strategies for new scenarios were divided into three categories: (1) changing the Type I parameters; (2) changing the Type II parameters; and (3) changing the rules used for setting parameters as shown in Table 5-3. For the Type I parameters (Table 5-1), the ranges of permissible values during the iterative calibration process (parameter constraints) were adjusted based on the characteristic differences between mCBP and mMPE. For example, the minimum allowed value of LZSN was lowered for the calibration using mMPE, given that average annual mMPE is less than that of mCBP. For the Type II parameters (Table 5-1), the sensitivities of parameters such as KVARY, DEEPER, and CEPSC were investigated. With all the parameter constraints fixed, a target Type II parameter was assigned a different value as each run was conducted. For scenarios conducted by modifying the rule shown in Table 5-3(c), the study modified the ratio of UZSN to LZSN by land use,

the monthly factors for UZSN for crop land, and/or the ratio of the hydrology parameters for “hom” to their corresponding values by land use. After experimenting with numerous scenarios, all the changes made for each scenario were set back to the original values which were set by the CBP, if the results showed insignificant improvements or inconsistent results with the sample variation among the 37 calibration sites. In this way, confounding effects could be avoided as much as possible.

5.4.2 Parameter comparison

When a model is forced with different precipitation data, the calibration procedure must adjust model parameters in order to match simulated outputs to observations as closely as possible according to the defined objective functions. Each calibration run results in a different parameter set for each land segment. This study investigated how the parameters changed to make the model match the same observed streamflow with different precipitation inputs. Parameters as calibrated in the different runs were compared graphically.

5.4.3 Spatial analysis of land parameters and model accuracies

External forcing variables (including precipitation and other meteorology) were supplied and land parameters were calibrated at individual land segments, whereas model accuracies were calculated at calibration sites. Since land segment parameters did not correspond directly to each calibration site in the river, results were spatially displayed using GIS tools. Land segments are shown as polygons and calibrations sites as points in the maps.

5.4.4 Comparison tools

This chapter uses the relative seasonal difference between mCBP and mMPE during the study period (2001-2005) calculated using Equation 3-5. The relative difference for winter precipitation was calculated based on the precipitation values from December, January, and February. The relative difference for summer precipitation was calculated based on the precipitation values from June, July, and August.

The relative bias in seasonal streamflow, WBIAS for winter and SBIAS for summer, were calculated using Equations 4-1. Notice that this study uses "relative difference" for the comparison of the two precipitation data sets and "relative bias" for the comparison of simulated stream flows with the observed flows, as mentioned in Section 4.2.2.

To understand differences between the two model versions (Runs 1 and 3), the components of surface and subsurface flow were investigated. The model output includes the outflow to streams for each land use from any given land segment. For this analysis, the CBW model's 24 land uses were grouped into Crop, Pasture, Forest, Pervious Urban, Impervious Urban, and Other. To determine the contribution of each land use group to the outflow for each flow component at a given land segment, a weighted average of runoff depth (flux) was calculated

$$Avg\ fx\ (LUg) = \sum_{i=1}^{n} fx(i)\left(\frac{area(i)}{total\ area(LUg)}\right) \tag{5-11}$$

where

fx = flux; SURO, IFWO, AGWO [in]
i = index of landuse
n= number of land uses included in general land use category:
13 land uses for CROP, 3 for PASTURE, 2 for FOREST,
2 for pURBAN, 2 for iURBAN, and 2 for OTHER

LUg = general land use code;
Crop, pasture, forest, pervious urban, impervious urban, other

5.4.5 Cross-validation

The CBP calibrates the CBW model over the Chesapeake Bay region based on information from the years 1985-2005 and makes the model available to the public. Model users may wish to simulate a different time period with a different type of precipitation data from that used in calibration. In Chapter 4, Run 1 was the best-calibrated model using mCBP and Run 3 was the best-calibrated model using mMPE during 2002-2004. These models were validated for 2005. In this part of the study, cross-validation was conducted by simulating Run 1 using mMPE and Run 3 using mCBP to evaluate possible errors occurring when the model was used to simulate streamflow using a different type of precipitation than the one used for calibration. The model accuracy for cross-validation was reported in terms of Se/Sy (Equation 4-12).

5.5 Results

5.5.1 Experiments with parameter rules, fixed parameter values, and constraints

As explained in Section 5.4.1, all of the scenarios explored to find the optimum parameter sets for the model using mMPE were constructed in one of three ways: (1) by changing the rules used for setting parameters as seen in Table 5-3; (2) by changing the Type II parameters; or (3) by changing the Type I parameters. Changing the rules (Table 5-3) did not show any significant contributions to the improvement of model calibration accuracy. Therefore, all rules set by the CBP were maintained throughout the study. Changes in Type II parameters did not noticeably affect calibration results either. For

instance, this study experimented with KVARY which affects the ground water recession, one of the fixed parameters during each calibration run. While Atkins et al. (2005) assert that KVARY is a sensitive parameter, a study guide for BASINS (USEPA, 2010) says that it is one of the last parameters to adjust. All parameters categorized as Type II, including KVARY in Table 5-1 also remained the same as set by the CBP throughout the study.

Adjusting the allowed ranges for the Type I parameters reveals that changing the upper and lower boundaries of the six parameters (Table 5-2) for the iterative calibration process does not seem to affect model calibration accuracy. The exception to this is LZSN. Even for LZSN, the user-defined boundaries have to be changed with caution because UZSN also changes in accordance with the rule [Table 5-3(a)] as LZSN changes. For example, when the lower boundary of LZSN is set too low, UZSN is also lowered based on the ratio shown in Table 5-3 (a). Then, an internal checking program in the CBW model system reviews the range of parameter values for computation stability and generates an error message before crashing. In this study, the LZSN parameter constraints were set so as not to violate these internal checks through trial and error.

After examining the results of numerous runs, this study accepted the parameter constraints for the calibration of the model using mMPE in Chapter 4: NEW. Except for LZSN, all other parameter constraints are the same as used in Run1 (Table 5-4). The calibration accuracies for these four Runs were discussed in Chapter 4.

Table 5-4. Parameter constraints of six parameters calibrated during the autocalibration procedure. Parameters constraints in Run 1 are set by the CBP

Precipitation	mCBP	mMPE	mMPE	mCBP
Scenario setting	Run 1	Run 2	Run 3	Run 4
Name for Parameter constraints	CBP (in Chapter4)		NEW (in Chapter4)	
AGWR	0.92 – 0.995		0.92 – 0.995	
INFILT	0.0125 – 0.25		0.0125 – 0.25	
INTFW	0.3 – 5.0		0.3 – 5.0	
IRC	0.3 – 0.85		0.3 – 0.85	
LAND_EVAP	0.75 – 1.25		0.75 – 1.25	
LZSN	**8.0 – 12.0**		**2.0 – 12.0**	

5.5.2 Comparison of calibrated parameters

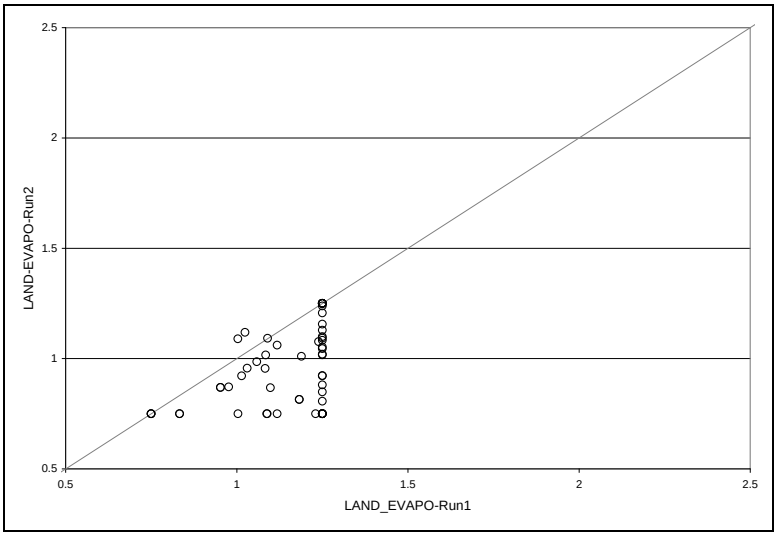
The parameters as estimated by the four calibration runs are compared graphically (Figures 5-3 through 5-9); in all figures, the results of calibration using CBP constraints are on the left panel, NEW constraints on the right.

There are 57 values of LAND_EVAP, LZSN, AGWR, and IRC, one value per land segment for each parameter. There are 4 different INTFW values depending on the land use [Table 5-3(c)] in any given land segment. The parameter INFILT takes 6 different values per land segment, depending on the land use at a given land segment. As noted in Section 5.2.3, however, only the "hom" land use parameters are adjusted in calibration, and these values are scaled to the other land uses by defined rules. In other words, each land segment has six parameters estimated in calibration, estimating 342 parameters in total. The figures may appear to show fewer than the total number of parameters estimated, because values overlap for many land segments.

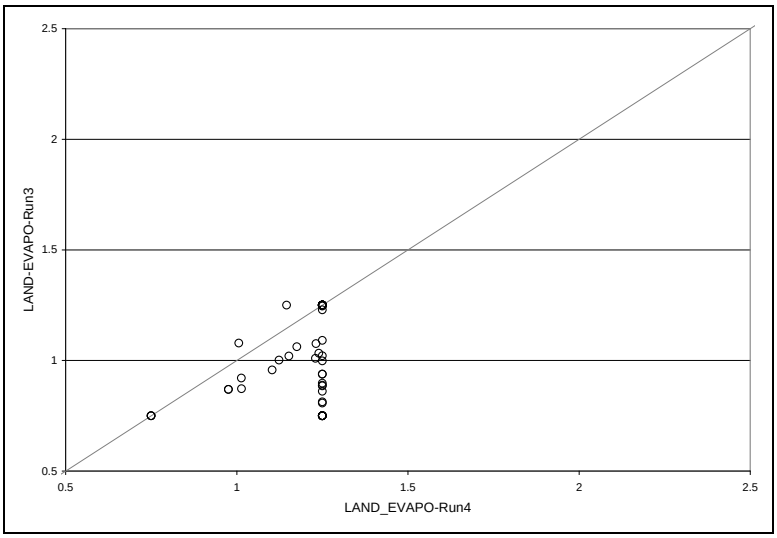
For LAND_EVAP under CBP [Figure 5-3(a)], 30 land segments in Run 1 converge to the maximum values of 1.25, while 16 land segments in Run 2 converge to the minimum values of 0.75. Using mMPE, the model tends to have a low LAND_EVAP at many land segments (Runs 2 on the y-axis). This pattern is the same with NEW [Figure 5-3(b)]: many land segments tend to have minimum values of 0.75 in

Run 3, whereas 41 out of the 57 land segments reach the upper boundary of 1.25 in Run 4. This indicates that more land segments reach the maximum values of LAND_EVAP when using mCBP rather than mMPE, enhancing the climatically-calculated potential evapotranspiration, and allowing more water to evaporate. The trend remains the same when using either CBP or NEW. This is consistent with the fact that the total volume of mCBP is larger than mMPE for the majority of counties in the PRB region, as shown in Chapter 4. The model that receives greater precipitation must evaporate more water in order to match observed streamflow and satisfy overall water balance.

Figure 5-3. Comparison of LAND_EVAP in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.



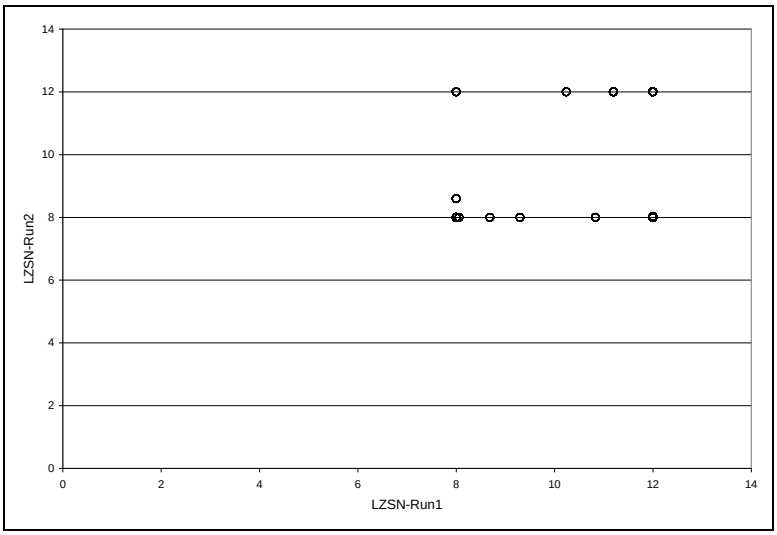
(a)



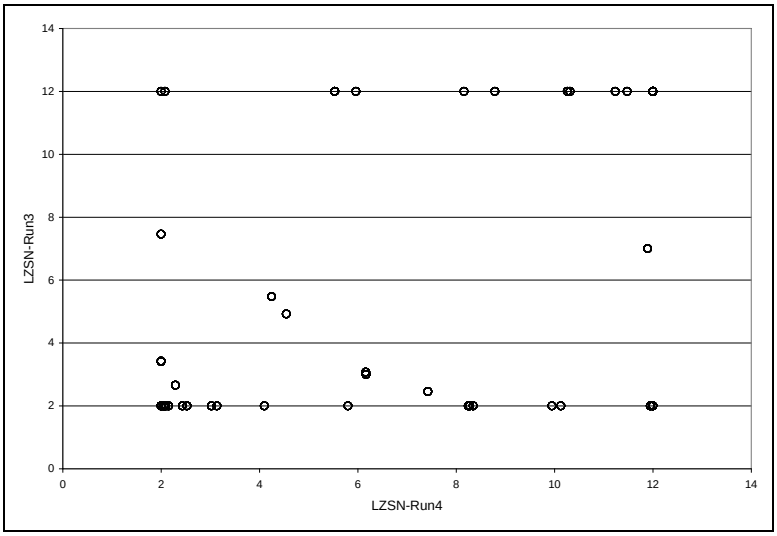
(b)

For the lower soil zone nominal storage (LZSN), both panels in Figure 5-4 show that most of the values tend to stay at the boundaries. Using CBP constraints [Figure 5-4(a)], 50% of the 57 land segments in Run 1 take the maximum value (12), whereas 56% of the land segments in Run 2 take the minimum value (8). Using mCBP, land segments tend to have a higher value of LZSN than when using mMPE. As mentioned earlier, the figure appears to show fewer than the total number of parameters estimated, because values overlap for many land segments. Using NEW constraints [Figure 5-4(b)], 40% of the land segments in Run 3 take the minimum value of 2, while in Run 4 only 15% of the land segments reach the lower boundary.

Figure 5-4. Comparison of LZSN in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.



