REFINEMENTS ON A GIS-BASED, SPATIALLY DISTRIBUTED RAINFALL-RUNOFF MODEL FOR A SMALL WATERSHED

by

Matthew T. Swensson

BA, Hamilton College, 1997

Submitted to the Graduate Faculty of

the School of Engineering in partial fulfillment

of the requirements for the degree of

Master of Science in Civil Engineering

University of Pittsburgh

2003

UNIVERSITY OF PITTSBURGH

SCHOOL OF ENGINEERING

This thesis was presented

by

Matthew T. Swensson

It was defended on

December 4, 2003

and approved by

Chao-Lin Chiu, Ph.D., Civil Engineering

Tin-Kan Hung, Ph.D., Civil Engineering

Jeen-Shang Lin., Sc.D., Civil Engineering

Thesis Advisor: Rafael G. Quimpo, Ph.D, Civil Engineering

REFINEMENTS ON A GIS-BASED, SPATIALLY DISTRIBUTED RAINFALL-RUNOFF MODEL FOR A SMALL WATERSHED

Matthew T. Swensson

University of Pittsburgh, 2003

This research addresses the problem of predicting the direct runoff generated by precipitation falling over a small watershed using an existing rainfall-runoff model. The model consists of several separate modules, which process data describing the spatial variation of watershed properties, and compute the runoff time series at the watershed outlet. Data processing and visualization is handled primarily by GIS software, while computer programs perform the bulk of the computations. The model is designed to operate a simply and generally as possible, requiring only four external data sets as input (Digital Elevation Model (DEM), land coverage, soil coverage, and incremental precipitation depths) to create all other data needed to compute the direct runoff for the watershed under study. Our goal is to improve the model by refining the existing programs and increasing the level of spatial distribution.

To achieve this goal, we focus our research on the following three objectives: testing and refining the existing model; developing a hydraulic flow routing module using the Kinematic Wave model; and exploring the possibility of calibrating the model results to the actual soil moisture conditions present in the watershed during the time of interest by modifying the SCS curve number values across the watershed. We consider Little Pine Creek watershed located in Allegheny County as our case study, and compare the model's predictions for two separate precipitation events with the direct runoff hydrographs given by a USGS stream gauge located at the outlet of the Little Pine Creek watershed.

Based on the results of our investigation we identify and correct several problems with the existing model, and provide a better understanding of how the model responds under varying conditions and assumptions. We show that the Kinematic Wave flow routing model does not work well with our model because is cannot account for storage within the watershed. We also develop an effective method for calibrating the volume of runoff predicted by the model to that given by the USGS stream gauge by adjusting the SCS curve number values in the watershed to reflect that actual soil moisture conditions present at the time of interest, guaranteeing that fluid mass is conserved between our model predictions and reality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take time here to acknowledge those provided me with assistance and guidance during the course of my research. First and foremost, I would like to thank my advisor Dr. Rafael Quimpo for his advice and guidance over the past year, and for providing me with the opportunity to work on this project. Dr. Quimpo always found the time to discuss my work and welcomed any questions I might have. His interest in my education and genuine passion for teaching was a great source of inspiration during my work. Secondly, I would like to acknowledge Chester Engineers for their generous fellowship that funded my research. I would also like to acknowledge the other three members of my advising committee, Dr. Chao-Lin Chiu, Dr. Tin-Kan Hung, and Dr. Jeen-Shang Lin. Thanks to Dr. Chiu for providing me with a solid background in hydrologic modeling. Thanks to Dr. Hung for advising me on hydraulic routing and programming. Thanks to Dr. Lin for providing me with advice on the thesis process, and a wealth of stimulating conversation.

I would also like to thank those whose work on this project preceded mine, especially, James Emerick and Khalid Khan. Thanks to Jim for getting me started on this research, and taking the time to respond to my deluge of email questions over the summer. Thanks to Khalid for providing me the answers to several key computer programming questions, without which I would have never been able to develop the new computer programs used during this study. I would also like to thank those who provided me with the data used during this study. Thank you to the employees of the Three Rivers Wet Weather Demonstration Program for providing the calibrated radar-rainfall data. Thank you to Raymond Siwicki and Lonnie Fekula of the United States Geological Survey's Pittsburgh office for providing me with data from the stream gauge at Little Pine Creek. Finally, thanks to Natalie Weaver for supporting me during these past years, and my father Richard Swensson, for subtly influencing me to follow the path I have chosen in life.

1.0 INTRODUCTION

Water is nature's primary agent for change on the earth. It has the erosive power to bring down entire mountain ranges, and carve vast canyons. Water is the single most important requirement for life. Without an abundant supply of water for hydration, irrigation, and transportation, human civilization would not exist on this plant today at any level. As it brings life, water can also take life away. Large volumes of flowing water carry tremendous energy. Every year thousands of people worldwide die as the result of flooding, or related phenomena such as mudslides. For these reasons, throughout history, humans have attempted to understand the behavior of water in order to reap the benefits it provides, while minimizing it's potential for destruction. Over time, mankind has managed to achieve this understanding by using predictive models to simulating the behavior of water flow.

Hydrologic models have been developed to simulate a variety of fluid behaviors including ocean waves, storm surge, tsunamis, currents and most importantly runoff. Understanding the rainfall-runoff process in a watershed is critical as it has a greater direct impact on more people and structures than any of the other process combined. From development projects, to culverts to bridges to sewers, successful engineering designs hinge on the availability of accurate information describing the volume of runoff that a particular watershed will generate in response to a certain depth of precipitation falling on the watershed. Rainfall-runoff models provide a critical predictive tool for flood forecasting as well, where they provide estimates not only of the volume of water input to the river, but how the flood wave will propagate down the river. The goal of any watershed rainfall-runoff model is to provide a hydrograph showing the variation of volume flow rate (Q) of direct runoff over time at a particular point of interest, usually taken as the watershed outlet.

The watershed rainfall-runoff model addressed by this research was developed to predict the runoff generated from precipitation falling over a small watershed more accurately and

consistently. The model consists of several separate modules each of which computes a particular set of data required as input by other modules to generate the direct runoff hydrograph for the watershed. All computations in the model are based on solving different variations of the governing equations and other empirical relationships. Over the years since its creation, the model has undergone many changes as each researcher involved in the project incorporates improvements in an effort to improve the accuracy of the model's predictions. This process of continuous testing and refining in an effort to improve accuracy, is arguable the most important concept in modeling. Since it is impossible to replicate nature exactly, there is always room for improvement in our attempts to do so.

Following this concept, in this study we attempt to improve upon the existing model by fulfilling the following objectives. First we seek to test and refine the existing model programs and procedures, focusing in particular on examining the effects of varying certain model parameters and threshold values on the resulting hydrograph at the watershed outlet. Second, we seek to continue the existing trend of moving the model towards an increasingly distributed representation of the watershed by developing a hydraulic flow routing module using the Kinematic Wave Model. Finally, will introduce and examine a new method for calibrating the model's results based on adjusting the soil moisture conditions assumed in the runoff generation process.

2.0 CONCEPTS

2.1 SURFACE WATER MODELING

2.1.1 Overview

All models that simulate fluid behavior are based on three governing equations: conservation of mass, conservation of momentum and conservation of energy. In their complete form, these equations are impossible to solve for a natural system. Thus, in order to reduce the equations to a form that we can manage, all models must involve some level of simplification which is achieved by making certain assumptions about the system's properties and or behavior. By incorporating such assumptions however, we create a model that describes an idealized version of the system's true behavior. Because of this, modelers must be sure to check whether the assumptions they incorporate into a model will yield results that agree with observations of the actual system. This process of evaluating the accuracy of model results by comparing them with actual data is called ground-truthing, and is an essential part of the model development process.

2.1.2 A Systems Approach to Watershed Modeling

Hydrologists often adopt a systems approach when developing a model to replicate a complex natural phenomenon such as the conversion of rainfall to runoff within a watershed. In such an approach, the system is defined to exist within boundaries that make up a control volume, within which conservation principles apply. The system's behavior within this control volume is simulated by a response function that transforms input to output in a way that best replicates the behavior of the actual system being modeled. In simplest form the response function relates input (I) and output (O) to the change in storage within the control volume (ΔS) by the relationship

$$I - O = \Delta S$$

In terms of watershed modeling, the control volume is taken as the watershed's boundary. The input to the system is defined as the volume of excess precipitation generated by rain falling over the watershed. The output exiting the system is defined as the volume of direct runoff that passes through the watershed's outlet. We consider conservation by checking that the volume of excess precipitation generated by the system (input) is equivalent to the volume of direct runoff that passes through the outlet (output). Here it is always a good idea to ensure that conservation holds not only within the system itself, but that the output of the system compares well with some outside source that provides actual data that describing the output produced by the natural process that the system represents. In the case of watershed modeling, the outside data source is almost always a stream gauge located at the watershed's outlet. Stream gauges record the variation of volume flow rate passing through the outlet with time, and can provide a fairly accurate estimate of the volume of direct runoff exiting the watershed during the period of interest.

2.1.3 Modeling the Rainfall-Runoff Process within a Watershed

When precipitation falls on a watershed it can either seep into the ground surface as infiltration, or flow across the surface as excess precipitation. The ratio of infiltration to excess precipitation for a particular area of a watershed depends on the characteristics of the surface upon which the rain falls, specifically the degree of perviousness or imperviousness of that surface. Impervious surfaces, such as asphalt, have a much lower ratio of infiltration to excess precipitation than do pervious surfaces, such as bare sandy soil. Infiltrated water may reach the stream channels of the watershed eventually, but this process takes a long time, and the component of stream flow contributed by groundwater is minor in comparison to the component of excess precipitation that reaches the stream as surface runoff during a precipitation event. Thus, when modeling a watershed's response to precipitation, we ignore the volume lost to infiltration, and focus on the volume of excess precipitation.

The volume of excess precipitation or runoff begins its journey to the watershed outlet by flowing across the land surface toward the closest stream channel in the direction of the steepest elevation gradient. Normally, overland flows terminate in tiny stream channels called rivulets. These rivulets feed into larger furrows and ditches, which in turn feed into successively larger and larger permanent stream channels, eventually ending up at the main channel that drains the entire watershed. Thus, watershed models must consider two components of flow, that which moves across the land surface (overland), and that conveyed by stream channels.

The point at which the main channel exits the watershed is called the outlet. Taken in a larger context, the outlet defines the point at which a smaller watershed flows into the next larger scale of watershed. The smaller watersheds that contribute flow to a larger watershed are referred to as sub-watersheds. The sub-watershed concept can also be applied on the micro scale to divide a small watershed into sub-areas, with the outlet of each sub-area located at the intersection of tributary streams with the main channel, or at the confluence of two tributary streams. The sub-watershed concept provides a convenient way to simulate the flow of runoff through a watershed by computing the runoff generated at the outlet of each sub-watershed individually, then combining these values at the outlet of the main watershed using a process of routing and summing flows.

The variation of runoff volume that reaches the outlet of a sub-watershed at particular increments of time during and following a precipitation event defines a runoff time series. In order to obtain the total runoff at the watershed outlet, the runoff time series computed for each sub-watershed area must first be routed through the reach connecting the outlet of that sub-watershed with that of the next downstream sub-watershed. During the routing process the storage characteristics of reach modify the incoming runoff time series by reducing the peak flow rates, and extending the series duration in a process known as dissipation. Following the routing process the runoff time series for each contributing upstream sub-watershed must be added at the downstream sub-watershed outlet to obtain the cumulative runoff time series at that point. Once this is accomplished, the process is repeated for each subsequent downstream sub-watershed through which the flow passes on its way to the main outlet. In this way, the model simulates the natural process by which flows from tributary channels draining different areas of the watershed combine downstream in a series of confluences leading to the main channel, and eventually the watershed outlet.

2.1.4 Deterministic Modeling Approaches: Lumped vs. Distributed

Deterministic hydrologic models provide a specific output (answer) for a particular input by solving a set of equations derived from either the three governing equations for fluid flow, or some other empirical relationship. There are two basic types of deterministic models, lumped, and distributed. Speaking in terms of watershed modeling, lumped models are those in which the properties of a watershed are assumed to remain constant spatially, where as distributed models consider spatial variability. Both models may consider flow as steady or unsteady. Since they do not account for spatial variability, lumped models are more simplistic than distributed models; but can still simulate the behavior of a watershed accurately, especially if the watershed's properties do not vary much spatially. In general though, distributed models produce much better results because they can account for such variations as the difference in runoff potential between a field and an asphalt parking lot. In simplest terms, the difference between the two models is a matter of resolution.

Our model began as a simple lumped model, but over the past ten years since its inception the increasing availability of spatially distributed data in the form of GIS raster files, coupled with a remarkable increase in computer processing speed have fostered its transition to a distributed representation. This transition occurred as a series of steps resulting from the work of several different investigators, each involving modifications to certain modules within the model to account for spatial variability of watersheds characteristics and precipitation input. As the result of this step-by-step transition, the present day model is almost entirely spatially distributed, but still retains much of the terminology and framework of the early lumped versions.

2.1.5 Lumped Runoff Models – Time-Area Method

The time-area method for calculating runoff at the outlet of a watershed is based on the time-area histogram concept. In this method, the watershed is divided into sub-areas containing cells with equivalent time of concentration (TOC) values, where TOC is simply the time required for water to flow from a particular cell to the watershed outlet. These divisions between cells with equivalent TOC values can be envisioned as contours of equal travel time, or isochrones. The

time area histogram is a plot of the number of cells in the watershed with a particular TOC value. Figure 2.1 (Ponce, 1989) illustrates an example watershed divided by isochrone intervals, and its corresponding histogram.

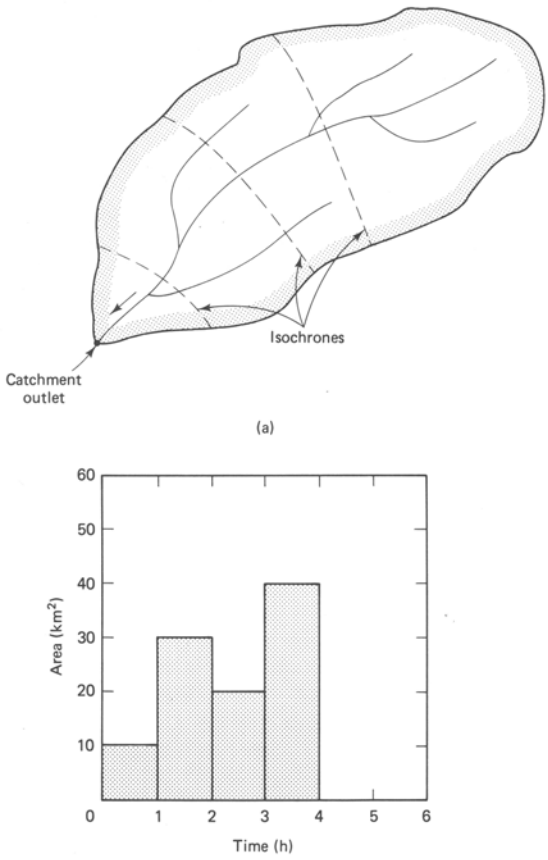


Figure 2.1 - Example of a watershed divided into isochrone increments and corresponding time-area histogram for the watershed (Ponce, 1989)

The incremental runoff at the end of a precipitation interval is computed as the product of the effective precipitation and the area of the cells with a particular TOC value. The total runoff hydrograph is obtained at the watershed outlet by lagging and summing these incremental runoff values appropriately over the duration of the precipitation event. The time-area method has no provision for storage within the watershed, so modelers normally simulate the effects of storage by routing the runoff hydrograph obtained for the watershed outlet through a linear reservoir with storage constant k, using the relationship $S = kQ$.

The original method for calculating runoff in our model involved generating a Clark unit hydrograph by applying the time area method as discussed above for a unit depth of precipitation, then multiplying the ordinates of this unit hydrograph by the actual values of excess precipitation to obtain the direct runoff hydrograph for the watershed. As the transition to a distributed model took place, the runoff module was modified to account for spatial variability of properties such as runoff potential, and precipitation input across the watershed area. What results in the present code is a sort of variation on the time-area method. The code still retains the concept of computing runoff at the watershed outlet based on cell TOC values, but considers the runoff contributed by each cell separately, and no longer employs a unit hydrograph in the computations.

2.1.6 Distributed Runoff Models - SCS Method for Abstractions

The SCS Method for Abstractions was incorporated into our model by Khan, 2001, to account for spatial variability of runoff potential across the watershed. This method uses the SCS Curve number to determine the amount of a given precipitation depth that infiltrates into the ground, and the amount that of excess that remains on the surface for the particular area. SCS Curve numbers describe a surface's potential for generating runoff as a function of the soil type and land use present across the surface. Curve numbers range between $0 < CN \leq 100$, with 0 as the theoretic lower limit describing a surface that absorbs all precipitation, and 100 the upper limit describing an impervious surface such as asphalt, or water, where all precipitation becomes runoff. The method computes the excess precipitation (P_e) generated for an incremental depth of precipitation falling on an area using the following relationship

$$P_e = \frac{(P - 0.2S)^2}{(P + 0.8S)} \qquad (2-1)$$

Here, P is the incremental precipitation depth, and S represents the potential maximum amount of water retained by the soil upon which the precipitation falls. The value S is related to the curve number by

$$S = \frac{1000}{CN} - 10 \qquad (2-2)$$

Equation 2-1 applies only for the condition that $P \geq 0.2S$, otherwise all the precipitation is assumed lost to infiltration.

The SCS method for abstractions considers three conditions of antecedent soil moisture, dry AMC(I), normal AMC(II), and saturated AMC(III). Equations 2-1 and 2-2 apply for normal antecedent moisture conditions (AMC II). To describe dry and saturated conditions the curve numbers must be modified as shown in equations 2-3 and 2-4. The discretization of soil moisture conditions into only three categories represents one of the major weaknesses of the SCS Method for Abstractions as a tool for modeling watershed runoff, as the actual soil moisture conditions present in the watershed at the time of interest are most likely somewhere between the three. We examine the effect that adjusting the curve number to represent soil moisture conditions between those given by the SCS Method has on the resulting volume of runoff predicted by the model in Section 7.3 of Results.

$$CN(I) = \frac{4.2\,CN(II)}{(10 - 0.058\,CN(II))} \qquad (2-3)$$

$$CN(III) = \frac{23\,CN(II)}{(10 + 0.13\,CN(II))} \qquad (2-4)$$

As its final product, the SCS method provides the depth of excess precipitation resulting from a given depth of precipitation falling over an area during a specific time interval. The volume of excess precipitation or runoff is simply the product of the runoff depth and the area.

When properly lagged, these incremental runoff values can be combined at the watershed outlet to produce a direct runoff hydrograph for the entire watershed area.

2.1.7 Lumped Flow Routing Models - Muskingum Method

Lumped flow routing models are also known as hydrologic models. The Muskingum Method is the most commonly used such model. The method involves routing an input runoff hydrograph through an idealized stream channel, assuming that storage is a function of both inflow and outflow by

$$I - O = \frac{dS}{dT} \qquad (2-5)$$

The method considers two components of storage, wedge and prism. Wedge storage can be visualized as the wedge of storage that results from the rise and fall of water level as a flood wave passes through a reach. The advancing wave creates a positive storage, while the receding wave creates a negative. Prism storage is the constant volume in the reach below the wedge (Chow, 1988). The relationship between flow and storage is represented as:

$$S = K[XI + (1-X)O] \qquad (2-6)$$

The constant K represents the travel time of the flood wave through the reach. X is a weighting factor for reach storage. Values for X may range between $0 \leq X \leq 0.5$; but values rarely exceed 0.2 in natural channels. Discretizing 2-5 with respect to x and t, substituting, and rearranging terms yields equation 2-7, which describes the routed outflow from the reach for a given inflow at a particular time increment.

$$O_2 = C_0I_2 + C_1I_1 + C_2O_1 \qquad (2-7)$$

with coefficients C_0, C_1 and C_2

$$C_0 = \frac{\frac{\Delta t}{K} - 2X}{2(1-X) + \frac{\Delta t}{K}} \qquad C_1 = \frac{\frac{\Delta t}{K} + 2X}{2(1-X) + \frac{\Delta t}{K}} \qquad C_2 = \frac{2(1-X) - \frac{\Delta t}{K}}{2(1-X) + \frac{\Delta t}{K}}$$

Obtaining values for the Muskingum K coefficient is relatively simple. X is more difficult since it depends of the storage characteristics of the reach that the model simulates, and is thus different for each.

2.1.8 Distributed Flow Routing Models Using the Saint-Venant Equations

Distributed flow routing models, also known as hydraulic models, describe the passage of a flood wave down a section of reach both in space and time. Such models are based on the solution of the Saint-Venant equations; partial differential equations of momentum and continuity that describe the behavior of one-dimensional unsteady flow in an open channel. In complete form the Saint-Venant equations are

$$\text{Continuity:} \quad \frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \qquad (2-8)$$

$$\text{Momentum:} \quad \frac{\partial Q}{\partial x} + \frac{\partial (\beta Q^2 / A)}{\partial x} + gA\left(\frac{\partial h}{\partial x} + S_f + S_e\right) - \beta q v_x + W_f B = 0 \qquad (2-9)$$

(Chow, 1988)

Simplifying the momentum equation by neglecting various terms produces three one-dimensional distributed flow routing models, the dynamic wave model, the diffusion wave model and the kinematic wave model.

The dynamic wave is the most complete of these models. For this case, only the losses due to the wind shear force component and the drag force from eddies in the flow are neglected in the momentum equation (Chow, 1988). The simplified version of the momentum equation for this model is presented below with each term labeled.

$$\frac{1}{A}\frac{\partial Q}{\partial t} + \frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + g\frac{\partial y}{\partial x} - g(S_0 - S_f) = 0 \quad (2-10)$$

Local Acceleration Term	Convective Acceleration Term	Pressure Force Term	Gravity Force Term	Friction Force Term

(Chow, 1988)

The dynamic wave model can account for acceleration, differences in pressure, and the effects of both gravity and frictional forces on the flow.

The diffusion wave represents the next step down in complexity from the dynamic wave model. The simplified version of the momentum equation used in the diffusion wave model is presented below.

$$g\frac{\partial y}{\partial x} - g(S_0 - S_f) = 0 \quad (2-11)$$

Pressure Force Term	Gravity Force Term	Friction Force Term

(Chow, 1988)

The diffusion wave neglects the effects of acceleration, but can still account for pressure differences as well as gravitation and frictional forces.

The Kinematic Wave model is the simplest of the distributed models. The momentum equation used in this model (2-12) neglects both acceleration and pressure terms, describing the flow behavior in terms of gravitational and frictional forces alone.

$$g(S_0 - S_f) = 0 \quad (2-12)$$

Gravity Force Term	Friction Force Term

(Chow, 1988)

In other words this model assumes that the slope of the energy grade line is parallel to that of the channel bed, $S_0 = S_f$.

2.1.9 The Kinematic Wave Model for Distributed Flow Routing

The literature recognizes two types of flood waves, kinematic and dynamic. Though both types of waves may be present in a particular flood, most are dominated and may thus be described by either one or the other (USACE 1979). Dynamic waves dominate where inertial and pressure forces are important, such as the case of a flood wave propagating down a wide river (USACE 1979). Kinematic waves dominate in situations where the effects of inertial and pressure forces are negligible, which makes sense as kinematics is the study of motion without the influence of external forces (USACE, 1979). Under these conditions, the component of flow weight due to gravity acting in the downstream direction of the channel axis is approximately balanced by that of the frictional force produced by resistance to flow along the channel's wetted perimeter. This balance of forces allows us to neglect acceleration and treat the flow as steady (USACE, 1979).

Since flow is assumed steady, we may approximate Q using Manning's equation

$$Q = \frac{1.49 S_0^{1/2}}{nP^{2/3}} A^{5/3} \qquad (2-13)$$

Solving this equation for Area (A) gives

$$A = \left(\frac{nP^{2/3}}{1.49 S_0^{1/2}} \right)^{3/5} Q^{3/5} \qquad (2-14)$$

This equation may be simplified to obtain

$$A = \alpha Q^{\beta} \qquad (2\text{-}15)$$

where, $\beta = 0.6$, and

$$\alpha = \left(\frac{nP^{2/3}}{1.49 S_0^{1/2}} \right)^{0.6}$$

Substituting this expression for A into the continuity equation (2-8) gives

$$\frac{\partial Q}{\partial x} + \alpha\beta Q^{\beta-1}\frac{\partial Q}{\partial t} = q \qquad (2-16)$$

Equation 2-16 may be solved for Q at any point along the reach length at any time, using either analytical or numerical techniques.

Though the assumptions involved in deriving the kinematic wave equations consider flow to be steady, the propagation of a flood wave down a channel reach is actually a gradually varied phenomenon. It might seem contradictory to model a gradually varied flow process with a steady flow model, but the Kinematic Wave Model is not a completely steady flow model. Though the Kinematic Wave Model considers flow to be steady with regards to momentum, the continuity equation can still account for variations in the flow rate by considering Q incrementally (Chaudry, 1993). Each value of Q in the input runoff time series becomes a separate initial condition, and is routed through the reach independently. The resulting model can be visualized as a succession of steady uniform flow pulses that travel through the reach as kinematic waves.

Ponce (1992) states that the kinematic wave assumptions are usually valid for relatively steeply sloping streams where $S_0 > 0.01$. Ponce (1992) also lists the following as important characteristics of kinematic waves:

- They transport mass
- They do not diffuse appreciably over distance
- They describe a 1-1 relationship between stage and discharge at any cross-section

The second point here is the most interesting, as it implies that the Kinematic Wave Model is one of almost pure translation. The absence of any significant wave diffusion is explained by considering the assumptions made to obtain the version of the momentum equation (2-12) used in the Kinematic Wave Model. Neglecting both the pressure gradient and inertial terms results in the 1-1 stage-discharge relationship mentioned above, which leaves the continuity equation (2-8) to describe the evolution of the flood wave as it passes through the reach. Since equation 2-8 is a first order partial differential equation it cannot describe diffusion, because this is a second order process. Therefore the Kinematic Wave Model has no way of accounting for hydrograph diffusion (Ponce, 1992).

2.1.10 Numerical Methods for Solving the Kinematic Wave Equations

The Kinematic Wave equations can be solved analytically, or numerically. Analytical methods provide the exact solution for any particular case, while numerical methods provide an approximation of the exact solution. Since our model uses computer programs to perform the majority of the computations, we opted for a numerical approach since these are simpler to program. The numerical technique of choice for solving flow routing problems is the method of approximation by finite differences.

In the finite difference method, partial differential equations are transformed into a set of algebraic equations where the derivatives are represented by incremental values of time (Δt) and distance (Δx). The solution proceeds iteratively across a computational grid defined by these distance and time increments, with Q^i_j values computed at each Δx,Δt node. There are two approaches for solving finite-difference equations, the explicit approach, and the implicit approach. The main difference between these two approaches is how the solution proceeds across the computational grid. In the explicit scheme, values for Q are computed for current time line in the computational grid Q^i_j, whereas in the implicit method, Q values are computed for the subsequent (unknown) time line as $Q^{i+1}_{j+1.}$

The explicit method has the advantage of being simpler, but can be computationally unstable (Chow, 1988). This issue of stability is minimized by ensuring that all explicit schemes use Δx and Δt increments which satisfy the Courant condition, listed by Chow (1988) for the kinematic wave equations as

$$\Delta t \leq \frac{\Delta x}{C_k} \qquad (2-17)$$

Here, C_k is the kinematic wave celerity. Physically, the Courant condition states that the wave celerity as defined by the computational grid ($\Delta x/\Delta t$) must be greater than the kinematic wave celerity through the reach. What results if this condition is not met can be envisioned as a piling up of water at the computational nodes along the reach (Chow, 1988). This causes the total volume of water being routed down the channel to increase, which is not desirable.

Implicit methods are slightly more mathematically complex, but this is not really an issue considering the processing speed of modern PCs. Implicit methods offer an additional advantage over explicit methods because they will remain stable for computational steps of any size (Chow, 1988). Despite their inherent stability, it is still advisable to use the Courant condition as a guideline for selecting computational increments when using an implicit scheme to ensure that the results are accurate. We selected an implicit finite difference method in our model because of the advantage in computational stability.

Both explicit and implicit methods can employ linear or non-linear solution schemes. A linear scheme is one in which Q^{i+1}_{j+1} is computed as a linear function of Q values (Chow, 1988). In a non-linear scheme this function is obviously non-linear. Our model employs a linear scheme simply because it is simpler to program, and the situation does not dictate the use of a more complex non-linear model.

2.1.11 Baseflow Separation

Stream flow can be separated into two main components, baseflow and direct runoff. The direct runoff component is relatively short lived and intense, appearing as large spikes in flow volume on the stream hydrograph. Baseflow is the more permanent component of flow in a river or stream derived from groundwater and tributary contributions rather than excess precipitation. This component of flow is usually quite small in comparison with that corresponding to direct runoff, especially in smaller streams. In order to ensure accuracy however, the baseflow component of flow must be separated out from the direct runoff component when studying the rainfall-runoff response of a watershed. Hydrologists employ three basic baseflow separation techniques, the straight line method, the fixed base method and the variable slope method. These are illustrated in Figure 2.2 (Chow, 1988). In the straight line method baseflow is considered to be the volume of the hydrograph below the straight line between the point of rise on the hydrograph and the falling limb. In the fixed base method runoff is defined to end at a time N after the hydrograph peak. The line defining baseflow recession before the onset of direct runoff is extrapolated forward to the hydrograph peak, where another straight line is drawn connecting this point with that intersecting the falling limb at time N. The baseflow is assumed as the volume beneath these two lines. The variable slope method is similar to the fixed base method

except that the second line is defined by interpolating the baseflow recession curve back to the point of inflection on the hydrograph's falling limb; then the two interpolated baseflow recession lines are connected by a third line.

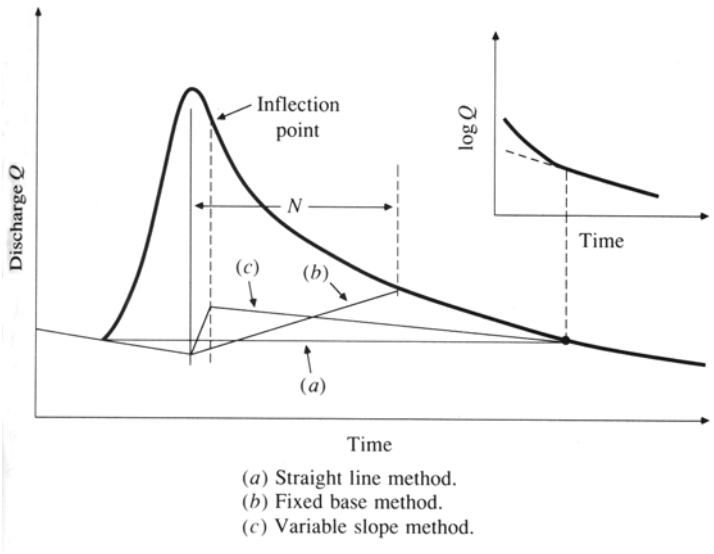


Figure 2.2: Illustration of baseflow separation methods (Chow, 1988)

2.1.12 Soil Moisture Accounting

Soil moisture content has a major effect on the ratio of runoff to infiltration within a watershed, and must therefore be considered when modeling the rainfall-runoff response of a watershed. There are several existing methods to describe the variation of soil moisture conditions within a particular basin. Two example are the Antecedent Moisture Condition (AMC) developed by the Soil Conservation Service, and the more versatile Antecedent Precipitation Index (API) developed by the National Weather Service (Quimpo, 2002). Though they might vary in complexity and number of variables considered, all methods of accounting for soil moisture conditions are based on an empirical equation of the form

$$I_{k+n} = b^n I_k + P_{k+n} \qquad (2\text{-}18)$$

(Quimpo, 2002)

Here I_k is the value of the soil moisture index at time k, P_{k+n} is the volume of precipitation that occurs at time k+n, and b is an empirical constant specific to a particular watershed area. This formula provides a way of continuously updating the soil moisture conditions within the area of interest, which in turn allows for more accurate predictions of the runoff volume that a watershed will generate during a particular precipitation event occurring at a certain time. Though this simple formula and its many more complicated extensions provide a nice way of accounting for soil moisture, its use is limited by the constant b, which must be determined experimentally, and only applies for a relatively small area. Figure 2.3 illustrates the variation of soil moisture conditions in a watershed as described by equation 2-18 (Quimpo, 2002). Here, the sudden increases in soil moisture content correspond to periods of rainfall.

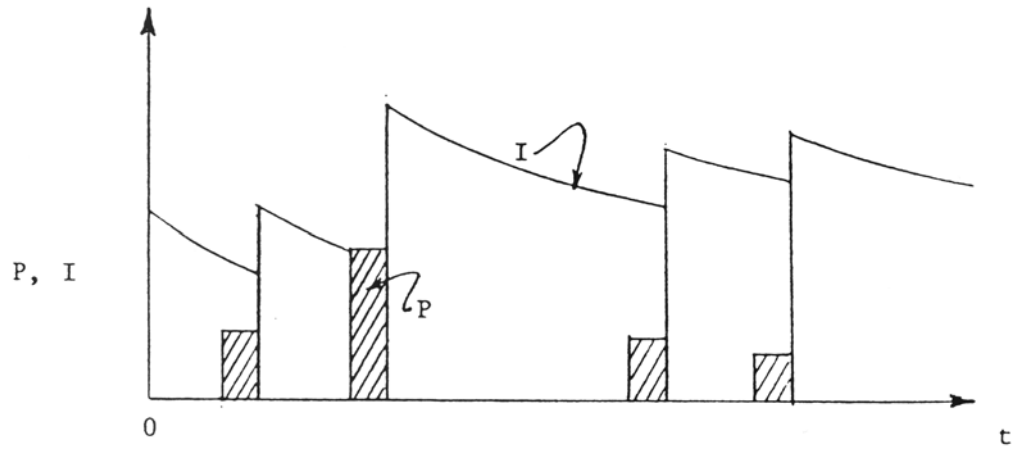


Figure 2.3: Example of soil moisture variation in a watershed by equation 2-18 (Quimpo, 2002)

2.2 PRECIPITATION DATA

The availability of accurate precipitation data to use as input is essential when modeling the relationship between rainfall and runoff in a watershed. Fortunately, the importance of rainfall data in meteorology and other branches of science was recognized early on, and people have

been recording these data for years. Rain gauges are the most common method of collecting precipitation data. The three most commonly used rain gauges are the weighing type, the float type, and the tipping bucket type (Chow, 1988). Though each type of gauge operates differently, all three perform the same basic function, measuring the depth of precipitation falling at a particular point in space. The precipitation estimates provided by rain gauges are useful because they represent actual measured values of rainfall depth; but are limited by the fact that they are only truly representative of the area immediately surrounding the gauge. Since rainfall is a highly spatially variable phenomenon, there is a major disadvantage involved with using rain gauge data to estimate the precipitation depth over an entire watershed, especially when the rain gauges are separated by large distances.

Calibrated radar data provides an attractive alternative to the rain gauge for estimating the depth of precipitation falling over a watershed, because it captures the spatial variation of precipitation at a higher resolution, and accounts for the temporal distribution as well. Emerick (2003) discusses the theory behind estimating precipitation using calibrated radar data in detail. In short, these data are obtained by using a Doppler radar system, which sends out a series of electromagnetic signals from a transmitter. When these signals scatter upon striking an object in the air, some are reflected back to the transmitter. Since the power of a reflected signal is proportional to the diameter of the raindrop it struck (assuming the object that caused the signal to scatter was actually a raindrop), the intensity of reflected signals can provide an estimate of the precipitation intensity. The spatial resolution of this intensity data is usually on the order of 1 km^2 for most Doppler radar systems. Since any object, not just a raindrop, will cause signal scatter, radar data estimates of precipitation intensity are more accurate when calibrated using actual precipitation measurements from near-by rain gauges, hence the term calibrated radar data. Calibrated radar data can be used to estimate the depth of precipitation falling over parcels of land at a resolution equal to that of the radar system.

