

**WATER-DRIVEN EROSION PREDICTION TECHNOLOGY FOR A
MORE COMPLICATED REALITY**

by

Josept David Revuelta Acosta

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy



School of Agricultural and Biological Engineering

West Lafayette, Indiana

May 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Bernard A. Engel, Co-Chair

Department of Agricultural and Biological Engineering

Dr. Dennis C. Flanagan, Co-Chair

Department of Agricultural and Biological Engineering

Dr. Margaret W. Gitau

Department of Agricultural and Biological Engineering

Dr. Laura C. Bowling

Department of Agronomy

Approved by:

Dr. Nathan S. Mosier

I would like to dedicate my work to God, my family, friends, and myself for their unconditional support.

ACKNOWLEDGMENTS

I want to express my gratitude to God for being with me this whole time, for loving me, for protecting me, and for helping to complete this work.

To my dear family, Adela, and Johnny. Thank you for your love, support, and words of wisdom during difficult times. To my friends, Sandra, Maria, Johann, Momoh, Erin, Anurag, and so many others who were with me in this adventure. An especial dedication to my grandfather, “Pipo”, who taught me lots of things, love you. I see you in heaven.

To my advisors, Dr. Engel and Dr. Flanagan, I thank you for wisdom, knowledge, time. Thank you for helping me to become independent and completing my studies. To Dr. Suresh Rao and Dr. Laura Bowling for the most exciting classes that helped me to generate ideas for my research objectives. Thank you to my committee members, Dr. Bowling and Dr. Gitau, for their knowledge and contribution to my research.

The data used in this research was through a collaboration with Dr. Kevin King, USDA scientist, and Dr. Anurag Srivastava, USDA post-doc. Thank you for sharing part of your work and contributing to my research. To the Consejo Nacional de Ciencia y Tecnologia (CONACYT) and Dr. Engel, thank you for the funds provided to complete my Ph.D.

Lastly, special gratitude to Dr. Eusebio Ventura, “El Apa,” for the motivation and the assistance to start this journey, you are always a great friend to me.

TABLE OF CONTENTS

LIST OF TABLES	8
LIST OF FIGURES	9
ABSTRACT	11
1. INTRODUCTION	13
1.1 Problem statement	13
1.2 Objectives	18
1.3 Dissertation organization	19
1.4 References	19
2. DEVELOPMENT OF A STOCHASTIC STORM GENERATOR USING HIGH- RESOLUTION PRECIPITATION RECORDS	30
Abstract	30
2.1 Introduction	30
2.2 Objectives and scope	32
2.3 Procedure	33
2.3.1 General approach	33
2.3.2 Data used	34
2.4 Characterization of rainstorm events	35
2.4.1 Definition of a rainstorm event	35
2.4.2 Characterization of storm arrival	36
2.4.3 Characterization of time between storms, storm duration, and storm depth	38
2.5 Stochastic simulation of rainstorm sequences	41
2.5.1 Monte Carlo correlated non-normal multivariate simulation	41
2.5.2 Modeling storm pattern	44
2.6 Model performance	49
2.6.1 Analysis	49
2.6.2 Comparison of frequency distributions	50
2.6.3 Model prediction efficiency	53
2.6.4 Storm pattern performance	56
2.7 Conclusions	57

2.8	References	58
3.	IMPROVEMENT OF THE WATER EROSION PREDICTION PROJECT (WEPP) MODEL FOR QUANTIFYING FIELD SCALE SUBSURFACE DRAINAGE DISCHARGE.....	63
	Abstract	63
3.1	Introduction	63
3.2	Materials and Methods	66
3.2.1	Model description	66
3.2.2	Experimental sites.....	73
3.2.3	Model input and calibration/validation.....	74
3.2.4	Meteorological data	75
3.2.5	Soil hydraulic parameters	75
3.2.6	Crop/plant management.....	76
3.2.7	Performance evaluation measures	77
3.3	Results and discussion.....	78
3.3.1	Uncalibrated performance at ES(1) (WEPP v2012.8 vs WEPP v2012.9).....	78
3.3.2	Drainage at uncalibrated sites ES(2-4)	82
3.3.3	Drainage at calibrated/validated sites ES(2-4)	84
3.3.4	Water balance at the experimental sites.....	87
3.4	Concluding remarks.....	90
3.5	References	90
4.	FORMULATION OF A ROUTING MODEL FOR THE WATER EROSION PREDICTION PROJECT (WEPP) MODEL BASED ON CATCHMENT GEOMORPHOLOGY AND TRAVEL TIME THEORY.....	99
	Abstract	99
4.1	Introduction	99
4.2	Objectives	101
4.3	Theoretical background	101
4.3.1	Mass transport theory	101
4.3.1	Travel time distribution formulations.....	103
4.4	Methods	107
4.4.1	General model approach	107

4.4.2	Description of the area of study and hydrologic data	110
4.4.3	Runoff generation	113
4.4.4	Calibration and validation.....	113
4.4.5	Performance evaluation	114
4.5	Results and discussion.....	115
4.5.1	Catchment geomorphology.....	115
4.5.2	Surface travel time distribution	116
4.5.3	Subsurface travel time distribution.....	119
4.5.4	Catchment total response	124
4.6	Concluding remarks.....	129
4.7	References	130
5.	CONCLUSIONS	139
5.1	Research overview.....	139
5.2	Major research findings.....	140
5.3	Limitations of the current study and future research.....	141
VITA.....		143
PUBLICATION.....		144

LIST OF TABLES

Table 2.1 Measured rainfall characteristics at Charlotte-Mecklenburg, North Carolina for USGS-CHM.	37
Table 2.2. Number of precipitation events by month.	38
Table 2.3. Comparison between observed and simulated time between storms at USGS-CHM at Charlotte-Mecklenburg.	50
Table 2.4. Comparison between observed and simulated storm duration and storm depth at USGS-CHM at Charlotte-Mecklenburg, North Carolina.	52
Table 2.5. Coefficient of determination (R^2) between observed and simulated storm patterns in Figure 2.11.	57
Table 3.1. Soil physical properties and hydraulic properties at the experimental sites from USDA-NRCS Soil Survey and ROSETTA model.	76
Table 3.2 Subsurface drainage system and crop management data.....	77
Table 3.3 Model accuracy evaluation for daily drainage flow estimates at test site ES(1) and calibrated/validated sites ES(2-4).	80
Table 3.4 Annual water balance at the sites ES(2-4).	87
Table 3.5 Annual WEPP v2012.9 predicted and observed subsurface drainage amounts for the sites ES(1-4).	89
Table 4.1 WEPP-CTM calibration parameters	114
Table 4.2 Performance evaluation of daily flow discharges for Cedar River Watershed for both calibration and validation periods.	126
Table 4.3 Annual water balance from the WEPP-CTM model for 2009-2011 water years.	127
Table 4.4 Comparison of model performance (NSE) of large scale hydrological models and WEPP model for flow discharge estimation.....	128

LIST OF FIGURES

Figure 2.1 Empirical distributions of monthly time between storms and the corresponding theoretical distributions calculated in the exponential method for USGS-CHM precipitation data (2008-2017) at Charlotte-Mecklenburg, North Carolina.	36
Figure 2.2 Monthly storm depth (red squares) and storm duration (black circles) empirical distributions for USGS-CHM at Charlotte-Mecklenburg, North Carolina. The distributions were truncated to the depth resolution of the data (1 mm).	40
Figure 2.3 Flowchart of a Monte Carlo correlated multivariate simulation.	44
Figure 2.4 Significant clusters for USGS-CHM at Charlotte-Mecklenburg, North Carolina.	47
Figure 2.5 Intrinsic variability of storm intensities within a storm pattern. 86 storm patterns (dashed-lines) in an M-type representative pattern (solid-line).	48
Figure 2.6 Comparison between simulated and observed time between storms at USGS-CHM at Charlotte-Mecklenburg, North Carolina, for the month with the longest average time between storms (May).	51
Figure 2.7 Comparison between simulated and observed storm duration at USGS-CHM at Charlotte-Mecklenburg for the month with longest storm duration (May).	53
Figure 2.8 Comparison between simulated and observed storm depths at USGS-CHM at Charlotte-Mecklenburg for the month with greatest storm depths (May).	53
Figure 2.9 Model predictions for the three storm variables based on number of years generated: a) time between storms; b) storm duration; c) storm depth.	54
Figure 2.10 Correlation estimations between storm duration and storm depth based on the number of years generated.	55
Figure 2.11 Average generated storms patterns (SD/SM/SC/SR type) vs. representative patterns (D/M/C/R type).	56
Figure 3.1 Input relationships for the WEPP model (v2012.9)	69
Figure 3.2 Experimental site locations. NW Station: Northwest station managed by OARDC; ES(1): Uncalibrated validation site; ES(2-4): Calibrated/Validated sites.	74
Figure 3.3 Model output at ES(1). (top) observed precipitation; (bottom) observed drainage (circles), WEPP v2012.8 simulated drainage (red solid-line), and WEPP v2012.9 daily drainage (black solid-line).	79
Figure 3.4 Model output at ES(1): (top) observed precipitation; (bottom) WEPP v2012.8 simulated water table depth (black solid-line), and WEPP v2012.9 simulated water table depth (grey solid-line).	80

Figure 3.5 Predicted vs. observed daily subsurface drainage water depth (cm) using WEPP model v2012.8 (hollow circles) and WEPP model v2012.9 (solid circles) for uncalibrated performance at ES(1).	82
Figure 3.6 Predicted vs. observed daily subsurface drainage depths (cm) using WEPP model v2012.9 for validation site ES(1) and calibrated (C-M)/validated(V-M) sites ES(2-4).	83
Figure 3.7 Daily drainage at experimental sites - ES(2-4): (Top) Simulated and observed drainage at ES(2); (Middle) Simulated and observed drainage at ES(3); (Bottom) Simulated and observed drainage at ES(4).....	86
Figure 3.8 Monthly drainage depths for sites ES(2-4): WEPP v2012.9 simulated and observed drainage.....	88
Figure 4.1 Hydrologic response as a combination of: a) surface response function and b) subsurface response function.	108
Figure 4.2 Subsurface travel time distribution and its derivation.....	109
Figure 4.3 Location map of the Upper Cedar River Watershed in western Washington, US. ...	112
Figure 4.4 Flow distance characterization: a) Distance to the closest stream; b) Distance to the outlet.	116
Figure 4.5 Area-width function of the catchment in the Upper Cedar River Watershed: a) distance to the outlet distribution, b) normalized width-area.....	117
Figure 4.6 Surface travel time distribution for different distance amplification factors $r = ucuh$	118
Figure 4.7 Subsurface travel time distributions for different mean travel times. Flow velocity, $u = 1.0$ m/s; and hydrodynamic dispersion, $D = 1000$ m ² /s.	120
Figure 4.8 Travel time distribution under different flow velocities in the channel.	122
Figure 4.9 Travel time distribution under different hydrodynamic dispersion coefficients in the channel.	123
Figure 4.10 Temporal variability of travel time distribution for subsurface and channel conditions in the Cedar River Watershed.	124
Figure 4.11 Daily flow discharge at the outlet of the simulated subcatchment of the Upper Cedar River Watershed.....	125
Figure 4.12 Monthly flow discharge at the outlet of the modeled subcatchment of the Upper Cedar River Watershed.....	126

ABSTRACT

Hydrological modeling has been a valuable tool to understand the processes governing water distribution, quantity, and quality of the planet Earth. Through models, one has been able to grasp processes such as runoff, soil moisture, soil erosion, subsurface drainage, plant growth, evapotranspiration, and effects of land use changes on hydrology at field and watershed scales. The number and diversity of water-related challenges are vast and expected to increase. As a result, current models need to be under continuous modifications to extend their application to more complex processes. Several models have been extensively developed in recent years. These models include the Soil and Water Assessment Tool (SWAT), Variable Infiltration Capacity (VIC) model, MIKE-SHE, and the Water Erosion Prediction Project (WEPP) model. The latter, although it is a well-validated model at field scales, the WEPP watershed model has been limited to small catchments, and almost no research has been introduced regarding water quality issues (only one study).

In this research, three objectives were proposed to improve the WEPP model in three areas where either the model has not been applied, or modifications can be performed to improve algorithms of the processes within the model (e.g. erosion, runoff, drainage). The enhancements impact the WEPP model by improving the current stochastic weather generation, extending its applicability to subsurface drainage estimation, and formulating a new routing model that allows future incorporation of transport of reactive solutes.

The first contribution was development of a stochastic storm generator based on 5-min time resolution and correlated non-normal Monte Carlo-based numerical simulation. The model considered the correlated and non-normal rainstorm characteristics such as time between storms, duration, and amount of precipitation, as well as the storm intensity structure. The model was tested using precipitation data from a randomly selected 5-min weather station in North Carolina. Results showed that the proposed storm generator captured the essential statistical features of rainstorms and their intensity patterns, preserving the first four moments of monthly storm events, good annual extreme event correspondence, and the correlation structure within each storm. Since the proposed model depends on statistical properties at a site, this may allow the use of synthetic storms in ungauged locations provided relevant information from a regional analysis is available.

A second development included the testing, improvement, and validation of the WEPP model to simulate subsurface flow discharges. The proposed model included the modification of the current subsurface drainage algorithm (Hooghoudt-based expression) and the WEPP model percolation routine. The modified WEPP model was tested and validated on an extensive dataset collected at four experimental sites managed by USDA-ARS within the Lake Erie Watershed. Predicted subsurface discharges show Nash-Sutcliffe Efficiency (NSE) values ranging from 0.50 to 0.70, and percent bias ranging from -30% to +15% at daily and monthly resolutions. Evidence suggests the WEPP model can be used to produce reliable estimates of subsurface flow with minimum calibration.

The last objective presented the theoretical framework for a new hillslope and channel-routing model for the Water Erosion Prediction Project (WEPP) model. The routing model (WEPP-CMT) is based on catchment geomorphology and mass transport theory for flow and transport of reactive solutes. The WEPP-CMT uses the unique functionality of WEPP to simulate hillslope responses under diverse land use and management conditions and a Lagrangian description of the carrier hydrologic runoff at hillslope and channel domains. An example of the model functionality was tested in a sub-catchment of the Upper Cedar River Watershed in the U.S. Pacific Northwest. Results showed that the proposed model provides an acceptable representation of flow at the outlet of the study catchment. Model efficiencies and percent bias for the calibration period and the validation period were $NSE = 0.55$ and 0.65 , and $PBIAS = -2.8\%$ and 2.1% , respectively. The WEPP-CMT provides a suitable foundation for the transport of reactive solutes (e.g. nitrates) at basin scales.

1. INTRODUCTION

1.1 Problem statement

Hydrological modeling is critical to help understand several ongoing issues including climate change, water quality, soil erosion, and land use change effects at different scales (Rosenzweig et al., 2013; Martinez-Martinez et al., 2014; Negm et al., 2016; Wu et al., 2017; Anache et al., 2018). The number and diversity of water-related challenges are large and expected to increase in the future. These issues are found to often have negative impacts on society. For instance, three examples can be mentioned: 1) soil deterioration and lower water quality due to erosion and surface runoff produces lands that no longer can be cultivated and must be abandoned, which involves economical and food security problems (Pimentel, 2000; Rhodes, 2014; Song et al., 2019; Zhang and Zhuang, 2019); 2) climate change significantly impacts global amount precipitation and its distribution. Increase of rainfall in the tropics and higher latitudes and decrease in rainfall within the already semi-arid to arid mid-latitudes and the interior of large continents (Dore, 2005; Turrall et al., 2011; Kunkel et al., 2013; Martel et al., 2018) will bring limited crop productivity, water scarcity in dry areas, and more recurrent floods and droughts along with human health consequences; and 3) The amount of pollutants generated by human activities such as destruction of wetlands, grasslands, and forest in favor of urban landscape and agricultural practices have increased hypoxia in estuaries and coastal waters. As a result, long term changes in ecology, weakening of species, economic losses, and freshwater availability are predicted (Levinton et al., 2011; Almeida, 2016; Jenny et al., 2016; Jabbari et al., 2019).

In response to these challenges, environmental models have been developed to understand, analyze, and propose solutions for sustainable water management. Conceptualization of a hydrologic system leads to the formulation of a variety of ways to classify rainfall-runoff models from deterministic to stochastic models, from physically-based models to empirical models, and the most distinctive, from lumped to distributed models. In terms of functionality, physically-based distributed models are often used in research to gain better understanding of the hydrological phenomena happening in a catchment and how changes in the catchment may affect the hydrology. Some of the physically-based distributed models found in the literature include the European Hydrological System (SHE) (Abbott et al., 1986a, 1986b), MIKE SHE (Refsgaard and Storm,

1995), TIN-based Real-time Integrated Basin Simulator (tRIBS) (Ivanov et al., 2004), HydroGeoSphere (HGS) (Therrien et al., 2010; Brunner and Simmons, 2012), and the Water Erosion Prediction Project (WEPP) model (Flanagan et al., 2012). Typically, the data and computation requirements for these models are enormous and demanding. To overcome this limitation, semi-distributed models partially allow catchment heterogeneities to vary in space by dividing the basin into several sub-basins, which in turn are treated as a single unit. Among these models, one could find the Variable Infiltration Capacity (VIC) model (Liang et al., 1994), TOPMODEL (Beven and Kirby 1984; Beven et al., 2001), and the Soil and Water Assessment Tool (SWAT) (Arnold et al., 2012).

Acknowledging the potential of physically-based distributed models to simulate hydrological processes at field and/or watershed scales (Daniel et al., 2011), great efforts have been made to improve representation of hydrologic processes in this type of model. Research directed to improve semi-distributed and distributed process-based models include work on: SWAT (Moriiasi et al., 2012; Bauwe et al., 2015; Zhang et al., 2016; Fuka et al., 2016; Bauwe et al., 2017; Raju and Nandagiri, 2017; Mittelstet et al., 2017), The Agricultural Policy/Environmental eXtender (APEX) model (Wang et al., 2008; Jeong et al., 2014; Ford et al., 2016; Pi et al., 2017; Hong et al., 2018), WEPP (Conroy, 2005; Mao et al., 2006; Wang et al., 2010; Frankenberger et al., 2011; Singh et al., 2011; Flanagan et al., 2012; Srivastava et al., 2017; Wang et al., 2018), and MIKE-SHE (Cui et al., 2005; Hansen et al., 2013; Zhang et al., 2015; Ma et al., 2016; De Schepper et al., 2017; Botero-Acosta et al., 2017; Sonnenborg et al., 2017). Among the available models cited above, one lagging in terms of water quantity and quality research at both the hillslope and watershed scales is the WEPP model.

The WEPP model is a physically-based soil erosion prediction technology. It integrates hydrology, plant science, hydraulics and erosion mechanics to predict erosion at hillslope and watershed scales. The WEPP model can simulate and assess a variety of land uses, climates and hydrologic conditions. The main hydrologic processes in the WEPP model include evapotranspiration, snow accumulation and melt, surface runoff, infiltration, subsurface flow, and deep percolation of soil water (Flanagan et al., 2012). The model was developed beginning in 1985 in an effort undertaken by the United States Department of Agriculture – Agricultural Research Service (USDA-ARS) to replace the Universal Soil Loss Equation (USLE; Wischmeier and Smith, 1978), which was the standard tool for soil loss estimation used by the USDA Soil Conservation

Service (SCS). Although WEPP has been well validated at the hillslope scale since its creation, there are still factors that contribute to the previously mentioned development lag in comparison to other hydrological models. The research described in this thesis deals with these factors and proposes three improvements to the WEPP model to promote its development and applicability. The enhancements impact the model by improving the stochastic weather generation, extending its applicability to improved subsurface drainage estimation, and formulating a new routing model that allows future incorporation of transport of reactive solutes. The following paragraphs describe the justification of these improvements and state the research objectives.

First, climate temporal resolutions, such as daily precipitation often used to simulate hydrologic processes and soil erosion in the WEPP model (Fashi et al., 2015; Brooks et al., 2016; Shen et al., 2016; Kinnell, 2017; Srivastava et al., 2017; Zemke et al., 2019), are not entirely appropriate since water erosion estimates are more dominated by rainfall intensities than total rainfall (Archer and Fowler, 2018; Dunkerley, 2019). Rainfall intensities occur by nature in small time increments and storm durations can be shorter or longer than a day. Additionally, processes such as infiltration and chemical transport, as well as assessment of Best Management Practices (BMPs), often require time steps as small as minutes to realistically capture the flow dynamics and sediment loads coming from upland areas. Although the WEPP model can simulate hydrology and erosion using sub-daily values of precipitation, most research has been conducted using daily amounts of precipitation, which is often more available for longer time periods than in sub-daily form or it is easier to obtain. In an attempt to overcome this limitation, the WEPP model integrates a stochastic weather generator, CLIGEN (CLimate GENerator; Nicks et al., 1995). CLIGEN generates one storm per day with individual parameters including storm depth, time to peak intensity, peak intensity, and storm duration. However, this still limits the applicability of the WEPP model to long-term evaluation of erosion estimates and agricultural practices where a more extensive description of the storm intensities is required.

As part of the first WEPP model enhancement, two facts are identified. First, in the United States, 3200+ meteorological stations record precipitation on a daily basis (<https://www.ncdc.noaa.gov/ghcnd-data-access>), 2000+ stations measure precipitation at 15-min time increments (<https://www.ncdc.noaa.gov/cdo-web/datasets>), and 1500+ stations measure 1-min precipitation (<https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/automated-surface-observing-system-asos>). One can see that daily precipitation is still

more available than sub-daily measurements. However, there are a significant number of stations providing high resolution, continuous measurement of precipitation, which can provide a more realistic view of storm events (Bonta, 2004). Although daily precipitation is often considered as continuous, it just reflects the total volume of precipitation in a day with no intensity associated with the data, while sub-daily measurements consider the entire occurrence of events and the associated characteristics of each rainfall event. Second, future estimates of the hydrologic processes and BMP design using environmental models often require the use of long-term measurements of rainfall (i.e. 50-100 yrs) or weather generators that reflect the temporal characteristics of the storm under stationarity or nonstationarity assumptions. As a result, many storm generators are available for generating synthetic weather. Nonetheless, the data availability in the past five decades limited storm generators to daily values of precipitation and storm characteristics generated by rainfall disaggregation methods (Vu et al., 2018). Although storm generators have been tested in many studies (Pui et al., 2012; Muller et al., 2018; Kossieris et al., 2018), it is still far from reality, especially in cases where storms last more than 24 hours or multiple storms happen on the same day. These issues indicate that there is an opportunity to enhance the prediction of hydrological models through improving the resolution of their inputs (i.e. precipitation) by taking advantage of the finer but still short available precipitation data and the statistical numerical tools such as the concept of stochastic storm generator models.

Second, in the Midwest US and other regions, subsurface drainage systems have drastically influenced the development of agriculture by allowing crop growth in otherwise unsuitable areas. Subsurface drains influence hydrology by substantially decreasing surface runoff, reducing periods of surface ponding and lowering the water table depth. However, one of the major concerns with the use of subsurface drainage systems is the adverse environmental effect as they modify soil water dynamics. In terms of water quality, these systems change volume and timing of sediment, nitrogen, and phosphorus losses, decreasing the residence time of the water within the soil, allowing direct passage of water and contaminants off-site. In drainage systems without surface inlets, subsurface drainage decreases sediment losses and potentially sediment-bound nitrates and phosphorus (David et al., 2010, King et al., 2015). Nonetheless, these systems enhance losses of soluble nitrates and phosphorus out of the tiles as subsurface flow moves into ditches, streams or lakes. In systems with integrated tile riser inlet pipes, sediments and sediment-bound nitrogen and phosphorus also travel in the subsurface flow and discharge directly into ditches and

downstream rivers and lakes. Large-scale studies show that the Upper Mississippi River Basin is the most significant source of nitrate flux contributing to the overall load of nitrates to the Gulf of Mexico, which contributes to the hypoxic dead zone (USEPA, 2007). Of this total contribution, the most prominent contributors of nitrate to the Mississippi River Basin are southwest Minnesota to across Iowa, Illinois, Indiana, and Ohio (David et al., 2010). Many natural resource models have been tested to predict surface and subsurface flow at different scales. These include DRAINMOD (Skaggs et al., 2012), DRENAFEM (Castanheira and Santos, 2009), GLEAMS (Knisel and Douglas-Mankin, 2012), WEPP (Flanagan et al., 2012), ADAPT (Gowda et al., 2012), and SWAT (Arnold et al., 2012). Of these, DRAINMOD and SWAT are widely used in North America to simulate outflows from subsurface drainage systems. In contrast, the WEPP model has been used in only a few studies on modeling subsurface drainage (Savabi, 1993; Maalim et al., 2013), which opens a window for testing and improving its capabilities as a model that integrates not only hydrology, soil loss estimation and plant growth components but also subsurface drainage, a practice heavily used in the midwestern U.S. The second contribution of this research was to test, enhance, and validate the WEPP model's DRAINMOD-based routines used for artificial subsurface flow prediction. This work will also extend the possibility of WEPP to simulate realistic model responses for subsurface drainage systems and provide a clearer understanding of the effect of these systems on water dynamics, water quality and soil erosion.

Lastly, a significant factor for the lag of application of the WEPP model at watershed scales is that the model was developed to function on hillslope and small catchments (< 260 ha) where residence times in channels are small in comparison with those in hillslopes and its use was mainly for soil erosion estimation at those scales (Flanagan et al., 2012). Whereas this limitation in regard to channel hydrology was addressed by the incorporation of a conventional routing model (i.e. Muskingum-Cunge) (Wang et al., 2010), only a few studies have used and validated it at relatively large catchment scales (Afshar and Hassanzadeh, 2017; Srivastava et al., 2017; Srivastava et al., 2019). In this research it is believed that “the best model” is not only the one that gives the closest results to reality but also a model where each improvement allows subsequent enhancements. As a result of the problem and model definition explained above, the last contribution of this research was the formulation of a new routing model dealing with both flow and transport, which extends the use of the WEPP model to larger scales and provides the groundwork for future transport of reactive solute applications. In this context, general transport models (i.e. advection-diffusion

models), especially those that follow a stochastic framework have been shown to be appropriate for large-scale applications (Rodriguez-Iturbe and Rinaldo, 2001; Rinaldo et al., 2005), especially based on travel time distributions (Rinaldo and Marani, 1987; Rinaldo and Rodriguez-Iturbe, 1996; Gupta and Cvetkovic, 2002; Botter and Rinaldo, 2003; Botter et al., 2005; Rinaldo et al., 2005; Botter et al., 2010; Rinaldo et al., 2011). In this research, a general mass transport model with a stochastic framework based on travel time distributions was developed for routing the hydrologic response. The proposed model integrates the element of general transport theory with the unique functionality of a process-based model such as WEPP through an objective manipulation of hydrologic, geomorphologic and soil properties, and land use data, and provides suitable groundwork for transport of reactive solutes using the WEPP model, which is currently limited to only one study (Wang et al., 2018).

1.2 Objectives

The objectives of this dissertation are to 1) develop a stochastic storm generator using high-resolution precipitation records to be used not only in the WEPP model but with other hydrological models; 2) improve the subsurface drainage flow algorithm in the WEPP hillslope model; and 3) formulate a new routing method for the WEPP model based on catchment geomorphology and travel time theory.

The sub-objectives were:

1. Introduce the concepts and the characterization of the variables defining precipitation events, present the mathematical and statistical methods to reproduce series of correlated storm variables and storm patterns using a machine learning tool, and conduct a preliminary test and analysis of the stochastic storm generator performance using 15-min precipitation data.
2. Test the subsurface drainage routine found in the current version of the WEPP model (v2012.8), present the basis of the new algorithm to enhance the subsurface drainage prediction of the model, and calibrate and validate the model using an extensive dataset collected at four experimental sites managed by the USDA-ARS within the Lake Erie Watershed.
3. Present the formulation of a new general transport model for routing of hydrologic response based on geomorphologic properties and travel time distributions at the basin scale, and

demonstrate a practical application of the WEPP model and the proposed mass transport model to predict the hydrologic response in a mountainous basin in western Washington state, U.S.

1.3 Dissertation organization

This dissertation contains five chapters. Chapter one provides an introduction about the importance of hydrological modeling and the current state of the WEPP model and other models as well as the motivation, research problems, gaps found in the model, and the general and specific objectives of this study. Chapter two covers the development of a stochastic storm generation approach using short time increment precipitation and its applicability. Chapter three describes the testing and improvement of the subsurface drainage routine in the current version of the WEPP model. The model was evaluated, calibrated, and validated using experimental data collected in four fields within the western Lake Erie basin watershed. Chapter four covers the formulation of a new hydrologic routing model for WEPP based on catchment geomorphology and travel time distributions. The general transport model theory is presented and performance of the model in a mountainous catchment is assessed. Lastly, the final chapter is structured to present the overall summary and conclusions of this research, as well as significant findings, uncertainties, and recommendations for further research.

1.4 References

- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986a). An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87(1–2), 45–59. [https://doi.org/10.1016/0022-1694\(86\)90114-9](https://doi.org/10.1016/0022-1694(86)90114-9)
- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986b). An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. *Journal of Hydrology*, 87(1–2), 61–77. [https://doi.org/10.1016/0022-1694\(86\)90115-0](https://doi.org/10.1016/0022-1694(86)90115-0)

- Afshar, A. A., & Hassanzadeh, Y. (2017). Determination of monthly hydrological erosion severity and runoff in Torogh Dam watershed basin using SWAT and WEPP models. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 41(2), 221–228.
- Almeida, L. Z. (2016). *Effects of Lake Erie hypoxia on fish habitat quality and yellow perch behavior and physiology*. M.S. Thesis, Purdue University. Open Access Theses. 741. https://docs.lib.purdue.edu/open_access_theses/741
- Anache, J. A. A., Flanagan, D. C., Srivastava, A., & Wendland, E. C. (2018). Land use and climate change impacts on runoff and soil erosion at the hillslope scale in the Brazilian Cerrado. *Science of the Total Environment*, 622–623, 140–151. <https://doi.org/10.1016/j.scitotenv.2017.11.257>
- Archer, D. R., & Fowler, H. J. (2018). Characterising flash flood response to intense rainfall and impacts using historical information and gauged data in Britain. *Journal of Flood Risk Management*, 11, S121–S133. <https://doi.org/10.1111/jfr3.12187>
- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., van Griensven, A., Van Liew, M. W., Kannan, N., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1491–1508.
- Bauwe, A., Kahle, P., & Lennartz, B. (2016). Hydrologic evaluation of the curve number and Green and Ampt infiltration methods by applying Hooghoudt and Kirkham tile drain equations using SWAT. *Journal of Hydrology*, 537, 311–321. <https://doi.org/10.1016/j.jhydrol.2016.03.054>
- Bauwe, A., Tiedemann, S., Kahle, P., & Lennartz, B. (2017). Does the temporal resolution of precipitation input influence the simulated hydrological components employing the SWAT model? *Journal of the American Water Resources Association*, 53(5), 997–1007. <https://doi.org/10.1111/1752-1688.12560>
- Beven, K., & Freer, J. (2001). A dynamic TOPMODEL. *Hydrological Processes*, 15(10), 1993–2011. <https://doi.org/10.1002/hyp.252>
- Beven, K. J., Kirkby, M. J., Schofield, N., & Tagg, A. F. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments. *Journal of Hydrology*, 69, 119–143. [https://doi.org/10.1016/0022-1694\(84\)90159-8](https://doi.org/10.1016/0022-1694(84)90159-8)

- Bonta, J. V. (2004). Stochastic simulation of storm occurrence, depth, duration, and within-storm intensities. *Transactions of the ASAE*, 47(5), 1573–1584.
- Botero-Acosta, A., Chu, M. L., Guzman, J. A., Starks, P. J., & Moriasi, D. N. (2017). Riparian erosion vulnerability model based on environmental features. *Journal of Environmental Management*, 203, 592–602. <https://doi.org/10.1016/j.jenvman.2017.02.045>
- Botter, G., Bertuzzo, E., Bellin, A., & Rinaldo, A. (2005). On the Lagrangian formulations of reactive solute transport in the hydrologic response. *Water Resources Research*, 41(4).
- Botter, G., Bertuzzo, E., & Rinaldo, A. (2010). Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox. *Water Resources Research*, 46(3), 1–18. <https://doi.org/10.1029/2009WR008371>
- Botter, G., & Rinaldo, A. (2003). Scale effect on geomorphologic and kinematic dispersion. *Water Resources Research*, 39(10), 1286. <https://doi.org/10.1029/2003WR002154>.
- Brooks, E. S., Dobre, M., Elliot, W. J., Wu, J. Q., & Boll, J. (2016). Watershed-scale evaluation of the Water Erosion Prediction Project (WEPP) model in the Lake Tahoe basin. *Journal of Hydrology*, 533, 389–402. <https://doi.org/10.1016/j.jhydrol.2015.12.004>
- Brown, L. R. (2016). World population growth, soil erosion, and food security. *American Association for the Advancement of Science*, 214(4524), 995–1002. Retrieved from <http://www.jstor.org/stable/1686685>
- Brunner, P., & Simmons, C. T. (2012). HydroGeoSphere: A fully integrated, physically based hydrological model. *Groundwater*, 50(2), 170–176. <https://doi.org/10.1111/j.1745-6584.2011.00882.x>
- Castanheira, P. J., & Santos, F. L. (2009). A simple numerical analyses software for predicting water table height in subsurface drainage. *Irrigation and Drainage Systems*, 23(4), 153–162. <https://doi.org/10.1007/s10795-009-9079-5>
- Conroy, W. J. (2005). Using CCHE1D to improve stream flow routing capabilities of the WEPP model. *Watershed Management to Meet Water Quality Standards and Emerging TMDL (Total Maximum Daily Load) Proceedings of the Third Conference 5-9 March 2005 (Atlanta, Georgia USA)*. ASABE: St. Joseph, MI. <https://doi.org/10.13031/2013.18136>
- Cui, J., Li, C., Sun, G., & Trettin, C. (2005). Linkage of MIKE SHE to Wetland-DNDC for carbon budgeting and anaerobic biogeochemistry simulation. *Biogeochemistry*, 72(2), 147–167. <https://doi.org/10.1007/s10533-004-0367-8>

- Daniel, E. B., Camp, J. V., LeBoeuf, E. J., Penrod, J. R., Dobbins, J. P., & Abkowitz, M. D. (2011). Watershed modeling and its applications: A state-of-the-art review. *The Open Hydrology Journal*, 5, 26–50. <https://doi.org/10.2174/1874378101105010026>
- David, M. B., Drinkwater, L. E., & McIsaac, G. F. (2010). Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, 39(5), 1657–1667. <https://doi.org/10.2134/jeq2010.0115>
- Davison, B., & van der Kamp, G. (2008). Low-flows in deterministic modelling: A brief review. *Canadian Water Resources Journal*, 33(2), 181–194. <https://doi.org/10.4296/cwrj3302181>
- De Schepper, G., Therrien, R., Refsgaard, J. C., He, X., Kjaergaard, C., & Iversen, B. V. (2017). Simulating seasonal variations of tile drainage discharge in an agricultural catchment. *Water Resources Research*, 53(5), 3896–3920. <https://doi.org/10.1002/2016WR020209>
- Dore, M. H. I. (2005). Climate change and changes in global precipitation patterns: What do we know? *Environment International*, 31(8), 1167–1181. <https://doi.org/10.1016/j.envint.2005.03.004>
- Dunkerley, D. L. (2019). Rainfall intensity bursts and the erosion of soils: An analysis highlighting the need for high temporal resolution rainfall data for research under current and future climates. *Earth Surface Dynamics*, 7(2), 345–360. <https://doi.org/10.5194/esurf-7-345-2019>
- Fashi, F. H., & Ejlali, F. (2015). Evaluating spatial distribution of soil erosion using WEPP erosion model and GIS tools: A review. *Agricultura*, 95(3–4), 123–133. <https://doi.org/10.15835/arspa.v95i3-4.11790>
- Flanagan, D. C., Frankenberger, J. R., & Ascough II, J. C. (2012). WEPP: Model use, calibration and validation. *Transactions of the ASABE*, 55(4), 1463–1477.
- Ford, W. I., King, K. W., Williams, M. R., & Confesor, R. B. (2017). Modified APEX model for simulating macropore phosphorus contributions to tile drains. *Journal of Environmental Quality*, 46(6), 1413–1423. <https://doi.org/10.2134/jeq2016.06.0218>
- Frankenberger, J. R., Dun, S., Flanagan, D. C., Wu, J. Q., & Elliot, W. J. (2011). Development of a GIS interface for WEPP model application to Great Lakes forested watersheds. *Proceedings of the International Symposium on Erosion and Landscape Evolution (ISELE)*, 18-21 September 2011, Anchorage, Alaska, 711P0311cd, Paper #11139. ASABE: St. Joseph, MI. <https://doi.org/10.13031/2013.39195>

- Fuka, D. R., Collick, A. S., Kleinman, P. J. A., Auerbach, D. A., Harmel, R. D., & Easton, Z. M. (2016). Improving the spatial representation of soil properties and hydrology using topographically derived initialization processes in the SWAT model. *Hydrological Processes*, 30(24), 4633–4643. <https://doi.org/10.1002/hyp.10899>
- Gelder, B., Sklenar, T., James, D., Herzmann, D., Cruse, R., Gesch, K., & Laflen, J. (2018). The Daily Erosion Project – daily estimates of water runoff, soil detachment, and erosion. *Earth Surface Processes and Landforms*, 43(5), 1105–1117. <https://doi.org/10.1002/esp.4286>
- Gowda, P. H., Mulla, D. J., Desmond, E. D., Ward, A. D., & Moriasi, D. N. (2012). ADAPT: Model use, calibration, and validation. *Transactions of ASABE*, 55(4), 1345–1352.
- Gupta, A., & Cvetkovic, V. (2002). Material transport from different sources in a network of streams through a catchment. *Water Resources Research*, 38(7), 1–3.
- Hansen, A. L., Refsgaard, J. C., Baun Christensen, B. S., & Jensen, K. H. (2013). Importance of including small-scale tile drain discharge in the calibration of a coupled groundwater-surface water catchment model. *Water Resources Research*, 49(1), 585–603. <https://doi.org/10.1029/2011WR011783>
- Hong, E. M., Park, Y., Muirhead, R., Jeong, J., & Pachepsky, Y. A. (2018). Development and evaluation of the bacterial fate and transport module for the Agricultural Policy/Environmental eXtender (APEX) model. *Science of the Total Environment*, 615, 47–58. <https://doi.org/10.1016/j.scitotenv.2017.09.231>
- Ivanov, V. Y., Vivoni, E. R., Bras, R. L., & Entekhabi, D. (2004). Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: A fully-distributed physically-based approach. *Journal of Hydrology*, 298(1–4), 80–111. <https://doi.org/10.1016/j.jhydrol.2004.03.041>
- Jabbari, A., Ackerman, J. D., Boegman, L., & Zhao, Y. (2019). Episodic hypoxia in the western basin of Lake Erie. *Limnology and Oceanography*, 64(5), 2220–2236. <https://doi.org/10.1002/lno.11180>
- Jenny, J. P., Francus, P., Normandeau, A., Lapointe, F., Perga, M. E., Ojala, A., & Zolitschka, B. (2016). Global spread of hypoxia in freshwater ecosystems during the last three centuries is caused by rising local human pressure. *Global Change Biology*, 22(4), 1481–1489. <https://doi.org/10.1111/gcb.13193>

- Jeong, J., Williams, J. R., Merkel, W. H., Arnold, J. G., Wang, X., & Rossi, C. G. (2014). Improvement of the variable storage coefficient method with water surface gradient as a variable. *Transactions of the ASABE*, 57(3), 791–801.
<https://doi.org/10.13031/trans.57.10369>
- Kinnell, P. I. A. (2017). A comparison of the abilities of the USLE-M, RUSLE2 and WEPP to model event erosion from bare fallow areas. *Science of the Total Environment*, 596–597, 32–42. <https://doi.org/10.1016/j.scitotenv.2017.04.046>
- Knisel, W. G., & Douglas-Mankin, K. R. (2012). CREAMS/GLEAMS: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1291–1302.
- Kossieris, P., Makropoulos, C., Onof, C., & Koutsoyiannis, D. (2018). A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures. *Journal of Hydrology*, 556, 980–992.
<https://doi.org/10.1016/j.jhydrol.2016.07.015>
- Kunkel, K. E., Karl, T. R., Easterling, D. R., Redmond, K., Young, J., Yin, X., & Hennon, P. (2013). Probable maximum precipitation and climate change. *Geophysical Research Letters*, 40(7), 1402–1408. <https://doi.org/10.1002/grl.50334>
- Levinton, J., Doall, M., Ralston, D., Starke, A., & Allam, B. (2011). Climate change, precipitation and impacts on an estuarine refuge from disease. *PLoS ONE*, 6(4), 1–8.
<https://doi.org/10.1371/journal.pone.0018849>
- Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, 13(1–4), 195–206. [https://doi.org/10.1016/0921-8181\(95\)00046-1](https://doi.org/10.1016/0921-8181(95)00046-1)
- Liu, Y., Evans, M. A., & Scavia, D. (2010). Gulf of Mexico hypoxia: Exploring increasing sensitivity to nitrogen loads. *Environmental Science and Technology*, 44(15), 5836–5841.
<https://doi.org/10.1021/es903521n>
- Ma, L., He, C., Bian, H., & Sheng, L. (2016). MIKE SHE modeling of ecohydrological processes: Merits, applications, and challenges. *Ecological Engineering*, 96, 137–149.
<https://doi.org/10.1016/j.ecoleng.2016.01.008>
- Maalim, F. K., Melesse, A. M., Belmont, P., & Gran, K. B. (2013). Modeling the impact of land use changes on runoff and sediment yield in the Le Sueur Watershed, Minnesota using GeoWEPP. *Catena*, 107, 35–45. <https://doi.org/10.1016/j.catena.2013.03.004>

- Mao, D., Cherkauer, K. A., & Flanagan, D. C. (2006). Large-scale impacts of frozen soil on soil erosion: Coupling the WEPP model to a macro-scale hydrologic model. Paper No. 062188, 2006 ASAE Annual International Meeting. ASAE: St. Joseph, MI. <https://doi.org/10.13031/2013.21101>
- Martel, J. L., Mailhot, A., Brissette, F., & Caya, D. (2018). Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales. *Journal of Climate*, 31(11), 4241–4263. <https://doi.org/10.1175/JCLI-D-17-0282.1>
- Martinez-Martinez, E., Nejadhashemi, A. P., Woznicki, S. A., & Love, B. J. (2014). Modeling the hydrological significance of wetland restoration scenarios. *Journal of Environmental Management*, 133, 121–134. <https://doi.org/10.1016/j.jenvman.2013.11.046>
- Mittelstet, A. R., Storm, D. E., & Fox, G. A. (2017). Testing of the modified streambank erosion and instream phosphorus routines for the SWAT model. *Journal of the American Water Resources Association*, 53(1), 101–114. <https://doi.org/10.1111/1752-1688.12485>
- Moriasi, D. N., Rossi, C. G., Arnold, J. G., & Tomer, M. D. (2012). Evaluating hydrology of the Soil and Water Assessment Tool (SWAT) with new tile drain equations. *Journal of Soil and Water Conservation*, 67(6), 513–524. <https://doi.org/10.2489/jswc.67.6.513>
- Mullan, D., Favis-Mortlock, D., & Fealy, R. (2012). Addressing key limitations associated with modelling soil erosion under the impacts of future climate change. *Agricultural and Forest Meteorology*, 156, 18–30. <https://doi.org/10.1016/j.agrformet.2011.12.004>
- Müller, H., & Haberlandt, U. (2018). Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. *Journal of Hydrology*, 556, 847–864. <https://doi.org/10.1016/j.jhydrol.2016.01.031s>
- Negm, L. M., Youssef, M. A., Chescheir, G. M., & Skaggs, R. W. (2016). DRAINMOD-based tools for quantifying reductions in annual drainage flow and nitrate losses resulting from drainage water management on croplands in eastern North Carolina. *Agricultural Water Management*, 166, 86–100. <https://doi.org/10.1016/j.agwat.2015.12.014>
- Nicks, A.D., L.J. Lane, and G.A. Gander. 1995. Chapter 2. Weather Generator. In (D.C. Flanagan and M.A. Nearing, eds.) *USDA –Water Erosion Prediction Project Hillslope Profile and Watershed Model Documentation*. NSERL Report No. 10, USDA-ARS National Soil Erosion Research Laboratory, West Lafayette, Indiana. pp. 2.1-2.22.