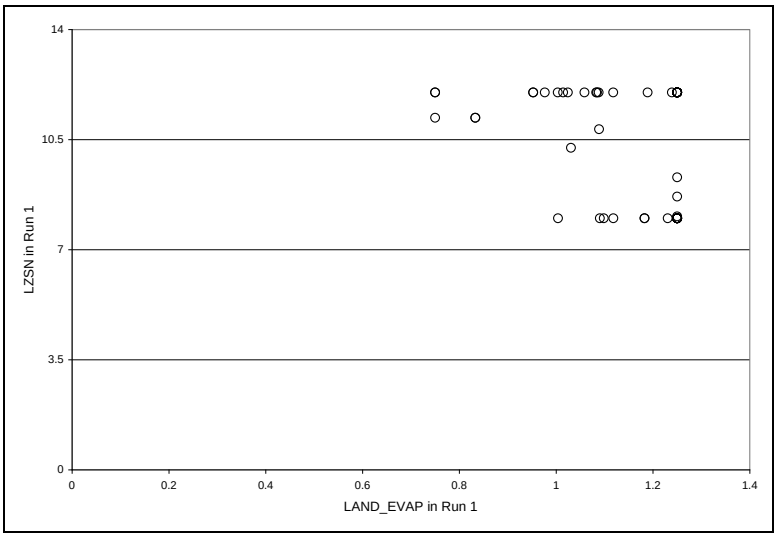
(a)



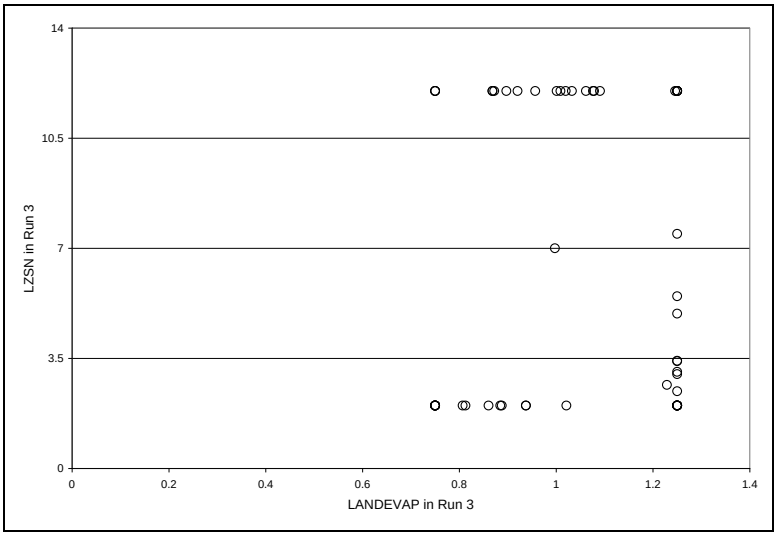
(b)

Both the CBP and NEW constraint results show that using mMPE results in a lower LZSN. A lower zone nominal storage reduces the amount of water that can be stored in the lower zone. This pattern is expected, because Runs 2 and 3 receive less precipitation than Runs 1 and 4 do. Values found on the boundaries indicate that parameters would continue to increase or decrease if the constraints are relaxed. When the constraints are relaxed, however, this study finds instances where the model run crashes because of computational instabilities and irrational parameter values.

Since both LAND_EVAP and LZSN are being adjusted to compensate for larger or smaller volume of precipitation, these two parameters may be correlated at a given land segment. However, the two parameters are not correlated (Figure 5-5). It is because LZSN is adjusted based on biases in seasonal stream flows, while LAND_EVAP is adjusted for the overall water balance: Chapter 4 shows that TBIAS (total bias) and seasonal biases, SBIAS and WBIAS, are very different within each Run.



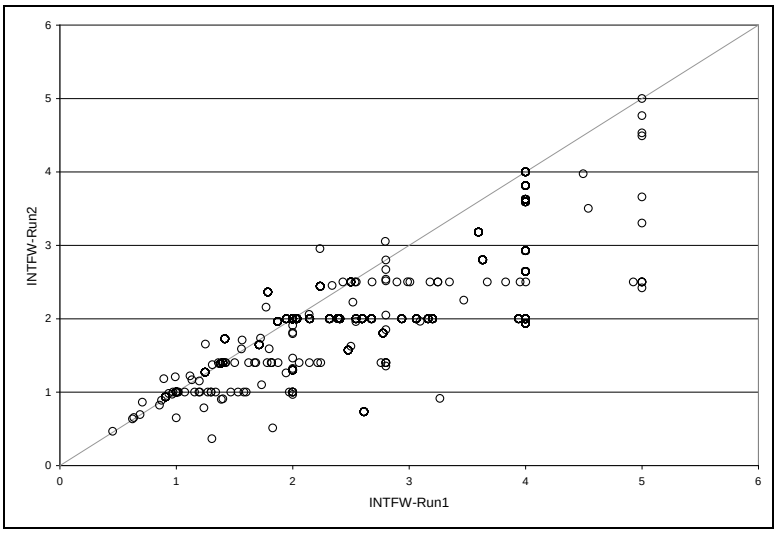
(a)



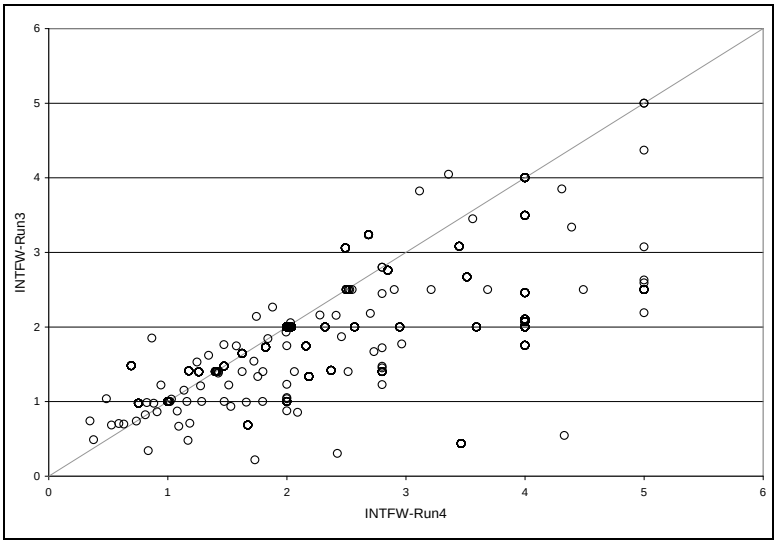
(b)

Figure 5-5. Relationship between LAND_EVAP and LZSN at each land segment using (a) Run 1; (b) Run 3.

In general, Runs 1 and 4 show higher values of INTFW and INFILT than Runs 2 and 3 (Figures 5-6 and 5-7). The model using mCBP tends to have a higher INTFW, partitioning more water to upper zone storage and interflow, and reducing direct overland flow to streams. The precipitation also infiltrates the ground at a higher rate (INFILT) when using mCBP.

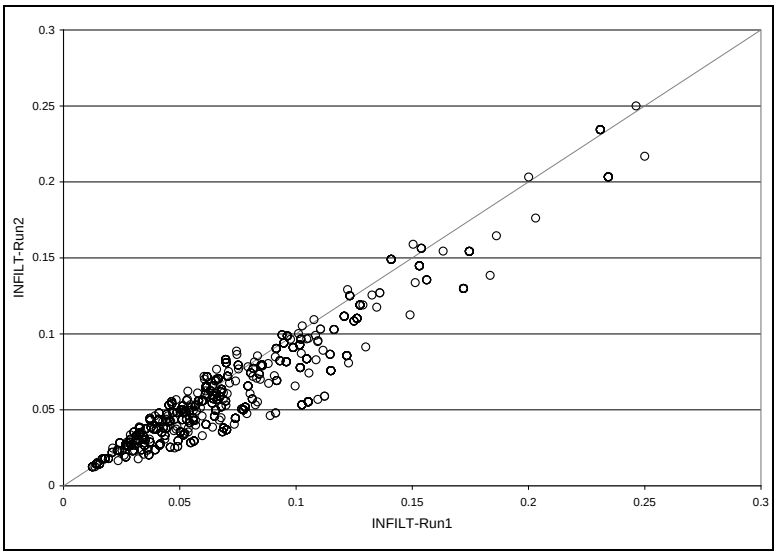


(a)

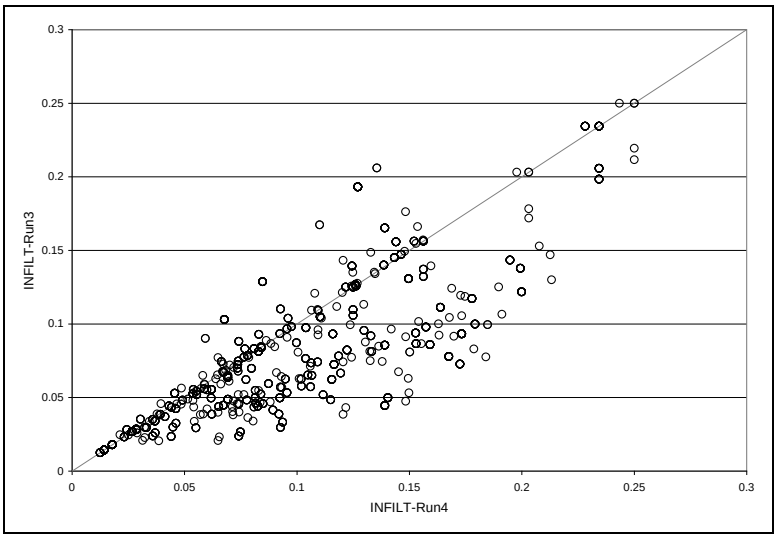


(b)

Figure 5-6. Comparison of INTFW in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.



(a)

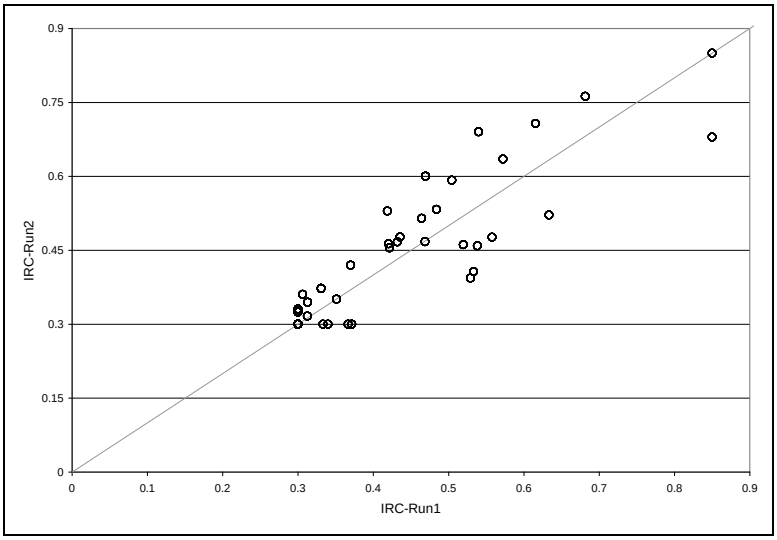


(b)

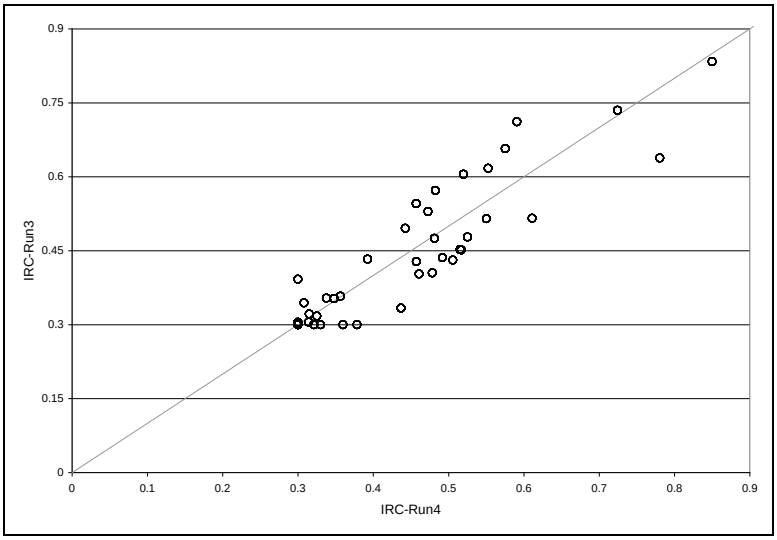
Figure 5-7. Comparison of INFILT in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.

The recession rates of interflow (IRC) and ground water (AGWR) in Runs 1 and 4 do not seem to be significantly different from those in Runs 2 and 3, respectively, based on visual inspection (Figures 5-8 and 5-9). This indicates that these parameters are not greatly affected by either different types of precipitation or different parameter

constraints. This finding is consistent with the analysis of calibration accuracy in Chapter 4.

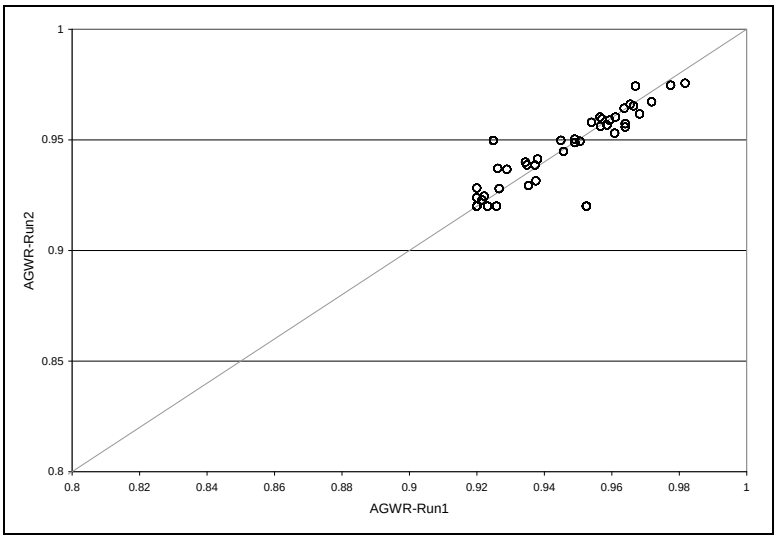


(a)

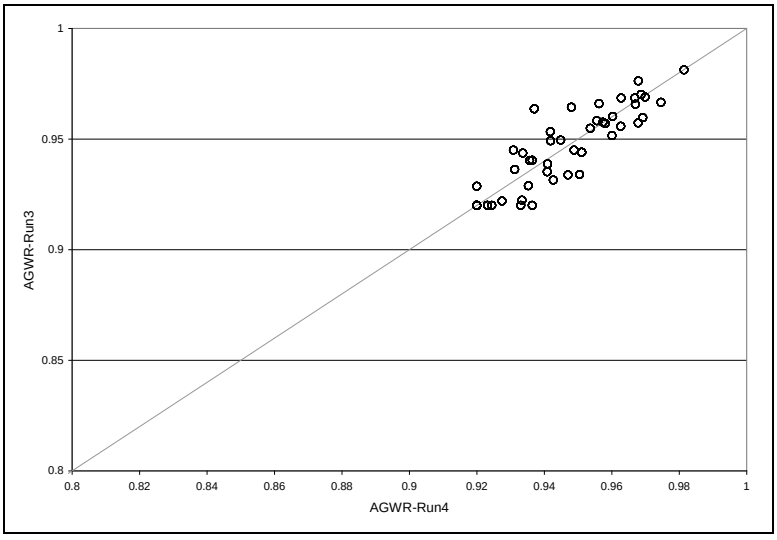


(b)

Figure 5-8. Comparison of IRC in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.



(a)



(b)

Figure 5-9. Comparison of AGWR in Run 1or Run 4 using mCBP (X-axis) and Run 2 or Run 3 using mMPE (Y-axis). (a) CBP constraints on LZSN; (b) NEW constraints on LZSN.