2.3 GIS-BASED HYDROLOGIC MODELING

2.3.1 GIS – an Overview

GIS stands for Geographic Information Systems. Worded in this way, GIS sounds like a method to accomplish a task, but it is really more of a concept. GIS describes the idea of presenting geographic data spatially. These data, referred to as attributes, can include values of ground surface elevation, land coverage, locations of department stores, population densities, and just about any type of information that varies spatially. Attributes that describe a single data type are called layers. A GIS database usually contains many such layers, each describing the spatial variation of a particular attribute. The key feature of any GIS program is the graphical interface, which allows a user to actually see how a certain type of data varies across a geographic area by displaying attribute layers. In the same way, a user can study the relationship between multiple attributes by displaying several layers at once. Since these systems are very visually oriented, most GIS databases employ a map or other image file as their base layer. GIS programs also allow users to modify existing attribute layers, or create new layers by performing mathematical operations on the data sets, merging one of more attributes together based on common properties, and a host of other options.

All attribute layers in a GIS database are linked together by a common coordinate system, and map projection. Without this link of common geographic reference it would be impossible to consider the relationship between multiple attributes. The Environmental Systems Research Institute (ESRI), creators of the popular software ArcGIS, define a coordinate system as “a reference system used to measure horizontal and vertical distances on a planimetric map” (ESRI, 2003). Coordinate systems consist of a map projection, a reference spheroid and a datum (ESRI, 2003). According to ESRI, a map projection is “a mathematical model that transforms the locations of features on the Earth's surface to locations on a two-dimensional surface. This is required because we wish to represent data from a three dimensional earth (the reference spheroid) on a two dimensional map. The two most common projections are the Universal Transverse Mercator (UTM), common to national data sets, and the State Plane Coordinate System (SPC), used mainly in state local data sets (Emerick, 2003). ESRI defines datum as “a set of parameters and control points used to accurately define the three-dimensional shape of the

Earth" (ESRI, 2003). Most data sets use either the North American Datum of 1927 or 1983, referred to as NAD1927, and NAD1983, respectively. In an ideal world, all data sets would reference the same coordinate system. Since in reality this is far from the case, most GIS programs contain modules for re-projecting data sets to a different coordinate system.

GIS programs use two main types of data layers, raster files, and vector files. A raster files is essentially a large matrix of cells or pixels, with each pixel containing a value of the data that the raster file represents. Raster files normally represent images, or large spatially variable data sets such as the intensity of precipitation shown by a Doppler radar image. The size of each pixel in the raster file is a measure of the file resolution. Files with high resolution have a greater number of smaller pixels than files with low resolution. Vector files consist of a series of lines connected by points called nodes. Areas defined by an enclosure of lines and nodes, are called polygons. GIS programs use vector files to represent linear features, such as rivers, highways, or watershed boundaries.

As simple collections of pixels, lines and nodes, both raster and vector files are completely meaningless in the context of a GIS database without some sort of information describing what they represent, and how they should appear. This descriptive information is contained within a metadata file, which accompanies every raster and vector file. Metadata files contain information about the data projection, datum, maximum and minimum values, scale, data collection procedures, the author of the data set, and pretty much anything else that might be useful to know about the data set at some point in time. GIS programs automatically generate an accompanying metadata file with each new raster or vector file, and will not display a layer without one. Thus, GIS data files should be considered as data/metadata pairs rather than individual raster or vector files.

2.3.2 Use of GIS in Hydrologic Modeling

GIS programs are incredibly useful tools in hydrologic modeling, providing modelers with a quick and easy way to manipulate and display large data sets that describe the variation of such characteristics as soil type and land coverage within a watershed. Presently, there are literally thousands of raster and vector data sets describing all matter of subjects available for download on the web sites of many Federal Government, state and local agencies. Using these published

GIS raster files as input data for a model permits the consideration of spatial variability within a watershed at a level impossible to achieve by any other means within a reasonable time frame. This permits us to move away from the use of modeling approaches that involve lumping of watershed characteristics, in favor of distributed approaches, which are more representative of the natural conditions.

2.4 PREVIOUS RESEARCH INVOLVING THIS MODEL

2.4.1 Mark Michelini

The model was conceived in 1994 as the result of research completed by Mark Michelini (1995), under the direction of Dr. Rafael Quimpo. Michelini developed a model that would generate a unit hydrograph for a watershed automatically by processing a Digital Elevation Model (DEM) file. The motivation behind his research involved reducing the human subjectivity involved in developing unit hydrographs and the resulting direct runoff hydrographs. Michelini examined two method of computing time of concentration values in the watershed. The first used the SCS method to compute the time of concentration for all cells in the watershed. The second considered overland and channel cells separately, using the Ramser equation to compute channel concentration time values, and the Kerby equation for overland cells. Michelini simulated the watershed's response to precipitation with a Clark unit hydrograph, created by routing a time-area histogram through a linear reservoir. This version of the model still required the user to determine a composite curve number for the entire watershed using soil and land coverage data.

2.4.2 Randal Bodnar

Bodnar (1995) expanded upon Michelini's work by testing the model's accuracy using data from an actual gauged watershed, and comparing the model results with data from the stream gauge. He also studied different methods of computing time of concentration values in the watershed comparing the SCS equation, the Kerby equation, the Ramser equation, an equation developed

by the Federal Aviation Administration, and one using the Kinematic Wave. In 1995, Bodnar concluded that the model results compared well with those described by the stream gauge.

2.4.3 Jaber Almedeij

Almedeij (1998) focused on reducing the influence of human subjectivity on the model results further by automating the curve number generation process. Almedeij, created vector files by digitizing land use and soil type data for a watershed of interest, then developed a program to compute a composite curve number for the watershed based on weighted averages of soil and land use values. In 1998, Almedeji concluded that his method of automating the curve number generation process produced excellent results.

2.4.4 Khalid Khan

Khan (2002) began the transition from a lumped to a distributed model. To this end, Khan made the process of curve number generation spatially distributed by developing a program that computed curve number values for each cell within the raster file representing the watershed area, using soil and land coverage raster files as inputs. He implemented the SCS overland flow velocity equations as the method to compute the velocity values required to obtain time of concentration values for overland flow. Khan was also the first to examine the effect of subdividing the watershed on the model's results, and to consider the effects of using spatially distributed temporally incremented precipitation data. In 2002, Khan concluded that distributed curve numbers generated greater runoff volume, higher peak discharge, and shorter time to peak flow values than those generated using lumped curve numbers.

2.4.5 James Emerick

Emerick (2003) continued to move the model from lumped to distributed by incorporating newly available spatially distributed, temporally incremented precipitation data as input. The modification required that Emerick change several of the existing programs used in the model; particularly that used to compute runoff. Emerick tested the accuracy of this new more

distributed version of the model by using it to simulate the response of an actual small, gauged watershed in the area. His work in testing the model included examining its response to subdividing the watershed, and using different time increments to compute time of concentration values. Emerick concluded that the model performed well in simulating the watershed's response to a precipitation event, and that using distributed precipitation data provided an improvement over the original method. He also determined that the model generated the best results using a 5-minute increment to express time of concentration values. Emerick finished his research in 2003.

3.0 RESEARCH OBJECTIVES

The objectives of this research are to test and refine the existing model, to design a new flow routing module using the Kinematic Wave method, and to investigate the idea of calibrating the model results to the runoff volume data given by the USGS stream gauge at Little Pine Creek by adjusting the soil moisture conditions. Some of the objectives that fall into the category of testing and refining the existing model were established at the beginning of the study, but as is usually the case in any research project, some also arose from the need to address questions raised during the study itself.

3.1 TESTING AND REFINING THE EXISTING MODEL

3.1.1 Identifying and Correcting Problems with Existing Codes and Methods

Writing computer programs is art. To code a program so that it behaves exactly as the author intends, under all conditions, efficiently, requires careful planning and much experience. Yet, even the most carefully planned, well-designed programs are bound to have errors somewhere. Convention states that it is up to the programmer to make sure that his or her program is bug free before unleashing it on the world; but this is nearly impossible, as is painfully obvious in many commercially available programs on the market today. Much of the time, it takes an outside source to stumble across some of the more subtle errors in a program which might cause it to behave incorrectly or inconsistently under certain conditions. Though it was not one of the research objectives initially, this editing process soon became an important component of the study as work concerning other objectives uncovered problems with the codes used in the model.

3.1.2 Effect of Watershed Subdivision on Model Results

In the two studies previous to this, both Khan, 2001, and Emerick, 2003, divided the watershed area into several smaller sub-areas (sub-watersheds). In his research, Emerick looked briefly at

the effect that the number and size of subdivisions have on the results by comparing the watershed hydrographs generated for the same precipitation event using two different sub-watershed arrangements (Emerick, 2003). Up to this point however, nobody has examined this effect in detail. In this study we examine how increasing the number of sub-watershed areas used to represent the whole affects the hydrograph at the watershed's outlet predicted by the model. What effect does increasing the number of sub-watersheds have on the peak flow rate (Q_p) and time to peak (T_p) of the hydrograph predicted by the model? What number of subdivisions will produce the best results?

3.1.3 Effect of Varying Muskingum Coefficient X on Model Results

The X coefficient in the Muskingum method acts as a weighting factor on storage within the reach. Values for X are specific to a particular section of reach, and depend on the shape of the wedge used to represent storage within the reach as a flood wave passes through. Thus, selecting a good value for X to accurately represent a reach requires accurate information about the characteristics of that reach, such as channel geometry. As Emerick points out, this is difficult to do with our model because we know practically nothing about the reaches in the watershed other than channel slope (Emerick, 2003). In his research, Emerick found that using a low value of X = 0.05 yielded results that compared best with the hydrograph developed from the USGS stream gauge data at Little Pine Creek for a given precipitation event. Emerick also suggests that the value of X selected has little effect on the resulting hydrograph. In this study we test this assertion by varying the Muskingum X value and comparing the resulting hydrographs predicted at the watershed's outlet.

3.1.4 Examining the Threshold Used to Define Channel Flow

In the model we distinguish between cells where water flow occurs in stream channels, and cells where water flows across the land surface (overland), by computing the number of cells that contribute water to any particular cell in the watershed, then establishing a threshold value for the number of contributing cells required to initiate channelized flow. Previous research considered the area defined by 100 contributing cells as a sufficient threshold value. 100 cells

correspond to an area of approximately 22 acres. By our definitions then, it takes 22 acres of land to generate enough runoff for flow to become channelized. This seems a bit large, considering the abundance of small creeks and rivulets in the wooded areas of Western Pennsylvania. Since our model considers overland and channel flows separately when computing runoff, increasing the number of channel flow cells by decreasing the contributing area required to initiate channelized flow might influence the shape of the hydrograph at the watershed outlet. We examine this idea by lowering the threshold number of contributing cells required for channel flow, and comparing the resulting hydrographs.

3.1.5 Examining the Method Used to Compute Channel Time of Concentration Values

As was the case with correcting errors discovered in other existing codes, this objective arose during other work, when we noticed a strange pattern in the time of concentration (TOC) values for channel cells predicted by the model. Time of concentration is the time required for water flowing from a particular cell in the watershed to reach the outlet. Our model employs a module called the isochrone (equal travel time) code to compute these values. While examining a close up view of a raster file output by this code, we noticed that the TOC values for cells defining a channel reach did not always decrease from upstream to downstream as expected. There were many instances where TOC values would increase for a short distance, before decreasing again. In other cases we encountered isolated cells with lower TOC values than their upstream and downstream neighbors. Whether increasing or decreasing, any such anomaly in the TOC values along a stream channel is undesirable because is does not accurately represent the natural process of stream flow. Out of this observation came the objective of identifying the cause of these anomalous TOC values for cells within the channel reaches, and correcting the problem.

3.2 INCORPORATING KINEMATIC WAVE ROUTING IN THE MODEL

The Kinematic Wave routing method is a completely different approach to flow routing than that used previously. The Muskingum method is a lumped or hydrologic method, considering only the variation of flow with time; but the Kinematic Wave is a distributed or hydraulic approach

that considers the variation of flow with time and space. The reason behind exploring this new flow routing method is to determine how the results from a distributed routing model compare with those generated using the lumped Muskingum method. The trend in the research leading up to this study has been toward making the model more distributed in an attempt to improve the accuracy of the results by considering the characteristics of each cell in the watershed independently. This was accomplished first with the isochrone and runoff computation codes. The next logical step involves switching to a distributed model for flow routing. First, we examine the theory behind kinematic wave routing to gain an understanding of how it might fit into our existing model. Then, we develop a module to perform flow routing using the Kinematic Wave method, and see how the results compare with those generated by the present hydrologic method, and those predicted by the USGS stream gauge at the outlet of the Little Pine Creek watershed.

3.3 MODEL CALIBRATION BY CURVE NUMBER ADJUSTMENT

This objective actually arose at the very end of our study in response to the need to address disagreement observed between the watershed hydrograph predicted by the model and that given by the USGS stream gauge data at Little Pine Creek. Comparing the volume of runoff predicted by the model with that given by the USGS stream gauge provides a measure of how well the model is calibrated to the actual conditions in the watershed at the time of study. Since we use the USGS data as our representation of the actual direct runoff produced by the watershed, any large difference between the volume of excess precipitation predicted by the model and the runoff volume given by the USGS gauge data violates conservation principles, since it implies that the watershed generates more or less runoff than it actually does. During his research Emerick, found that these two volumes were within 5%, for the August 17, 2002 precipitation event, signifying that the model was properly calibrated. In light of these findings, we completed all of our research under the assumption that Emerick's findings held true. However, after implementing the various changes and modification to the programs, we discovered that our results differed substantially from those shown by the USGS hydrograph data for the same precipitation event. Investigating the cause of this discrepancy, we discovered that the volume of

runoff predicted by the model was much greater than that given by the USGS hydrograph data. Thus, the model was no longer calibrated properly. In an attempt to recalibrate the model we investigate adjusting the soil moisture conditions in the watershed as given by the SCS curve number value using a weighted sum of the SCS curve number for AMC(II) conditions and AMC(I) conditions. Since the SCS curve number is an indication of the runoff potential, adjusting it will likely provide a means by which to calibrate our model to various precipitation events in the future.

4.0 METHODS

4.1 OVERVIEW

Our model is the product of ten years of research completed by six different investigators. It is GIS-based in the sense that it employs a combination of GIS and computer methods to accomplish the conversion of rainfall to runoff. On the whole, the model can be classified as a distributed model, as most of the computations consider the characteristics of each individual cell within the raster file that represents the watershed, instead of lumping them all together. Model inputs consist mainly of GIS raster data files, which describe the characteristics of the cells that make up the watershed area, and the spatial variation of precipitation depths that fall over the watershed. The model computes runoff by considering the watershed as a series of sub-areas (sub-watersheds), and computing the runoff time series at the outlet of each on a cell-by-cell basis. The runoff hydrograph at the main watershed outlet is obtained by routing the runoff time series data through the reaches of downstream sub-watersheds using either hydrologic (Muskingum Method), or hydraulic (Kinematic Wave) channel routing methods, and adding the data together at downstream sub-watershed outlet. The following subsections explain how GIS and computer methods are used in the model, and provide a detailed explanation of each step involved in the modeling process in chronological order.

4.1.1 GIS

We use the IDRISI software, developed by Clark Labs, to perform the majority of GIS data analysis. The GIS software is used mainly for processing and manipulating the raster data sets required as inputs by the computer programs, and for visualizing some results. The model uses following ten raster data files as inputs: the elevation DEM file, flow direction file, stream file, watershed file, soil type file, land use file, SCS curve number file, precipitation grid file, incremental precipitation files, and the isochrone file. All data used in the model are developed from the DEM elevation layer, the soil coverage layer, the land coverage layer, or the precipitation layer. These four layers are the only external data required to run the model.

4.1.2 Computer Programs

All computer programs were written in FORTRAN 77, and complied using the FORCE FORTRAN compiler. FORTRAN programs perform the bulk of the model calculations, such as delineating the watershed boundaries, generating the isochrone file and runoff time series, and routing the runoff hydrographs through downstream channel reaches. These programs use the raster files and other data files as input, and create additional data files and or raster files as output. The programs used in this study include:

- DIRECT – used to compute flow directions for each cell
- WATERSHED – defines the watershed and sub-watershed areas
- CURVENUM – computes the SCS curve number value for each cell
- ISOSTYLE_MTS – calculates the time of concentration for each cell
- GRIDPRECIP – generates incremental precipitation files
- RUNOFF_MTS – calculates runoff
- MUSKINGUM_REACH_MTS – routes runoff using Muskingum Method
- K-WAVE_MTS – routes runoff using Kinematic Wave Method
- ADD_HYDRO_MTS – Adds routed and total hydrographs at sub-watershed outlets

All programs that were modified during this study and all new programs are identified by the extension MTS (Matthew T Swensson) to distinguish them from previous versions, and identify the author.

4.2 DEM PROCESSING

The elevation raster file represents the model's base layer. It is created from the Digital Elevation Model (DEM) file for the area of interest, Glenshaw, in the case of this study. The DEM provides a set of surface elevations in meters at 30-meter intervals. Thus the 30 m by 30 m resolution of the individual cells used in our model grid is defined by the DEM. The corrected elevation raster file created from the DEM is the main input used to create the flow direction file, stream file, and watershed and sub-watershed files.

4.2.1 DEM Correction

Once the DEM is imported into IDRISI and converted into a raster file, the file must be cleaned up, so to speak. Why? As with any large data set gathered by remote sensing equipment, DEM files can contain errors, which originate during the data gathering and processing phase. These errors usually show up as abnormally high (spikes) or low (pits) points in the finished DEM data set. In regards to water flow spikes are not a major problem, tending only to divert the flow of water slightly. Pits, on the other hand, are of particular concern, as they can capture large volumes of flow, which adversely affects the model results. For this reason, the elevation raster file generated from the DEM must be corrected to remove any anomalously high or low values using IDRISI's PIT REMOVE feature, which replaces these anomalous values with vlaues interpolated from neighboring cells. An example of a corrected DEM file is presented as Figure 4.1.

4.2.2 Flow Direction File

In order to compute runoff the model must have knowledge of how water will flow through the watershed, specifically, the direction that water entering into, or precipitation landing on any one cell will flow out of that cell. This is accomplished by creating a flow direction file, using the program DIRECT. This program examines each cell in the corrected elevation raster file, and determines the direction of flow for that particular cell based on the elevations of the eight neighboring cells. The program writes one of eight values corresponding to the eight possible directions that flow may exit the cell (one for each neighboring cell) to every cell in the file. The integer values and the flow directions they represent are shown in Figure 4.3. Figure 4.2 provides an example of a flow direction raster file.

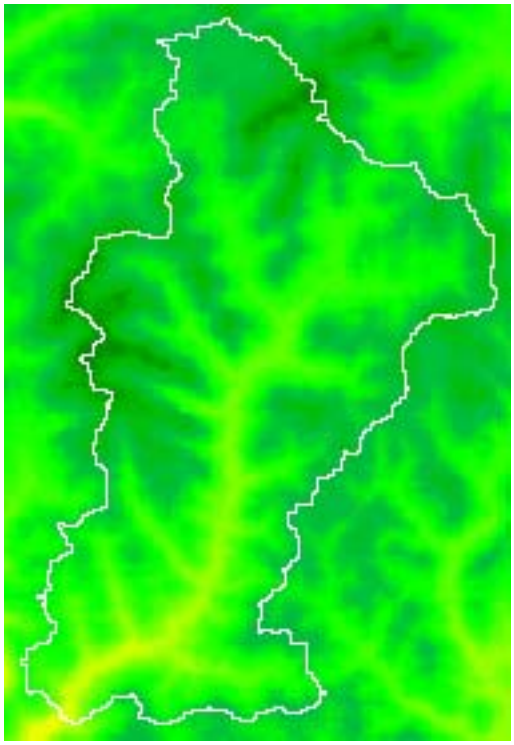


Figure 4.1: Example DEM raster file

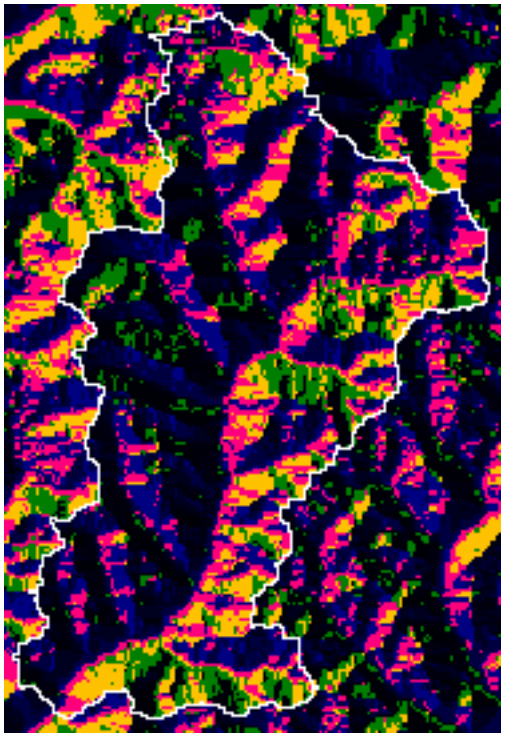


Figure 4.2: Example flow direction raster file

64	**128**	**1**
32		**2**
16	**8**	**4**

Figure 4.3: Example of integers used to represent flow directions and the directions they represent

4.2.3 Stream Network File

Once the flow path of water through the watershed is established, the next question is how it travels along the path. Is the flow traveling across the land surface or in a stream channel? In the model these two methods of conveyance are distinguished, by defining a particular cell as either a channel cell, or overland cell (non-channel), based on the number of other cells in the

watershed that will contribute water to it. This is accomplished by using IDRISI's Runoff function to create a raster file where each cell contains the number of the cells that will eventually pass water to it. Then, an appropriate threshold in terms of number of contributing cells is selected for distinguishing between channelized and overland flow. In this study, as in other previous work, the default threshold value for channelized flow is set at 100 contributing cells (roughly 22 acres of drainage area). To represent this distinction numerically, the existing raster file is updated by assigning all cells with fewer than 100 flow contributing cells a value of 0, and all those with 100 or greater a value of 1. The resulting dendritic stream network is illustrated in Figure 4.4.

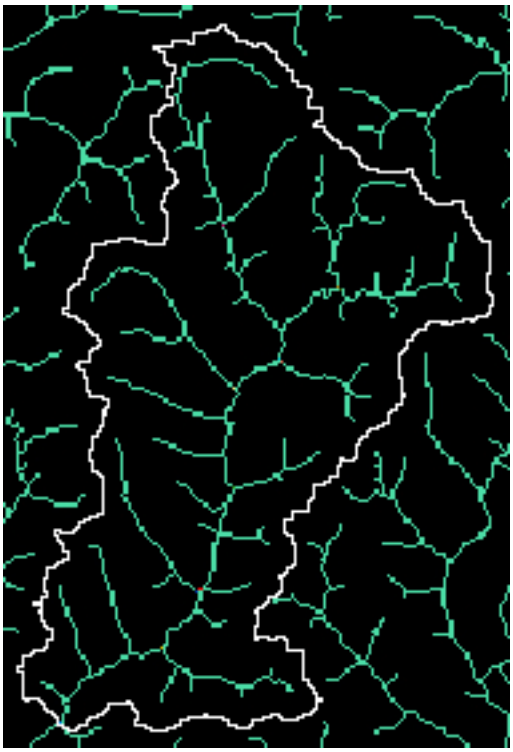


Figure 4.4: Example stream raster file

4.2.4 Watershed and Sub-watershed Delineation

With the flow directions, and flow conveyance methods specified for the entire area defined by the DEM, the area of focus must be narrowed down to only that portion of the region which is important to the study, the watershed. The watershed boundaries are defined by specifying a particular cell to serve as the watershed outlet, which, obviously, must be a channel cell. This is accomplished by assigning the outlet cell in the stream raster file a value of 2 in IDRISI to distinguish it from the other channel flow cells, which have values of 1, then running the program WATERSHED. The WATERSHED program examines the flow direction raster file to determine which cells will eventually flow through the outlet cell. The program outputs a raster file where each of these contributing cells has a value equal to that of the outlet cell in the stream raster file, 2 in this case. Thus the watershed consists only of those cells that contribute flow to the outlet cell. Following previous work on the model, the Little Pine Creek watershed, located in Shaler Township in Allegheny County, was selected for this study. The USGS maintains a stream gage at the outlet of the Little Pine Creek watershed, and the outlet of the modeled watershed was defined at the same location. The Little Pine Creek watershed as defined by the WATERSHED program is visible as the white outline in Figures 4.1, 4.2, and 4.4.

Since most rainfall-runoff models achieve their greatest accuracy for smaller to medium sized watersheds, it is beneficial to divide the main watershed into sub-watersheds to increase the accuracy of the model results. We accomplish this by repeating the process for delineating the watershed described above, this time using several outlets within the existing watershed to generate a raster file containing several sub-watershed areas. The WATERSHED program labels sub-watersheds using the same convention described above. The resulting file with 11 sub-watershed divisions is shown overlaid on the stream file in Figure 4.5. The process of subdividing the watershed is a bit subjective in that it is up to the user to identify proper locations for sub-watershed outlets. However, if these outlet locations are selected sensibly, the sub-watershed arrangement should have little effect on the model results.

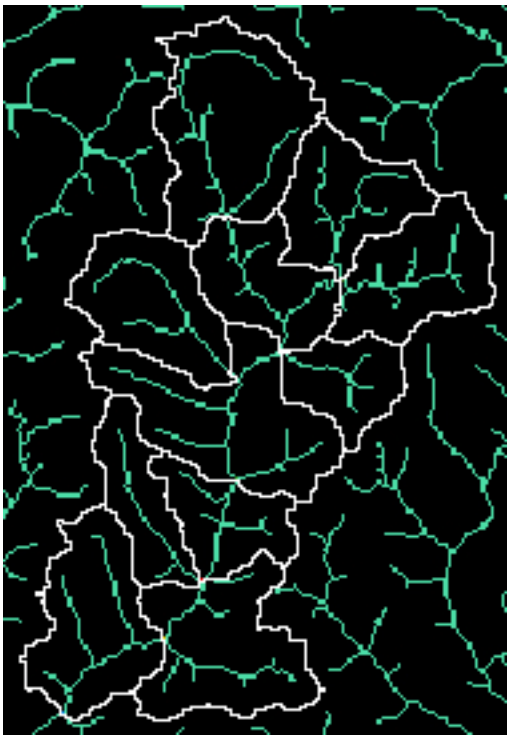


Figure 4.5: Example of stream raster file with 11 sub-watershed divisions

4.3 CURVE NUMBER GENERATION

The next essential step in constructing the model is computing the SCS curve number for every cell in the watershed, as these values are used to calculate both the time it takes for water to reach the watershed outlet and the runoff during later steps. Curve numbers are generated from the soil type and land use raster data layers as described below.

4.3.1 Soil Type File

Soil coverage data in the form of Shapefiles for all counties in Pennsylvania are available for download through the Pennsylvania Spatial Data Access Center's (PASDA) web site. Please refer to Section 5.3 of Data Collection, for more information regarding downloading soil coverage Shapefiles. The Little Pine Creek watershed modeled in this study is located within the Allegheny County soil coverage Shapefiles. The area of interest is clipped from this large file using an outline of the watershed file for Little Pine Creek. The NRCS identifies over 100 soil

groups within Allegheny County. These must all be reclassified into the four hydrologic soil groups (A, B, C and D) identified by the SCS for use in the curve number generation program. Table 4.1 provides the descriptions of these four soil types provided by the Hydraulics Section of the National Engineering Handbook (NRCS, 2001). The resulting raster file of reclassified soil types for the Little Pine Creek watershed area is shown in Figure 4.6. Each cell in this raster file contains an integer value of 1, 2, 3 or 4, which are used to represent soil types A, B, C and D, respectively.

4.3.2 Land Coverage File

The land coverage raster file is created following the same basic steps used to create the soil type file. The land coverage data for Allegheny County used in this study were obtained from the Southwest Pennsylvania Commission. Land uses are divided into 22 categories based on such criteria as vegetative cover type, percent imperviousness, and zoning. Table 4.2 presents a breakdown and description of the land use categories used in the model. In the resulting raster file, Figure 4.7, each cell contains an integer number used to identify its land use in the curve number generation program.

Table 4.1: Description of SCS soil groups A, B, C and D

Hydrologic Soil Group	Description
A	Sandy, loamy sand, or sandy loam. Soils in this group are made up of deep sands or gravels which are well drained to excessively drained with a high infiltrations rate even when fully wetted. Consequently these soils have a low runoff potential.
B	Silt loam or loam. Soil in this group have a moderate infiltration rate. They consist mainly of moderately deep to deep, moderately drained to well drained sands or gravels.
C	Sandy clay loam. This group is made up of soils either having a layer that impedes the downward movement of water or soils consisting of a moderately fine to fine texture. These soils have a slow rate of infiltration.
D	Clay loam, silty clay loam, sandy clay, silty clay, or clay. Soils in this group have a low infiltration rate when fully wetted. This group consists of clays with a high swelling potential, soils existing in areas with a high water table, soils with a clay pan or clay layer near the surface, and shallow soils over nearly imprevious material. The soils in this group have a high runoff potential.

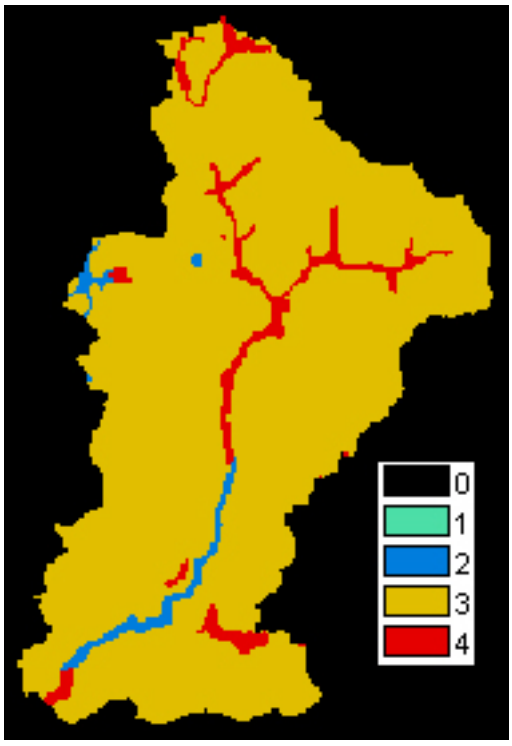


Figure 4.6: Raster file showing soil type distribution across Little Pine Creek watershed

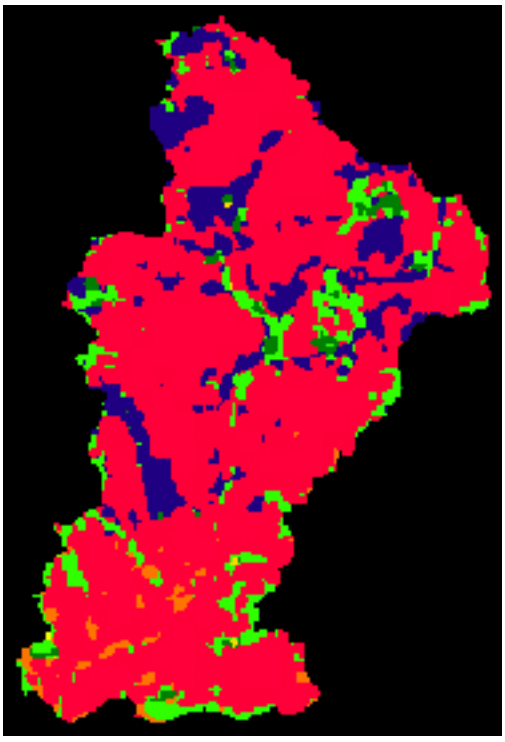


Figure 4.7: Raster file showing land use distribution across Little Pine Creek watershed

Table 4.2: Land coverage descriptions

Type of land use		Land use No. in program
Cultivated Land:	without conservation treatment	1
	with conservation treatment	2
Pasture or range land:	poor condition	3
	good condition	4
Meadow:	good condition	5
Wood or forest land:	thin stand, poor cover, no mulch	6
	good cover	7
Open spaces, lawns, parks	good condition (75% grass cover)	8
golf courses, cemeteries, etc.	fair condition (50%-75% grass cover)	9
Commercial and business areas:	(85% imprevious)	10
Industrial districts:	(72% imprevious)	11
Residential:	1/8 acre or less (65% impervious)	12
	1/4 acre (38% imprevious)	13
	1/3 acre (65% impervious)	14
	1/2 acre (65% impervious)	15
	1 acre (65% impervious)	16
Paved parking lots, roofs, driveways, etc…		17
Streets and roads:	Paved with curbs and storm sewers	18
	Gravel	19
	Dirt	20
Bare soil		21
Water		22

4.3.3 Curve Number File

The curve number raster file is generated using the program CURVENUM. This program reads the soil type and land use values for each cell in the watershed from the respective raster files, and computes an SCS curve number value for the cell using a lookup table. This lookup table of soil type and land use values can be visualized as shown in Table 4.3. The resulting raster file, shown in Figure 4.8, contains SCS curve number values for each cell in the watershed. The curve number raster file is an essential part of computing both the time of concentration values for cells in the watershed, and the runoff time series.

Table 4.3: SCS Curve Number look up table

	Hydrologic Soil Group			
	A = 1	B = 2	C = 3	D = 4
Land Use No.	Curve Number			
1	72	81	88	91
2	62	71	78	81
3	68	79	86	89
4	39	61	74	80
5	30	58	71	78
6	45	66	77	83
7	25	55	70	77
8	39	61	74	80
9	49	69	79	84
10	89	92	94	95
11	81	88	91	93
12	77	85	90	92
13	61	75	83	87
14	57	72	81	86
15	54	70	80	85
16	51	68	79	84
17	98	98	98	98
18	76	85	89	91
19	76	85	89	91
20	72	82	87	89
21	77	86	91	94
22	100	100	100	100

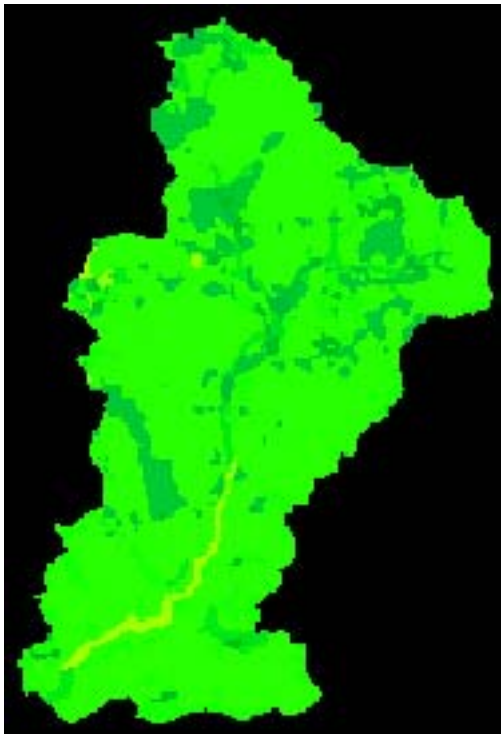


Figure 4.8: Example curve no. raster file for Little Pine Creek

4.4 ISOCHRONE FILE GENERATION

The isochrone raster file contains values of the time required for water to flow from any particular cell located within a watershed or sub-watershed area to the outlet of that watershed or sub-watershed. This travel time is known as the time of concentration (TOC). The term isochrone is a relic from past versions of the model, which used the lumped, time-area method to compute runoff. Though this lumped version of the model is several years removed, the term isochrone has been carried over.

The program computes TOC values using input from the elevation, flow direction, stream, watershed, and curve number raster files. During this study, the existing isochrone generation program, ISOCHRONE, was modified in an attempt correct inconsistent channel cell

TOC values observed within the raster file output. This problem and the steps taken to correct it are discussed in Section 7.1.5 of Results.

The new code, ISOSTYLE_MTS, is similar to the original code. The user must first specify the time increment at which to report the TOC results written to the cells within each sub-watershed For this study, TOC values were reported at 5-minute increments, as this was found to produce good results during previous work (Emerick, 2002). For a case in which the watershed is subdivided, the code repeats the calculations for each sub-watershed area, proceeding from highest to lowest. First, each cell is examined to determine whether it belongs to the sub-watershed in question. If so, the code then uses the direction raster file to trace the flow path from that particular cell to the sub-watershed's outlet, computing overland and channel flow distances, and average slopes for each type of flow along the flow path. The code computes the TOC for overland and channelized portions of the flow path separately; then adds the two values together to obtain the final TOC value for the cell in question. This process continues for all cells within a sub-watershed, and all sub-watersheds with the main watershed.