- Pi, H., Sharratt, B., Feng, G., & Lei, J. (2017). Evaluation of two empirical wind erosion models in arid and semi-arid regions of China and the USA. *Environmental Modelling and Software*, 91, 28–46. <https://doi.org/10.1016/j.envsoft.2017.01.013>
- Pimentel, D. (2000). Soil erosion and the threat to food security and the environment. *Ecosystem Health*, 6(4), 221–226.
- Praskievicz, S., & Chang, H. (2009). A review of hydrological modelling of basin-scale climate change and urban development impacts. *Progress in Physical Geography*, 33(5), 650–671. <https://doi.org/10.1177/0309133309348098>
- Pribadi, D. O., Vollmer, D., & Pauleit, S. (2018). Impact of peri-urban agriculture on runoff and soil erosion in the rapidly developing metropolitan area of Jakarta, Indonesia. *Regional Environmental Change*, 18(7), 2129–2143. <https://doi.org/10.1007/s10113-018-1341-7>
- Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., & Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, 470–471, 138–157. <https://doi.org/10.1016/j.jhydrol.2012.08.041>
- Raju, B. C. K., & Nandagiri, L. (2018). Assessment of variable source area hydrological models in humid tropical watersheds. *International Journal of River Basin Management*, 16(2), 145–156. <https://doi.org/10.1080/15715124.2017.1372446>
- Refsgaard, J. C., & Storm, B. (1995). MIKE SHE. In: Singh, V. P., ed., *Computer Models of Watershed Hydrology*. Water Resources Publications: Highlands Ranch, CO. pp. 809–846.
- Rhodes, C. J. (2014). Soil erosion, climate change and global food security: Challenges and strategies. *Science Progress*, 97(2), 97–153. <https://doi.org/10.3184/003685014X13994567941465>
- Rinaldo, A. (1987). Basin scale model of solute transport. *Water Resources Research*, 23(11), 2107–2118.
- Rinaldo, A., Beven, K. J., Bertuzzo, E., Nicotina, L., Davies, J., Fiori, A., & Botter, G. (2011). Catchment travel time distributions and water flow in soils. *Water Resources Research*, 47(7), 1–13. <https://doi.org/10.1029/2011WR010478>
- Rinaldo, A., Benettin, P., Harman, C. J., Hrachowitz, M., McGuire, K. J., Velde, Y. Van Der, & Botter, G. (2015). Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes. *Water Resources Research*, 51(6), 4840–4847. <https://doi.org/10.1002/2015WR017273>

- Rinaldo, A., Botter, G., Bertuzzo, E., Uccelli, A., Settin, T., & Marani, M. (2006). Transport at basin scales: 1. Theoretical framework. *Hydrology and Earth System Sciences*, 10(1), 19–29. <https://doi.org/10.5194/hess-10-19-2006>
- Rinaldo, A., & Rodriguez-Iturbe, I. (1996). Geomorphological theory of the hydrological response. *Hydrological Processes*, 10(6), 803–829. [https://doi.org/10.1002/\(SICI\)1099-1085\(199606\)10:6<803::AID-HYP373>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1099-1085(199606)10:6<803::AID-HYP373>3.0.CO;2-N)
- Rodriguez-Iturbe, I., & Rinaldo, A. (2001). *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press. 570 pp.
- Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., & Winter, J. M. (2013). The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, 170, 166–182. <https://doi.org/10.1016/j.agrformet.2012.09.011>
- Savabi, M. R. (1993). Modeling subsurface drainage and surface runoff with WEPP. *Journal of Irrigation and Drainage Engineering*, 119(5), 801–813. [https://doi.org/10.1061/\(ASCE\)0733-9437\(1995\)121:2\(217\)](https://doi.org/10.1061/(ASCE)0733-9437(1995)121:2(217))
- Shen, N., Wang, Z., & Wang, S. (2016). Flume experiment to verify WEPP rill erosion equation performances using loess material. *Journal of Soils and Sediments*, 16(9), 2275–2285. <https://doi.org/10.1007/s11368-016-1408-3>
- Singh, R. K., Panda, R. K., Satapathy, K. K., & Ngachan, S. V. (2011). Simulation of runoff and sediment yield from a hilly watershed in the eastern Himalaya, India using the WEPP model. *Journal of Hydrology*, 405(3–4), 261–276.
- Sivapalan, M., & Blöschl, G. (2017). The growth of hydrological understanding: Technologies, ideas, and societal needs shape the field. *Water Resources Research*, 53(10), 8137–8146. <https://doi.org/10.1002/2017WR021396>
- Skaggs, R. W., Youssef, M. A., & Chescheir, G. M. (2012). DRAINMOD: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1509–1522.
- Song, Z., Seitz, S., Li, J., Goebes, P., Schmidt, K., Kühn, P., & Scholten, T. (2019). Tree diversity reduced soil erosion by affecting tree canopy and biological soil crust development in a subtropical forest experiment. *Forest Ecology and Management*, 444(July), 69–77. <https://doi.org/10.1016/j.foreco.2019.04.015>

- Sonnenborg, T. O., Christiansen, J. R., Pang, B., Bruge, A., Stisen, S., & Gundersen, P. (2017). Analyzing the hydrological impact of afforestation and tree species in two catchments with contrasting soil properties using the spatially distributed model MIKE SHE SWET. *Agricultural and Forest Meteorology*, 239, 118–133.
<https://doi.org/10.1016/j.agrformet.2017.03.001>
- Srivastava, A., Brooks, E. S., Dobre, M., Elliot, W. J., Wu, J. Q., Flanagan, D. C., & Link, T. E. (2019). Modeling forest management effects on water and sediment yield from nested, paired watersheds in the interior Pacific Northwest, USA using WEPP. *Science of The Total Environment*, 701(2020), 134877. <https://doi.org/10.1016/j.scitotenv.2019.134877>
- Srivastava, A., Wu, J. Q., Elliot, W. J., Brooks, E. S., & Flanagan, D. C. (2017). Modeling streamflow in a snow-dominated forest watershed using the Water Erosion Prediction Project (WEPP) model. *Transactions of the ASABE*, 60(4), 1171–1187.
- Therrien, R., McLaren, R. G., Sudicky, E. A., & Panday, S. M. (2010). *HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport*. Groundwater Simulations Group, University of Waterloo, Waterloo, ON.
- Todini, E. (2007). Hydrological catchment modelling: Past, present and future. *Hydrology and Earth System Sciences*, 11(1), 468–482. <https://doi.org/10.5194/hess-11-468-2007>
- Turrall, H., Burke, J., & Faures, J. M. (2011). Climate change, water and food security. *Food and Agriculture Organization*, 66, 37–39.
- USEPA. 2007. *Hypoxia in the northern Gulf of Mexico: An update by the EPA Science Advisory Board*. EPA-SAB-08-004. USEPA, Washington, DC.
- Vu, T. M., Raghavan, S. V., Liong, S. Y., & Mishra, A. K. (2018). Uncertainties of gridded precipitation observations in characterizing spatio-temporal drought and wetness over Vietnam. *International Journal of Climatology*, 38(4), 2067–2081.
<https://doi.org/10.1002/joc.5317>
- Wang, L., Wu, J. Q., Elliot, W. J., Dun, S., Lapin, S., Fiedler, F. R., & Flanagan, D. C. (2010). Implementation of channel-routing routines in the Water Erosion Prediction Project (WEPP) model. *Proceedings of the 2009 SIAM Conference on “Mathematics for Industry” The Art of “Mathematics for Industry”*, 120–127. SIAM.

- Wang, L., Flanagan, D., Wang, Z., & Cherkauer, K. (2018). Climate change impacts on nutrient losses of two watersheds in the Great Lakes region. *Water*, 10(4), 442. <https://doi.org/10.3390/w10040442>
- Wang, X., Gassman, P. W., Williams, J. R., Potter, S., & Kemanian, A. R. (2008). Modeling the impacts of soil management practices on runoff, sediment yield, maize productivity, and soil organic carbon using APEX. *Soil and Tillage Research*, 101(1–2), 78–88. <https://doi.org/10.1016/j.still.2008.07.014>
- Wischmeier, W. H., & Smith, D. D. (1978). Predicting rainfall erosion losses. A guide to conversation planning. In Science and Education Administration. Agriculture Handbook No. 537, USDA, Washington, D.C.
- Wu, J., Miao, C., Zhang, X., Yang, T., & Duan, Q. (2017). Detecting the quantitative hydrological response to changes in climate and human activities. *Science of the Total Environment*, 586, 328–337. <https://doi.org/10.1016/j.scitotenv.2017.02.010>
- Zemke, J. J., Pöhler, J., & Stegmann, S. (2019). Modeling runoff-formation and soil erosion after pumice excavation at forested Andosol-sites in SW-Germany using WEPP. *Soil Systems*, 3(3), 48. <https://doi.org/10.3390/soilsystems3030048>
- Zhang, D., Chen, X., & Yao, H. (2016). SWAT-CSenm: Enhancing SWAT nitrate module for a Canadian Shield catchment. *Science of the Total Environment*, 550, 598–610. <https://doi.org/10.1016/j.scitotenv.2016.01.109>
- Zhang, D., Madsen, H., Ridler, M. E., Refsgaard, J. C., & Jensen, K. H. (2015). Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model. *Advances in Water Resources*, 86, 400–413. <https://doi.org/10.1016/j.advwatres.2015.07.018>
- Zhang, H., & Zhuang, L. (2019). The impact of soil erosion on internal migration in China. *PLoS ONE*, 14(4), 1–17. <https://doi.org/10.1371/journal.pone.0215124>

2. DEVELOPMENT OF A STOCHASTIC STORM GENERATOR USING HIGH-RESOLUTION PRECIPITATION RECORDS

Abstract

Sophisticated field and watershed scale environmental models for runoff, erosion control, environmental, and global-change investigations require detailed continuous temporal and spatial inputs of precipitation to drive the hydrologic processes. For accurate estimates of these processes, the resolution of the input data must allow the representation of the variability of precipitation as it represents a major source of variability in the model outputs. Currently, the use of stochastic weather generators is widespread to generate continuous series of meteorological data at gauged and ungauged locations. These weather simulators are designed to replicate the statistical properties of real weather data at monthly or daily time resolutions. However, daily values of precipitation do not represent the variability of storm parameters within a day, which is assumed to significantly influence the predictions of environmental or agricultural models where processes are sensitive to sub-daily values. This research proposes a parsimonious stochastic storm generator based on 5-minute time resolution and correlated non-normal Monte Carlo-based numerical simulation. The model considers correlated non-normal random rainstorm characteristics such as time between storms, duration, and amount of precipitation, as well as the storm intensity structure. The accuracy of the model was verified by comparing the generated rainfall with rainfall data from a randomly selected 5-min weather station in North Carolina. Current results have shown that the proposed storm generator can capture the essential statistical features of rainstorms as well as their patterns followed by their intensities, preserving the first four moments of monthly storm events, good annual extreme event correspondence, and the correlation structure within each storm. Finally, as the proposed model depends on statistical properties at a site, this may allow the use of the synthetic storms in ungauged locations provided relevant information from a regional analysis is available.

2.1 Introduction

Advanced field and watershed scale hydrological models for engineering design, soil erosion prediction, land use planning, and global-change research require high-resolution continuous

temporal and spatial inputs of precipitation to execute the hydrologic processes logic integrated into their formulations (Stone et al., 1995; Arnold et al., 2012). Accurate estimates of processes such as infiltration, soil erosion, and chemical transport in water need precipitation increments on the order of minutes. However, observed precipitation data series with such resolution are rarely available or lack wide-spread spatial coverage.

Extensive research has been conducted in several disciplines using daily or longer time-scale precipitation, due in large part to the availability of high-quality global land-surface precipitation datasets (Tank et al., 2002; Menne et al., 2012; Schamm et al., 2014). However, this time increment does not represent the natural variability of storm events. Storms can occur at different times throughout a day with different durations and intensities, including the possibility of individual storm events lasting longer than a day. Scientists overcome these problems through the use of complex transformation methods known as rainfall disaggregation or temporal downscaling (Pui et al., 2012; Lu and Qin, 2014; Lee and Park, 2017). Additionally, observed precipitation data show dependency and correlation of the random variables defining the storms, which makes it difficult to reproduce daily precipitation records containing the natural variability observed in natural storm records.

Research on climate generators allows scientists to produce long series of daily synthetic precipitation values to use as meteorological inputs to hydrological models. These models often provide series of continuous daily precipitation and characteristics defining a single rainstorm event such as storm duration, storm intensities, and wet or dry day occurrence. Although daily storm generators are widely used and define the minimum storm variables, they are still an unrealistic representation and do not capture the intrinsic variability of storms and the statistical properties discussed above. Among all storm generators found in the literature, the Weather Generator (WGEN) (Richardson and Wright, 1984), USCLIMATE (Hanson et al., 1994), CLimate GENERator (CLIGEN) (Nicks et al., 1995); Long Ashton Research Station-Weather Generator (LARS-WG) (Semenov and Barrow, 2002), and a MATLAB-based daily scale weather generator (WeaGETs) (Chen et al., 2010, 2012) are widely used along with a disaggregation procedure to provide climate inputs for a variety of hydrological models with daily time step simulations.

Stochastic generation of hyetographs or within-storm intensities given a total daily amount of precipitation is complicated, as a single event can span less or more than 24 hours or several individual storms can occur throughout a day. Studies have presented two approaches to deal with

this complexity: 1) the use of so-called design storms to represent the variability of storm rainfall intensities, and 2) daily and sub-daily storm generators using rainfall disaggregation methods and high-resolution precipitation records. On one hand, the design storm methods provide an entire description of storm characteristics, which are often derived from Intensity-Duration-Frequency (IDF) curves; Chicago type, Huff curves; or observed large storm events. However, although these methods may be a good representation for design purposes, they do not represent the natural variability of storm intensities, which in fact present infinite variations and are intrinsically correlated. On the other hand, current sub-daily storm generators contain different variations of Poisson point process and cluster-based models (i.e., Neyman-Scott and Barlett-Lewis rectangular pulses). These models have shown different degrees of accuracy and applicability. However, one drawback is that they assume that intensities and durations are mutually independent, and independent of the Poisson process. These assumptions are invalid as storm duration and depth show strong correlation as well as the time distribution of intensities (Bonta, 2004).

Based on the previous discussion, research on stochastic storm generation is needed to overcome these limitations. A stochastic storm generator has been developed at the National Soil Erosion Research Laboratory (NSERL) at West Lafayette, Indiana. The model stochastically simulates storm occurrence, time between storms, storm duration, and storm depth, characteristics found in the current daily and sub-daily storm generators. However, the formulation of the model accounts for the intrinsic correlation among the random variables defining the storms, which is not considered in the current models. Additionally, the storm generator simulates the variability of intensities in natural storms rather than the selection of fixed patterns (i.e., design storm methods) or the assumptions in the rectangular pulses Poisson process or cluster-based models. The formulation of the model involves the application of statistical techniques in some of the models found in the literature, as well as the use of a correlated non-normal multivariate Monte Carlo simulation integrated into both the simulation of the random variables defining the storm and its pattern. The flexibility of its formulation allows for future development of spatial and temporal storm generation in ungauged locations.

2.2 Objectives and scope

The objectives of this study were: (1) to present the concepts and the characterization of the variables defining precipitation events; (2) to present the mathematical and statistical methods to

reproduce series of correlated storm variables and storm patterns using a machine learning tool; and (3) to conduct a preliminary test and analysis of the model performance. A more extensive and detailed application of the model for other locations is a matter for future studies.

2.3 Procedure

2.3.1 General approach

The approach in this model was statistical and did not relate storm characteristics to physics. The model initialization consists of the parameterization of the geographical location (e.g, precipitation gauge) where the synthetic storms are to be generated. The information needed to create the parameters file include: station ID, geographical location, time resolution, predefined number of storm patterns, number of years to be synthesized, number of years of observed data available, and input filename containing the observed precipitation record. Input data required by the model must be fine-resolution precipitation records, ranging from 1-min interval to 1-h interval at maximum. A minimum length of 10 years of observed data is recommended to ensure the accuracy of storm temporal definition. The generated synthetic storms will have the same resolution of the input data. The initial model execution generates the parameter files defining the storm distribution random variables, actual number of storm patterns, and the general descriptive statistics of the observed records.

The main model tasks include the identification of statistically independent storms in the input precipitation dataset, characterization of the random variables defining storm events such as time between storms, storm depth, storm duration, and storm patterns, parameterization of suitable distributions and their correlation, and the generation of a synthetic time series of precipitation. Storm identification required the well-known assumption of independent storm events following a Poisson process with exponentially distributed arrival times (Todorovic and Yevjevich, 1979; Rodriguez-Iturbe et al., 1987a). Although stationarity might not be correct for long records of precipitation (e.g. Midwestern U.S. systems), a monthly characterization of storms assumes stationarity within a month. Thus, monthly descriptions of storm variables minimize the amount of data characterization and represent the most essential inputs required for model execution. Alternatively, the model could be set up for seasonal characterization of storms in its basic configuration.

The mathematical formulation of the model treats the random variables defining storms as correlated. Data analyses have shown a strong correlation between storm duration and storm depth, whereas time between storms is weakly correlated with the other two variables. Model parametrization does not require any optimization or calibration, as it only depends on the resolution of the input variables in the available precipitation record.

The generation of hyetographs uses the advantages offered by clustering algorithms to identify similar storm patterns at a location. The inner structures of the storms, that is to say, each of the pulses of precipitation defining the storm pattern, are treated as random variables with individual frequency distributions. Similar to the generation of the storm variables mentioned above, the model assumes that there are correlations and dependencies among the storm depths. This assumption allows capturing the variability seen in natural storms and avoids the use of a fixed shape as in several existing storm generators.

In this study, each storm variable, the methods for identification and characterization, model conceptualization and parametrization, and simulation concepts are presented. At the end of this article, a preliminary test of model performance is presented. Evaluation of the outputs includes comparisons of measured and simulated frequency distributions, monthly averages, and storm pattern accuracies. All algorithms and functions integrated in the storm generator proposed in this study were coded in the R-programming language version 3.4.4.

2.3.2 Data used

The model proposed requires time series of precipitation at sub-daily time resolution (e.g., 1 h, 15 min, 10 min) to generate synthetic series of precipitation. Consequently, any location with precipitation measured at this resolution can be used. Since the main objective of this investigation was to determine the potential of the proposed model to synthesize storms, data inputs were set to minimal requirements (e.g. minimum number of years for storm separation). In this study, ten years of data from rain gauge “USGS351320080502645” at Charlotte-Mecklenburg, North Carolina, here named USGS-CHM, were used to illustrate the concepts of the stochastic storm generator. The data ranged from 2008 to 2017 in increments of 5 min with a rainfall depth resolution of 1 mm. Average annual precipitation for Charlotte-Mecklenburg is 1023 mm. The current state of the model does not deal with months that contain precipitation in the form of snow, so months with snow values in the records were discarded. However, this is the subject of ongoing

research and will be evaluated in subsequent papers. Additionally, a minimum number of 60 storms was required to maintain the statistical power of the test for goodness of fit (e.g., Kolmogorov-Smirnov test). As a result of these two conditions, and with the intention to simulate a continuous period, months from March to August were selected for this study.

2.4 Characterization of rainstorm events

2.4.1 Definition of a rainstorm event

Analysis of the probabilistic properties of rainstorm characteristics requires prior identification of rainstorm events. In meteorological terms, a storm is an event where rainfall is falling intermittently (Wu et al., 2006). In several studies on stochastic storm generation, storms were recommended to be statistically independent events (Restrepo and Eagleson, 1982; Rodriguez-Iturbe et al., 1984; Marien and Vandewiele, 1986; Acreman, 1990; Bonta, 2003, 2004). This assumption preserves the stochastic representation of a Poisson arrival process and facilitates the characterization of storms, parameter identification, and modeling the occurrence of events by avoiding the use of a conditional probability density function. Some of the methods for storm identification focus on the definition of an arbitrary threshold (e.g., 6 h in Wischmeier and Smith, 1978), autocorrelation (Wenzel and Voorhees, 1981), rank correlation (Bonta and Rao, 1988), and exponential distribution (Bonta, 2003). All of these methods use the concept of a critical duration (CD), which means a minimum dry period separates individual storms in the precipitation record.

Bonta and Rao (1988) found that the exponential method by Restrepo and Eagleson (1982) provided a better characterization for storm generation purposes. The assumptions in this method define the time between storms as exponentially distributed, and the storm generation equivalent to a Poisson process, making events statistically independent and simplifying simulation. Bonta and Rao (1988) estimated that at least ten years of precipitation records were needed for a reliable estimation of the critical duration for rain gauges in Ohio. The variability of the CD was very high if the calculations were performed on fewer than six years of data. Detailed studies of the procedure are given in Bonta (2001, 2004).

In this study, the identification of independent storms is based on the exponential model presented by Restrepo and Eagleson (1982). However, the present model allows the user to select the rank correlation method or an arbitrary threshold as alternative options. For the USGS-CHM

precipitation data, the CD values ranged from 295 to 1270 min. August presented the longest average time between storm (ATBS = 7798 min), whereas July had the shortest rainless time (ATBS = 3869 min). Figure 2.1 shows the exponentially distributed time between storms (TBS) plotted in a logarithmic scale for the months with the greatest number of storms and precipitation in the form of rainfall (March-August). To validate the Poisson-process assumption, the ratio of the simulated critical duration to the ATBS is calculated in Table 2.1 for months from March to August. Values much smaller than unity suggest that the Poisson-process analog is appropriate (Restrepo and Eagleson, 1982). Consequently, values close to unity were considered as months with storm events in which the Poisson process assumption was not valid, and therefore they were discarded from this study.

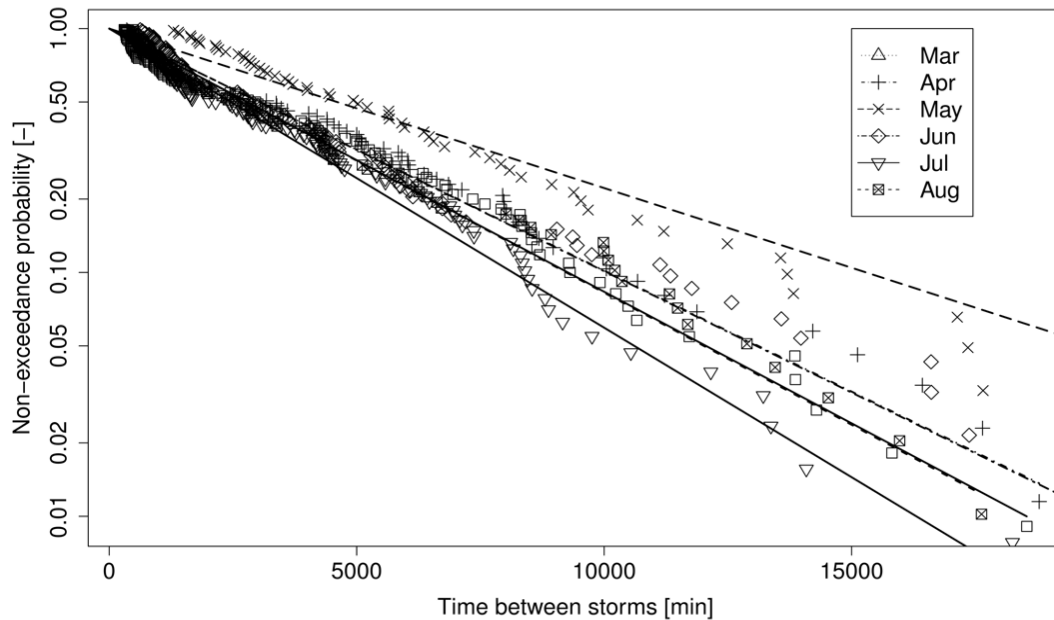


Figure 2.1 Empirical distributions of monthly time between storms and the corresponding theoretical distributions calculated in the exponential method for USGS-CHM precipitation data (2008-2017) at Charlotte-Mecklenburg, North Carolina.

2.4.2 Characterization of storm arrival

The generation of the storm sequences over a period of years to simulate accounts for the temporal distribution of the monthly number of storms found in the data. The Poisson distribution is commonly used to describe the variability of annual occurrence of hydrologic events (Eagleson

et al., 1987; Clark, 1998). Thus, the occurrence of storms at USGS-CHM was characterized for each month over the entire period of years and fitted into a Poisson distribution with rate $1/\mu$, where μ is the average of the monthly occurrence over the period of years analyzed.

The accuracy of the Poisson distribution to replicate storm arrivals is measured by the Fisher dispersion index, D_i , defined as:

$$D_i = \sum_{i=1}^n \frac{(s_i - \bar{s})^2}{\bar{s}} \quad (2.1)$$

where s_i is number of storms in the month i ($i = 1, 2, \dots, n$), and \bar{s} the sample mean of s_i . The Fisher dispersion index follows a χ^2 distribution. The hypothesis that the storm occurrence follows a Poisson distribution is not rejected if the p -value associated with the sample D_i is larger than the 5% significance level.

Table 2.1 Measured rainfall characteristics at Charlotte-Mecklenburg, North Carolina for USGS-CHM.

Month	Critical Duration (min)	Average Time Between Storms (ATBS) (min)	Ratio of Storm Duration to ATBS	Average Storm Duration (min)	Average Storm Depth (mm)
March	285	3930	0.075	395.1	7.6
April	310	4240	0.083	350.5	10.1
May	1270	6469	0.114	737.2	16.0
June	520	4206	0.051	214.2	9.8
July	295	3443	0.037	128.9	8.1
August	460	3903	0.065	253.8	11.0
Average			0.061		

Table 2.2 shows the number of storms by month over the entire record. July had the shortest time between storms (Table 2.1), and it also had the maximum number of storms, which means that a small TBS will generate a greater number of storms after sampling its corresponding distribution. January, February, September and October had the fewest number of storms, so they were discarded from this study as well as months when snow may occur (November, December). A total of 590 out of 779 storms were contained in the selected period (March-August).

Table 2.2. Number of precipitation events by month.

Period	Number of storms
January	59
February	64
March	109
April	86
May	60
June	92
July	127
August	97
September	47
October	51
November	52
December	81
Total	925

2.4.3 Characterization of time between storms, storm duration, and storm depth

Storm modeling requires the characterization of three main storm variables: time between storms, storm duration, and storm depth. The time between storms is defined as the period between two individual storm events. Storm duration represents the total time from the beginning of a storm to its end. Depending on how storms are identified, storm duration can include periods where rainfall intensities are zero. The exponential method proposed by Restrepo and Eagleson (1982) allows storms with zero-intensity ordinates. Storm depth quantifies the total amount of precipitation measured over the duration of an individual storm.

From the database of 590 storms previously identified, storm durations and storm depths were characterized into series with length equal to the number of storms. Marginal distributions for each variable were then defined, and the goodness of fit tested. The proposed model allows selection of the best distribution among five options: Exponential, Gamma, Log-normal, Pareto, and Weibull. Two parameter estimation methods are used: a) Maximum Likelihood and b) L-moments. The goodness of fit of the marginal distributions is measured with the Kolmogorov-Smirnov test, where the null hypothesis that the distribution comes from the selected option is rejected if the associated p-value is lower than the 5% significance level.

The time between storms (TBS) for USGS-CHM at Charlotte-Mecklenburg characterized in the previous section was found to be exponentially distributed because of an assumption explicitly considered in the identification of storms. As a result, the marginal distribution of TBS could be easily described by the one-parameter exponential distribution with the parameter set equal to the average time between storms (ATBS in Table 2.1). However, the model also has the option to use other distributions for TBS simulation.

Monthly storm duration frequency distributions were created using the database of storms previously identified with the exponential method. The measured monthly average storm durations ranged from 128.9 min (July) to 737.2 min (May) (Table 2.1). Storms longer than 24 h were more frequent in May (20%) than in all other months considered in this study (~10%). The empirical distributions plotted in Figure 2.2 quantify the effect of time of year on the storm duration variability. At this particular location, storm duration tended to fit the two-parameter Weibull distribution which has been widely used to simulate hydrological variables such as the depth of precipitation (Abas et al., 2014) and storm duration (Raudkivi and Lawgun, 1970; Bonta and Rao, 1992; Bonta, 2004).

Analogous to storm duration, empirical distributions of the storm depth were prepared from the database of storms and fitted to a Weibull distribution function. A visual analysis of Figure 2.2 indicates that there is a high correlation between storm duration and storm depth. Storms in May had the longest durations and delivered the greatest precipitation depths. A more descriptive analysis of the storm characteristics series revealed that the Pearson correlation coefficients for the data ranged from 0.49 to 0.75. Hence, it is relevant that the stochastic storm generation considers the intrinsic correlation that exists between storm duration, storm depth, and time between storms.

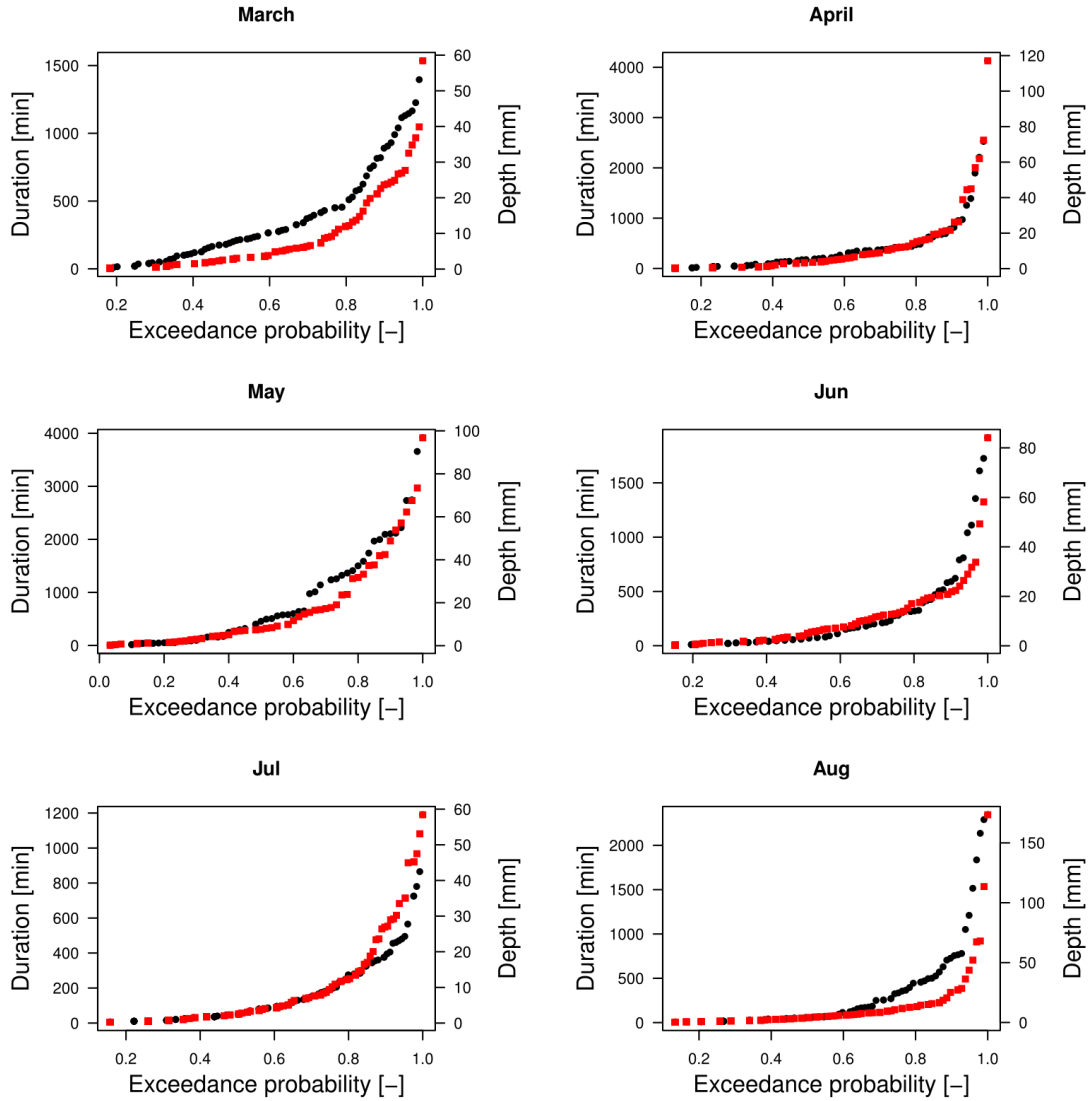


Figure 2.2 Monthly storm depth (red squares) and storm duration (black circles) empirical distributions for USGS-CHM at Charlotte-Mecklenburg, North Carolina. The distributions were truncated to the depth resolution of the data (1 mm).

Among the methods to simulate correlated random variables, Bonta (2004) defined three ways to deal with the aforesaid correlation. The first method, regression, uses a simple power equation to remove the positive correlation between the storm depths and durations. Storm depth is found after sampling the storm duration distribution function using the regression equation and the frequency distributions of the residuals. A drawback of this method is that it is influenced by the variability of the residuals. The second method consists of determining regions of statistical

independence between storm depth and duration by iteratively taking a minimum set of points and testing them for statistical dependence. Once the regions are defined, storm depths contained in the region are collapsed and fitted to empirical distributions. One problem in this method is that the conditional storm depths do not generate unique independent regions since the method depends on the selected initial set of values. The last method is a slight variation of the second one where the problem of non-unique independent duration regions is solved. However, the correlation between quantiles is only possible after collapsing the data from the estimated depth distributions. To avoid the introduction of uncertainties that strictly depend on the conceptualization of specific parameters and user interpretation as cited above, a more sophisticated method has been adopted in this study to deal with the correlation among random variables. A Monte Carlo numerical simulation of multivariate distributions that receives information from the marginal distributions and correlation from the three storm characteristics is applied. This method has the advantage of requiring only basic information of the random variables, and does not require their joint distributions, which are generally difficult to establish but necessary when simulating correlated random variables. More details of the model and its implementation are discussed in the next section.

2.5 Stochastic simulation of rainstorm sequences

2.5.1 Monte Carlo correlated non-normal multivariate simulation

Storm duration, rainfall depth, and time between storms are inherently correlated, and their probabilistic distributions are likely to be non-normal. Therefore, the generation of these three rainstorm characteristics should preserve their respective marginal statistical properties and correlation. The model proposed here implements a procedure that replicates storm variable characteristics by using a Monte Carlo numerical simulation to generate random variables given the distribution properties and their correlation.

Statistically and physically dependent random variables are not unique to storm generation but in many practical engineering situations. Furthermore, distribution types for the random variables involved can be a mixture of different distributions. To accurately replicate such systems, the Monte Carlo approach used here maintains not only the first moments of the random variables but also preserves the correlation relationships between the stochastic parameters and their

distributions. Examples of this type of Monte Carlo simulation include structural reliability analysis (Li et al., 2013), probabilistic load flow (Chen et al., 2012, 2015) and probabilistic safety analysis (Karanki et al., 2010); however, they have been rarely applied in hydrological modeling.

Because of the nature of engineering problems and limited statistical data, the joint probability density function (PDF) of all input variables is hard to acquire, so that direct implementation of a Monte Carlo experiment becomes unsuitable. Here, the Nataf transformation is used to overcome the limited data issue. The mathematical model implemented in the storm generator allows the transformation from original space to mutually independent standard normal space. As a result, the joint PDF of the random variables is not required, making the marginal distributions of the input variables as well as their correlation matrix the only two requirements to run the model. Both inputs were obtained from the datasets characterized in the previous section. A brief explanation of the Nataf transformation is provided next, and more detailed information can be found in Nataf (1962).

The Nataf transformation or Nataf model considers standard normal variates $Z = (Z_1, \dots, Z_n)$ calculated by marginal transformations of $X = (X_1, \dots, X_n)$ with a correlation matrix R .

$$Z_i = \Phi^{-1}[F_{X_i}(X_i)] \quad i = 1, \dots, n \quad (2.2)$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution. The Nataf distribution for X is obtained by assuming that Z has a normal joint distribution (Nataf, 1962). The joint PDF of X after transformation is of the form:

$$f_X(x) = f_{X_1}(x_1)f_{X_2}(x_2) \dots f_{X_n}(x_n) \frac{\varphi_n(z, R')}{\varphi(z_1)\varphi(z_2) \dots \varphi(z_n)} \quad (2.3)$$

where $z_i = \Phi^{-1}[F_{X_i}(x_i)]$, $\varphi(\cdot)$ is the standard normal PDF, and $\varphi_n(z, R')$ is the n-dimensional normal PDF of zero means, unit standard deviations, and correlation matrix R' . The elements ρ'_{ij} of R' are expressed in terms of the correlation coefficients ρ_{ij} (correlation coefficients of the random variables X_i) through the integral relation:

$$\rho_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x_i - \mu_i}{\sigma_i} \right) \left(\frac{x_j - \mu_j}{\sigma_j} \right) \varphi_{ij}(z_i, z_j, \rho'_{ij}) dz_i dz_j \quad (2.4)$$

To solve Equation (2.4), several alternatives were explored. Chang et al. (1994) overcame the required computation for solving ρ'_{ij} when the correlation coefficient ρ_{ij} and the marginal distribution are given using a set of semi-empirical formulas proposed by Der Kiureguian and Liu (1986) for a variety of probability distributions. However, some of the distributions analyzed in Der Kiureguian and Liu (1986) were not the same as the ones used here. Furthermore, the use of Der Kiureguian and Liu's expressions are restricted to correlation values between 0.1 and 0.5.