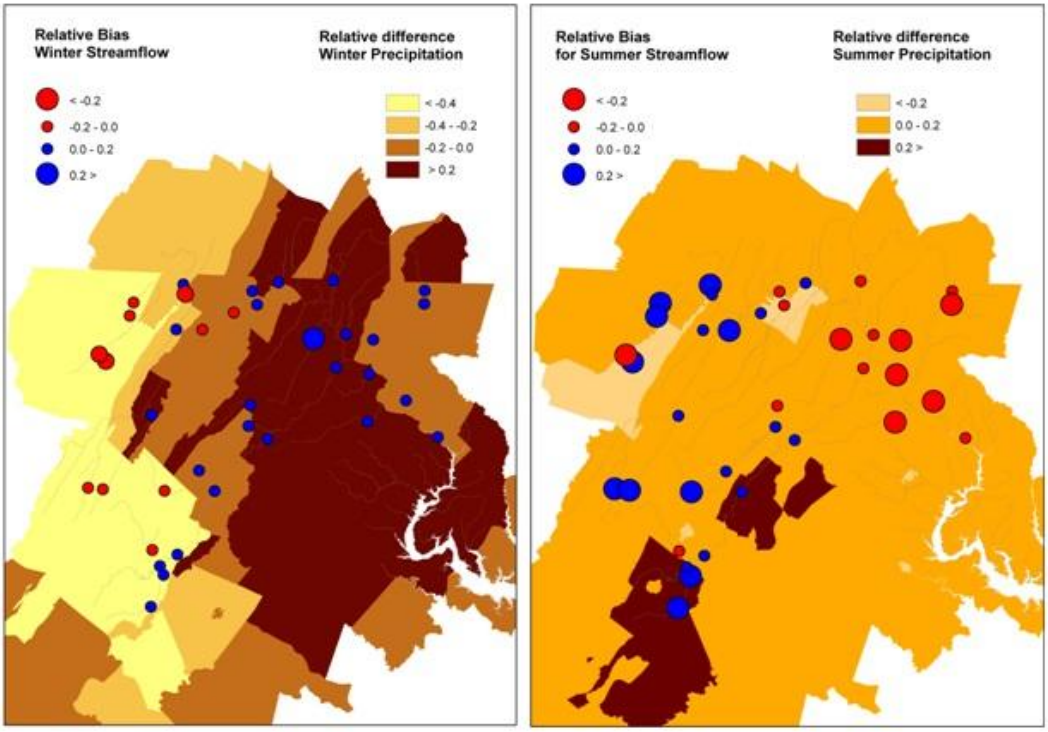
Overall, LZSN and LAND_EVAP are the parameters that tend to converge to the boundaries of the permissible ranges during the calibration process. Other parameters show more continuous distributions of values within the defined ranges of values under either CBP or NEW parameter constraints.

5.5.3 Spatial analysis of land parameters and model accuracies

The calibration of LZSN is controlled by biases of summer and winter flows (Table 4-1): LZSN is adjusted in calibration to equalize these biases. The relative biases in seasonal streamflow were calculated to evaluate the distinctive behavior of LZSN during autocalibration. The relative difference between mMPE and mCBP is relevant, given that the type of precipitation is the only different input among Runs. Since the biases for streamflow calculated at calibration sites in the river do not correspond to the relative difference between mMPE and mCBP at individual land segments, spatial analysis was used in this study. This provides a general idea of how the model simulates streamflow, producing different land simulations in response to the two different forcing inputs.

In Run 3, the relative bias between simulated and observed flow during the winter is consistent with the volume of precipitation [Figure 5-10 (a)]. In other words, in the regions where mMPE underestimates mCBP, the simulated flow underestimates the observed flow, and vice versa. During the summer, however, the relative bias between simulated and observed flow has little relation to the difference between the two forcing inputs [Figure 5-10 (b)]. On the eastern side of the basin, the simulated streamflow using Run 3 underestimates the observed flow, which is to be expected based on the relative difference between the two precipitation inputs. On the western side of the basin, Run 3 oversimulates the observed summer flow, although the mMPE forcing input provides less water than mCBP does. This indicates that mMPE underestimates mCBP, but still provides the model with enough water to simulate a similar summer flow to the observed summer flow. This is confirmed by showing that Run 1using mCBP also oversimulates

the observed summer flow in the western side of the basin [Figure 5-11 (b)]. It is also noticed that the spatial pattern of the relative biases in seasonal streamflow using Run 3 (Figure 5-10) is more obvious than the pattern obtained using Run 1for both seasons, winter and summer.

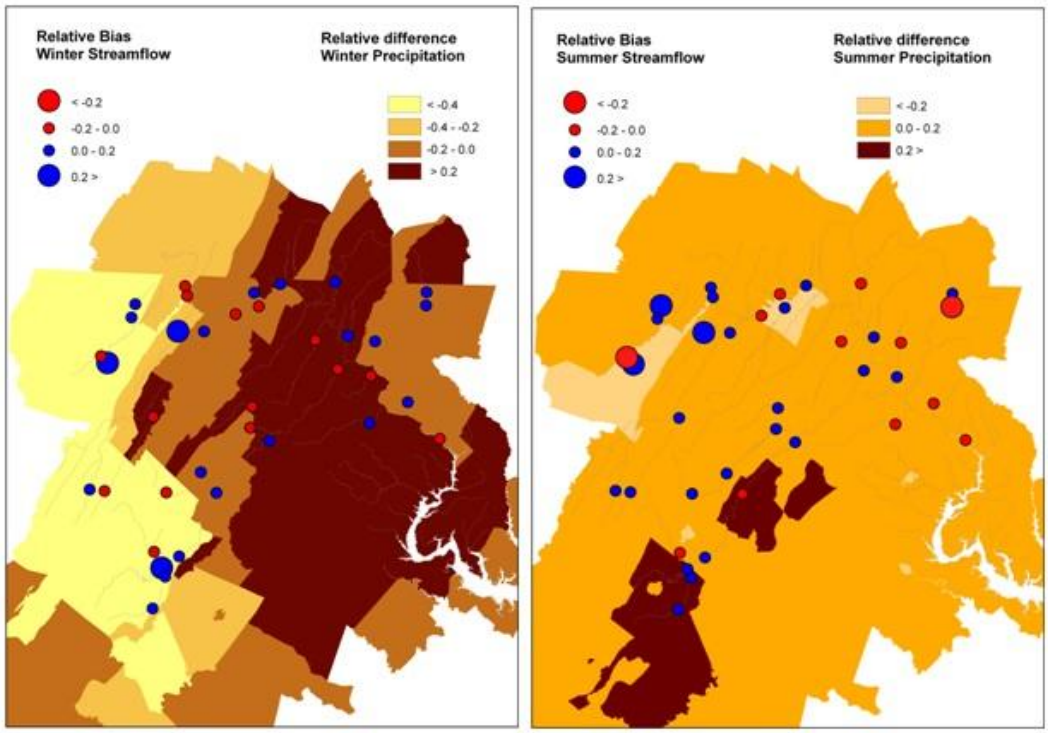


(a) (b)

Figure 5-10. Relative difference between mMPE and mCBP shown with the relative bias in streamflow during (a) winter; and (b) summer in Run 3.

The meteorological data sets (Chapter 2) were investigated to further identify causes of the summer streamflow bias in Runs (Figure 5-12). It is of interest that the meteorological effect by regions on the seasonal flow biases using Run 1 [Figure 5-12 (a) and (b)] does not appear as clear as the effect observed using Run 3 [Figure 5-12 (c) and (d)], even though the regional meteorological data were created and the mCBP

precipitation data sets prepared for the same purpose (Chapter 2): to feed the CBW model.

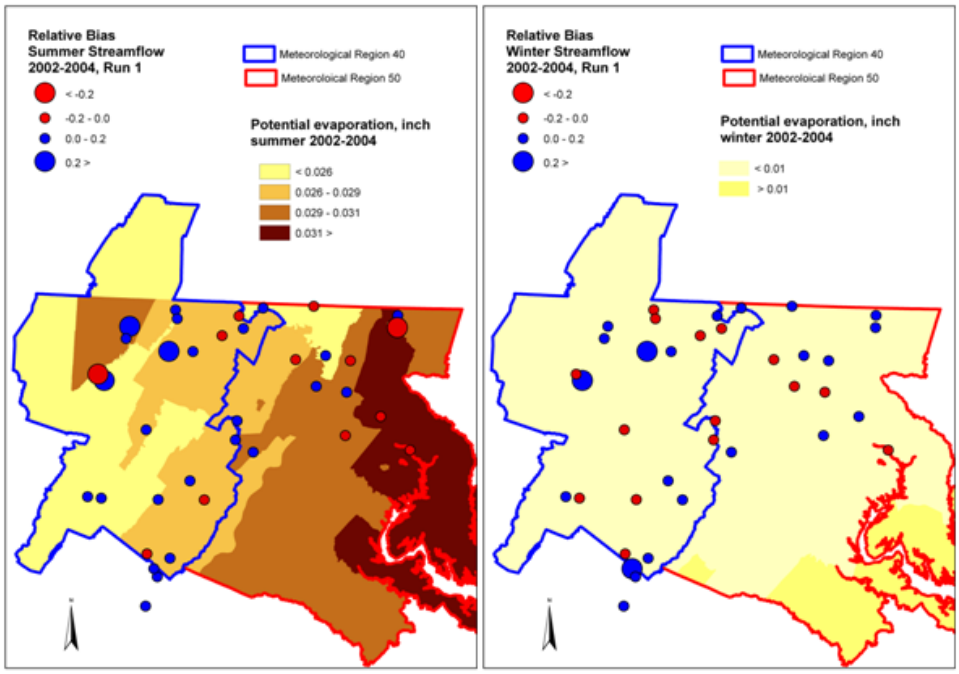


(a) (b)

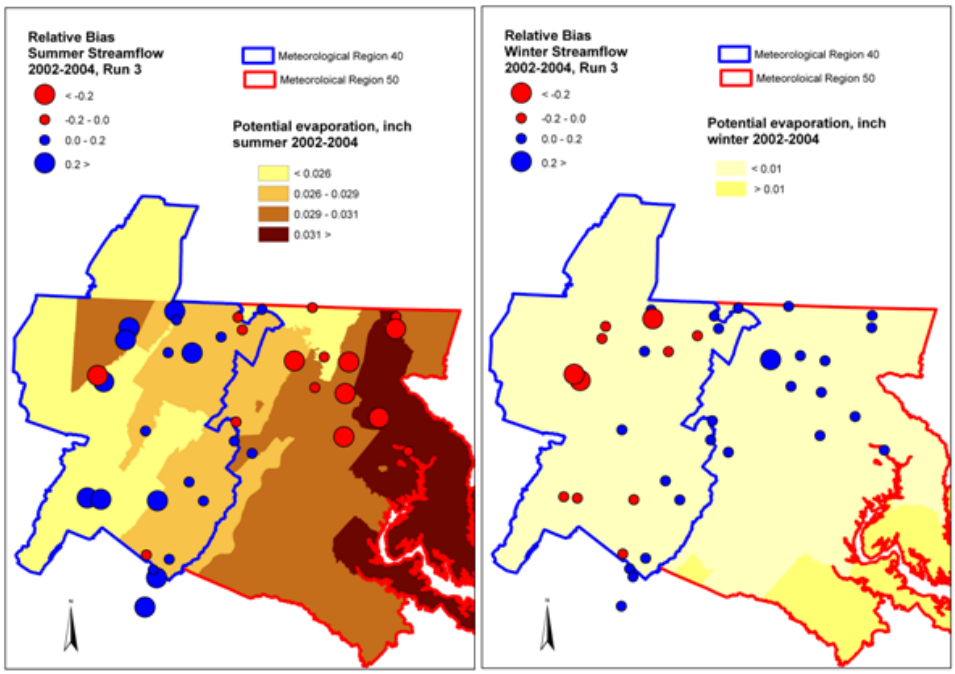
Figure 5-11. Relative difference between mMPE and mCBP shown with the relative bias in streamflow during (a) winter, and (b) summer in Run 1.

During the summer, the PET values are smaller in Meteorological Region 40 than in Region 50. This PET is adjusted in the model by the land parameter, LAND_EVAP (Chapter 4). However, the spatial pattern of the parameter LAND_EVAP does not seem to follow meteorological regions (Figure 5-13). This is expected, because LAND_EVAP is adjusted to control annual water balance, rather than seasonal water balance: this parameter is adjusted by the updating factor which is a function of TBIAS (Table 4-1). The pattern of the relative bias in summer streamflow using Run 3 follows the meteorological regions [Figure5-12(c)]. During the winter, there are no distinctive

patterns found in the spatial distribution of PET across the meteorological regions 40 and 50 for either of the Runs [Figures 5-12 (b) and (d)].

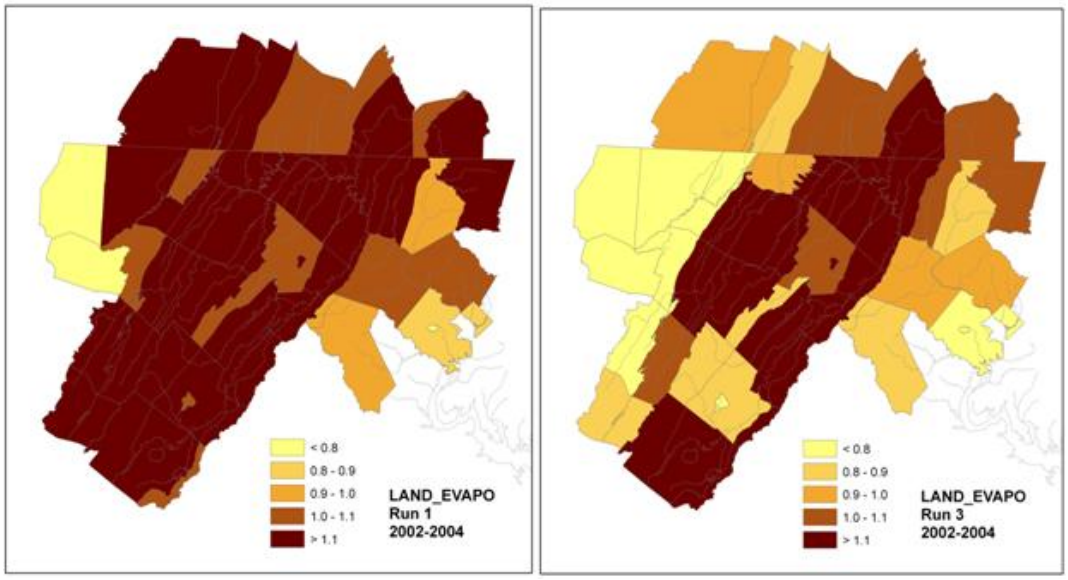


(a) (b)



(c) (d)

Figure 5-12. Spatial distribution of PET for the meteorological regions 40 (blue polygon) and 50 (red polygon). The relative bias in streamflow is also shown during (a) summer; (b) winter in Run 1; and (c) summer; (d) winter in Run 3.



(a) (b)

Figure 5-13. Spatial distribution of LAND_EVAP for (a) Run 1; (b) Run 3.

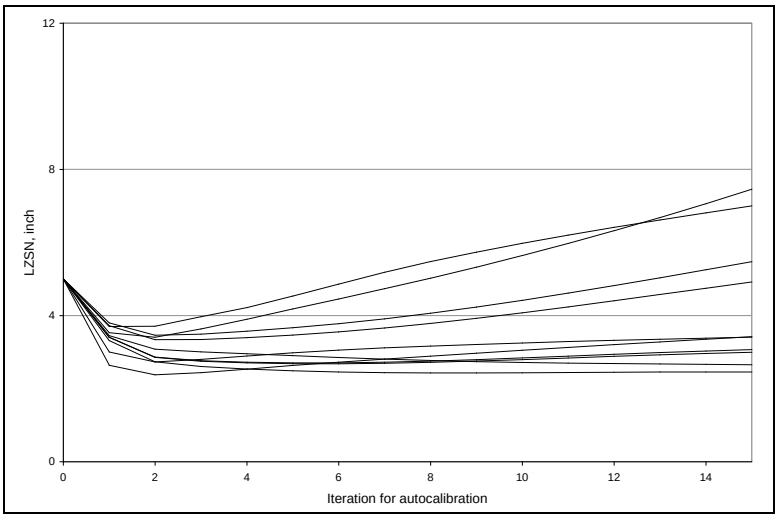
Another meteorological factor affecting the hydrological process is wind, when the model includes a snow simulation. The wind speed changes by meteorological region, not by individual land segments (Table 5-5). During both winter and summer, wind is weaker in Meteorological Region 40 than in 50. According to the model manual (USEPA, 2010), more melted snow seems to be available for streamflow than evaporation in Region 40. Wind shows greater discrepancy between meteorological regions in the summer than in the winter. Not only does the seasonal pattern of mMPE make the model simulate seasonal stream flow differently, but also meteorological data. The same meteorological data is applied to all of the calibration runs. Therefore, their effect should not be a dominant contributor to the differences among Runs. However, it does help to explain the biases between seasonal flows within a given Run.

