Calibration and Analysis of the MESH Hydrological Model applied to Cold Regions

by

Angela Jean MacLean

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Civil Engineering

Waterloo, Ontario, Canada, 2009

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

A. J. MacLean

Abstract

Concerns regarding climate change have brought about an increased interest in cold region hydrology, leading to the formation of the IP3 research network. This work is part of the IP3 Network, which has the overall goal to evaluate and demonstrate improved predictions of hydrological and atmospheric fields for cold regions. As such this thesis involves a series of calibration and validation experiments on the MESH hydrological model (used by IP3 for predictions) with two cold region case studies. The first case study is the very well instrumented Reynolds Creek Experimental Watershed in Idaho, USA and the second case study is the Wolf Creek watershed in the Yukon Territory. As the MESH model is still in the development phase, a critical component of model development is a thorough analysis of model setup and performance. One intention of this research is to provide feedback for future development of the MESH hydrological model.

The Reynolds Creek site was modeled as part of this thesis work. This site was chosen based on the long term, highly distributed and detailed data set. The second site, Wolf Creek, was used for a simplified case study. Models of both case study sites were calibrated and validated to carefully evaluate model performance. Reynolds Creek was calibrated as a single objective problem as well as multi-objective problem using snow water equivalent data and streamflow data for multiple sites.

The hydrological simulations for Wolf Creek were fair; further calibration effort and a more detailed examination of the model setup would have likely produced better results. Calibration and validation of Reynolds Creek produced very good results for streamflow and snow water equivalent at multiple sites though out the watershed.

Calibrating streamflow generated a very different optimal parameter set compared to calibrating snow water equivalent or calibrating to both snow water equivalent and streamflow in a multi-objective framework. A weighted average multi-objective approach for simultaneously calibrating to snow water equivalent and streamflow can be effective as it yields a reasonable solution that improves the single objective snow water equivalent results without degrading the single objective streamflow results.

Acknowledgements

I would like to acknowledge my supervisor, Dr. Bryan Tolson, who provided me with guidance and helpful suggestions over the past two years.

I would like to thank Dr. Frank Selenieks, who played a large role in helping me with my thesis work and my undergraduate research. Frank provided much guidance and feedback on the how to use the MESH model and was always there for debugging assistance.

I would like to acknowledge the assistance of Dr. James Craig who helped with the solar radiation inputs for the Reynolds Creek and was always around to help with trouble shooting. Thanks to Dr. Nick Kouwen, who provided guidance with the data processing for Reynolds Creek and also started me down the path of research and introduced me to the world of hydrological modeling.

The USDA provided me with seemingly endless amounts of data for Reynolds Creek. They also have taken the time to very thoroughly document the research efforts throughout the watershed. These documents became invaluable resources to me throughout the course of this project and many of the figures that they created for these documents are used in this thesis.

Finally I would like to thank and acknowledge my friends and family who supported and encouraged me over the course of this research work. Particularly my husband, Tyler Inwood, who always made sure I did my homework and my mother, Pamela MacLean, who has been great help with reviewing my spelling and grammar over the last 20 years.

Table of Contents

Appendices

List of Figures

List of Tables

1. Introduction

1.1 Background

Climate change is currently one the largest issues currently under investigation by researchers and scientists (Lemmen et al., 2008). The impacts of climate change are not just limited to an increase of the global temperature but they will also have a substantial impact on the global water cycle. The impacts of the changing climate have been documented in every region of Canada (Lemmen et al., 2008). A particularly large concern for the Canadian Prairie provinces (Alberta, Saskatchewan, and Manitoba) as well as the North Western USA is water scarcity due to the changing climate. A large source of water for these regions comes from glacial and snow melt processes in mountainous regions. Glacial melt and water stored as snow are presently used to help achieve water demand during the dryer summer season. However, changes in temperature will impact the timing of the melt water and interrupt the water supply for these populations. These concerns have brought about an increased research interest in cold region hydrology and led to the formation of the IP3 Network.

The IP3 Network work is a group of researchers funded by the Canadian Foundation for Climate and Atmospheric Sciences. IP3 stands for "Improved Processes and Parameterisation for Prediction in Cold Regions" (IP3 Network, 2008). The three objectives of IP3 Network are:

- To improve understanding of cold region hydrometeorological processes, this includes physical processes such as blowing snow transportation, snow melt, infiltration of water into the frozen ground, and other processes;
- Improve the mathematical parameterization of cold region processes, this step tries to mathematically simulate the individual processes that occur in cold regions; and
- Evaluate and demonstrate improved prediction of hydrological and atmospheric fields at regional and smaller scales in the cold regions under varying conditions (IP3 Network, 2008).