Similarly, Xiao (2014) analyzed the use of two possible methods for computing ρ'_{ij} : (1) interpolation using the Gauss-Hermite quadrature and (2) a Monte Carlo simulation. For three mixtures of correlated distributions, Xiao (2014) concluded that equivalent results were obtained by both techniques. One advantage of the Monte Carlo method is that it can be used when an analytical expression between ρ_{ij} and ρ'_{ij} is unobtainable. As a result, a Monte Carlo technique like the one presented by Xiao (2014) was adopted for a mixture of the three distributions defining storm characteristics.

Finally, after all coefficients of the matrix R' in the standard normal space were obtained, the inverse transformation was calculated by the following expression:

$$X_i = F_i^{-1}[\Phi(Z_i)] \quad (2.5)$$

where $F_i^{-1}[\cdot]$ is performed using the known marginal distributions. Figure 2.3 is a flow chart outlining the proposed procedure.

The generation of storm characteristics by using the Monte Carlo Multivariate method involves the following steps:

Step 1. Identify statistically independent rainstorms. This can be done using the exponential method, rank correlation, or user-defined critical duration.

Step 2. Based on the storms identified in Step 1, summarize relevant monthly statistics of rainstorm characteristics (occurrence, the time between storms, duration, and depth) and determine statistical moments, marginal distribution parameters, and correlation relationships.

Step 3. Generate the number of storms for each month, year by year over the simulation period.

Step 4. For each generated storm event, simulate the time between storms, storm duration, and storm depth using the procedure for simulating correlated multivariate non-normal variables shown in Figure 2.3.

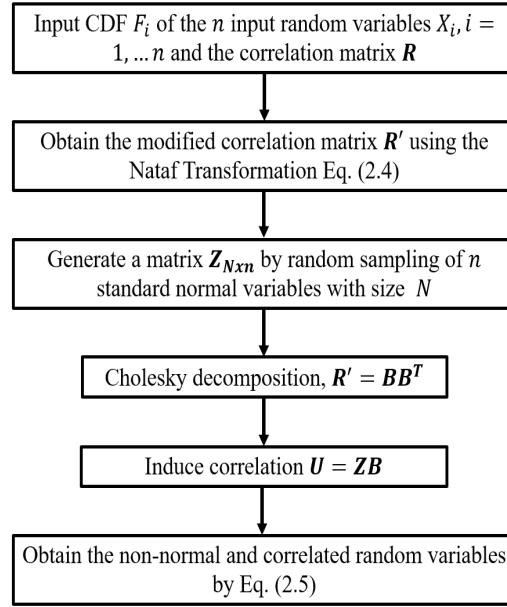


Figure 2.3 Flowchart of a Monte Carlo correlated multivariate simulation.

2.5.2 Modeling storm pattern

A storm pattern or hyetograph is a representation of the time distribution of rainfall depth or intensity within a storm event. Storm distributions are required for the design of small drainage systems, which are more sensitive to short-duration and high-intensity storm events than large basins. However, some processes such as infiltration and soil erosion at basin scales often require high-resolution rainfall. As storm patterns change from one event to another, it is necessary that the model introduced here not only generates rainfall variables such as time between the storm, duration, and depth but also simulates the rainstorm patterns that represent the stochasticity of storm depths or intensities.

Synthetic hyetographs or storm patterns are often generated by four methods: (1) particular geometrical shapes fixed to a peak intensity; (2) intensity-duration-frequency (IDF) curves; (3) dimensionless mass curves; and (4) stochastic models. Geometrical shapes are used to represent a simple local storm distribution that preserves rainfall characteristics of a particular zone (Ellozue et al., 2009). A fixed shape for hyetographs represents by far the greatest simplification of storm

pattern when data are not available; however, the approach does not agree with the behavior of natural rainfall. The well-known rainfall IDF method such as the Chicago method (Keifer and Chu, 1957) and the Frequency Based Hypothetical Storm (Feldman, 2000) are focused on intense bursts within rainfall events rather than series of events. Average duration and rainfall intensity (with a particular frequency) are extracted from IDF curves and used to obtain the final synthetic hyetograph. The resulting pattern is an envelope curve which tends to be conservative for the general design of hydro-system infrastructure. Nonetheless, they do not resemble the rainfall patterns that are seen in natural precipitation records.

More realistic approaches include expressions of the random and highly variable character of within storm intensities known as Huff curves (Huff, 1967) and the Soil Conservation Service (SCS) rainfall type curves (Cronshey, 1986). Huff (1967) developed this method by considering heavy rains in the Midwestern United States, with durations ranging from 3 to 48 h over an area up to 400 square miles. After deriving the dimensionless cumulative mass curves, the distributions were grouped into quartiles in which the peak intensity was located. These quartiles describe when the peak intensity occurs in a given storm event.

Early stochastic models of the rainfall process included point process rainfall models created to replicate statistical characteristics like the rainfall used to fit the model. Since daily records were the data available for these models, they often integrated disaggregation methods to produce synthetic rainfall at finer resolutions (e.g., hourly). The first novel developments were the Poisson counting process for rainfall time series by Rodriguez-Iturbe et al. (1984) and its later improvement by proposing a cluster-based rectangular model with two main versions: the Neyman-Scott process and the Barlett-Lewis process (Rodriguez-Iturbe et al., 1987a, b). These two models rely on the same basic theory and only differ in the generation of rainfall pulses. More recent studies have used the information contained in Huff curves to stochastically estimate rainfall distributions (Bonta, 2004). Although the method was a promising way to use Huff curves for stochastic generation of storms, further research is needed to deal with the apparent correlation of adjacent depths and serial correlation incorporated into the mass curves using Huff curves.

As each actual rainfall event is unique concerning its duration, depth, and pattern, this study adopted a more practical method to identify a few characteristic patterns and characterize the variability of their intensities rather than use a single curve as in previous studies. Hence, cluster

analysis objectively groups the entities by their similarities and differences (Tyron and Bailey, 1970).

In general, clustering algorithms can be divided into two categories: (1) hierarchical clustering and (2) non-hierarchical clustering. The proposed storm generator is based on a non-hierarchical clustering approach where the number of clusters is fixed in advance, and several iterations are performed to identify the best cluster and their cluster centers. Applications of clustering techniques have provided evidence that non-hierarchical clustering approaches are better than hierarchical clustering ones (Punj and Steward, 1983). Lin et al. (2017) used a two-step clustering technique and pattern recognition to propose a model for a typhoon hourly rainfall forecasting model using historical data from the Tamsui River Basin in northern Taiwan. Wu et al. (2006) applied the Euclidean distance-based K-means clustering method (MacQueen, 1967) to classify 8289 rainfall events into representative patterns occurring in Hong Kong territories in Southeast Asia. In their research, they constructed synthetic hyetographs by assuming that intensities were correlated as non-normal random variables under a unit-sum constraint. Because of this assumption, the log-ratio transformation was adopted to simplify the simulation into unconstrained multivariate non-normal variates of correlated variables. Wu et al. (2006) concluded that clustering analysis combined with a constrained Monte Carlo numerical simulation offered a suitable procedure for simulating storm patterns with the limitation that the log-ratio transformation was only acceptable for non-zero values of rainfall intensities.

To synthesize hyetographs in this study, representative storm patterns for the USGS-CHM rain gauge at Charlotte-Mecklenburg, North Carolina, were defined by using the divisive K-means clustering method, previously shown to be satisfactory at classification of rainstorm patterns (Ramos, 2001; Wu et al., 2006). However, the limitation of the non-zero precipitation restriction was overcome.

Before cluster definition, cumulative mass curves were derived from the 925 storms identified from the USGS-CHM rain gauge. Dimensionless curves were obtained by normalizing each storm with respect to its total duration and total depth. This allowed rainfall events with different durations and depths to be combined, examined, and categorized into representative rainfall patterns. Normalized duration D_t was obtained by dividing the cumulative time in a storm by the total duration, D , and normalized depth, F_t , was calculated as the cumulative rainfall depth

divided by the total depth, D_d . Incremental rainfall or the hyetograph ordinates, P_τ , were computed as $F_\tau - F_{\tau-1}$.

Clustering analysis required a D-dimensional data set of length n , and a k value that indicates how many cluster means should be produced. To produce inputs with the same dimension for all the storms, mass curves were divided into m number of elements. Although m was an arbitrary number, its selection had to assure capture of the overall shape of the dimensionless mass curves. A high number of elements might capture unnecessary details of the curves, while too small of a number could miss important characteristics of the storm patterns. In this study, 15 elements were considered sufficient to capture the variability of the wide number of storm patterns. Since the appropriate number of clusters (k) was not known in advance, it is commonly estimated by a trial-and-error process. For USGS-CHM, six significant clusters were identified based on visual inspection of the centroid curves (Figure 2.4). A rule of thumb is that overestimation of the number of representative patterns shows curves overlapping existing cluster centroids, while representative clusters with significant space in between means that more clusters are necessary.

Among the clusters shown in Figure 2.4, four basic types can be observed. R-1 patterns represent storms events having the earliest peak intensity at the beginning of the storms compared to R-2. Storms adopting D-2 patterns are events with the most delayed peak intensity at the end of the storm compared to the ones in D-1 patterns. The M-1 pattern represents storms that attain their peak intensity at half of their durations. Lastly, the C-1 pattern represents storms with relatively uniform intensity over the entire duration.

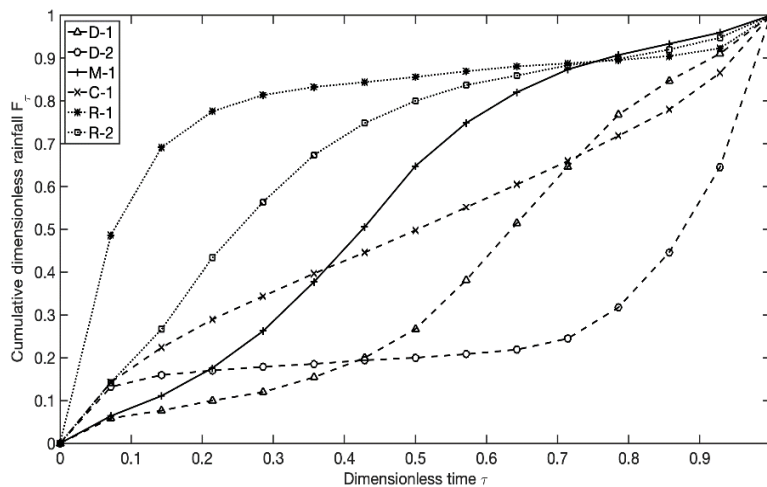


Figure 2.4 Significant clusters for USGS-CHM at Charlotte-Mecklenburg, North Carolina.

Storm pattern generation studies previously mentioned lack a procedure that examines the intrinsic variability of the storm intensities within an event assuming that storms follow a particular shape, or they are limited to a modification of the Neyman-Scott or the Barlett-Lewis processes. A similar procedure as the one in Wu et al. (2006) was followed in this study. However, former limitations such as non-zero values of precipitation were overcome in the proposed model because this is often seen in natural rainfall. To simulate the hyetograph more suitably, the intensity variability inside each representative pattern needs to be considered. Consequently, to stochastically simulate Pattern M-1 in Figure 2.5 containing 86 storm shapes (dashed-lines), ordinates of the corresponding curves need to be fit into a suitable distribution function. The variety of shapes offered by the Johnson distribution makes it the best one for this study (Kottegoda, 1987). The goodness of fit of the distribution was measured by the Kolmogorov-Smirnov test where the null hypothesis that the data comes from the selected distribution is rejected when the associated p-value is less than the 5% significance level.

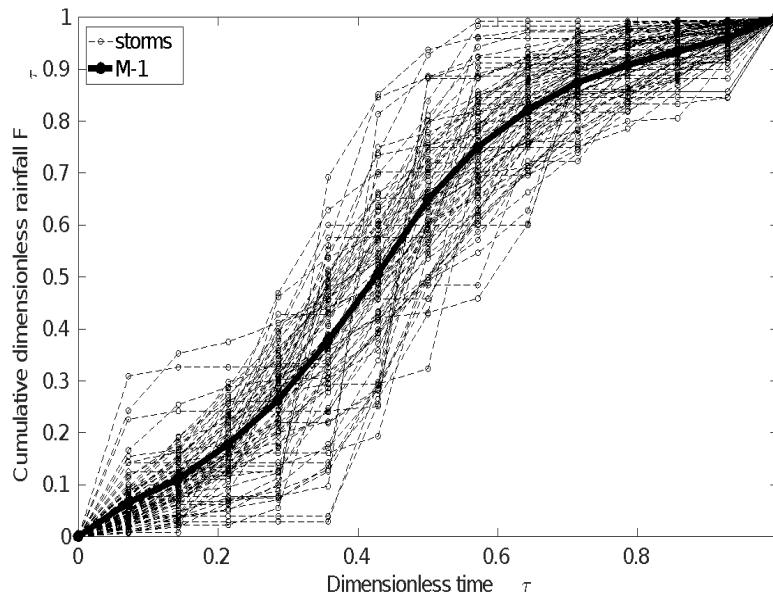


Figure 2.5 Intrinsic variability of storm intensities within a storm pattern. 86 storm patterns (dashed-lines) in an M-type representative pattern (solid-line).

Analogous to the previous section, a multivariate non-normal Monte Carlo simulation was proposed to generate storm patterns. However, the Monte Carlo simulation was modified to the following constraints: (1) unit sum, namely, the addition of the dimensionless hyetograph

intensities must be one: $P_1 + P_2 + \dots + P_m = 1.0$, where m was, as previously defined, equal to the number of ordinates that defines the dimensionless hyetograph; (2) non-negativity: $P_\tau \geq 0$ for $\tau = 1, 2, \dots, m$. These conditions assured that the dimensionless mass curve ordinates were bounded between 0 and 1.

The generation of storm patterns by using the Constrained Monte Carlo Multivariate method involves the following steps:

Step 1. Based on the database of the independent storms, the pattern for each storm is characterized and normalized by the corresponding duration and depth.

Step 2. Run the divisive K-means clustering method through the dimensionless mass curve to identify representative curves.

Step 3. Summarize relevant statistical information of each ordinate of the representative dimensionless hyetographs, determine statistical moments, and fit the data contained in each ordinate into a Johnson-type distribution.

Step 4. For each storm generated in the last section, randomly assign a type of storm from the six representative patterns identified in Step 2. Each generated storm is assumed to have the same probability of adopting any of the representative patterns. It has been observed that storm depth, duration, and patterns are correlated (Wu et al., 2006). However, the assumption mentioned here was adopted for simplicity in this study.

Step 5. Generate m ordinates of the dimensionless mass curves for the same number of storms in Step 4 by the Monte Carlo simulation procedure shown in Figure 2.3. This simulation is constrained to the following conditions:

$$P_1 + P_2 + \dots + P_i = 10 \quad i = 1 \dots m \quad (2.6)$$

$$P_i \geq 0 \quad i = 1 \dots m \quad (2.7)$$

2.6 Model performance

2.6.1 Analysis

The model proposed in this study was evaluated by comparing observed and simulated monthly averages for the simulated variables: time between storms, storm duration, and storm depth. The simulation period was 100 years of continuous precipitation data. Furthermore, an assessment of empirical and simulated frequency distributions for the month with the greatest

storm duration and depth illustrated the degree of approximation provided by the model formulation. The generated storm patterns contained in each of the selected clusters were summarized and compared with the six representative dimensionless mass curves shown in Figure 2.4. These comparisons allowed measuring the proper functioning of the model and documenting the overall performance of all components in the model. Hence, future research will provide substantial improvements in the model characterization and simulation. The goodness of fit between frequency distributions of the simulated storm variables and the corresponding empirical distribution was evaluated using the Kolmogorov-Smirnov (K-S) test and its probability value. Average monthly precipitation values for the selected period (March-August) were compared with the simulated monthly average, and difference and percent deviation are shown. For storms that span months, time between storms, storm duration, and storm depth for those cases were assigned to the month when the storms began. Table 2.1 shows the monthly averages at USGS-CHM at Charlotte-Mecklenburg, North Carolina.

Table 2.3. Comparison between observed and simulated time between storms at USGS-CHM at Charlotte-Mecklenburg.

Month	Time Between Storms			
	Average Observed (min)	Average Simulated (min)	Deviation (%)	K-S p-value
March	3930	3768	-4.1	0.054
April	4240	4259	0.5	0.053
May	6469	6306	-2.5	0.071
June	4206	4068	-3.3	0.088
July	3443	3382	-1.7	0.038
August	3903	4020	3.0	0.186
Average			-1.4	

2.6.2 Comparison of frequency distributions

The numerical distribution of time between storms was well simulated (Figure 2.6). The smallest K-S probability value belonged to July (0.038). However, most of the probability values were greater than the 0.05 significance value considered here, meaning that the null hypothesis that the simulated and the observed TBS values come from the same distribution cannot be rejected.

Average monthly measured and simulated time between storms values agreed well for most of the months (Table 2.3). The largest percent deviation occurred for March (-4.1%). Percent monthly deviations ranged from -4.1% to 3.0%. The excellent agreement of time between storms can be partly attributed to the characterization of dry times between storm events that can span months (Bonta, 2004).

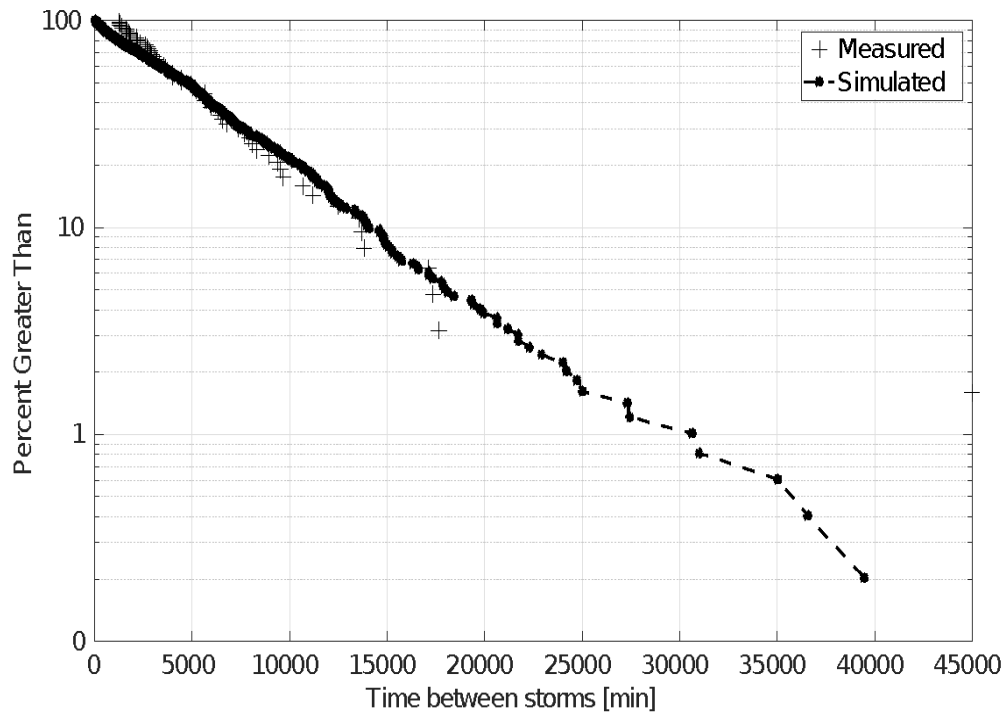


Figure 2.6 Comparison between simulated and observed time between storms at USGS-CHM at Charlotte-Mecklenburg, North Carolina, for the month with the longest average time between storms (May).

Storm duration was also reasonably well approximated (Table 2.4). Probability values were the smallest in March and the most significant percent deviation occurred in July, 0.035 and -14.9%, respectively. Percent monthly deviations varied from -14.9% in July to 2.2% in May, and average deviation was only -3.0%. Figure 2.7 shows the empirical and numerical Weibull distribution function superimposed, which indicates the good agreement between the observed and simulated values. While the empirical distribution was truncated at the minimum storm duration, the numerical distributions allowed for simulation of storms with shorter durations with a very low probability of occurrence.

Table 2.4. Comparison between observed and simulated storm duration and storm depth at USGS-CHM at Charlotte-Mecklenburg, North Carolina.

Month	Storm duration				Storm depth			
	Average Measured (min)	Average Simulated (min)	Deviation (%)	K-S p-value	Average Measured (mm)	Average Simulated (mm)	Deviation (%)	K-S p-value
March	295.1	290.8	-1.5	0.035	7.6	7.0	-7.9	0.056
April	350.5	346.8	-1.0	0.642	10.1	10.4	3.0	+0.000
May	737.2	753.2	2.2	0.576	16.0	16.5	3.0	0.606
June	214.2	210.6	-1.6	0.088	9.8	9.9	1.5	0.153
July	128.9	109.7	-14.9	0.118	8.1	7.2	-10.6	+0.000
August	253.8	251.3	-1.0	0.096	11.0	11.2	1.8	0.020
Average			-3.0				-1.5	

Average monthly storm depths were less well simulated by the model (Table 2.4). Although significant probability values were found smaller than 0.05, rejecting that the simulated and observed depths come from the same distribution, the percent deviations ranging from -10.6% in July to 3.0% in April and May seemed to be acceptable, which meant that the model captured ~90% of the monthly average for all months. For storm depth, July was the month with the most significant deviation percentage and the smallest probability value, -10.6% and +0.000, respectively. Similar to storm duration, the Gamma distribution used here to simulate storm depths might include storms with smaller depths than the resolution of the gauge (1 mm). However, those depths have a very low probability of occurrence throughout the simulation (Figure 2.8).

The considerable difference between observed and simulated storm duration and depths in July can be attributed to two main factors: (1) the distribution selected to simulate storm duration and depths does not describe the distribution of the observations of these two variables adequately. This claim is supported by the low K-S probability for both storm duration and storm depth. (2) Cases with probability values equal to +0.000 presented the least observed correlation between storm duration and depth. This may affect the simulation of these variables since the assumption of correlation is less appropriate for these months. The first factor, however, is more likely to cause the significant deviations in Table 2.4 than the second. Nonetheless, the simulated values still capture at least 90% of the variability of these correlated variables.

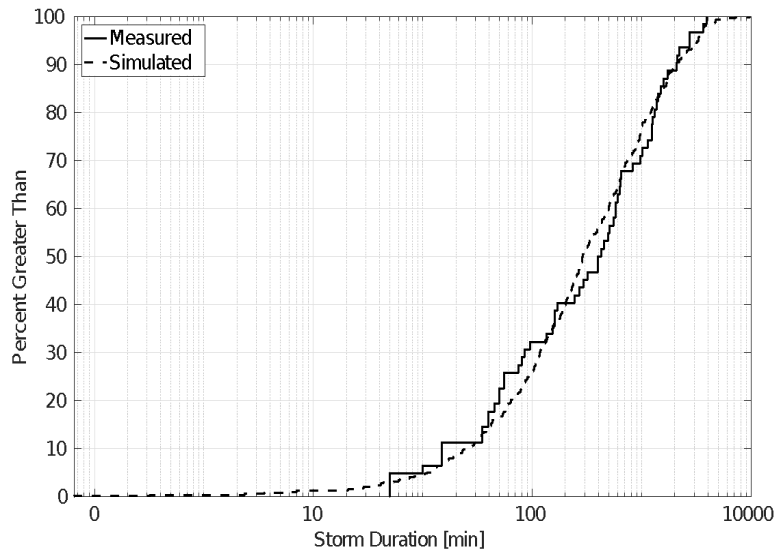


Figure 2.7 Comparison between simulated and observed storm duration at USGS-CHM at Charlotte-Mecklenburg for the month with longest storm duration (May).

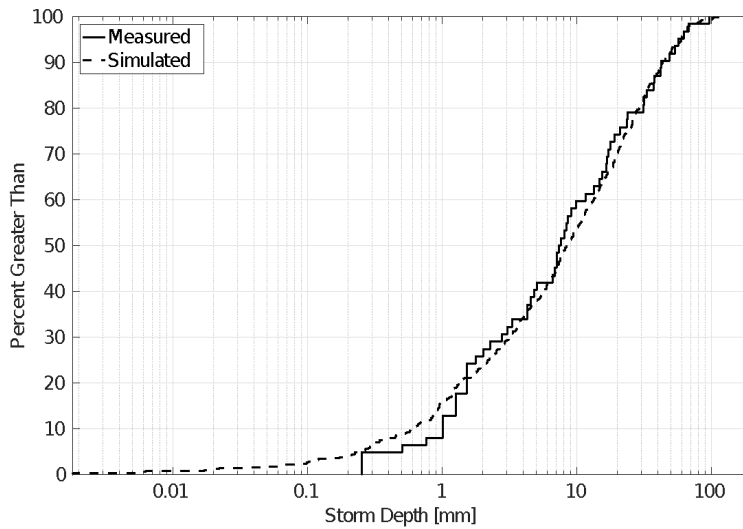


Figure 2.8 Comparison between simulated and observed storm depths at USGS-CHM at Charlotte-Mecklenburg for the month with greatest storm depths (May).

2.6.3 Model prediction efficiency

The formulation of the model proposed here replicated the first four storm characteristics with at least 80% accuracy for the USGS-CHM rain gauge at Charlotte-Mecklenburg, North Carolina. However, this accuracy highly depended upon the simulated number of years of continuous precipitation. To evaluate the minimum number of simulated years required to preserve

at least 80% of the monthly averages of time between storms, storm duration, and storm depth, 10 100-year precipitation series were generated, and then the average monthly precipitation values were calculated from the simulations (Figure 2.9a). The monthly time between storms averages were well simulated for 100 years of precipitation. Even though the accuracy percentage decreased for shorter periods of simulation, the replicability was still greater than 90%. There were no significant deviations for time between storms in series longer than 10 years. The more substantial discrepancy occurred for May when ten years of precipitation were generated with an accuracy of only ~70%. These deviations could be explained because May had the smallest number of storms in the precipitation record and a low K-S probability value (0.22, one-sample K-S test). At least 20 years of synthetic precipitation is recommended to capture the essential statistical features of the time between storms.

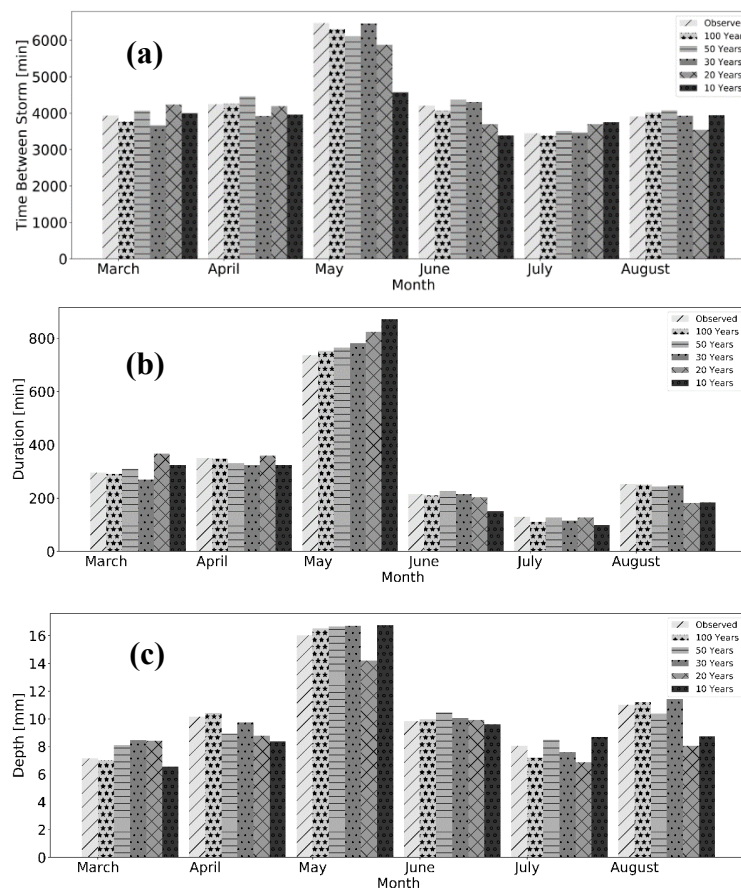


Figure 2.9 Model predictions for the three storm variables based on number of years generated: a) time between storms; b) storm duration; c) storm depth.

Monthly storm duration averages were also well simulated for a 100 year-precipitation period (Figure 2.9b). For periods longer than 20 years of precipitation, the replicability was still above 85% accuracy. The greatest differences ranged from ~18% to ~30% for 10 years of simulated precipitation for the months of May to August. The goodness of fit test (one-sample K-S test) revealed that these months had the lowest but still acceptable probability values that ranged from 0.27 to 0.77. To be able to capture the statistical features of storm durations, a series of precipitation greater than 20 years must be simulated.

Deviations ranging from ~18% to ~20% were found for 10 years and 20 years of precipitation depths, respectively (Figure 2.9c). These differences arise from two main possible sources. First, since the number of storms in May were the lowest, greater variability of the predicted variable is expected. This can be reduced once a longer observed series of precipitation data are used to parameterize the model. A second reason includes the variability introduced by the random sampling of the model, meaning that the high variability for the 20-year run for May in Figure 2.9c might decrease in a second 20-year run. Although the one-sample K-S test indicated that the generated depth data came from the associated Gamma distribution, series of precipitation longer than 20 years need to be simulated to capture enough statistical features of storm depths.

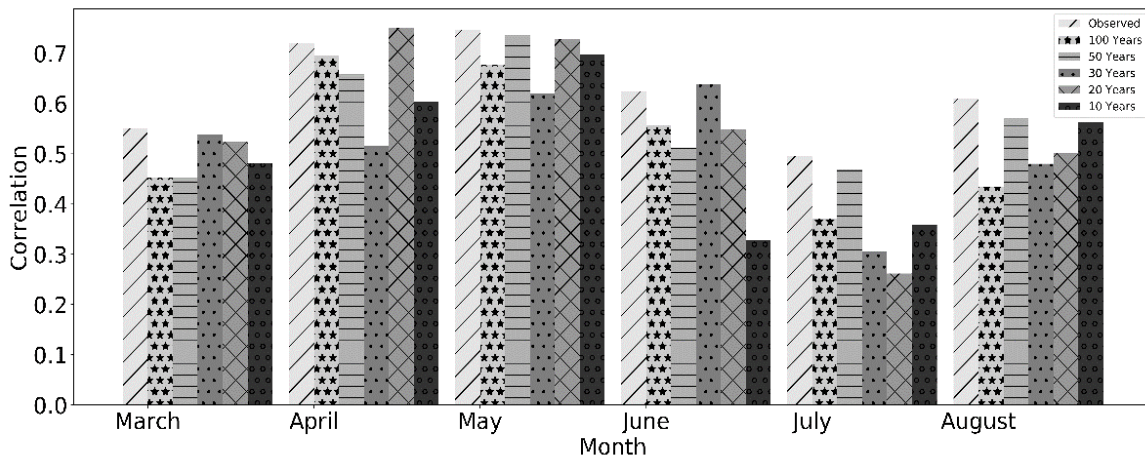


Figure 2.10 Correlation estimations between storm duration and storm depth based on the number of years generated.

Lastly, the correlation relationship between storm duration and storm depth was calculated for the simulated time periods as was done for the storm variables above (Figure 2.10). Correlation coefficients were not dependent on the number of years simulated but instead were a feature of the

random sampling involved in the formulation of the model. As a result, this was not affected by the length of the simulation but rather the multivariate model used here.

2.6.4 Storm pattern performance

Storm pattern (hyetograph) simulation was evaluated by aggregating all simulated storms to their corresponding representative pattern, and the coefficient of determination (R^2) was calculated to determine how accurately the simulated storms replicated the associated cluster. For the six representative storm patterns (D, M, C, and R types) identified in this study, the model showed good replicability (Figure 2.11) after aggregating the simulated dimensionless mass curves (SD, SM, SC, and SR types). The calculated R^2 values ranged from 0.922 to 0.997 (Table 2.5).

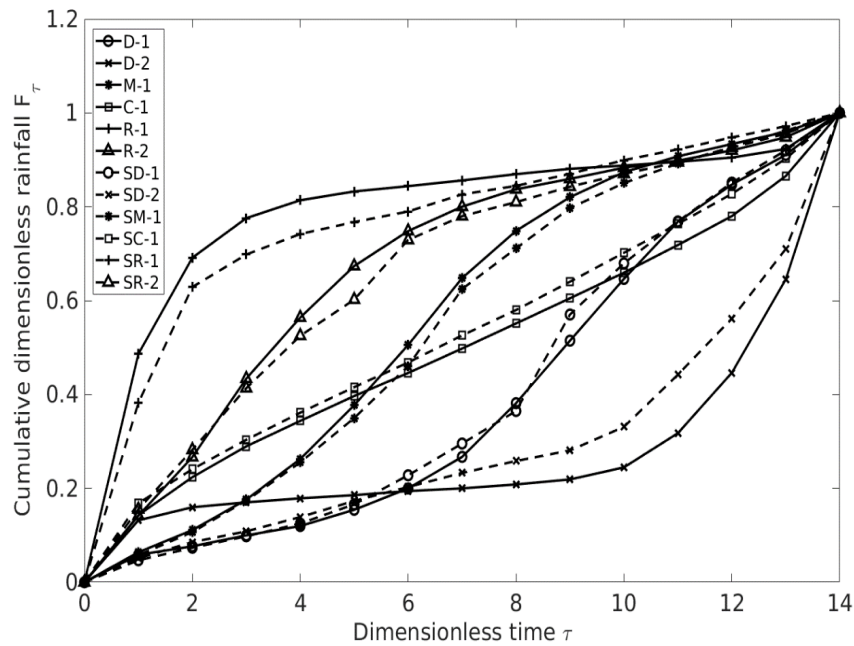


Figure 2.11 Average generated storms patterns (SD/SM/SC/SR type) vs. representative patterns (D/M/C/R type).

Table 2.5. Coefficient of determination (R^2) between observed and simulated storm patterns in Figure 2.11.

Storm Pattern	R^2
D-1 / SD-1	0.996
D-2 / SD-2	0.922
M-1 / SM-1	0.997
C-1 / SC-1	0.988
R-1 / SR-1	0.953
R-2 / SR-2	0.993

2.7 Conclusions

This article presents a practical stochastic model for generating sub-daily rainfall storms according to the statistics of rainstorm characteristics, including the storm occurrence, time between storms, storm duration, storm depth, and storm pattern. The occurrence of statistically independent storms involves the use of a Poisson distribution and a Monte Carlo simulation for non-normal and correlated time between storms, storm duration, and rainfall depth, as well as a constrained Monte Carlo simulation for the corresponding rainfall pattern or hyetograph.

The model was capable of reproducing the essential statistical features of the storm characteristics. The moment estimation efficiency showed that enough details of storm features were captured for periods of continuous precipitation longer than 20 years. However, a 100-year simulation had the best ability to replicate storm variables. The proposed model shows promising potential applicability to studies where high temporal resolution of rainfall is required, and within-storm intensities play an essential role in the estimation of other hydrological or environmental variables.

In this study, the simulation of storm events was based on monthly frequency distributions and 5-min increment precipitation data. The proposed model, however, can simulate storms over seasonal or annual frequency distributions and can be extended to shorter or longer time resolutions of precipitation data. Additionally, the assumption of equal probability for the storm pattern occurrence can be modified by assigning a proper distribution function to it. Lastly, as the model requires the specification of statistical properties of storm characteristics at a specific location, a proper regional analysis could extend its applicability to ungauged sites.