Overland TOC values are calculated by dividing the total distance of overland flow along a particular flow path, by the flow velocity. The code calculates overland flow velocity using a method obtained from the Hydraulics Section of the SCS National Engineering Handbook (NRCS, 2001), in which the velocity is computed using one of six different equations depending on the curve number value of the cell in question. Table 4.4 lists these equations and the curve number values corresponding to each. For channel TOC values, the code employs the Ramser equation (4-1) to calculate the velocity of channel flow

$$T_c = 0.0195 L_c^{0.77} S_c^{-0.385} \qquad (4\text{-}1)$$

Here, T_c represents TOC, L_c is the length (distance) of channel flow, and S_c is the average slope of the channel cells.

In addition to a raster file of TOC values, the program outputs a data file (data.txt), which contains a summary of several of the more important variables used in the calculations for each sub-watershed area. Table 4.5 presents an example of this summary file for a single sub-watershed. Figure 4.9 presents an example of a raster file for a case with 11 sub-watersheds, with vector representations of the stream network and sub-watershed boundaries overlaid.

Table 4.4: SCS overland flow velocity formulas

Curve Number	Velocity Formula
0 - 50	$y = 0.041e^{0.2951x}$
51 - 60	$y = 0.0813e^{0.2836x}$
61 - 70	$y = 0.1217e^{0.2861x}$
71 - 75	$y = 0.1686e^{0.2882x}$
76 - 85	$y = 0.2533e^{0.2925x}$
86 - 100	$y = 0.3378e^{0.2885x}$

Table 4.5: Example of data.txt summary file

Subwatershed Number 9
Number of Cells in sbws = 1234
Max TOC in sbws = 60.56 min
Longest Flowpath in sbws = 1815.81 meters
Max Channel TOC in sbws = 23.24 min
Length of Longest Channel in sbws = 1531.25 meters
Average Overland Slope for sbws = 0.12043
Average Channel Slope = 0.02593

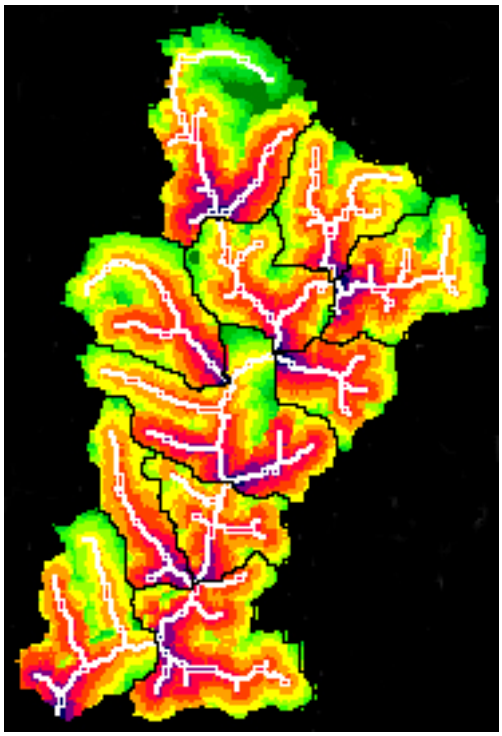


Figure 4.9: Example of isochrone raster file for Little Pine Creek watershed

4.5 PRECIPITATION RASTER FILES

As its final external input, the model requires a set of data describing the spatial and temporal variation of precipitation over the watershed. We use calibrated radar rainfall data at 15-minute time increments available online from the 3 Rivers Wet Weather Demonstration Project's (TRWWDP, 2003) web site to create this input. For more information concerning precipitation data, or the concept of calibrated radar rainfall, please refer to Section 5.5 of Data Collection, or Section 2.2 of Concepts.

The incremental radar rainfall data is only available at a 1-kilometer cell resolution, which is larger than the 30-meter cell resolution of the model. To resolve this resolution conflict, we create a raster file containing only those cells of interest by clipping the grid of calibrated radar rainfall cells with an outline of the watershed. Each 30-meter cell in this file contains an integer value that identifies which of the larger calibrated radar cells it lies within. Figure 4.10 illustrates the precipitation grid, visible as the squares on the series of incremental raster file.

The program GRIDPRECIP accepts an input file containing a series of precipitation depths incremented every 15 minutes over the duration of the precipitation event under consideration for each calibrated radar cell listed in the precipitation grid raster file. The program uses this input file to create a separate raster file of precipitation depths for each time increment listed in the input file. We use 15-minute precipitation increments in this study because this is the finest resolution available for the calibrated radar rainfall data. The model however, will work for any increment of precipitation specified by the user.

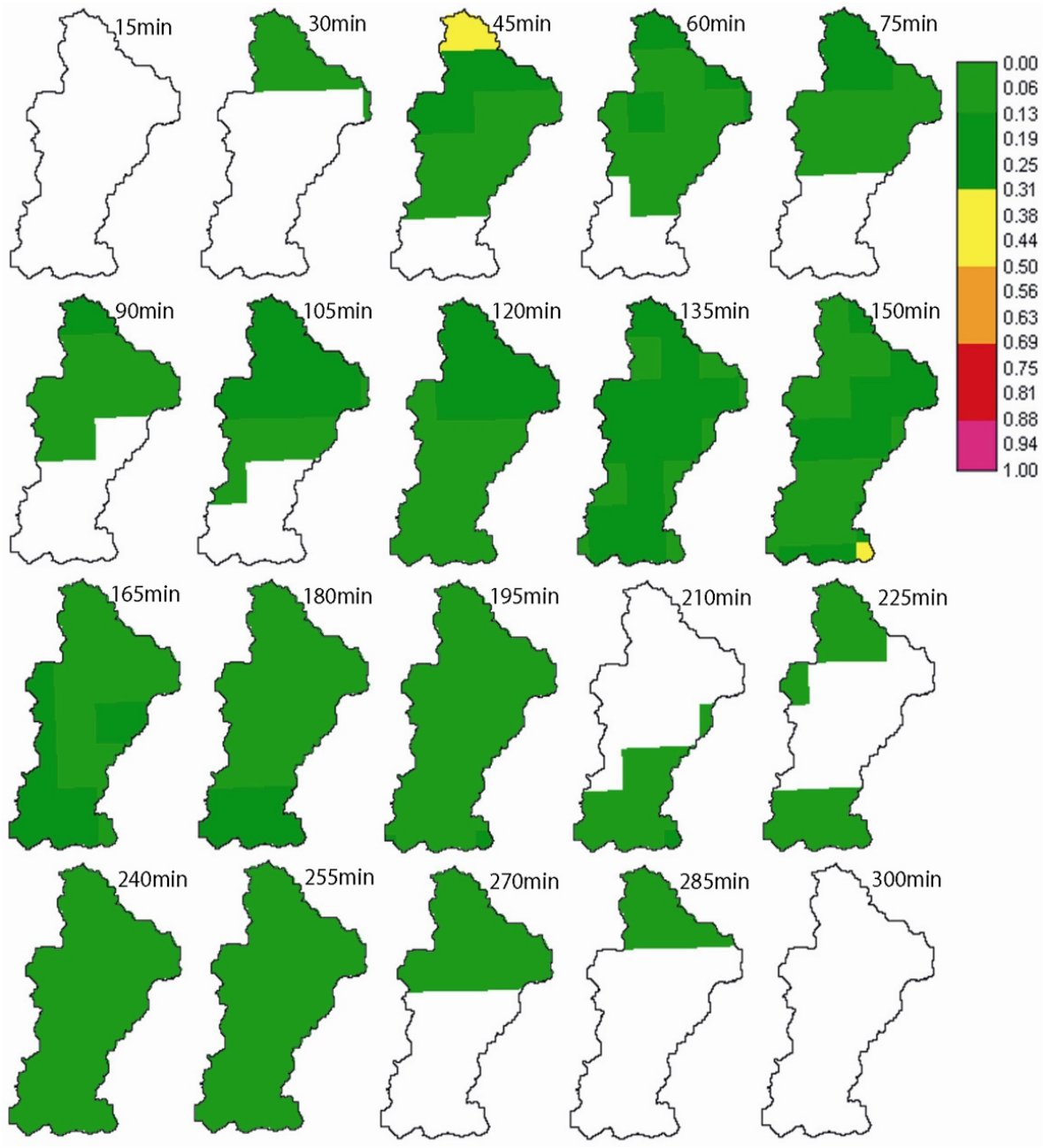


Figure 4.10: Example of incremental precipitation raster files for Little Pine Creek watershed

4.6 GENERATING THE RUNOFF TIME SERIES

Once all of the required raster files are assembled, we can finally simulate the runoff produced within the watershed during a given precipitation event. The objective of this process is to generate a hydrograph of flow rate vs. time at the outlet of the entire watershed area, by computing the runoff time series at the outlet of each sub-watershed, then passing these runoff time series data through successive downstream sub-watersheds to the main watershed outlet by a process of routing and adding flows, as illustrated by Figure 4.11. The process consists of three steps: generating runoff, routing runoff, and adding runoff.

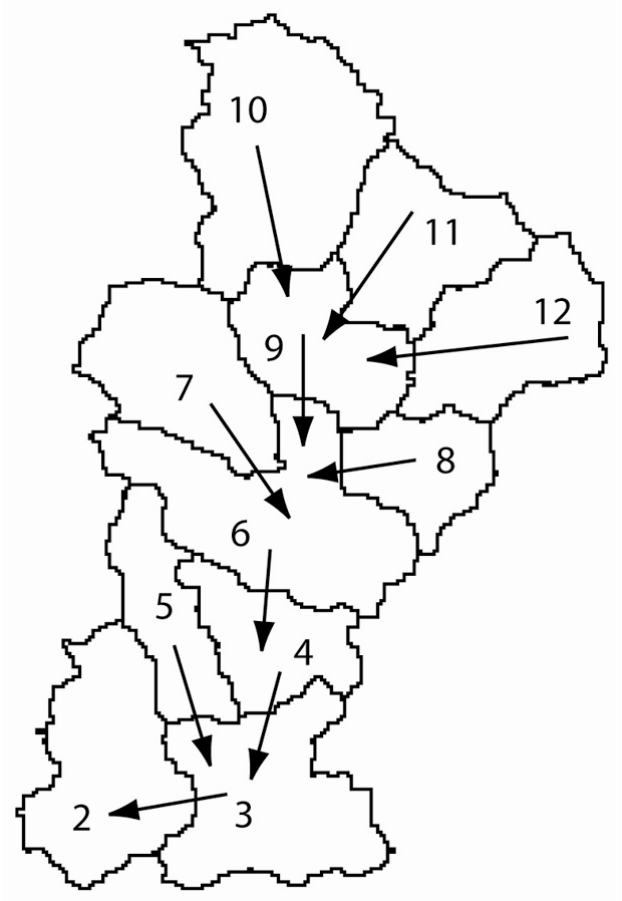


Figure 4.11: Illustration of process hydrographs routing and adding process used in the model

4.6.1 Generating Runoff

The method we use to compute runoff is conceptually similar to the time-area method, but different in that it considers spatial variability within the watershed in two ways. First, the precipitation excess is computed for individuals cell by considering such characteristics as curve number independently for each, rather than using an average value to represent the entire watershed area. Second, the raster files of input precipitation depths vary spatially at a grid spacing that is completely independent of the size and orientation of the sub-watershed areas. This means that different cells within the same sub-watershed area may be assigned different precipitation depth values during a particular precipitation increment. The original code used to compute runoff in the model, RUNOFF, was developed during the previous study (Emerick, 2002). During the course of this research, this original code was replaced by a modified version, RUNOFF_MTS, of similar design, that produced more consistently accurate results. Please refer to Section 7.1.1.2 of Results for an account of these modifications.

RUNOFF_MTS uses data from the corrected elevation, flow direction, stream, sub-watershed, curve number, precipitation, and isochrone raster files, to compute runoff for each sub-watershed area, moving from highest to lowest. For each increment of precipitation spanning the entire precipitation event, the code proceeds cell-by-cell through the watershed, looking for cells located within the sub-watershed in question. For each such cell, the code computes the depth of precipitation excess resulting from the total depth of precipitation falling on that cell during the given time increment, using the SCS method for abstractions. The depth of excess precipitation generated by any cell is converted into a volume flow rate (runoff); then added to a time series array that contains a running total of runoff generated by all cells during the precipitation increment in question. This process is repeated for all cells within the sub-watershed, and for all precipitation increments. At the end, the program adds the runoff time series arrays for each precipitation increment, into a single lagged runoff time series file. What results are as many text files containing runoff time series data as sub-watersheds in the watershed area.

The code offers users the option of computing the runoff time series at the same time increment used to calculate time of concentration values in the isochrone program, or at the precipitation time increment. This choice was added to the original program to accommodate the

kinematic wave routing code developed during this study, which requires runoff time series data at the isochrone increments. The Muskingum flow routing program accepts runoff time series data at time increments equal to that of the precipitation data.

The code also provides the user with a choice of three Antecedent Moisture Conditions (AMC) to use when performing the calculations. AMC I for dry soil moisture conditions, AMC II for normal conditions, or AMC III for wet conditions. During this study, all runoff is computed assuming AMC II, normal soil moisture conditions unless otherwise noted.

4.6.2 Routing the Runoff Time Series

4.6.2.1 Hydrologic Routing with the Muskingum Method

The concept of hydrologic routing using the Muskingum Method was explored during the previous work on this model with the development of the code MUSKINGUM_REACH. During the present study this code was modified slightly to permit routing of runoff time series data from as many as 25 sub-watersheds. Other than this slight modification the modified code, MUSKINGUM_REACH_MTS, is identical to the original. The code takes an input of runoff time series data as a text file for a specified sub-watershed, and routes it through the sub-watershed located immediately downstream. During this process the input file is modified slightly to account for the effects of storage with the reach according to the Muskingum K and X constants specified by the user. When finished, the code outputs a text file containing the routed runoff time series for the upstream sub-watershed.

The value for constant K, the time of travel of flow through the reach under consideration, is assumed equivalent to the time of concentration value of the first channel cell downstream from the outlet of the sub-watershed in question. Selecting the wedge storage shape-weighting factor X is a slightly more subjective process, as described in Section 2.1.7 of Concepts. For the purposes of this study, the X parameter is assumed to be 0.05 for all cases unless otherwise noted, as this was the value selected during the previous study (Emerick, 2002).

4.6.2.2 Hydraulic Routing with the Kinematic Wave Method

The kinematic wave routing process is based on solving a simplified version of the Saint Venant equations, which assume that flow is uniform along a channel reach. Under these assumptions, both the acceleration and pressure terms of the momentum equation are reduced to zero, and the

solution reduces to solving the continuity equation, under the condition that the energy slope (S_f) is equivalent to the channel bed slope (S_0). The theory behind the Kinematic Wave equations and their use in hydraulic flow routing is detailed in Sections 2.1.8 and 2.1.9 of Concepts.

The code developed to study kinematic wave routing, K-wave_MTS, solves equation 4-2, developed by combining the kinematic wave versions of the momentum and continuity equations, numerically using a linear, implicit finite difference scheme to find the unknown value of flow $Q^{j+1}{}_{i+1}$. Please refer to Section 2.1.9 of Concepts, for more details on the finite difference method used to solve the Kinematic Wave equation.

$$\frac{\partial Q}{\partial x} + \alpha\beta Q^{\beta-1}\frac{\partial Q}{\partial t} = q \qquad (4-2)$$

Approximating the space and time derivatives of Q with a backward difference

$$\frac{\partial Q}{\partial t} = \frac{Q_{i+1}{}^{j+1} - Q_{i+1}{}^{j}}{\Delta t} \qquad (4-3)$$

$$\frac{\partial Q}{\partial x} = \frac{Q_{i+1}{}^{j+1} - Q_{i}{}^{j+1}}{\Delta x} \qquad (4-4)$$

and the value Q in the term $\alpha\beta Q^{\beta-1}$ as the average

$$Q = \frac{Q_{i+1}{}^{j} + Q_{i}{}^{j+1}}{2} \qquad (4-5)$$

gives the following equation used by the code to compute $Q^{j+1}{}_{i+1}$, with the assumption that lateral inflow q = 0.

$$Q_{i+1}{}^{j+1} = \frac{\frac{\Delta t}{\Delta x} Q_{i}{}^{j+1} + \alpha\beta Q_{i+1}{}^{j}\left(\frac{Q_{i+1}{}^{j} + Q_{i}{}^{j+1}}{2}\right)^{\beta-1}}{\frac{\Delta t}{\Delta x} + \alpha\beta\left(\frac{Q_{i+1}{}^{j} + Q_{i}{}^{j+1}}{2}\right)^{\beta-1}} \qquad (4-6)$$

In equation 4-6, $\beta = 0.6$, and

$$\alpha = \left(\frac{nP^{2/3}}{1.49 S_0^{1/2}} \right)^{0.6}$$

To solve for α then, we must calculate the channel's wetted perimeter, and estimate a value for Manning's n. This presents a problem, because neither the model results nor the data available for the watershed can describe channel geometry or roughness. It would be possible to collect this type of data directly from the watershed, but the amount of time involved in performing such a field survey make this option impractical. A solution to this problem is achieved by making the following assumptions. All channel reaches in the model are represented using a 45-90-45 triangle, or half-square cross-section. Under this assumption the wetted perimeter of the channel cross section reduces to $2D\sqrt{2}$. The value for the hydraulic mean depth (D) is approximated using the continuity equation $Q = VA$, where $A = D^2$ for the given cross section, and $D = \sqrt{(Q/V)}$. The velocity term in this last expression is estimated using the ratio of the distance to the time step used in the finite difference calculations, $V = (\Delta x/\Delta t)$. Assuming a value for Manning's n is a difficult and subjective process for any channel, especially a natural channel. Several sources exist that provide guidelines for selecting Manning's n values in natural stream channels (USDOT, 1984, and SCS, 1963), but these require careful documentation of channel characteristics in the field. Thus, due to time constraints, and for simplicity of programming, the average value for Manning's in natural rivers and streams suggested by Chow (1959), n = 0.045, was selected to represent the roughness for all channels in the watershed.

The code K-Wave_MTS allows the user to route the runoff time series data for as many as four sub-watersheds though a single downstream sub-watershed at once. For each runoff series, the code first identifies the section of reach in the downstream sub-watershed that it will travel through; then computes an average slope for the reach. Next the space and time steps Δx and Δt used in the finite difference computations are defined in one of two ways. Here, the code provides the user the option of either selecting the Δx and Δt manually, or selecting Δx and allowing the computer to select the Δt required to satisfy the Courant condition. To find the Δt required to satisfy the Courant Condition the program estimates the velocity of a kinematic wave (C_k) though a particular reach using the equation

$$C_k = \frac{0.9935}{n} S_0^{1/2} D^{2/3} \qquad (4-7)$$

In equation 4-7, D is approximated using the relationship

$$D = \sqrt{(Q / V)} = \sqrt{[(Q)\ (\text{Reach travel time} / \text{Reach length})]} \qquad (4\text{-}8)$$

where the reach travel time is taken as TOC value of the first cell in the reach downstream from the upstream sub-watershed's outlet.

With the computational grid established, the runoff time series to be routed passes through a piece-wise linear interpolation routine to generate flow data at the time step increments required by the code. Then, with all the input correctly formatted, the code computes the routed runoff time series at the outlet of the downstream sub-watershed by solving the finite difference representation of the Kinematic Wave equations. The code solves for the flow at each node Q^{j+1}_{i+1} for all Δx increments within each Δt interval using known boundary and initial conditions at $t = 0$ and $x = 0$. The result is a solution matrix of Δt rows and Δx columns, where the last column represents the values of Q at the outlet of the downstream sub-watershed, or the routed runoff. As the final step, the routed runoff values are re-interpolated back to the original time increment of the input runoff time series.

When finished, the code outputs a text file containing the routed runoff time series data as it appears at the outlet of the downstream sub-watershed for each runoff time series input. During this study, all TOC increments were set at 5 minutes, thus all routed runoff values were incremented at 5 minutes as well.

4.6.3 Adding Routed Hydrographs at Sub-watershed Outlets

Once the runoff time series data flowing into a particular sub-watershed are routed through that sub-watershed, we must add them to the runoff total for the sub-watershed to obtain the total runoff time series at the outlet before continuing the routing process downstream. This simple linear process is accomplished using the code ADD_HYDRO_MTS. This code is an updated version of a code developed by Emerick (2003). The new code is essentially the same as the original, but incorporates the following minor changes

- The code will now add up to 4 routed hydrographs to the sub-watershed's own runoff time series
- The code now recognizes up to 25 sub-watershed divisions within a watershed area

Both of these modifications were completed to accommodate an investigation of the effect that subdividing the watershed has on the resulting runoff hydrograph for the entire watershed area.

5.0 DATA COLLECTION

5.1 OVERVIEW

Our model requires four external data sets, the Digital Elevation Model (DEM), soil use and land coverage files, and incremental precipitation data. With the exception of the land coverage data file, the other three are available online, which greatly simplifies matters. The process of assembling all required external data sets, which used to take days to complete using digitization tables, we can now accomplish in an hour. The only other external data considered in this study are those from the stream gauge at Little Pine Creek, against which we compare our results.

5.2 DEM DATA

A DEM is basically a large set of land surface elevation data for a particular coverage area. DEM files for the entire state of Pennsylvania are available from the Pennsylvania Spatial Data Access center's (PASDA) web site, which is a part of the Pennsylvania Geospatial Data Clearinghouse.

www.pasda.psu.edu

DEM files have a cell resolution of 30 meters by 30 meters, or 0.22 acres. The DEM files used in this study are referenced to the NAD 1927 datum and projected using UTM Zone 17.

5.3 SOIL DATA

Data sets describing the variation of soil types for the each county in Pennsylvania are available in the form of Shapefiles online from the PASDA web site. These Shapefiles were created in 1999 by digitizing data from the most recent soil survey completed in 1981. In our research we use the Allegheny County shape file. The coordinate system for this file is keyed to the NAD 1927 datum, and projected in UTM Zone 17 N. Figure 5.1 illustrates the soil file for Allegheny County.

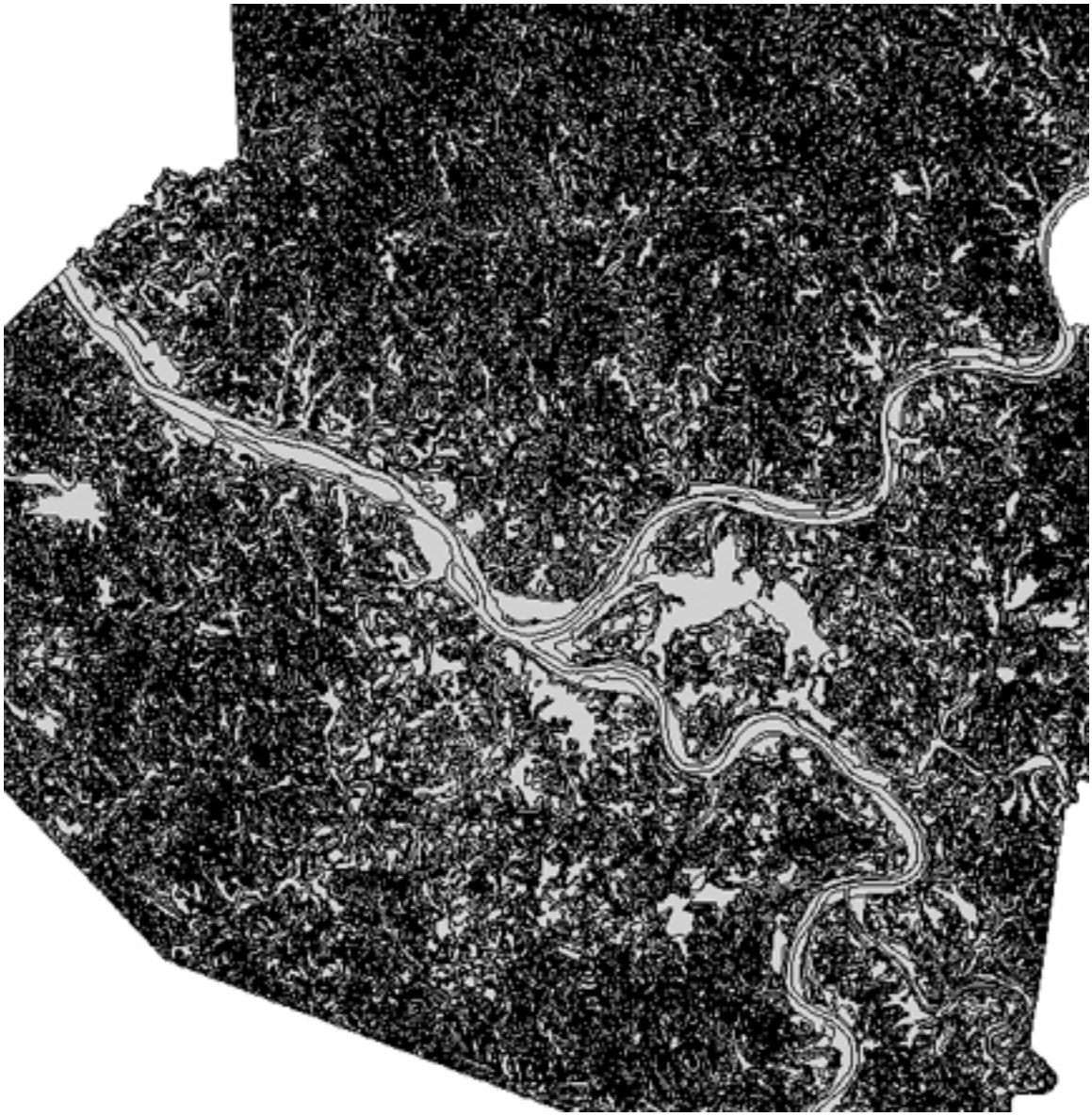


Figure 5.1: View of soil data file for Allegheny County

5.4 LAND COVERAGE DATA

The land coverage data used in this study is a digitized version of a land cover survey completed by the Southwestern Pennsylvania Commission using Landsat imagery in 1992. The portion of this data set covering Allegheny County consists of three Shapefiles, Allegheny County North, South and East. The Little Pine Creek watershed is located within the northern Shapefile. Emerick (2002) obtained these data sets from the government documents section of the University of Pittsburgh's Hillman Library. Though other land coverage data sets are available for our area of interest online at the PASDA web site, Emerick reasoned that a data set generated from a local survey is likely to be more accurate than one completed at the national scale. The land coverage Shapefile is referenced to the NAD 1983 datum, and projected in Pennsylvania State Plane South. Thus, it must be re-projected to UTM Zone 17 N before use in the model. Figure 5.2 illustrates the land coverage data for Allegheny County.

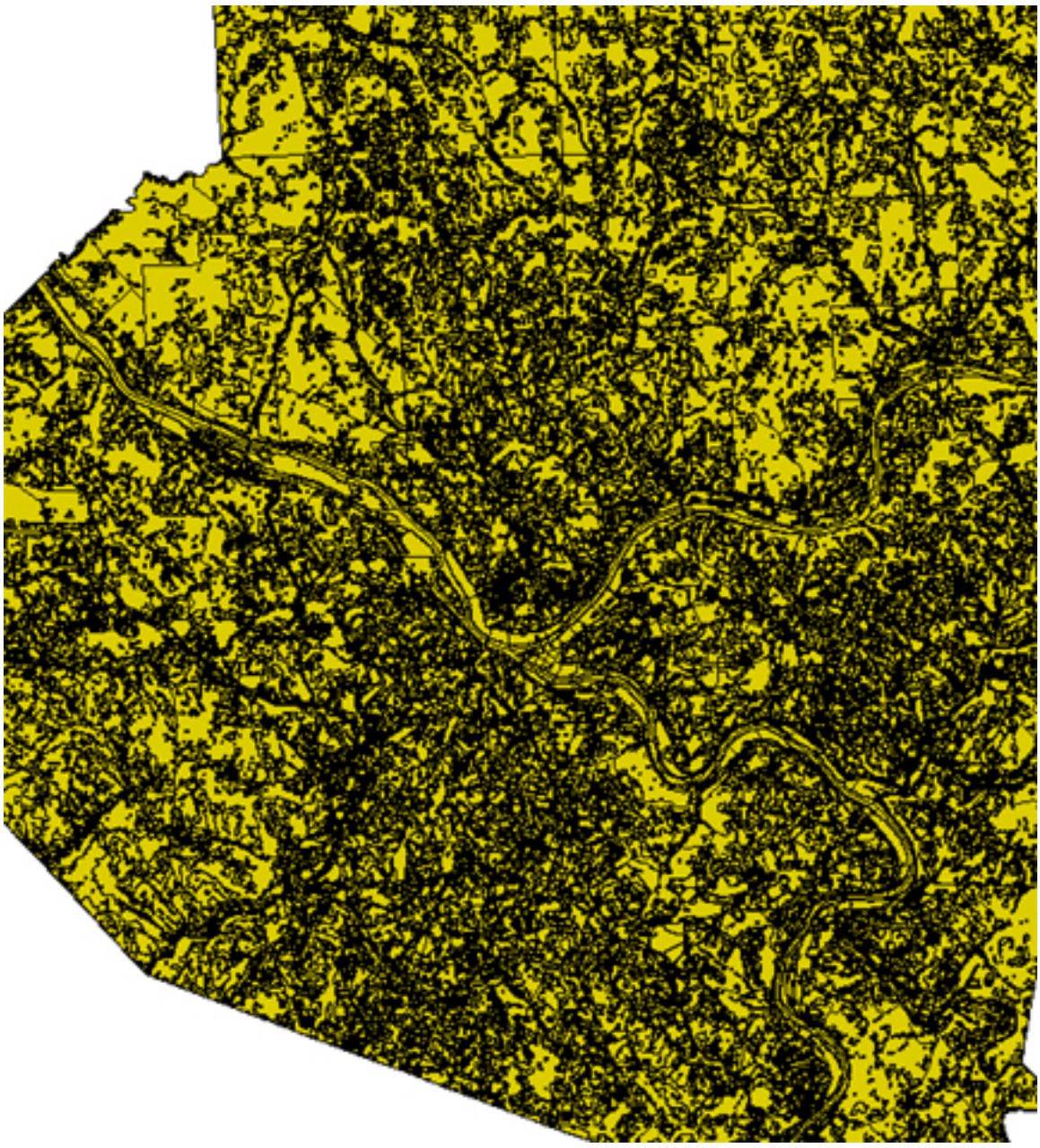


Figure 5.2: Land coverage data for Allegheny County

5.5 PRECIPITATION DATA

The calibrated radar-rainfall data used to create our precipitation input comes from the Three Rivers Wet Weather Demonstration Project's (TRWWDP) web site.

http://3riverswetweather.org

TRWWDP provides access to two forms of precipitation data on its web site; depth values measured by a network of rain gauges located across the county, and those created from a Doppler radar system covering Allegheny County at a grid resolution of 1 km^2. We are interested in the later. TRWWDP makes this data available for viewing and download free of charge, but does require that first time users create a login ID and password to access the database. The calibrated radar rainfall data is available by selecting a particular pixel of interest from a figure depicting the entire coverage grid, or for all pixels of interest, by entering a formula containing the row and column numbers of the desired pixels. The formula for the 28 cells that fall within the boundaries of the Little Pine Creek watershed is

151,129;151,130;152,125;152,126;152,127;152,128;152,129;152,130;153,123;153,124;153,125;153,126;153,127;153,128;153,129;153,130;154,124;154,125;154,126;154,127;154,128;154,129;154,130;155,125;155,126;155,127;156,125;156,126

To obtain these data, the user must specify the desired time duration and the time increment at which to present the data. The system offers the option of daily, hourly, and 15-minute time increments. In this study we use 15-minute increments to obtain the best temporal distribution possible. There are two items worth mentioning here. First, calibrated radar-rainfall data is not available for every single precipitation event. The system only operates during events that are considered significant rain-makers. The company that provides TRWWDP with this data, Vieux and Associates, Inc., publishes monthly reports of the precipitation events recorded by the system, which are also available on the TRWWDP web site. Second, calibrated radar-rainfall data is not available online for all pixels within the coverage grid for all precipitation events. Until July 1, 2003, the coverage grid did not include all cells shown on the map. Unfortunately for us, several of these cells excluded from the original grid lie within the northern reaches of the

Little Pine Creek watershed. Figure 5.3 illustrates the radar coverage grid for Allegheny County, and the locations of the rain gauges used to calibrate the data.

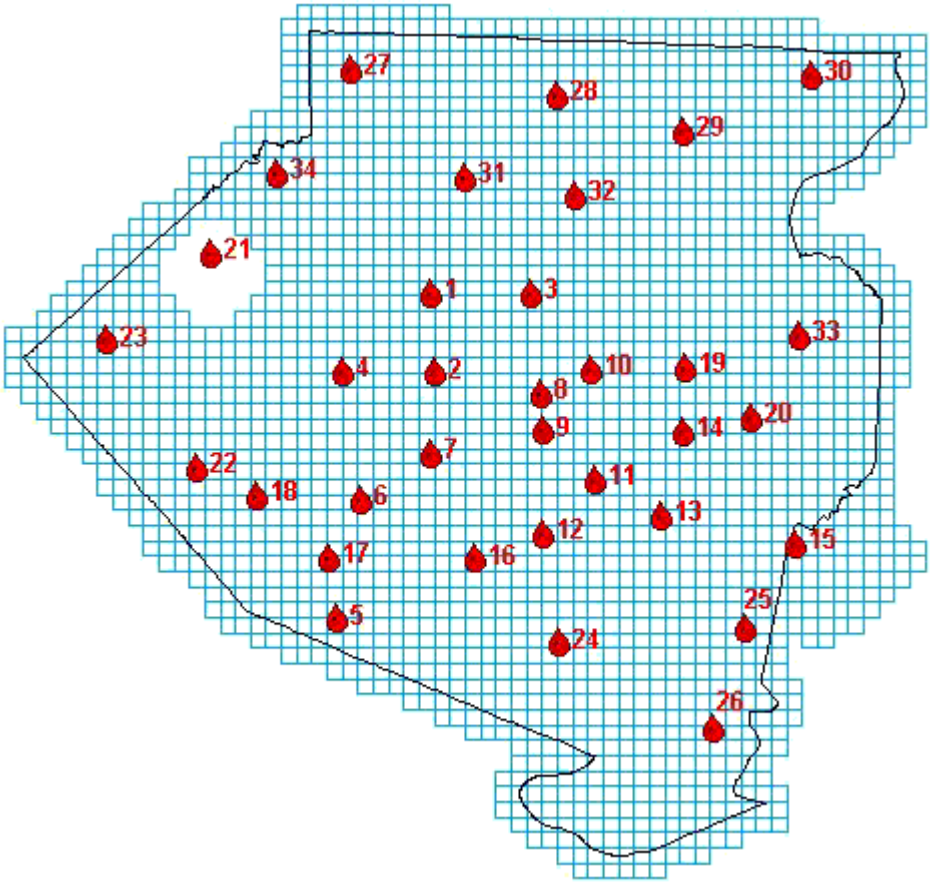


Figure 5.3: Calibrated radar data coverage grid for and locations of stream gauges across Allegheny County

5.6 STREAM GAUGE DATA

The United States Geological Survey (USGS) maintains a stream gauge (ID# 05010009) at the outlet of the Little Pine Creek watershed. The gauge sits at 40°31'13" N Latitude, 79°56'18" E Longitude, at 774.26 feet above mean sea level. Please refer to Figure 6.9 and 6.10 for photos of this stream gauge. Information about this stream gauge and actual flow data are available online through the Water Resources Division of the USGS

http://water.usgs.gov

The gauge at Little Pine Creek is a real-time gauge, which means that it shows the stage and discharge data reported by the gauge online as these measurements come in. While this is interesting, data reported by real-time gauges are, unfortunately, only available online for the past 31 days. To obtain gauge data more than 31 days old one must contact the USGS Pittsburgh office and request it. The contacts at the USGS who provided the data used in this study were Ray Siwicki and Lonnie Fekula.