The IP3 project focuses on eight official study sites located in Canada and one unofficial study site located in the United States of America. The sites cover a variety of geographic locations and physical conditions. Sites range from high altitude to low altitude and high latitudes to lower latitude. The definition of a cold region site is not strictly limited to higher latitude sites, as mountain sites or higher altitude sites also experiences many similar processes. This thesis work focuses on two sites, Wolf Creek, a high altitude, high latitude site located in the Yukon, Canada, and Reynolds Creek, a high altitude, low latitude site located in Idaho, USA.

The MESH hydrological model is being used to evaluate and demonstrate improved predictions of hydrological and atmospheric fields for the IP3 project. This research will provide feedback for MESH model development though a series of calibration and validation experiments.

1.2 Objectives

The overall goal of this research work was to contribute to improved predictions of the MESH hydrological model by improving the calibration process for the model. This was done using two case studies, particularly Reynolds Creek which was set up as a MESH hydrological modeling case study with a long term highly distributed and detailed data set. Part of the Reynolds Creek model setup included developing a set of site specific initial parameters to be used for model calibration experiments.

To improve model predictions, calibration was used to find the best parameter set for the prediction of snow water equivalent (SWE) and streamflow at multiple sites. Both single objective and multiple objective calibration strategies were implemented for SWE and streamflow. The model was validated with data from multiple sites to verify the quality of the model predictions.

To contribute to future MESH modeling projects, a number of experiments were performed to compare alternative model calibration strategies that could be implemented with the MESH model. These experiments provided feedback on the approximate number of function evaluations for model calibration, the discretization of different land cover configuration, the initial parameter set for calibration, the parameter ranges for used for calibration, the methods used for soil calibration, and the estimation techniques for the solar radiation.

Using information from the calibration experiments, conclusions and recommendations will be drawn regarding the setup and calibration process for the MESH hydrological model. Recommendations will also be made to improve the overall performance and usability of the MESH model.

1.3 Outline of Thesis

This remainder of the thesis has been organized into the following chapters:

Chapter 2: This chapter provides a review of literature relevant to this research work.

Chapter 3: This chapter provides detailed information on the two case studies used for this work, Reynolds Creek and Wolf Creek.

Chapter 4: This chapter provides details on the data processing done to set up the MESH model for Reynolds Creek and Wolf Creek.

Chapter 5: This chapter provides information on the calibration strategy used to calibrate the MESH model.

Chapter 6: This chapter details the model experiments performed and the results for each set of model experiments. Hydrographs are presented for model calibration and validation periods as well as tables of the hydrological statistics.

Chapter 7: This chapter presents the conclusions and recommendations for this report.

2. Literature Review

2.1 Hydrological Models

A hydrological model is a mathematical model used to simulate the movement and redistribution of water within a defined area, typically a watershed. Models can be physically based, empirically based or semi-empirically based (a combination of the two). The basic principles used in a hydrological model are typically physically based in nature, derived from representations of the water cycle and the conservation of water (or the water balance).

Hydrological modeling has many engineering applications such as land use planning, flood forecasting, and water quality. There are many different types of hydrological models as well as different applications. This thesis deals with large scale hydrological model applied to simulating streamflow and case studies to determine the most applicable model setup and configuration process.

The required model inputs vary for each hydrological model. The required data can be as simple as temperature and precipitation for a water balance model or much more complex when other processes such as evapotranspiration, evaporation, snowmelt, and other physically based processes are incorporated into the model. Typically, when more processes are modeled, more data are required for model simulation and validation.

The hydrological models discussed in this section can be set up as a distributed model or a lumped model. For this thesis a lumped hydrological is a model that uses the same meteorological data (e.g. rainfall, temperature, etc.) and model parameters over the entire area being modelled. A distributed model varies the meteorological data and model parameters spatially. To use a distributed model such as MESH or WATFLOOD in a lumped fashion, the model can be set up so that the entire watershed is encompassed in one grid cell as the single model response unit that has the same meteorological forcings and model parameter values.

Numerous hydrological models have been developed by researchers. Some of the better known models include SWAT (Soil and Water Assessment Tool) (Neitsch, et. al, 2005), MIKE SHE (DHI, 1999), and TOPMODEL (Beven and Freer, 2001). Two hydrological models designed and used in Canada are WATFLOOD (Kouwen et al., 1993) and WATCLASS (Soulis et al., 2000). MESH is the next generation of the WATCLASS model. The MESH model is currently under development by the research and development group at Environment Canada.