2.8 References

- Abas, N., Daud, Z. M., & Yusof, F. (2014). A comparative study of mixed exponential and Weibull distributions in a stochastic model replicating a tropical rainfall process. *Theoretical and Applied Climatology*, 118(3), 597–607. <http://doi.org/10.1007/s00704-013-1060-4>
- Acreman, M. C. (1990). A simple stochastic model of hourly rainfall for Farnborough, England. *Hydrological Science Journal*, 35(2), 119–148. <http://doi.org/10.1080/02626669009492414>
- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1491–1508. <http://doi.org/10.13031/2013.42256>
- Bonta, J. V. (2001). Characterizing and estimating spatial and temporal variability of times between storms. *Transactions of the ASAE*, 44(6), 1593–1601. <http://doi.org/10.13031/2013.7045>
- Bonta, J. V. (2003). Estimation of parameters characterizing frequency distributions of times between storms. *Transactions of the ASAE*, 46(2), 331–343. <http://doi.org/10.13031/2013.12984>
- Bonta, J. V. (2004). Stochastic simulation of storm occurrence, depth, duration and within-storm intensities. *Transactions of the ASAE*, 47(5), 1573–1584. <http://doi.org/10.13031/2013.17635>
- Bonta, J. V., & Rao, A. R. (1988). Factors affecting the identification of independent storm events. *Journal of Hydrology*, 98(3), 275–293. [http://doi.org/10.1016/0022-1694\(88\)90018-2](http://doi.org/10.1016/0022-1694(88)90018-2)
- Bonta, J. V., & Rao, A. R. (1992). Estimating peak flows from small agricultural watersheds. *Journal of Irrigation and Drainage Engineering*, 118(1), 122–137. [https://doi.org/10.1061/\(ASCE\)0733-9437\(1992\)118:1\(122\)](https://doi.org/10.1061/(ASCE)0733-9437(1992)118:1(122))
- Bras, R. L., & Rodriguez-Iturbe, I. (1976). Rainfall generation: A nonstationary time-varying multidimensional model. *Water Resources Research*, 12(3), 450–456. <https://doi.org/10.1029/WR012i003p00450>
- Chang, C. H., Tung, Y. K., & Yang, J. C. (1994). Monte Carlo simulation for correlated variables with marginal distributions. *Journal of Hydraulic Engineering*, 120(3), 313–331. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1994\)120:3\(313\)](https://doi.org/10.1061/(ASCE)0733-9429(1994)120:3(313))
- Chen, C., Wu, W., Zhang, B., & Sun, H. (2015). Correlated probabilistic load flow using a point estimate method with Nataf transformation. *Int. J. Electrical Power Energy Syst.*, 65, 325–333. <http://doi.org/10.1016/j.ijepes.2014.10.035>

- Chen, J., Brissette, F. P., & Leconte, R. (2010). A daily stochastic weather generator for preserving low-frequency of climate variability. *Journal of Hydrology*, 388(3), 480–490.
<http://doi.org/10.1016/j.jhydrol.2010.05.032>
- Chen, J., Brissette, F. P., & Leconte, R. (2012). WeaGETS – a Matlab-based daily scale weather generator for generating precipitation and temperature. *Procedia Environ. Sci.*, 13, 2222–2235. <http://doi.org/10.1016/j.proenv.2012.01.211>
- Chen, Y., Wen, J., & Cheng, S. (2013). Probabilistic load flow method based on Nataf transformation and Latin hypercube sampling. *IEEE Trans. Sustainable Energy*, 4(2), 294–301. <https://doi.org/10.1109/TSTE.2012.2222680>
- Clarke, R. T. (1998). *Stochastic processes for water scientists – Developments and applications*. New York, NY: John Wiley.
- Cronshey, R. (1986). *Urban hydrology for small watersheds*. Washington, DC: USDA, Soil Conservation Service, Engineering Division
- Der Kiureghian, A., & Liu, P. (1986). Structural reliability under incomplete probability Information. *J. Eng. Mechanics*, 112(1), 85–104. [https://doi.org/10.1061/\(ASCE\)0733-9399\(1986\)112:1\(85\)](https://doi.org/10.1061/(ASCE)0733-9399(1986)112:1(85))
- Eagleson, P. S., Fennessey, N. M., Qinliang, W., & Rodriguez-Iturbe, I. (1987). Application of spatial Poisson models to air mass thunderstorm rainfall. *J. Geophys. Res.:Atmospheres*, 92(D8), 9661–9678. <http://doi.org/10.1029/JD092iD08p09661>
- Ellouze, M., Abida, H., & Safi, R. (2009). A triangular model for the generation of synthetic hyetographs. *Hydrological Science Journal*, 54(2), 287–299.
<http://doi.org/10.1623/hysj.54.2.287>
- Feldman, A. D. (2000). *Hydrologic modeling system HEC-HMS: technical reference manual*. Davis, CA: U.S. Army Corps of Engineers, Hydrologic Engineering Center.
- Grace, R. A., & Eagleson, P. S. (1966). *Synthesis of short-time-increment rainfall sequences*. Report (Massachusetts Institute of Technology. Hydrodynamics Laboratory) No. 91.
- Hanson, C. L., Cumming, K. A., Woolhiser, D. A., & Richardson, C. W. (1994). Microcomputer program for daily weather simulation in the contiguous United States. *USDA–ARS Publ. ARS–114*, Washington, D.C.: USDA–ARS.
- Huff, F. A. (1967). Time distribution of rainfall in heavy storms. *Water Resources Research*, 3(4), 1007–1019. <http://doi.org/10.1029/WR003i004p01007>

- Karanki, D. R., Jadhav, P. A., Chandrakar, A., Srividya, A., & Verma, A. K. (2010). Uncertainty analysis in PSA with correlated input parameters. *Int. J. Syst. Assur. Eng. Manag.*, 1(1), 66–71. <http://doi.org/10.1007/s13198-010-0012-y>
- Keifer, C. J., & Chu, H. H. (1957). Synthetic storm pattern for drainage design. *J. Hydraulics Division*, 83(4), 1-25.
- Kottegoda, N. T. (1987). Fitting Johnson S_B Curve by the method of maximum likelihood to annual maximum daily rainfalls. *Water Resources Research*, 23(4), 728–732. <http://doi.org/10.1029/WR023i004p00728>
- Kottegoda, N. T., Natale, L., & Raiteri, E. (2014). Monte Carlo simulation of rainfall hyetographs for analysis and design. *Journal of Hydrology*, 519, 1–11. <http://doi.org/10.1016/j.jhydrol.2014.06.041>
- Lee, T., & Park, T. (2017). Nonparametric temporal downscaling with event-based population generating algorithm for RCM daily precipitation to hourly: Model development and performance evaluation. *Journal of Hydrology*, 547, 498–516. <http://doi.org/10.1016/j.jhydrol.2017.01.049>
- Li, D.-Q., Jiang, S.-H., Wu, S.-B., Zhou, C.-B., & Zhang, L.-M. (2013). Modeling multivariate distributions using Monte Carlo simulation for structural reliability analysis with complex performance function. *Proceedings of the Institution of Mechanical Engineers, Part O: J. Risk Reliability*, 227(2), 109–118. <http://doi.org/10.1177/1748006X13476821>
- Lin, F. R., Wu, N.-J., & Tsay, T.-K. (2017). Applications of cluster analysis and pattern recognition for typhoon hourly rainfall forecast. *Advances in Meteorology*, 2017, 5019646, 17 pp. <http://doi.org/10.1155/2017/5019646>
- Lu, Y., & Qin, X. S. (2014). Multisite rainfall downscaling and disaggregation in a tropical urban area. *Journal of Hydrology*, 509, 55–65. <http://doi.org/10.1016/j.jhydrol.2013.11.027>
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proc. 5th Berkeley Symp. on Mathematical Statistics and Probability*, 1, pp. 281–297.
- Marien, J. L., & Vandewiele, G. L. (1986). A point rainfall generator with internal storm structure. *Water Resources Research*, 22(4), 475–482. <http://doi.org/10.1029/WR022i00475>
- Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *J. Atmos. Oceanic Technol.*, 29(7), 897–910. <http://doi.org/10.1175/JTECH-D-11-00103.1>

- Nataf, A. (1962). Determination des distribution dont Les marges sont donnees. *Comptes Rendus de Lacademie Des Sciences*, 225, 42–43.
- Nicks, A. D., Lane, L. J., & Gander G. A. (1995). *Chapter 2. Weather generator*. In D. C. Flanagan, & M. A. Nearing (Eds.), USDA–Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation. NSERL Report No. 10. West Lafayette IN: USDA–ARS National Soil Erosion Research Laboratory
- Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., & Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, 470–471, 138–157. <http://doi.org/10.1016/j.jhydrol.2012.08.041>
- Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *J. Marketing Res.*, 20(2), 134-148. <http://doi.org/10.2307/3151680>
- Ramos, M. C. (2001). Divisive and hierarchical clustering techniques to analyse variability of rainfall distribution patterns in a Mediterranean region. *Atmos. Res.*, 57(2), 123–138. [http://doi.org/10.1016/S0169-8095\(01\)00065-5](http://doi.org/10.1016/S0169-8095(01)00065-5)
- Raudkivi, A. I., & Lawgun, N. (1970). Synthesis of urban rainfall. *Water Resources Research*, 6(2), 455–464. <https://doi.org/10.1029/WR006i002p00455>
- Restrepo-Posada, P. J., & Eagleson, P. S. (1982). Identification of independent rainstorms. *Journal of Hydrology*, 55(1), 303–319. [http://doi.org/10.1016/0022-1694\(82\)90136-6](http://doi.org/10.1016/0022-1694(82)90136-6)
- Richardson, C. W., & Wright, D. A. (1984). *WGEN: A model for generating daily weather variables*. ARS-8. Washington, DC: USDA, ARS.
Retrieved from <https://archive.org/details/CAT85844304>.
- Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1987a). Some models for rainfall based on stochastic point processes. *Proc. Royal Soc. London. A. Math., Phys. Sci.*, 410(1839), 269–288. <http://doi.org/10.1098/rspa.1987.0039>
- Rodríguez-Iturbe, I., de Power, B. F., & Valdes, J. B. (1987b). Rectangular pulses point process models for rainfall: Analysis of empirical data. *J. Geophys. Res.: Atmos.*, 92(D8), 9645-9656. <http://doi.org/10.1029/JD092iD08p09645>
- Rodriguez-Iturbe, I., Gupta, V. K., & Waymire, E. (1984). Scale considerations in the modeling of temporal rainfall. *Water Resources Research*, 20(11), 1611–1619. <https://doi.org/10.1029/WR020i011p01611>

- Schamm, K., Ziese, M., Becker, A., Finger, P., Schneider, U., Schroder, M., & Stender, P. (2014). Global gridded precipitation over land: a description of the new GPCC First Guess Daily product. *Earth Syst. Sci. Data*, 6(1), 49–60. <https://doi.org/10.5194/essd-6-49-2014>
- Semenov, M. A., & Barrow, E. M. (2002). *A stochastic weather generator for use in climate impact studies. User Manual*. Hertfordshire, UK. Retrieved from: <http://resources.rothamsted.ac.uk/sites/default/files/groups/mas-models/download/LARS-WG-Manual.pdf>
- Stone, J. J., Lane, L. J., Shirley, E. D., & Hernandez, M. (1995). *Chapter 4. Hillslope surface hydrology*. In D. C. Flanagan & M. A. Nearing (Eds.), USDA-Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation, NSERL Report No. 10. West Lafayette, IN: USDA-ARS National Soil Erosion Research Laboratory.
- Tank, A. M., Wijngaard, J. B., Konnen, G. P., Bohm, R., Demaree, G., Gocheva, A., ... Petrovic, P. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.*, 22(12), 1441–1453. <http://doi.org/10.1002/joc.773>
- Todorovic, P., & Yevjevich, V. (1969). *Stochastic process of precipitation*. 35,1-61. Fort Collins: Hydrology Papers, Colorado State University. Retrieved from: https://dspace.library.colostate.edu/bitstream/handle/10217/61315/HydrologyPapers_n35.pdf?sequence=1
- Tyron, R. C., & Bailey, D. E. (1970). *Cluster Analysis*. New York, NY: McGraw-Hill.
- Wenzel, J. H., & Voorhees, M. L. (1981). An evaluation of the urban design storm concept. *Res. Rep. 164*. University of Illinois, Water Resources Center.
- Wischmeier, W. H., & Smith, D. D. (1978). *Predicting rainfall erosion losses-a guide to conservation planning*. USDA, ARS, Agric. Handbook 537, Washington DC: USDA, ARS.
- Wu, S. J., Yang, J. C., & Tung, Y. K. (2006). Identification and stochastic generation of representative rainfall temporal patterns in Hong Kong territory. *Stochastic Environ. Res. and Risk Assessment*, 203 (171–183). <http://doi.org/10.1007/s00477-005-0245-5>
- Xiao, Q. (2014). Evaluating correlation coefficient for Nataf transformation. *Probabilistic Eng. Mech.*, 37, 1–6. <https://doi.org/10.1016/j.probengmech.2014.03.010>
- Yen, B. C., & Chow, V. T. (1980). Design hyetographs for small drainage structures. *J. Hydrol. Div. 106 (HY6)*, 1055–1076.

3. IMPROVEMENT OF THE WATER EROSION PREDICTION PROJECT (WEPP) MODEL FOR QUANTIFYING FIELD SCALE SUBSURFACE DRAINAGE DISCHARGE

Abstract

In the poorly drained regions of the world, subsurface drainage systems are required to remove excess water for crop growth. Plastic drains alter a field's hydrology by lowering the water table, reducing surface ponding, and reducing surface runoff. One of the significant concerns with the use of subsurface drainage systems is adverse environmental effects because of the modification of the soil water dynamics. Some effects include the reduction of ecological services since wetlands change to croplands, water quality concerns, particularly sediment, nitrogen, and phosphorus losses in agricultural subsurface discharge water, as well as changes in the volume and timing of off-site discharges. Hydrological simulation models predict surface and artificial subsurface flow at different scales. Often in these models, Hooghoudt-based expressions are adapted in their internal algorithms. In this study, the Water Erosion Prediction Project (WEPP) model, developed by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) for soil and water conservation planning activities, was tested and improved to simulate surface and subsurface discharges. The modified WEPP model was tested and validated on an extensive dataset collected at four experimental sites managed by USDA-ARS within the Lake Erie Watershed. Predicted drainage discharges show Nash-Sutcliffe Efficiency (NSE) values ranging from 0.50 to 0.70, and Percent Bias ranging from -30% to +15% at daily and monthly resolutions. Evidence suggests that the WEPP model can be used to produce reliable estimates of subsurface flow with minimum calibration. Future work includes the extension of the model for quantifying subsurface drainage under controlled water table and watershed-scale simulations.

3.1 Introduction

Humid areas in Northern Europe, Canada, and the Midwest U.S are artificially drained by lowering the water table of perennially or seasonably wet soils (Pavelis, 1987; Gilliam et al., 1999). In the humid Upper Midwestern portion of the U.S., about 37% of the land (20.6 million ha) has been artificially drained to transform wet soils into highly productive cropland (Zucker and Brown,

1998). Drainage impacts hillslope and watershed hydrology as it increases the water storage capacity and improves physical structure of the upper layers of the soils (Skaggs and Broadhead, 1982; Fraser and Flemming, 2001). Higher storage capacity reduces surface runoff and promotes infiltration (Skaggs et al., 1987; Robinson and Rycroft, 1999). Research shows that artificially drained soils act as a buffer for rainfall and spread runoff over longer time periods. Thus, peak flows decrease, and less flooding occurs as a result (Robinson, 1990; Konyha et al., 1992; Skaggs et al., 1994; Schilling and Helmers, 2008; Henine et al., 2010).

Crop production benefits from subsurface drainage since it affects physical and chemical soil properties and strongly influences crop growth and quality (Messing and Westrom, 2006; Vopravil et al., 2017). As the goal of proper subsurface drainage is to remove excess gravitational water, consisting of shallow groundwater flow and water moving downward from upper soil horizons, constant aeration of the crop rooting zone promotes root health and microbial activity, which are important for maximizing nutrient availability (Kladivko et al., 1999; Tiemeyer et al., 2006). Additionally, artificial drainage allows earlier planting, which increases the yield potential for crops (Kornecki and Fouss, 2001). Research has additionally demonstrated that subsurface drainage improves workability, reduces the potential for soil compaction and decreases surface ponding (Aldabagh, 1971; Kornecki and Fouss, 2001).

Mathematical formulations have been developed to predict subsurface flow and associated transport focused on determining the optimal size and spacing of pipes in the field under steady or non-steady flow conditions. Most investigations assume systems with parallel drainage and equally spaced pipes (Wesseling, 1964; Kirkham, 1966; van Schilfgaarde, 1970), with Hooghoudt's equation (Hooghoudt, 1940) applied to approximate flow into two parallel drains. However, when parallel spacing is not a suitable assumption (e.g., fields with irregular systems), the actual spacing is substituted by an effective spacing, which approximates the system to equivalent parallel drains of the same length. Effective spacing is not site-specific but somewhat depends on the soil type (Kurien et al., 1997) and accretion rate (Northcott et al., 2001).

Hydrological simulation models estimate subsurface flow to drains at different scales. These models include the large scale Soil and Water Assessment Tool (SWAT, Arnold et al., 2012), field-scale models such as DRAINMOD (Skaggs, 2012c), GLEAMS (Leonard et al., 1987), and ADAPT (Gowda et al., 2012), as well as the multiscale Water Erosion Prediction Project (WEPP, Flanagan and Nearing, 1995; Flanagan and Livingston, 1995) model. Other models

include mathematical algorithms such as DRENAFEM (Castanheira and Santos, 2009) and HYDRUS2D (Simunek et al., 1999). Although the latter two models are not hydrologic formulations, they offer a more complex and complete representation of the transient saturated/unsaturated vertical flow in the soil profile. In practice, however, the hydrological models previously mentioned are solely based on the Kirkham (Kirkham, 1966) and Hooghoudt's equation (Hooghoudt, 1940). At watershed and hillslope scales, respectively, SWAT and DRAINMOD have been extensively validated, whereas GLEAMS, ADAPT, and WEPP have limited evidence of their success for subsurface drainage estimation.

As subsurface drainage influences infiltration and surface runoff, it is an important factor in the soil erosion process. Drainage reduces surface runoff, but questions abound on its influence on sediment transport through hillslopes and watersheds. Skaggs et al. (1982) and Maalim and Melesse (2013) used the DRAINMOD and WEPP models, respectively, to estimate average annual rates of erosion. They found that annual rates of erosion are up to 10 times lower when subsurface drainage is used as a practice to control water table depths. However, none of these predictions were validated with observed data. From a modeling standpoint, DRAINMOD has been widely applied in the United States for subsurface flow prediction. However, it neglects the effects of management operations, agricultural practices, and subsurface drainage on hillslope hydrology, which are conditions that greatly influence soil erosion by water.

Management practices, such as soil disturbance by tillage operations, residue management and decomposition, and soil erosion control practices, are incorporated in the WEPP model (Flanagan et al., 2007), an environmental model for soil erosion prediction used by the United States Department of Agriculture (USDA) in soil and water conservation planning activities. WEPP includes mechanisms such as weather generation, winter processes, irrigation, hydrology, plant growth, management operations, and erosion for both hillslope and watershed scales (Migliaccio and Srivastava, 2007). However, model inconsistencies have been found when validating the subsurface drainage algorithms (Oztek et al., 2004). Such discrepancies include overprediction of daily runoff for daily storms, large deviations between observed and simulated subsurface flow conditions, and poor representation of the water table depth variations. Although these inconsistencies were addressed by comparing the model predictions with those using DRAINMOD, there is no validation of the model using observed field drainage data and no updates have been included in the current public version of the WEPP model (v2012.8).

A comprehensive understanding of the hydrology of subsurface drainage fields and its impact on sediment delivery and transport are major knowledge gaps (Sims et al., 1998; King et al., 2014a,b) that restrict suitable cropland management strategies, improved water quality, soil erosion mitigation, and selection and implementation of best management practices. The goal of this study was to evaluate and enhance the WEPP model subsurface drainage routines to provide a more realistic and holistic representation of the conditions observed in agricultural fields under subsurface drainage systems. The model improvements will not only extend the applicability of the current version of the WEPP model to artificially drained croplands and contribute to reducing the gap seen between agricultural practices and subsurface drainage but will also better account for these effects on soil erosion predictions. A comprehensive dataset of hourly weather variables and daily subsurface flow collected at four experimental sites managed by the USDA-ARS in the Lake Erie watershed were used for testing and validation as part of this study. Particularly, this research provides details on 1) subsurface drainage prediction calculated by the current version of the WEPP model (v2012.8), 2) modifications of the source code in subsurface drainage, percolation, water balance, and infiltration algorithms, and the incorporation of soil water movement by capillary rise from the water table into the model code, and 3) evaluation, calibration, and validation of the new subsurface drainage algorithms implemented in the model.

3.2 Materials and Methods

3.2.1 Model description

The WEPP model (Flanagan et al., 2012, Ascough et al., 2013) is a soil erosion prediction tool developed since 1985 to replace the well-known empirical Universal Soil Loss Equation (USLE, Wischmeier and Smith, 1978). Contrary to the USLE, the WEPP model is a process-based technology that estimates the spatial and temporal distribution of soil loss and sediment deposition at both hillslope and watershed scales (Flanagan et al., 2007). It simulates continuous processes with a daily time step including weather generation, hydrologic processes, soil physics, plant growth, and erosion mechanics. The model water balance, based on soil water content in the root zone, includes processes such as infiltration, runoff, percolation, soil and residue evaporation, plant transpiration, lateral flow, snowmelt, subsurface drainage, and seepage. For more details on those processes and the model structure, see the model documentation (Flanagan and Nearing, 1995).

The current subsurface drainage component of WEPP (Savabi, 1993) simulates subsurface flow to pipes or ditches. Drainage fluxes are estimated for a given water table depth with the steady-state Hooghoudt equation (Bouwer and van Schilfgaarde, 1963) as follows:

$$Q_d = \frac{4K_e m(2d_e + m)}{L^2} \quad (3.1)$$

where K_e is the soil's equivalent lateral hydraulic conductivity (cm h^{-1}), m is the midpoint water table elevation above the pipes (cm), d_e is the equivalent depth from the drain to the impervious layer (cm), and L is the drain spacing (cm). The equivalent depth in the equation can be calculated with the Moody (1967) equations:

$$d_e = \frac{d}{1 + \frac{d}{L} \left[\frac{8}{\pi} \ln \left(\frac{d}{r} \right) - 3.4 \right]} \quad \frac{d}{L} < 0.3 \quad (3.2)$$

$$d_e = \frac{L\pi}{8 \left[\ln \left(\frac{L}{r} \right) - 1.15 \right]} \quad \frac{d}{L} \geq 0.3 \quad (3.3)$$

where d is the distance between the impervious layer and the drain (cm), and r is the effective drain radius (cm).

Equation (3.1) in the WEPP model (v.2012.8) calculates the drainage rates when water table depth is located at and below the surface. For the latter condition, the Hooghoudt equation has been widely applied and validated (Skaggs, 2012). However, for the water table at the surface, the Kirkham equation (1957) has shown a more accurate representation. This condition and the inconsistencies previously mentioned were addressed by modifications in the model source code as described in the next subsections. The WEPP model (v2012.9), the enhanced version of the model, is used throughout this research for comparison purposes with the current version of the model (v2012.8).

3.2.1.1 Subsurface drainage modifications

For ponding of water at the soil surface, drainage flow in the WEPP model (v2012.9) is computed as (Kirkham, 1957):

$$Q_d = 4\pi K_e(d_p + d_d + r)/gL \quad (3.4)$$

where d_p is the depth of ponded water (cm), d_d is the depth to the drain (cm), and g is the estimated Kirkham's coefficient (Kirkham, 1957) calculated as:

$$g = 2 \ln \left[\frac{\tan \frac{(2d-r)\pi}{4h}}{\tan \frac{\pi r}{4h}} \right] + 2 \sum_{m=1}^{\infty} \ln \left[\left(\frac{\cosh \frac{\pi mL}{2h} + \cosh \frac{\pi r}{2h}}{\cosh \frac{\pi mL}{2h} - \cosh \frac{\pi r}{2h}} \right) \left(\frac{\cosh \frac{\pi mL}{2h} - \cosh \frac{\pi r}{2h}}{\cosh \frac{\pi mL}{2h} + \cosh \frac{\pi r}{2h}} \right) \right] \quad (3.5)$$

where h is equal to the hydraulic head (cm). Drainage at the soil surface and below are restricted to the drainage coefficient (d_c) of the system (cm/day).

3.2.1.2 Soil water characteristics modifications

The soil water distribution in the enhanced version (v2012.9) of the WEPP model is assumed to be divided into wet and dry zones. A wet zone develops from the water table level up to the root zone and possibly extending to the soil surface, whereas a dry zone occurs when water is removed from the root zone storage once the maximum capillary rise from the water table is not sufficient to supply the evapotranspiration amount (Skaggs, 2012c). A drained-to-equilibrium condition is assumed to control the water distribution in the soil domain (Skaggs, 1980), which represents a major change in the percolation routine in the WEPP model (v2012.8), which originally used storage routing techniques to percolate water content exceeding the soil field capacity.

The soil water distribution above the water table and the volume of water-free pore space is calculated at each time step of the model (Figure 3.1a). Inputs required to make these calculations include the soil water retention curve (or soil moisture characteristic) for each soil layer (Figure 3.1b), the volumetric water content lower limit available to the crop or permanent wilting point for the soil layer where roots are located, and the maximum upward water movement vs. water table depth relationships (Figure 3.1c).

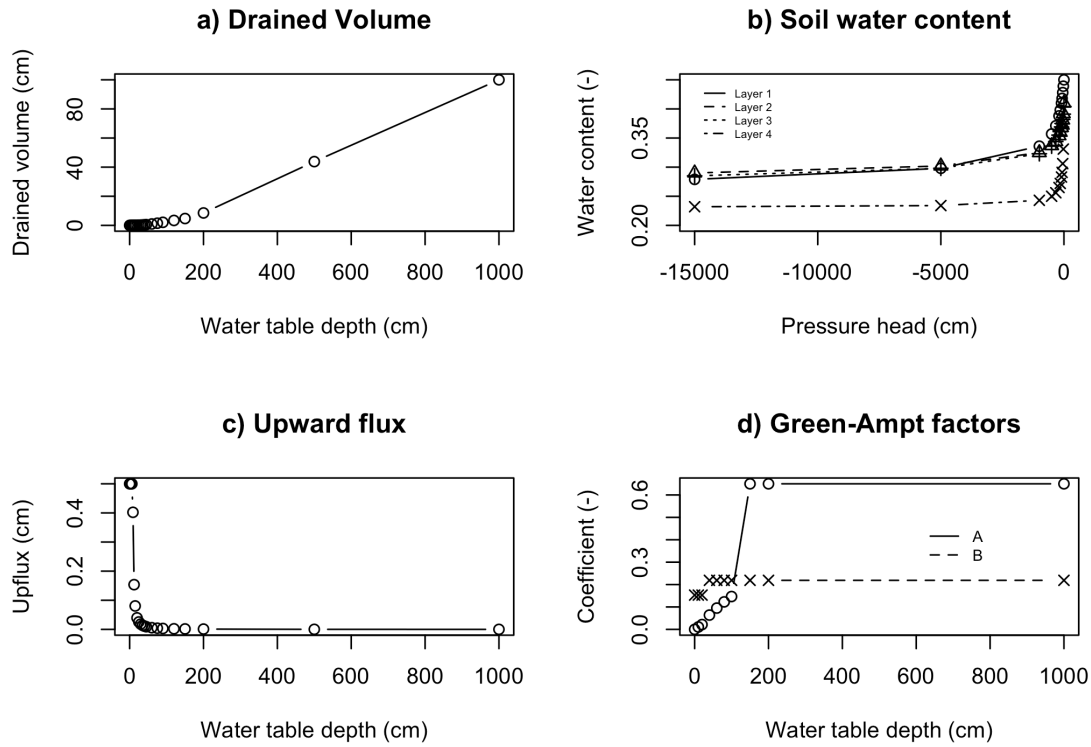


Figure 3.1 Input relationships for the WEPP model (v2012.9)

3.2.1.3 Infiltration modifications

WEPP model v2012.8 calculates the cumulative infiltration using the Green-Ampt Mein-Larson (Mein and Larson, 1973) model for unsteady rainfall. The Mein and Larson (1973) results showed good agreement with rates obtained from solutions of the Richard's equation for a wide range of soil types. In this study, a simplistic solution of the Green-Ampt model was adopted under the assumption that the infiltration rate vs cumulative infiltration relationship is independent of the injection rate, and the only input parameters are those reflecting the initial conditions of the soil water content distribution (Skaggs, 1980).

The Green-Ampt equation adopts the following form (Green and Ampt, 1911):

$$f = K_s \left(1 + \frac{MS_w}{F} \right) \quad (3.6)$$

where K_s is the vertical soil hydraulic conductivity in the transmission zone, M is the fillable porosity, S_w is the effective suction at the wetting front, and F is the cumulative infiltration. The original form of the Green-Ampt expression was derived under total saturation behind the wetting

front. However, this assumption was relaxed by Philip (1954) so that K_s is expected to be less than the saturated hydraulic conductivity.

For a given soil with an initial water content, Equation (3.6) can be written as,

$$f = \frac{A}{F} + B \quad (3.7)$$

where the parameters $A = K_s MS_{av}$ and $B = K_s$ depend on soil properties, initial water content distribution, surface conditions, and water table depth. As in other Green-Ampt formulations, infiltration rates are set equal to rainfall rates until infiltration capacity is exceeded.

For shallow water table soils, a typical initial condition is an unsaturated profile in equilibrium with the water table. Skaggs (1980) demonstrated that infiltration rate vs cumulative infiltration is nearly independent of the rainfall rate for a wide range of sandy loam soils under different initial water table depths. The latter is consistent with the results obtained by Mein and Larson (1973) for the same soils. As a result, the parameters A and B can be derived using regression methods from numerical calculations from the Richards equation or experimental data (Skaggs, 1980).

In this study, the WEPP model was modified to input a set of parameters A and B versus water table depths (Figure 3.1d). For water table depths in between input values, parameters can be calculated by linear interpolation. Skaggs (1980) presents detailed methods to calculate the parameters A and B required for the WEPP model from both experimental and numerical-derived data. The numerical method described here was implemented and adopted in the WEPP model (v2012.9) using the input parameters derived from the application of a soil properties generator model (Schaap et al. 2001) based on soil properties used in the WEPP model (v2012.8) so that no additional inputs are required.

3.2.1.4 Water balance

A primary WEPP model water balance is calculated at daily time increments (Flanagan and Nearing, 1995). In this study, a secondary water balance within the soil profile was performed on an hourly basis to capture the variation of drainage fluxes in a more realistic fashion. A section of unit surface area extending from the impermeable layer to the soil surface and located midway

between parallel drains is idealized. The water balance within the soil profile is then expressed by the following:

$$\Delta V_p = Q_{dh} + ET_h - IF_h \quad (3.8)$$

where ΔV_p is the variation in water-free pore space in the soil profile (cm), Q_{dh} is the drainage flux leaving the soil section (cm), ET_h is the evapotranspiration including soil and residue evaporation, and plant transpiration (cm), and IF_h is the infiltrated water (cm). As evapotranspiration rates are calculated daily in the WEPP model (v2012.8), daily evapotranspiration is distributed during a 12-h period (6 am to 6 pm) to ensure an hourly water balance. It is assumed that evapotranspiration is only significant during daylight. Additionally, a computational sub-time step equal to 0.05 h is used to calculate infiltration when rainfall intensity exceeds the infiltration capacity. Once change in air volume in the profile is calculated, the water distribution in the soil is estimated at each time step, and each component of the primary water balance is updated to produce daily outputs.

3.2.1.5 Upward water movement

Capillary rise from the water table depth supplies water to the plant roots. As a result, it is an important process in models simulating water management systems such as artificial subsurface drainage. As the water table rises, water can be transmitted by capillary rise to the root zone to satisfy the ET demand of the crops. Capillary rise affects the water table drawdown since soil profile water is lost by upward flux. Consequently, upward flux has a direct effect on evapotranspiration. Originally, the WEPP model did not take upward water movement into account. Modifications to the model were done to determine the steady upward flux from different water table depths for the soils examined in this study. The method used here was illustrated by Memon et al. (1986), where the concept of matrix flux potential (MFLP) was defined by Shaykewich and Stroosnijder (1977):

$$MFLP(h) = \int_{h_o}^{h_{max}} K(h)dh \quad (3.9)$$

where h_{max} is the maximum pressure head (cm) allowed at the center of the effective root zone, and h_o is the pressure head at the water table (cm). Taylor and Ashcroft (1972) presented values

of h_{max} for a wide range of crops. These values vary with crop growth. Using the integral from the Darcy-Buckingham equation as:

$$(q + K(h)) \int_{z_0}^z dz = \int_{h_0}^h K(h) dh \quad (3.10)$$

where q is the rate of upward movement (cm/day), and z is the water table depth (cm). Substituting Equation (3.10) into Equation (3.9), a finite difference form of Equation (3.9) is obtained:

$$z_{i+1} = \frac{MFLP_{i+1} - MFLP_i}{q + [K(h_{i+1}) + K(h_i)]/2} \quad (3.11)$$

For a range of soil matric potentials (i.e., $h=-500$ cm to $h=0$, soil matric potentials at the effective root zone and at the water table, respectively) an assumed flux upward, z_{i+1} , can be estimated. The accumulated value of z_{i+1} represents the water table depth as the corresponding upward water movement occurs. These calculations are repeated for several values of q , which represents theoretical values of the maximum amount of water supplied by capillary rise of the water table that can be used for plant transpiration.

Calculations of the maximum upflux of water require previous knowledge of the soil unsaturated hydraulic conductivity when the water table is at h (cm). The unsaturated hydraulic conductivity, $K(h)$, for each soil layer is estimated using the Mualem and Dagan (1978) model:

$$K(\theta) = S_e^n \frac{\int_0^\theta \frac{(\theta - \epsilon)}{h^{2+b}} d\epsilon}{\int_0^{\theta_s} \frac{(\theta - \epsilon)}{h^{2+b}} d\epsilon} \quad (3.12)$$

where n is the pore size distribution, b is the soil tortuosity, S_e is the effective saturation defined as θ/θ_s , θ is the effective water content defined as water content minus residual water content θ_r , and θ_s is the water content at saturation. This model ensures that $K(\theta_s)$ corresponds to the vertical saturated hydraulic conductivity, K_s . Soil water retention curves were fitted using the van Genuchten (1980) model:

$$S_e = [1 + (\alpha h)^n]^{-m} \quad (3.13)$$

where α is the air entry value, and m is a fitting parameter set to one. All parameters required by the Mualem and Dagan and van Genuchten models are numerically calculated from soil texture, soil density, field capacity and wilting point.

3.2.1.6 Additional modifications

Supplementary modifications to the WEPP model (v2012.8) source code were performed to completely adapt the model to the new subsurface drainage routines and all the changes previously discussed. WEPP input routines were modified to allow inclusion of field and laboratory measurements such as soil water retention curves, upward movement of water or capillary rise, lateral hydraulic conductivity, and relevant information used for drainage flux estimation. Lastly, output routines were also modified to provide a temporal description of subsurface drainage. A new output file now includes information on water table level, subsurface drainage, and water-free pore space.

3.2.2 Experimental sites

A field selected to perform an uncalibrated validation of the current version of the WEPP model (v2012.8) and its latter modifications, from now on called “ES(1)”, is located near Hamilton in Steuben County, northeast Indiana (Figure 3.2). The USDA-ARS National Soil Erosion Research Laboratory (NSERL) at Purdue University, West Lafayette, Indiana collects crop management and subsurface drainage data to assess water and soil conservation practice implementation effects. Dominant soils in the area of study are Blount loam and Pewamo silty clay loam (Table 3.1). The drainage system consists of corrugated plastic drains separated every 10 m at a depth of 1 m from the soil surface. Meteorological data and drainage fluxes were measured from 2015 to 2017 at the experimental site (Table 3.2). Drainage flow characteristics such as discharge and speed were measured with a Thelmar 12” Compound Weir and a Teledyne ISCO TieNet 350 Area Velocity Sensor attached to a Teledyne ISCO 330 flowmeter, respectively. Crop and soil management at ES(1) are a two-year corn-soybean rotation and annual primary and secondary tillage operations, respectively (See Crop Management section and Table 3.2).

For calibration and validation evaluations, fields, named here “ES(2-4)” were selected to demonstrate the efficiency of the enhanced version of the WEPP model (v2012.9) after the

modifications proposed in this paper. Sites ES(2-4) are located in northwest Ohio and managed by local farmers (Figure 3.2). The area has dominant soils such as Hoytville silty clay loam and Hoytville clay loam (See Table 3.1 for physical properties). For these sites, pipe spacing ranges from 6 to 10 m, at depths varying from 0.76 to 1 m (See Table 3.2). Drainage fluxes and precipitation from 2014 to 2018 were measured at the experimental sites. Additional weather variables were obtained from the Northwest Station (NW-Station) maintained by the Ohio Agricultural Research and Development Center (OARDC) (<https://www.oardc.ohio-state.edu/weather1/>). Crop management data are detailed in the Crop Management section and Table 3.2.

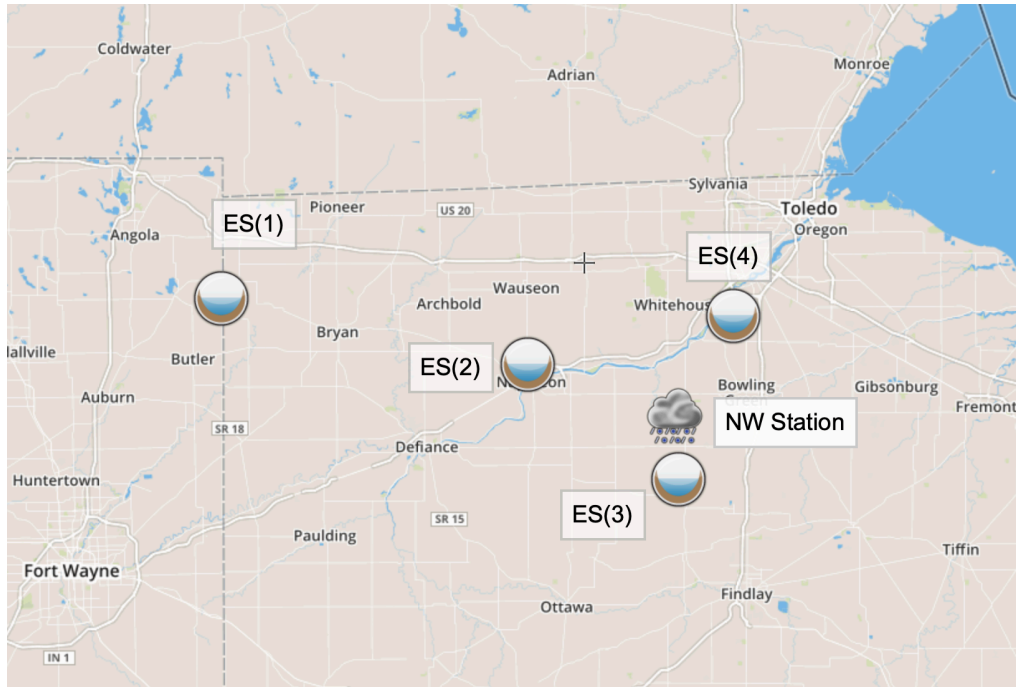


Figure 3.2 Experimental site locations. NW Station: Northwest station managed by OARDC; ES(1): Uncalibrated validation site; ES(2-4): Calibrated/Validated sites.

3.2.3 Model input and calibration/validation

The hillslope WEPP model simulation uses four basic inputs, 1) climate inputs describing temporal variability of precipitation, temperature, radiation, and wind speed and direction; 2) hydraulic and physical properties of single or layered soils (i.e., depths, soil textures, effective hydraulic conductivity, etc.); 3) detailed description of the steepness along the hillslope profile; and 4) plant/crop management data. Purposely, the ES(1) site was only used for an uncalibrated

validation, and calibration was not performed since the goal was to evaluate the accuracy of the current version and the improved version of the WEPP model under the same input conditions and parameters. To evaluate the modifications of the WEPP model (v2012.9) under different soil and drainage conditions, fields ES(2-4) were used for calibration and validation. A sensitivity analysis using the Model-Independent Parameter Estimation & Uncertainty Analysis (PEST) (Doherty, 2015) software revealed that the saturated vertical hydraulic conductivity and saturated lateral hydraulic conductivity were the most sensitive parameters. Calibration of those parameters was performed until the best efficiency of the model to predict observed drainage fluxes was obtained. DRAINMOD-based model studies have demonstrated that model response is sensitive to multiple soil parameters among other variables (i.e. monthly potential evapotranspiration), and usually calibrated. However, in this study, only the first two most sensitive parameters were adjusted to maintain simplicity and practicability of the model (Singh et al., 2006; Singh et al., 2007; Morrison et al., 2014; Qi et al., 2015, Liang et al., 2018). After the model calibration, a one-year period was selected for validation purposes. For the simulations, one-year (2017) and two-year (2014-2015) warmup periods were used for WEPP to develop reliable initial conditions for the ES(1) and ES(2-4) sites, respectively.

3.2.4 Meteorological data

Meteorological data including hourly precipitation and relative humidity, maximum and minimum daily air temperatures, and wind speed and direction measured at the site ES(1) were collected as part of the Conservation Effects Assessment Project (CEAP). Similarly, weather data at ES(2-4) were collected at each experimental site. All these data were used to calculate potential evapotranspiration using the Penman-Monteith method (Allen et al., 1998). Missing values in the climatic data were obtained from the OARDC weather system within a 30-km radius from the sites used for calibration. Daily solar radiation for the same period was obtained from the National Solar Radiation Data Base (NSRDB).

3.2.5 Soil hydraulic parameters

Most field soil properties relevant in subsurface flow estimation in the current version of the WEPP model were input as described in the NRCS Soil Survey for the soil found at the

experimental sites (Table 3.1). These properties included soil texture, field capacity, wilting point, bulk density, and vertical hydraulic conductivities per each soil layer. However, the improved version of the model required additional parameters to generate three main functions: soil water retention, drainage volume, and maximum capillary movement (more details on these functions in Appendix A). Although these additional relationships are commonly measured under laboratory conditions, in this paper, pedotransfer functions (PTFs) were used to predict the additional soil hydraulic parameters and subsequently develop the three input parameters (Table 3.1). Schaap et al. (2001) presents the ROSETTA software, a hierarchical neural network analysis model, that uses soil textures and soil water content at -33 and -1500 kPa and PTFs for generating sets of soil hydraulic properties. This method is convenient since it allows the generation of the additional soil hydraulic parameters using the same information obtained from the NRCS Soil Survey. In addition to the parameters shown in Table 3.1, the WEPP model soil input file requires information on organic matter and cation exchange capacity, both available in the NRCS Soil Survey.

Table 3.1. Soil physical properties and hydraulic properties at the experimental sites from USDA-NRCS Soil Survey and ROSETTA model.

Site	Soil	Layer	BD ^a (g cm ⁻³)	Particle fraction (%)			θ_r^b (cm ³ cm ⁻³)	θ_s^b (cm ³ cm ⁻³)	α^b	n^b	K_{sv}^a	K_{sl}^a (cm h ⁻¹)
				Sand ^a	Silt ^a	Clay ^a						
ES(1)	Loam	1	1.45	32	45	23	0.0611	0.4117	0.0142	1.3133	0.920	1.38
		2	1.45	16	42	42	0.1058	0.4345	0.0394	1.2534	0.092	0.138
		3	1.45	25	40	35	0.2275	0.4065	0.0485	1.8536	0.092	0.138
		4	1.45	22	45	33	0.2051	0.4065	0.0465	1.9721	0.074	0.111
ES(2,4)	Clay loam	1	1.32	25	39	36	0.0664	0.4162	0.0112	1.2628	0.773	1.1595
		2	1.55	18	36	46	0.1056	0.4224	0.0433	1.1814	0.282	0.423
		3	1.55	23	39	38	0.0819	0.4159	0.0269	1.1959	0.920	1.38
		4	1.55	25	38	37	0.2120	0.3864	0.0521	1.6368	0.005	0.0075
ES(3)	Silty clay loam	1	1.43	19	42	39	0.0773	0.4497	0.0139	1.2614	0.100	0.15
		2	1.60	18	36	46	0.0978	0.4098	0.0412	1.1694	0.282	0.423
		3	1.60	23	39	38	0.0766	0.4037	0.0252	1.1865	0.092	0.138
		4	1.60	25	38	37	0.1984	0.3736	0.0529	1.5767	0.005	0.0075

Notes: 1) BD, bulk density; θ_r , residual soil water content; θ_s , saturated soil water content; α , inverse of air-entry value; n , pore size distribution index, K_{sv} saturated vertical hydraulic conductivity, $K_{sl} = 1.5K_s$, saturated lateral hydraulic conductivity. 2) ^a Values as described in the NRCS Soil Survey, ^b ROSETTA model generated values.

3.2.6 Crop/plant management

Crop parameters for WEPP model simulations were obtained from the extensive database available within the model installation. The management input file contains all information related to plants/crops grown, growth parameters, residue characteristics, residue decomposition parameters, tillage implements, soil disturbance parameters, as well as their temporal and spatial

description. Management information collected at each of the sites included tillage operations, planting and harvesting dates, fertilizer application, and details on the subsurface drainage system installation (Table 3.2). Site ES(1) consisted of a two-year corn-soybean rotation with primary and secondary tillage operations using a plowing disk and field cultivator, respectively. Field ES(2) had a two-year corn-soybean rotation with no tillage operations. Tillage operations in ES(3) included primary tillage with a combination subsoiler (disk ripper 30.48 cm (12”) deep) and land leveler equipment, and secondary tillage using a seedbed conditioner (turbo tillage). Lastly, ES(4) included primary tillage with a disk chisel 20.32-25.40 cm (8-10”) deep and secondary tillage using a field cultivator with a rolling harrow 5.08 cm (2”) deep.