6.0 CASE STUDY

6.1 LITTLE PINE CREEK WATERSHED

During his research, Emerick chose to model the Little Pine Creek watershed because it was small, located within the gridded radar rainfall coverage area, has a UGSG stream gauge at the outlet, and contains no significant detention ponds or other storage. The watershed covers an area of approximately 5.8 mi^2 located mostly within Shaler and O'Hara Townships in north-central Allegheny County. Figure 6.1 identifies the Little Pine Creek watershed area within a map of Allegheny County watersheds. Little Pine Creek is a tributary of Pine Creek, which flows into the Allegheny River at the town of Etna. The watershed drains a forested hilly terrain for the most part, as indicated by Figure 6.2. Figure 6.3 shows that the dominant soil type in the area is a sandy clay loam, possessing a slow infiltration rate. Figure 6.4 illustrates that the SCS Curve Number values are mostly in the seventies for this watershed, which indicates a high potential for runoff. Photos showing the actual Little Pine Creek from its upstream reaches to the watershed outlet are presented as Figures 6.5, 6.6, 6.7 and 6.8. These figures show how the Creek develops from a small, brush-choked streamlet (Figure 6.5), to a wider permanent channel downstream (Figure 6.8).

Since we use the data from the USGS stream gauge at Little Pine Creek to ground truth our model results, we define the outlet of the Little Pine Creek watershed at the same location as the gauge. We use the fixed base method to separate the baseflow from the direct runoff in the hydrograph defined by the stream gauge data. Here we define the time beyond the hydrograph peak at which direct runoff ends (N) to be equivalent to the time between the peak and 660 minutes, which represent the time that direct runoff ends in most of our model results. Photos of the stream gauge are presented as Figures 6.9 and 6.10. Please refer to Section 5.6 of Data Collection, for more information about the stream gauge, and the data it collects.

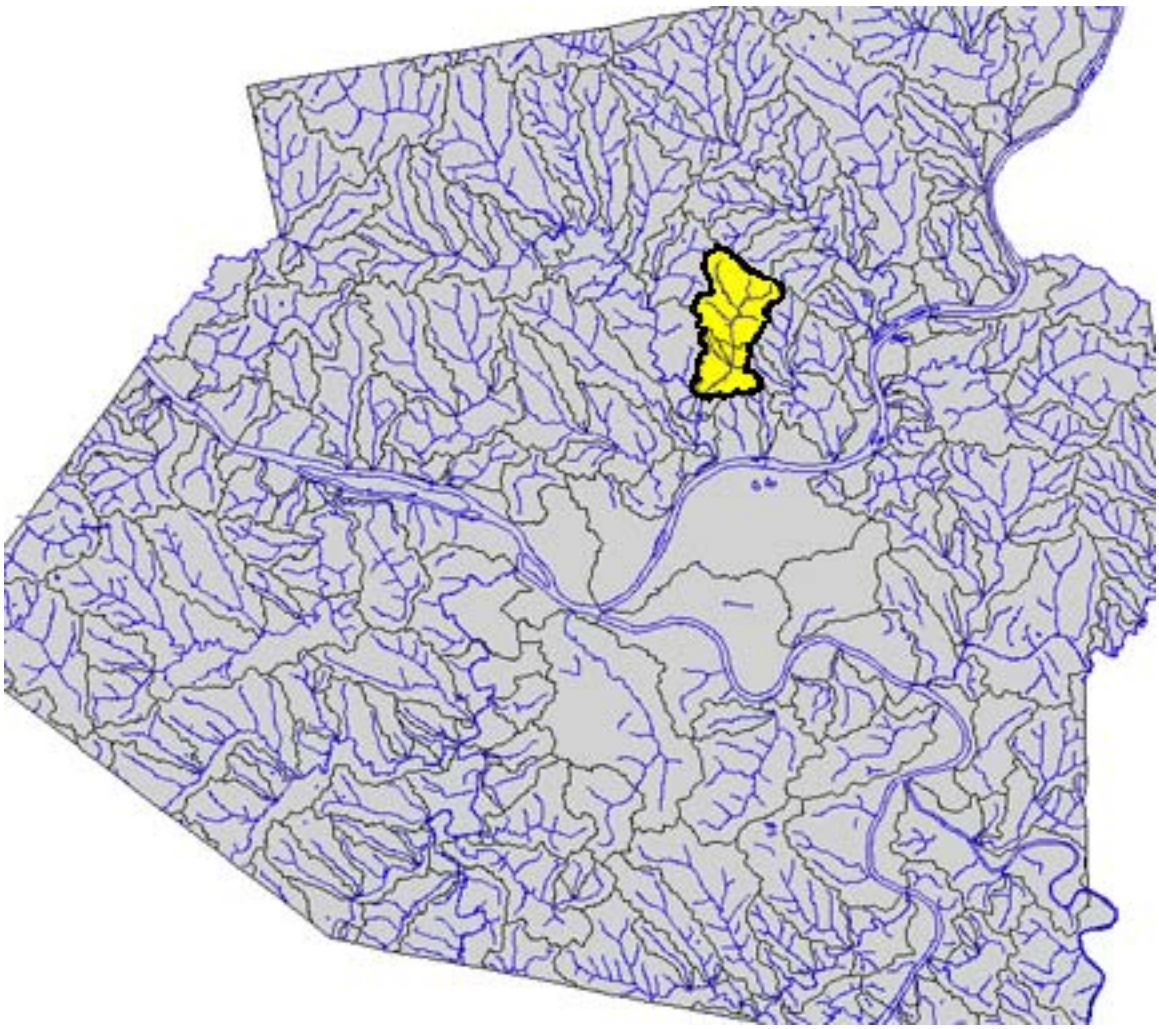


Figure 6.1: Location of Little Pine Creek in Allegheny County

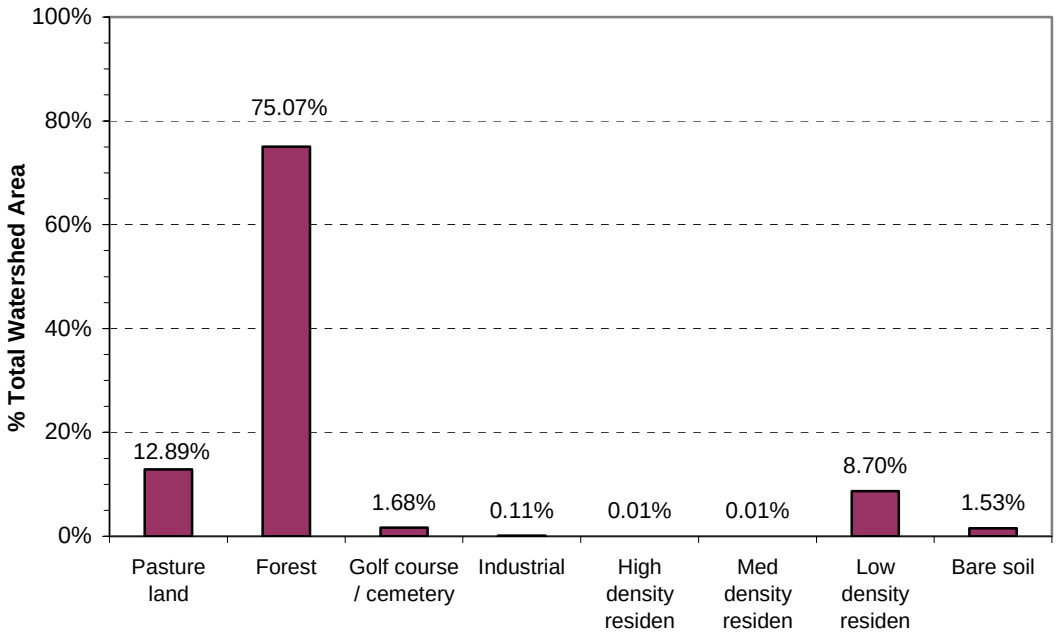


Figure 6.2: Histogram showing distribution of land coverage across Little Pine Creek watershed

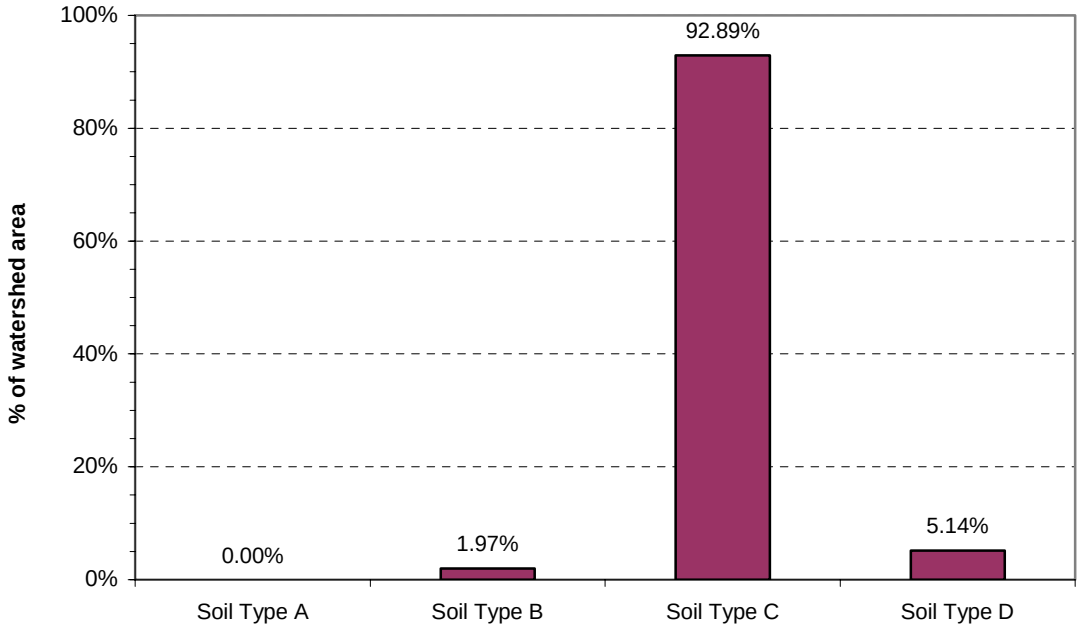


Figure 6.3: Histogram showing distribution of soil types across Little Pine Creek watershed

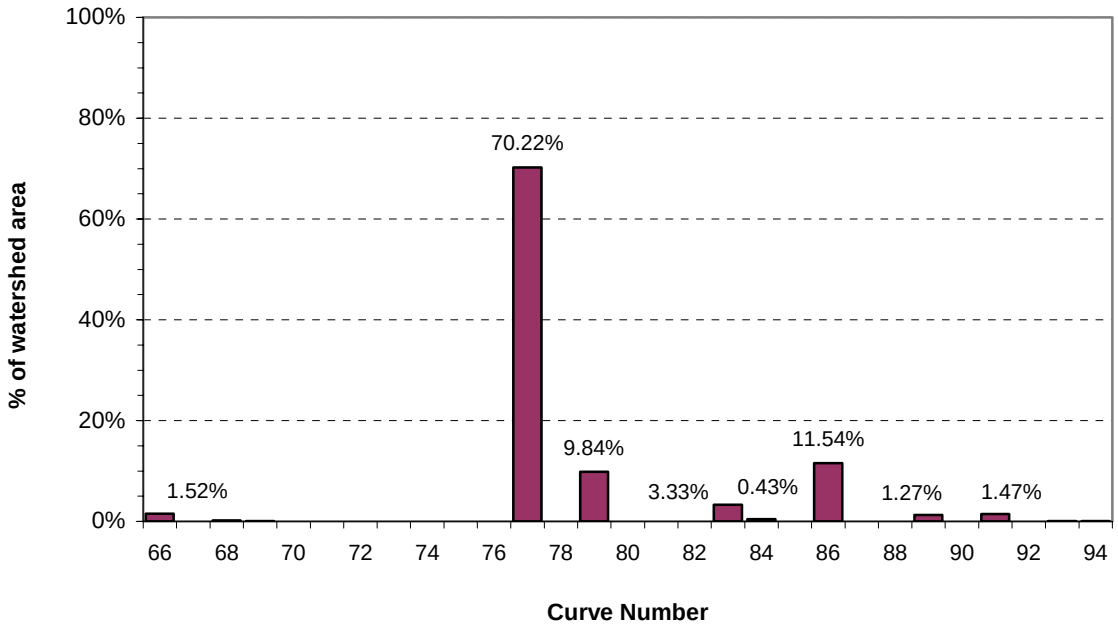


Figure 6.4: Histogram showing distribution of SCS curve numbers across Little Pine Creek watershed

Figure 6.5: Photo of furthest upstream reaches of Little Pine Creek

Figure 6.6: Photo of Little Pine Creek further downstream

Figure 6.7: Photo of Little Pine Creek still further downstream

Figure 6.8: Photo of Little Pine Creek at watershed outlet

Figure 6.9: Photo of USGS stream gauge at Little Pine Creek watershed outlet

Figure 6.10: Photo of creek section measured by UGSG stream gauge

6.2 PRECIPITATION EVENT 1

Since we use the USGS stream gauge data to evaluate the accuracy of our results, we must ensure that our model can describe the conditions in the watershed at the time of interest. In other word, keep it simple! In regards to selecting a precipitation event to use as input, this means we need an event of relatively short duration and considerable magnitude, which occurs on the heels of a relatively long dry period. We examine the model's response for two such precipitation events during the research. The first, also used by Emerick in his 2003 study, occurred during the early morning hours of August 17, 2002. Vieux describes this event as "a small cluster of thunderstorms which developed along a west to east oriented outflow boundary to the north of Allegheny County" (Vieux, 2002). Precipitation began in the northern reaches of the Little Pine Creek watershed at approximately 1:00, and lasted for about 5 hours. Figure 6.11 presents a histogram of 15-minutes incremental precipitation depth values during the course of the storm. Figure 6.12 illustrates the distribution of precipitation depths across the watershed area as given by the precipitation grid at 15-minute intervals. These figures show that greatest precipitation depths occurred over the northern reaches of the watershed during the first half of the storm, while the southern areas received more rainfall during the later part of the event.

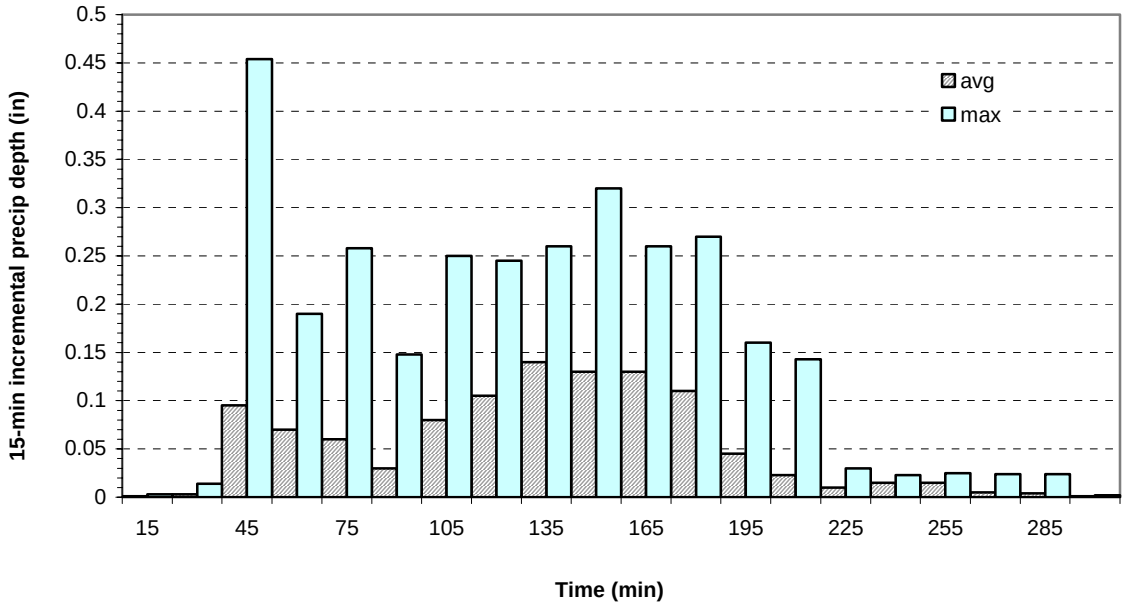


Figure 6.11: Histogram of avg and max precipitation values for 15-minute precipitation increments (8/17/02)

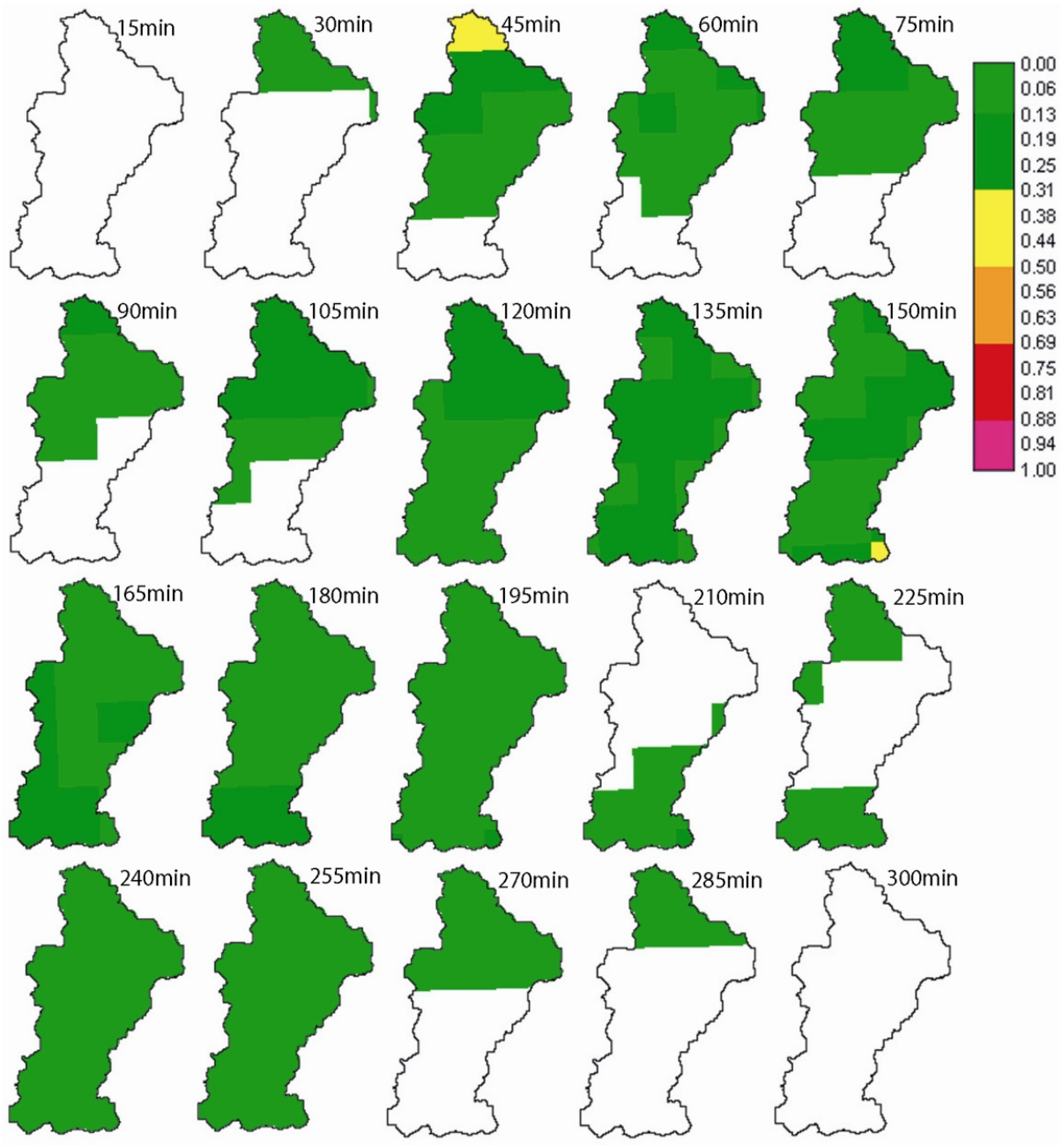


Figure 6.12: Incremental precipitation raster files showing precipitation distribution over time (8/17/02)

6.3 PRECIPITATION EVENT 2

The second event occurred during the late night and early morning hours of July 21, and 22, 2003. Vieux describes this event as "multiple waves of moderate to strong showers and thunderstorms as a system slowly moved through the Ohio Valley" (Vieux, 2003). Precipitation began at approximately 10:00 PM with a brief period of light rain across the watershed outlet area, followed by a period of more intense rainfall lasting until approximately 1:30 AM. By 2:00 AM the event had ended, lasting a total of 255 minutes. Figure 6.13 presents a histogram of 15-minutes incremental precipitation depth values during the course of the storm. Figure 6.14 illustrates the distribution of precipitation depths across the watershed area as given by the precipitation grid at 15-minute intervals. Figure 6.13 shows that that the maximum precipitation depths occurred during the middle of the event, between 105 and 120 minutes. This period was followed by an interval of lesser depths, then by a second peak. Figure 6.14 shows that the northern reaches of the watershed received the greatest precipitation depths during the periods of intense rainfall, but that precipitation depths were distributed fairly evenly across the watershed during the remainder of the event.

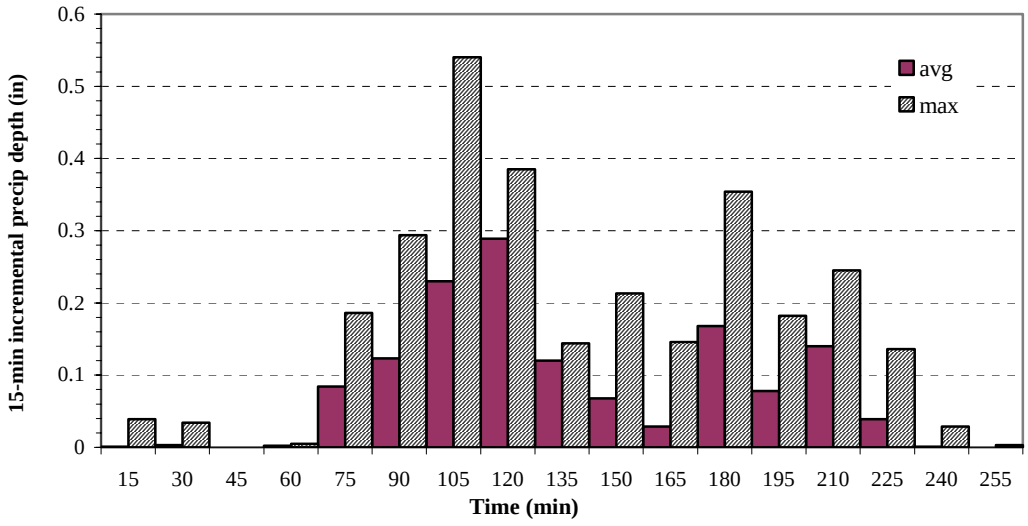


Figure 6.13: Histogram of avg and max 15-minute incremental precipitation depth values (7/21-22/03)

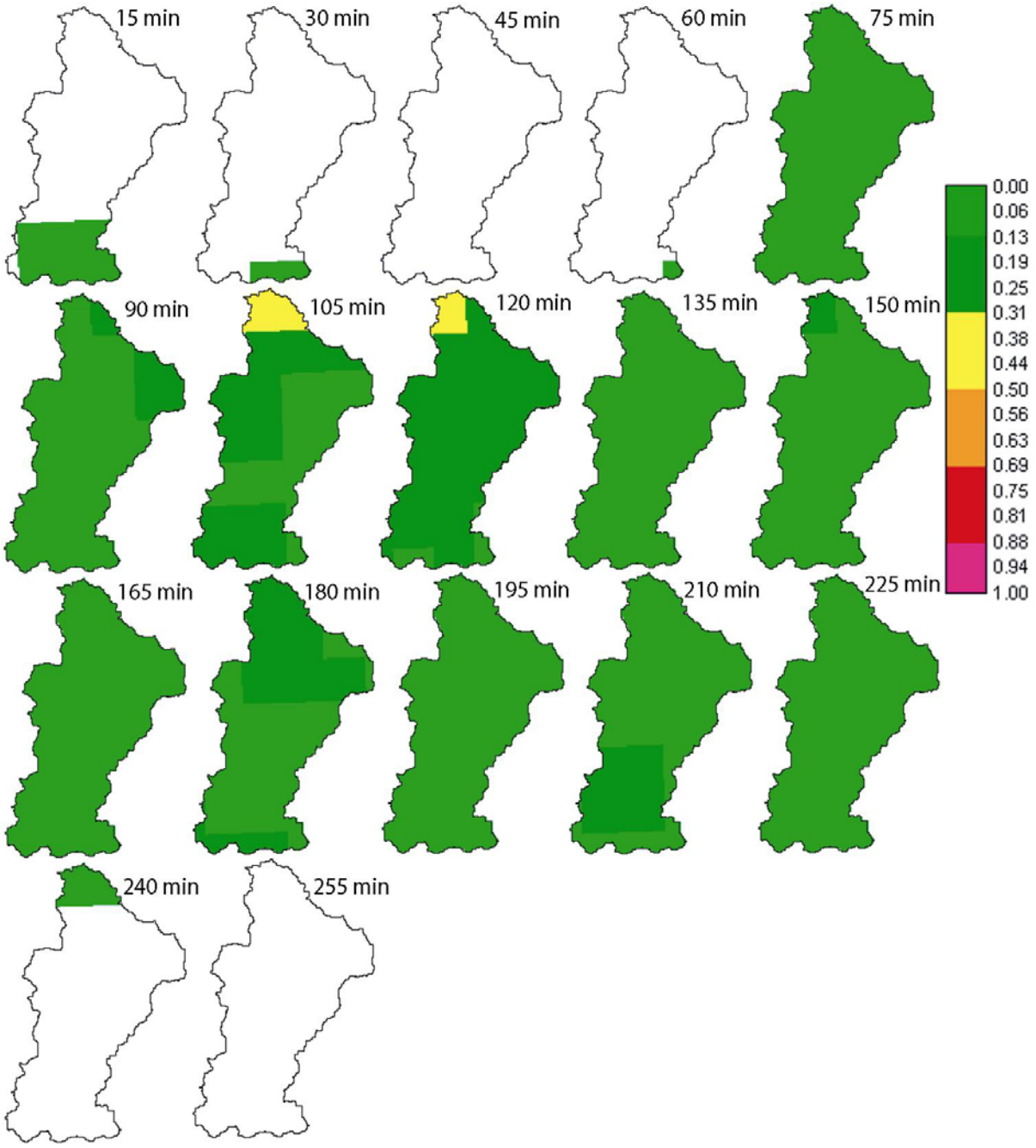


Figure 6.14: Incremental precipitation raster files showing precipitation distribution over time (7/21-22/03)

7.0 RESULTS

7.1 TESTING AND REFINING THE EXISTING MODEL

7.1.1 Correcting Problems with Existing Codes

7.1.1.1 Original vs. New Precipitation Raster File

At the end of his research, Emerick discovered that the watershed boundary for Little Pine Creek generated by the WATERSHED program was inaccurate because it did not include the entire watershed area as described in the published data. He recommended that future investigators use the watershed boundary produced by the program Geo-HMS instead. During the course of the present research, we were unable to reproduce this problem. The Little Pine Creek watershed boundary generated by the WATERSHED program provides a perfectly accurate description of the true watershed area.

Since this problem with the watershed boundary was discovered at the end of the previous research, all the precipitation raster files from during the last study were developed from a precipitation grid file that was created using an inaccurate watershed outline. Figures 7.1 and 7.2 show that the greatest difference between the areas described by the correct and incorrect watershed boundaries occurs in the northern most reaches of the watershed, where the old file underestimates the actual watershed area. This difference in this omitted area is small when compared to that of the entire watershed, so one would expect that any differences between the model's results obtained using either file would be small as well. This might be true under most conditions, but not for data generated using the August 17, 2002 precipitation event. Figure 6.12 shows that the greatest precipitation depths during the entire event occur during the first hour over the northernmost reaches of the watershed. Since the old precipitation grid file does not include some of this area, the results generated using the old and new watershed outlines should differ noticeably.

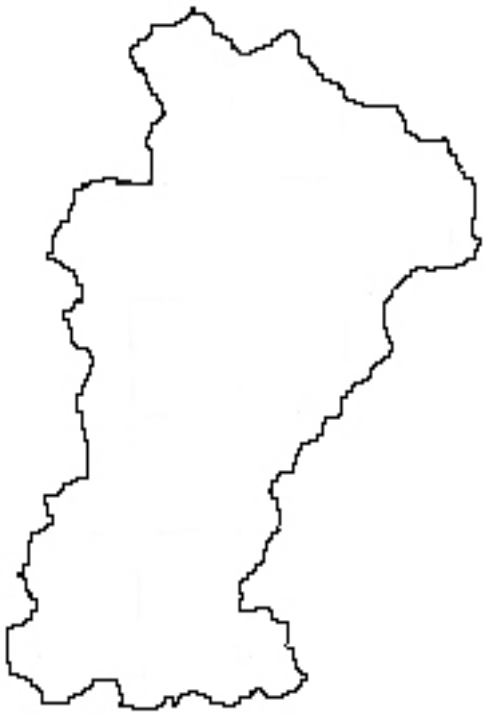


Figure 7.1: Previous watershed outline

Figure 7.2: Corrected watershed outline

Figure 7.3 provides a comparison of the runoff hydrographs at the watershed outlet generated using precipitation grid raster files created from the previous and new watershed boundaries. Table 7.1 compares the peak flow (Q_p) and time to peak (T_p) values for each case, and shows that the resulting values generated using the new file are 14% and 5.5% greater, respectively, than the originals. In light of this discovery, all results from our present research were generated using the correct delineation of the Little Pine Creek watershed area, as defined by the WATERSHED program.

Table 7.1: Hydrograph Qp and Tp values using 2 different Little Pine Creek watershed outlines

Case	Qp (cfs)	Tp (min)
Original precipitation data	154.36	255
New precipitation data	179.88	270
% Difference between data	14.19%	5.56%

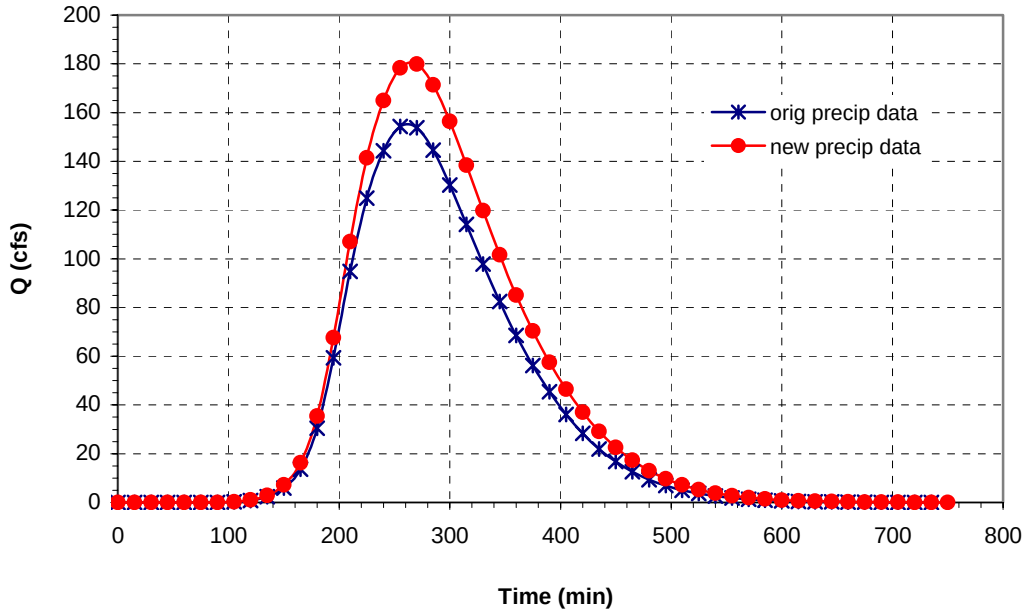


Figure 7.3: Hydrographs resulting from 2 different outlines of Little Pine Creek watershed

7.1.1.2 Improvements to the Runoff Code

While exploring the effect of subdividing the watershed on the hydrograph at the outlet, the code RUNOFF began behaving inconsistently. Cases run using the same number of sub-watershed divisions yielded substantially different hydrograph Q_p results depending on how the sub-watershed divisions were arranged. Further investigation into this matter revealed that the runoff code would generate different runoff volumes for the same area of the watershed, receiving the same depth of precipitation input depending on how it was subdivided. As it turns out, the source of this discrepancy is the way in which the original runoff code computes the 15-minute precipitation time increments required by the Muskingum routing module by averaging the 5-minute incremented time-series used to perform calculations in the code. In the original scheme, the code uses only the first and last of the 5-minute incremental Q values within any 15-minute period to compute the average, neglecting the weights of the second and third values. Averaging in this way, without considering the weights of all four 5-minute incremental values that fall within a particular 15-minute time interval will produce different total runoff values for different sub-watershed arrangements, because each considers different 5-minute incremental Q values

when creating the averages. The updated version of the runoff code, RUNOFF_MTS, fixes this problem by using all four 5-minute incremental Q values (t = 0, 5, 10 and 15) to compute any one 15-minute average. Table 7.2 compares the total runoff generated using the original and new runoff codes for an area in the northwest portion of the watershed shown in Figure 7.4, assuming two and four sub-watershed divisions. The total runoff volume predicted by the new code for this area is essentially the same for both cases, while the volume predicted by the original code differs substantially between the two cases.

The new RUNOFF_MTS code operates with a greater degree of independence from the user as well. The original design requires that the user run the code for each individual sub-watershed to generate runoff. The new design offers the user the option of computing runoff for all sub-watershed areas at once, to save time, or individually.

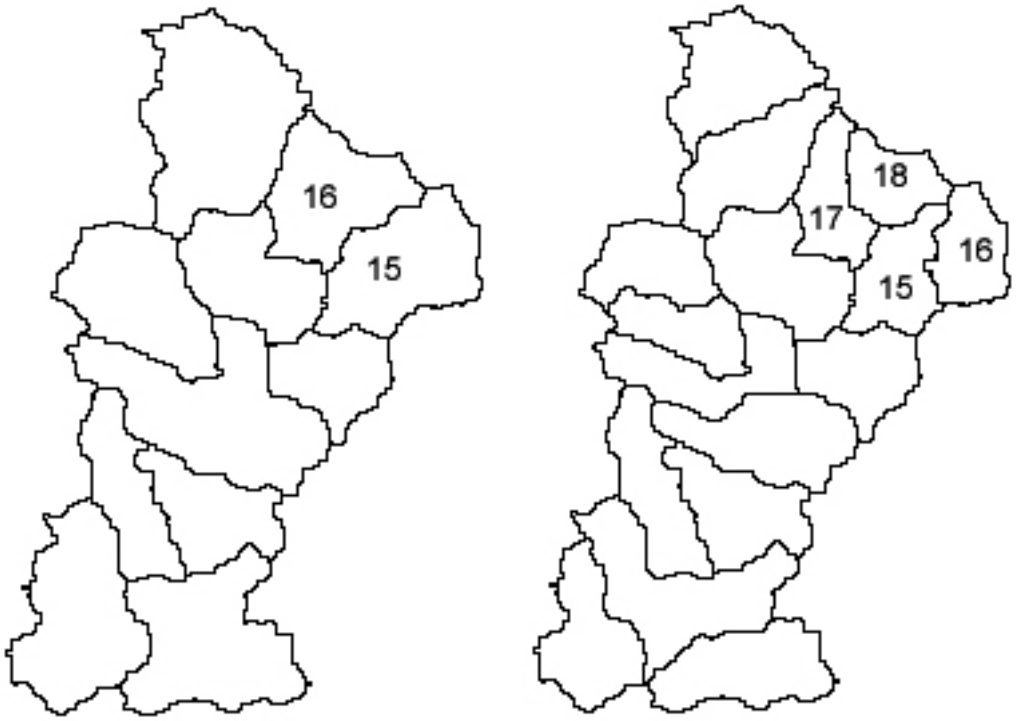


Figure 7.4: Sub-watershed arrangements used to generate runoff data presented in Table 7.2

Table 7.2: Comparison of runoff values generated by code

	Original RUNOFF code						New RUNOFF_MTS code					
	Area 1		Area 2				Area 1		Area 2			
Sbws No.	sbws-15	sbws-16	sbws-15	sbws-16	sbws-17	sbws-18	sbws-15	sbws-16	sbws-15	sbws-16	sbws-17	sbws-18
Runoff Vol (ft^3)	274810	153644	96131	103831	139031	60670	244513	171086	104851	112815	139294	58207
Total Runoff Vol (ft^3)	**428454**		**399663**				**415598**		**415168**			

7.1.2 Effect of Subdividing the Watershed

We examine the effect of increasing the number of watershed subdivisions on the hydrograph at the watershed's outlet by considering the model's response to two separate precipitation events using 5, 7, 9, 11, 13, 15, 17, 19, 21, 23 and 25 sub-watershed divisions. Figure 7.5 illustrates the orientations of the sub-watersheds for each case. All model inputs were held constant for each of the 11 cases. All results were generated using Muskingum flow routing, and the most recent code versions and raster files.