Table 5-5. Statistics of seasonal wind time series [mile/hour] by meteorological regions during 2002-2004.

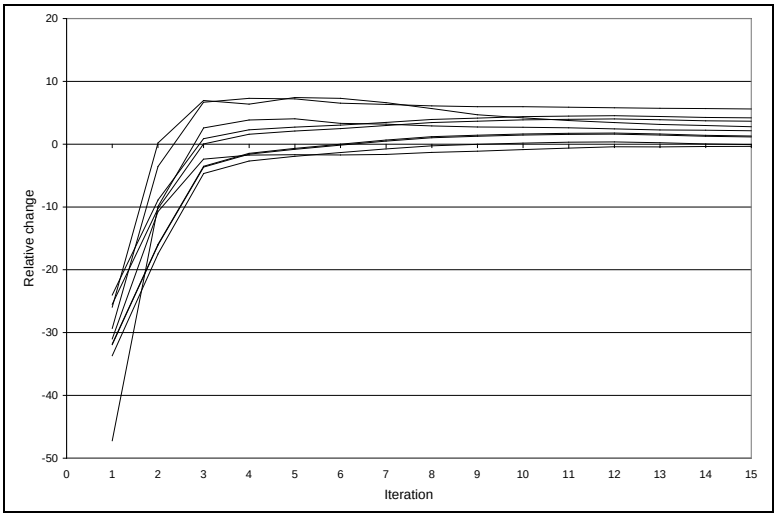
Season	Winter		Summer	
Region	40	50	40	50
Mean	5.618	7.551	2.980	5.457
Median	5.197	6.940	2.476	5.016
75th	8.009	10.113	3.749	6.689
Max	21.940	27.490	14.200	23.450

To investigate how the model adjusts LZSN based on these summer and winter flow biases during the autocalibration procedure, the trajectory of LZSN change in Run 3 during the iterative calibration process was examined (Figure 5-14). Out of the 57 land segments, 10 segments' LZSN values gradually converge to certain values [Figure 5-14 (a)] within 5% of the final values during the iteration [Figures 5-14 (b)]. According to the EPA document (2010), converging to a certain value within 5% is acceptable. At the majority of the land segments, the LZSN values are quickly adjusted to the extreme allowed values, 2 or 12. Then, the value remains the same for the rest of the iterations [Figure 5-14 (c)]. This suggests that large discrepancies in summer and winter flow biases cause the LZSN values to be constrained by user-defined lower and upper boundaries. Except for the ten land segments, the change in LZSN stops after the third or fourth iterations. This is consistent with the change in calibration accuracy by measure of the seasonal flow bias, which showed no change after the third or fourth iterations for the majority of the land segments [Figure 4-3 (b)].

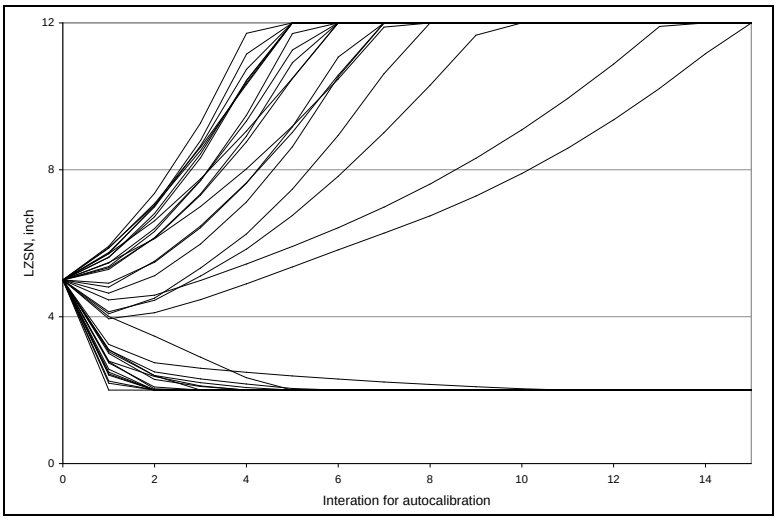
Figure 5-14. LZSN change during the autocalibration process using Run 3. The initial value is 5 at all land segments. (a)10 land segments show the final LZSN values between 2 and 12; (b) relative % change in LZSN values in each iteration for land segments mentioned in (a); (c) 24 land segments show a final value of 2, and 23 land segments have final values of 12.


LZSN, inch
Iteration for autocalibration

(a)

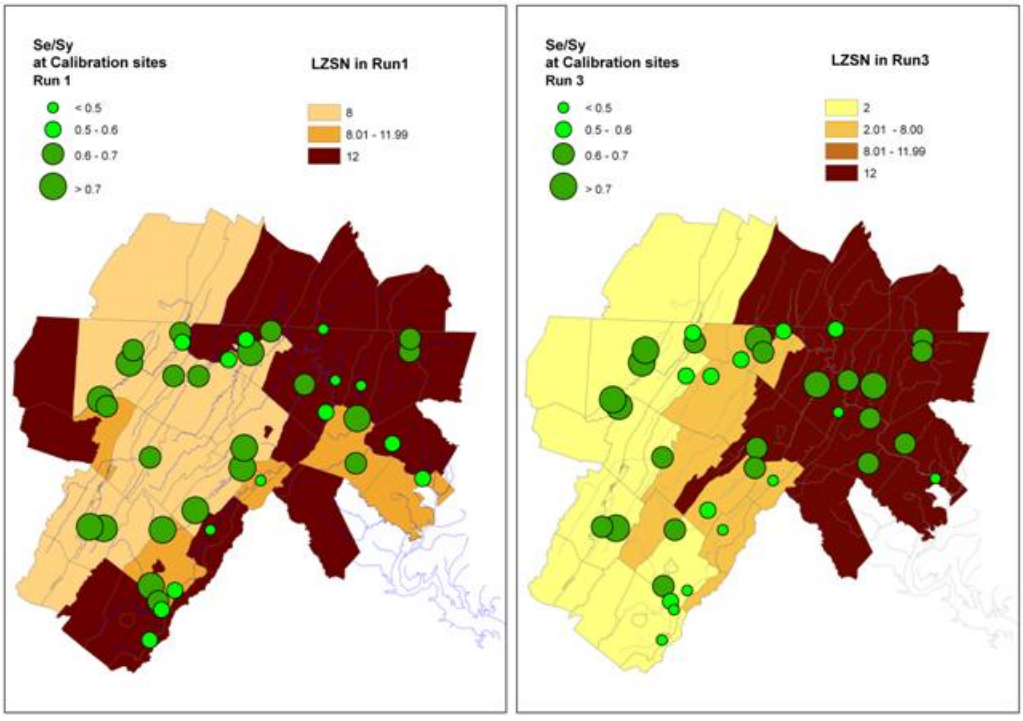

Relative change
Iteration

(b)


LZSN, inch
Interation for autocalibration

(c)

It was suspected that calibration accuracy would be poor when the LZSN values were constrained by user-defined lower and upper boundaries and the iteration process could not improve calibration accuracy. The spatial distribution of LZSN, along with the model calibration accuracy at each calibration site, supports this proposition (Figure 5-15). The model accuracy of Run1 shows little relation to the LZSN values, but the spatial pattern of LZSN of Run 3 is more distinctive than that of Run 1. In Run 3, the LZSN values increase from west to east of the PRB [Figure 5-15(b)]. The model accuracies are generally poor in the region where the LZSN values are the extremes, 2 or 12. The model accuracy is moderate where the values of LZSN lie between 2 and 12.



(a) (b)

Figure 5-15. Model calibration accuracy in Se/Sy calculated in Chapter 4 shown with the LZSN land parameters in (a) Run 1; (b) Run 3.

The spatial pattern of LZSN in Run 3 was further investigated by utilizing the calibration statistics calculated in Chapter 4. In the course of autocalibration, if the

model undersimulates the observed flow, then the LZSN should be lowered to produce more streamflow. For the same reason, if the model oversimulates the observed flow, the LZSN value should be increased for the overall water balance. Based on the spatial distribution of LZSN however, this conceptual relationship does not hold true across the PRB (Figure 5-16). This occurs because LZSN is updated based on the discrepancy between summer and winter seasonal flow biases (Table 4-1), not responding to the total bias. The updating factor, φ_2 is expressed as a function of TBIAS, but TBIAS appears in the numerator and denominator and is cancelled out. Therefore, TBIAS does not affect the calculation of φ_2. At the calibration site inside the box shown in the figure, for example, Run 3 oversimulates the observed flow (relative total bias of 0.0481). LZSN should be increased to reduce streamflow, but instead LZSN is decreased with a φ_2 of 0.93, calculated based on a summer flow bias of 0.1663, and a winter flow bias of 0.055.

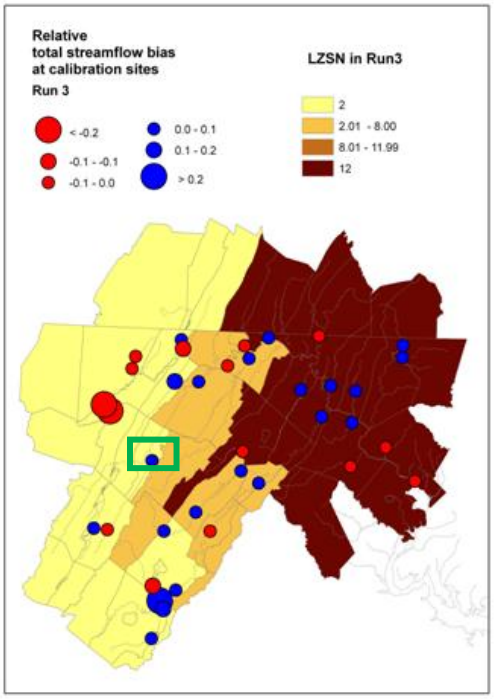


Figure 5-16. Spatial distribution of LZSN values and the relative total bias in streamflow using Run3.

5.5.4 Fluxes

By adjusting the LZSN value in autocalibration, the model responds differently to different types of precipitation inputs. This study examined the simulated hydrologic fluxes to understand how the model hydrology is affected by LZSN as the model is adjusted to different precipitation inputs. The model generates active groundwater flow (AGWO), interflow (IFWO), and surface runoff (SURO) for each pervious land use type in each land segment. The average annual total flux in Run1 was compared with that in Run 3 (Figure 5-17). Despite the different precipitation inputs and the different parameters, the total land outflow to streams is equal or nearly equal at 36 out of the 57 land segments, lying close to the 1-to-1 line +/- 10%. Ten land segments have greater flux using mMPE, whereas eleven land segments have greater flux using mCBP.

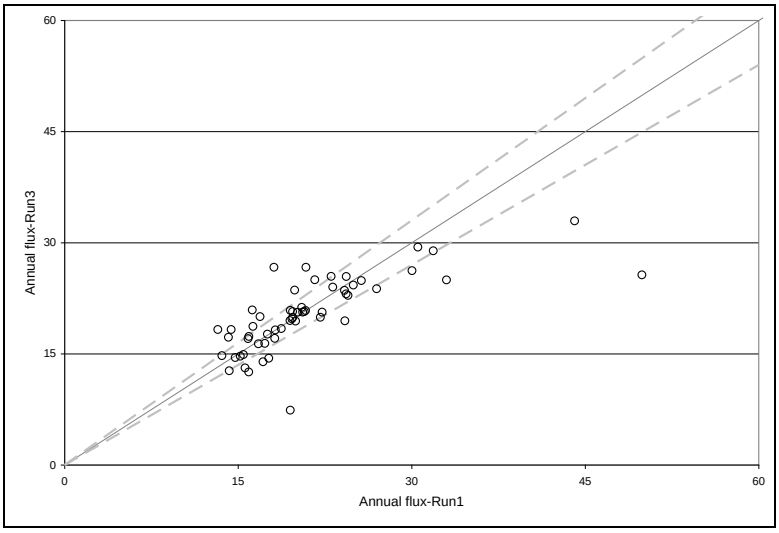


Figure 5-17. Comparison of the average annual total flux [in] between Run 1(using mCBP) and Run 3(using mMPE) of the 57 land segments. The two dashed lines fall 10% above and below the 1-to-1 line.

For the closer look at the annual total flux seen in Figure 5-17, the total flux from the pervious and impervious areas were compared (Figure 5-18). The annual total runoff

from the impervious area in any land segment is the same for Runs 1 and 3, except for one land segment (solid circles in Figure 5-18). The parameters for the impervious land simulation are the same in all Runs; the total runoff from impervious area is, therefore, greatly affected by the discrepancy between the types of precipitation. To explain the large flux discrepancy at the land segment (inside a box shown in Figure 5-18), this study examined the precipitation difference. That land segment has significant average annual difference between mCBP and mMPE: mMPE underestimates mCBP by 18 inches, the second largest difference within the entire basin. Compared to Run1, the volume of precipitation in Run3 is too small to generate as much outflow.

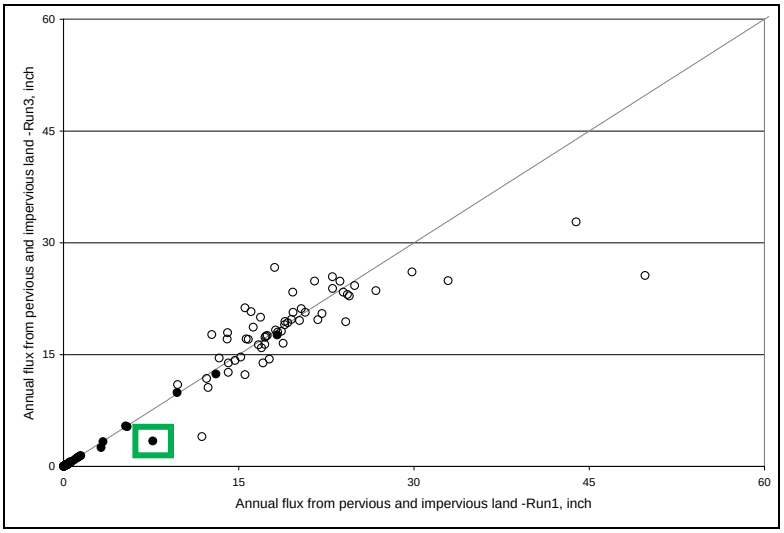
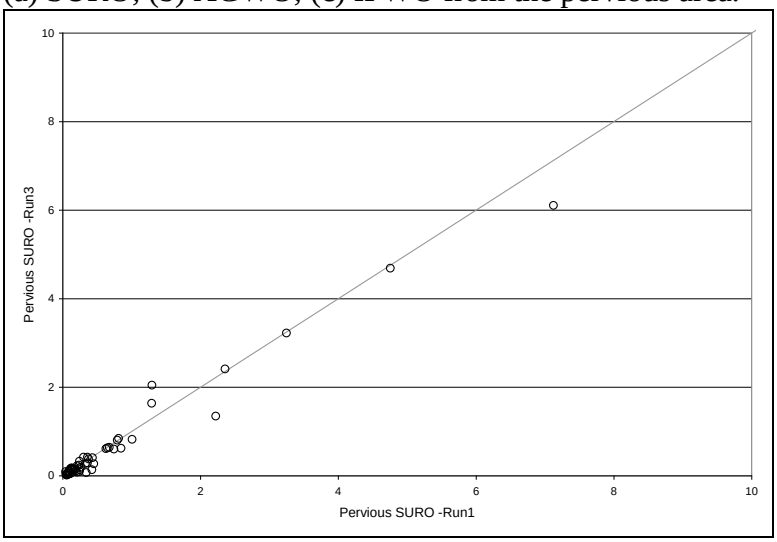


Figure 5-18. Comparison of annual total flux between Runs 1 and 3 from the pervious (○) and impervious areas (●).