Table 3.2 Subsurface drainage system and crop management data

	ES(1)	ES(2)	ES(3)	ES(4)
<i>System design</i>				
Depth to drain (cm)	100	91	91	91
Spacing between drains (cm)	900	1200	760	1200
Depth to impervious layer (cm)	200	200	200	200
Drain radius (cm)	10.16	10.16	10.16	10.16
Drainage coefficient (cm)	3.80	1.75	1.25	0.870
Initial water table depth (cm) ^a	90	70	90	74
<i>Plant/crop management</i>				
Crop	Corn-soybean	Corn-soybean	Soybean-wheat-corn	Corn-soybean
Tillage	Plowing disk and field cultivator	No-till	Disk ripper, land leveler and turbo tillage	Disk chisel and field cultivator

Note: ^a A reasonable initial water table was selected for the beginning of the warmup period. No water table depth data were available.

3.2.7 Performance evaluation measures

Statistical measures can be used to quantify the differences between the predicted and observed values and evaluate the model efficiency. In this study, four statistical measures were used to assess the quality and reliability of WEPP model predictions of subsurface drainage: Root Mean Square Error (RMSE) (Liang et al., 2018), Mean Absolute Error (MAE) (Willmott and Matsuura, 2005), Percent Bias (PBIAS) (Yapo et al., 1996), and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - S_i)^2}{n}} \quad (3.14)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |O_i - S_i| \quad (3.15)$$

$$PBIAS = 100 \frac{\sum_{i=1}^n (S_i - O_i)}{\sum_{i=1}^n O_i} \quad (3.16)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (3.17)$$

where n is the number of observed-simulated values, O_i is the i^{th} observed value, and S_i is the i^{th} simulated value. RMSE represents the absolute difference between observed and simulated values; MAE measures the average magnitude of the errors between the observations and predictions without considering their direction; PBIAS quantifies the average tendency of the simulated values to be larger or smaller than the observed values, and NSE measures the model performance. Evaluation metrics are compared with the calibration criteria for statistical measures of agreement between predicted and measured drainage volume in Moriasi et al. (2015) and Skaggs et al. (2012).

3.3 Results and discussion

3.3.1 Uncalibrated performance at ES(1) (WEPP v2012.8 vs WEPP v2012.9)

Daily drainage measurements at ES(1) were initiated in late 2015 and collected until 2017. Two years (2016-2017) of drainage flux, weather variables, and crop management data were used to compare the uncalibrated performance of the current WEPP model (v2012.8) and the enhanced WEPP model (v2012.9). Simulated daily drainage values using the current WEPP version showed a poor match with measured drainage flux distribution and peaks (Figure 3.3). The model predicted values presented delayed responses with respect to the observed values. Model accuracy measures indicated both large differences between simulated and observed conditions and poor model performance (Table 3.3). For instance, the RMSE and MAE for uncalibrated conditions were 0.686 cm and 0.402 cm, respectively. These measures represent substantial differences since drainage values at ES(1) varied from 0.0 to 2.65 cm. Percent bias was 69.8% meaning that ~70% of the simulated values tended to be greater than the observations. Overall, model performance was poor ($NSE < 0$) with an NSE of -1.17.

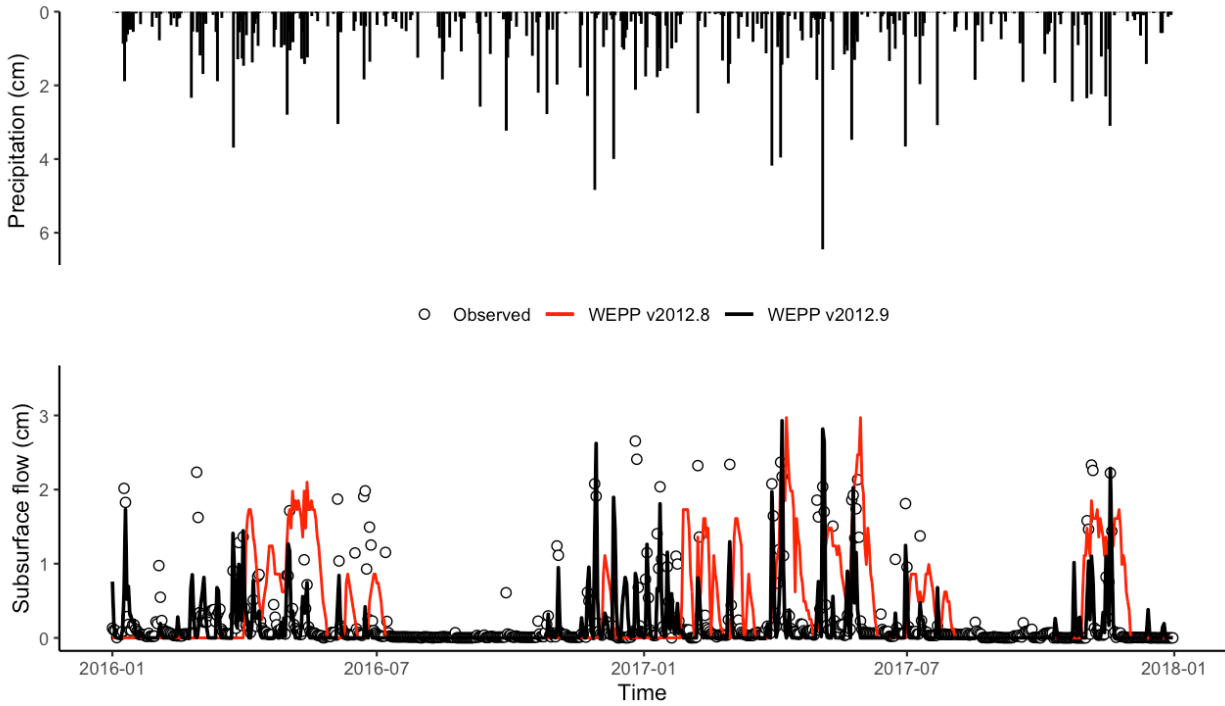


Figure 3.3 Model output at ES(1). (top) observed precipitation; (bottom) observed drainage (circles), WEPP v2012.8 simulated drainage (red solid-line), and WEPP v2012.9 daily drainage (black solid-line).

After analyzing the results, it was difficult to compare predicted to observed values as the water table depth in the current WEPP version did not present a realistic temporal distribution (Figure 3.4) as often observed in previous subsurface drainage investigations (Skaggs et al., 2012a,b; Saadat et al., 2017). Simulations showed water table depth gradients changed very quickly under the presence of small amounts of rainfall or when water was removed from the soil through evapotranspiration. For instance, in the 2016 growing season, a sudden decline of the water table depth was observed (Figure 3.4, red solid line). When water was above the depth to the pipe (100 cm), the water table depths were limited to the soil depth (~80 cm). This behavior seemed to be related to the assumption of the WEPP model moving all water above the field capacity at the end of the day, which forces the model to maintain the water table depth only to a certain soil depth (Figure 3.4, red solid line). In other words, these two behaviors cause transitions between saturated and unsaturated conditions to happen very quickly, which is unusual for many soils.

Table 3.3 Model accuracy evaluation for daily drainage flow estimates at test site ES(1) and calibrated/validated sites ES(2-4).

Site/Model	RMSE (cm)			MAE (cm)			PBIAS (%)			NSE		
	U-M	C-M	V-M	U-M	C-M	V-M	U-M	C-M	V-M	U-C	C-M	V-M
ES(1) – WEPP ^a	0.686	-	-	0.402	-	-	69.8	-	-	-1.17	-	-
ES(1) – WEPP ^b	0.337	-	-	0.154	-	-	-33.2	-	-	0.47	-	-
ES(2) – WEPP ^a	0.203	-	-	0.123	-	-	-12.4	-	-	-0.20	-	-
ES(2) – WEPP ^b	0.157	0.106	0.100	0.069	0.053	0.051	-68.9	-39.3	-28.2	0.27	0.65	0.73
ES(3) – WEPP ^a	0.299	-	-	0.183	-	-	88.4	-	-	-4.72	-	-
ES(3) – WEPP ^b	0.100	0.702	0.112	0.046	0.034	0.045	-53.4	38.2	-15.4	0.36	0.59	0.42
ES(4) – WEPP ^a	0.155	-	-	0.096	-	-	20.0	-	-	-0.18	-	-
ES(4) – WEPP ^b	0.104	0.072	0.089	0.052	0.036	0.045	-30.5	-32.2	-15.3	0.47	0.71	0.69

Note: ^a WEPP v2012.8, ^b WEPP v2012.9. Uncalibrated Model (U-M); Calibrated Model (C-M); and Validated Model (V-M).

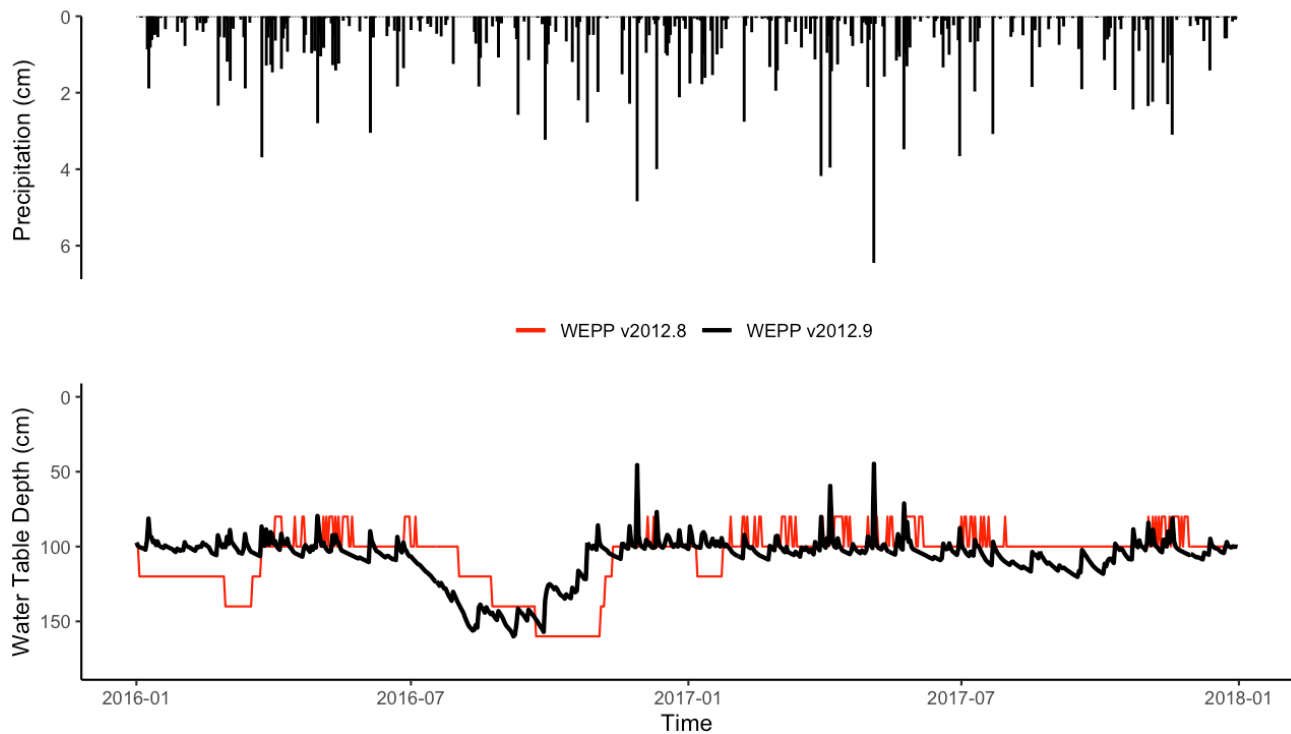


Figure 3.4 Model output at ES(1): (top) observed precipitation; (bottom) WEPP v2012.8 simulated water table depth (red solid-line), and WEPP v2012.9 simulated water table depth (black solid-line).

There are several possible reasons why WEPP drainage fluxes and water table depths were poorly simulated. First, the 24-hour time step used to calculate drainage in the current model is large, and water table depths can change rapidly during this period. Second, the WEPP model seemed to predict a perched water table, and not a water table with saturated conditions from the bottom of the soil profile. Internally, WEPP runs its water routing percolation process through nine layers and creates saturated conditions in upper layers while lower layers may not be saturated.

However, WEPP assumes that water table depth was created from the bottom to top of the soil profile. As a result, lower layers may be in an unsaturated condition, and the transitions between saturated and unsaturated states can occur quickly. Lastly, the WEPP model appeared to underestimate snow melt for the observed field conditions, which reduced infiltration water, and along with the saturated condition of the topsoil layer, that water then tended to leave the hillslope in the form of runoff later in the simulations. As a result, a flat-water table was observed during winter periods of the simulations (Figure 3.4, red solid line). Although this might not affect previous studies where the main purpose was to predict runoff-induced erosion (Savabi, 1993; Savabi et al., 1995; Zhang et al., 1996), this turned out to be important for drainage flux estimation where soil water distribution played an important role. The problem discussed above confirmed some of the findings previously reported by Oztekin et al. (2004) and re-emphasized that the WEPP model needed improvements in its drainage computations.

After all model modifications, drainage predictions were statistically improved (Table 3.3) at site ES(1). Water table depth showed more expected variations (Figure 3.4, black solid line), similar to those produced in soils under shallow water tables (Skaggs et al., 2012c). Predicted drainage peaks matched observed peaks in most cases (Figure 3.3, black solid line) and deviations between predicted and observed drainage improved (Figure 3.5, solid circles). In periods where plants were actively growing (May-Oct), the water table depths were deeper than the pipe installation depth, and consequently drainage stopped. These periods also had greater rates of evapotranspiration so that infiltrated water decreased to its minimum. Differences between observed and predicted daily drainage depth amounts were reduced by more than 50% after the proposed modifications. RMSE and MAE were 0.337 and 0.154 cm, respectively; whereas the current version of WEPP overpredicted drainage (PBIAS = 69.8%), the modified version underpredicted it (PBIAS = -33.2%). Model efficiency, NSE=0.47, greatly increased after the model code changes. The coefficient of determination (R^2) between simulated event drainage amount values and observations was equal to 0.50 (Figure 3.6a).

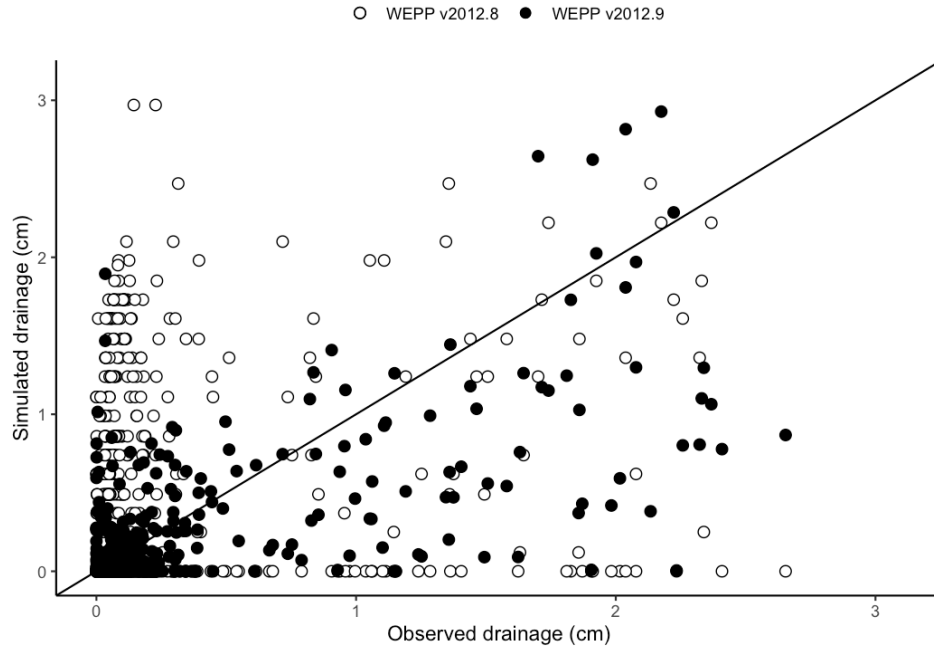


Figure 3.5 Predicted vs. observed daily subsurface drainage water depth (cm) using WEPP model v2012.8 (hollow circles) and WEPP model v2012.9 (solid circles) for uncalibrated performance at ES(1).

3.3.2 Drainage at uncalibrated sites ES(2-4)

Performance measures of the enhanced WEPP model (v2012.9) for uncalibrated conditions (U-M) are shown in Table 3.3. Site ES(2) presented the largest errors of all three locations, with RMSE and MAE equal to 0.157 cm and 0.069 cm, respectively. ES(3) and ES(4) presented similar values of RMSE and MAE ranging from ~ 0.100 cm to MAE ~ 0.046 cm. The WEPP model (v2012.9) applied to ES(2-4) tended to underestimate drainage water depth amounts, with PBIAS = -68.9%, -53.4% and, -30.5%, respectively. NSE values for individual drainage event depths ranged from 0.27 to 0.47 for the three uncalibrated fields (Table 3.3). These performance levels indicate that the uncalibrated version of the improved WEPP model predicted daily drainage discharge depths with an accuracy ranging from $\sim 30\%$ to $\sim 50\%$ for the experimental fields analyzed in this study.

Similar to the uncalibrated validation at ES(1), the values of RMSE and MAE calculated with the current version of the model (v2012.8) at ES(2-4) were about twice as high as ones using the improved version of the model (Table 3.3). RMSE ranged from 0.155 to 0.299 cm, whereas MAE ranged from 0.096 to 0.183 cm. Average values of subsurface drainage depths tended to be

overpredicted at sites ES(3) and ES(4) with PBIAS = 88.4% and 20%, respectively, while at ES(2) values were underpredicted with PBIAS = -12.4%. Lastly, poor behavior of the model was indicated by the negative Nash-Sutcliffe efficiency (NSE) values varying from -4.72 to -0.18 for sites ES(3-4). These results contrasted with those obtained using the improved WEPP model (v2012.9), that increased the confidence in the model enhancements to predict daily estimates of drainage flow.

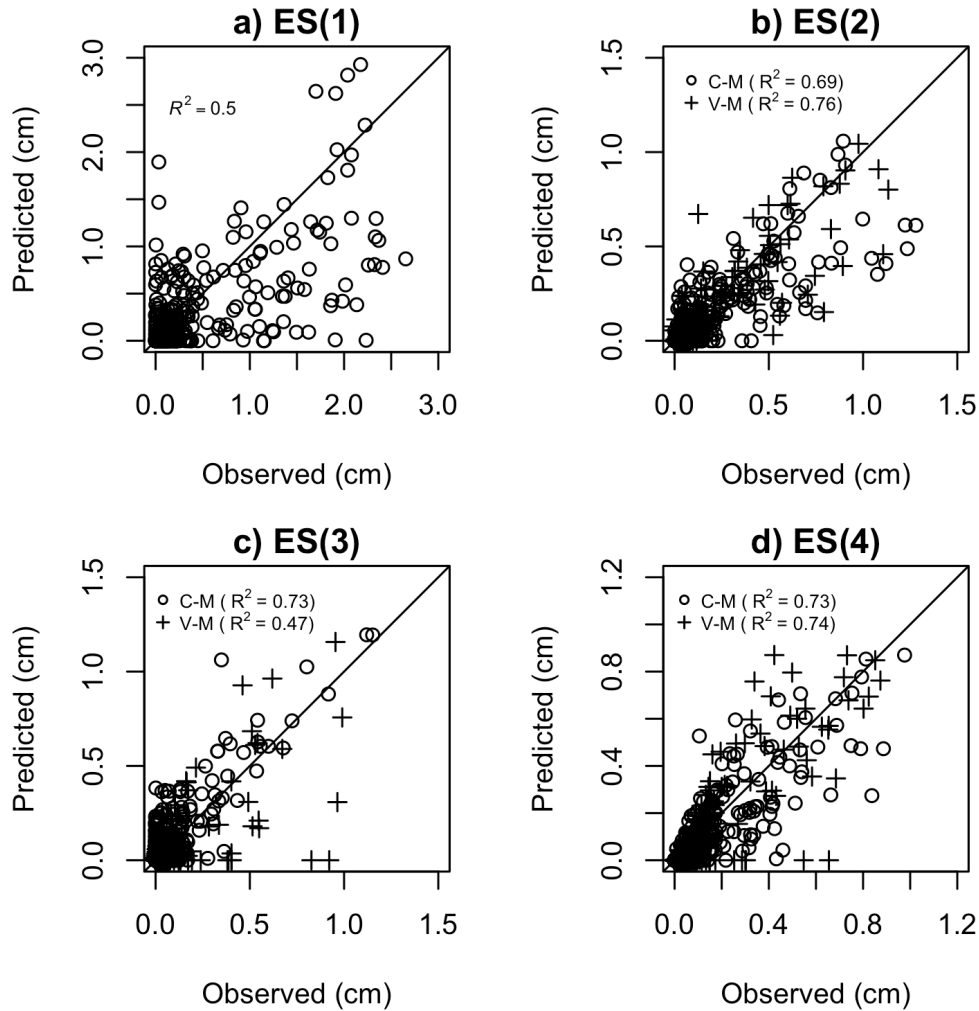


Figure 3.6 Predicted vs. observed daily subsurface drainage depths (cm) using WEPP model v2012.9 for validation site ES(1) and calibrated (C-M)/validated(V-M) sites ES(2-4).

One can easily notice that for the uncalibrated conditions all sites ES(1-4) were in the acceptable range of performance according to Moriasi et al. (2015) and Skaggs et al. (2012), which means the modifications performed in the WEPP model allow suitable subsurface drainage

estimation for the field conditions found at the study sites. In the following section, the saturated vertical and lateral hydraulic conductivity values of the soils at sites ES(2-4) were adjusted manually to improve the model performance during the calibration period. A validation period was then evaluated to observe the model behavior under the calibration parameters.

3.3.3 Drainage at calibrated/validated sites ES(2-4)

Simulated drainage amounts at the calibrated sites ES(2-4) using the enhanced WEPP model (v2012.9) were expected to agree well with observations since the two soil input parameters, K_{sv} and K_{sl} , were adjusted. Although more parameters could be adjusted to reach an optimal agreement, changing many input parameters can lead to non-meaningful model reliability. Skaggs et al. (2012c) presented a list of possible soil and physical parameters that may be adjusted due to uncertainty found in field-scale simulations. Here, only hydraulic conductivities (most sensitive soil parameters) were adjusted to allow for assessment of the model potential for predicting drainage amounts under limited calibration.

Table 3.3 shows the differences and model efficiencies between observed and predicted drainage amounts at the calibrated sites. ES(2) had the greatest errors among all sites, with RMSE and MAE equal to 0.106 cm and 0.053 cm, respectively. ES(3) and ES(4) had similar error magnitudes (RMSE \sim 0.071 cm and MAE \sim 0.035 cm). Although RMSE is in the same drainage units, it does not give either the relative size or the nature of the errors. The updated WEPP model applied to ES(2) and ES(4) tended to underestimate drainage amounts, with PBIAS = -39.3% and -32.4%, respectively. WEPP applied in ES(3) overpredicted drainage (PBIAS = 38.2%). Model efficiencies ranged from 0.52 to 0.71 for all sites (Table 3.3). Even though these efficiencies might be improved by a more detailed calibration procedure, values between $0.50 \leq NSE \leq 1.00$ are considered acceptable when predicting subsurface drainage amounts (Wang et al., 2006; Skaggs et al., 2012c).

Average values of calibrated K_{sv} and K_{sl} ranged from 0.223 to 0.481 cm h⁻¹ and from 0.291 to 1.487 cm h⁻¹, respectively. Clay loam soils at ES(2) and ES(4) showed similar magnitudes of both vertical and lateral hydraulic conductivities, 0.481 and 0.435 cm h⁻¹, respectively. This behavior was expected since soils at those sites had similar physical and hydraulic properties (Table 3.1). K_{sv} and K_{sl} for ES(3) were 0.223 and 0.290 cm h⁻¹, respectively. Figure 3.6 shows substantial correlation between predicted and observed drainage amounts with coefficients of

determination (R^2) ranging from 0.69 to 0.73; WEPP v2012.9 model predictions at the ES(3-4) sites had the greatest correlation. Figure 3.6 also shows how model predictions at ES(3) tended to overpredict drainage amounts, whereas simulations at ES(2) and ES(4) showed the opposite behavior. Predicted daily drainage fluxes at ES(2-4) for the calibration period (2016-2017) matched most of the peaks in the observed data (Figure 3.7). Also, most of the drainage occurred during the spring months (March-May) when crops were not in the field or at initial growth stages and did not have any significant effect on drainage rates.

For the validation period at sites ES(2-4), RMSE and MAE were similar in magnitude to the ones in the calibration period. RMSE average values were ~0.10 cm and MAE values were ~0.47 cm. The latter shows that the mean errors tended to be greater in the validation period. Percent bias in the same period, showed improvement with respect to the calibration period, being equal to -28.2%, -15.4%, and -15.3% for ES(2), ES(3), and ES(4), respectively. While the modified WEPP model (v2012.9) was overpredicting daily drainage flux at ES(3) for the calibration period, estimates were underpredicted for the validation period. NSE values decreased during the validation period at sites ES(3) and ES(4) to 0.42 and 0.69, respectively. However, for site ES(2), NSE improved from 0.65 to 0.73. In summary, the modified version of the WEPP model showed an acceptable response to predict daily drainage depths under the conditions found at experimental sites ES(2-4) for the validation period selected in this study.

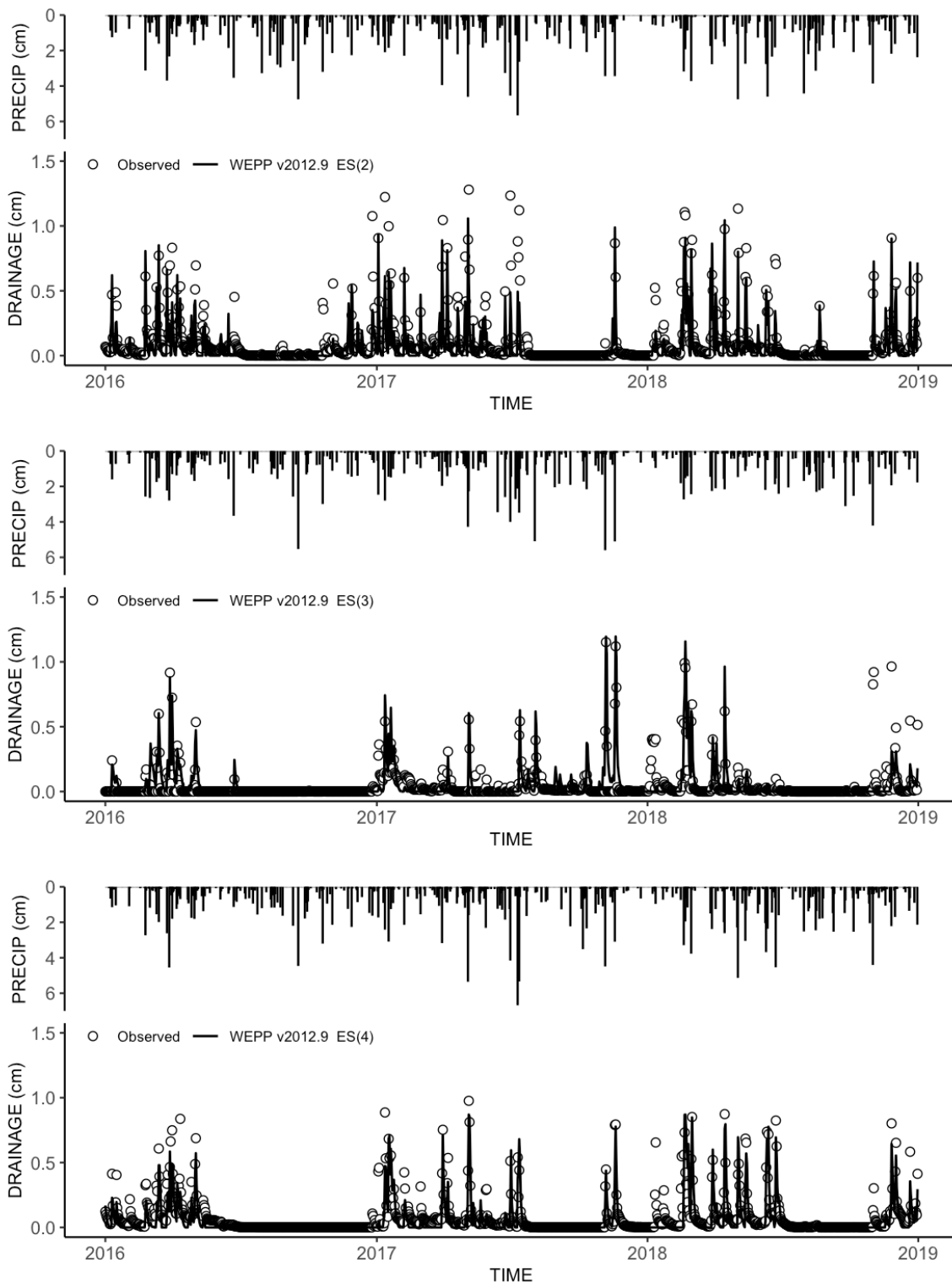


Figure 3.7 Daily drainage at experimental sites - ES(2-4): (Top) Simulated and observed drainage at ES(2); (Middle) Simulated and observed drainage at ES(3); (Bottom) Simulated and observed drainage at ES(4).

3.3.4 Water balance at the experimental sites

Average infiltrated water contributing either to evapotranspiration or subsurface drainage at ES(2-4) for the entire time period (2016-2018) corresponded to 78%, 99%, and 92% of the annual precipitation, respectively. Evapotranspiration amounts at the experimental sites ES(2-4) accounted for approximately 53%, 78%, and 71% of the annual precipitation, respectively (Table 3.4). No seepage was calculated since we assumed that all infiltrated water not contributing to evapotranspiration was drained through the pipes. ES(3) had the greatest rates of infiltration and evapotranspiration. The latter might have occurred due to the presence of a crop during the winter period which induced water infiltration and increased water demand by evapotranspiration. The rates of evapotranspiration for ES(2-4) were compared to the mean annual values for northwestern Ohio and northeast Indiana calculated in the Sanford and Selnick (2013) study. The estimated mean annual ratio of actual evapotranspiration to precipitation (ET/P) at sites ES(2-4) were 0.61, 0.75, 0.70, respectively. Those ET/P values are comparable to a mean value equal to 0.65 found for the field locations in Sanford and Selnick (2013).

Accumulated monthly amounts of drainage for the experimental sites are shown in Figure 3.8. Similar to the daily amounts, model simulations for ES(2) and ES(4) underpredicted monthly drainage amounts, while simulations at ES(3) overpredicted the drainage amounts. For most months, the improved version of the WEPP model simulated ~70% of the subsurface drainage at the sites accurately.

Table 3.4 Annual water balance at the sites ES(2-4).

Sites	Year	Precipitation (cm) ^c	Runoff (cm) ^{a,d}	Infiltration (cm) ^{a,d}	ET (cm) ^{b,d}	Drainage (cm) ^{b,d}
ES(2)	2016	85.22	14.37 (17%)	70.85 (83%)	52.38 (74%)	18.60 (26%)
	2017	92.84	23.14 (25%)	69.70 (75%)	47.33 (68%)	22.55 (32%)
	2018	98.43	24.15 (25%)	74.27 (75%)	45.96 (62%)	27.43 (38%)
ES(3)	2016	76.26	0	76.26 (100%)	65.13 (85%)	13.45 (15%)
	2017	100.51	3.58 (4%)	96.93 (96%)	71.61 (74%)	23.87 (26%)
	2018	86.64	0	86.64 (100%)	67.22 (78%)	16.85 (22%)
ES(4)	2016	80.37	2.91 (4%)	77.46 (96%)	67.80 (88%)	11.83 (12%)
	2017	98.79	11.52 (12%)	87.26 (88%)	65.83 (75%)	21.35 (25%)
	2018	109.94	9.69 (9%)	99.95 (91%)	68.66 (69%)	28.15 (31%)

^a Percent as a fraction of the total precipitation

^b Percent as a fraction of the infiltrated water

^c Measured variable

^d Simulated variable

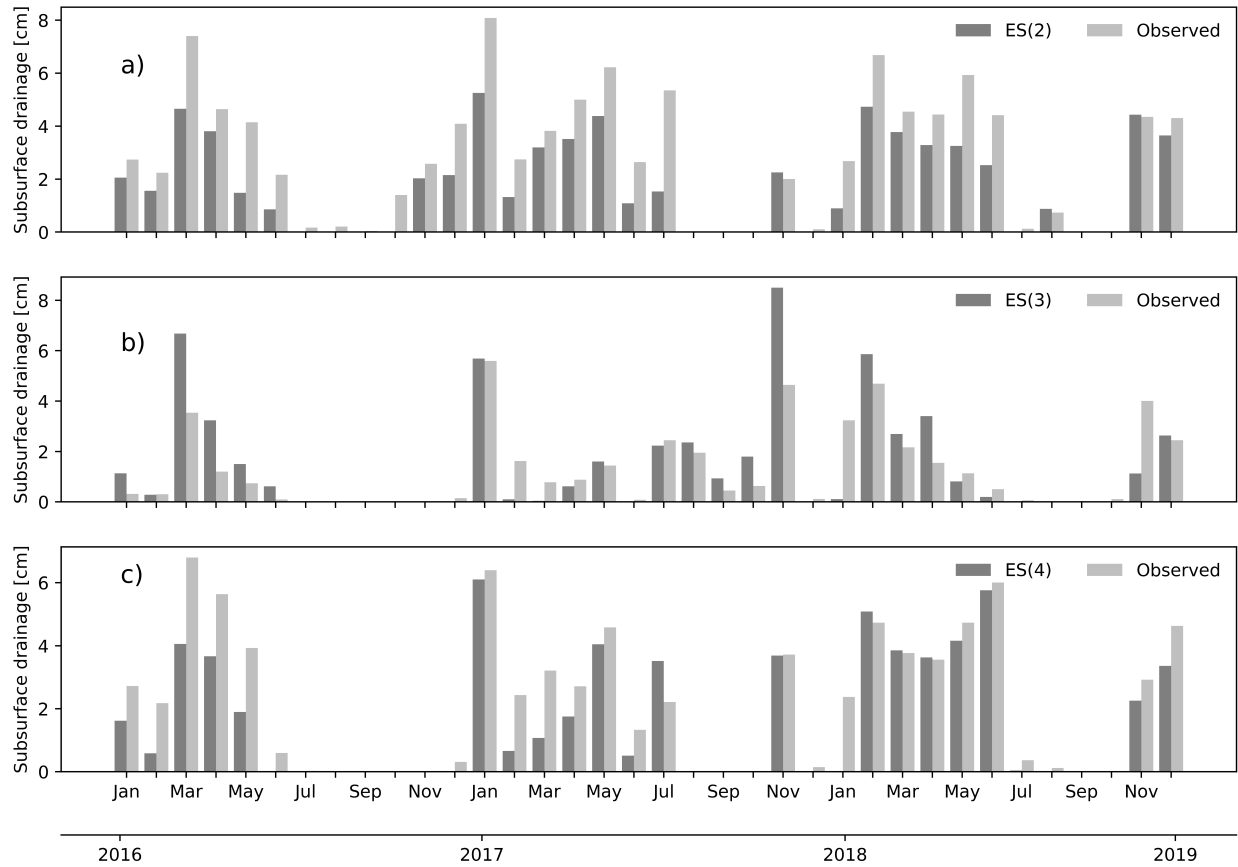


Figure 3.8 Monthly drainage depths for sites ES(2-4): WEPP v2012.9 simulated and observed drainage.

Model simulations for ES(2-4) showed fairly good agreement for drainage fluxes during the winter months for the first two years (2016-2017) (Figure 3.8). However, drainage for the last year of the calibration period was not as accurately simulated. This difference can be attributed to the fact that the assumption of drained-to-equilibrium, which was used here to calculate drainage, does not consider the effects of the freezing and thawing processes on water movement in soils. This issue can be solved by incorporating the unfrozen water content curve derived from the soil freezing characteristic curve measured on-site or approximated using air temperature as a reference (Luo et al., 2000).

Annual observed and predicted drainage amounts are reported in Table 3.5. WEPP v2012.9 simulations at ES(1), ES(2), and ES(4) predicted ~67% of the observed total drained water for the entire simulated period. Model predictions at ES(3) overestimated (~20%) subsurface flow. About 70% of the annual drainage amount at ES(1) was accurately predicted, which indicates that the

modified WEPP model can provided acceptable annual estimates of subsurface drainage for similar soils and conditions.

Table 3.5 Annual WEPP v2012.9 predicted and observed subsurface drainage amounts for the sites ES(1-4).

Year	ES(1)		ES(2)		ES(3)		ES(4)	
	Predicted (cm)	Observed (cm)	Predicted (cm)	Observed (cm)	Predicted (cm)	Observed (cm)	Predicted (cm)	Observed (cm)
2016	49.35	70.92	18.60	31.82	13.45	6.37	11.83	22.19
2017	56.54	87.57	22.55	35.97	23.87	20.63	21.35	26.75
2018	-	-	27.42	38.20	16.85	19.92	28.15	33.21

Among the models for subsurface drainage simulations for both field and watershed scales, DRAINMOD has been leading research at field scales. For a comparison between the WEPP model (v2012.9) and DRAINMOD (v6.0) performances, recent field studies using DRAINMOD can be divided into two categories depending on the nature of the input soil properties: 1) studies using experimental-based inputs (Yang et al., 2007; Salazar et al., 2009; Nabi-Sichani and Sepaskhah, 2012; Skaggs et al., 2012c; Ale et al., 2013; Negm et al., 2014; El Hawary et al., 2015; Golmohammadi et al., 2016; Negm et al., 2016; Negm et al., 2017; Liang et al., 2018; Wilson et al., 2019), and 2) studies using numerical-based inputs (Singh et al., 2006; Singh et al., 2007; Morrison et al., 2014; Qi et al., 2015). The overall efficiency (i.e. NSE) of the studies contained in the first category was 0.65. However, about 50% of those studies were focused on monthly drainage estimates. Average Nash-Sutcliffe Efficiency for monthly drainage fluxes was ~0.75, whereas for the daily drainage fluxes NSE ~ 0.60. The performance of the model in the second category provided an overall efficiency of ~0.60, whereas Nash-Sutcliffe Efficiency values for the daily and monthly drainage studies were 0.72 and 0.40, respectively. The NSE values found in this study for daily and monthly drainage estimates were 0.63 and 0.70, respectively. The improved WEPP model (v2012.9) thus showed a comparable efficiency for both monthly and daily drainage estimates among the studies found in the literature. This indicates that the new version of the WEPP model presented in this study may be a suitable tool for computing estimates of subsurface drainage without any increase in the amount of input data required in previous versions of the model.

3.4 Concluding remarks

The field-scale hydrologic models WEPP v2012.8 and v2012.9 were tested to simulate the effect of artificial subsurface drainage on hydrology of four fields within the western Lake Erie basin, U.S. Initial testing of WEPP showed that the current version (v2012.8) did not accurately predict the daily amount of subsurface drainage. Modifications in the subsurface drainage, percolation, water balance, and soil water distribution algorithms were made to the source code to improve the model drainage calculations. Calibration and validation of the improved version of the model (v2012.9) in three agricultural fields showed that WEPP could predict subsurface drainage amounts well with $NSE > 0.6$. On average 21% of the precipitation left the field through drain pipes at the sites studied here. The monthly and annual amounts of water leaving through drains were also predicted well ($NSE \sim 0.70$). The combined use of the enhanced WEPP model and a suitable soil property estimation model (ROSETTA) showed that an acceptable amount of drainage could be estimated with limited field data. Additional future modifications can be performed to improve the WEPP model estimates of subsurface drainage under shallow water tables during winter conditions.

The results of this study when implemented will increase confidence in the WEPP model predictions for agricultural fields with installed drainage systems in the Midwest U.S. Further investigations using field data and modification of the model to controlled water table conditions are recommended to understand the impacts of shallow water tables on hydrology and crop production in fields with subsurface drainage systems.

3.5 References

- Ale, S., Gowda, P. H., Mulla, D. J., Moriasi, D. N., & Youssef, M. A. (2013). Comparison of the performances of DRAINMOD-NII and ADAPT models in simulating nitrate losses from subsurface drainage systems. *Agricultural Water Management*, 129, 21–30.
<https://doi.org/10.1016/j.agwat.2013.07.008>
- Allen, R.G., L.S. Pereira, D. Raes, D., and M. Smith. (1998). *Crop evapotranspiration, guidelines for computing crop water requirements*. FAO Irrig. and Drain. Paper 56, Food and Agric. Orgn. of the United Nations, Rome, Italy. 300 pp.