7.1.2.1 August 17, 2002 Precipitation Event

Figure 7.6 illustrates the effect that sub-dividing the watershed has on the outlet hydrograph for cases with 5, 11, 15, 21 and 25 sub-watershed divisions. Between 5 and 11 divisions, the peak discharge (Q_p) decreases and shifts to the right, as the hydrograph spreads out. Between 11 and 15 sub-watershed divisions, the hydrograph peak shifts slightly to the right while Q_p remains constant. Between 15 and 21 divisions, the hydrograph Q_p decreases slightly while T_p remains essentially constant. Between 21 ands 25 divisions Q_p increases slightly, while the hydrograph's time to peak (T_p) remains constant. Figure 7.7 presents a comparison of hydrograph Q_p and T_p values for all 11 cases. This figure shows that Q_p decreases from 223 cfs to 203 cfs between cases with 5 and 9 sub-watershed divisions, respectively. Q_p increases again to 210 cfs for the case with 11 divisions, then varies randomly between cases as the number of divisions increase to 25, while showing an overall decrease to 201 cfs. Figure 7.7 shows that T_p increases from 225 minutes to 255 minutes between cases 5 and 9; then remains fairly constant for the remaining cases.

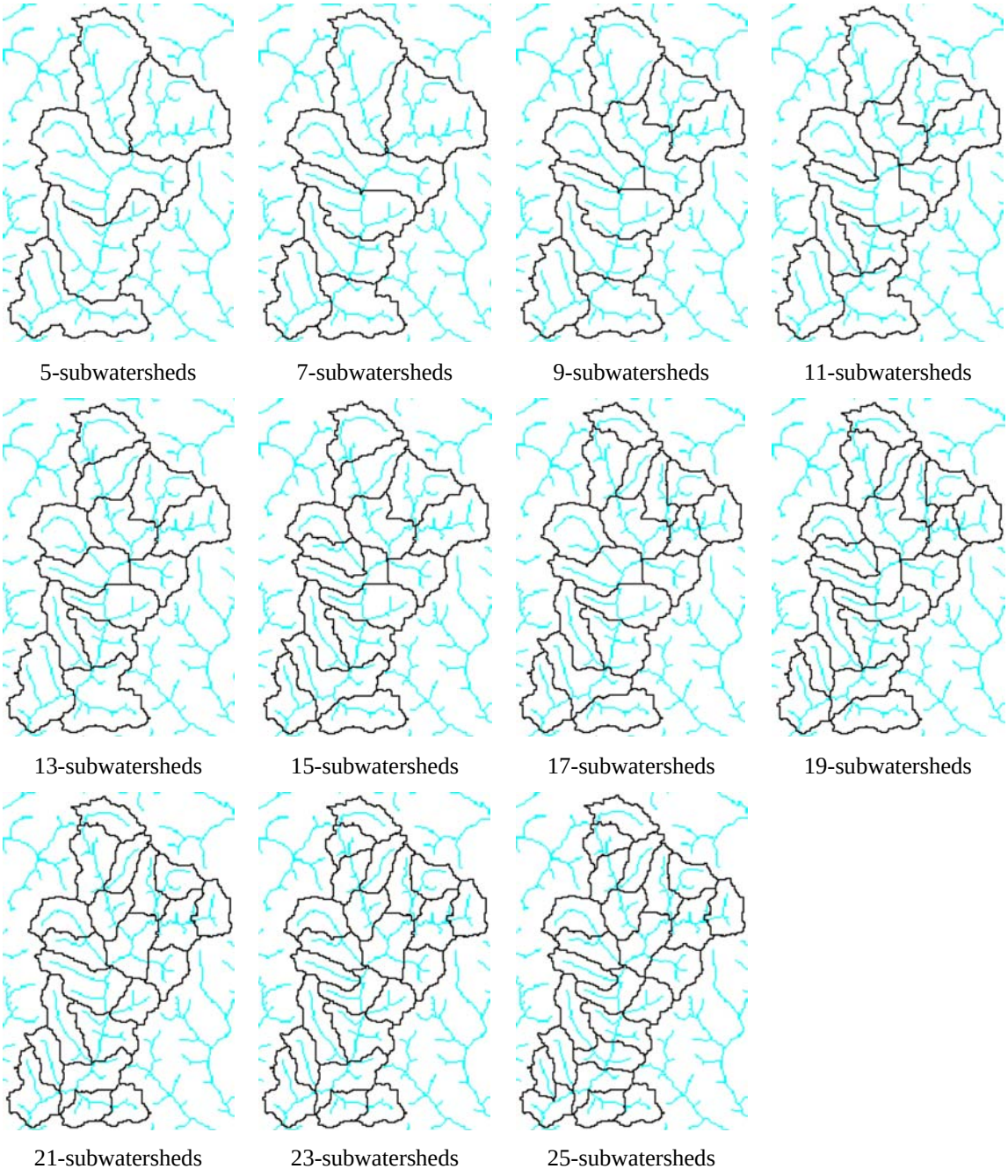


Figure 7.5: Orientation of sub-watersheds in cases used to study effect of watershed subdivision

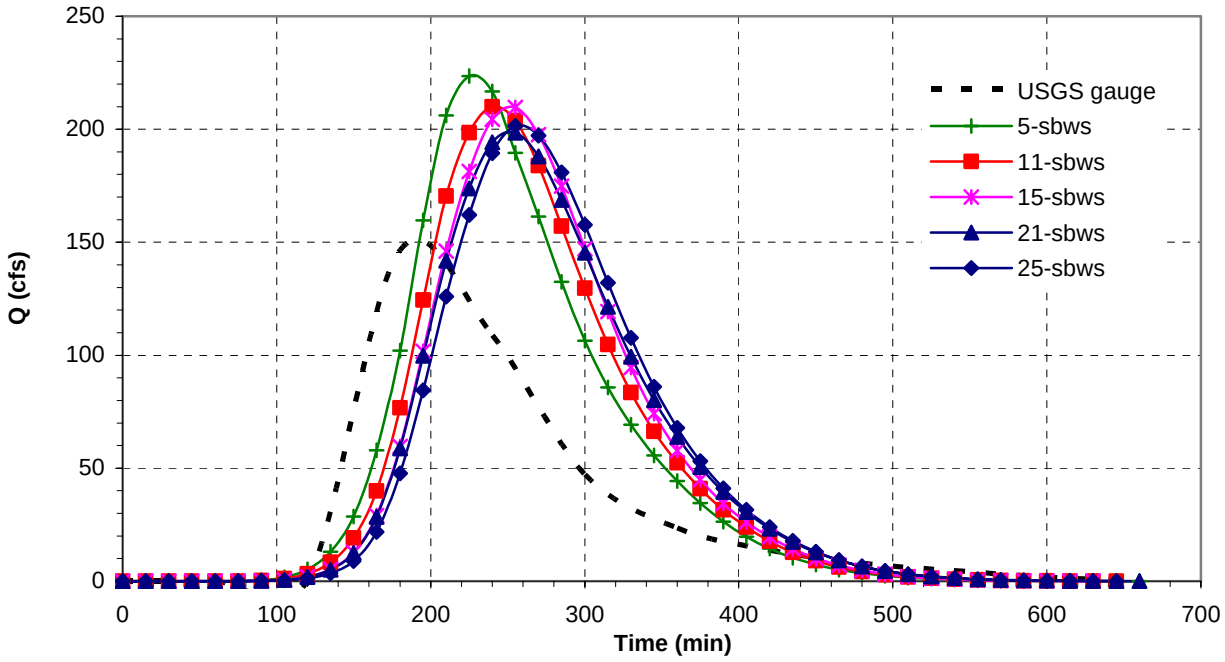


Figure 7.6: Watershed hydrographs generated using 5, 11, 15, 21 and 25 sub-watershed divisions (8/17/02)

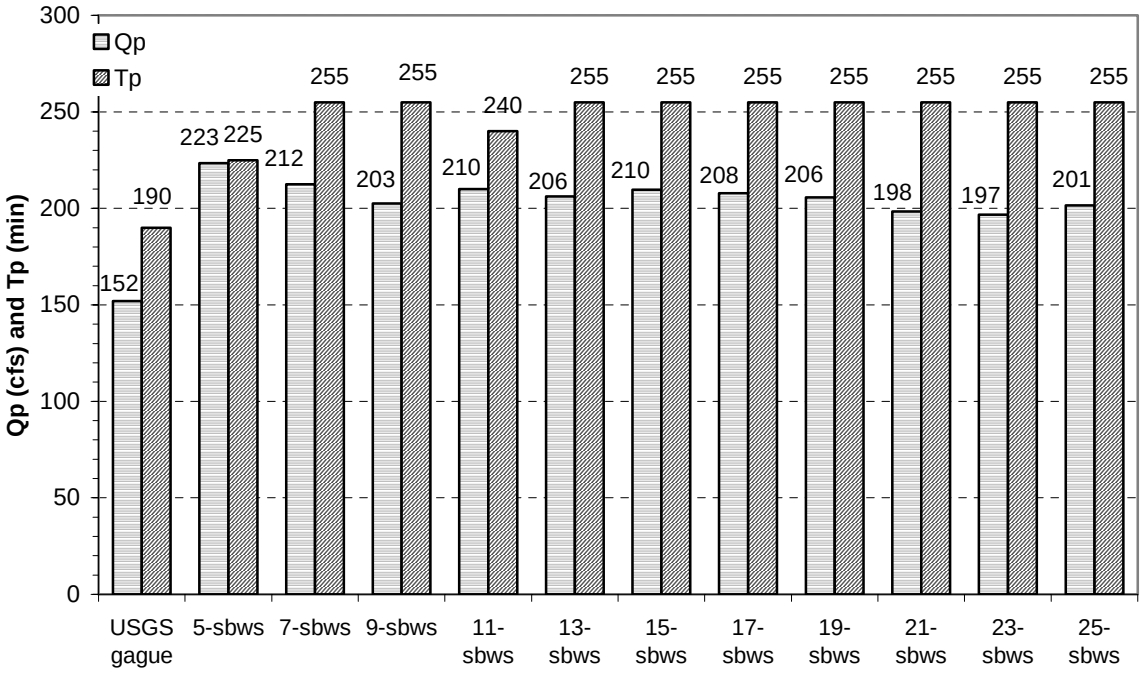


Figure 7.7: Hydrograph Qp and Tp values (8/17/02)

7.1.2.2 July 21-22, 2003 Precipitation Event

Figure 7.8 shows the same general trend of peak attenuation and rightward shift observed in the model hydrographs for the August 17, 2002 event. In this case however, the trend continues through the case with 25 sub-watershed divisions. Figure 7.9 shows that Q_p decreases overall from 355 cfs to 293 cfs as the number of sub-watershed divisions increases from 5 to 25, respectively. T_p increases from 240 minutes to 270 minutes as the number of watershed subdivisions increases from 5 to 11; then remains fairly constant before increasing again to 285 minutes at 25 divisions.

7.1.2.3 Discussion

Considering the data, it appears that increasing the number of sub-watershed divisions will cause Q_p to decrease overall, but this becomes far less significant when the number of sub-watersheds increases beyond 9 to 11 sub-watersheds divisions. Similarly, T_p values are more-or-less stable for cases with 7 sub-watershed divisions or greater. Thus, we can conclude that increasing the number of sub-watershed divisions beyond 9, or 11 at the very most, has little effect on the resulting hydrograph for the Little Pine Creek Watershed. Table 7.3 lists the areas of each sub-watershed for both the cases with 9 and 11 divisions. The average area for case 9 is 412 acres, while the average for case 11 is 337 acres.

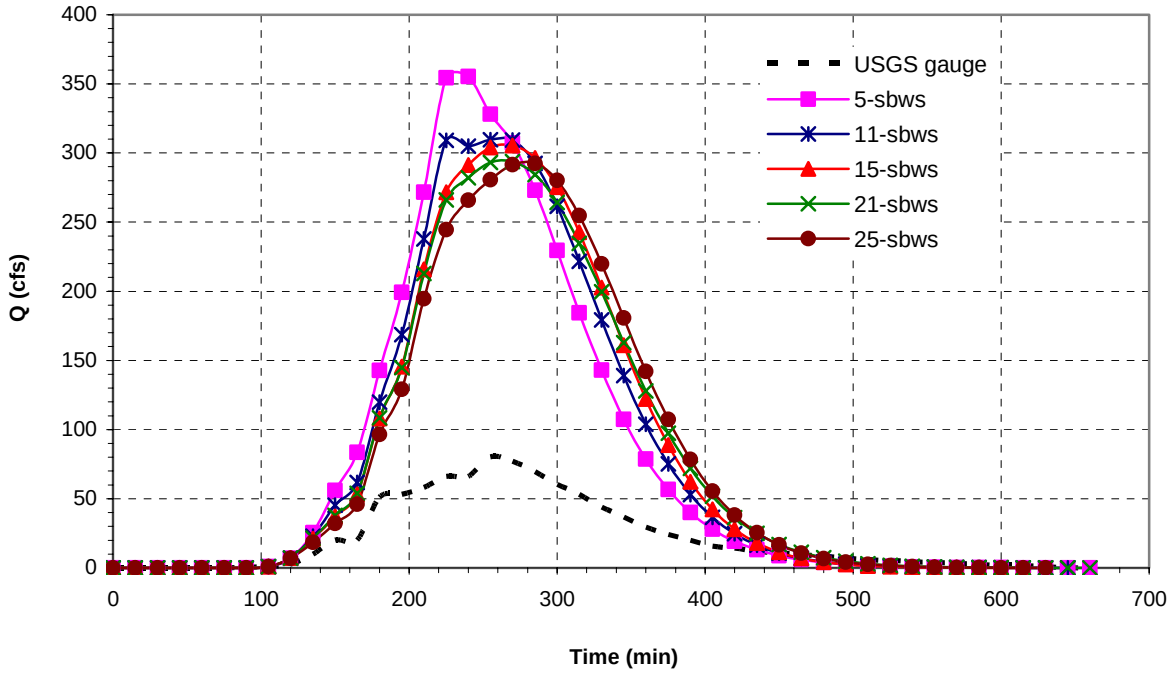


Figure 7.8: Watershed hydrographs generated using 5, 11, 15, 21 and 25 sub-watershed divisions (7/21-22/03)

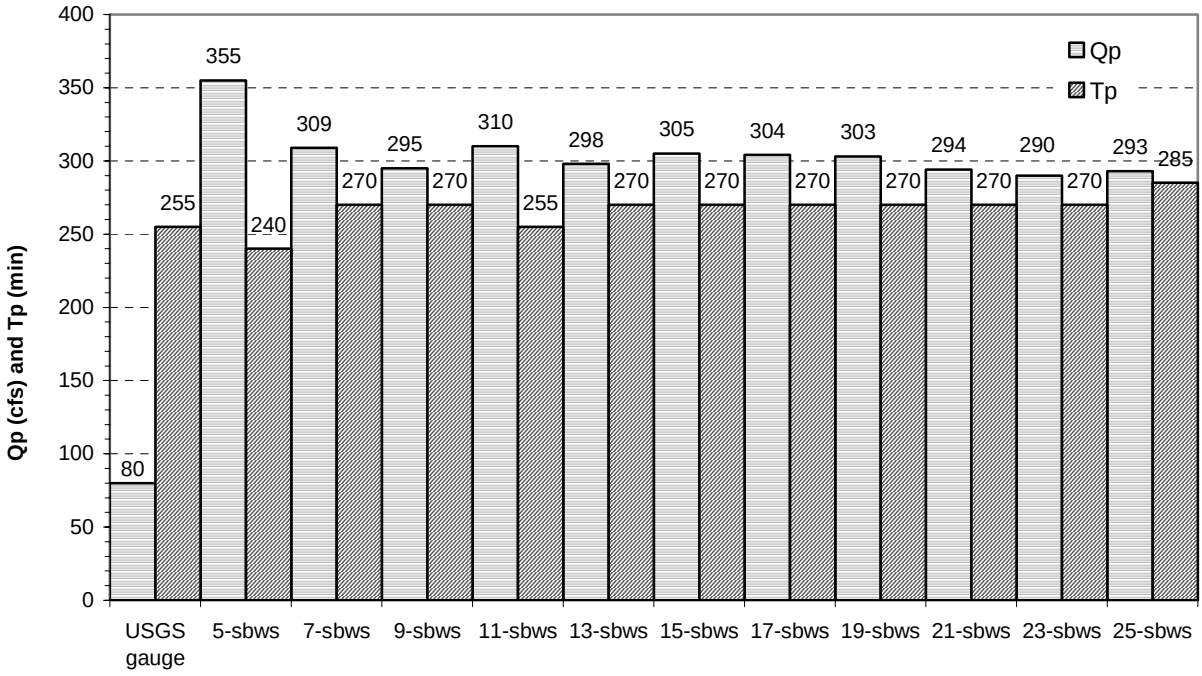


Figure 7.9: Hydrograph Qp and Tp values (7/21-22/03)

Table 7.3: Size of sub-watershed areas for cases with 9 and 11 sub-watershed divisions

Sub-watershed number	Area (acres)	
	9 sub-watershed divisions	11 sub-watershed divisions
sbws-2	390	390
sbws-3	367	418
sbws-4	423	219
sbws-5	467	215
sbws-6	419	494
sbws-7	477	338
sbws-8	535	196
sbws-9	333	274
sbws-10	299	535
sbws-11	-	333
sbws-12	-	299
Avg Area	**412**	**337**

7.1.3 Effect of Muskingum X value on Outlet Hydrograph

Chow, 1988, states that values of the Muskingum X coefficient may vary between 0 and 0.3 in natural streams. According to Dr. Quimpo however, an upper limit of X = 0.2 is more realistic. We examine the effect that the Muskingum X coefficient has on the predicted watershed hydrograph by comparing several such hydrographs generated using X values of 0.0, 0.04, 0.08, 0.12, 0.16 and 0.2, for cases with 5, 11, 15, 21 and 25 sub-watershed divisions. We examine the results from two separate precipitation events here, but hold all other model parameters constant between cases. All results are computed using the most recent code versions.

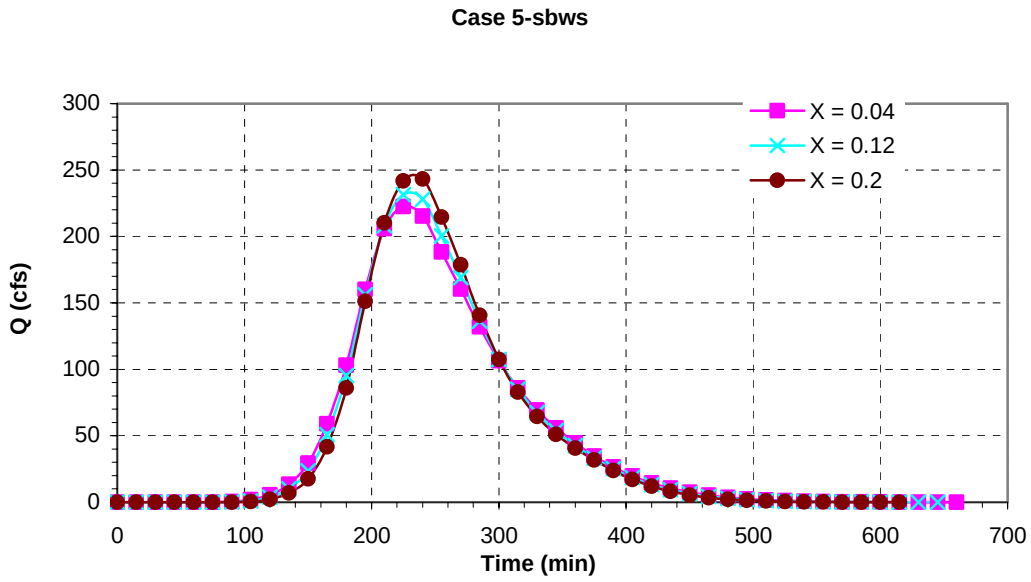
7.1.3.1 August 17, 2002 Precipitation Event

Figure 7.10 compares the hydrographs generated using Muskingum X values of 0.04, 0.12 and 0.2, for cases with 5 and 21 sub-watershed divisions. Table 7.4 presents the hydrograph Q_p and T_p values generated using different X values for each of the five cases. Figure 7.11 illustrates the variation of Q_p with X and the number of sub-watershed divisions. These data show that Q_p increases with increasing X, and decreases with an increasing number of sub-watershed divisions; but what is of interest here is the magnitude of the difference between the peak flow values (ΔQ_p) generated by each X value. Table 7.4 lists these ΔQ_p values for each sub-watershed arrangement expressed numerically and as a percentage difference. The Q_p difference between

cases using X = 0 and X = 0.2 ranges between 10.23% and 13.41% for the five sub-watershed cases. Figure 7.12 shows that varying X has no effect on the hydrograph T_p values.

Table 7.4: Hydrograph Qp and Tp values at varying Muskingum X values, and percentage difference in Qp (8/17/02)

Case	5-sbws		11-sbws		15-sbws		21-sbws		25-sbws	
X value	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)
X = 0	218	225	204	240	203	255	192	255	196	255
X = 0.04	222	225	209	240	208	255	197	255	200	255
X = 0.08	227	225	214	240	214	255	203	255	205	255
X = 0.12	231	225	220	240	220	255	208	255	210	255
X = 0.16	236	225	226	240	226	255	215	255	216	255
X = 0.2	243	240	232	240	233	255	222	255	222	255
ΔQp	25 10.23%		28 12.26%		30 12.85%		30 13.41%		26 11.77%	



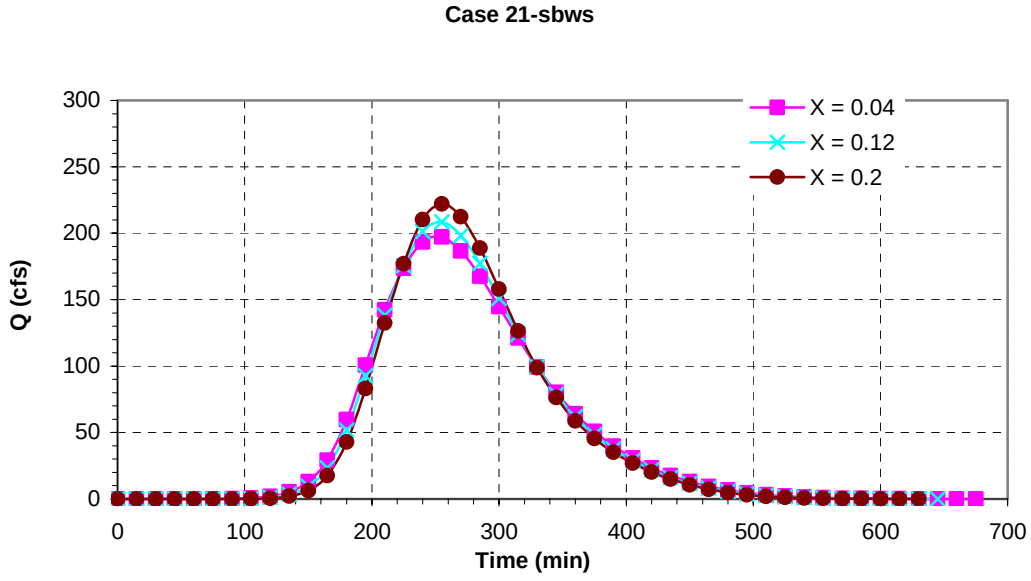


Figure 7.10: Hydrographs at different X values for 5 and 21 sub-watershed divisions (8/17/02)

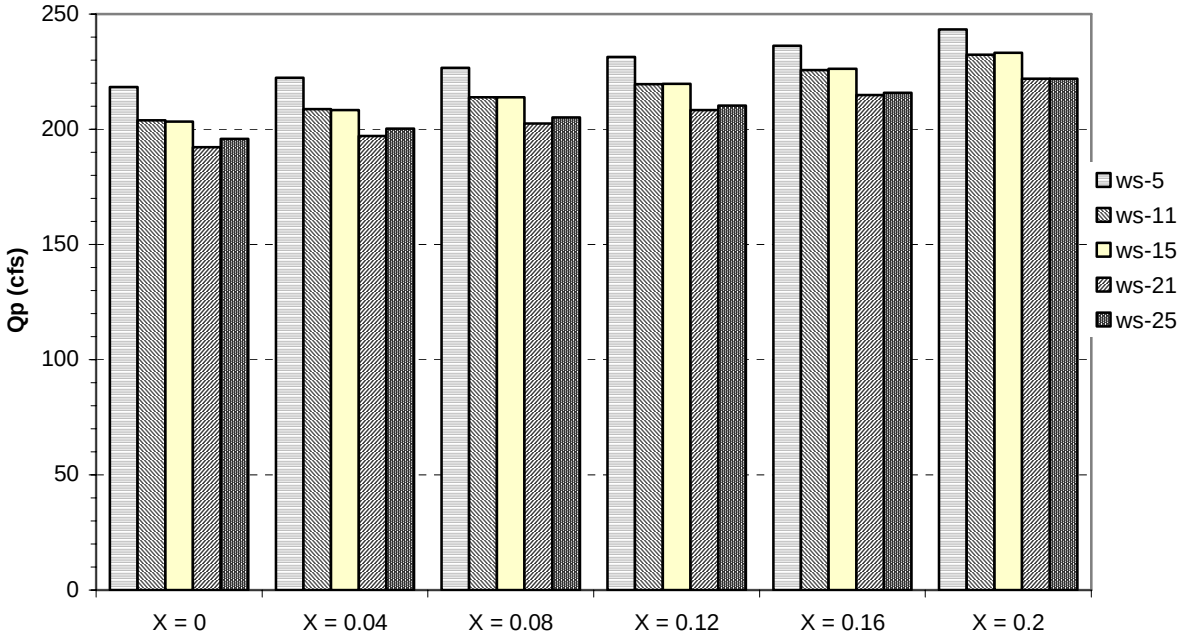


Figure 7.11: Variation of hydrograph Qp values with X for 5, 11, 15, 21 and 25 sub-watershed divisions (8/17/02)

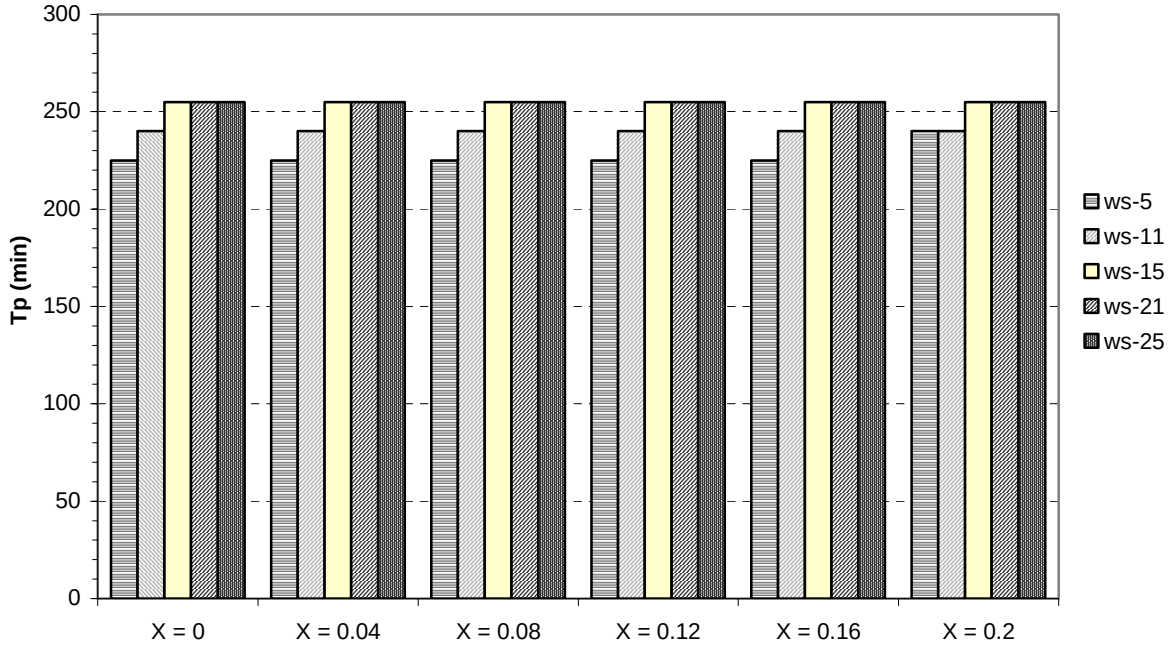


Figure 7.12: Hydrograph Tp values for varying X using 5, 11, 15, 21 and 25 sub-watershed divisions (8/17/02)

7.1.3.2 July 21-22, 2003 Precipitation Event

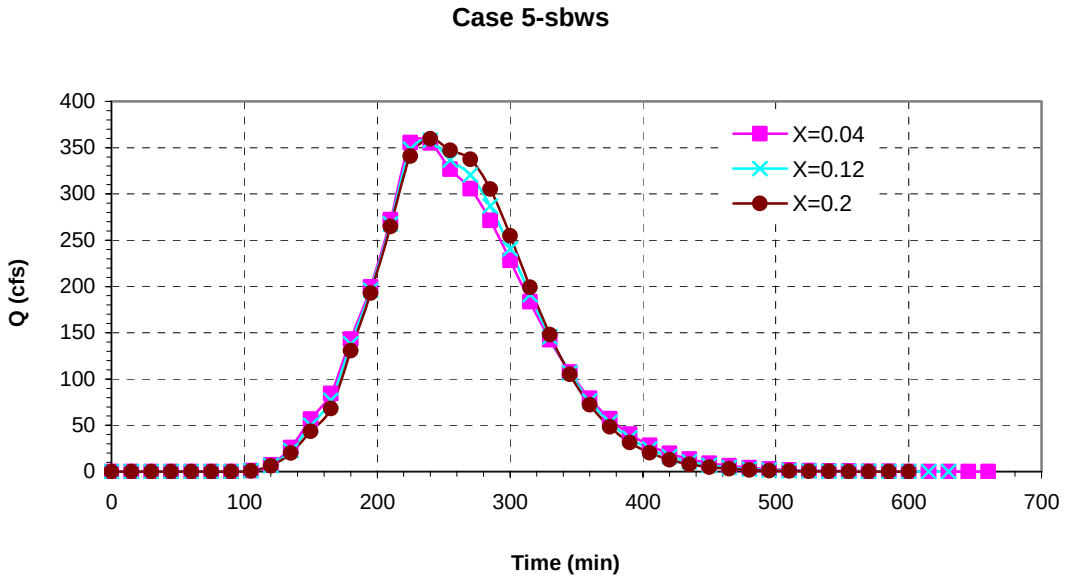
Figure 7.13 compares the hydrographs generated using Muskingum X values of 0.04, 0.12 and 0.2, for the cases with 5 and 21 sub-watershed divisions. Table 7.5 presents the hydrograph Q_p and T_p values generated using different X values for all five sub-watershed arrangements. Figure 7.14 illustrates the variation of Q_p with X and the number of sub-watershed divisions. Like the results from the first precipitation event, these data show that Q_p increases with increasing X, and decreases with an increasing number of sub-watershed divisions. For this precipitation event however, ΔQ_p is about half of what it was for the results from the first event. These values show a greater range (0.56% - 7.85%) as well. Figure 7.15 shows that hydrograph T_p values will increase slightly as X increases between 0.0 and 0.08, but then remain constant.

7.1.3.3 Discussion

It is apparent for both precipitation events that lower values of X produce greater attenuation of the hydrograph Q_p. This results is expected, considering that X is the variable used to describe the effect of inflow on reach storage, and that at X = 0 the system reduces to that describing a linear reservoir, with maximum storage. In looking at the range of ΔQ_p values for each event though, it is apparent that the magnitude of change in peak flow with X relatively small. The results for both event show that varying Muskingum X has very little effect on hydrograph T_p. Thus, we can conclude that varying the value of the Muskingum X constant between 0 and 0.2 has only a limited effect on the watershed outlet hydrograph predicted by our model.