The WATFLOOD model was developed by Dr. N. Kouwen at the University of Waterloo. It was initially developed as an early flood forecasting system. It is a combination of a physically-based routing model and a conceptual hydrological simulation model. The processes incorporated into WATFLOOD include interception, infiltration, evaporation, snow accumulation and ablation, interflow, recharge, baseflow, and overland and channel routing (Kouwen et al., 1993).

The WATCLASS model uses the distributed hydrological routing of WATFLOOD coupled with a land surface model; Canadian Land Surface Scheme (CLASS) (Kouwen, et al., 2009). CLASS (Verseghy, 1991; Verseghy et al., 1993) was developed by Diana Verseghy to be used with the Canadian Global Climate Model. A land surface model, such as CLASS, is used to simulate the energy balance and water balance equations (Graham and Bergstrom, 2000). CLASS models the energy balances and water balances of the soil, snow, and vegetation canopy.

By coupling of the WATFLOOD and CLASS models (for WATCLASS), the estimation of the water balance is improved by allowing water to move horizontally and vertically though the soil layers. This also incorporates a slope component to the model and enhances the hydrology while maintaining the atmospheric representation of the model. Other improvements that were made to the hydrology of the WATCLASS model during the development phases of WATCLASS to include treatment of surface runoff and soil drainage processes, the inclusion of a groundwater routing and storage process, and the effects of temperature, soil, ice, and snow pack of the surface and soil flow characteristics (Bastien, 2004).

The WATCLASS model has recently evolved and was renamed to MESH (Pietroniro et al., 2007). MESH includes all the features of WATCLASS but additional hydrological functions continue to be incorporated into the model. Pietroniro et al. (2007) describes the current configuration of the MESH as, individual grid cells subdivided into a number of grouped response units (GRUs) and then the land surface model (CLASS) is run on each GRU independently. After each GRU has been run, the overall fluxes and prognostic variables are calculated for each grid cell using the weighed area of each GRU. This process currently does not account for any fluxes beyond streamflow that may occur between grid cells (e.g. the redistribution of snow) and therefore still has the potential to be improved with further research.

2.1.1 Hydrological Model Studies with MESH and WATCLASS

The WATCLASS 3.0 model was used as part of the Mackenzie GEWEX study (MAGS) (Soulis and Seglenieks, 2007) as an integrated modeling system of CLASS and WATFLOOD. It was used as part of MAGS to model several research basins and for the Mackenzie River and its major tributaries (Soulis and Seglenieks, 2007). The research basins included the Smoky River (in Alberta), Wolf Creek watershed (also used in this thesis work), and Trail Valley Creek (in the Northwest Territories) as well as many of the larger tributaries of the Mackenzie River.

Bastien (2004) applied WATCLASS 3.0 using two Canadian watersheds and focused on the ability of the model to simulate streamflow and volumetric soil moisture for the Wolf Creek watershed and the Grand River Watershed. It also included an assessment of the effects of the cold soil algorithms used in WATCLASS 3.0. All model calibration focused on streamflow as a performance indicator for model.

Davison et al. (2006) used WATCLASS 3.0 in a study examining an approach to improve the simulations of the spatial variability of snowmelt during the spring melt period for a small arctic watershed, Trail Valley Creek. Landscape based parameters were estimated via calibration to streamflow and snow-cover area data. Davison et al. (2006) report that when streamflow was used for a calibration objective, the corresponding snow-cover area results were poor and vice versa, demonstrating the need for multi-objective calibration. The study also notes that the difficulty to obtain correct discharge and snow-cover areas simultaneously indicates that there are some problems with the snow processes in the model. Davison et al. (2006) conclude that difficulties associated with modeling the snow processes could be problems modeling meltwater retention in the snow pack and snow damming processes in the stream channels.

Pietroniro, et al. (2007) applied the MESH model to regional scale hydrological forecasting for the Laurentian Great Lakes. The goal of this research work was to better understand behavior of different land-surface models, and testing different schemes for producing streamflow forecasts.

Dornes et al. (2008) applied the MESH model to two artic case studies, Granger Basin (a subwatershed of Wolf Creek) and Trail Valley Creek. Dornes et al. (2008) evaluated how well landcover-based model parameters can be transferred between watersheds in a northern environment. A selection of snowmelt parameters were calibrated to snow water equivalent data in Granger Basin using CLASS and then transferred to Trail Valley Creek using a landscape similarity criterion. Findings in Dornes et al. (2008) demonstrate that when effective landscape-based parameters are defined, better snow-cover depletion and snowmelt runoff simulations are achieved than using a default parameterisation scheme.