- Arnold J. G., Moriasi, D. N., Gassman, P.W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., van Griensven, A., Van Liew, M. W., Kannan, N., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1491–1508.
- Ascough II, J. C., Flanagan, D. C., Nearing, M. A., & Engel, B. A. (2013). Sensitivity and first-order/Monte Carlo uncertainty analysis of the WEPP hillslope erosion model. *Transactions of the ASABE*, 56(2), 437–452.
- Bauwe, A., Kahle, P., & Lennartz, B. (2016). Hydrologic evaluation of the curve number and Green and Ampt infiltration methods by applying Hooghoudt and Kirkham tile drain equations using SWAT. *Journal of Hydrology*, 537, 311–321.
<https://doi.org/10.1016/j.jhydrol.2016.03.054>
- Bouwer, H., & van Schilfgaarde, J. (1963). Simplified method of predicting fall of water table in drained land. *Transactions of the ASAE*, 6(4), 288-291, 296.
<https://doi.org/10.13031/2013.40893>
- Castanheira, P. J., & Santos, F. L. (2009). A simple numerical analyses software for predicting water table height in subsurface drainage. *Irrigation and Drainage Systems*, 23(4), 153–162.
<https://doi.org/10.1007/s10795-009-9079-5>
- Cooke, R., & Verma, S. (2012). Performance of drainage water management systems in Illinois, United States. *Journal of Soil and Water Conservation*, 67(6), 453–464.
<https://doi.org/10.2489/jswc.67.6.453>
- Dale, V., Bianchi, T., Blumberg, A., Boynton, W., Conley, D. J., Crumpton, W., ... Kling, C. (2007). Hypoxia in the northern Gulf of Mexico: An update by the EPA Science Advisory Board. *EPA-SAB-08e003*. *EPA Science Advisory Board, Washington, DC from* [https://yosemite.epa.gov/sab/SABPRODUCT.NSF/C3D2F27094E03F90852573B800601D93/\\$File/EPA-SAB-08-003complete.unsigned.pdf](https://yosemite.epa.gov/sab/SABPRODUCT.NSF/C3D2F27094E03F90852573B800601D93/$File/EPA-SAB-08-003complete.unsigned.pdf). (Accessed 3 October 2019).
- David, M. B., Drinkwater, L. E., & McIsaac, G. F. (2010). Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, 39(5), 1657-1667.
<https://doi.org/10.2134/jeq2010.0115>
- Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models*. Watermark Numerical Computing, Brisbane, Australia. ISBN: 978-0-9943786-0-6

- El Hawary, A., Atta, Y., & El Saadi, A. (2015). Assessment of the DRAINMOD-N II model for simulating nitrogen losses in newly reclaimed lands of Egypt. *Alexandria Engineering Journal*, 54(4), 1305–1313. <https://doi.org/10.1016/j.aej.2015.05.020>
- Flanagan D. C., Gilley J. E., & Franti, T. G. (2007). Water Erosion Prediction Project (WEPP): Development history, model capabilities, and future enhancements. *Transactions of the ASABE*, 50(5), 1603–1612. <https://doi.org/10.13031/2013.23968>
- Flanagan, D. C., Frankenberger, J. R., & Ascough, J. C. II. (2012). WEPP: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1463–1477. <https://doi.org/10.13031/2013.42254>
- Golmohammadi, G., Rudra, R. P., Prasher, S. O., Madani, A., Goel, P. K., & Mohammadi, K. (2016). Modeling the impacts of tillage practices on water table depth, drain outflow and nitrogen losses using DRAINMOD. *Computers and Electronics in Agriculture*, 124, 73–83. <https://doi.org/10.1016/j.compag.2016.03.031>
- Gowda, P. H., Mulla, D. J., Desmond, E. D., Ward, A. D., & Moriasi, D. N. (2012). ADAPT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1345–1352. <https://doi.org/10.13031/2013.42246>
- Green, W. H., & Ampt, G. A. (1911). Studies on soil physics. *The Journal of Agricultural Science*, 4(1), 1–24.
- Green, C. H., Tomer, M. D., di Luzio, M., & Arnold, J. G. (2006). Hydrologic evaluation of the Soil and Water Assessment Tool for a large tile-drained watershed in Iowa. *Transactions of the ASABE*, 49(2), 413–422. <https://doi.org/10.13031/2013.20415>
- Hermesmeier, L. F. (1968). Yield of tile and surface drains and their effect on the water table in a wet soil. *Transactions of the ASAE* 11(1):86-89, 93.
- Hooghoudt, S. B. (1940). Contributions to knowledge of some natural properties of soil. VII. General consideration of the problem of local draining of soil, and of infiltration from parallel running drains and open drain furrows, ditches and drainage channels. *Verslagen van Landbouwkundige Onderzoekingen*, 515–707.
- Keller, C. K., Butcher, C. N., Smith, J. L., & Allen-King, R. M. (2008). Nitrate in tile drainage of the semiarid Palouse Basin. *Journal of Environmental Quality*, 37(2), 353–361. <https://doi.org/10.2134/jeq2006.0515>

- King, K. W., Fausey, N. R., & Williams, M. R. (2014). Effect of subsurface drainage on streamflow in an agricultural headwater watershed. *Journal of Hydrology*, 519(PA), 438–445. <https://doi.org/10.1016/j.jhydrol.2014.07.035>
- King, K. W., Williams, M. R., Macrae, M. L., Fausey, N. R., Frankenberger, J., Smith, D. R., Brown, L. C. (2015). Phosphorus transport in agricultural subsurface drainage: A review. *Journal of Environmental Quality*, 44(2), 467. <https://doi.org/10.2134/jeq2014.04.0163>
- Kirkham, D. (1957). Theory of land drainage. *Drainage of agricultural lands* (Vol. 7, pp. 139–181). American Society of Agronomy, Madison, Wisconsin.
- Kirkham, D. (1966). Steady-state theories for drainage. *ASCE Journal of the Irrigation and Drainage Division*, 92(1), 19–40.
- Kurien, V. M., Cooke, R. A., Hirschi, M. C., & Mitchell, J. K. (1997). Estimating drain spacing of incomplete drainage systems. *Transactions of the ASAE*, 40(2), 377–382.
- Leonard, R. A., Knisel, W. G., & Still, D. A. (1987). GLEAMS: Groundwater Loading Effects of Agricultural Management Systems. *Transactions of the ASAE*, 30(5), 1403–1418. <https://doi.org/10.13031/2013.30578>
- Liang, H., Qi, Z., Hu, K., Li, B., & Prasher, S. O. (2018). Modelling subsurface drainage and nitrogen losses from artificially drained cropland using coupled DRAINMOD and WHCNS models. *Agricultural Water Management*, 195, 201–210. <https://doi.org/10.1016/j.agwat.2017.10.011>
- Luo, W., Skaggs, R. W., & Chescheir, G. M. (2000). DRAINMOD modifications for cold conditions. *Transactions of the ASAE*, 43(6), 1569–1582.
- Maalim, F. K., & Melesse, A. M. (2013). Modelling the impacts of subsurface drainage on surface runoff and sediment yield in the Le Sueur Watershed, Minnesota, USA. *Hydrological Sciences Journal*, 58(3), 570–586. <https://doi.org/10.1080/02626667.2013.774088>
- Mein, R. G., & Larson, C. L. (1973). Modeling infiltration during a steady rain. *Water Resources Research*, 9(2), 384–394.
- Memon, N. A., Madramootoo, C. A., Prasher, S. O., & Broughton, R. S. (1986). A method for estimating the steady upward flux from a water table. *Transactions of the ASAE*, 29(6), 1646–1649.
- Migliaccio, K. W., & Srivastava, P. (2007). Hydrologic components of watershed-scale models. *Transactions of the ASABE*, 50(5), 1695–1703. <https://doi.org/10.13031/2013.23955>

- Moody, W. T. (1967). Nonlinear differential equation of drain spacing. *Journal of the Irrigation and Drainage Division*, 92(2), 1–10.
- Moriasi, D. N., Rossi, C. G., Arnold, J. G., & Tomer, M. D. (2012). Evaluating hydrology of the Soil and Water Assessment Tool (SWAT) with new tile drain equations. *Journal of Soil and Water Conservation*, 67(6), 513–524. <https://doi.org/10.2489/jswc.67.6.513>
- Moriasi, D. N., Gowda, P. H., Arnold, J. G., Mulla, D. J., Ale, S., & Steiner, J. L. (2013a). Modeling the impact of nitrogen fertilizer application and tile drain configuration on nitrate leaching using SWAT. *Agricultural Water Management*, 130, 36–43. <https://doi.org/10.1016/j.agwat.2013.08.003>
- Moriasi, D. N., Gowda, P. H., Arnold, J. G., Mulla, D. J., Ale, S., Steiner, J. L., & Tomer, M. D. (2013b). Evaluation of the Hooghoudt and Kirkham tile drain equations in the Soil and Water Assessment Tool to simulate tile flow and nitrate-nitrogen. *Journal of Environmental Quality*, 42(6), 1699–1710. <https://doi.org/10.2134/jeq2013.01.0018>
- Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*, 58(6), 1763–1785. <https://doi.org/10.13031/trans.58.10715>
- Morrison, J., Madramootoo, C. A., & Chikhaoui, M. (2014). Modeling agricultural land drainage under spring snowmelt conditions with DRAINMOD. *Canadian Journal of Civil Engineering*, 41(4), 275–284. <https://doi.org/10.1139/cjce-2013-0416>
- Mualem, Y., & Dagan, G. (1978). Hydraulic conductivity of soils: unified approach to the statistical models 1. *Soil Science Society of America Journal*, 42(3), 392–395.
- Nabi-Sichani, N., & Sepaskhah, A. R. (2012). Evaluation of different methods for the determination of saturated hydraulic conductivity for simulation of water table depth and drainage discharge using the DRAINMOD computer simulation model. *Archives of Agronomy and Soil Science*, 58(8), 887–901. <https://doi.org/10.1080/03650340.2010.550282>
- Nash, J. E., & Sutcliffe, J. V. (1970). River forecasting using conceptual models, 1. A discussion of principles. *Journal of Hydrology*, 10, 280–290.
- Negm, L. M., Youssef, M. A., Skaggs, R. W., Chescheir, G. M., & Jones, J. (2014). DRAINMOD-DSSAT model for simulating hydrology, soil carbon and nitrogen dynamics, and crop growth for drained crop land. *Agricultural Water Management*, 137, 30–45. <https://doi.org/10.1016/j.agwat.2014.02.001>

- Negm, L. M., Youssef, M. A., Chescheir, G. M., & Skaggs, R. W. (2016). DRAINMOD-based tools for quantifying reductions in annual drainage flow and nitrate losses resulting from drainage water management on croplands in eastern North Carolina. *Agricultural Water Management*, 166, 86–100. <https://doi.org/10.1016/j.agwat.2015.12.014>
- Negm, L. M., Youssef, M. A., & Jaynes, D. B. (2017). Evaluation of DRAINMOD-DSSAT simulated effects of controlled drainage on crop yield, water balance, and water quality for a corn-soybean cropping system in central Iowa. *Agricultural Water Management*, 187, 57–68. <https://doi.org/10.1016/j.agwat.2017.03.010>
- Northcott, W. J., Cooke, R. A., Walker, S. E., Mitchell J.K., & Hirschi M. C. (2001). Application of DRAINMOD-N to fields with irregular drainage systems. *Transactions of the ASABE*, 44(2), 241–249.
- Oztekin, T., Brown, L. C., & Fausey, N. R. (2004). Modification and evaluation of the WEPP hillslope model for subsurface drained cropland. Paper No. 042274, *ASAE/CSAE Annual International Meeting 2004*. ASAE: St. Joseph, Mich.. <https://doi.org/10.13031/2013.16450>
- Philip, J. R. (1954). An infiltration equation with physical significance. *Soil Science*, 77(2), 153–158.
- Qi, Z., Helmers, M. J., Christianson, R. D., & Pederson, C. H. (2011). Nitrate-nitrogen losses through subsurface drainage under various agricultural land covers. *Journal of Environmental Quality*, 40(5), 1578–1585. <https://doi.org/10.2134/jeq2011.0151>
- Randall, G. W., & Goss, M. J. (2008). Chapter 6. Nitrate losses to surface water through subsurface, tile drainage. In (Hatfield, J. R. & Follett, R. F., eds.) *Nitrogen in the Environment: Sources, Problems, and Management*, 2nd ed. Elsevier: Amsterdam. pp. 145–175. <https://doi.org/10.1016/B978-0-12-374347-3.00006-8>
- Saadat, S., Bowling, L., Frankenberger, J., & Brooks, K. (2017). Effects of controlled drainage on water table recession rate. *Transactions of the ASABE*, 60(3), 813–821. <https://doi.org/10.13031/trans.11922>
- Salazar, O., Wesström, I., Youssef, M. A., Skaggs, R. W., & Joel, A. (2009). Evaluation of the DRAINMOD-N II model for predicting nitrogen losses in a loamy sand under cultivation in south-east Sweden. *Agricultural Water Management*, 96(2), 267–281. <https://doi.org/10.1016/j.agwat.2008.08.008>

- Sanford, W. E., & Selnick, D. L. (2013). Estimation of evapotranspiration across the conterminous United States using a regression with climate and land-cover data. *Journal of the American Water Resources Association*, 49(1), 217–230. <https://doi.org/10.1111/jawr.12010>
- Savabi, M. R. (1993). Modeling subsurface drainage and surface runoff with WEPP. *Journal of Irrigation and Drainage Engineering*, 119(5), 801–813. [https://doi.org/10.1061/\(asce\)0733-9437\(1993\)119:5\(801\)](https://doi.org/10.1061/(asce)0733-9437(1993)119:5(801))
- Savabi, M. R., Flanagan, D. C., Hebel, B., & Engel, B. A. (1995). Application of WEPP and GIS-GRASS to a small watershed in Indiana. *Journal of Soil and Water Conservation*, 50(5), 477–483.
- Schaap, M. G., Leij, F. J., & Van Genuchten, M. T. (2001). Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, 251(3–4), 163–176.
- Shaykewich, C. F., & Stroosnijder, L. (1977). The concept of matrix flux potential applied to simulation of evaporation from soil. *Netherlands Journal of Agricultural Science*, 25, 63–82.
- Shokri, A., & Bardsley, W. E. (2015). Enhancement of the Hooghoudt drain-spacing equation. *Journal of Irrigation and Drainage Engineering*, 141(6), 04014070. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000835](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000835)
- Singh, R., Helmers, M. J., Crumpton, W. G., & Lemke, D. W. (2007). Predicting effects of drainage water management in Iowa's subsurface drained landscapes. *Agricultural Water Management*, 92(3), 162–170. <https://doi.org/10.1016/j.agwat.2007.05.012>
- Singh, R., Helmers, M. J., & Qi, Z. (2006). Calibration and validation of DRAINMOD to design subsurface drainage systems for Iowa's tile landscapes. *Agricultural Water Management*, 85(3), 221–232. <https://doi.org/10.1016/j.agwat.2006.05.013>
- Skaggs, R. W. (1980). Drainmod. Reference report. Methods for design and evaluation of drainage-water management systems for soils with high water tables. USDA-SCS, South National Technical Center, Fort Worth, TX, USA, 130 pp., unpublished.
- Skaggs, R. W., Brevé, M. A., & Gilliam, J. W. (1994). Hydrologic and water quality impacts of agricultural drainage*. *Critical Reviews in Environmental Science and Technology*, 24(1), 1–32. <https://doi.org/10.1080/10643389409388459>

- Skaggs, R. W., Youssef, M. A., Gilliam, J. W., & Evans, R. O. (2010). Effect of controlled drainage on water and nitrogen balances in drained lands. *Transactions of the ASABE*, 53(6), 1843–1850. <https://doi.org/10.13031/2013.35810>
- Skaggs, R. W., Youssef, M. A., & Chescheir, G. M. (2012a). A DRAINMOD-based method to estimate effects of drainage water management on annual nitrogen loss to surface water. *Transactions of the ASABE*, 55(3), 799–808. <https://doi.org/10.13031/2013.41515>
- Skaggs, R. W., Fausey, N. R., & Evans, R. O. (2012b). Drainage water management. *Journal of Soil and Water Conservation*, 67(6), 167A-172A. <https://doi.org/10.2489/jswc.67.6.167A>
- Skaggs R. W., Youssef, M. A., & Chescheir, G. M. (2012c). DRAINMOD: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1509–1522. <https://doi.org/10.13031/2013.42259>
- Taylor, S. A., & Ashcroft, G. L. (1972). *Physical Edaphology. The Physics of Irrigated and Nonirrigated Soils*. W.H. Freeman and Company, San Francisco, Calif., USA. 533 pp.
- van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5), 892–898.
- van Schilfgaarde, J. (1970). Theory of flow to drains. In *Advances in Hydrosience* (Vol. 6, pp. 43–106). Academic Press, Inc.: New York. <https://doi.org/10.1016/B978-0-12-021806-6.50008-3>
- Wang, X., Frankenberger, J. R., & Klavivko, E. J. (2006). Uncertainties in DRAINMOD predictions of subsurface drain flow for an Indiana silt loam using the GLUE methodology. *Hydrological Processes*, 20(14), 3069–3084. <https://doi.org/10.1002/hyp.6080>
- Wesseling, J. (1964). A comparison of the steady state drain spacing formulas of Hooghoudt and Kirkham in connection with design practice. *Journal of Hydrology*, 2(1), 25–32. [https://doi.org/10.1016/0022-1694\(64\)90062-9](https://doi.org/10.1016/0022-1694(64)90062-9)
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79–82. <https://doi.org/10.3354/cr030079>
- Wilson, G. L., Mulla, D. J., Galzki, J., Laacouri, A., Vetsch, J., & Sands, G. (2019). Effects of fertilizer timing and variable rate N on nitrate–N losses from a tile drained corn-soybean rotation simulated using DRAINMOD-NII. *Precision Agriculture*, (0123456789). <https://doi.org/10.1007/s11119-019-09668-4>

- Wischmeier, W. H., & Smith, D. D. (1978). Predicting rainfall erosion losses – a guide to conservation planning. *USDA Agriculture Handbook No. 537*. U.S. Govt. Printing Office, Washington, D.C.
- Yang, C. C., Prasher, S. O., Wang, S., Kim, S. H., Tan, C. S., Drury, C., & Patel, R. M. (2007). Simulation of nitrate-N movement in southern Ontario, Canada with DRAINMOD-N. *Agricultural Water Management*, 87(3), 299–306.
<https://doi.org/10.1016/j.agwat.2006.07.009>
- Yapo, P. O., Gupta, H. V., & Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *Journal of Hydrology*, 181(1–4), 23–48.
[https://doi.org/10.1016/0022-1694\(95\)02918-4](https://doi.org/10.1016/0022-1694(95)02918-4)
- Zhang, X. C., Nearing, M. A., Risse, L. M., & McGregor, K. C. (1996). Evaluation of WEPP runoff and soil loss predictions using natural runoff plot data. *Transactions of the ASAE*, 39(3), 855-863.

4. FORMULATION OF A ROUTING MODEL FOR THE WATER EROSION PREDICTION PROJECT (WEPP) MODEL BASED ON CATCHMENT GEOMORPHOLOGY AND TRAVEL TIME THEORY

Abstract

This paper describes the theoretical framework for a new hillslope and channel-routing routine for the Water Erosion Prediction Project (WEPP) model. The routing model (WEPP-CMT) underlies the continuous mass transport modeling (CMT) theory for flow and transport of reactive solutes. The model structure uses the unique functionality of WEPP to simulate hillslope responses under diverse land use and management conditions and a Lagrangian description of the carrier hydrologic runoff at two stages: hillslope and channel domains where flow and transport are represented in the model through travel time distributions (TTD) derived from catchment geomorphology using a stochastic framework. An example of the model functionality was tested in a sub-catchment of the Upper Cedar River Watershed in the U.S. Pacific Northwest. Results showed that the model offers an acceptable representation of flow at the outlet of the studied catchment. Model efficiencies and percent bias for the calibration period and the validation period were $NSE = 0.55$ and 0.65 , and $PBIAS = -10\%$ and -3.1% , respectively. A comparison with an existing routing method (WEPP-MK) found in the geospatial version of the model, GeoWEPP, showed similar results. Nonetheless, the latter was able to capture peaks of the hydrologic response slightly more accurately, which suggests that WEPP-CMT needs a more appropriate representation of the temporal and spatial variability of the parameters controlling the routing model. The WEPP-CMT provides a suitable foundation for the transport of reactive solutes (e.g. nitrates) at basin scales.

4.1 Introduction

Understanding of hydrological systems is an important subject for people and their environment. It allows grasping the relationship of the waters of the earth within each phase of the hydrologic cycle, their occurrence, circulation and distribution, their chemical and physical properties, and their reaction with the environment, including their relation to living things (Linsleuy et al., 1975). Particularly, the transformation of natural ecosystems into agricultural and

urban areas impacts water quantity and water quality. These physical perturbations along with ongoing climate change and soil heterogeneities modify flow discharges of many rivers around the world and bring collateral changes in other hydrologic processes where water draining into streams interacts. One of the techniques to quantify these changes is to model the effect of human disturbances on water quantity and quality, and to compare the outputs under different management conditions (Zhou et al., 2013; Moriasi et al., 2012; Moriasi et al., 2013; Hoghooghi et al., 2017; Wang et al., 2018; Aboelnour et al., 2019).

Various natural resource models (Arnold et al., 2012; Flanagan et al., 2012; Flanagan et al., 2013; Bhandari et al., 2017) continue efforts to improve the representation of the hydrological processes in large complex basins. Each model possesses its unique characteristics and requires inputs such as precipitation, air temperature, soil properties, topography, vegetation, and other physical parameters. Among available hydrologic models, one lagging in terms of water quantity and quality investigations at the catchment scale is the Water Erosion Prediction Project (WEPP) (Flanagan et al., 2012). A significant factor for this lag is that the WEPP model was developed to function on hillslopes and small catchments (~260 ha) where residence times in channels are small in comparison with those on hillslopes (Flanagan et al., 2012). To overcome this limitation, Wang et al. (2010) implemented additional channel-routing routines in the WEPP source code, but only a few researchers have used and validated it at relatively large catchment scales (Afshar and Hassanzadeh, 2017; Srivastava et al., 2017; Srivastava et al., 2019), and the model interface is not set up to function for large catchments. Additionally, water quality research using the WEPP model has been limited to hillslope scales, and only one recent investigation on transport of reactive solutes has been published (Wang et al., 2018).

As a result of the problems explained above, the need for a routing model dealing with both flow and solute transport is an open opportunity to extend the use of the WEPP model to larger scales. In this context, general transport models (i.e. advection-diffusion models), especially those that follow a stochastic framework, have been shown to be appropriate for large-scale applications (Rodriguez-Iturbe and Rinaldo, 2001; Rinaldo et al., 2005). Formulations of transport by travel time distribution have been widely supported by several studies (Rodriguez-Iturbe and Valdes, 1979; Gupta et al., 1980; Dagan, 1989; Rinaldo and Rodriguez-Iturbe, 1996) and applied to investigations at basin scales (Rinaldo and Marani, 1987; Rinaldo and Rodriguez-Iturbe, 1996; Gupta and Cvetkovic, 2002; Botter and Rinaldo, 2003; Botter et al., 2005). However, these have

largely been applications with several simplifications of the hydrologic response (e.g. direct runoff, movement of water in soils, and dispersion mechanisms) to fully comprehend a single component of the hydrologic response. In this study, a general mass transport model with a stochastic framework based on travel time distributions is developed for routing the hydrologic response. The proposed model integrates the element of general transport theory with the unique functionality of a process-based model such as WEPP through an objective manipulation of hydrologic, geomorphologic and soil properties, and land use data, and provides suitable groundwork for transport of reactive solutes using the WEPP model.

The structure of this paper is divided into four sections. The first section provides a brief introduction of the general transport theory used for the formulation of the model; the second section describes the underlying model structure and assumptions; a practical scenario is provided to evaluate the model performance in section three; and lastly, concluding remarks as well as future research for model improvements are mentioned in section four.

4.2 Objectives

The two main objectives of this study were to: (1) present the formulation of a new general transport model for routing of hydrologic response based on geomorphologic properties and travel time distributions at the basin scale, (2) demonstrate a practical application of the WEPP model and the proposed mass transport model to predict the hydrologic response in a mountainous basin in western Washington state, U.S.

4.3 Theoretical background

4.3.1 Mass transport theory

Flow and transport phenomena in the hydrologic response are complex because of the interaction of physical and geochemical properties, acting over a multiplicity of spatial and temporal scales (Rubin, 2003). Further complexities at the basin scale increase prediction uncertainties due to time-dependent flow paths and variability of chemical conditions. Thus, stochastic models might be a better option than deterministic models, as they embody both spatial variability and uncertainty into a coherent mathematical framework with a limited number of parameters such as those using probability distributions. Stochastic models also require fewer

assumptions and can provide a suitable description of the physical processes involved in these complex systems (Rinaldo et al., 2005).

A practical description of the journey of a water particle in a control volume (i.e., basin) can be seen as surface and subsurface trajectories the particle travels within natural formations before reaching the basin's outlet without significant change in its mass. Assuming that the mass of an injected particle m_w at time $t_0 = 0$ is carried by a single particle, trajectories can be described through a Lagrangian scheme as follows:

$$X(t) = x_0 + \int_0^t v(X(\tau), \tau) d\tau \quad (4.1)$$

where $v(x, t)$ is known as the velocity vector. In a transport volume, the spatial distribution of water concentration after the injection of a single particle is described by (Taylor, 1922):

$$c_w(x, t) \approx m_w \delta(x - X(t)) \quad (4.2)$$

where $\delta(*)$ is Dirac's delta function. Under the scenario of water transport, a unit and constant porosity are assumed so that the water concentration in the transport volume at any specific location along the flow paths remains constant.

At basin scales, the trajectory of an injected water particle can be seen as a random function due to the heterogeneity of natural formations as well as the interaction with physical and geochemical processes. Assuming that $g(X)dX$ is the probability that the particle is found within the infinitesimal volume dX located around the position X at time t , the average concentration $\langle c_w(x, t) \rangle$ is given by (Taylor, 1922, Dagan, 1989):

$$\langle c_w(x, t) \rangle = \int_{-\infty}^{\infty} m_w \delta(x - X) g(X) dX = m_w g(x, t) \quad (4.3)$$

where $g(X)$ is called the displacement probability function, which under the kinematics of the particle injected in the control volume rules the travel time distribution $f(t)$. It is worthy to note that the time of all particles injected through the control section is exclusive, and the flux passing through the outlet is proportional to the travel time distribution. Thus, the total mass storage inside the control volume V at time t is:

$$M_w(t) = m_v \int_V g(X, t) dx = m_v P(T \geq t) \quad (4.4)$$

Here, $P(T \geq t)$ represents the probability that the residence time is larger than the actual time. Using the concept of mass continuity, one can obtain:

$$\frac{M_w(t)}{dt} = I - Q_w \quad (4.5)$$

where I is the mass injected in the control volume, and Q_w is mass flux at the basin outlet. Replacing Equation (4.4) in (4.5) and assuming an instantaneous pulse of water, $I(t) = m_w \delta(t)$, a final expression for Q_w as a function of the travel times is as follows:

$$Q_w(t) = -m_w \frac{dP(T \geq t)}{dt} = m_w f(t) \quad (4.6)$$

A particular challenge appears under this formulation, that is the definition of a proper travel time distribution function $f(t)$. Studies have tried to address this challenge adopting particular functions and deriving their parameters with relatively physical meanings (Rodriguez-Iturbe and Valdes, 1979); others have derived travel time distributions from an extensive analysis of the equation of motion (Rinaldo and Rodriguez-Iturbe, 1996). The following section provides a brief analysis of the two approaches for characterizing travel time distributions used in this paper.

4.3.1 Travel time distribution formulations

In transport models, the determination of travel time distributions follows an analysis of the detailed motion of water particles in space and time over hillslopes and through the channel networks. In this context, a catchment or basin is a nested structure of geomorphologic states in which the hydrologic transport occurs. Catchment hydrology identifies two main states: hillslope and channel. Transport processes taking place on hillslopes are central in defining the features of the catchment travel time distributions even at large scales (Wood et al., 1988; Botter et al., 2005). The relatively small flow velocities on hillslopes are responsible for relatively large contact times between the soil and the hydrologic surface and subsurface flows, with significant implications for the resulting carrier concentrations.

In contrast, flow velocity in channels is usually between 1-5 m/s, while the overland flow (runoff) rates on hillslopes or in unchannelized valleys are at least 10 times slower (Emmett, 1978; Huff et al. 1982; van der Tak and Bras, 1990; McGuire et al, 2005). Although one could think solute generation mostly occurs on hillslopes, it can be significant from channels if travel times become large due to exchanges with hyporheic zones, riparian vegetation, and organic matter

decay. As a result of these differences in travel times, independent travel time distributions for surface and subsurface stages need to be defined for suitable mass transport model implementation at basin scales.

4.3.1.1 Surface hydrologic response function

The transport of the fast component of the hydrologic response has been investigated through linear systems theory. One of the first contributions was attributed to Sherman (1932) developing the theory of the unit hydrograph, which simulates hydrologic response as the sum of pulses generated by different inputs of precipitation. Later, Rodriguez-Iturbe and Valdes (1979) reformulated this theory and provided a geomorphological interpretation of the basin response function and named it the Geomorphological Instantaneous Unit Hydrograph (GIUH). This formulation has been extensively analyzed in many studies (Gupta et al., 1980; Rinaldo et al., 1991, Rinaldo and Rodriguez-Iturbe, 1996; Saco and Kumar, 2002a, 2002b; White et al., 2004; D’Odorico and Rigon, 2003; Paik and Kumar, 2004), as it provides a suitable approach to study dispersion mechanisms at catchment scales; however, only a few studies have had practical applications (Rinaldo et al., 2005; Nicótina et al., 2008).

In this study, the travel time distribution $f(t)$ governing the overland response routing relates somehow to the instantaneous unit hydrograph (Gupta et al., 1980). However, wider representation of hydrologic response is adopted through a functional description of the distances from any point of the catchment to its outlet. This descriptor is known as the area-width function (Beven and Wood, 1993), which provides a complete coverage of all possible paths in a basin of both hillslope and channel states, which are intrinsically heterogeneous media due to heterogeneity of soil texture, land cover, stream roughness, and other properties of a catchment. Although this function has the advantage of being directly derived from Digital Elevation Models (DEM), a transfer function is required to express it as a travel time distribution. A residence time function accounting for both convective and storage-diffusive effects in channels is a solution of the Fokker-Plank equation (e.g. Equation (4.8), Rinaldo and Rodriguez-Iturbe, 1996). These expressions do not consider the fractal sinuosity of individual channels, yet capture the fractal properties of the network structure. Thus, the travel time distribution $f(t)$ for the fast-hydrologic response in Equation (4.6) can be expressed as:

$$f_s(t) = \int_0^l W(x) f_c(t) \quad (4.7)$$

where $W(x)$ is the width-area function, l is the longest distance from the basin outlet, and $f_c(t)$ is the advection-dispersion equation of all possible paths defined as nested transitions between hillslope and channel states contained in $W(x)$ and described by the following expression:

$$f_c(t) = \frac{L}{\sqrt{4\pi Dt^3}} e^{\left[-\frac{(x-at)^2}{4Dt}\right]} \quad (4.8)$$

where L is the maximum length from the basin outlet, D is the hydrodynamic dispersion coefficient, and a is the velocity of propagation of the flood wave or wave celerity. A common equivalence of wave celerity in terms of channel velocity u is $a = 1.5u$. The difference in velocities between hillslopes and channels are quantified by assuming that distances in hillslope states are amplified by a factor $r = u_c/u_h$ representing to the number of times that velocity in channels u_c exceeds the ones in hillslopes u_h . Thus, distances x in $W(x)$ are expressed by the following equation:

$$x = x_c + rx_h \quad (4.9)$$

where x_c is the distance of the trajectory that water travels in the channel state, and x_h is the distance in the hillslope state.

4.3.1.2 Subsurface hydrology response function

A particular case of Lagrangian approaches is presented by Rinaldo et al. (2005) to couple the transport of water and matter in the framework of transport by travel time distributions and extend to the case of reactive components in the context of the geomorphologic unit hydrograph (GIUH) (Rodriguez-Iturbe and Valdes, 1979). Similar to the overland response, the idea stems from assuming that a particle of water spends a certain amount of time in each state defining each path (i.e. sub-surficial hillslopes and channels) while traveling to the basin outlet. The travel time T_X of the particle in state X is a random variable described by the probability density function $f_X(t)$. Similarly, a particle in a different state Y with travel time T_Y can have a different travel time distribution $f_Y(t)$, where T_X and T_Y are statistically independent. Any path γ from the source to the outlet contained in Γ (i.e. basin) is composed of a set of states $\gamma = (X_1, \dots, X_O)$ where X_O is the control section at the outlet of the basin. Thus, by geomorphological classification, one can

categorize the states corresponding to hillslopes (h) and channels (c). The entire set of states and their corresponding transitions $X_1, \dots, X_O \in (h_1, \dots, h_O, c_1, \dots, c_O)$ provides a detailed description of the entire structure of the river network. The random travel time through the entire path γ , $T_\gamma = T_{X_1} + \dots + T_{X_O}$, follows the path travel time distribution $f_\gamma(t)$, and is defined as the integral convolution over all of the state travel time distributions forming the entire path as follows:

$$f_\gamma(t) = f_{X_1} * \dots * f_{X_O} \quad (4.10)$$

Hence, the travel time distribution $f(t)$ for the slow component of the hydrologic response in Equation (4.6) at the outlet is obtained by randomizing over the set Γ of all possible paths to the basin outlet (Rinaldo et al., 2005):

$$f_{sb}(t) = \sum_{\gamma \in \Gamma} p(\gamma) f_\gamma(t) \quad (4.11)$$

where $p(\gamma)$ is the path probability and $\sum_{\gamma \in \Gamma} p(\gamma) = 1$. Under the assumption that the rainfall input $I(t)$ is uniformly distributed in space, $p(\gamma)$ is simply A_{h_i}/A_b , with A_{h_i} being the area of the injection area at hillslope state h_i (i.e. sub-basin area) and A_b the area of the basin. Although this might be true for small catchments, precipitation in large catchments tends to present spatial and temporal variability. For those cases $p(\gamma)$ presents temporal variability and can be estimated as:

$$p(\gamma, t) = \frac{I(\gamma, t)}{I(t)} \quad (4.12)$$

where $I(\gamma, t)$ is the injected rainfall in state h_i in time t , and $I(t)$ is the total rainfall injected in the basin at time t . Here, h_i represents each of the hillslopes discretized within the basin and where the precipitation is considered to be spatially uniform, but not the same among all h_i 's.

It is important to note that by the latter assumption, one is relaxing the condition of never-changing travel time distributions such as the case of the unit hydrograph theory (Sherman, 1932). Distribution functions adopted for the channel states are in the form of Equation (4.8), whereas for hillslope states an exponential function with mean residence time (D'Odorico and Rigon, 2003; Nicotina et al., 2008) can be used:

$$\bar{t}_{\gamma, i} = \alpha_i A_\gamma^{0.38} \quad (4.13)$$

where A_γ is the sub-basin area of the hillslope state h_i , and α_i is a calibration parameter. In Equation (4.13), the average travel time is a function of the basin size. Such empirical expression has been proposed by Boyd (1978) and Bathurst (1993) for basins ranging between 0.388 km² and

251 km². Additionally, the selection of a transport model by travel time distributions used in this study offers a suitable framework to link travel times with the global contact time controlling mass exchanges and consequently provides reasonable operational capabilities to describe large-scale transport processes where hydrologic transport occurs.

4.4 Methods

4.4.1 General model approach

The general approach of the WEPP model and its integration with a Continuous Mass Transport model (WEPP-CMT) follows the sequence in Figure 4.1. The partition into the surface and subsurface pathways was performed by using the WEPP hillslope model for all the hillslopes found within the catchment. The WEPP watershed model treated a watershed as a collection of hillslopes and channels draining to the basin outlet (Baffaut et al., 1997). Watershed discretization was performed using a geospatial interface of the WEPP model such as GeoWEPP (Renschler, 2003). This representation allowed the incorporation of a mass transport model and travel time distribution theories as a possible routing routine at watershed scales such as the one described in the previous section. Following the watershed spatial discretization, information describing topography, soils, land use, and climate was input to the WEPP model to predict both surface and subsurface rainfall partitioning. Outputs from each hillslope were then used as inputs for the proposed mass transport model (WEPP-CMT), which routes daily partitions of rainfall traveling from the surface and subsurface states to the catchment outlet (Figure 4.1).

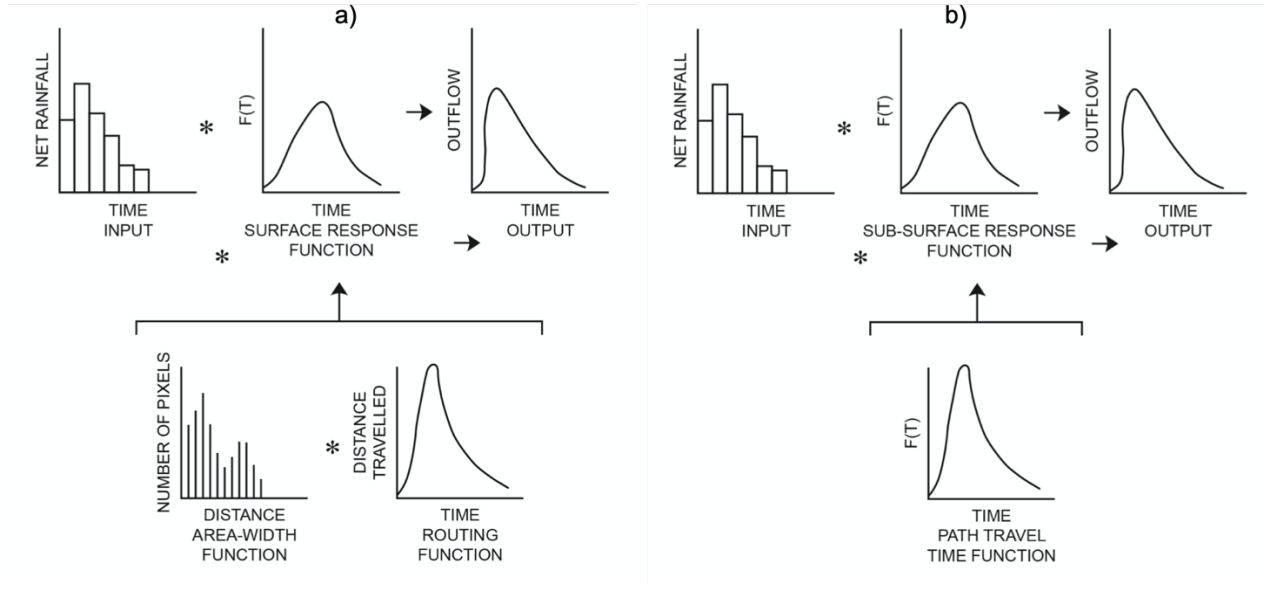


Figure 4.1 Hydrologic response as a combination of: a) surface response function and b) subsurface response function.

Once the rainfall was partitioned (i.e. net rainfall for both surface and subsurface states in Figure 4.1), hillslope response functions were obtained following Equations (4.7) and (4.11) described in the previous section for both surface and subsurface components of the response, respectively. Description of the distance from any point on the land surface to the closest stream was required to obtain the width-area function in Equation (4.7), whereas the hillslope areas were required in Equation (4.13). The spatial analysis of these properties was performed using the geospatial software GRASS GIS (v7.5). The transfer function for the surface response function (Figure 4.1a) for the surface runoff adopted the form shown in Equation (4.8), which depended on the hillslope distances and two hydrodynamic parameters. Hillslope distances were modified according to the calibration factor r in Equation (4.9), which changes the number of times that channel velocities exceed hillslope velocities. For the subsurface flow, hillslope travel time distributions took an exponential form with the mean travel time calculated using Equation (4.12). This expression required the hillslope area in km^2 (Figure 4.2b) and a calibration parameter α in $\text{hours}/(\text{km}^2)^{0.38}$.