Table 7.5: Qp and Tp values at varying Muskingum X values, and percentage difference in Qp (7/21-22/03)

Case	5-sbws		11-sbws		15-sbws		21-sbws		25-sbws	
X value	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)	Qp (cfs)	Tp (min)
X = 0	358	225	305	255	301	255	290	255	288	270
X = 0.04	355	225	309	255	304	270	293	270	291	285
X = 0.08	356	240	314	270	308	270	297	270	296	285
X = 0.12	358	240	319	270	312	270	301	270	301	285
X = 0.16	359	240	325	270	316	270	305	270	306	285
X = 0.2	360	240	331	270	319	270	309	270	311	285
ΔQp	2 0.56%		26 7.85%		18 5.64%		19 6.15%		23 7.40%	



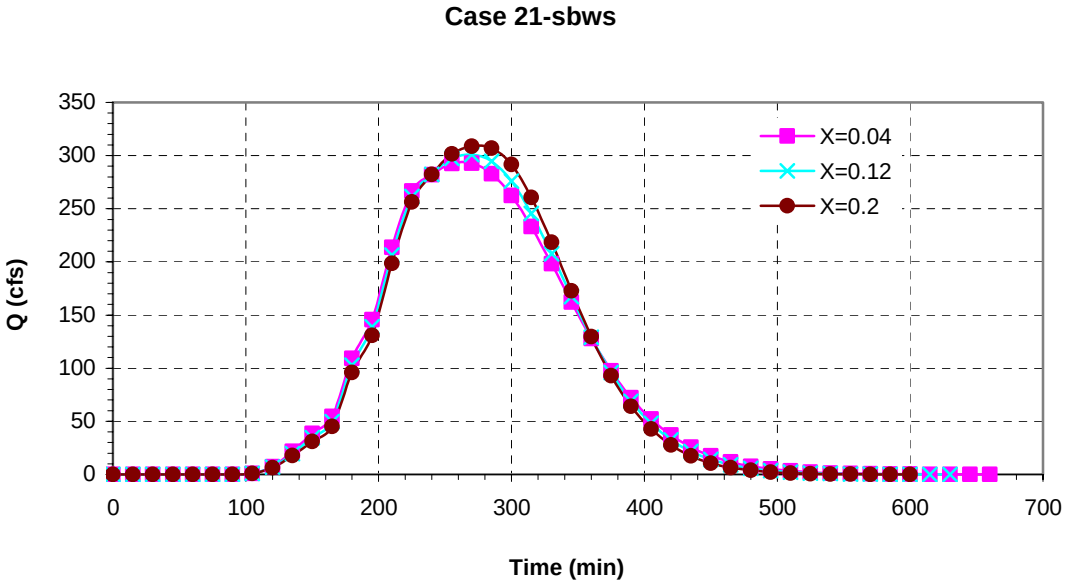


Figure 7.13: Hydrographs at different X values for 5 and 21 sub-watershed divisions (7/21-22/03)

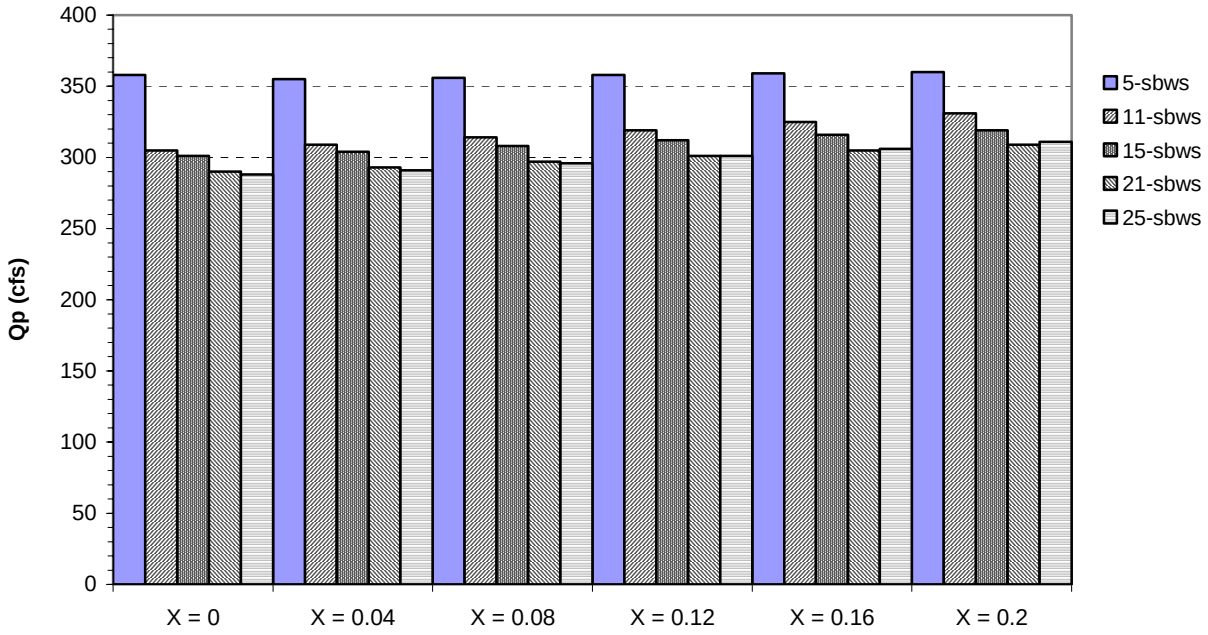


Figure 7.14: Variation of hydrograph Qp values with X for 5, 11, 15, 21, 25 sub-watersheds (7/21-22/03)

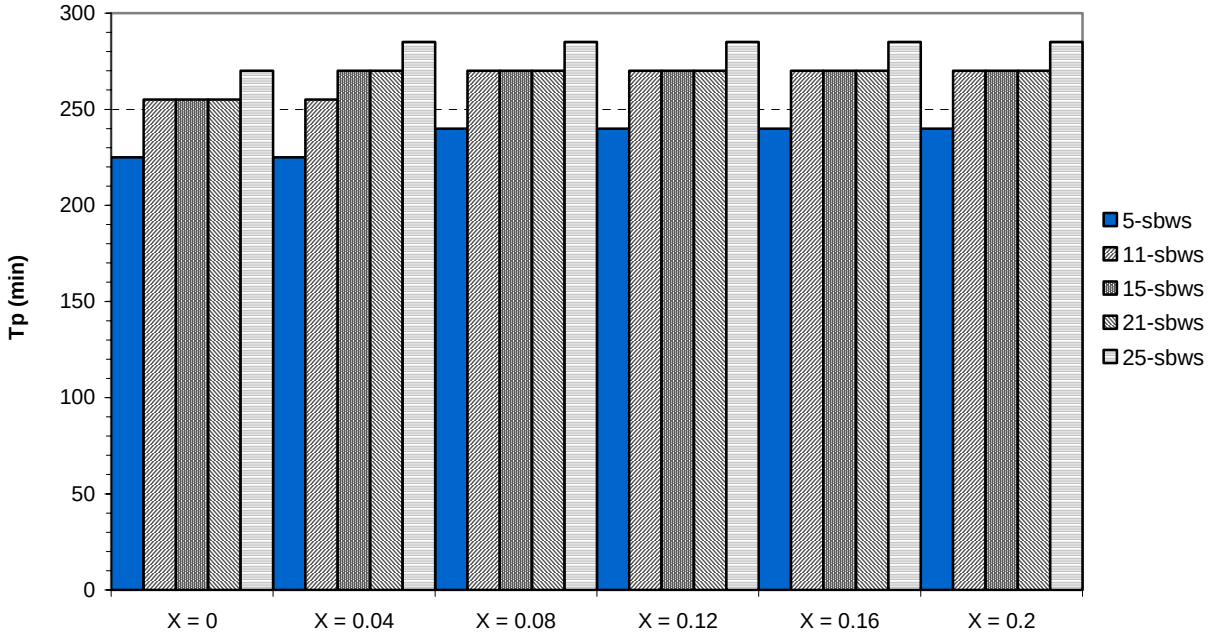
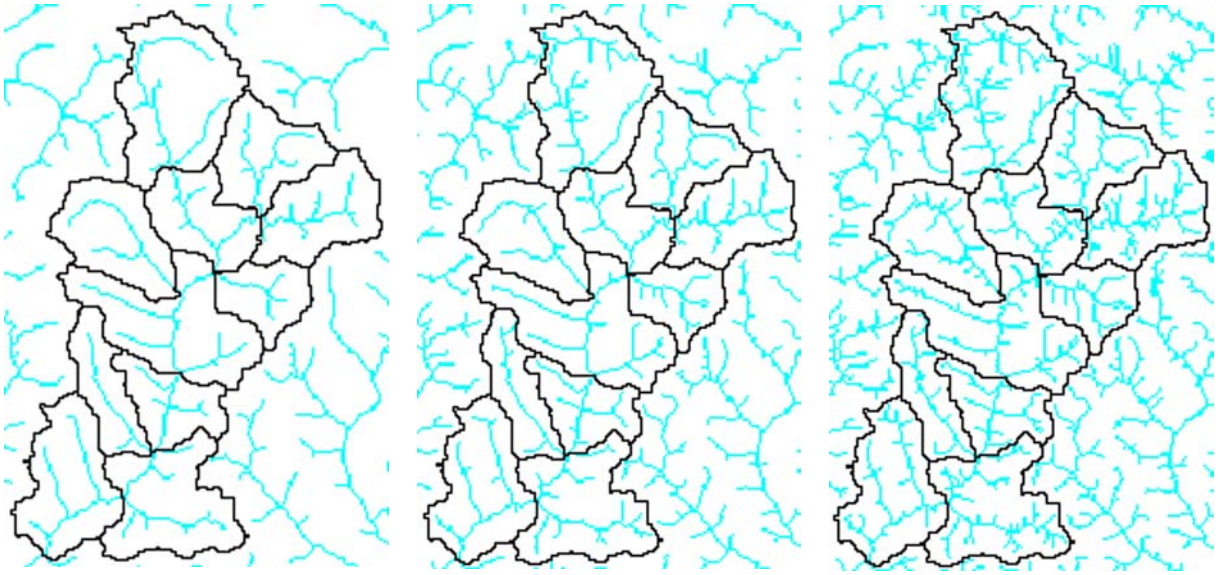


Figure 7.15: Hydrograph Tp values for varying X using 5, 11, 15, 21 and 25 sub-watersheds (7/21-22/03)

7.1.4 Effect of Varying Threshold for Channel Cell Flow

During this portion of the study, we examine the effect that lowering the threshold number of flow contributing cells required to distinguish between channel and overland flows has on the resulting hydrograph at the watershed outlet. Figure 7.16 compares the stream file generated using the standard cut off of 100 flow contributing cells used to classify flow within a cell as channelized, with those that result from using cut off values of only 40 and 20 cells. We examined the models response for both precipitation events using the case with 11 sub-watershed divisions, and Muskingum routing. As expected, the stream files created using a lower channel flow threshold contain more stream channels. Figures 7.17 and 7.18 shows the effect this has on the watershed outlet hydrograph for the August 17, 2002, and July 21-22, 2003 precipitation events, respectively. Figure 7.19 compares the Q_p and T_p values for the three cases for both precipitation events. For the August 17, 2002 precipitation event, as the threshold value drops from 100 to 40 cells, Q_p rises from 210 cfs, to 211 cfs; but when the threshold is lowered further to 20 cells, Q_p remains essentially constant at 211 cfs. The total difference between these Q_p values is only 0.59%. Lowering the channel cell threshold has no effect on the hydrograph T_p, which remains steady at 240 cfs for all three cases. The results for the July 21-22, 2003 event show much the same pattern. Here, Q_p increase slightly from 310 cfs to 312 cfs upon lowering the threshold from 100 to 40 cells; then remains constant. T_p is steady at 255 min for all three cases. These data indicate that lowering the threshold number of flow-contributing cells used to define a channel produces little change in the hydrograph at the watershed outlet, and certainly nothing that can be considered an improvement over the standard 100-cell threshold used in the present model.



100-cell threshold 40-cell threshold 20-cell threshold

Figure 7.16: Stream files for 100, 40 and 20-cell channel flow thresholds

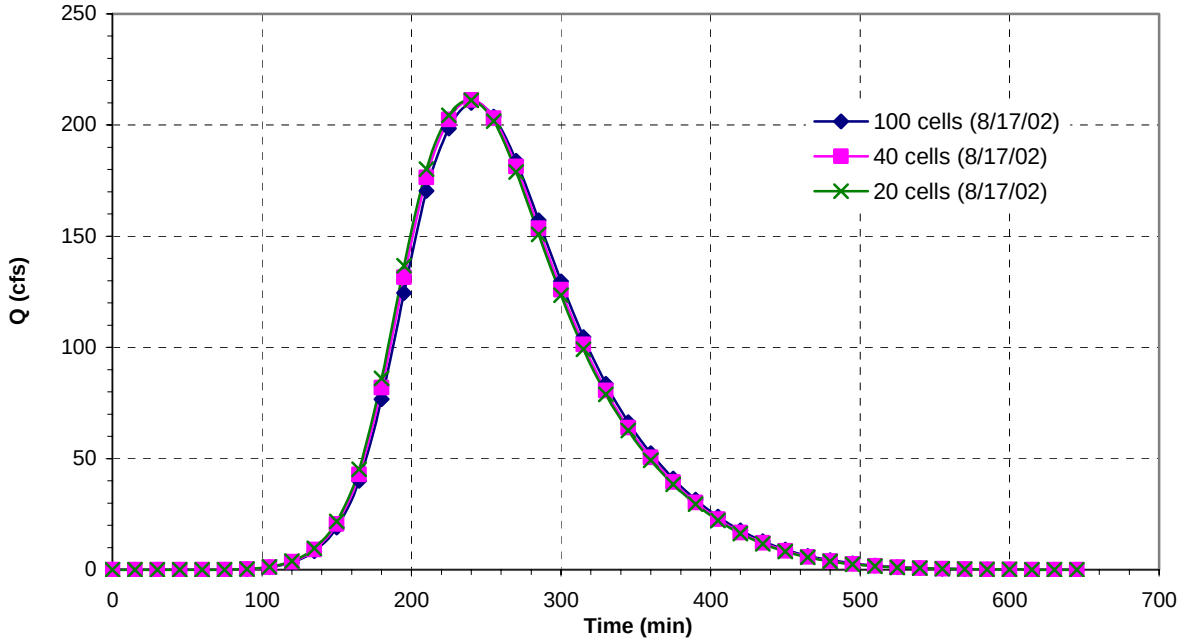


Figure 7.17: Watershed hydrograph for 100, 40 and 20-channel flow cell thresholds (8/17/02)

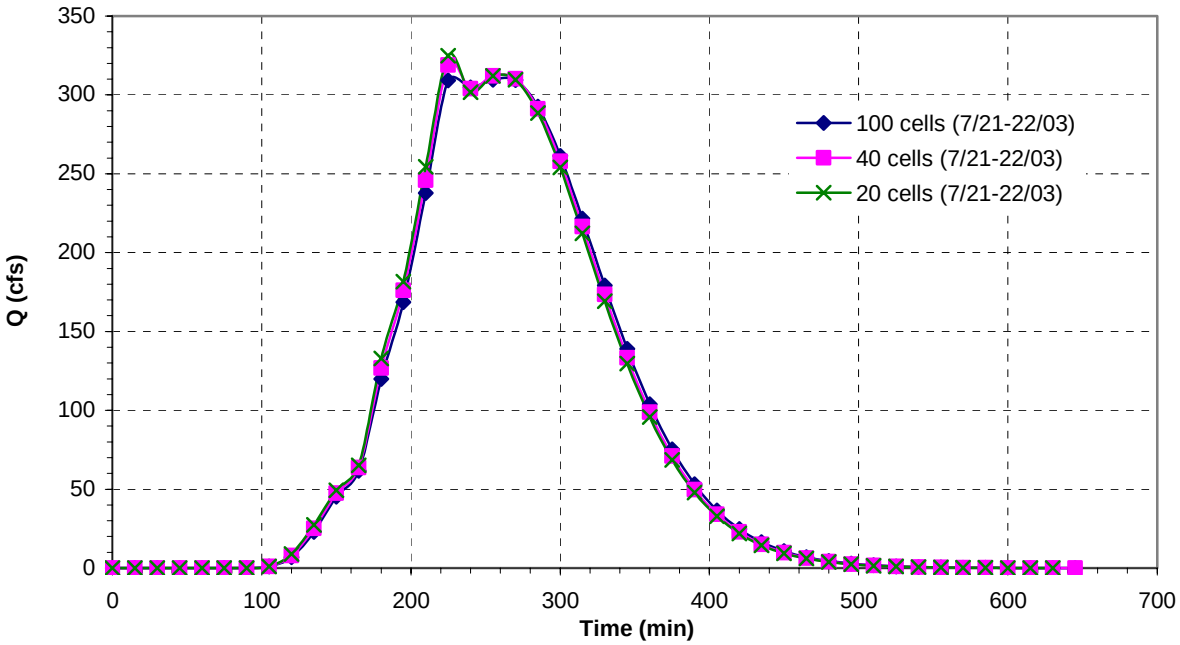


Figure 7.18: Watershed hydrograph for 100, 40 and 20-channel flow cell thresholds (7/21-22/03)

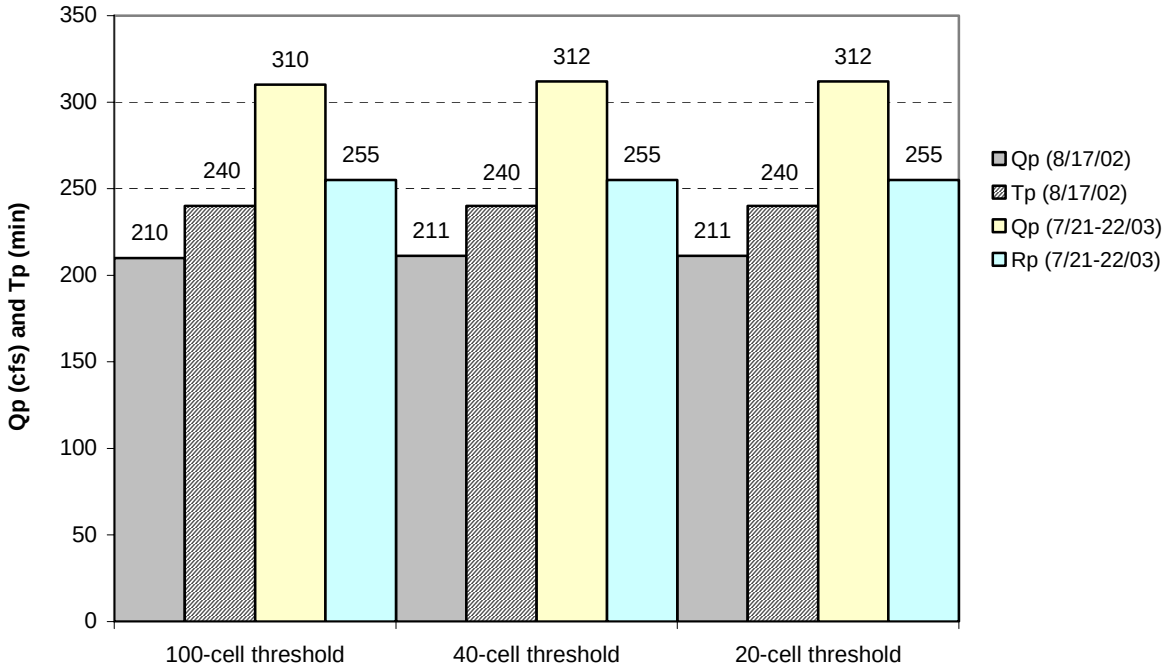


Figure 7.19: Watershed hydrograph Qp and Tp values for 100, 40 and 20-channel flow cell thresholds

7.1.5 Studying Inconsistencies with the Isochrone Code

The motivation behind this portion of the study was to explain and resolve problems with the raster file of TOC values output by the isochrone code. While reviewing results produced during other research, we observed that the time of concentration values for channel cells would suddenly increase at certain locations moving from upstream section of the reach toward the outlet of a sub-watershed. Figure 7.20 illustrates this phenomenon, showing a close up view of the isochrone raster file for the case with 11 sub-watershed divisions. The TOC values of the channel cells are noted on the figure at locations where they increase or decrease. It is obvious from looking at the figure that the problem is better described as one of inconsistent channel TOC values, than simply the occasionally high values downstream. Though the majority of problem cells contain higher TOC values than their upstream and downstream neighbors, some are out of place because they contain low values within a zone of higher values. Why does this happen? The answer lies in the way that the original code computes TOC values for channel cells.

The original ISOCHRONE code computes TOC values for each channel cell within a particular sub-watershed individually, using equation 7-1, which is a relationship between velocity, given by the Ramser equation, and the distance from the cell in question to the outlet of the sub-watershed.

$$\text{TOC Channel} = 0.0195(\text{Length of Channel})^{0.77} / (\text{Avg. Channel Slope})^{0.385} \qquad (7\text{-}1)$$

In this equation, the average channel slope is obtained by summing the slopes values between the individual cells along the flow path from the cell in question to the outlet, and dividing this by the number of cells in the flow path. This method works fine for areas where the slope between cells remains fairly constant; but breaks down at locations where large changes in slope occur. For example, consider the point where a steeply sloping reach (say 1/20) flows into a zone with a gentle slope (say 1/500). Above this break in slope, the average slope values used in equation 7-1 are weighted by the larger slope values of the steeper reach section, which results in lower TOC values. Below the slope break however, the average slope values are much smaller because they are not weighted by values from the steeply sloping portion of the reach. Thus TOC values

downstream from the break in slope will be greater than those upstream. This situation will also occur at sudden drops within section of reach with a gradual slope.

Upon identifying the source of the anomalous TOC values, we attempted correct the problem by creating a new version of the isochrone code that employed a new method by which to compute the slope used in the channel TOC calculations. Several such methods were examined during this process, including:

- Using different value of slope (S_0 = 0.01, 0.001, 0.0001) assigned by the code to a horizontal sections of reach (i.e. where two adjacent cells have the same elevation).
- Computing the average slope value used in equation 7-1 as simply the elevation difference between the channel cell in question and the cell representing the sub-watershed outlet.
- Computing channel TOC values for several cells defining a section of a reach at once rather than individually.

Some of these methods corrected part of the problem, and some simply made things worse. In short nothing we tried got rid of the anomalous isochrone values completely. We obtained the best results by changing the slope value assigned to horizontal sections of reach from S_0 = 0.001, used in the original code, to S_0 = 0.01. The new isochrone code resulting from this study, ISOCHRONE_MTS, incorporates this and other improvements over the original. To correct the few remaining problem cells, the code includes a section that checks the final channel TOC cell values, replacing any anomalously high or low values encountered with those of the neighboring upstream and/or downstream cells.

Figure 7.21 illustrates the success of the new code at eliminating inconsistencies in the channel TOC data for the same area of the watershed shown in Figure 7.20. Figure 7.22 presents a comparison of the watershed hydrographs generated using the original and new isochrone codes for cases with 5 and 21 sub-watershed divisions for the August 17, 2002 precipitation event. The hydrographs generated by the code ISOCHRONE_MTS have large peak flow rates and a shorter time to peak than those generated using the original code. For comparison, all plots in Figure 7.22 include the hydrograph developed from the USGS stream gauge data at Little Pine Creek for the August 17, 2002 precipitation event. Unfortunately, the improvements incorporated into the new isochrone code also cause the disagreement between our model's results and those shown by the stream gauge to increase.

The results of this research showed that isolated anomalously high or low TOC values observed in the raster file are inherent in the cell-by-cell approach used to compute channel TOC values in the isochrone code.

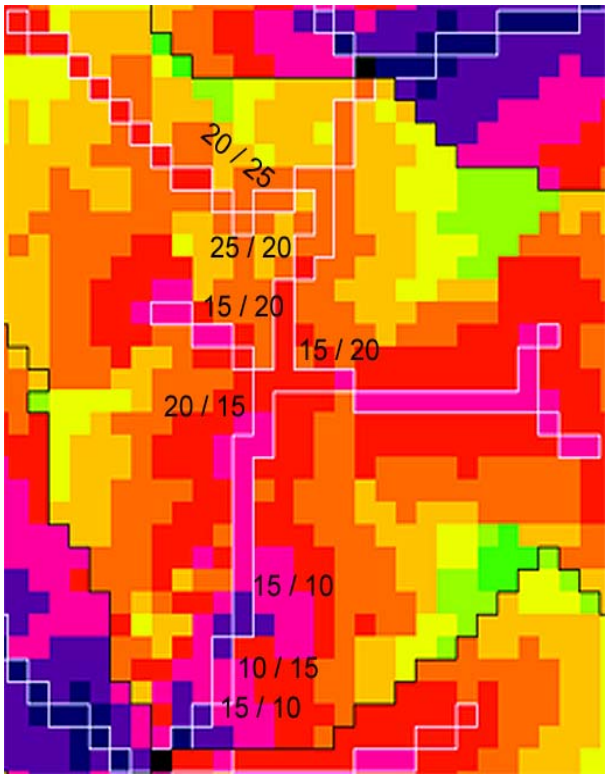


Figure 7.20: Close up view of isochrone raster file generated using original code

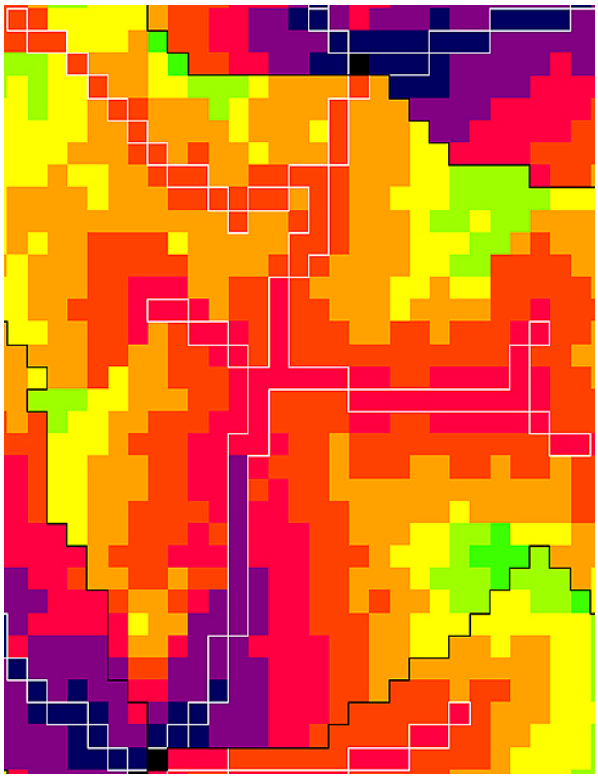
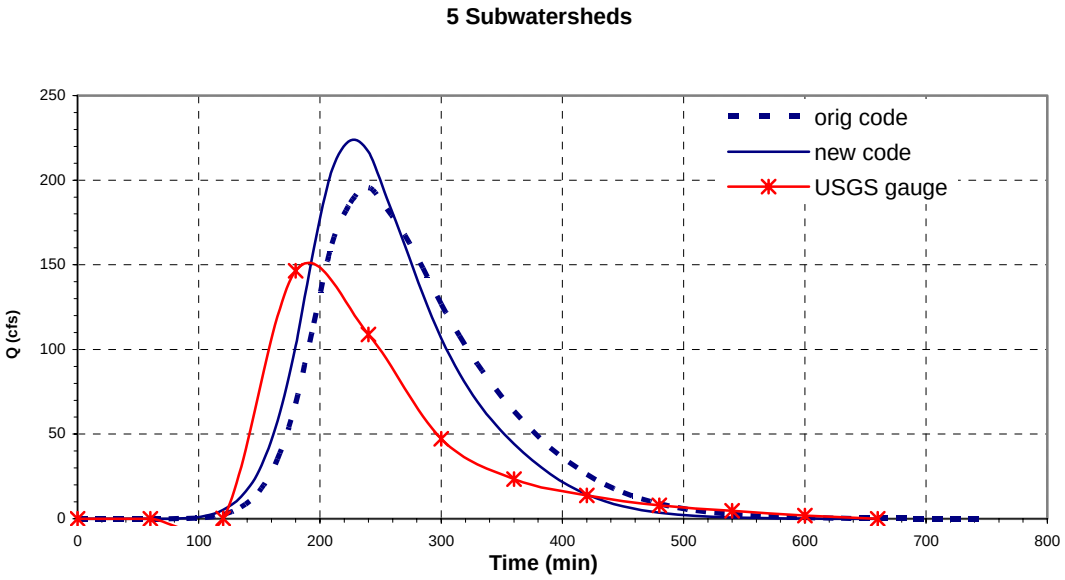


Figure 7.21: Close up view of same area of same file generated using new code



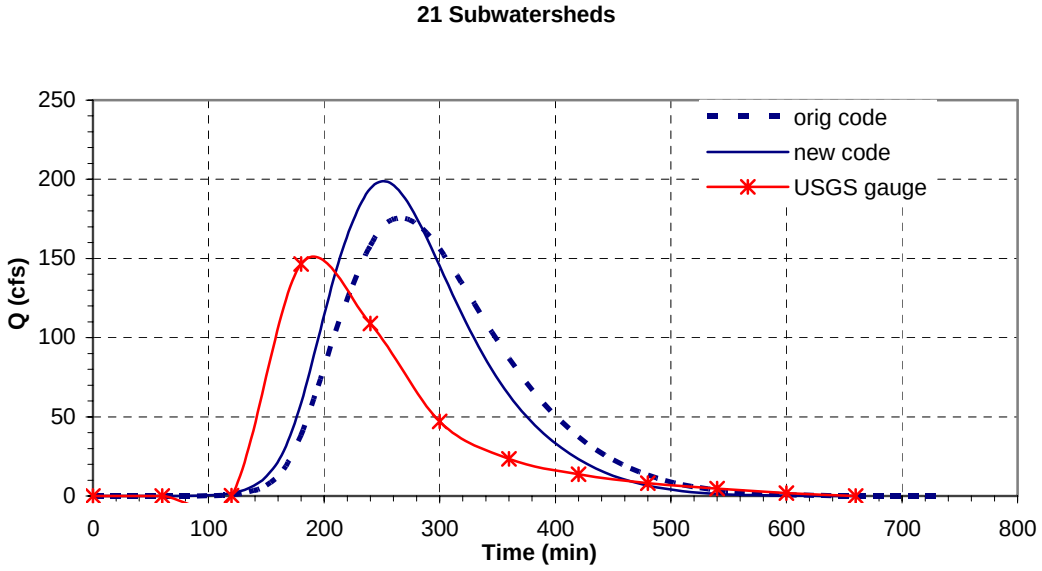


Figure 7.22: Watershed hydrographs generated using original and new isochrone codes

7.2 INCORPORATING KINEMATIC WAVE ROUTING

7.2.1 Overview

We investigate kinematic wave flow routing using the code K-wave_MTS, which solves the Kinematic Wave equations with a linear, implicit, finite difference scheme. The code requires that the user select the number of distance step increments (Δx), and provides the option of selecting the time step (Δt) manually, or computing it automatically based satisfying the Courant condition ($\Delta t < \Delta x/C_k$). Though the Courant condition is only a stability requirement in explicit finite difference routing schemes, it is recommended that implicit finite difference routing schemes also satisfy this condition to achieve the most accurate results. Thus, our study examines cases where Δt is provided by the program. We look first at the effect of varying the assumed Δx increment on the model output; then compare the resulting watershed hydrographs with those generated using Muskingum routing, and that developed from the USGS stream gauge data at Little Pine Creek.

7.2.2 Hydrograph Predicted by Kinematic Wave Routing

Stability is always a concern in any finite difference model. Does the model generate results that make sense for all cases? We address this issue by comparing the total volume of routed runoff predicted at the watershed outlet with the total volume of runoff generated by the code for all sub-watersheds. Table 7.6 presents the runoff volume predicted by the code using 4, 8, 16, and 32 Δx increments, for cases with 5 and 11 sub-watershed divisions, and the percent difference of each from the total runoff input for the watershed area. This comparison of input and output volumes shows that the kinematic wave routing method over-predicts the amount of runoff that reaches the watershed outlet, but not by much. The volume difference is less than 1 % for all cases, which is very minor. Thus, we can conclude that the kinematic wave routing code is stable.

Table 7.6: Comparison of volume differences

Case	No. Δx increments	Hydrograph Volume (ft^3)	% Difference from Total Runoff Volume
Total watershed runoff Aug 17, 2002 event		**1789360**	-
5 subwatershed divisions (8/17/02 event)	4	1803873	0.80%
	8	1797716	0.46%
	16	1796511	0.40%
	32	1793825	0.25%
11 subwatershed divisions (8/17/02 event)	4	1804348	0.83%
	8	1801310	0.66%
	16	1789491	0.01%
	32	1793094	0.21%
Total watershed runoff July 21-22, 2003 event		**2994467**	-
5 subwatershed divisions (7/21-22/03 event)	4	3022600	0.93%
	8	3011816	0.58%
	16	3012997	0.61%
	32	3002943	0.28%
11 subwatershed divisions (7/21-22/03 event)	4	2983335	-0.37%
	8	3003430	0.30%
	16	2996176	0.06%
	32	3003575	0.30%

7.2.2.1 August 17, 2002 Precipitation Event

Figures 7.23 and 7.24 illustrate the peak flow (Q_p) and time to peak (T_p) values of the hydrographs, respectively, generated using 4, 8, 16, and 32 Δx increments, for cases with 5 and 11 sub-watershed divisions. Figure 7.25 presents the hydrographs that result from using 4 and 32 Δx increments for the same cases. Together, these three figures show that the resulting watershed hydrograph Q_p values are very large as compared with those generated using Muskingum routing for the same cases, while the T_p values are quite small. Figure 7.23 reveals that Q_p values increase from 333cfs to 342 cfs, and from 331 cfs to 339 cfs, for the cases with 5, 11 sub-watershed divisions, respectively. Figure 7.25 illustrates how the hydrographs predicted using Kinematic Wave routing become steeper with as the number of Δx increments increases. This trend is particularly obvious in the rising limb. Figure 7.26 illustrates two additional hydrograph trends. First, that the small hump present in the falling limb becomes more pronounced with increasing Δx. Second, that that the hydrograph shape takes on a more choppy appearance with increasing Δx. Figure 7.24 shows that increasing the resolution of the computational grid has no effect on the hydrograph T_p values.