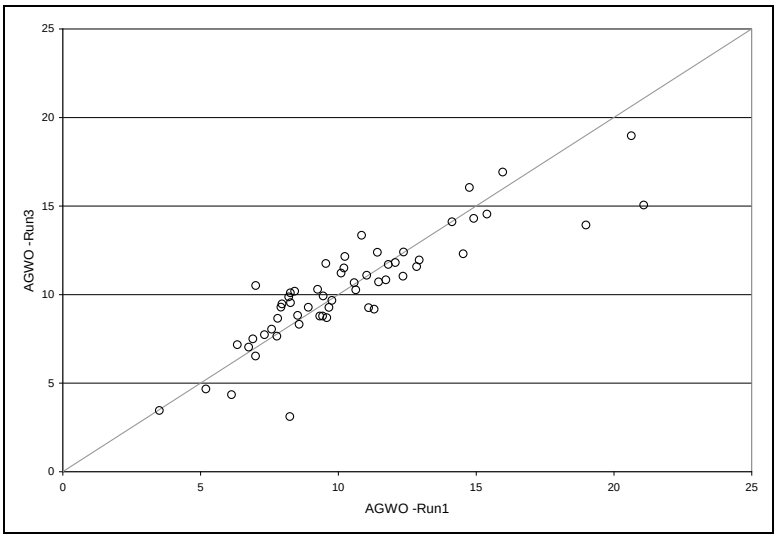
Since the pattern shown in Figure 5-17 is similar to the difference in the total flux from the pervious areas (open circles in Figure 5-18) between Runs 1 and 3, each flow type from the pervious area was compared to identify the source of the difference in annual total flux from pervious area between Runs 1 and 3 [Figures 5-19 (a) – (c)].

SURO from the pervious area is less than 8 inches at the majority of the land segments and consistent between the two types of precipitation data [Figure 5-19(a)].

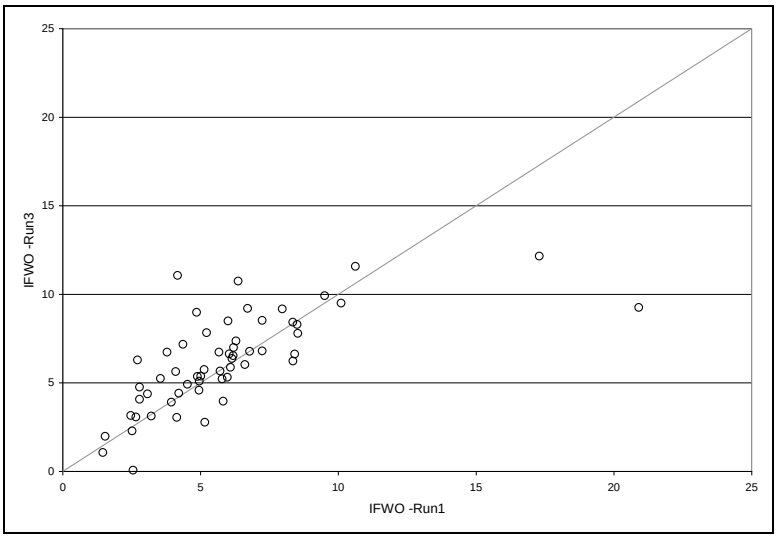
Figure 5-19. Comparison of annual total flux [in] between Runs 1 and 3 based on (a) SURO; (b) AGWO; (c) IFWO from the pervious area.



(a)



(b)

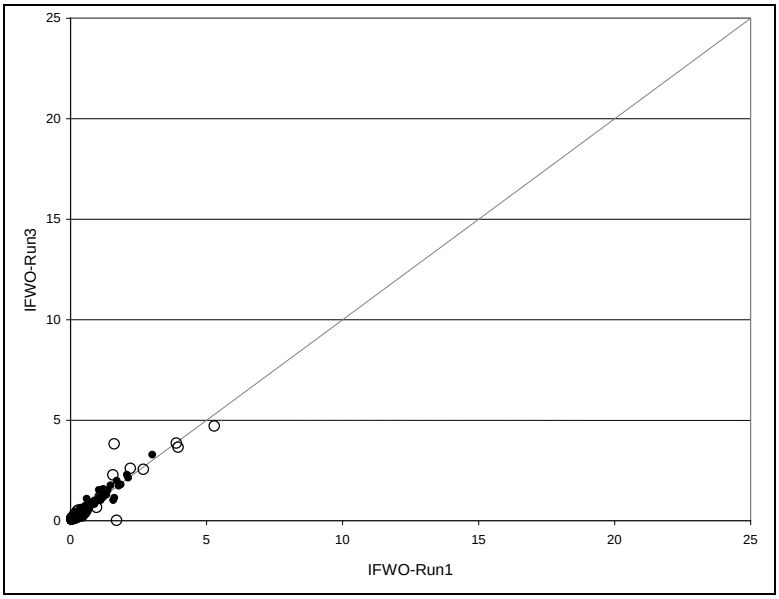


(c)

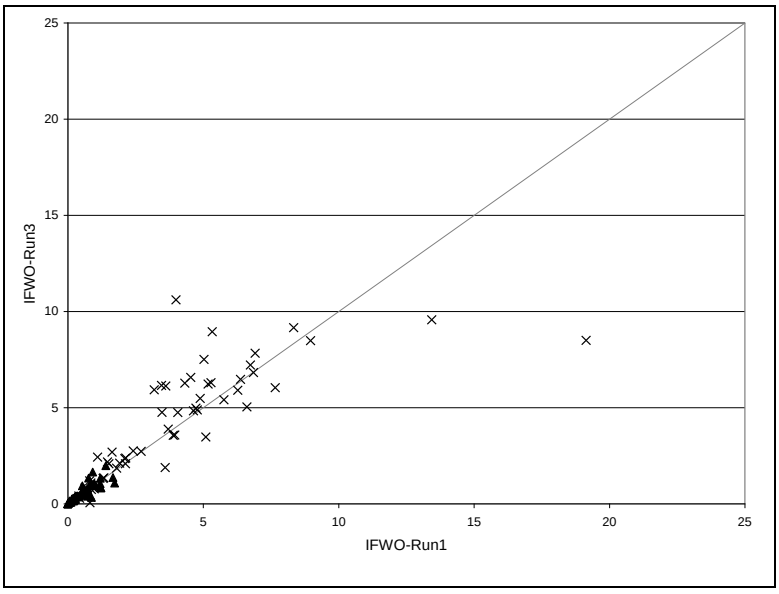
Based on the graphical analysis of Figure 5-19, AGWO and IFWO are the contributing components to the difference in annual total flux. To further analyze AGWO and IFWO, this study examined the effect of land use on each outflow (Figure 5-20).

To simplify the visual presentation, this study grouped the land uses into four general categories: crop, pervious urban, pasture, and forest. Except for forest, the types of land use do not seem to affect interflow [IFWO in Figures 5-20 (a) and (b)] and ground water flow [AGWO in Figures 5-20 (c) and (d)], showing no noticeable changes between Runs 1 and 3. In addition, fluxes from other land uses are at most 30% of flux from forest [Figures 5-20 (b) and (d)]. The noticeable changes between Runs 1 and 3 are found in forest IFWO and AGWO, showing that they are affected by the type of precipitation data and the different parameter sets determined in calibration.

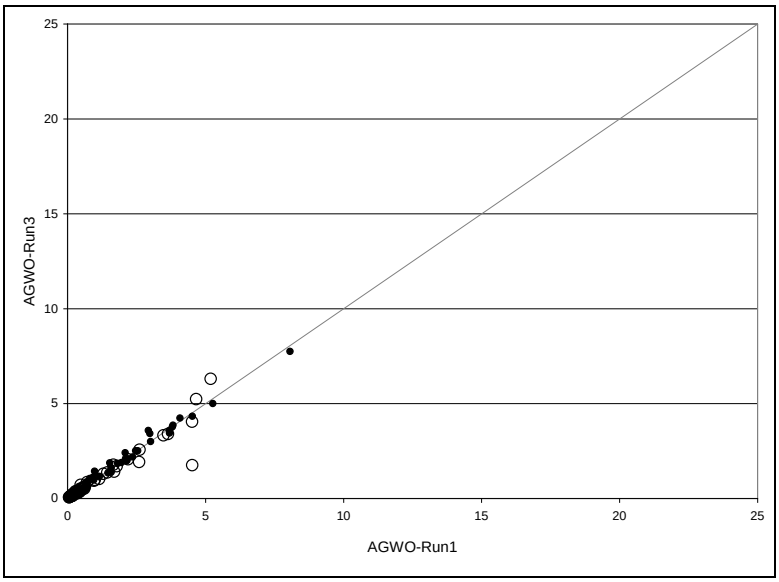
Figure 5-20. Comparison of IFWO and AGWO by land use: (a) IFWO by crop and pervious urban; (b) IFWO by forest and pasture; (c) AGWO by crop and pervious urban; (d) AGWO by forest and pasture. Crop (○); pervious urban (●); pasture (▲); forest (x)



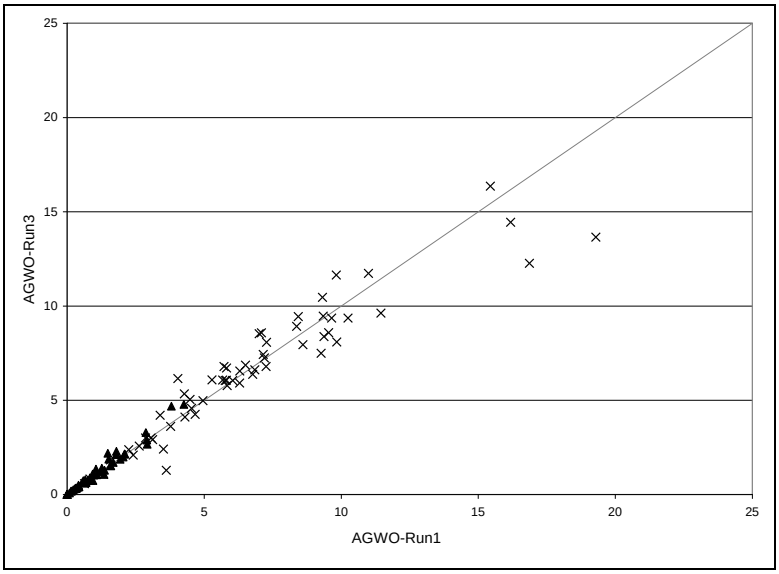
(a)



(b)



(c)



(d)

To examine the hydrologic process for AGWO and IFWO in the model, the ratio of the time varying lower zone storage (LZS) to the parameter LZSN was calculated. The ratio (LZRAT) controls infiltration and inflow (Equations 5-6 to 5-9) and the fraction of the infiltrated and percolated water entering to ground water storages (Equations 5-2 to 5-4). For LZS, the average annual value was calculated for each land segment. The

spatial pattern of LZS indicates that Run 1 does not have an average LZS value of less than 7.5 inches, whereas half of the land segments using Run 3 show such low values (Figure 5-21). On the eastern part of the basin where both Runs 1 and 3 have the maximum LZSN value of 12 (Figure 5-15), LZS values are also high. As a result, the ratio of LZS to LZSN for Run 1 is the same as that for Run 3 in this region, depicted as circles near the one-to-one line in Figure 5-22. In general, LZRAT for Run 3, which ranges from 1.12 to 1.92, is higher than that for Run 1(0.9 ~1.8). These different LZRAT values are used in Equations from 5-6 to 5-9 to generate different potential interflow inflow for Run 1 and Run 3. This explains why the interflow component (IFWO) is affected by the volume of precipitation.

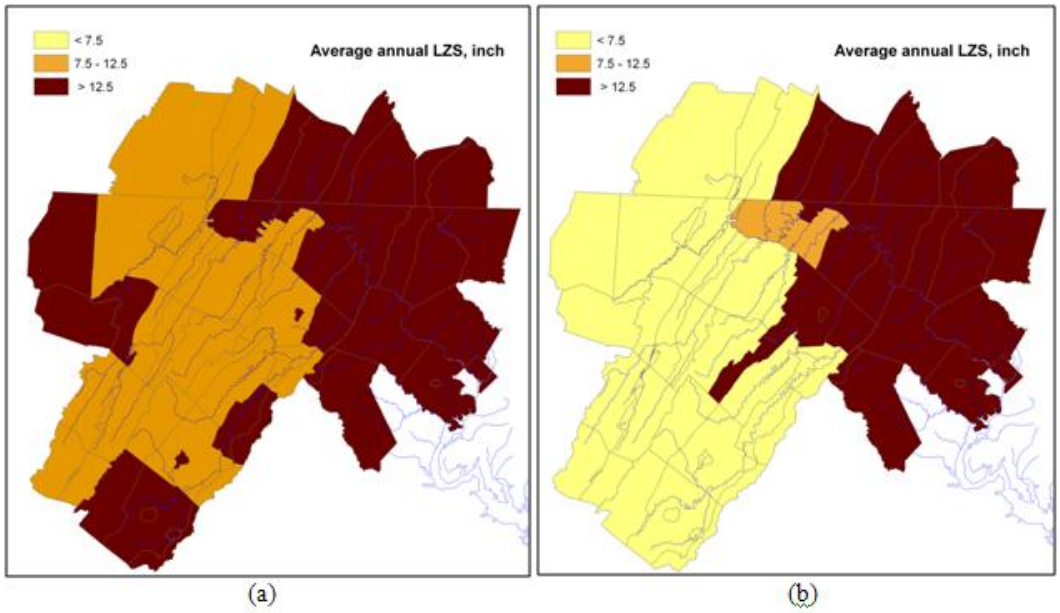


Figure 5-21. Spatial distribution of the average annual LZS for (a) Run 1; (b) Run 3.

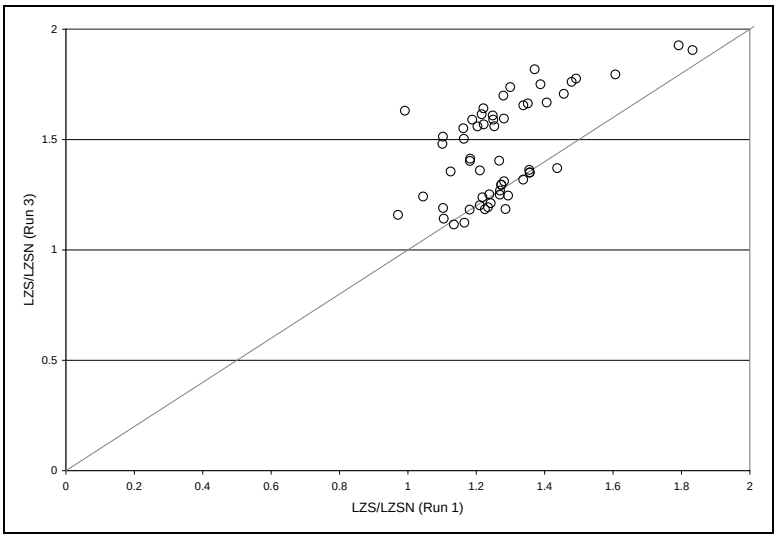


Figure 5-22. Plot of LZRAT (LZS/LZSN) for Run 1 and Run 3. The LZRAT was calculated using the average annual state variable LZS (2002-2004) over the calibrated LZSN at a given land segment. Each circle represents a land segment.