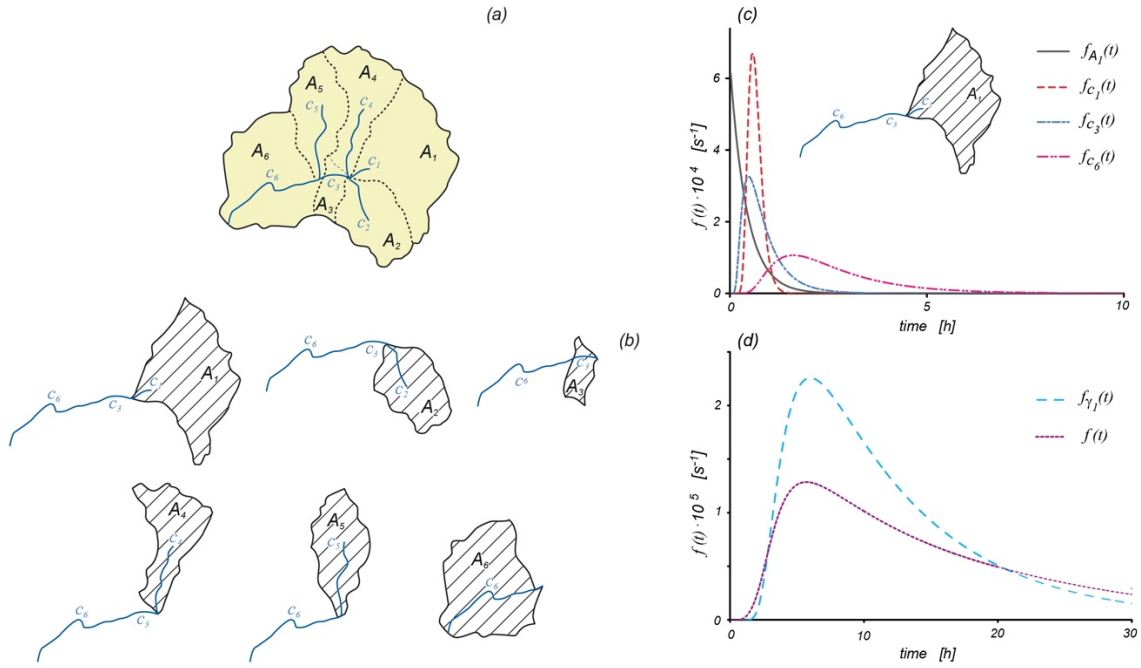


Figure 4.2 Subsurface travel time distribution and its derivation

After the hydrologic runoff leaves the hillslopes, the water is routed in channels as a mixture of surface and subsurface flows. As a result, a unique set of parameters (e.g. wave celerity and hydrodynamic dispersion) were required to route both responses so that no distinction needs to be added between the surface and subsurface states once the water reaches the channels. A detailed analysis of the channel morphology was performed using the information generated after the watershed delineation and stream generation of the study area (see the next section for more details). For the overland contribution to streamflow, distances from each reach/channel to the outlet were obtained from the extraction of the corresponding channel area (i.e. fraction of pixels) from the previously obtained area-width function (Figure 4.1a), which is called the width function when simplifying to channels. The routing function associated with each channel distance is as shown in Equation (4.8) (Figure 4.1a). This expression depends on two main physical parameters: wave celerity u in m/s, and hydrodynamic dispersion D in m^2/s . The network response function for surface runoff is then obtained by the integration of the width function and the routing function in the form of Equation (4.7). The total catchment response for the surface contribution, Q_S , is obtained by the following expression:

$$Q_s(t) = \int_0^t I_s(\tau) f_s(t - \tau) \quad (4.14)$$

where $I_s(\tau)$ is the net rainfall going to the surface partition of the basin.

The subsurface flow going to channels was routed by the routing function, $f_c(t)$, expressed in Equation (4.8) (Figure 4.1b), and an example of this function is shown in Figure 4.2c. This expression for the travel time function was selected to preserve consistency in the hydrodynamic parameters governing the channel routing for both surface and subsurface flows. Flow paths for each hillslope characterized are exemplified in Figure 4.2a-b. Path travel time distributions, $f_\gamma(t)$ (Figure 4.2d), were then generated by numerical convolution of the travel times function of all states in the path (Equation (4.10)) and later added as indicated in Equation (4.11). The total catchment hydrologic discharge was obtained through numerical integration of the fraction of the net rainfall going to the subsurface hydrologic component with its respective travel time distribution as follows:

$$Q_{sb}(t) = \sum_{\gamma \in \Gamma} \int_0^t I_{sb}(\gamma, \tau) f_\gamma(t - \tau) \quad (4.15)$$

where $I_{sb}(\gamma, \tau)$ is the net rainfall weighted by the respective probability of occurrence, and sb indicates the subsurface transport condition.

To better highlight the difference between the two schemes followed for routing the hydrologic runoff, it is necessary to understand that surface runoff is mainly controlled by the geomorphology of the catchment, which is described by an area-width function, while subsurface flow experiences soil random physical heterogeneities that can be described using a lumped-based concept such as a hillslope's mean travel time.

4.4.2 Description of the area of study and hydrologic data

The WEPP-CTM model was assessed in a mountainous forested sub-catchment of the Upper Cedar River Watershed in western Washington state US. The catchment drains a 105 km² surface area, and its elevation ranges from 477 to 1660 m (USDA, 2014). Precipitation in the watershed falls into three zones depending on elevation, aspect, and location. The first zone corresponds to lower elevations and precipitation primarily is in the form of rain, which drives the peak response of the streamflow during winter. A transition zone contains precipitation as a mixed form of snow and rain, which contributes to hydrograph peaks in winter and spring periods and highly depends

on temperature and precipitation type. Areas at higher elevations are characterized as a snow-dominated zone where snowmelt controls the peaks during the spring season (Srivastava et al., 2017). The dominant soils in the watershed include upland areas with loamy sand soils and gravelly sandy soils covering 56% and 15% of the total area of the catchment, respectively. In lower elevations, sandy loam soils are found and cover 29% of the area of the catchment (USGS, 2014; See Srivastava et al., 2017 for more details).

Vegetation within the watershed is a conifer forest of two different growing phases. The old-growth phase includes forest between 250 to 680 years old, mainly located at higher elevations, while the second phase is comprised of younger forests due to reforestation after timber harvest occurring at the lower elevations (Seattle, 2015). Of these forest areas, three main vegetation zones were characterized according to their elevation and climate (Franklin and Dyrness 1988; Henderson et al., 1992). The first forest zone, a rain-dominated area, is the Western Hemlock Zone encompassing primarily western hemlock and Douglas fir located at elevations below 800 m. At elevations between 800 and 1200 m, forests including Pacific silver fir and western hemlock constitute the Pacific Silver Zone, which is a snow-dominated area, and snow interception is about 60%. A third forest zone, the Mountain Hemlock Zone, is dominated by a less dense stand of Pacific silver fir and mountain hemlock located above 1200 m (Srivastava et al., 2017).

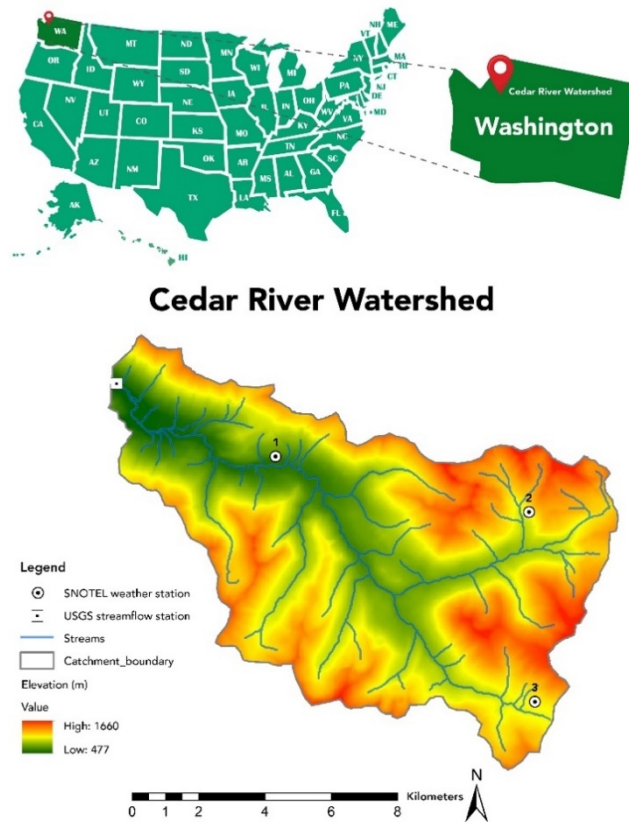


Figure 4.3 Location map of the Upper Cedar River Watershed in western Washington, US.

Daily weather within the watershed is recorded at different elevations (Figure 4.3). These three stations are managed by the Natural Resources Conservation Service (NRCS) Snow Telemetry (SNOTEL) network. Variables measured include daily precipitation, snow depth, snow water equivalent, maximum temperature, and minimum temperature. The annual precipitation from the three weather stations varies from 1727 to 3929 mm. Srivastava et al. (2017) found that precipitation in the watershed varies with both elevation and air temperature. While precipitation increases at $\sim 0.13 \text{ mm m}^{-1}$ during the winter and decreases $\sim 0.01 \text{ mm m}^{-1}$ in the summer, mean air temperature during both winter and summer decreases with elevation by $0.004 \text{ }^{\circ}\text{C m}^{-1}$. These two conditions indicate that spatial variability plays an important role in the hydrologic response of the catchment in this study. Streamflow at the outlet of the watershed was measured by USGS station 12115000 ($47^{\circ} 22' 13'' \text{ N}$, $121^{\circ} 37' 26''$) located upstream of Cedar Lake.

4.4.3 Runoff generation

The WEPP model is a complex process-based hydrology and erosion model designed to not only simulate hillslope hydrology but also to assess the impact of management practices on soil erosion and sediment delivery. The model has been applied and validated in different parts of the United States and around the world (Pandey et al., 2009; Pieri et al., 2007; Singh et al., 2011). Hillslope hydrologic processes simulated in the WEPP model include infiltration, surface runoff, lateral flow, and overland erosion (detachment and deposition) (Flanagan and Livingston, 1995; Flanagan and Nearing, 1995). Only one investigation related to water quality and transport has been found at the hillslope scale (Wang et al., 2017) and none at the watershed scale.

The watershed application of the WEPP model, GeoWEPP (Renschler, 2003), or online interfaces allow the discretization of the watershed into multiple hillslopes and channels under different drainage densities. Hydrology is calculated for each hillslope, and then flow is transported from the channels to the catchment outlet by a routing algorithm (Wang et al., 2010). However, only limited studies have assessed the model at watershed scales (Abaci and Papanicolau, 2008; Abaci and Papanicolau, 2009; Srivastava et al., 2017; Brooks et al., 2016; Srivastava et al., 2019). Two of the primary limitations are that the model routing algorithm and the baseflow contributions due to groundwater interactions have not been fully integrated into the model.

Acknowledging that the hillslope version of the WEPP model is well validated and its ability to predict the hydrologic response at hillslope scales as well as the potential of the current geospatial applications of the model to discretize a watershed into multiple hillslopes and channels, this study focused on developing a new channel routing approach through the proposed model (WEPP-CTM). As a result, a previously calibrated and validated hillslope simulation for the Cedar River Watershed was used to test the suitability of the WEPP-CTM to route the hydrologic response within hillslopes and channels; details on calibration and validation can be found in the Srivastava et al. (2017) study.

4.4.4 Calibration and validation

WEPP-CMT model simulations were run at a daily time step with a computational increment of 15 minutes. Daily observed streamflow records were divided into two periods: model calibration period (2009-2010) and the validation period (2011). A warm-up period for the model was not

required since hillslope hydrologic processes were already calibrated in a previous study. The calibrated parameters in WEPP-CMT are listed in Table 4.1. Manual calibration of those parameters was performed until the best efficiency of the model to predict daily flow discharge was obtained.

Table 4.1 WEPP-CTM calibration parameters

Parameter	Notes	Initial value	Parameter range	Calibrated parameter
α_i	Sub-surficial hillslope's mean residence time (hours/(km ²) ^{0.38})	10.0	5-100	14.2
\bar{u}_i	Uniform-flow bank full velocity (m/s)	1.0	0.5-5.0	0.63
D_i	Hydrodynamic dispersion coefficient (m ² /s)	500.0	100-5000	1000
r	Velocity ratio between overland and channel wave celerity (m/s)	1.0	1.0-5.0	1.77

Calibration parameters such as r and α_i govern the surface and subsurface hydrology at hillslopes states, respectively. Routing at channel states are dominated by \bar{u}_i and D_i . Suitable ranges for the last two parameters were taken from available literature (Deng et al., 2001; Sahin, 2014).

4.4.5 Performance evaluation

Statistical measures can be used to quantify the differences between the predicted and observed values and evaluate the model efficiency. In this study, three statistical measures were used to assess the quality and reliability of the WEPP-CTM model to predict flow discharge at the sub-basin of the Upper Cedar River Watershed: Root Mean Square Error (RMSE) (Liang et al., 2018), Percent Bias (PBIAS) (Yapo et al., 1996), and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - S_i)^2}{n}} \quad (4.16)$$

$$PBIAS = 100 \frac{\sum_{i=1}^n (S_i - O_i)}{\sum_{i=1}^n O_i} \quad (4.17)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2} \quad (4.18)$$

where n is the number of observed-simulated values, O_i is the i th observed value, and S_i is the i th simulated value. RMSE represents the absolute difference between observed and simulated values; MAE measures the average magnitude of the errors between the observations and predictions without considering their direction; PBIAS quantifies the average tendency of the simulated values to be larger or smaller than the observed values, and NSE measures the model performance.

The study of similarity between travel time distributions, or the lack thereof, was assessed by using a divergence measure. In particular, a measure of distances between two or more distribution functions is required (Wolpert and Macready, 2007). In this study, the generalized Jensen-Shannon divergence (JSD) was used to evaluate the similarity among multiple travel time distributions. The JSD is defined as (Grosse et al., 2002):

$$JSD(p_1, p_2, \dots, p_n) = H\left[\frac{1}{n} \sum_{i=1}^n p_i\right] - \frac{1}{n} \sum_{i=1}^n H[p_i] \quad (4.19)$$

where p_i are the probability distributions ($i = 1, \dots, n$), $H[p] = -\sum_{j=1}^k p(j) \log_2 p(j)$ is the Shannon entropy of p (Shannon, 1948). The JSD adopts values from 1-0, meaning 1 has perfect similarity, whereas 0 has a completely dissimilar set of probability distributions.

4.5 Results and discussion

4.5.1 Catchment geomorphology

Discretization of the catchment in the Upper Cedar River Watershed using GeoWEPP generated 253 hillslopes and 103 channels. Based on the Horton-Strahler number method, the catchment was composed of 52 first-order channels, 21 second-order channels, 7 third-order channels, and 23 fourth-order channels. The number of paths found was 253, which matches the number of hillslopes. The longest transitions consisted of one hillslope and 30 channels with a travel distance of 22 km, while the shortest path encompassed two transitions (1 hillslope and 1 channel) and a travel distance of 0.43 km.

Characterization of the travel distances included flow length to streams and distances from any point of the catchment to the outlet. Thus, distances at hillslope and channel network states were calibrated to simulate differences in velocity in both states. A full description of both

hillslopes and channels is shown in Figure 4.4. Flow lengths from the land to the streams showed a maximum distance of 2.39 km (Figure 4.4a), while the maximum distance to the catchment outlet was 22.74 km (Figure 4.4b). This spatial description was used for the parametrization of the WEPP-CMT model to predict the surface hydrologic response as follows in the next sub-section.

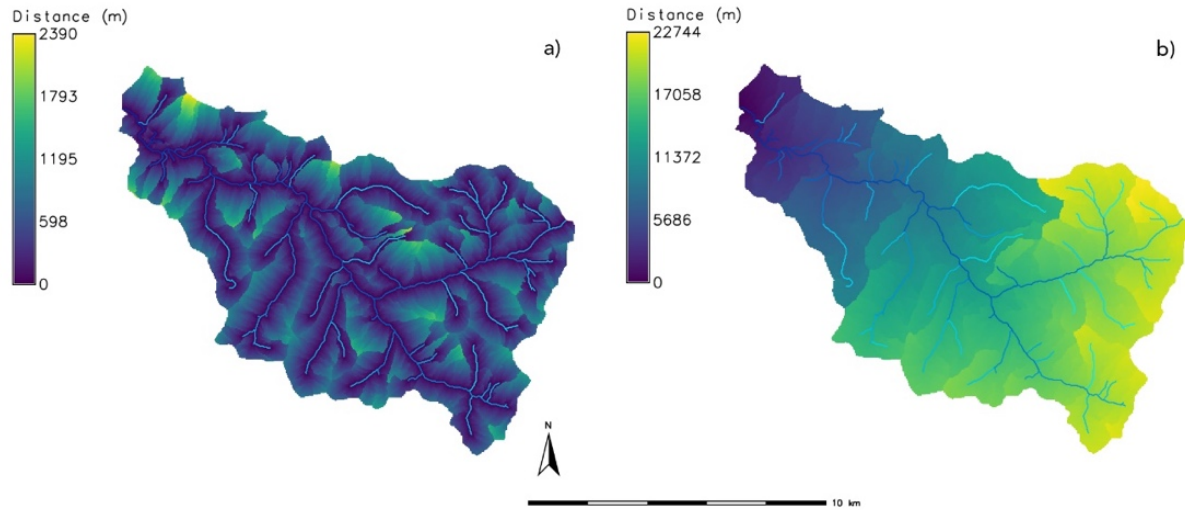


Figure 4.4 Flow distance characterization: a) Distance to the closest stream; b) Distance to the outlet.

4.5.2 Surface travel time distribution

4.5.2.1 Area-width function

As previously discussed, the area-width function can be seen as the map of the distances from any point in a catchment to the outlet following the drainage direction. In Figure 4.5a, the distribution of flow distances is described. Notice that the function redefined a two-dimensional basin characteristic (i.e. flow distances) into a one-dimensional descriptor independent of climate, geology, and vegetation. Additionally, the area-width functions resemble the basin shape, being narrow closer to its outlet, wider at distances of about 20 km, and narrow at maximum distances (Figure 4.4b). Studies on extracting basin boundaries from area-width functions can be found in the literature (Marani et al, 1991; Rinaldo et al., 1995).

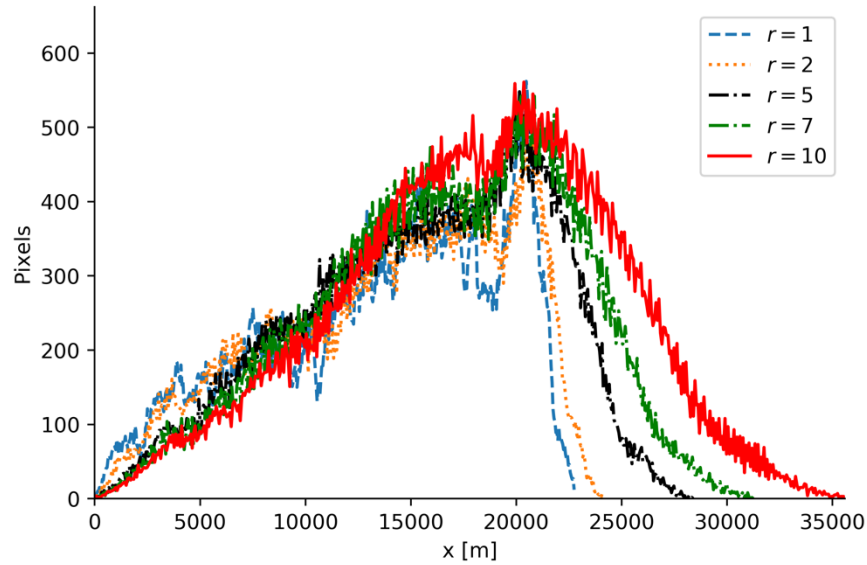


Figure 4.5 Area-width function of the catchment in the Upper Cedar River Watershed: a) distance to the outlet distribution, b) normalized width-area.

The width-area function for the catchment in the Upper Cedar River Watershed, when no difference between hillslope and channel states was imposed, overlooked the role of transitions from areas of divergent topography to areas of convergent topography with or without channelization ($r = 1$, dashed blue line in Figure 4.5). The addition of this distinction through the factor r in Equation (4.9) allowed modifying this condition. Width-area functions especially show a skewed-right noisy shape when velocities in hillslope and channels are assumed equal (dashed blue line in Figure 4.5). This behavior changed when flow velocities in the hillslopes were assumed to be slower than in the channels, especially for those values of r greater than 2.0, making the hillslopes' distance increase and creating smoother rescaled width-area functions (Figure 4.5). However, these functions do not entirely resemble the skewed left feature observed in natural hydrologic response. As a result, it can be claimed that the smooth shape and pronounced skewness of real-life hydrographs is not solely because of the addition of different velocities between hillslope and channels but it also involves dispersion mechanisms. This finding contradicts and reinforces the conclusions by Rinaldo et al. (1995) and Rinaldo et al. (2005). A contradiction was found since the distinction of hillslope and channel velocities was the only factor that changed the skewness of the hydrologic response as for the watershed in Rinaldo et al. (1995). However, once the dispersion mechanism was included through the introduction of a transfer function, the shape

of a more realistic hydrologic response was obtained as will be shown in the next section (Figure 4.6). As a result, both hydrodynamic dispersion and variable wave celerities between hillslope and channels need to be included in the formulation of transport by travel time distributions, especially for cases where hillslopes are where most solute generation within the carrier runoff occurs. It is important to mention that no distinction between hillslopes and channels has been introduced in both the unit hydrograph approach and its geomorphological version.

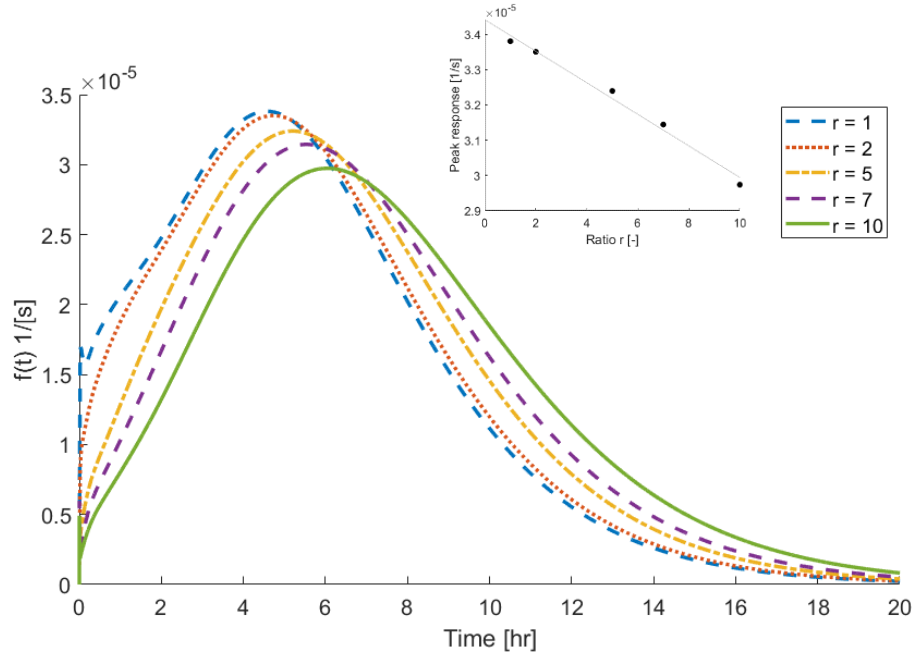


Figure 4.6 Surface travel time distribution for different distance amplification factors $r = u_c/u_h$.

4.5.2.2 Travel time distribution

The presence of both dynamic effects (i.e. wave celerity and hydrodynamic dispersion) through Equation (4.8) and its numerical integration over the domain of all catchment distances and probabilities obtained in the previous section are shown in the travel time distributions in Figure 4.6. All the distributions show a relatively fast rising limb and a long tail, which are typical characteristics in a basin's hydrologic response. The case of uniform velocities over the catchment ($u_c/u_h = 1$) had a higher peak among all the overland response functions. Once the slower velocities were imposed for overland runoff ($u_c/u_h = 2, \dots, 10$), decreases in the peak response and time to peak were obtained (Figure 4.6). The peak response of the travel time distribution

showed an exponential decrease for higher values of the amplification factor r as can be seen in the inset of Figure 4.6.

The analysis of the probability density functions for the surface component of the hydrologic runoff contributed to a more comprehensive understanding of travel times in hillslope and channel states. However, it is important to mention that hillslopes in this approach were assumed to have the same overland velocity to maintain the simplicity of the model. The relaxation of this assumption is a topic for another paper, where a more detailed description of spatial distribution of hillslope velocities can be characterized. Nonetheless, this analysis served as a tool for model calibration.

Manual calibration of the model showed that the suitable r factor was equal to 1.77, which means that hillslope flow velocities are about half the velocities seen in channels. The Jensen-Shannon divergence for the surface travel time functions was equal to 0.03, which indicates that travel time distributions for the surface runoff do not differ significantly and suggests that the assumption of differences in velocities in this approach plays a relatively small role in the selection of an appropriate distribution function. However, this may change for other watersheds where travel time becomes significantly larger on hillslope surfaces.

4.5.3 Subsurface travel time distribution

As explicitly expressed in the methods section, subsurface travel times in this approach were divided into two states: hillslopes and channel contributions. Whereas surface response was described in terms of the geomorphologic characteristics of the landscape, subsurface conditions are, in reality, more complex domains where soil heterogeneities vary significantly within relatively small areas. Consequently, flow movement on hillslopes was expressed as a lumped-based exponential model with mean travel time calculated as a function of the catchment area. Discussion of results is presented in the following sub-sections.

4.5.3.1 Hillslope contribution

Mean travel times in subsurface conditions played a significant role since hillslopes represent the state where solute generation within the hydrologic runoff occurs. As a result, the implementation of the distributions governing the transit of water in the subsurface hillslope states

used in this study was also adapted to extend this framework to solute transport in future research. In Figure 4.7, the distribution functions of travel times for the subsurface component of the hydrologic runoff for different mean residence times ($\alpha = 10 \dots 100 \text{ hours}/(\text{km}^2)^{0.38}$) are presented. These curves were estimated using Equation (4.11) after the numerical convolution of the 253 paths' travel time distributions identified within the catchment and flow velocity and their respective probabilities. Dispersion coefficient values adopted were 1 m/s and 1000 m²/s, respectively, to facilitate analysis of travel times in Equation (4.12).

Probability density functions showed significant variability under different values of the factor α , which controls the mean travel time for all hillslopes in the watershed. The higher the magnitude of α , the longer the residence time of water particles in hillslopes. For $\alpha = 10 \text{ hours}/(\text{km}^2)^{0.38}$, travel times tended to reflect a hillslope under a predominance of convection and relatively low storage-diffusion effects (dashed blue line in Figure 4.7). When higher values of α were adopted ($\alpha = 20, \dots, 100$), flow and transport were storage-diffusion dominated. The mean travel time that presented the greatest difference was $\alpha = 100 \text{ hours}/(\text{km}^2)^{0.38}$ (dotted blue line in Figure 4.7). The time to peak of the functions occurred at similar times, while the peak response decreased exponentially with greater values of α (inset in Figure 4.7). Manual calibration of the model showed that the optimal α factor was equal to 14.2 hours/ $(\text{km}^2)^{0.38}$.

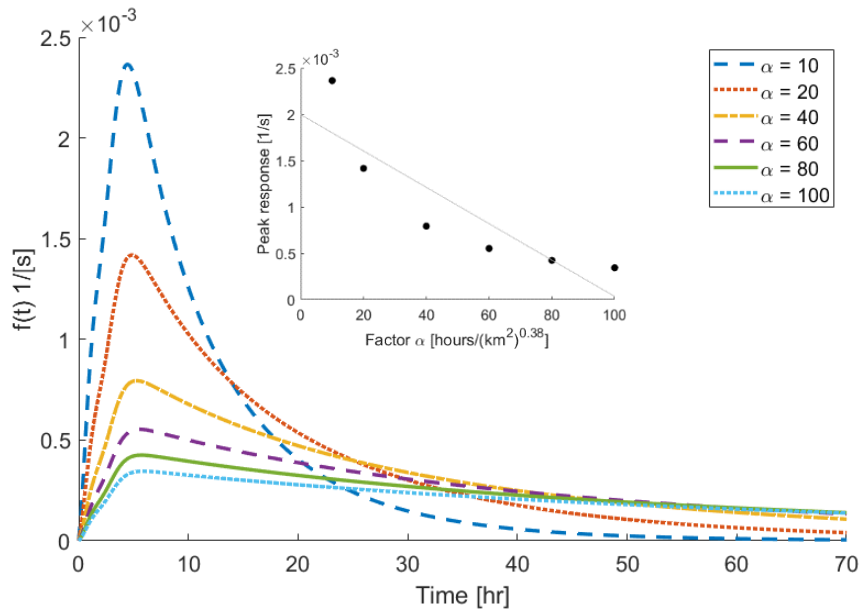


Figure 4.7 Subsurface travel time distributions for different mean travel times. Flow velocity, $u = 1.0 \text{ m/s}$; and hydrodynamic dispersion, $D = 1000 \text{ m}^2/\text{s}$.

A significant difference was found among the travel time distributions with a Jensen Shannon divergence equal to 0.13. As expected, travel times highly depended on hillslope areas, and the average residence time for a range of hillslope areas was equal to 12.89 hours. A limitation found in this approach is that the factor α was assumed to be the same for all hillslopes within the catchment. Having said that, this does not mean that travel time was assumed the same as it depends on the hillslope area, making the analysis presented here a contribution to understanding the travel time theory underlying the model and its calibration.

4.5.3.2 Channel contribution

For channel states, the hydrodynamic conditions governing flow are the same as those previously discussed for surface conditions. Although the discussion of the difference between hillslope and channel velocities and its incorporation into the area-width function was established, the effect of the parameter r was ignored and it was kept equal to its calibrated value. In this section, the sensitivity of the probability density function of the travel times in channels to the flow velocity and hydrodynamic dispersion are discussed. To isolate the effect of hillslope travel time distributions, the factor α was set equal to its calibrated value (14.2 hours/(km²)^{0.38}).

An obvious hydrodynamic variable that modifies the travel time in channels is the flow velocity. In this particular case, the uniform flow bank full velocity in the channel modified the travel time distribution, as shown in Figure 4.8. It can be noticed that Equation (4.8), used as the routing function, depends directly on the wave celerity and not on flow velocity. A common equivalence between these parameters was adopted as $a = 1.5u$, where u is the flow velocity (Paik and Kumar, 2004). Higher velocities tend to make a more convective channel network (solid green line in Figure 4.8). On the contrary, when low values of the bank full velocities were imposed, a channel network dominated by storage and diffusive effects was obtained.

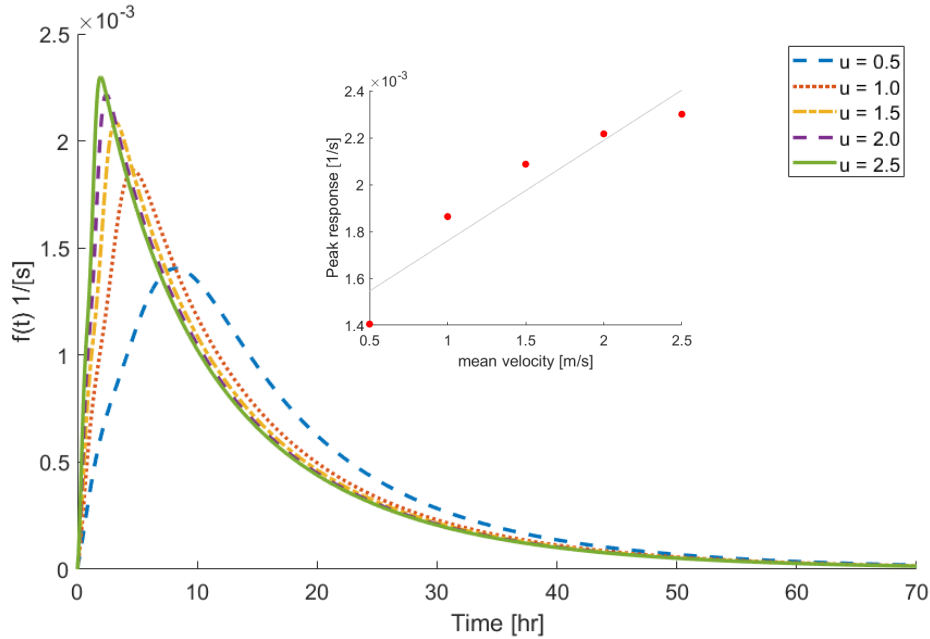


Figure 4.8 Travel time distribution under different flow velocities in the channel.

Both time to peak and peak response were significantly modified when flow velocities were varied. Peak response increased exponentially with respect to flow velocities (inset in Figure 4.8), with $u = 2.5$ m/s the most convective response and $u = 0.5$ m/s the smallest response. The Jensen Shannon divergence among the travel time distributions for the different flow conditions was equal to 0.27, which suggests that the travel time in channels was significantly modified by the bank full velocity in channels, with the latter equal to 0.63 m/s after manual calibration.

A second hydrodynamic variable important for the prediction of travel time in channels was the hydrodynamic dispersion as can be seen in Equation (4.8). Figure 4.9 shows the travel time distributions for a range of dispersion coefficients. Flow velocity u and parameter α were set equal to their calibration values. Graphically, travel time distributions did not seem to be significantly sensitive to changes in hydrodynamic dispersion, being the peak response was the most affected. The Jensen Shannon divergence of the probability density functions was equal to 0.00053, which corroborates the claim that dispersion coefficient changes do not significantly affect the peak response of the hydrologic response. However, as previously discussed, the dispersion mechanism is responsible for the skewed left shape of the travel time distribution. The dispersion coefficient used for the subcatchment in the Upper Cedar River Watershed was set to 1000 m²/s for calibration purposes.

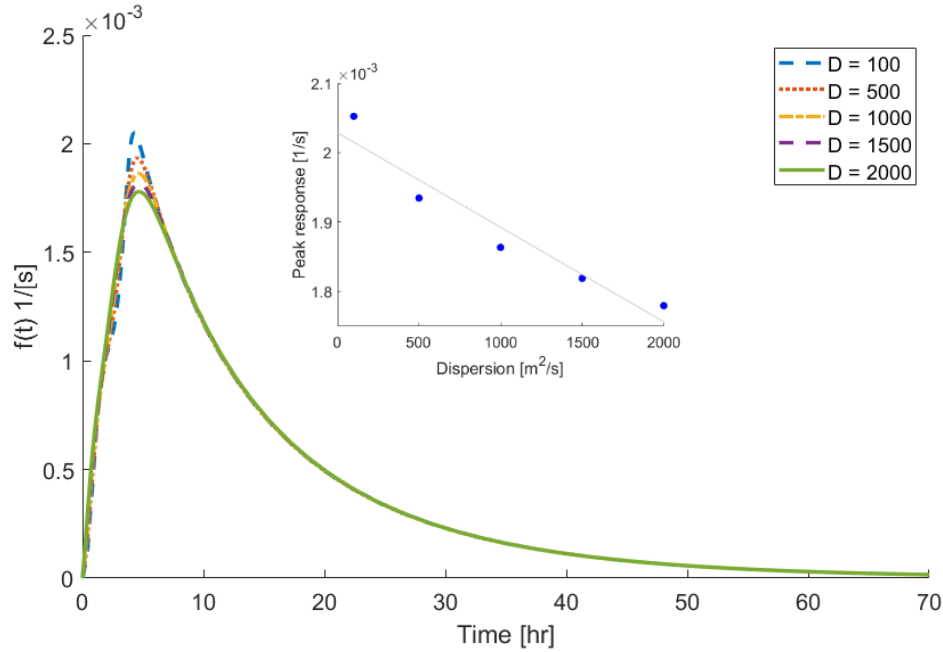


Figure 4.9 Travel time distribution under different hydrodynamic dispersion coefficients in the channel.

4.5.3.3 Rainfall heterogeneity

In previous analyses, rainfall variability was not incorporated into the sensitivity analysis in order to maintain simplicity and avoid the generation of multiple scenarios for each rainfall event. The effect of rainfall heterogeneities in the stochastic framework adopted here through Equation (4.12) showed that temporal variability of path probabilities played a relatively small role in shaping the hydrologic response for the subcatchment in the Upper Cedar River Watershed (Figure 4.10). However, previous analysis of the spatial distribution of the precipitation in this catchment varied with elevation and temperature (Srivastava et al., 2017), which is a common characteristic in mountainous basins. As a result, precipitation heterogeneity is therefore transferred to the surface and subsurface response of the basin.

Figure 4.10 shows the calibrated travel time distribution for subsurface and channel conditions under different precipitation events. The time to peak and peak responses depend on both rainfall spatial distribution and the magnitude of the rainfall event (Figure 4.10). A significant similarity among the probability density functions was found with a JSD = 0.012.

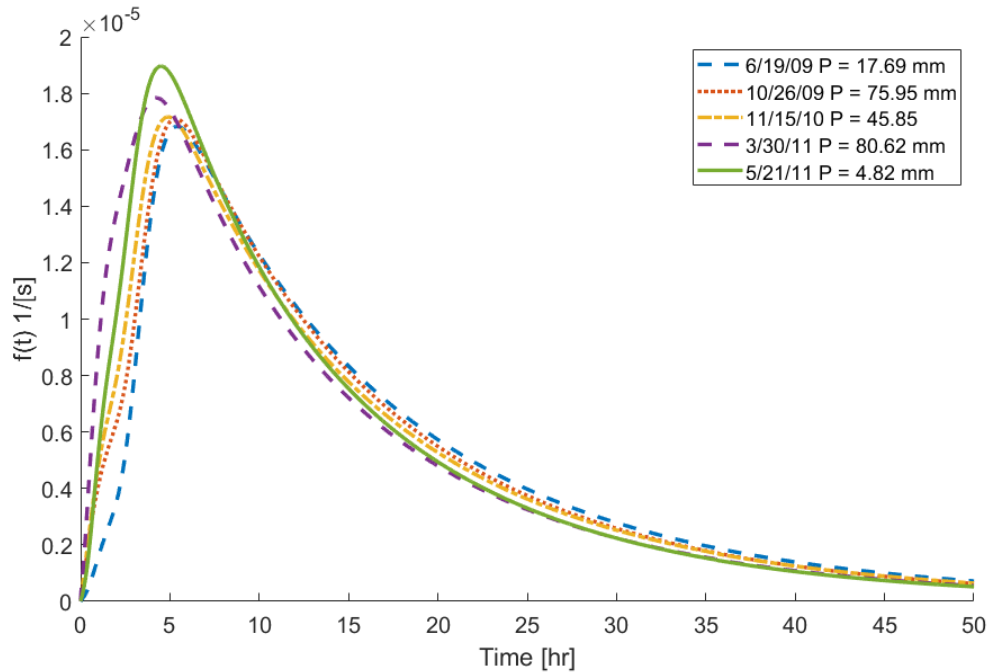


Figure 4.10 Temporal variability of travel time distribution for subsurface and channel conditions in the Cedar River Watershed.

4.5.4 Catchment total response

The total hydrologic response of the catchment was calculated by numerical convolution of each component of the hydrologic response function (i.e. surface and surface compartments) and their corresponding rainfall partitions (i.e. net rainfall) through Equations (4.14-4.15). The temporal distribution of flow discharge at the outlet of the Cedar River Watershed for both calibration and validation periods is displayed in Figure 4.11. Predictions using WEPP-CTM agreed well with the observations, except for months during the spring where snowmelt occurred. Hydrograph peaks for the spring season tended to be underpredicted for the calibration period while during the validation period they were overpredicted. Prediction results obtained by using the Muskingum-Cunge method (WEPP-MK) developed by Wang et al. (2010) were included for comparison. Both methods showed similar predictions, with the WEPP-MK more accurate in estimating peak discharges.

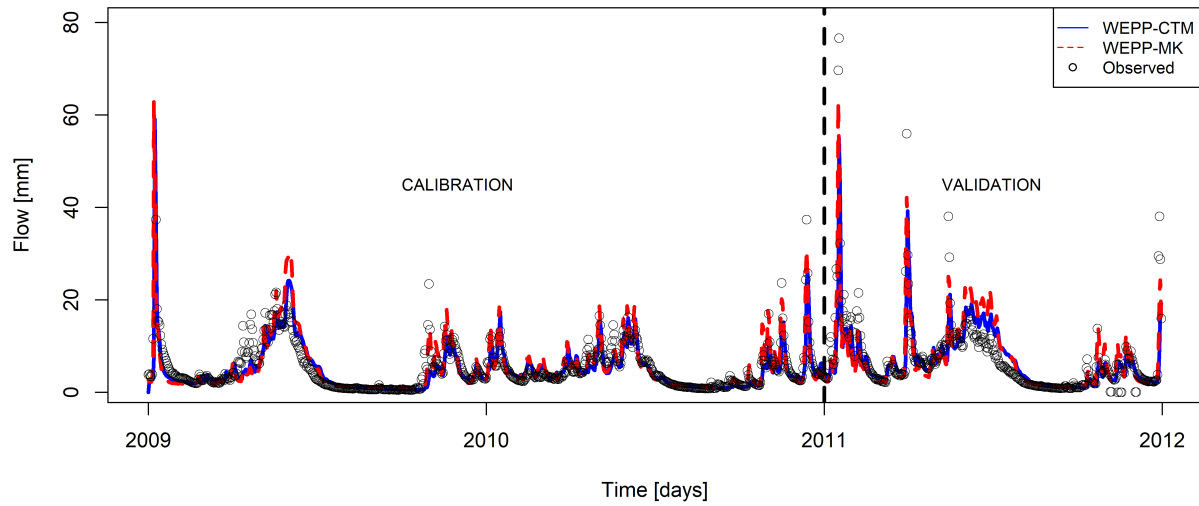


Figure 4.11 Daily flow discharge at the outlet of the simulated subcatchment of the Upper Cedar River Watershed.