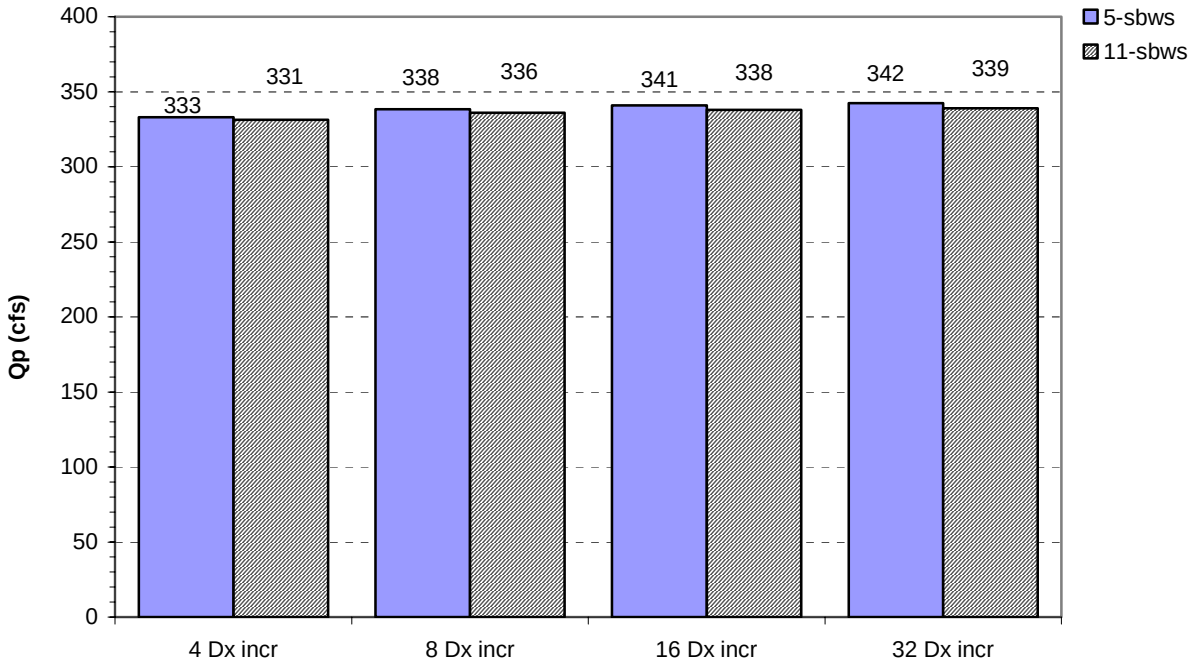


Figure 7.23: Comparison of hydrograph Qp values with increasing Δx increments (8/17/02)

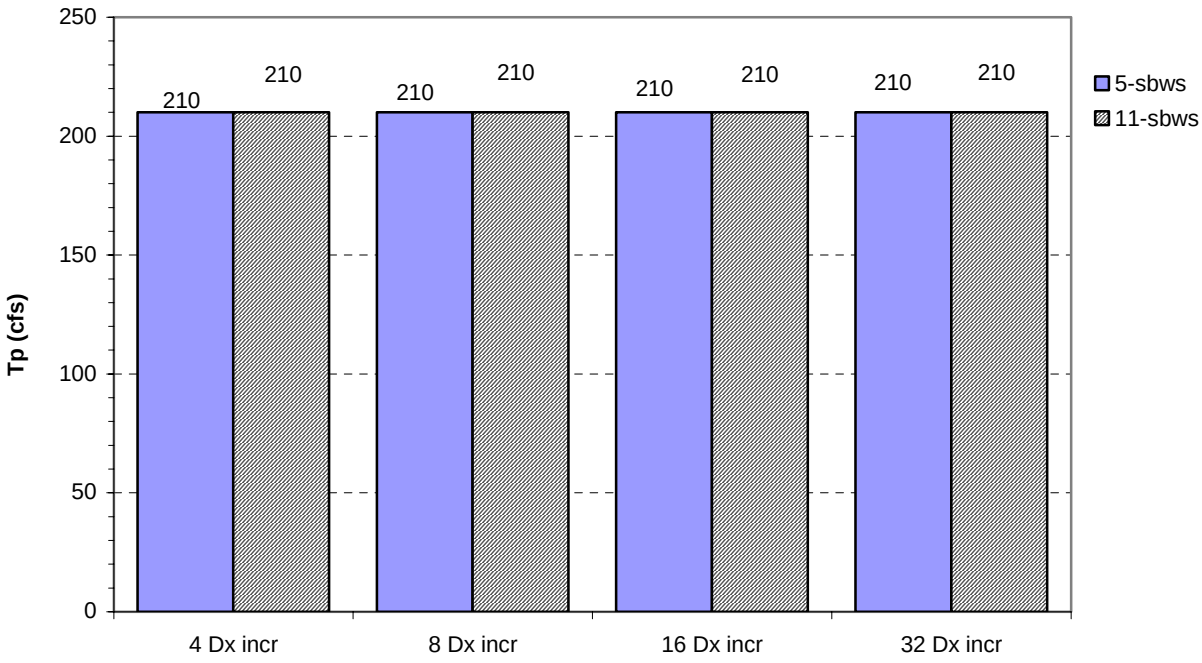


Figure 7.24: Comparison of hydrograph Tp values with increasing Δx increments (8/17/02)

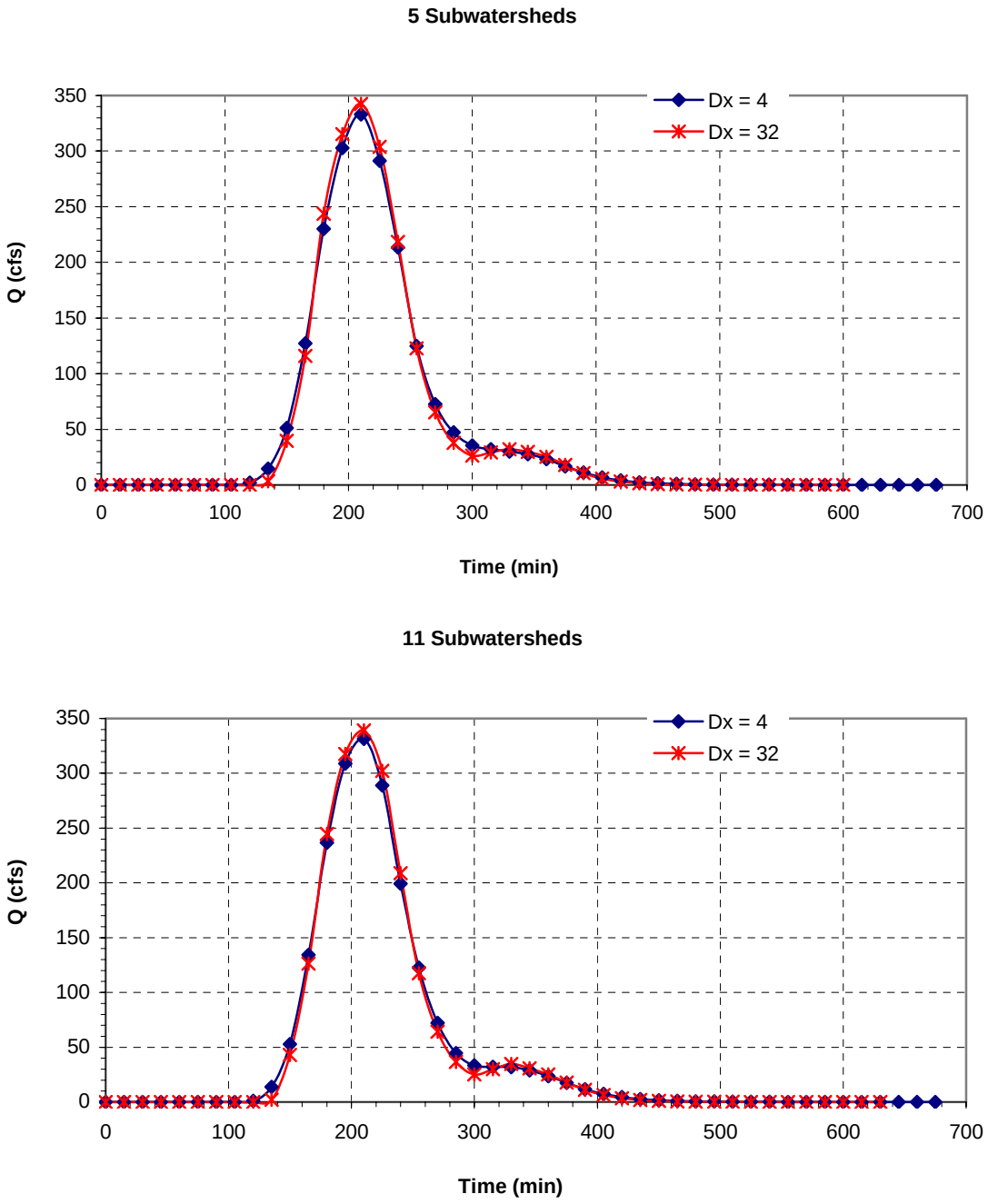


Figure 7.25: Hydrographs for 4 and 32 Δx divisions (8/17/02)

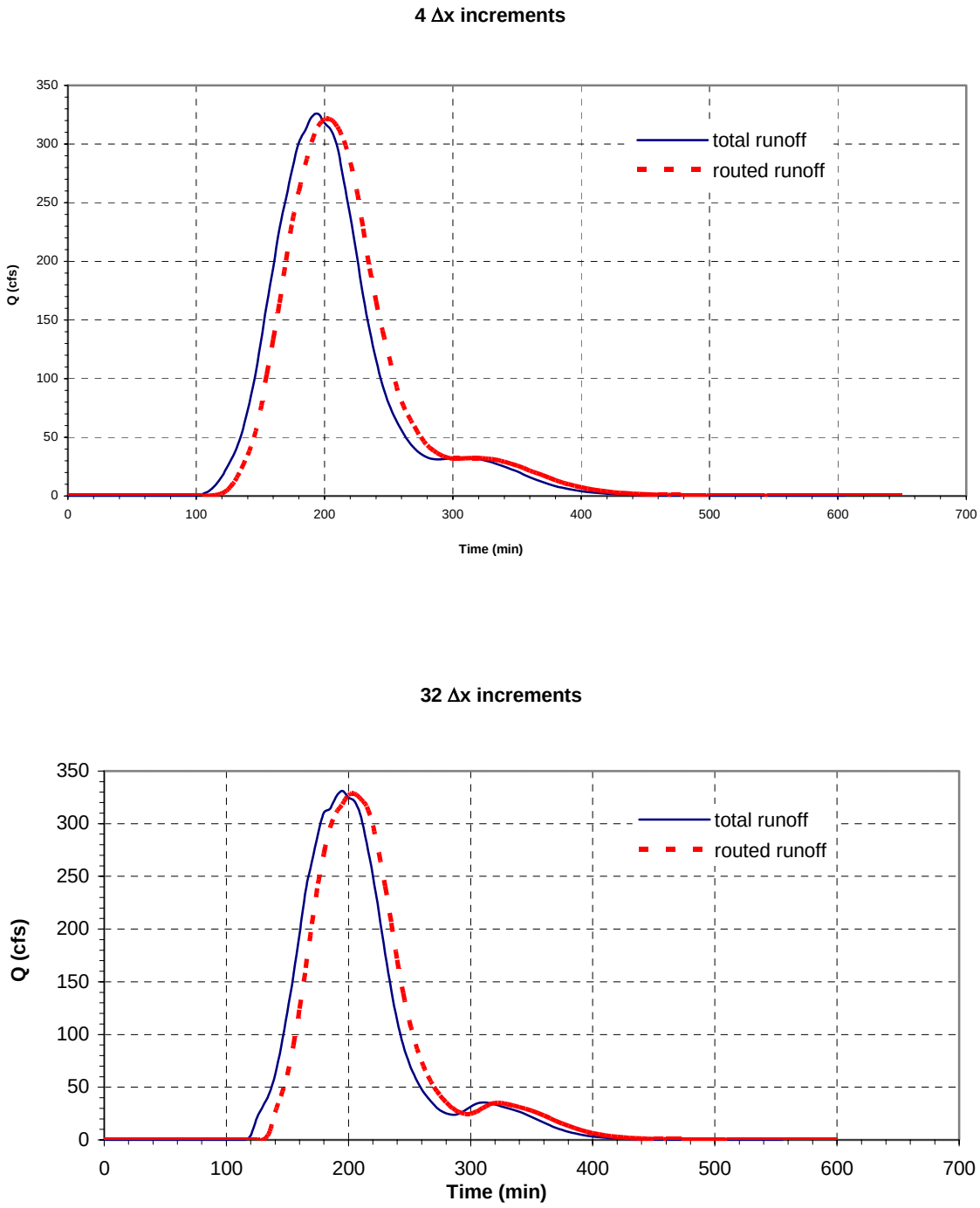


Figure 7.26: Comparison of runoff and routed runoff hydrographs for varying Δx values (8/17/02)

7.2.2.2 July 21-22, 2003 Precipitation Event

Figures 7.27 and 7.28 illustrate the peak flow (Q_p) and time to peak (T_p) values of the hydrographs, respectively, using 4, 8, 16, and 32 Δx increments, for cases with 5 and 11 sub-watershed divisions. Figure 7.29 compares the watershed hydrographs predicted using 4 and 32 Δx increments for the same cases. Similar to the first precipitation event, these figures show that the hydrograph Q_p values predicted using the kinematic wave routing are very large, approximately twice the size of those predicted using Muskingum routing for the same cases. Figure 7.27 reveals that Q_p values increase from 597cfs to 623 cfs (5-sbws case), and from 577 cfs to 603 cfs (11-sbws case), as the number of Δx increments is increased from 4 to 32. In contrast, varying Δx has little effect on the hydrograph T_p values. Figures 7.29, and 7.30 illustrates the same trends of hydrograph steepening and increasing choppiness with decreasing Δx distance noticed in the data for the first precipitation event.

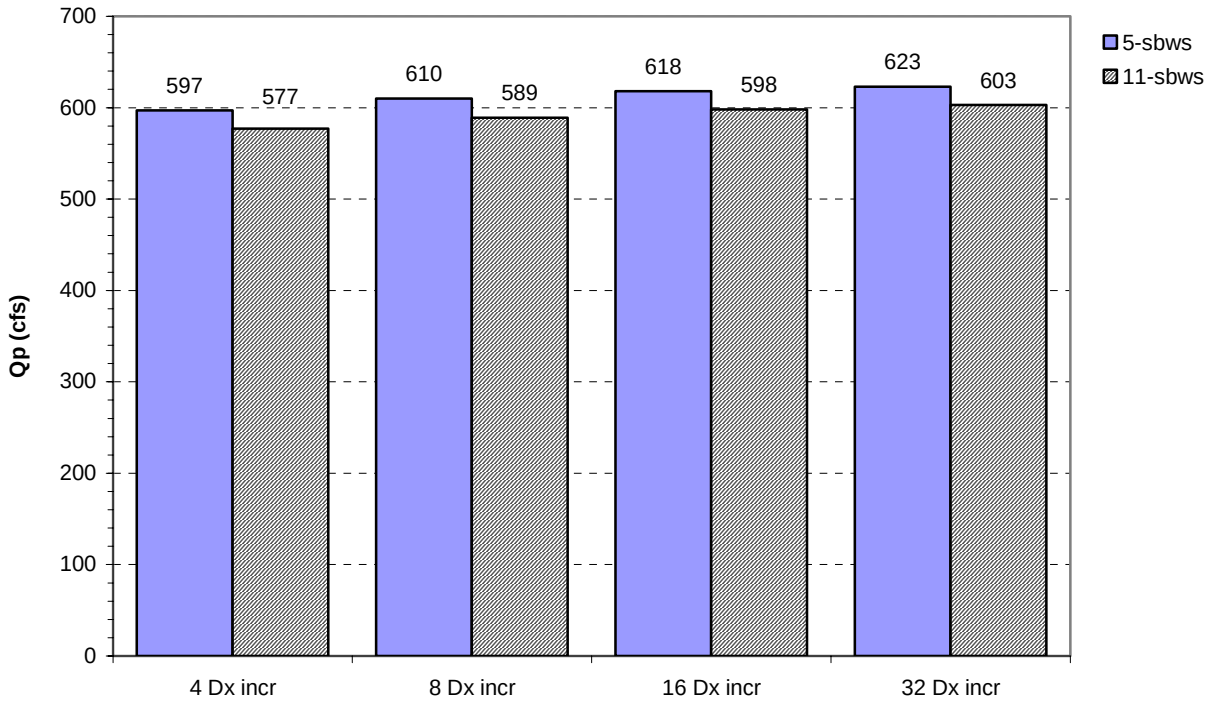


Figure 7.27: Comparison of hydrograph Qp values for varying Δx increments (7/21-22/03)

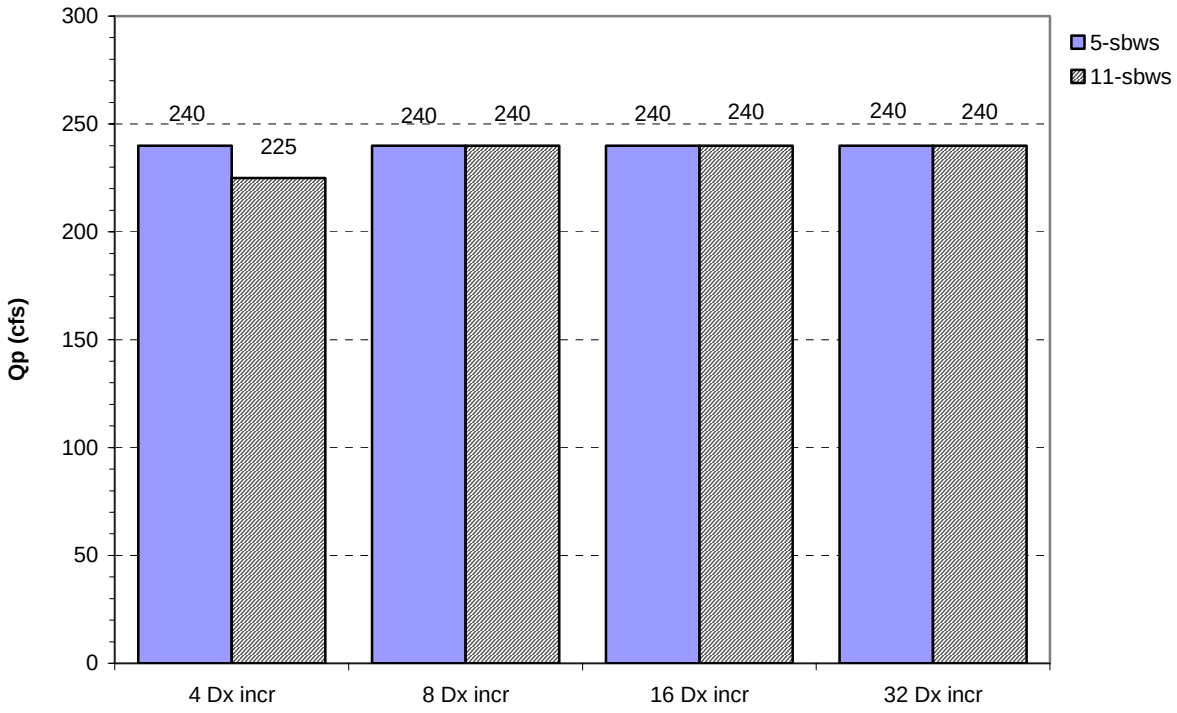


Figure 7.28: Comparison of hydrograph Tp values for varying Δx increments (7/21-22/03)

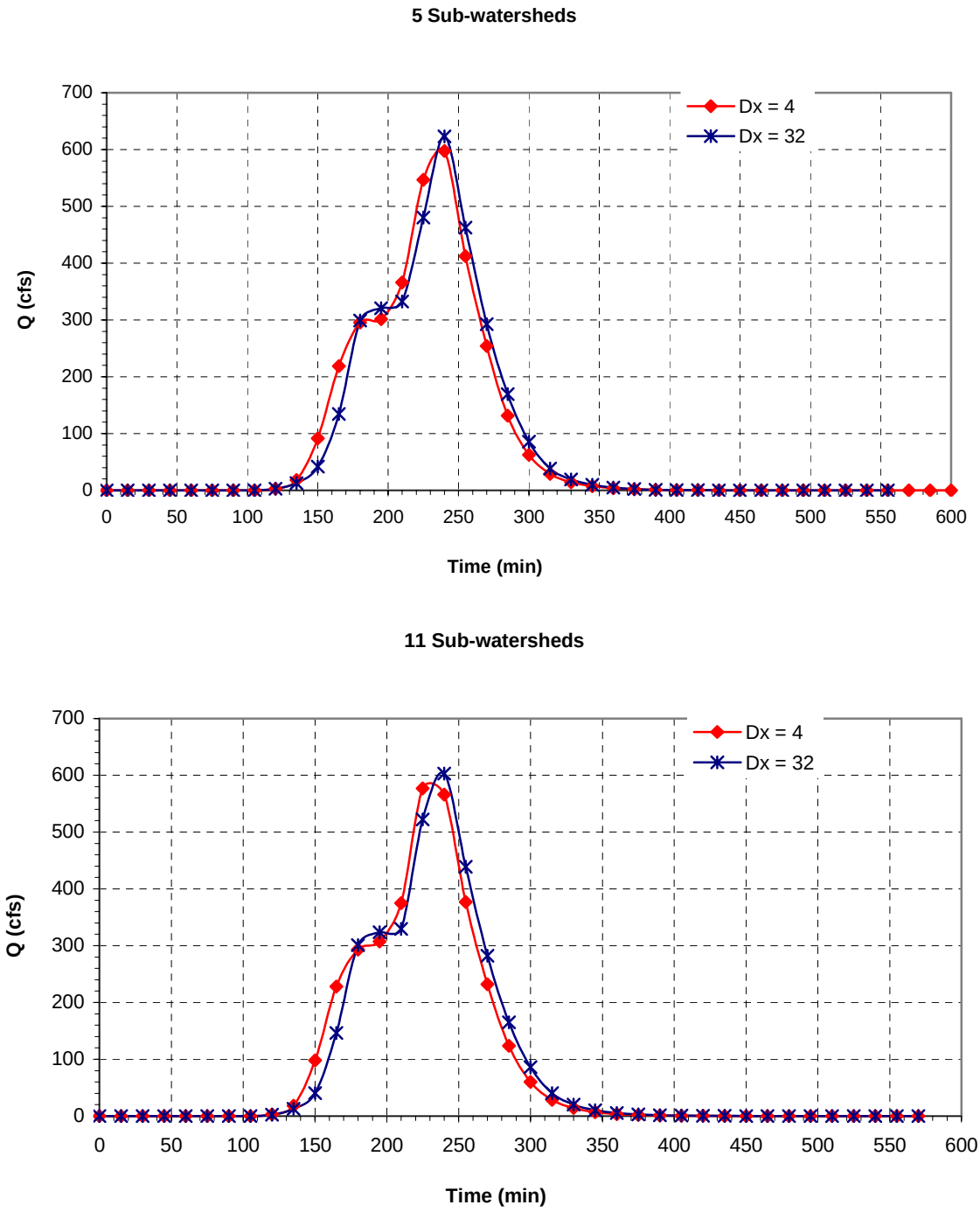
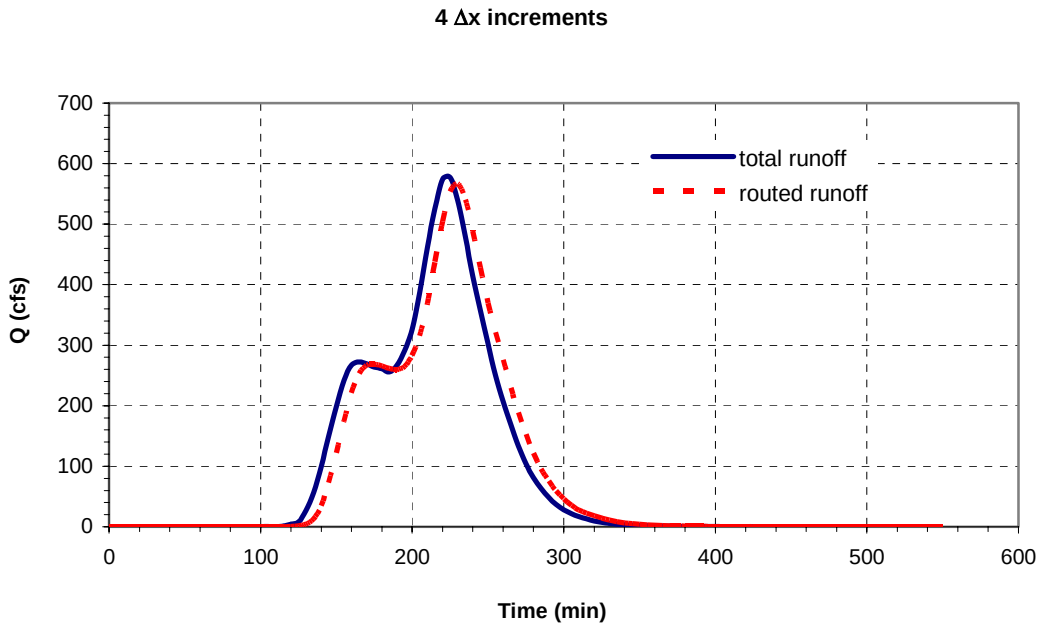


Figure 7.29: Hydrographs for 4 and 32 Δx divisions (7/21-22/03)



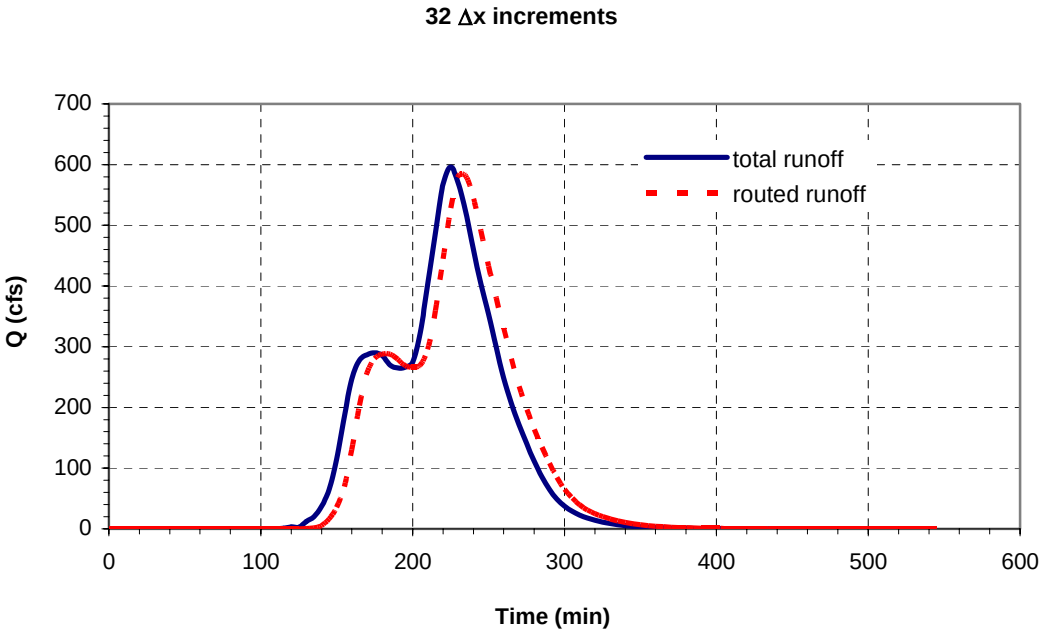


Figure 7.30: Comparison of total and routed runoff hydrographs for 4 and 32 Δx divisions (7/21-22/03)

7.2.2.3 Discussion

This trend of increasing Q_p values with an increasing number of Δx increments shown by our results highlight the most important characteristic of kinematic wave flow routing; it is an almost purely translational process that produces very little dissipation of the hydrograph through either attenuation of the peak, or dispersion of the shape. Any attenuation observed in the hydrograph peak is an artifact of the computational grid resolution. Since the finite difference method solves the partial differential Kinematic Wave equations by approximating the solutions at each node (Δt,Δx) in the computational grid, the accuracy of the numerical solution with respect to the true analytical solution depends on the computational grid selected. As grid becomes finer, resolution increases and the solution approaches that of the analytical solution, which in the case of the Kinematic Wave equation, is pure translation of the hydrograph with no dissipation. Thus, it makes sense that Q_p should increase with the number of Δx increments. Figures 7.26 and 7.30 illustrate the translational effect that kinematic wave routing has on a runoff time series data set, by comparing the total runoff and routed runoff hydrographs of a single sub-watershed area generated using 4 and 32 Δx increments for the case with 11 subdivisions.

The non-diffusive properties of the Kinematic Wave Model can also explain the observation that the hydrographs become increasingly steep and irregular with decreasing Δx distance. Since, as Δx increases the model provides a closer and closer approximation to the analytical solution of pure translation, any alteration of the hydrographs shape during the routing process due to the effects of dissipation will disappear. As a result the shape of the routed hydrograph retains more of the irregularities presents in the runoff time series data from which it was created.

The observation that kinematic wave routing has little effect on hydrograph T_p is expected, since the dissipation produced by decreasing grid spacing functions primarily to expand the hydrograph slightly and reduce its peak, with only a very minor decrease in T_p (Chow, 1988).

7.2.3 Kinematic Wave vs. Muskingum Routing

Figure 7.31 compares the hydrographs predicted by our model using both Kinematic Wave and Muskingum routing methods for the case with 11 sub-watersheds, to that given by the USGS

stream gauge at the outlet of the Little Pine Creek watershed on August 17, 2002. Figure 7.32 provides the exact same comparison for July 21-22, 2003. Figures 7.33 and 7.34 compare the Q_p and T_p values of these hydrographs for the August 17, 2002 and July 21-22, 2003 precipitation events, respectively. It is obvious from looking at these hydrographs that the Muskingum routing method does a much better job than the Kinematic Wave method at approximating the actual peak flow given by the USGS hydrograph. Table 7.7 presents a comparison of hydrograph Q_p, T_p and total volume, and the percentage that the predicted values differ from those shown by the USGS hydrograph for both precipitation events. The data for the August 17, 2002 event show that the model results over-predicts the Q_p value given by the gauge by 25.12 % using Muskingum routing, and 54.13 % with kinematic wave routing. The degree of over-prediction of Q_p by both routing methods is even more extreme for the July 21-22, 2003 event. The reason why the kinematic wave routing method does such a poor job here is because it cannot account for the effects of storage within a channel reach. Since the current model does not consider storage effect in the runoff module, the task must be handled during the routing process. If we were modeling a system where storage was unimportant, this would not be a problem; however, in the case of the Little Pine Creek watershed it clearly is. In contrast to the Kinematic Wave, the Muskingum method is able to simulate the effects storage within a reach by dissipating the input hydrograph according the values of constants K and X, and consequently the Q_p results it produces provide a better approximation of reality in this case, at least for the August 17, 2002 event.

Table 7.7 shows that the T_p value predicted by kinematic wave routing differs from that of the USGS data reported on August 17, 2002 by only 9.52%, which is, in fact, a better approximation than that given by the Muskingum method (20.83% difference). This observation is a bit puzzling, as one would expect either one routing method or the other to produce superior results for both Q_p and T_p. The T_p values predicted by the model for the July 21-22, 2003 event do not show this same trend, and in fact, both routing methods predict T_p = 255 min, which is identical that given by the USGS data.

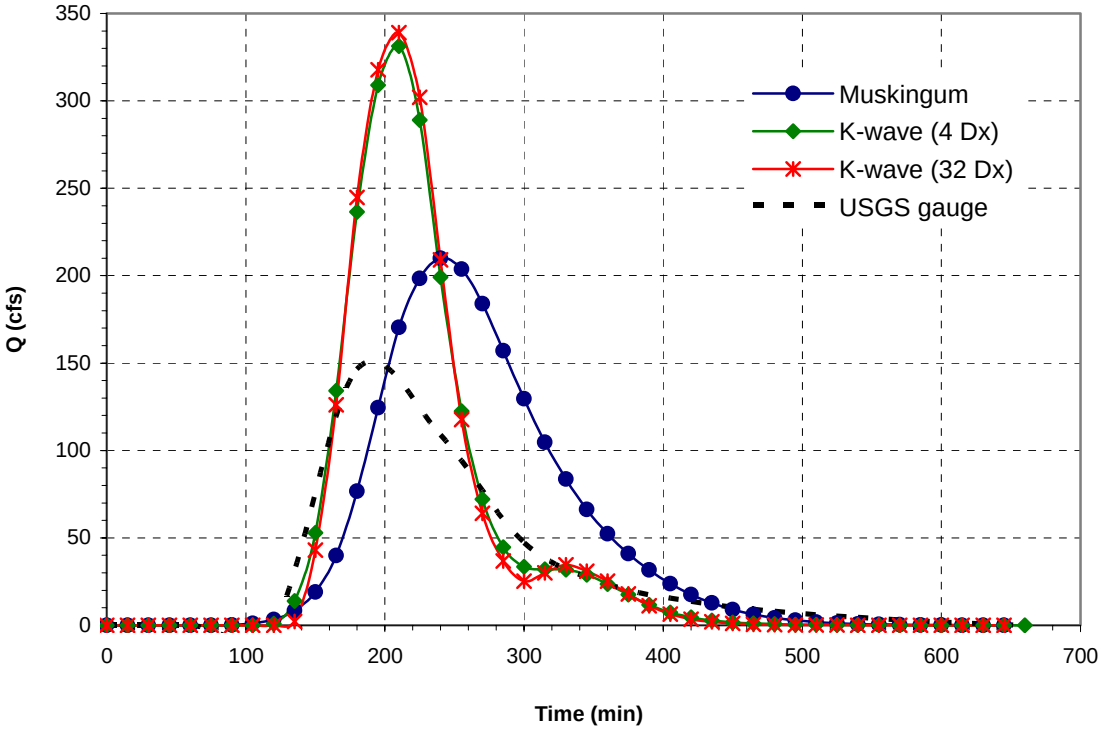


Figure 7.31: Comparison of Kinematic Wave, Muskingum and USGS hydrographs (8/17/02)

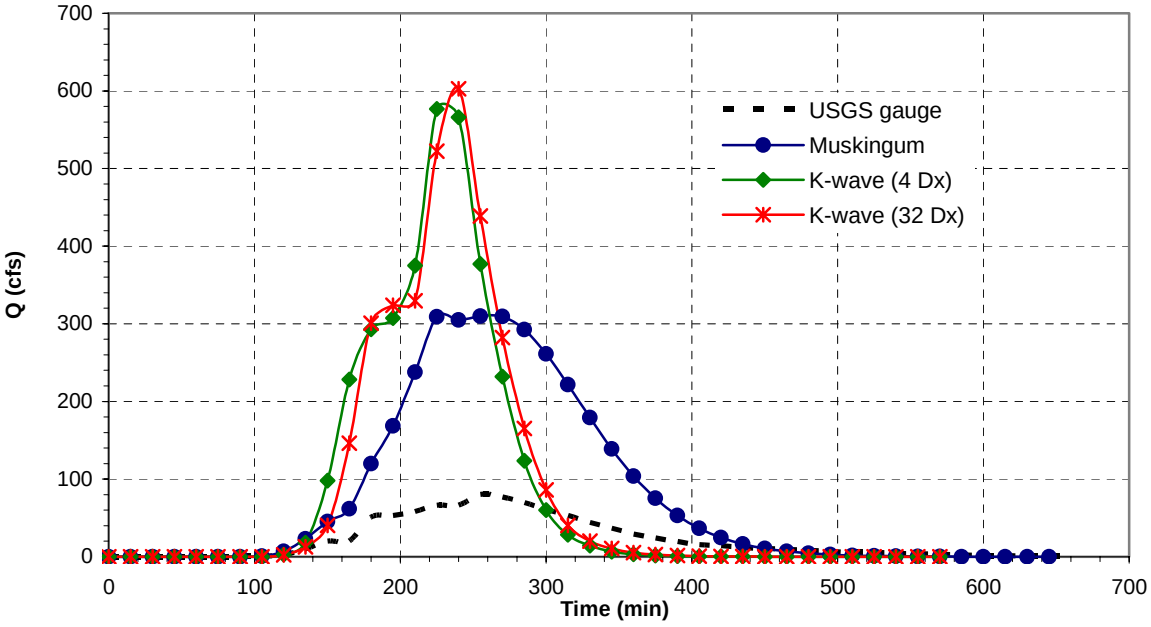


Figure 7.32: Comparison of Kinematic Wave, Muskingum and USGS hydrographs (7/21-22/03)

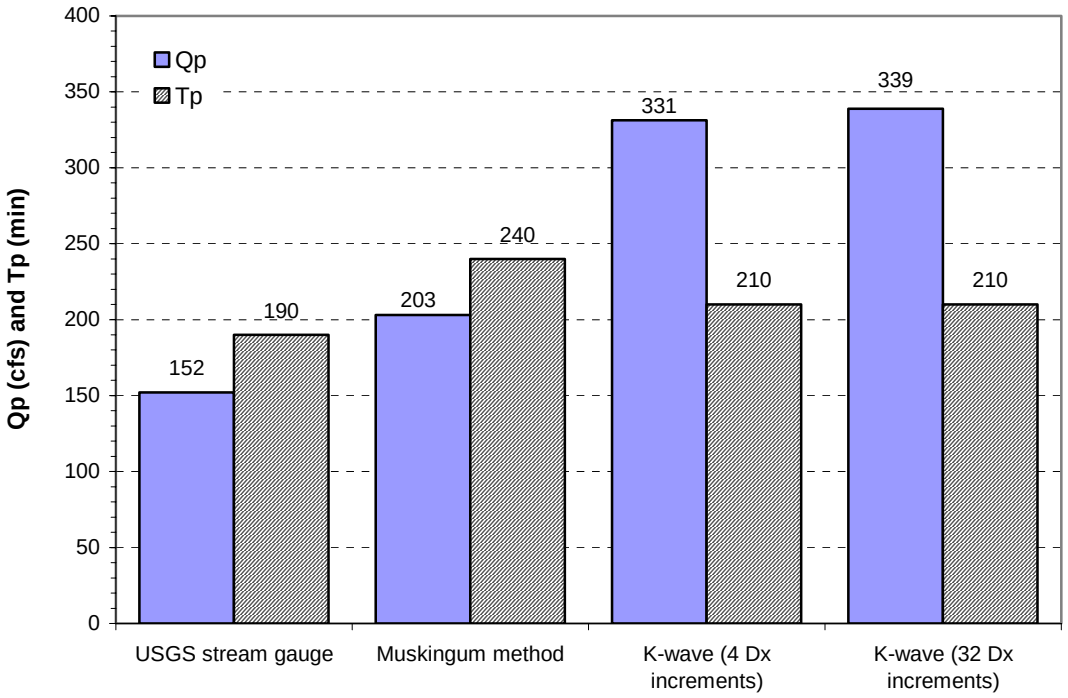


Figure 7.33: Comparison of hydrograph Qp and Tp values (8/17/02)

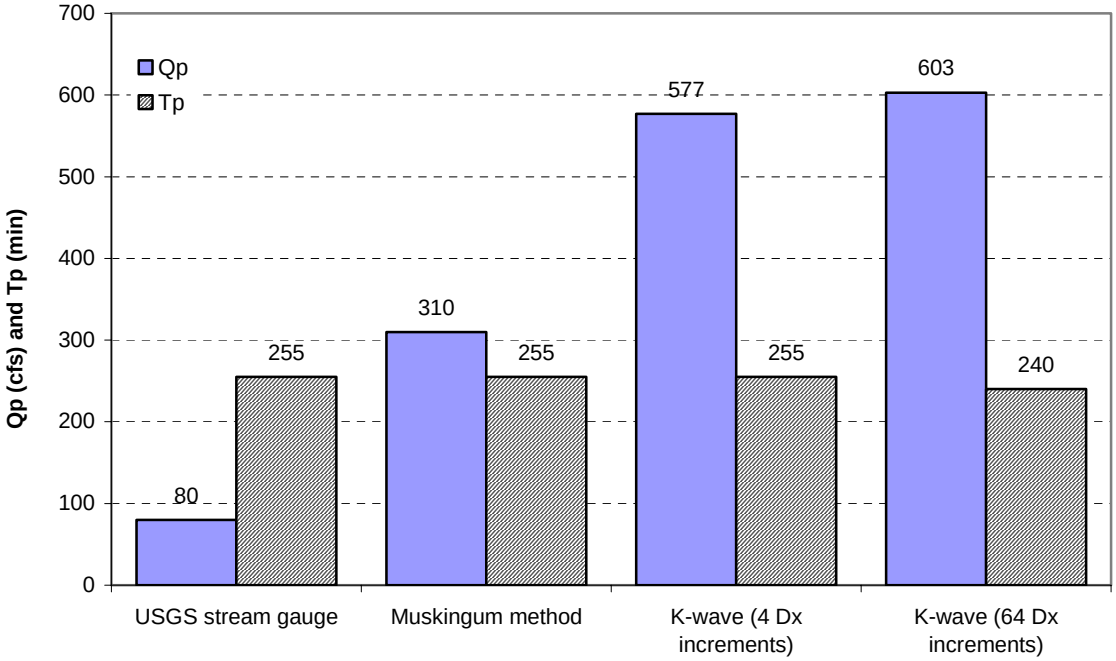


Figure 7.34: Comparison of hydrograph Qp and Tp values (7/21-22/03)

Table 7.7: Comparison of hydrograph Qp and Tp differences

Data set	Q_p		T_p		Total Volume	
	(cfs)	(Δ %)	(min)	(Δ %)	(ft^3)	(Δ %)
August 17, 2002 event						
USGS stream gauge	152	-	190	-	1276007	-
Muskingum method	203	25.12%	240	20.83%	1789231	28.68%
K-wave (4 Δx increments)	331	54.13%	210	9.52%	1804348	29.28%
K-wave (32 Δx increments)	339	55.16%	210	9.52%	1791135	28.76%
July 21-22, 2003 event						
USGS stream gauge	80	-	255	-	856600	-
Muskingum method	310	74.19%	255	0.00%	2996873	71.42%
K-wave (4 Δx increments)	577	86.14%	255	0.00%	2983335	71.29%
K-wave (32 Δx increments)	603	86.73%	240	-5.88%	3003575	71.48%

7.3 CURVE NUMBER MODEL CALIBRATION

In the last section we stated that the Muskingum routing method produced a much better approximation of hydrograph peak flow than the Kinematic Wave method. Actually, this statement is only partly true. While the degree to which the model results generated using Muskingum routing over-predict the gauge values is definitely less than those generate using kinematic wave routing, the Muskingum Q_p results are still substantially off the mark. As shown in Table 7.7 the hydrograph Q_p value predicted using Muskingum routing is 25% larger than the USGS value on August 17, 2002, and an alarming 74% larger for July 21-22, 2003. While this caught our attention, the true problem indicator is the disagreement between the volume of total runoff described by the USGS hydrograph, and the total volume of precipitation excess computed by the model. Table 7.7 shows that the model over-predicts the volume of direct runoff given by the USGS hydrograph data by 29% for the August 17, 2002 event, and 71% for the July 21-22, 2003 event. Such a large discrepancy between predicted and actual runoff volumes represents a blatant violation of conservation principles, and is a clear indication that our model is not well calibrated in regards to its ability to reproduce the actual volume of direct runoff reported at the watershed outlet, particularly for the July 21-22, 2003 event. Since runoff generation in the model is controlled by the SCS curve number, we decided that this was the likely cause of the problem. The SCS Method for Abstractions has a major weakness in that it only allows for three Antecedent Soil Moisture Conditions, dry AMC(I), normal AMC(II), and

wet AMC(III). Since it is very unlikely that the actual soil moisture conditions in a watershed at the time of interest will be identical to one of these three, the method is not likely to provide accurate excess precipitation depth predictions across the watershed in most situations.