Groundwater storage and the groundwater flow component (AGWO) follow the partitioning curve as illustrated in Bicknell et al. (2001). Using LZRAT - the average annual LZS values over calibrated LZSN, the fraction for ground water storages from infiltrated and percolated water was calculated (Equations 5-2, 3 and 4 or Figure 5-23). In Figure 5-23, the portion below the curve is assigned to the lower zone and the portion above the curve is assigned to ground water storage. For Run 1, the range of LZRAT (0.9 ~ 1.8) allows 46% ~ 92% of infiltrated and percolated water to be available for ground water. For Run 3, the fraction of the infiltrated and percolated water for ground water is higher, ranging from 60% to 95% with higher LZRAT than Run 1. This explains why AGWO is affected by the volume of precipitation.

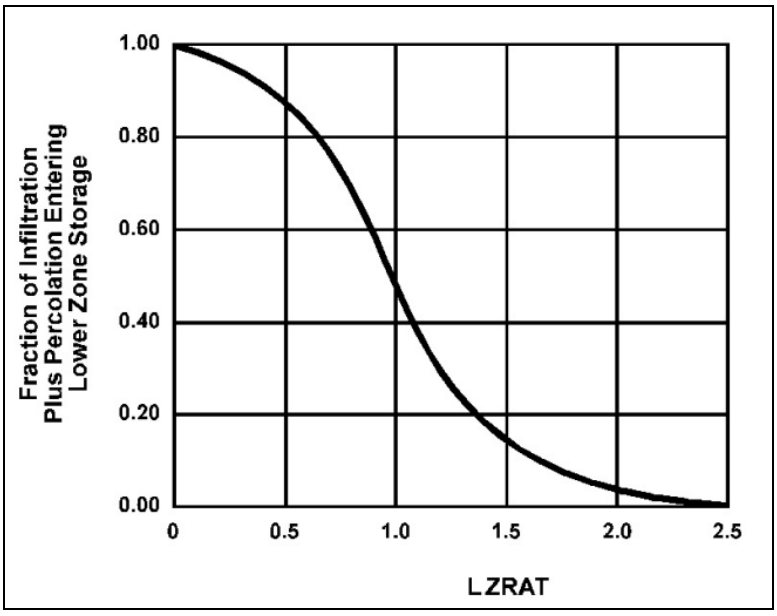
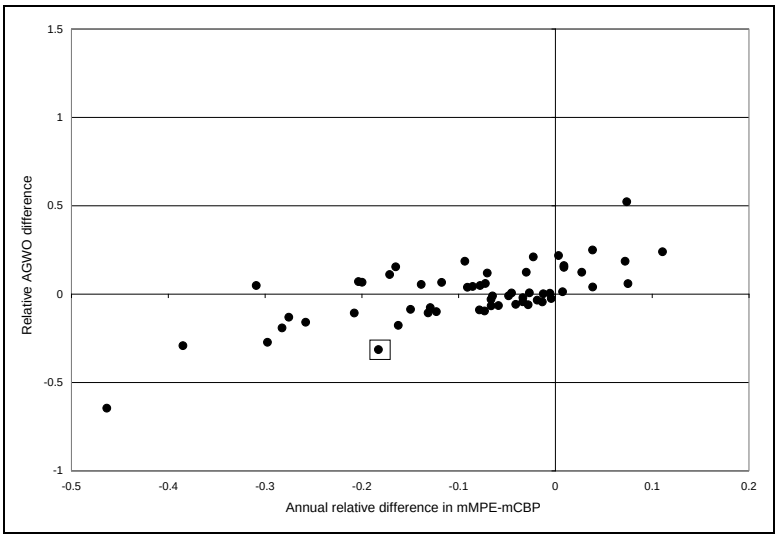
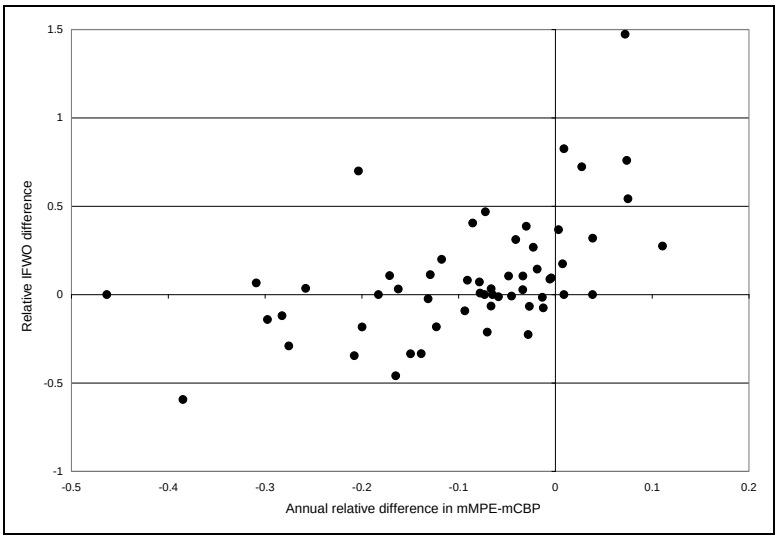


Figure 5-23. Fraction of infiltration plus percolation entering lower zone storage. Adapted from Bicknell et al. (2001).

The difference between Run 1-flux and Run 3-flux is plotted against the average annual difference between mCBP and mMPE (Figure 5-24). The difference in average annual precipitation from Chapter 3 was used for comparison with the difference in fluxes from forest at any given land segment. Both fluxes illustrate the same tendency, but AGWO (r^2=0.489) is more strongly affected by the type of precipitation than IFWO (r^2=0.259). Both AGWO and IFWO are affected by parameter constraints. For example, the AGWO value inside the box in Figure 5-24 (a) has the third largest difference in AGWO between Runs 1 and 3, although it has moderate difference between the two precipitation data sets. It is observed that the difference in mCBP and mMPE between Run 1 and Run 3 is reflected in differences in subsurface flow of the same sign: the greater volume of precipitation the model has, the greater AGWO or IFWO the model release. On the other hand, the surface flow component is generally the same for the different precipitation time series.



(a)



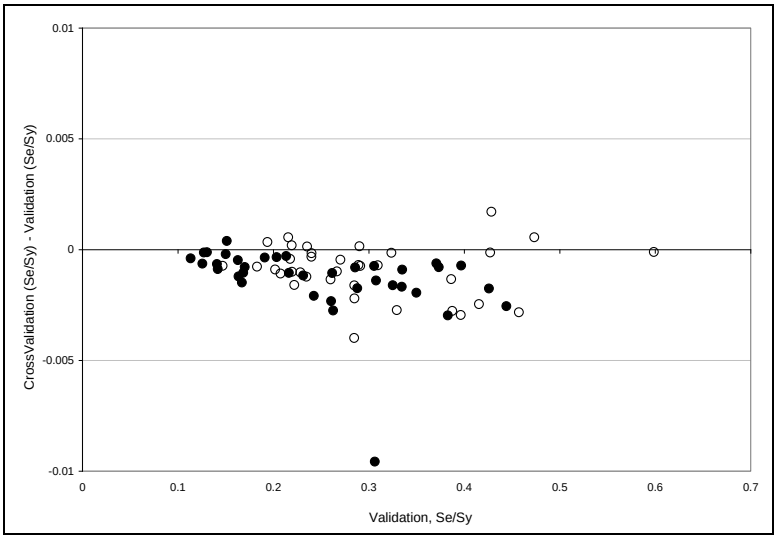
(b)

Figure 5-24. Plot of the relative annual difference between the two types of precipitation and the relative difference in fluxes [(flux from Run 3-flux from Run 1)/flux from Run1]: (a) forest-AGWO; (b) forest-IFWO.

This study demonstrates how the different types of precipitation data affect model hydrology, causing changes in partitioning of different components of flow. The model calibrations attempt to match the same streamflow time series with different precipitation inputs. Direct surface runoff, a small fraction of the total, agrees well. The differences are found in the subsurface components.

5.5.5 Cross-validation

Using the model calibrated to mCBP and recalibrated to mMPE, this study evaluated the errors occurring when the model simulated streamflow using switched inputs. Cross-validation was conducted as explained in Section 5.4.5. The change in Se/Sy between cross-validation and validation accuracies is plotted (Figure 5-25). The plot shows that driving the mCBP-calibrated model with mMPE (Run 1) increases model accuracies at all of the stations (solid circles). When the calibration accuracy using mCBP in Run 1 is poor (high value of Se/Sy on the x-axis), cross-validation using mMPE improves (larger negative values on the y-axis). The results for driving the mMPE-calibrated model with mCBP (Run 3) show a similar pattern with more variations and increased Se/Sy in a few cases. Overall, the magnitude of the change (y-axis) is not significant, based on the change in Se/Sy of less than 1%. It is hard to show a certain trend in model performance based on a single-year validation and cross-validation analyses.



5-25. Change in Se/Sy between the validation and cross-validation using Run 1 (●) and Run 3 (○). The study period is 2005. Each point represents a calibration site.

5.6 Discussion and Conclusion

This chapter examined how hydrologic processes in the CBW model vary when the model is calibrated with different precipitation data sets. The calibrated model parameters and fluxes were compared using Run1 and Run 3. The former is calibrated using mCBP and the latter using mMPE. Questions proposed at the beginning of the study are discussed in this section.

5.6.1 Changes in parameters and fluxes when mCBP is replaced with mMPE

In the CBW model, six parameters were calibrated during the autocalibration process: potential evapotranspiration adjustment factor (LAND_EVAP), lower zone nominal storage (LZSN), infiltration index (INFILT), interflow inflow index (INTFW), interflow recession index (IRC), and groundwater recession index (AGWR). Overall, with both precipitation data sets, LZSN and LAND_EVAP tend to take values at the boundaries of their permissible ranges during the calibration process. Other parameters remain within the specified ranges. When mCBP is replaced with mMPE, the LAND_EVAP parameter changes as expected: with a smaller volume of precipitation, the model must reduce evapotranspration in order to match the observed streamflow and satisfy overall water balance. Therefore, Run 3 has lower LAND_EVAP values than Run 1 at the majority of the calibration sites.

This study finds that assigning boundaries for LZSN is crucial to keep parameter values reasonable during calibration. Without constraining LZSN in the calibrations, models show numerical instability and non-convergence of LZSN, disturbing the other parameters which then must compensate.

Generally, using mMPE (Run 3) results in lower LZSN values, to reduce the amount of water that can be stored in the lower zone. However, this conceptual process does not always hold true over the entire PRB. Biases in seasonal streamflow are driven largely by the characteristics of precipitation and also affected by meteorological conditions. With large differences in summer and winter flow biases, LZSN values become constrained by user-defined lower and upper parameter boundaries. When the parameter value has reached one of these boundaries, the iteration process does not affect further improvement of the calibration accuracy. As a result, model accuracies are generally poor in regions of extreme LZSN values. Based on the results from Run 1 (using mCBP), this study cannot explain the model's spatial responses due to a lack of spatial patterns.

This study also examined the simulated hydrologic fluxes to understand how the model hydrology is affected by LZSN as it is adjusted to different precipitation inputs. Based on the analysis of fluxes by various land uses, interflow and active groundwater flow (IFWO and AGWO) from forest are mainly affected by the difference in forcing inputs, whereas the surface flow component is generally the same for the different precipitation time series.

Despite the different precipitation inputs and the different parameters, the total land outflow to streams is equal or nearly equal in the comparison of Runs 1 and 3 at 36 of the 57 land segments. The spatially distinctive difference in model accuracy between Runs 1 and 3 arises from differing degrees of simultaneous agreement with observed flow, due to different partitioning of flow components.

Calibration accuracy is improved with recalibration using revised parameter constraints. Based on the 1-year cross-validation analysis, however, it is difficult to draw a conclusion whether or not the model requires the same type of precipitation data that it is calibrated to for its application uses.

5.6.2 Considerations in selecting precipitation inputs for the CBW model

This study reveals that driving the CBW model with different types of precipitation data results in different partitioning of hydrologic components. This observation may provide useful information to water quality modelers using the CBW model. The model was originally developed to provide scientific information to restore the Chesapeake Bay by reducing nutrients and to develop Total Maximum Daily Loads (Chapter 2). Different partitioning of a hydrograph may affect the prediction of fate and transport of nutrients and pollutants by affecting their residence time in the subsurface or surface water.

Considering that mCBP time series contain data up until 2005, MPE is a strong alternative data source for the CBW model. The critical phase in preparing the mCBP time series would be collecting data from more than 500 daily and hourly gauge stations and filling the missing data. As this study illustrates in Chapter 3, a significant portion of station data are missing or invalid, whereas in MPE, the missing data is filled in by experts in both the meteorological and hydrological fields before the data becomes available to the public. Further, the spatial domain of MPE data available in MARFC covers the entire Chesapeake Bay model domain, meaning no spatial interpretation or

mosaics of the data are needed for the model. This is a critical advantage of using MPE for model users.

As this study indicates, mMPE can be implemented into the CBW model at a basin scale without major technical issues. Compared to the model using mCBP, the model using mMPE responds more distinctively to the spatial pattern of other forcing inputs over the Potomac River Basin. Therefore it is easier to identify spatial regions for calibration.

Model users decide which data sets are appropriate to use, depending on the purposes of studies and other factors. With 1-year cross-validation analysis, this study could not draw a clear conclusion whether or not the model calibrated to mCBP (mMPE) can be used in predicting streamflow or in any other applications using switched precipitation inputs, mMPE and mCBP, respectively. However, this study demonstrates how the different types of precipitation data affect model hydrology, causing changes in the partitioning of precipitation into different components of flow. For future model applications, such as TMDL studies, the model may still need to be calibrated with the same type of data as is used in simulation and for the specific location, following the findings of Shirmohammadi et al. (2006). They reported that the model parameters resulting from the calibration process generally apply only to that watershed and corresponding measured data. For the CBW model users, the model as currently calibrated can be driven with mCBP, which would entail preparing the gauge-based precipitation data sets after 2005 for future use. Or, model users can recalibrate the model with mMPE and drive it for various applications using mMPE that are readily available.

CHAPTER SIX

Research Summary, Contributions, and Recommendations for Future Work

6.1 Summary

The main goal of this study was to examine the impact of different precipitation data sources on the hydrologic calibration of the Chesapeake Bay Watershed (CBW) model. To achieve this goal, three specific objectives were developed: (1) compare gauge-based and NEXRAD radar-based data at the (a) point (COOP/ASOS)-pixel (MPE) level and (b) spatially aggregated level (mCBP and mMPE); (2) evaluate the calibration accuracy of the CBW model using the two types of precipitation data averaged to the spatial scale of the model, mCBP and mMPE; (3) examine how model hydrology responds differently to the two forcing inputs in the CBW model. The investigation was conducted in the non-tidal Potomac River Basin, a subwatershed of the Chesapeake Bay.

Using MPE in hydrologic and water quality modeling appears promising, due to its spatial and temporal coverage. MPE does not suffer from the temporal data gaps, which are concerns in the gauge records. Unlike for gauge-based precipitation data, spatial interpolation is not required. At the 80 station locations analyzed, the average annual reported MPE pixel and gauge-point precipitation values (when both are valid) agree well. This study, however, finds significant differences in average annual precipitation between the spatially aggregated mMPE and mCBP over the 114 land segments analyzed.