Table 4.2 shows the differences and model efficiencies between the observed daily flow and predicted daily flow for the Upper Cedar River Watershed for both the calibration and validation periods using both WEPP-CMT and WEPP-MK. The RMSE for both models were of similar magnitude. WEPP-CMT presented slightly higher errors (RMSE = 5.2 mm) in comparison with the ones obtained with WEPP-MK (RMSE=4.0 mm). The proposed transport model in this study tended to underestimate the daily flow amount with PBIAS= -2.8%. Similar percent bias was found for the conventional routing method (PBIAS=-3.1%), which is a condition expected since as the water balance for both models are equivalent (Table 4.3). Small differences were introduced during the calculation of the numerical convolution but considered negligible for this study. Correlation between observations and predictions were in the acceptable range for both the WEPP-CMT and WEPP-MK models, being $R^2=0.57$ and $R^2=0.71$, respectively (Ibarra-Zavaleta et al., 2017). Nash-Sutcliffe model efficiency values for predicted daily flow discharge values using WEPP-CTM and WEPP-MK were 0.57 and 0.70, respectively. This difference was due to the capacity of the conventional method (WEPP-MK) to capture peaks more accurately than the model based on travel time distributions (WEPP-CTM) (Figure 4.11). However, both models were in the acceptable range of prediction (Ibarra-Zavaleta et al., 2017; Abouelnour et al., 2019).

Table 4.2 Performance evaluation of daily flow discharges for Cedar River Watershed for both calibration and validation periods.

Model	RMSE (mm)		PBIAS (%)		R ²		NSE	
	Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
WEPP-CMT	5.20	5.09	-2.8	2.1	0.57	0.66	0.55	0.65
WEPP-MK	4.00	4.00	-3.1	2.6	0.71	0.77	0.70	0.77

For the validation period, RMSE values were similar to those found in the calibration period; the WEPP-CMT model had RMSE = 5.09 mm while the WEPP-MK model had a RMSE = 4.00 mm (Table 4.2). The mass transport model showed improvement but still underestimated daily flow with a PBIAS = -3.1%, whereas the conventional routing shifted to daily flow overprediction with PBIAS = 2.6%. Improved correlation between observed values and predictions for both models considered here was observed. Correlation coefficients showed acceptable ranges for performance evaluation for hydrological models. Model efficiencies improved for the validation period. Nash-Sutcliffe model efficiency values were 0.65 and 0.77 for WEPP-CMT and WEPP-MK, respectively.

Accumulated monthly streamflow for both calibration and validation periods are shown in Figure 4.12. Similar to daily conditions, streamflow was often under predicted by WEPP-CMT, except for winter months where WEPP-MK also showed similar underprediction. However, for most months the routing model based on travel time distributions predicted ~80% of the monthly streamflow that occurred in the subcatchment of the Upper Cedar River Watershed accurately.

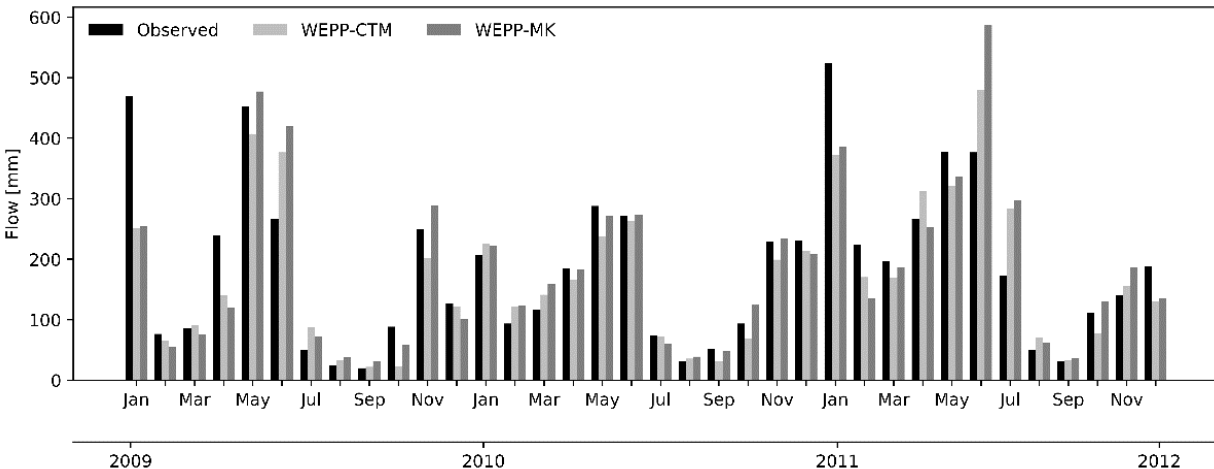


Figure 4.12 Monthly flow discharge at the outlet of the modeled subcatchment of the Upper Cedar River Watershed.

Water balance for water years 2009-2010 is shown in Table 4.3. About 20% of the input water left the watershed in the form of evapotranspiration for the three water years. Link et al (2004) reported a similar ET (21%) for old growth coniferous trees exceeding 450 years and located 405 km south the Cedar River Watershed and mean annual temperature of 8.7 °C. For younger ages of coniferous forests, 255 km northwest of the Cedar River Watershed, Jassal et al. (2009) measured ET using the eddy covariance technique. The measured ET was about 23% of the annual precipitation. Therefore, the simulated ET in this study is comparable with values found in literature considering the similarities in climate. The fraction of the input water contributing to baseflow was very small in comparison to the amount of subsurface flow (~70%). Surface flow was ~10% of the precipitation in the analyzed water years. No groundwater recharge was assumed in the watershed as an impervious layer was present at a depth of 0.60 m from the soil surface, so that the major contribution to the hydrologic runoff was by subsurface flow.

Table 4.3 Annual water balance from the WEPP-CTM model for 2009-2011 water years.

Water year	P (mm)	RM (mm)	ET (mm)	D (mm)	Δ TSW (mm)	Δ GW (mm)	Q _B (mm)	Q _S (mm)	Q _{SB} (mm)
2009	2590	2538	551	43	-68	0	43	291	1721
2010	2494	2467	485	43	151	0	43	96	1692
2011	3348	3281	499	43	-66	0	43	335	2470

Note: P = precipitation, RM = rain plus snowmelt, ET = evapotranspiration, D = deep percolation from the soil profile, Δ TSW and Δ GW = yearly change in soil water and groundwater, Q_B = baseflow, Q_S = surface runoff, and Q_L = lateral flow.

Among the available studies dealing with hydrology at a large scale, some of the most similar investigations to the case study here are shown in Table 4.4. The WEPP model performed well for relatively small catchments (0.09-40 km²) with NSE varying from 0.42 to 0.94 for both calibration and validation scenarios. Larger catchment applications using WEPP included studies where NSE ranged from 0.55 to 0.82, and only three studies were found. Other well-known models, such as SWAT, commonly applied in agricultural watersheds include studies at watershed scales with drainage areas varying from 44 to 400 km². Overall, SWAT efficiencies were in acceptable ranges (0.34-0.94) and several studies could be found in the literature. Between 2016 and 2019, 1538 SWAT publications were identified (https://www.card.iastate.edu/swat_articles). This indicates that the SWAT model has experienced continuous development and application for water quantity and quality research at catchment scales.

Table 4.4 Comparison of model performance (NSE) of large-scale hydrological models and WEPP model for flow discharge estimation.

Reference	Watershed	Model	Drainage area (km ²)	Time period		NSE	
				CAL	VAL	CAL	VAL
Conroy et al. (2006)	North Fork Caspar Creek Experimental Watershed (U.S)	WEPP	4.24	1989-1991	1992-1997	0.42	0.70
Pandey et al. (2009)	Karso Watershed (India)	WEPP	27.93	1996	1992, 1993, 1995, 1997, 2000	0.92	0.88
Dun et al. (2009)	Hermada watershed (U.S.)	WEPP	0.09		1995-2000		0.45
Singh et al. (2011)	Umroi Watershed (India)	WEPP	2.39	2003	2004	0.94	0.87
Brooks et al. (2016)	Lake Tahoe Basin (U.S)	WEPP	19.12-36.75	1989-2005	1989-2005		0.37-0.42
Gould et al. (2016)	Little Salmon River Watershed (U.S.)	VIC-WEPP	1491	1979-1994	1995-2010	0.76	0.75
Afshar and Hassanzadeh (2017)	Torohg Dam Watershed (Iran)	WEPP	163.12	2004-2005	2006-2007	0.854	0.824
Srivastava et al. (2017)	Cedar River Basin (U.S.)	WEPP	105	1997-2003	2004-2011	0.55	0.60
Rouhani et al. (2007)	Grote Nete River (Belgium)	SWAT	383	1986-1989	1990-1995	0.67	0.66
Geza and McCray (2008)	Turkey Creek (U.S.)	SWAT	126	1998-2001		0.27-0.77	-0.01-0.70
Glavan et al. (2012)	River Axe (England)	SWAT	400	1988-1997	1998-2005	0.62	0.47
Mishra and Kar (2012)	Banha Watershed (India)	SWAT	16.95	1996	2000, 2001	0.70	0.62-0.70
Cerro et al. (2014)	Alegria Watershed (Spain)	SWAT	53	2009-2010	2010-2011	0.68	0.49
Fohrer et al. (2014)	Kielstau Watershed (Germany)	SWAT	50	2003-2005	2006-2009	0.83	0.76
Oliver et al. (2014)	Big Haynes Creek (U.S)	SWAT	44	2003-2006	2007-2010	0.49	0.37
McMichael et al. (2006)	Jameson catchment (U.S)	MIKE-SHE	34	1962-1981	1983-1992	0.80	0.60
Ping et al. (2017)	Gujiao Watershed (China)	MIKE-SHE	1584	1961-1975	1976-1980	0.87	0.94
Sonnenborg et al. (2017)	Skjern and Lejre catchments (Denmark)	MIKE-SHE	1098, 465	2008-2010, 2002-2008	2011, 2009		0.67, 0.49
Dan et al. (2012)	Heihe River basin (China)	VIC	11.6 × 10 ⁴	1986-1990	1993-1997	0.74	0.72
Meng et al. (2017)	Jinshajiang River Basin (China)	VIC	3.35 × 10 ⁴	2002-2007	2008-2010	0.78	0.78
Lu et al. (2017)	Yellow River, Yangtze River (China)	VIC	12.19 × 10 ⁴ , 13.77 × 10 ⁴	1981-1996	1997-2000	0.72, 0.75	0.75, 0.79

CAL: Calibration period
VAL: Validation period

Other model applications for MIKE-SHE and VIC can be also found in the literature with acceptable performance (NSE ~ 0.70), with VIC applied to very large areas. Comparing model performances in Table 4.4 and the WEPP-CTM implementation presented in this research, one can conclude that WEPP-CTM was in the range of acceptable efficiency (NSE ~ 0.62) among well-known models for watershed scale applications in this very limited case study application.

4.6 Concluding remarks

This paper presented the formulation of a new approach for hillslope and channel routing based on catchment geomorphology and travel time distributions. The geomorphological approaches of the hydrologic response pursued herein showed a suitable stochastic framework for streamflow routing and potential transport of reactive solutes, which is mainly controlled by contact times of the hydrological carrier between the mobile and immobile phases. The model was tested to simulate the transport of the hydrologic response in a snow-dominated mountainous watershed, a subcatchment of the Upper Cedar River Watershed, which is located in western Washington state, U.S. Travel times on hillslopes had a relatively small effect on shaping of the surface hydrologic response functions. Surface travel time distributions showed relatively low sensitivity when imposing different hillslope flow velocities (JSD = 0.03), being that the flow velocities in the channels were almost twice as high as those on the hillslopes ($r = 1.77$). However, response functions tend to be dissimilar when channel velocities increase.

Subsurface response functions of the catchment highly depended on the hillslope mean travel time. Long travel times shaped a response function dominated by storage-effects ($\alpha = 100$ hours/(km²)^{0.38}), whereas convection-dominated responses were observed when travel times decreased ($\alpha = 10$ hours/(km²)^{0.38}). Probability distributions were dissimilar for small changes in hillslope mean travel times (JSD = 0.3). The channel routing for water leaving the hillslope as subsurface flow seemed to be highly sensitive to changes in flow velocity. Hydrologic response functions had shorter times to peaks and higher peaks for small increments in flow velocities. The dispersion coefficient in channels was the least sensitive parameter (JSD=0.00053) for both surface and subsurface stochastic approaches. However, the inclusion of the dispersion mechanism in the hydrologic response function provided the invariant skewed-left shape seen in natural hydrographs. Additionally, rainfall spatial heterogeneities had a relatively small effect on temporal variability

of the hydrologic response for the study area but only in overall characteristics such as time to peak and peak response.

The applicability of the WEPP-CTM to predict flow discharges at the outlet of the subcatchment of the Upper Cedar River Watershed showed an acceptable performance. Overall, good correlation between observed and predicted flow discharges was achieved ($R^2 = 0.62$). Relatively small underprediction of streamflow was detected for both calibration and validation periods with PBIAS = -2.8% and 2.1%, respectively. The overall Nash Sutcliffe Efficiency was 0.60, which is an acceptable value in hydrologic modeling. Nonetheless, an existing routing model, WEPP-MK, had a greater model efficiency (NSE = 0.73). This suggests that the new model assumptions might need to be relaxed, especially the assumption of a uniform wave celerity over the entire catchment, which seems to be inaccurate for mountainous watersheds where velocity gradients are higher than in relatively flat watersheds.

In summary, the routing model presented in this study showed an acceptable performance in predicting streamflow at the outlet of the subcatchment of the Upper Cedar River Watershed. The interaction of a routing method based on travel time distributions and the WEPP hillslope model opens a window for modeling of transport of reactive solutes (e.g. nitrates), which is an unexplored field for the WEPP model at watershed scales.

4.7 References

- Abaci, O., & Papanicolaou, A. N. (2008). Evaluating the performance of the Water Erosion Prediction Project (WEPP) model for larger watersheds. *World Environmental and Water Resources Congress 2008: Ahupua'a*, 1–10.
- Abaci, O., & Papanicolaou, A. N. (2009). Long-term effects of management practices on water-driven soil erosion in an intense agricultural subwatershed: Monitoring and modelling. *Hydrological Processes*, 23(19): 2818-2837.
- Aboelnour, M., Gitau, M. W., & Engel, B. A. (2019). Hydrologic response in an urban watershed as affected by climate and land-use change. *Water*, 11(8), 1603.
- Afshar, A. A., & Hassanzadeh, Y. (2017). Determination of monthly hydrological erosion severity and runoff in Torogh Dam watershed basin using SWAT and WEPP models. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 41(2), 221–228.

- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., van Griensven, A., Van Liew, M. W., Kannan, N., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1491–1508.
- Baffaut, C., Nearing, M. A., Ascough II, J. C., & Liu, B. (1997). The WEPP watershed model: II. Sensitivity analysis and discretization on small watersheds. *Transactions of the ASAE*, 40(4), 935–943.
- Beven, K., & Wood, E. F. (1993). Flow routing and the hydrological response of channel networks. In Beven, K. and Kirkby, M. J. (Eds), *Channel Network Hydrology*. Wiley, Chichester, pp. 99-128
- Bhandari, A. B., Nelson, N. O., Sweeney, D. W., Baffaut, C., Lory, J. A., Senaviratne, A., & Barnes, P. L. (2017). Calibration of the APEX model to simulate management practice effects on runoff, sediment, and phosphorus loss. *Journal of Environmental Quality*, 46(6), 1332–1340.
- Botter, G., Bertuzzo, E., Bellin, A., & Rinaldo, A. (2005). On the Lagrangian formulations of reactive solute transport in the hydrologic response. *Water Resources Research*, 41(4), W04008. DOI: <https://doi.org/10.1029/2004WR003544>
- Botter, G., & Rinaldo, A. (2003). Scale effect on geomorphologic and kinematic dispersion. *Water Resources Research*, 39(10), 1286. DOI: <https://doi.org/10.1029/2003WR002154>
- Brooks, E. S., Dobre, M., Elliot, W. J., Wu, J. Q., & Boll, J. (2016). Watershed-scale evaluation of the Water Erosion Prediction Project (WEPP) model in the Lake Tahoe basin. *Journal of Hydrology*, 533, 389–402.
- Cerro, I., Antigüedad, I., Srinivasan, R., Sauvage, S., Volk, M., & Sanchez-Perez, J. M. (2014). Simulating land management options to reduce nitrate pollution in an agricultural watershed dominated by an alluvial aquifer. *Journal of Environmental Quality*, 43(1), 67–74. <https://doi.org/10.2134/jeq2011.0393>
- Conroy, W. J., Hotchkiss, R. H., & Elliot, W. J. (2006). A coupled upland-erosion and instream hydrodynamic-sediment transport model for evaluating sediment transport in forested watersheds. *Transactions of the ASABE*, 49(6), 1–10.
- D’Odorico, P., & Rigon, R. (2003). Hillslope and channel contributions to the hydrologic response. *Water Resources Research*, 39(5), 1–9. <https://doi.org/10.1029/2002WR001708>

- Dagan, G. (1989). *Flow and Transport in Porous Formations*. Springer-Verlag, Berlin-Heidelberg.
- Dan, L., Ji, J., Xie, Z., Chen, F., Wen, G., & Richey, J. E. (2012). Hydrological projections of climate change scenarios over the 3H region of China: A VIC model assessment. *Journal of Geophysical Research Atmospheres*, 117(11), 1–17. <https://doi.org/10.1029/2011JD017131>
- Deng, Z.-Q., Singh, V. P., & Bengtsson, L. (2001). Longitudinal dispersion coefficient in straight rivers. *Journal of Hydraulic Engineering*, 127(11), 919–927.
- Dun, S., Wu, J. Q., Elliot, W. J., Robichaud, P. R., Flanagan, D. C., Frankenberger, J. R., & Xu, A. C. (2009). Adapting the Water Erosion Prediction Project (WEPP) model for forest applications. *Journal of Hydrology*, 366(1–4), 46–54. <https://doi.org/10.1016/j.jhydrol.2008.12.019>
- Emmet, W. (1978). Overland Flow. In (Kirkby, M. J., ed.) *Hillslope Hydrology*, pp. 145 – 176, John Wiley, New York.
- Flanagan, D. C., Frankenberger, J. R., & Ascough II, J. C. (2012). WEPP: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1463–1477.
- Flanagan, D. C., & Livingston, S. J., eds. (1995). USDA-Water Erosion Prediction Project: WEPP User Summary. NSERL Report #11. West Lafayette, Ind.: USDA-ARS National Soil Erosion Research Laboratory.
- Flanagan, D. C., & Nearing, M. A., eds. (1995). USDA-Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation. NSERL Report #10. West Lafayette, Ind.: USDA-ARS National Soil Erosion Research Laboratory. 298 pp.
- Flanagan, D. C., Frankenberger, J. R., Cochrane, T. A., Renschler, C. S., & Elliot, W. J. (2013). Geospatial application of the Water Erosion Prediction Project (WEPP) model. *Transactions of the ASABE*, 56(2), 591–601.
- Fohrer, N., Dietrich, A., Kolychalow, O., & Ulrich, U. (2014). Assessment of the environmental fate of the herbicides flufenacet and metazachlor with the SWAT model. *Journal of Environmental Quality*, 43(1), 75–85. <https://doi.org/10.2134/jeq2011.0382>
- Franklin, J. F., & Dyrness, C. T. (1988). *Natural vegetation of Oregon and Washington*. Corvallis, Oregon: Oregon State University Press.

- Geza, M., & McCray, J. E. (2008). Effects of soil data resolution on SWAT model stream flow and water quality predictions. *Journal of Environmental Management*, 88(3), 393–406.
<https://doi.org/10.1016/j.jenvman.2007.03.016>
- Glavan, M., White, S. M., & Holman, I. P. (2012). Water quality targets and maintenance of valued landscape character - Experience in the Axe catchment, UK. *Journal of Environmental Management*, 103(2012), 142–153. <https://doi.org/10.1016/j.jenvman.2012.03.009>
- Gould, G. K., Liu, M., Barber, M. E., Cherkauer, K. A., Robichaud, P. R., & Adam, J. C. (2016). The effects of climate change and extreme wildfire events on runoff erosion over a mountain watershed. *Journal of Hydrology*, 536, 74–91. <https://doi.org/10.1016/j.jhydrol.2016.02.025>
- Grosse, I., Bernaola-Galván, P., Carpena, P., Román-Roldán, R., Oliver, J., & Stanley, H. E. (2002). Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E*, 65(4), 41905.
- Gupta, A., & Cvetkovic, V. (2002). Material transport from different sources in a network of streams through a catchment. *Water Resources Research*, 38(7), 1–3.
- Gupta, V. K., Waymire, E., & Wang, C. T. (1980). A representation of an instantaneous unit hydrograph from geomorphology. *Water Resources Research*, 16(5), 855–862.
- Henderson, J. A., Leshner, R. D., Peter, D. H., and Shaw, D. C. (1992). *Field Guide to the Forested Plant Associations of the Mt. Baker-Snoqualmie National Forest*. Technical Paper R6ECOLTP028-91, USDA, Forest Service, Pacific Northwest Region. 196 pp.
- Hoghooghi, N., Radcliffe, D. E., Habteselassie, M. Y., & Jeong, J. (2017). Modeling the effects of onsite wastewater treatment systems on nitrate loads using SWAT in an urban watershed of metropolitan Atlanta. *Journal of Environmental Quality*, 46(3), 632–640.
- Huff, D. D., O'Neill, R. V., Emanuel, W. R., Elwood, J. W., & Newbold, J. D. (1982). Flow variability and hillslope hydrology. *Earth Surface Processes and Landforms*, 7(1), 91–94.
- Ibarra-Zavaleta, S. P., Landgrave, R., Romero-López, R., Poulin, A., & Arango-Miranda, R. (2017). Distributed hydrological modeling: Determination of theoretical hydraulic potential & streamflow simulation of extreme hydrometeorological events. *Water (Switzerland)*, 9(8).
<https://doi.org/10.3390/w9080602>

- Jassal, R. S., Black, T. A., Spittlehouse, D. L., Brümmer, C., & Nesic, Z. (2009). Evapotranspiration and water use efficiency in different-aged Pacific Northwest Douglas-fir stands. *Agricultural and Forest Meteorology*, 149(6–7), 1168–1178.
<https://doi.org/10.1016/j.agrformet.2009.02.004>
- Liang, H., Qi, Z., Hu, K., Li, B., & Prasher, S. O. (2018). Modelling subsurface drainage and nitrogen losses from artificially drained cropland using coupled DRAINMOD and WHCNS models. *Agricultural Water Management*, 195, 201–210.
- Link, T. E., Flerchinger, G. N., Unsworth, M., & Marks, D. (2004). Simulation of water and energy fluxes in an old-growth seasonal temperate rain forest using the Simultaneous Heat and Water (SHAW) model. *J. Hydrometeorol.*, 5(3), 443–457.
[https://doi.org/10.1175/1525-7541\(2004\)005<0443:SOWAEF>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0443:SOWAEF>2.0.CO;2)
- Linsleuy, R. K. Jr., Kohler, M. A., & Paulhus, J. L. H., eds. (1975). *Hydrology For Engineers*, 2nd Edition. McGraw Hill Kogakusha Ltd., Tokyo. 482 pp.
- Lu, W., Wang, W., Shao, Q., Yu, Z., Hao, Z., Xing, W., & Li, J. (2018). Hydrological projections of future climate change over the source region of Yellow River and Yangtze River in the Tibetan Plateau: A comprehensive assessment by coupling RegCM4 and VIC model. *Hydrological Processes*, 32(13), 2096–2117. <https://doi.org/10.1002/hyp.13145>
- Marani, A., Rigon, R., & Rinaldo, A. (1991). A note on fractal channel networks. *Water Resources Research*, 27(12), 3041–3049.
- Meng, C., Zhou, J., Dai, M., Zhu, S., Xue, X., & Ye, L. (2017). Variable infiltration capacity model with BGSA-based wavelet neural network. *Stochastic Environmental Research and Risk Assessment*, 31(7), 1871–1885. <https://doi.org/10.1007/s00477-017-1413-0>
- McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., & Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research*, 41(5).
- McMichael, C. E., Hope, A. S., & Loaiciga, H. A. (2006). Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. *Journal of Hydrology*, 317(3–4), 307–324. <https://doi.org/10.1016/j.jhydrol.2005.05.023>
- Mishra, A., & Kar, S. (2012). Modeling hydrologic processes and NPS pollution in a small watershed in subhumid subtropics using SWAT. *Journal of Hydrologic Engineering*, 17(3), 445–454. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000458](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000458)

- Moriasi, D. N., Wilson, B. N., Douglas-Mankin, K. R., Arnold, J. G., & Gowda, P. H. (2012). Hydrologic and water quality models: Use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1241–1247.
- Moriasi, D. N., Gowda, P. H., Arnold, J. G., Mulla, D. J., Ale, S., & Steiner, J. L. (2013). Modeling the impact of nitrogen fertilizer application and tile drain configuration on nitrate leaching using SWAT. *Agricultural Water Management*, 130, 36–43.
<https://doi.org/10.1016/j.agwat.2013.08.003>
- Nash, J. E., & Sutcliffe, J. V. (1970). River forecasting using conceptual models, 1. A discussion of principles. *Journal of Hydrology*, 10, 280–290.
- Nicótina, L., Alessi Celegon, E., Rinaldo, A., & Marani, M. (2008). On the impact of rainfall patterns on the hydrologic response. *Water Resources Research*, 44(12), W12401.
<https://doi.org/10.1029/2007WR006654>
- Oliver, C. W., Radcliffe, D. E., Risse, L. M., Habteselassie, M., Mukundan, R., Jeong, J., & Hoghooghi, N. (2014). Quantifying the contribution of on-site wastewater treatment systems to stream discharge using the SWAT model. *Journal of Environmental Quality*, 43(2), 539–548. <https://doi.org/10.2134/jeq2013.05.0195>
- Paik, K., & Kumar, P. (2004). Hydraulic geometry and the nonlinearity of the network instantaneous response. *Water Resources Research*, 40, W03602.
<https://doi.org/10.1029/2003WR002821>
- Pandey, A., Chowdary, V. M., Mal, B. C., & Billib, M. (2009). Application of the WEPP model for prioritization and evaluation of best management practices in an Indian watershed. *Hydrological Processes*, 23(21), 2997–3005. <https://doi.org/10.1002/hyp.7411>
- Pieri, L., Bittelli, M., Wu, J. Q., Dun, S., Flanagan, D. C., Pisa, P. R., Ventura, F., & Salvatorelli, F. (2007). Using the Water Erosion Prediction Project (WEPP) model to simulate field-observed runoff and erosion in the Apennines mountain range, Italy. *Journal of Hydrology*, 336(1–2), 84–97. <https://doi.org/10.1016/j.jhydrol.2006.12.014>
- Ping, J., Yan, S., Gu, P., Wu, Z., & Hu, C. (2017). Application of MIKE SHE to study the impact of coal mining on river runoff in Gujiao mining area, Shanxi, China. *PLoS ONE*, 12(12), e0188949. <https://doi.org/10.1371/journal.pone.0188949>
- Renschler, C. S. (2003). Designing geo-spatial interfaces to scale process models: the GeoWEPP approach. *Hydrological Processes*, 17(5), 1005–1017. <https://doi.org/10.1002/hyp.1177>

- Rinaldo, A., Botter, G., Bertuzzo, E., Uccelli, A., Settin, T., & Marani, M. (2005). Transport at basin scales: 1. Theoretical framework. *Hydrol. Earth Sys. Sci. Discuss.*, 2(4), 1613–1640. Available at: <https://hal.archives-ouvertes.fr/file/index/docid/301519/filename/hessd-2-1613-2005.pdf>
- Rinaldo, A., & Marani, A. (1987). Basin scale model of solute transport. *Water Resources Research*, 23(11), 2107–2118.
- Rinaldo, A., Marani, A., & Rigon, R. (1991). Geomorphological dispersion. *Water Resources Research*, 27(4), 513–525.
- Rinaldo, A., & Rodriguez-Iturbe, I. (1996). Geomorphological theory of the hydrological response. *Hydrological Processes*, 10(6), 803–829.
- Rinaldo, A., Vogel, G. K., Rigon, R., & Rodriguez-Iturbe, I. (1995). Can one gauge the shape of a basin? *Water Resources Research*, 31(4), 1119–1127.
- Rodriguez-Iturbe, I., & Rinaldo, A. (2001). *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press. 570 pp.
- Rodríguez-Iturbe, I., & Valdes, J. B. (1979). The geomorphologic structure of hydrologic response. *Water Resources Research*, 15(6), 1409–1420.
- Rouhani, H., Willems, P., Wyseure, G., & Feyen, J. (2007). Parameter estimation in semi-distributed hydrological catchment modelling using a multi-criteria objective function. *Hydrological Processes*, 21(22), 2998–3008. <https://doi.org/10.1002/hyp.6527>
- Rubin, Y. (2003). *Applied Stochastic Hydrology*. Oxford University Press. New York.
- Saco, P. M., & Kumar, P. (2002a). Kinematic dispersion in stream networks 1. Coupling hydraulic and network geometry. *Water Resources Research*, 38(11), 1244. <https://doi.org/10.1029/2001WR000695>
- Saco, P. M., & Kumar, P. (2002b). Kinematic dispersion in stream networks 2. Scale issues and self-similar network organization. *Water Resources Research*, 38(11), 1245. <https://doi.org/10.1029/2001WR000694>
- Sahin, S. (2014). An empirical approach for determining longitudinal dispersion coefficients in rivers. *Environmental Processes*, 1(3), 277–285. <https://doi.org/10.1007/s40710-014-0018-6>
- Seattle. (2015). Forests. Seattle, WA: Seattle Public Utilities. Retrieved from <https://www.seattle.gov/utilities/environment-and-conservation/our-watersheds/cedar-river-watershed/biodiversity/habitats/forests>.

- Shannon, C. (1948), A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423).
- Sherman, L. K. (1932). Streamflow from rainfall by unit-graph method. *Engineering News-Record*, 108, 501–506.
- Singh, R. K., Panda, R. K., Satapathy, K. K., & Ngachan, S. V. (2011). Simulation of runoff and sediment yield from a hilly watershed in the eastern Himalaya, India using the WEPP model. *Journal of Hydrology*, 405(3–4), 261–276.
- Sonnenborg, T. O., Christiansen, J. R., Pang, B., Bruge, A., Stisen, S., & Gundersen, P. (2017). Analyzing the hydrological impact of afforestation and tree species in two catchments with contrasting soil properties using the spatially distributed model MIKE SHE SWET. *Agricultural and Forest Meteorology*, 239, 118–133.
<https://doi.org/10.1016/j.agrformet.2017.03.001>
- Srivastava, A., Brooks, E. S., Dobre, M., Elliot, W. J., Wu, J. Q., Flanagan, D. C., & Link, T. E. (2019). Modeling forest management effects on water and sediment yield from nested, paired watersheds in the interior Pacific Northwest, USA using WEPP. *Science of The Total Environment*, 701, 134877. <https://doi.org/10.1016/j.scitotenv.2019.134877>
- Srivastava, A., Wu, J. Q., Elliot, W. J., Brooks, E. S., & Flanagan, D. C. (2017). Modeling streamflow in a snow-dominated forest watershed using the Water Erosion Prediction Project (WEPP) model. *Transactions of the ASABE*, 60(4), 1171–1187.
- Taylor, G. I. (1922). Diffusion by continuous movements. *Proceedings of the London Mathematical Society*, 2(1), 196–212.
- USDA. (2014). USDA Geospatial Data Gateway. Washington, DC: USDA Natural Resources Conservation Service. Retrieved from <https://datagateway.nrcs.usda.gov/>
- USGS. (2014). *The National Atlas of the United States*. Reston, VA: U.S. Geological Survey. Retrieved from <https://pubs.er.usgs.gov/publication/fs08200>
- van der Tak, L. D., & Bras, R. L. (1990). Incorporating hillslope effects into the geomorphologic instantaneous unit hydrograph. *Water Resources Research*, 26(10), 2393–2400.
- Wang, L., Wu, J. Q., Elliot, W. J., Dun, S., Lapin, S., Fiedler, F. R., & Flanagan, D. C. (2010). Implementation of channel-routing routines in the Water Erosion Prediction Project (WEPP) model. *Proceedings of the 2009 SIAM Conference on “Mathematics for Industry” The Art of “Mathematics for Industry,”* pp. 120–127. Society for Industrial and Applied Mathematics.

- Wang, L., Flanagan, D. C., Wang, Z., & Cherkauer, K. A. (2018). Climate change impacts on nutrient losses of two watersheds in the Great Lakes region. *Water*, 10(4), 442.
- White, A. B., Kumar, P., Saco, P. M., Rhoads, B. L., & Yen, B. C. (2004). Hydrodynamic and geomorphologic dispersion: scale effects in the Illinois River Basin. *Journal of Hydrology*, 288(3–4), 237–257.
- Wolpert, D. H., & Macready, W. (2007). Using self-dissimilarity to quantify complexity. *Complexity*, 12(3), 77–85.
- Wood, E. F., Sivapalan, M., Beven, K., & Band, L. (1988). Effects of spatial variability and scale with implications to hydrologic modeling. *Journal of Hydrology*, 102(1–4), 29–47.
- Yapo, P. O., Gupta, H. V., & Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology*, 181(1–4), 23–48.
- Zhou, F., Xu, Y., Chen, Y., Xu, C.-Y., Gao, Y., & Du, J. (2013). Hydrological response to urbanization at different spatio-temporal scales simulated by coupling of CLUE-S and the SWAT model in the Yangtze River Delta region. *Journal of Hydrology*, 485, 113–125.

5. CONCLUSIONS

5.1 Research overview

The studies in this dissertation dealt with three major factors limiting advancement of the WEPP model (v.2012.8). The resolution of the precipitation input of the stochastic simulation option in WEPP has been upgraded, a new subsurface drainage procedure integrated into the main routines of the WEPP model, and a flow routing method based on general transport adapted to extend research on solute transport.

In the first study, a parsimonious stochastic storm generator based on 5-min time resolution and correlated non-normal Monte Carlo-based numerical simulation was created. The model considered correlated non-normal random rainstorm characteristics such as arrival time, storm duration, and amount of precipitation, as well as the storm intensity distribution. The accuracy of the model was verified by comparing the generated rainfall with rainfall data from a randomly selected 5-min weather station in North Carolina, U.S.

In the second study, the subsurface drainage routines in the WEPP model (v.2012.8) were tested and validated. The current version revealed serious errors in the routines governing subsurface flow estimation, so the model was improved through the modification of subsurface drainage, percolation, water balance, and soil water distribution algorithms, providing, thus, a more realistic estimation of the drainage at field scales. The modified WEPP model was tested and validated on an extensive dataset collected at four experimental sites managed by the USDA-ARS within the western Lake Erie Basin.

In the third study, the theoretical framework for a new hillslope and channel-routing routine for the WEPP model (v.2012.8) was introduced. A routing model (WEPP-CMT) was developed to allow for continuous mass transport modeling for flow and transport of reactive solutes. The model structure used the functionality of WEPP to simulate hillslope responses under diverse land use and management conditions and a Lagrangian description of carrier hydrologic runoff for hillslope and channel states, both represented in the model through travel time distributions derived from catchment geomorphologic properties using a stochastic scheme. Model functionality was tested in a sub-basin of the Upper Cedar River Watershed in the U.S. Pacific Northwest.

Overall, the studies in this dissertation have improved the WEPP model's ability to predict more accurate hydrology, improved implicitly the soil loss estimates by including not only better input of precipitation but also integrating a new subsurface drainage algorithm, and finally increased the potential of the model to be used at larger scales as well as for transport of reactive solutes.

5.2 Major research findings

Major research findings from the study are provided below.

The stochastic storm generator proposed in the first study was capable of reproducing the essential statistical properties of the storm characteristics such as time between storms, storm duration, amount of precipitation, and intensity distribution. The moment estimation efficiency showed that details of the storm features were captured for periods of continuous precipitation longer than 20 years. However, a 100-year simulation had the best representation of the storm variables. The proposed model shows promising potential in studies where high temporal resolution rainfall is required and within storm intensities play essential roles in estimation of other hydrological or environmental variables.

The current version of the WEPP model (v2012.8) did not accurately predict the daily subsurface drainage fluxes for the four experimental sites managed by the USDA-ARS within the western Lake Erie Basin analyzed in this study. The NSE values ranged from -4.72 to -0.20. The modifications in the water percolation, water table generation, and evapotranspiration algorithms of the source code (WEPP model, v2012.9) improved subsurface drainage calculations at the same sites. Calibration and validation of the improved version at three experimental sites showed that WEPP could predict subsurface drainage amounts well with NSE \sim 0.65 and NSE \sim 0.61, respectively. On average 21% of the precipitation at the study sites left the field as subsurface flow. The monthly and annual amounts of subsurface drainage were also well calculated by the modified WEPP model (NSE \sim 0.70). The combined use of the enhanced WEPP model and a soil property estimation model (ROSETTA) showed that an acceptable amount of drainage can be estimated with a limited amount of field data.

The hydrologic response based on catchment geomorphology and travel time distributions (WEPP-CMT) was demonstrated as a suitable stochastic framework for streamflow routing and a potential tool for transport of reactive solutes. For the sub-basin in the Upper Cedar River

Watershed, surface travel distributions had relatively small effects on shaping the hydrologic response function ($JSD = 0.03$), which suggested that velocity in the channels and hillslopes did not differ significantly. Subsurface response functions, on the contrary, were highly dominated by the hillslope travel time ($JSD = 0.3$), given that the mean travel time equaled 12.89 hours for the catchment in the study. The channel travel distributions were highly dependent on the flow velocity ($JSD = 0.27$), whereas hydrodynamic dispersion showed a small effect on the channel travel distributions ($JSD = 0.00053$) but primarily contributed to the overall shape of the response function. The applicability of the WEPP-CTM to predict daily flow discharges for the study site showed an acceptable performance with an overall $NSE = 0.60$.

5.3 Limitations of the current study and future research

A primary limitation found in the stochastic storm generator developed in the first study was the need for at least 10 years of data to reproduce acceptable storm characteristics. However, for most U.S precipitation databases (e.g. ASOS, NCDC, USGS), this limitation should not represent any restriction since most of the stations in these databases contain precipitation periods longer than 10 years. Another limitation was the assumption that storm generation is performed using stationary monthly probability density functions, which may not be entirely accurate in locations where climate change has been shown to have a significant effect on the amount of precipitation and intensities. Future work on this issue can be done by relaxing this assumption and characterizing the trend followed by the data and by adjusting the distribution parameter to this trend. Additional work would include an extensive test with more weather stations measuring precipitation at different time increments and spatial characterization of the parameters defining the distributions for those stations, which may allow the use of synthetic storms in ungauged locations provided relevant information from a regional analysis is available.

Limitations of the study on the implementation of a new subsurface drainage routine in the WEPP model include the fact that the model was only tested under three soil types (i.e. loam, clay loam, silty clay loam), and the additional soil properties required by the new model were generated using a soil property generator (ROSETTA) rather than measured field data, which is often recommended. Future work would include a test of the model under different soil types and slope conditions, model adaptation to controlled water table scenarios, and the extension of the model to WEPP watershed simulations.

A drawback of the routing model (WEPP-CMT) found in its application to the sub-basin of the Upper Cedar River Watershed was that uniform channel flow velocity was assumed for the entire watershed, which seems to be inaccurate for mountainous watersheds where velocity gradients are greater than in relatively flat watersheds. This limitation could be addressed by relaxing the above assumption and characterizing the wave celerity and dispersion coefficient at several points in the channels if discharge information is available. Additionally, the extension of this model to allow for transport of reactive solutes needs to be addressed along with an expanded hillslope component that includes all sorption/desorption processes determining the temporal variability of the solute mass transported by the moving particles interacting between the mobile and immobile phases.

VITA

Josept D. Revuelta-Acosta was born and raised in the City of Coatzacoalcos, Veracruz, Mexico. In August 2008, he graduated from the University of Veracruz in Mexico with a degree in Civil Engineering. In January 2009, he began graduate studies at the University of Queretaro in Mexico in the Structural Engineering Department. His master's work focused on the thermal behavior of passive houses as an alternative for energy saving in central Mexico. He completed his Master's of Science in Structural Engineering in 2011 and was hired by the University of Queretaro as a faculty member. Later, he joined the University of Veracruz in 2014 as a faculty member. He started his Ph.D. studies in the Agricultural and Biological Engineering Department at Purdue University in January 2015. His dissertation research has primarily focused on hydrological modeling and software development, with particular emphasis on the Water Erosion Prediction Project (WEPP) model. He will complete his Doctor of Philosophy degree in May 2020, and return to the University of Veracruz, Mexico.

PUBLICATION

Revuelta-Acosta, J. D., Flanagan D. C., & Engel B. A. (2019). Development of a stochastic storm generator using high-resolution precipitation records. *Applied Engineering in Agriculture*, 35(4), 461-473.