The SCS method accounts for dry and wet soil moisture conditions by adjusting the curve number accordingly. By the same logic we reasoned that we might be able to calibrate our model results for runoff volume to those given by the USGS stream gauge at the Little Pine Creek watershed outlet by adjusting the curve number in the Little Pine Creek watershed using a weighted average of the curve number for AMC(I) and AMC(II) conditions. We investigated this idea by computing a series of adjusted curve number raster files for the Little Pine Creek watershed using equation 7-2 for both the August 17, 2002 and July 21-22, 2003 precipitation events. We were able to generate the required curve number raster files with relative ease using the program IDRSI, which contains a math module that allows a user to modify some or all of the cells within a raster file according to an equation.

$$CN_{mod} = (wt)(\text{CN for AMC(II)}) + (1\text{-}wt)(\text{CN for AMC(I)}) \qquad (7\text{-}2)$$

For each case, we decreased the weighting factor (wt) from 1.0, by increments of 0.1; then used this new curve number raster file to generate runoff across the watershed, and compared the total volume of runoff with the value given by the UGSG hydrograph. Table 7.8 shows the versions of equation 7-2 used during this investigation, the corresponding total volume of precipitation excess predicted by the model for each, and the percentage difference between the predicted value and the gauge value for both precipitation events.

Table 7.8: Comparison of excess precipitation volume generated for various curve number weightings

Case	Composite CN formula	Total runoff volume (ft^3)	% Difference from gauge volume
August 17, 2002 Precipitation Event			
USGS gauge data	-	1276007	-
Model results			
AMC(II) Curve No.	-	1789360	28.69%
Composite Curve No. 1	CN = (0.9)(CN2) + (0.1)(CN1)	1473743	13.42%
Composite Curve No. 2	CN = (0.8)(CN2) + (0.2)(CN1)	1178893	-7.61%
Composite Curve No. 2	CN = (0.85)(CN2) + (0.15)(CN1)	1317312	3.14%
Composite Curve No. 3	CN = (0.84)(CN2) + (0.16)(CN1)	1308166	2.46%
Composite Curve No. 4	CN = (0.83)(CN2) + (0.17)(CN1)	1307126	2.38%
Composite Curve No. 5	CN = (0.82)(CN2) + (0.18)(CN1)	**1266384**	**-0.75%**
July 21-22, 2003 Precipitation Event			
USGS gauge data	-	856600	-
Model results			
AMC(II) Curve No.	-	2994467	71.39%
Composite Curve No. 1	CN = (0.9)(CN2) + (0.1)(CN1)	2491876	65.62%
Composite Curve No. 2	CN = (0.8)(CN2) + (0.2)(CN1)	2016140	57.51%
Composite Curve No. 3	CN = (0.7)(CN2) + (0.3)(CN1)	1655471	48.26%
Composite Curve No. 4	CN = (0.6)(CN2) + (0.4)(CN1)	1399556	38.79%
Composite Curve No. 5	CN = (0.5)(CN2) + (0.5)(CN1)	1116527	23.28%
Composite Curve No.6	CN = (0.4)(CN2) + (0.6)(CN1)	893330	4.11%
Composite Curve No. 7	CN = (0.35)(CN2) + (0.65)(CN1)	771622	-9.92%
Composite Curve No. 8	CN = (0.375)(CN2) + (0.625)(CN1)	793567	-7.36%
Composite Curve No. 9	CN = (0.39)(CN2) + (0.61)(CN1)	885831	3.30%
Composite Curve No. 10	CN = (0.385)(CN2) + (0.615)(CN1)	**850965**	**-0.66%**

Table 7.8 shows that the volume of excess precipitation generated by the model decreases with the weight assigned to the curve number value corresponding to AMC(II) conditions, as we expected it would. By increasing the precision of our weighting, we were able to get the model volume to within 0.75% of that given by the USGS stream gauge for the August 17, 2002 precipitation event, and within 0.66% of the gauge volume for the July 21-22, 2003 event. Figures 7.35 and 7.36 compare the hydrographs predicted by our model using the adjusted curve number for the case with 11 sub-watershed divisions, with those given by the USGS stream gauge on August 17, 2002, and July 21-22, 2003, respectively. To study the effect that increasing the number of sub-watershed divisions has on the hydrographs predicted by the calibrated model, we repeated part of the sub-watershed division study using cases with 11, 15, 21 and 25 divisions. The hydrograph Q_p and T_p values for these 4 cases and those given by the

USGS gauge data for August 17, 2002 are presented in Figure 7.37. Figure 7.38 shows the same data for the July 21-22, 2003 event. Table 7.9 shows the percentage difference between the Q_p and T_p values predicted by the calibrated model and those given by the USGS gauge for both events.

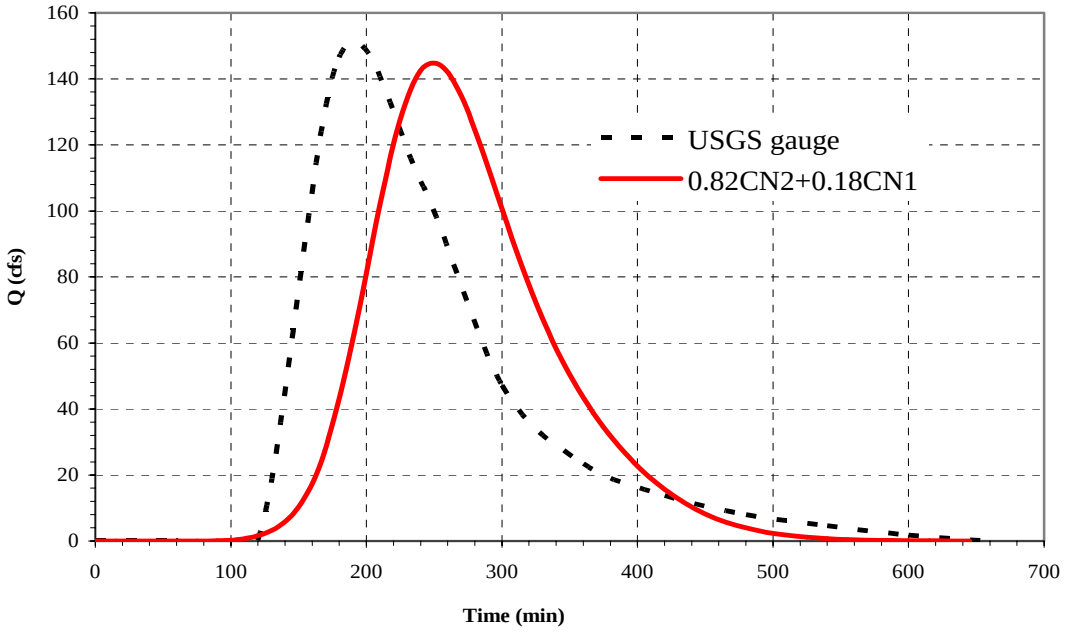


Figure 7.35: Comparison model and USGS hydrographs (8/17/02)

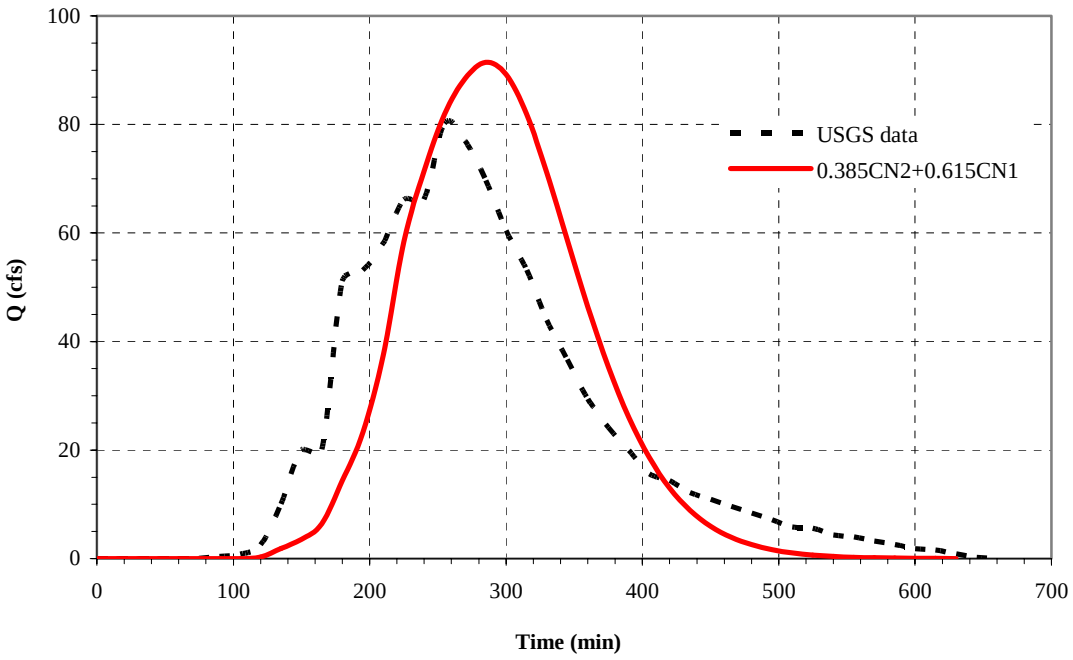


Figure 7.36: Comparison of model and USGS hydrographs (7/21-22/03)

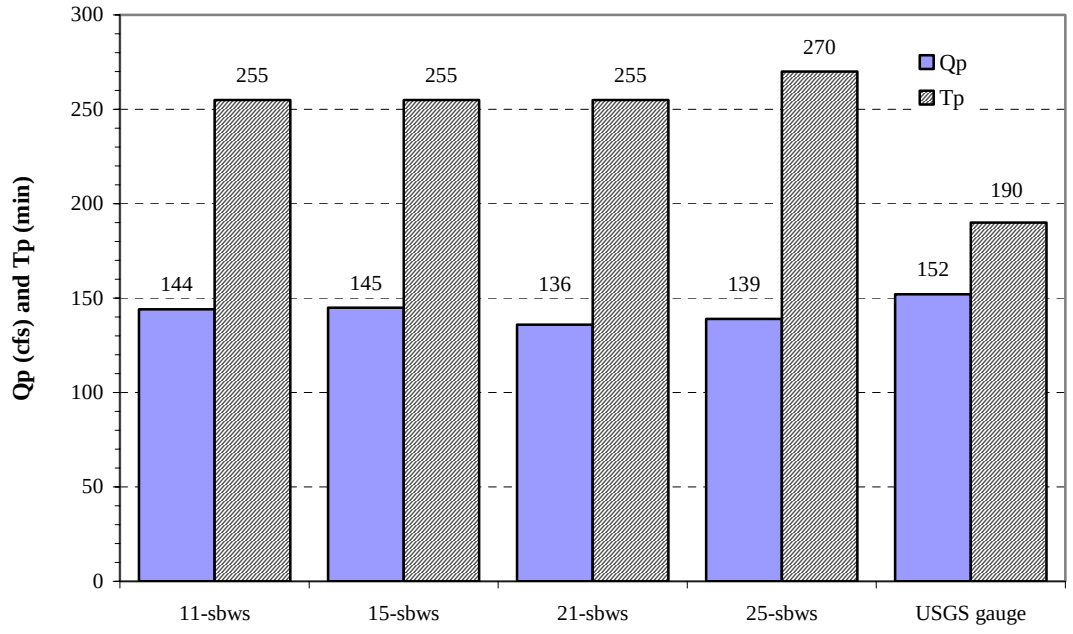


Figure 7.37: Comparison of model and USGS hydrograph Qp and Tp data (8/17/02)

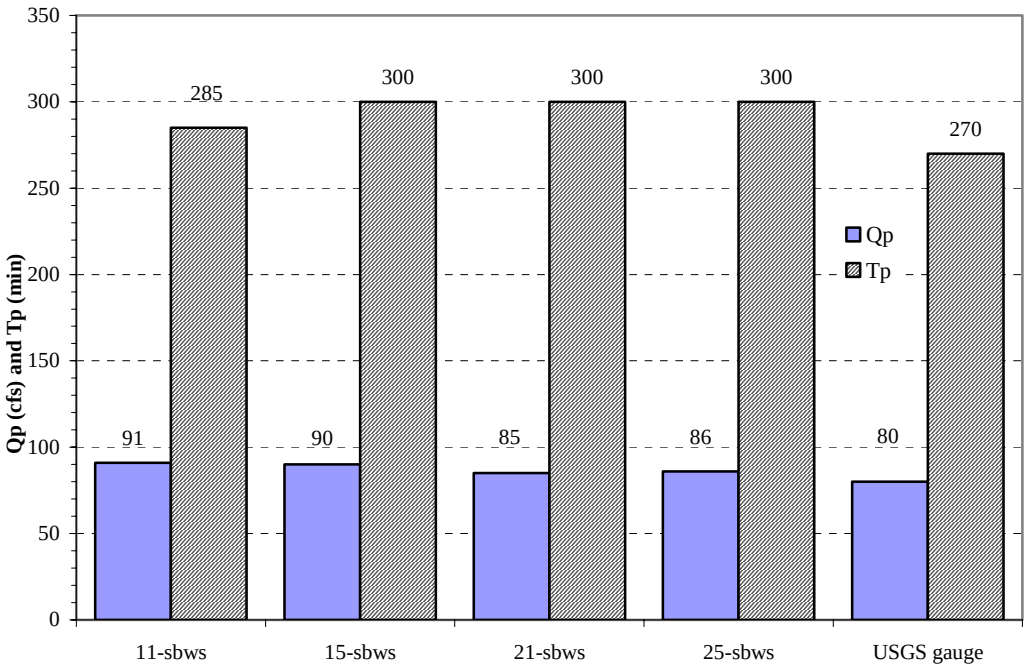


Figure 7.38: Comparison of model and USGS hydrograph Qp and Tp data (7/21-22/03)

Table 7.9: Percentage difference between predicted and gauge Qp and Tp values

Case	Q_p (cfs)	% Q_p difference model and USGS data	T_p (min)	% T_p difference model and USGS data
August 17, 2002 Precipitation Event				
USGS gauge	152	-	190	-
AMC(II) Curve No.	210	27.62%	240	20.83%
Calibrated Curve No. CN = (0.82)(CN2) + (0.18)(CN1)				
11-sbws	144	-5.26%	255	25.49%
15-sbws	145	-4.61%	255	25.49%
21-sbws	136	-10.53%	255	25.49%
25-sbws	139	-8.55%	270	29.63%
July 21-22, 2003 Precipitation Event				
USGS gauge	80	-	255	-
AMC(II) Curve No.	310	74.19%	255	0.00%
Calibrated Curve No. CN = (0385)(CN2) + (0.615)(CN1)				
11-sbws	91	12.09%	285	10.53%
15-sbws	90	11.11%	300	15.00%
21-sbws	85	5.88%	300	15.00%
25-sbws	86	6.98%	300	15.00%

Table 7.9 shows that the calibrated model still over-predicts the hydrograph peak flow values given by the USGS stream gauge at the watershed outlet for the July 21-22, 2003 event; but this difference has been reduced from 74% to 6% – 12%. In contrast, the calibrated model now under-predicts the gauge Q_p results for August 17, 2002 by 5% - 11%, down from a 28% over-prediction using AMC(II) conditions. The calibrated model does not fair quite as well with respect to hydrograph T_p predictions. In this case the values predicted using AMC(II) conditions provide a closer approximation to the T_p values given by the stream gauge data. Table 7.9 shows that the calibrated model performs worse for the August 17, 2002 event, over-predicting T_p by 25% - 30%, than the July 21-22, 2003 event (11% - 15%). This discrepancy between predicted and actual hydrograph T_p values is, however, relatively minor compared to the increase in agreement between Q_p values.

These results show the success of our method of calibrating the model with respect to runoff volume by adjusting the curve number to obtain an accurate representation of the soil moisture conditions present in the watershed at a particular period in time. In different terms this method is simply a way of representing soil moisture conditions that fall between the three identified by the SCS Method for Abstractions, and can be considered an advance upon the soil moisture accounting process of this method in regards to modeling the rainfall-runoff relationship of a watershed.

8.0 CONCLUSIONS

8.1 TESTING AND REFINING THE EXISTING CODE

8.1.1 Modifications to Existing Codes

The process of testing aspects of the existing model revealed two significant problems. First, the raster file used to generate the incremental precipitation files did not include the extreme northern part of the Little Pine Creek watershed. The problem was corrected by replacing this file with one created using the true Little Pine Creek watershed outline. Comparing results of hydrograph peak flow rate (Q_p) and time to peak (T_p) for both cases show that these values increase by 14.19% and 5.56%, respectively for the new file. This obviates the impact that a relatively minor increase in area can have on the volume of runoff generated in a watershed, and highlights why it is important to always try to obtain the most accurate data available when constructing a watershed model.

The second problem involved the method by which the original code RUNOFF computed averages when constructing the 15-minute incremental runoff time series required to match the increments of the precipitation input. The original method, considered only the first and last of each set of five minute incremental runoff values to compute the 15-minute average, which created large differences in runoff volume generated by the model for a given area depending on the way in which it was subdivided. The averaging problem was resolved in a new version of the code RUNOFF_MTS, which produced more consistent results.

8.1.2 Increasing the Number of Watershed Subdivisions

This phase of the study examined the effects of sub-dividing the watershed on the watershed hydrograph for two separate precipitation events. The results show that increasing the number of

sub-watershed divisions used to represent the whole from 5 to 25, produces an overall decrease in hydrograph Q_p values; but that this trend becomes far less significant when the number of subdivisions exceeds 11. Subdividing the watershed has an even slighter effect on hydrograph T_p values, which increase initially as the number subdivisions increases from 5 to 7 subdivisions then remain essentially constant. These results indicate that, for Little Pine Creek watershed at least, the model provides the best results when the area of the sub-watersheds is kept at or below a range of 412 to 337 acres, which corresponds to 9 to 11 sub-watershed divisions.

8.1.3 Examining the Effects Produced by Varying Muskingum X

Here the model results for the August 17, 2002 and July 21-22, 2003 precipitation events show a maximum increase in Q_p values between X = 0 and X = 0.2 of 13.41% and 7.85%, respectively. The results also show that T_p remains essentially constant for all X values for a given sub-watershed arrangement. Thus, we can conclude that increasing X has the expected effect of increasing the hydrograph peak flow, but the effect is not particularly severe.

8.1.4 Effect of Decreasing the Channel Flow Threshold

The results of this study objective show that lowering the threshold number of flow-contributing cells required to define channel flow from 100 to 20 decreased the hydrograph Q_p values by less than 1% for both precipitation events considered, and had no effect on the time to peak. Thus we conclude that our assumption that a 22-acre area of land is required to produce one stream channel is acceptable for the Little Pine Creek watershed.

8.1.5 Resolving Problems in the Isochrone Code

Examining the isochrone code's structure revealed that problem of anomalously increasing and decreasing time of concentration (TOC) values in stream channels is due to a fundamental flaw in the method we use to compute these values, or, more specifically, the values for channel slope used in the computation. Presently, the code computes an average value for channel slope by summing the incremental slopes defined as the difference between every two consecutive cells

along the flow path between the cell in question and the sub-watershed outlet, and dividing this sum by the total number of cells that make up the flow path. While this method works well for reaches with gradual transitions in slope, it breaks down at discontinuities such as a sudden elevation drop between two cells. In light of our lack of success resolving this problem, we conclude that the existence of anomalous TOC values in stream channels is unavoidable under the present method used to compute these values. We managed to get around this problem in the new isochrone code, ISOSTYLE_MTS, by designing the code to limit the instances of problem channel TOC values to a minimum, and including a routine to identify and replace any remaining such values.

8.2 INCORPORATING KINEMATIC WAVE ROUTING

This first part of this objective was accomplished by creating the code K-Wave_MTS, designed to route direct runoff hydrographs from the outlet of one sub-watershed through the reach and to the outlet of the next downstream sub-watershed. The results of changing number of distance increments defining the computational grid in the code on the watershed hydrograph revealed that Q_p increases with the number of distance (Δx) increments. We also observed that the hydrographs themselves become more irregular or choppy in appearance as the number of Δx increments increases. The observations are explained by considering that the kinematic wave routing method is a purely translational process. If solved analytically the kinematic wave solution will simply translate the hydrograph without affecting its shape by diffusion or dissipation at all. Any hydrograph attenuation generated by a numerical kinematic wave routing scheme is purely an artifact of increasing the size of the computational steps Δx and Δt. Thus, it makes sense that Q_p increases and that the hydrographs become more irregular in shape as the number of Δx increments is increased.

Comparing the watershed hydrographs predicted by the model using kinematic wave and Muskingum routing with those given by the USGS stream gauge for two precipitation events, shows that the kinematic wave can provide a good approximation of T_p, but grossly over-predicts the Q_p values reported by the stream gauge at Little Pine Creek. These results lead to the rather unfortunate conclusion that the kinematic wave flow routing method does not work well with our

existing model, and should no longer be considered as a flow routing option. Any routing method employed in this model must be able to describe the effects of storage within the reach on the hydrograph shape, since this is not accounted for during the runoff generation process. The Muskingum routing method, developed during the previous research, does a much better job at simulating that actual watershed response in this case, and should remain the preferred flow routing method for now.

8.3 CALIBRATING THE MODEL BY ADJUSTING CURVE NUMBERS

In our efforts to explain the discrepancy between the volume of excess precipitation computed by our model and that reported by the USGS stream gauge at Little Pine Creek during the July 21-22, 2003 precipitation event, we created a new method for calibrating the model by modifying the curve number. In this method, the model is calibrated with respect to runoff volume by adjusting the curve number value to obtain an accurate representation of the soil moisture conditions present in the watershed. Such calibration of the volume of excess precipitation predicted by the model with the volume of direct runoff given by the USGS stream gauge hydrograph is essential, because only by achieving this agreement can we ensure that our model satisfies conservation principles. If we assume that the direct runoff hydrograph shown by the stream gauge data is an accurate representation of reality, and we must assume this because we have no other means of ground-truthing our model results; then by continuity, the model must generate the same volume of excess precipitation as the direct runoff that passes through the watershed outlet.

We created modified curve numbers for every cell in the watershed by summing weighted AMC(I) and AMC(II) curve numbers values. By varying the weights assigned to each curve number and comparing the volume of excess precipitation generated by the model for each corresponding modified curve number, we were able to reduce the volume discrepancy between the model and the gauge results to less than 1%. Thus, we conclude that this method is an excellent way of ensuring conservation of mass in terms of the volume flow rate of direct runoff between the results predicted by our model and those that occur naturally in the Little Pine Creek watershed. Additionally, this new method offers a significant advance over the three-value soil

moisture accounting system used by the SCS Method for Abstractions, providing the modeler greater flexibility in selecting a soil moisture condition that best matches those present in the watershed at the time of interest, improving model results.

9.0 RECOMMENDATIONS

Our results can be used as a basis for several additional topics of future research with this watershed runoff model. Specifically, we recommend the following topics as those we consider most interesting and most beneficial to the model.

9.1 USING ANOTHER CASE STUDY

This recommendation applies to both precipitation events, and watersheds. Research to this point has only examined the response of the Little Pine Creek watershed for two precipitation events. As was obviated by our present study, the watershed's response in terms of runoff generation varies significantly between events. At our present level of knowledge we are unable to describe the variation in runoff generation with precipitation for the watershed at any greater detail than the obvious fact that more rain falling over a longer duration will produce more runoff than less rain falling over a shorter duration. Modeling the watershed's response to many other precipitation events of different characteristics (i.e. long, short, intense, light) might bring to light certain trends in the runoff response behavior of the Little Pine Creek watershed, which would definitely increase the use of our model as a predictive tool, if only for one watershed.

Along these same lines, the current version of our model has only been used to simulate one watershed, Little Pine Creek. The reason for this is simply that no other small, gauged watersheds exist within the calibrated radar coverage area for Allegheny County. The calibrated radar-rainfall data that the Three Rivers Wet Weather Demonstration Project makes available for free on their web site is a unique situation. To our knowledge such data is not available, at least for free, anywhere else in the country. During the coming years however, this is likely to change, as more people realize the usefulness of calibrated radar rainfall data for modeling

purposes. Thus, future researchers should begin looking for other sources of free calibrated radar-rainfall data in other areas of the country and the world, with the goal of finding another small gauged watershed on which to test the model.

9.2 IMPROVING EXISTING COMPUTER PROGRAMS

The present version of our model employs a total of nine separate FORTRAN programs. Most of the smaller programs used to complete the initial data processing, such as CUVRENUM, and DIRECT, accomplish their respective tasks quickly. Those designed to compute and route runoff however, tend to be far more painful to use in regards to time efficiency. During the present research, we modified the structure of several such codes in an effort to make them run more efficiently. For example, the runoff code now permits users to generate the runoff time series for all sub-watershed areas during a single run, as opposed to one for each sub-watershed. The isochrone code now accomplishes its task in seconds instead of several minutes. Increasing the computational speed of a program can be a rather difficult task to accomplish, requiring and extensive knowledge of the programming language in which the code is written; but it is comparatively simple to improve the structure of an existing code so that it operates more efficiently by installing a batching routine for example. Increasing the degree to which a code operates independently from the user and the speed at which it operates can make a world of difference in terms of productivity when running multiple cases. Thus, as a by-product of future research, modelers should always look for ways to improve the efficiency of the existing computer programs. It is important to remember that an improvement that might require a days-worth of work could end up saving far more than a days-worth of time in the long run.

9.3 IMPROVING THE ISOCHRONE CODE

One of the objectives of our research involved studying the problem of anomalously high and low time of concentration (TOC) values present in channels cells. Here, we concluded that this problem can be lessened, but not completely solved, as these anomalous values are inherent to

the present computation method. However, our investigation into the channel TOC problem considered only those alternatives that fit into the structure of the existing isochrone code. The existing code is set up to compute the TOC values for each cell individually by tracing the flow path of water from the cell to the sub-watershed outlet. It appears that in order to resolve this problem, we must consider a completely different approach for computing the TOC values of channel cells. One such new method that holds promise involves going about the computation in reverse, so to speak. The computations should begin at the sub-watershed outlet, and precede cell-by-cell upstream. The TOC value for each consecutive cell along the reach would be computed as the sum of the time required for water to travel between the two consecutive cells and the cumulative travel time to the outlet at that point. This method, though certainly more difficult to program than the previous, would provide a nice solution to the problem, and eliminate the need for routine in the program to replace anomalous TOC values.

9.4 MUSKINGUM-CUNGE ROUTING

Our results showed that the Kinematic Wave routing is not appropriate for our model, as it cannot describe the affects of reach storage, which are manifested in the peak flow attenuation and dispersion observed in the direct runoff hydrograph for the Little Pine Creek watershed. In the interest of continuing to incorporate distributed characteristics into our model, future research should examine other hydraulic routing methods. In particular, we recommend the Muskingum-Cunge Method, a distributed routing model that can describe hydrograph diffusion and dispersion that occur during the routing process. The Muskingum-Cunge Method is basically a combination of the Muskingum and Kinematic Wave Models. In brief, the model considers the Muskingum constants K and X to vary in time and space, and solves the Muskingum equation for flow rate (Q) numerically at each node in a computational grid defined by time and space steps Δt and Δx (Chow, 1988). Muskingum-Cunge offers an advantage over the regular Muskingum Method in that the constants K and X are computed using channel characteristics, such as Q and slope. This would represent an improvement of over the Muskingum Method presently used in our model where a user must assume a value for X at random. By using the constants K and X to describe the effects of reach storage, the Muskingum-Cunge model

provides a physical description of hydrograph diffusion and dispersion, in contrast to the Kinematic Wave model, where these processes are simply artifacts of the computational grid spacing.

9.5 MODEL CALIBRATION BY CURVE NUMBER ADJUSTMENT

During this study we developed an effective method of calibrating the volume of excess precipitation predicted by the model to the volume of direct runoff given by the USGS stream gauge hydrograph by adjusting the watershed curve number values to reflect the soil moisture conditions present at the time of the study. This avenue of research holds considerable potential for improving the accuracy of the model, and should certainly be pursued during future research. One example of a potential improvement would be to find a way to simplify the calibration process. The present method requires that the user compute a series of modified curve numbers, generate the runoff for each case, compare the runoff predicted by the model to that given by the stream gauge data, and then repeat the process until the volumes agree within a certain percent. This can take a while. Thus, it might make sense to automate the calibration process by developing a program to adjust the values of the watershed curve number raster file until the volume of runoff produced by the model agrees with that given by the USGS gauge to within a specified tolerance.

Another interesting idea would involve examining the watershed's rainfall-runoff response over a prolonged period of time rather than for a single event, in an attempt to gain a better understanding of the soil moisture variability within the Little Pine Creek watershed. This type of research could potentially lead to the development of an equation of the form (2-18), which could be used to provide more accurate estimations of the soil moisture conditions present in the watershed during the time of study, and result in more accurate runoff predictions.

9.6 VARYING MANNING'S N PARAMETER IN KINEMATIC WAVE ROUTING

During the presentation of this research, Dr. Chao-Lin Chiu raised a very interesting and valid point concerning the results of our kinematic wave routing investigation. Dr. Chiu pointed out that the value of Manning's n assumed in the kinematic wave computation has a considerable influence on the resulting hydrograph predicted by the model. He suggested that the large discrepancy between the hydrograph peak flow values show by hydrographs predicted by our model using kinematic wave flow routing, and the peak flow values shown by the USGS stream gauge hydrographs might be reduced if we ran the model for a different n value. In light of this suggestion, we recommend that future research look into this idea. What effect does varying the value of n that we assume for the watershed have on the direct runoff hydrograph predicted by our model? Taking this one step further, future investigators could even develop a method by which to calibrate the hydrographs predicted by our model using kinematic wave routing for individual precipitation events by adjusting Manning's n.

BIBLIOGRAPHY

Cited References

Acrement, G. J., and Schneider, V. R., "Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains", United States Geological Survey Water Supply Paper No. 2339 (1984).

Almedeij, Jaber H., "Automatic Generation of Runoff curve Numbers With Geographic Information Systems" (unpublished M.S. thesis, School of Engineering, University of Pittsburgh, 1998)

Bodnar, Randel F., "Verification of Automatically Generated Unit Hydrographs from DEM's with Field Data" (unpublished M.S. thesis, School of Engineering, University of Pittsburgh, 1995)

Chaudry, M. Hanif. Open-Channel Flow. New Jersey: Prentice Hall, 1993.

Chow, Ven Te, Maidment, David R. and Mays, Larry W. Applied Hydrology. New York: McGraw-Hill, 1988.

Chow, Ven Te. Open Channel Hydraulics. New York: McGraw-Hill, 1959.

Emerick, James, "Improving Upon A Geographical Information Systems Based Spatially Distributed Rainfall-Runoff Model" (unpublished M.S. thesis, School of Engineering, University of Pittsburgh, 2003).

Fasken, Guy B. Guide for Selecting Roughness Coefficient "n" Values for Channels. Lincoln, Nebraska: Soil Conservation Service (U.S.D.A), 1963.

Kennedy, Heather, ed., *ESRI home page, Glossary of GIS Terms*, ESRI Press, 2001. 12 Nov. 2003. < http://www.esri.com/library/glossary/glossary.html>.

Khan, Khalid Naseem, "A Geographic Information Systems Based Spatially Distributed Rainfall-Runoff Model" (unpublished M.S. thesis, School of Engineering, University of Pittsburgh, 2002).

Michelini, Mark, “Automatic Generation of Unit Hydrographs Using a Digital Elevation Model” (unpublished M.S. thesis, School of Engineering, University of Pittsburgh, 1995).

Natural Resources Conservation Service, Hydrology, Section 4, National Engineering Handbook, (2001).

Ponce, Victor M., “The Kinematic Wave Controversy.” Journal of Hydraulic Engineering 117(4) (1990a): 511-525.

Ponce, Victor M., “Kinematic Wave Modeling Where Do We Go From Here?” Proceedings of the International Symposium on Hydrology of Mountainous Areas, May 28-30 (1992), Shimla, India, 485-495.

Quimpo, Rafael G. Hydrology and Hydrualics. Pittsburgh, Pennsylvania: University of Pittsburgh, 2002.

USACE. Introduction and Application of Kinematic Wave Routing Techniques Using HEC-1, Training Document 10. Davis, CA: Hydrologic Engineering Center, (1979).

Vieux and Associates Allegheny County Radar Rainfall Analysis August 2002 Summary Report. Prepared for 3RWWDP, September 12, 2002.

Vieux & Associates Allegheny County Radar Rainfall Analysis July 2003 Summary Report. Prepared for 3RWWDP, August 25, 2003.

References Not Cited

Bronson, Gary. Modular FORTRAN 77 for Science and Engineering. El Granada, CA: Scott/Jones Inc., (1990).

Mojena, Richard, and Ageloff, Roy. FORTRAN 77. Belmont, CA: Wadsworth, (1990).

Nyhoff, Larry, and Leestma, Sanford. FORTRAN 77 and Numerical Methods for Engineers and Scientists. New Jersey: Prentice Hall, (1995).

USACE. Hydrologic Modeling System HEC-HMS Technical Reference Manual. Davis, CA: Hydrologic Engineering Center, (2000).