When the CBP's precipitation input is replaced with mMPE and the model is recalibrated (Run 3), the calibration accuracy is higher than using mCBP (Run 1) at 65% of the 37calibration sites. On the other hand, when the model is run using mMPE without recalibration, the calibration accuracy decreases. This indicates that when one type of precipitation input is replaced with another, the model needs to be recalibrated. The study also finds that revised parameter constraints are necessary in recalibration.

The model response is controlled mainly by seasonal difference in precipitation inputs in two aspects. First, even a recalibration with revised constraints cannot improve calibration accuracy because the model calibration process is unable to compensate for large difference in seasonal flow bias caused by the seasonal difference in precipitation data. When the Lower Zone Nominal Storage (LZSN) value reaches the user-defined lower and upper parameter boundaries because of large differences in seasonal flow bias, the iteration process during the autocalibration does not affect further improvement of calibration accuracy and the calibration accuracy is therefore, generally poor.

Second, the CBW model responds differently in calibration by changing the partitioning of hydrograph components to match the observed hydrograph when forced with two different types of precipitation data. Depending on seasonal flow biases introduced by the seasonal difference between the two types of precipitation data sets, the lower zone nominal storage parameter (LZSN) changes sensitively, according to the autocalibration rules (Table 4-1) set for the CBW model. Because LZSN is used to calculate interflow (IFWO) and active groundwater flow (AGWO) in the HSPF model hydrology, IFWO and AGWO are affected as LZSN changes. This finding is model-specific. If hydrologic processes are differently represented in a given model, the model

response to different types of precipitation data might be different from the CBW's response.

The comparison of the model responses to mMPE and mCBP indicates that the model seems to be affected more by the total volume of precipitation than by the average characteristics of individual storm events. More land segments in the model reach the maximum allowed values of LAND_EVAP when using mCBP rather than mMPE, reflecting the fact that the average annual mCBP is greater than the average annual mMPE (Table 3-7). The model does not seem to respond to the fact that the total depth per mMPE event slightly exceeds the total depth per mCBP event as seen in Table 3-9. This study's findings about the model response to the different types of precipitation inputs are model-specific: the findings in this study may hold true only for the CBW model.

Because of the nature in hydrology, this study suggests that non-linearity and autocorrelation should be accounted in the process of the model evaluation. GIS-based spatial analysis is also a useful tool when model units for land and river simulations are not necessarily the same. Visualizing the spatial pattern of input data, simulated outputs, and derived outputs (calibration statistics, for example) helps model users interpret model results.

6.2 Contributions to the Modeling Community

This study demonstrated potential contributions to the modeling community. First, it provided descriptive statistics that identified the similarities and differences between gauge-based and radar-based precipitation data sets. The results of the data

comparison help modelers assess the implication of using one data type over the other in hydrologic modeling.

Second, this study showed the beneficial aspects of implementing the radar-based precipitation into the CBW model. Using MPE in hydrologic and water quality modeling appears promising, due to its spatiotemporal coverage, rare missing records, and improved calibration accuracy of the CBW model. Considering that mCBP time series contain data through 2005 and that major analysis will be required to complete the data set to present time, MPE is a strong alternative data source for the CBW model as it is being produced operationally. By revealing the fact that the model needs to be recalibrated with revised parameter constraints when one type of precipitation input is replaced with another, this study suggests the following: For the CBW model users, the model as currently calibrated can be driven with mCBP, which would entail preparing the gauge-based precipitation data sets after 2005 for future use. Or, model users can recalibrate the model with mMPE and drive it for various applications using mMPE that are readily available.

Third, this study offers guidance for model's application uses using the MPE product. The surface flow component in this study is generally the same for the different precipitation time series. This finding indicates that the current mCBP-calibrated CBW model can use mMPE without recalibration for surface flow-dominated processes such as floods and sediment transport. However, the effects of the different forcing inputs on interflow and groundwater flow components may have implications for water quality simulations using the CBW model or similar HSPF-based models, because internal

hydrologic process mechanisms control the residence time, fate and transport of nutrients and pollutants.

Fourth, this study also demonstrated that spatially aggregated MPE precipitation data (mMPE) can be implemented into the CBW model at a basin scale without major technical issues. Beyond technical feasibility, the study also demonstrates an improvement of calibration accuracy for the time period analyzed. This is a practical and technical contribution to the CBW modeling community, since the community has expressed its need for the implementation of radar-based precipitation inputs (Chapter 1).

6.3 Recommendations for Future Work

In the course of this study, a number of questions were raised, and possible alternative approaches were identified. Further investigation of the following issues is recommended.

First, although the model calibration is largely affected or biased by the characteristics of the major forcing inputs, additional biases can be reduced through the calibration of the model at a sub-basin scale. Not only statistical indices, but also tools for spatial analysis are useful in identifying the new sub-basin scale when using mMPE. For example, the PRB can be divided into two regions according to the meteorological boundaries. Such a division might help to reduce seasonal streamflow biases that control the calibration of LZSN.

Second, the data for land use is recommended to be updated for the Phase 5.2, the model version used in this study. The calibration statistics were calculated based on results from 2002 to 2004, but the model used the land use data set created by the CBP

based on data sources for 2000. Since hydrologic fluxes are simulated per acre by land use at a given land segment, information on land use may significantly affect the outflows to streams and therefore calibration and prediction accuracy as well. Otherwise, the use of Phase 5.3, the newer model version, is recommended.

Third, the calibration of river parameters to compensate for the difference in land variables is also worthy of experimenting. The river parameters include a ratio of the maximum velocity to the mean velocity in the stream channel cross section or a weighting factor for hydraulic routing. The scope of this study was limited to follow the autocalibration approach that the CBP currently uses. Since the autocalibration procedure set by the CBP does not include river parameters, the model setting for the river simulation remains the same throughout the entire experiment in this study. If land and river parameters compensate for each other, however, the rationality of parameters should be evaluated because they may not reflect physical differences between the watershed regions and river reaches they are meant to represent. This experiment can end up as a mechanical fitting exercise.

Fourth, the model performance for the validation period should be evaluated using longer precipitation time series. This study used a single year (2005) for the validation and cross-validation analyses, and it is difficult to draw any conclusions due to the small sample size. As the availability of MPE data sets improves, more rigorous validation and cross-validation can be conducted in the future.

REFERENCES

AOPA (Aircraft Owners and Pilots Association) (AOPA), 1999. Automated Surface Observing System. Safety Advisor series. Air Safety Foundation, Frederick, MD, U.S.A..

Atkins, J. T., Jr., Wiley, J. B., Paybins, K. S., 2005. Calibration parameters used to simulate streamflow from application of the Hydrologic Simulation Program-FORTRAN Model (HSPF) to mountainous basins containing coal mines in West Virginia. Scientific Investigations Report 2005-5099. U.S. Geological Survey, Reston, VA, U.S.A..

Baker, D. B., Richards, R. P., Loftus, T. T, Kramer, J. W., 2004. A new flashiness index: characteristics and applications to Midwestern rivers and streams. American Water Resources Association (JAWRA) 40(2), 503-522.

Band, L., Campbell, K., Kinerson, R., Reckhow, K., Welty, C., 2005. Review of the Chesapeake Bay Watershed modeling effort. Chesapeake Research Consortium, Inc., Edgewater, MD, U.S.A..

Band, L., Dillaha, T., Duffy, C., Reckhow, K., Welty, C., 2008. Chesapeake Bay Watershed model Phase 5 review. Chesapeake Research Consortium, Inc., Edgewater, MD, U.S.A..

Bedient, P. B., Holder, A. W., Thompson, J. F., Fang, Z., 2007. Modelling of storm-water response under large tailwater conditions: case study for the Texas medical center. Journal of Hydrologic Engineering 12(3), 256-266.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L., Jr., Jobes, T.H., Donigian, A.S., Jr., 2001. Hydrological Simulation Program - Fortran (HSPF): User's Manual for Release 12. U.S. Environmental Protection Agency. Athens, GA, U.S.A..

Borga, M., Degli Esposti, S., Norbiato, D., 2006. Influence of errors in radar rainfall estimates on hydrological modeling prediction uncertainty. Water resources research 42(8), W08409.

Bradley, A. A., 1997. Use of Weather Radar with Lumped Parameter Hydrologic Models Preprints, 13th Conference on Hydrology, 77th Annual AMS Meeting, J87–J88.

Bradley, A. A., Kruger, A., 1998. Recalibration of hydrologic models for use with WSR-88D precipitation estimates. Special Symposium on Hydrology; 78th Annual AMS Conf., American Meteorological Society, Boston, U.S.A..

Cashman, E., Potter, K., 2006. Modeling the impacts of climate variability on sediment transport. Sediment Dynamics and the Hydromorphology of Fluvial Systems. Proceedings of a symposium. Dundee, UK. 306, 611-619.

Chang, M., Harrison, L., 2005. Field assessments on the accuracy of spherical gauges in rainfall measurements. Hydrological Processes 19, 403-412.

CBP (Chesapeake Bay Program), Chesapeake community modeling program. URL: http://ches.communitymodeling.org/models.php. Accessed in 2009.

Choi, J., Scott, S., Olivera, F., 2008. Hourly disaggregaton of daily rainfall in Texas using measured hourly precipitation at other locations. Journal of Hydrologic Engineering 13(6), 476-487.

Ciach, G.J., 2003. Local random errors in tipping-bucket rain gauge measurements. Journal of Atmospheric and Oceanic Technology 20, 752-759.

Ciach, G. J., Krajewski, W. F., Villarini, G., 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. American Meteorological Society 8, 1325-1347.

Daly, C., Gibson, W. P., Taylor, G. H., Doggett, M. K., Smith, J. I., 2007. Observer bias in daily precipitation measurements at United States Cooperative network stations. Bulletin of the American Meteorological Society 88(6), 899-912.

Das, T., Bardossy, A., Zehe, E., He, Y., 2008. Comparison of conceptual model performance using different representations of spatial variability. Journal of Hydrology 356, 106-118.

Debele, B., Srinivasan, R., Parlange, J., 2007. Accuracy evaluation of weather data generation and disaggregation methods at finer timescales. Advances in water resources 30, 1286-1300.

Dingman, S. L., 2002. Physical hydrology. Waveland Press, Inc., Illinois, U.S.A..

Engelmann, C. J. K., Ward, A. D., Christy, A. D., Bair, E. S., 2002. Application of the BASINS database and NPSM model on a small Ohio watershed. Journal of the American Water Resources Association 38(1), 289-300.

Fulton, R. A., Breidenbach, J. P., Seo, D. -J., Miller, D. A., 1998. WSR-88D rainfall algorithm. Weather Forecasting 13, 377-395.

Garcia-Pintado, J., Barbera, G. G., Erena, M., 2009. Rainfall estimation by rain gauge-radar combination: a concurrent multiplicative-additive approach. Water resources Research 45, W01415.

Geer, I. W., 1996. Glossary of weather and climate:with related oceanic and hydrologic terms. American Meteorological Society, Boston, MA, U.S.A..

Giangrande, S. E., Ryzhkov, A. V., 2008. Estimation of rainfall based on the results of polarimetric echo classification. Journal of applied meteorology and climatology 47(6), 2445-2462.

Goudenhoofdt, E., Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrology and Earth System Sciences. 13 (2), 195-203.

Guo, J., Liang, X., Leung, L. R, 2004. Impacts of different precipitation data sources on water budgets. Journal of hydrology 298, 311-334.

Gupta, H. V., Sorooshian, S., Yapo, P. O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. Journal of hydrologic engineering 4(2), 135-143.

Gutierrez-Magness, A. L., McCuen, R. H., 2005. Effect of flow proportions on HSPF model calibration accuracy. Journal of Hydrologic Engineering 10(5), 343-352.

Gyasi-Agyei, Y., Mahbub, S. M. P. B., 2007. A stochastic model for daily rainfall disaggregation into fine time scale for a large region. Journal of hydrology 347, 358-370.

Habib, E., 2009. Validation of NEXRAD multisensory precipitation estimates using an experimental dense rain gauge network in south Louisiana. Journal of Hydrology 373(3-4), 463-478.

Habib, E., Aduvala, A. V., Meselhe, E. A., 2008. Analysis of radar-rainfall error characteristics and implications for streamflow simulation uncertainty. Hydrological Sciences-Journal-DES Sciences Hydrologiques 53(3), 568-587.

Habib, E., Ciach, G. J., Krajewski, W. F., 2004. A method for filtering out raingauge representativeness errors from the verification distributions of radar and raingauge rainfall. Advances in Water Resources 27, 967-980.

Habib, E., Krajewski, W. F., Kruger, A., 2001. Sampling errors of tipping-bucket rain gauge measurements. Journal of Hydrologic Engineering 6(2), 159-166.

Habib, E., Messelhe E. A., Aduvala, A. V., 2008. Effect of local errors of tipping-bucket rain gauges on rainfall-runoff simulations. Journal of hydrologic engineering 13(6), 488-496.

Hay, L., Clark, M., Leavesley, G., 2000a. Use of atmospheric forecasts in hydrologic models-Part 2: Case study. In proceedings of the American Water Resources Association's Spring Specialty Conference on Water Resources in Extreme Environments, Anchorage, Alaska, April 2000.

Hay, L. E., Leavesley, G. H., Clark, M. P., Markstrom, S. L., Viger, R. J., Umemoto, M., 2006. Stepwise, multiple objective calibration of a hydrologic model for a snowmelt dominated basin. Journal of the American Water Resources Association 42(4), 0877-0890.

Hay, L. E., McCabe, Jr., G. J., Wolcok, D. M., Ayers, M. A., 1991. Simulation of precipitation by weather type analysis. Water Resources Research 27(4), 493-501.

Hay, L. E., Wilby, R. L., Leavesley, G. H., 2000b. A comparison of delta change and downscale GCM scenarios for three mountainous basins in the United States. Journal of the American Water Resources Association 36(2), 387-397.

Humphrey, M. D., Istok, J. D., Lee, J. Y., Hevesi, J. A., Flint, A. L., 1997. A new method for automated dynamic calibration of tipping-bucket rain gauges. Journal of Atmospheric and Oceanic Technology 14, 1513-1519.

Hunter, S. M., 1996. WSR-88D Rdar rainfall estimation: Capabilities, limitations and potential improvements. National Weather Digest 20, 26-38.

Jacobs, J. H., Srinivasan, R., 2005. Effects of curve number modification on runoff estimatioin using WSR-88D rainfall data in Texas watershed. Journal of soil and water conservation 60(5), 274-279.

Jain, S. K., Sudheer, K. P., 2008. Fitting of hydrologic models: a close look at the Nash-Sutcliffe index. Journal of Hydrologic Engineering 13(10), 981-986.

Jayakrishnan, R., Srinivasan, R., Arnonld, J. G., 2004. Comparison of raingauge and WSR-88D Stage III precipitation data over the Texas-Gulf basin. Journal of Hydrology 292, 135-152.

Jayakrishnan, R., Srinivasan, R., Santhi, C., Arnold, J. G., 2005. Advances in the application of the SWAT model for water resources management. Hydrological Processes 19, 749-762.

Jayawickreme, D. H., Hyndman, D. W., 2007. Evaluating the influence of alnd cover on seasonal water budgets using NEXRAD rainfall and streamflow data. Water Resources Research 43(2), W02408.

Ji, Z. G., 2008. Hydrodynamics and water quality: modeling rivers, lakes, and estuaries. John Wiley & Sons, Inc., New Jersey, U.S.A..

Kalin, L., Hantush, M. M., 2006. Hydrologic modeling of an eastern Pennsylvania watershed with NEXRAD and rain gauge data. Journal of Hydrologic Engineering 11(6), 555-569.

Kalinga, O. A., Gan, T. Y., 2007. Small storm precipitation estimation using merged radar and gauging data. International Journal of Remote Sensing 28(6), 1101-1112.

Kim, D., 2008. Optimized merging of hourly precipitation with daily coop precipitation data. 19th conf. on probability and statistics, 20-24 January 2008, New Orleans, LA, U.S.A..

Knebl, M. R., Yang, Z. -L., Hutchison, K., Maidment, D. R., 2005. Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin summer 2002 storm event. Journal of Environmental Management 75, 325-336.

Kondragunta, C., Kitzmiller, D., Seo, D. J., Shrestha, K., 2005. Objective integration of Satellite, rain gauge, and radar precipitation estimates in the multisensor precipitation estimator algorithm. 19th Conference on Hydrology. American Meteorological Society (AMS) Annual Meeting. San Diego, CA, U.S.A..

Kondragunta, C., Seo, D. J., 2004. Toward integration of satellite precipitation estimates into the multi-sensor precipitation estimator algorithm. 18th Conference on Hydrology. American Meteorological Society (AMS) Annual Meeting. Seattle, WA, U.S.A..

Krajewski, W. F., Smith, J. A., 2002. Radar hydrology: rainfall estimation. Advances in Water Resources 25, 1387-1394.

Kuzmin, V., Seo, D. J., Koren, V., 2008. Faze and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search. Journal of Hydrology 353(1-2), 109-128.

Legates, D. R., 2000. Real-time calibration of radar precipitation estimates. Professional Geographer 52(2), 235-246.

Legates, D. R., Deliberty, T. L., 1993. Precipitation measurement biases in the united-states. Water Resources Bulletin 29(5), 855-861.

Lu, J., Sun, G., McNulty, S. G., Amatya, D. M., 2005. A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. Journal of the American Water Resources Association 41(3), 621-633.

Martucci, S., Krstolic, J. L., Raffensperger, J. P., Hopkins, K. J., 2006. Development of land segmentation, stream-reach network, and watersheds in support of hydrological simulation program-Fortran (HSPF) modeling, Chesapeake Bay Watershed, and adjacent parts of Maryland, Delaware, and Virginia: U.S. Geological Survey Scientific Investigations report 2005-5073.

McCuen, R. H., 1993. Microcomputer Application in Statistical Hydrology. Prentice Hall, Inc., NJ, U.S.A..

McCuen, R. H., 1998. Hydrologic analysis and design. Prentice-Hall, Inc. NJ, U.S.A..

McCuen, R. H., 2003. Modeling hydrologic change: statistical methods. Lewis Publishers, NY, U.S.A..

McCuen, R. H., Knight, Z., Cutter, A. G., 2006. Evaluation of the Nash-Sutcliffe Efficiency index. Journal of hydrologic engineering 11(6), 597-602.

Meischner, P. (Ed), 2004. Weather Radar: Principles and advanced applications. Springer, Heidelberg, Germany.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American regional reanalysis. Bulletin of the American Meteorological Society. 87, 343-360.

Molini, A., Lanza, L. G., Barbera, P. L., 2005. The impact of tipping-bucket raingauge measurement error on design rainfall for urban-scale applications. Hydrological Processes 19, 1073-1088.

Moon J., Srinivasan R., Jacobs, J. H., 2004. Stream flow estimation using spatially distributed rainfall in the Trinity River Basin, Texas. Transactions of the ASAE 47(5), 1445-1451.

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

Munoz-Carpena, R., Vellidis, G., Shirmohammadi, A., Wallender, W., 2006. Evaluation of modeling tools for TMDL development and implementation. Transactions of the ASABE 49(4), 961-965.

NCDC (National Climatic Data Center). Hourly precipitation Data. Available at: http://www.ncdc.noaa.gov/oa/climate/stationlocator.html. Accessed in 2009.

Neary, V. S., Habib, E., Fleming, M., 2004. Hydrologic modeling with NEXRAD precipitation in middle Tennessee. Journal of hydrologic engineering 9(5), 339-349.

Nigro, J., Toll, D., Partington, E., Wenge, N. -M., Lee, S., Gutierrez-Magness A., Engman, T., Arsenault, K., 2010. NASA-Modified precipitation products to

improve USEPA nonpoint source water quality modeling for the Chesapeake Bay. Journal of Environmental Quality 39(4), 1388-1401.

NWSMC (National Weather Service Modernization Committee)/Commission on Engineering and Technical Systems/National Research Council, 1998. Future of the national weather service cooperative observer network. National Academy press, Washington, D.C., U.S.A..

Over, T. M., Murphy, E. A., Ortel, T. W., Ishii, A. L., 2007. Comparisons between NEXRAD radar and tipping-bucket gageg rainfall data: A case study for DuPage County, Illinois. In World Environmental and Water Resources Congress 2007: Restoring our natural habitat, Tampa, Florida, American Society of Civil Engineers, p. 6971.

Ramkellawan, J., Gharabaghi, B., Winter, J. G., 2009. Application of weather radar in estimation of bulk atmospheric deposition of total phosphorus over lake Simcoe. Canadian Water Resources Journal 34(1), 37-60.

Reed, S. M., Maidment, D. R., 1999. Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. Journal of Hydrologic Engineering 4(2), 174-182.

Ribarova, I., Ninov, P., Cooper, D., 2008. Modeling nutrient pollution during a first flood event using HSPF software: Iskar River case study, Bulgaria. Ecological Modeling 211, 241-246.

Rogalus III, M. J., Ogden, F. L., 2007. Comparison of GCIP and stage III radar-rainfall estimates over the Mississippi river basin for 1997. Journal of Hydrology 341, 177-185.

Rohli, R. V., Vega, A. V., 2007. Climatology. Jones &Bartlett Publishers, Canada.

Rutledge, A. T., 1998. Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records-update. Water-Resources Investigations Report 98-4148. U.S. Geological Survey.

Ryu, J. H., 2009. Application of HSPF to the distributed model intercomparison project: case study. Journal of Hydrologic Engineering 14(8), 847-857.

Schuurmans, J. M., Bierkens, M. F., 2007. Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. Hydro.Earth Syst. Sci. 11, 677-693.

Searcy, J. K., Hardison, H. H., 1966. Double-Mass Curve. Manual of hydrology: part 1, General surface water techniques. Geological Survey Water-Supply Paper 1541-B.

Seo, D. J., Breidenbach, J. P., 2002. Real-time correction of spatially non-uniform bias in radar rainfall data using rain gauge measurements. Journal of Hydrometeorology 3, 93-111.

Seo, D. J., Breidenbach, J. P., Fulton, R., Miler, D., 2000. Real-time Adjustment of range-dependent biases in WSR-88D rainfall estimates due to non-uniform vertical profile of reflectivity 1, 222-240.

Seo, D. J., Breidenbach, J. P., Johnson, E. R., 1999. Real-time estimation of mean field bias in radar rainfall data. Journal of Hydrology 223, 131-147.

Seo, D. J., Cajina, L., Corby, R., Howieson, T., 2009. Automatic state updating for operational streamflow forecasting via variational data assimilation. Journal of Hydrology 367(3-4), 255-275.

Seo, D. J., Seed, A., Delrieu, G., 2010. "Radar and multisensory rainfall estimation for hydrologic applications" in Rainfall: State of the Science, Testik, F. Y., and M. Gebremichael (Eds.). Geophysical Monograph Series 191, 79-104, American Geophysical Union, Washington, DC.

Sevruk, B., Nespor, V., 1998. Empirical and theoretical assessment of the wind induced error of rain measurement. Water Science and Technology 37(11), 171-178.

Sexton, A. M., Sadeghi, A. M., Zhang, X., Srinivasan, R., Shirmohammmadi, A., 2010. Using NEXRAD and rain gauge precipitation data for hydrologic calibration of SWAT in a northeastern watershed. Transactions of the ASABE 53(5), 1501-1510.

Sharif, H. O., Sparks, L., Hassan, A. A., Zeitler, J., Xie, H., 2010. Application of a distributed hydrologic model to the November 17, 2004, flood of Bull creek watershed, Austin, Texas. Journal of hydrologic engineering 15(8), 651-657.

Shenk, G., 2010. Model operations manual. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis MD, U.S.A.. Available at ftp://ftp.chesapeakebay.net/.

Shirmohammadi, A., Chaubey, I., Harmel, R. D., Bosch, D. D., Munoz-Carpena, R., Dharmasri, C., Sexton, A., Arabi, M., Wolf, M. L., Frankenberger, J., Graff, C., Sohrabi, T. M., 2006. Uncertainty in TMDL models. Transactions of the ASABE 49(4), 1033-1049.

Shirmohammadi, A., Chu, T. W., Montas, H. J., 2008. Modeling at catchment scale and associated uncertainties. Boreal Env Res. 13, 185-193.

Sieck, L. C., Burges, S. J., Steiner, M., 2007. Challenges in obtaining reliable measurements of point rainfall. Water Resources Research 43, W01420.

Silvestro, F., Rebora, N., Ferraris, L., 2009. An algorithm for real-time Rainfall rate estimation by using polarimetric radar: RIME. Journal of Hydrometeorology 10(1), 227-240.

Skinner, C., Bloetscher, F., Pathak, C. S., 2009. Comparison of NEXRAD and rain gauge precipitation measurements in south Florida. Journal of Hydrologic Engineering 14(3), 248-260.

Smith, M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J. J., Moreda, F., 2004. Runoff response to spatial variability in precipitation: an analysis of observed data. Journal of Hydrology 298, 267-286.

Sokol, Z., 2003. Utilization of regression models for rainfall estimates using radar-derived rainfall data and rain gauge data. Journal of Hydrology 278, 144-152.

Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V., Darden, R. W., 1999. Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. Water Resources Research 35(8), 2487-2503.

Stisen, S., Sandholt, I., 2010. Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modeling. Hydrological Rocesses. 24: 879-891.

Sun, X., Mein, R. G., Keenan, T. D., Elliott, J. F., 2000. Flood estimation using radar and raingauge data. Journal of Hydrology 239, 4-18.

Tadesse, A., Anagnostou, E. N., 2005. A statistical approach to ground radar-rainfall estimation. Journal of Atmospheric and Oceanic Technology 22(11), 1720-1732.

Tiruneh, Nebiyu, 2007. Basin-wide annual basedflow analysis for the fractured bedrock unit in the Potomac River Basin. ICPRB (Interstate Commission on the Potomac River Basin) Report No. 07-6.

Tobin, K. J., Bennett, M. E., 2009. Using SWAT to model streamflow in two river basins with ground and satellite precipitation data. Journal of the American Water Resources Association 45(1), 253-271.

Troutmann, B. M., 1982. An analysis of input errors in precipitation-runoff models using regression with errors in the independent variables. Water Resources Research, 18(4), 947–964.

Tuppad, P., Douglas-Mankin, K. R., Koelliker, J. K., Hutchinson, J. M. S., 2010. SWAT discharge response to spatial rainfall variability in a Kansas watershed. Transactions of the ASABE 53(1), 65-74.

USDC (U.S. Department of Commerce), NOAA, National Weather Service-Cooperative Weather Observer Program/Observing Services Division, 2007. Fischer-Porter Upgrade FPU Operational manual. Available at http://www.nws.noaa.gov/ops2/Surface/documents/FPU_OperationsMan05Feb2007.pdf.

USDC (U.S. Department of Commerce), NOAA, National Weather Service/Office of Operational Systems, Field Systems Operations Center/Observing Systems Branch, 2003. ASOS product improvement implementation plan for all weather precipitation accumulation gauge. Available at http://www.nws.noaa.gov/ops2/Surface/documents/awpagimplan.pdf.

USEPA (U.S.EPA), 2000. The quality of our nation's waters: a summary of the national water quality inventory: 1998. Report to Congress. EPA 841-S-00-001. Washington, D.C: Office of water.

USEPA (U.S.EPA), 2010. BASINS 4 lectures, data sets, and exercises (August 2010). Available at http://water.epa.gov/scitech/datait/models/basins/training.cfm. Accessed in 2010.

USEPA (U.S.EPA), 2010. Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002-CBP/TRS-303-10. Chesapeake Bay Program Office, Annapolis MD, U.S.A..

USGS (U.S. Geological Survey). USGS Surface-water data for the nation. Available at http://waterdata.usgs.gov/nwis/sw Accessed in 2009.

Villarini, G., Krajewski, W. F., Ciach, G. J., 2009. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. Water Resources Research 45, W01404.

Vivoni, E. R., Entekhabi, D., Bras, R. L., Ivanov, V. Y., Van Horne, M. P., Grassotti, C., Hoffman, R. N., 2006. Extending the predictablility of hydrolmeteorological flood events using radar rainfall nowcasting. Journal of Hydrometeorology 7(4), 660-677.

Vrugt, J. A., Diks, C. G. H, Gupta, H. V., Bouten, W., Verstraten, J. M., 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. Water Resources Research 41, W01017.

Wang, J., Fisher, B. L., Wolff, D. B., 2008. Estimating Rain Rates from Tipping-Bucket Rain Gauge Measurements. American meteorological Society. 25, 43-52.

Wang, X., Xie, H., Sharif H., Zeitler, J., 2008. Validating NEXRAD MPE and Stage III precipitation products for uniform rainfall on the Upper Guadalupe River basin of the Texas Hill country. Journal of Hydrology 348, 73-86.

Wang, Y., 2011. Calibrating Shenandoah watershed SWAT model using a nonlinear groundwater algorithm. M.S. dissertation. University of Maryland, College Park. MD, U.S.A.. Retrieved in 2011.

Watkins, D. W., Li, H., Cowden, J. R., 2007. Adjustment of radar based precipitation estimates for Great Lakes hydrologic modeling. Journal of Hydrologic Engineering 12(3), 298-305.

Weisstein, E. W., 2011. "Partial Sum" from Mathworld—A Wolfram Web Resource. http://mathworld.wolfram.com/PartialSum.html Accessed in 2011.

Westcott, N. E., Knapp, H. V., Hilberg, S. D., 2007. Comparison of gafe of multi-sensor precipitation estimates over a range of spatial and temporal scales in the Midwestern United States. Journal of Hydrology 351, 1-12.

Xie, H. J., Zhou, X. B., Hendrickx, J. M. H., Vivoni, E. R., Guan, H. D., Tian, Y. Q., Small, E. E., 2006. Evaluation of NEXRAD Stage III precipitation data over a semiarid region. Journal of the American Water Resources 42(1), 237-256.

Xie, H. J., Zhou, X. B., Vivoni, E. R, Hendricks, J.M.H, Small, E.E., 2005. GIS-based NEXRAD Stage III precipitation database: automated approaches fro data processing and visualization. Computers & Geosciences 31(1), 65-76.

Young, C.B., Bradley, A. A., Krajewski, W. F., Kruger, A., Morrissey, M. L., 2000. Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. American Meteorological Society 1, 241-248.

Young, C. B., Brunsell, N. A., 2008. Evaluating NEXRAD estimates for the Missouri river basin: analysis using daily raingauge data. Journal of Hydrologic Engineering 13(7), 549-553.

Yilmaz, K. K., Hogue, T., Hsu, K., Sorooshian, S., Gupta, H. V., Wagener, T., 2005. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. American Meteological Society 6, 497-517.

Zhang, Y., Adams, T., Bonta, J. V., 2007. Subpixel-scale rainfall variability and the effects on separation of radar and gauge rainfall errors. Journal of Hydrometeorology 6, 1348-1363.