2.2 Model Calibration

Historically, model calibration was conducted by trial and error, with modelers changing parameter values and reviewing the changes in hydrograph accuracy and statistical measures or model fit. This trial and error (or manual) calibration approach can be applied to any hydrological model calibration. This approach generally leads to a strong understanding of model parameters and can be highly effective when conducted by experienced modelers (Madsen, 2000).

With automatic calibration, parameters are adjusted according to a specified search algorithm and numerical measures asses the goodness-of-fit for the new parameter set (Madsen, 2000). Generally automatic calibration is much faster than manual calibration, however often a single objective measure is not adequate to properly estimate the simulation of all the important hydrological characteristics (Madsen, 2003).

With the ever-increasing number of physical and conceptual parameters applied in a distributed hydrological model, the use of an effective automatic calibration scheme is more important than ever. The current configuration of the MESH model used in this study has approximately thirty-one parameters per GRU that could be calibrated. With the use of a distributed multi-GRU model setup the number of parameters that could be calibrated quickly reaches more than a hundred. As a result of the large number of calibration parameters, care should be taken to ensure that automatic calibration is properly implemented.

The key components of automatic calibration include the selection of an appropriate objective function, appropriate calibration data (initial inputs and parameters), and an effective algorithm to optimize the objective function (Gupta et al. 1998, and Singh and Woolhiser, 2002).

The objective function is the only measure of solution quality (or model performance) and is used by the optimization algorithm as the measure for goodness-of-fit. A proper objective function should be chosen to reflect the purpose of the optimization. Different objective functions put more emphasis on different components of the hydrograph, such as matching the peaks, shape, volume or low flows. A very commonly used indicator of model performance used by hydrologists is the Nash–Sutcliffe efficiency measure (Nash and Sutcliffe, 1970). With the Nash–Sutcliffe efficiency measure, values are normalized between negative infinity and one, where a value of one indicates a perfect model fit and a value greater than zero indicates that the simulated streamflow will produce better results than the average observed streamflow. While this measure can properly capture the shape and volume of the hydrograph, a large emphasis is placed on matching the peak flows and it does not easily account for timing errors that often occur in hydrological models. The Nash-Sutcliffe value (*NS*) compares a simulated and observed hydrograph over *N* time steps is calculated as follows:

$$NS = 1 - \frac{\sum_{i=1}^{N}(S_i - O_i)^2}{\sum_{i=1}^{N}(O_i - O_{avg})^2}$$

Where S_i is the simulated discharge at time step *i*, O_i is the observed value at time step *i*, and O_{avg} is the average observed discharge for the *N* time steps.

Another measure of model performance is the absolute percent bias (APB) which focuses on the total volume difference between the observed and simulated time series. This measure does not examine the corresponding timing of the flows, and is therefore not a good measure for a single objective function, but can still be used in a multi-objective framework.

$$APB(\%) = \frac{\left|\sum_{i=1}^{N} S_i - \sum_{i=1}^{N} O_i\right|}{\sum_{i=1}^{N} O_i} * 100$$

The coefficient of determination (R^2) describes the proportion of the total variance in the observed data (Legates and McCabe, 1999). A perfect model has a correlation coefficient equal to 1.0. High values of the R coefficient indicate better agreement between observations and simulations.

$$R^2 = \left(\frac{\sum_{i=1}^{N}(O_i - O_{avg}) * (S_i - S_{avg})}{\sqrt{\sum_{i=1}^{N}\left(O_i - O_{avg}\right)^2} * \sqrt{\sum_{i=1}^{N}\left(S_i - S_{avg}\right)^2}} \right)^2$$

As with the Nash-Sutcliffe statistic, the correlation coefficient is more sensitive to outliers than to values near the observed mean.

2.2.1 Single Objective Calibration Methods

Optimization algorithms are often applied to automatically calibrate hydrological models. Early algorithms focused on finding locally optimal solutions; however, given the complexity of present day hydrological models, more advanced algorithms are necessary to find high quality solutions that are close to the global optimum (Tolson and Shoemaker, 2007). This section discusses three algorithms that are often applied for the automatic calibration of hydrological models. These algorithms are the genetic algorithm (GA) (Wang, 1991), shuffle complex evolution (SCE) (Duan et al., 1992), and dynamically dimensioned search (DDS) (Tolson and Shoemaker, 2007).

The GA is a popular global optimization algorithm that is an evolutionary optimization strategy. Wang (1991) was one of the first to apply the GA to hydrologic model calibration and describes the GA as follows. The method combines processes witnessed in nature, such as the process of natural selection, survival of the fittest and natural mutations. Since the initial development of the GA many other calibration strategies have used this framework to develop other global search algorithm. A set of random solutions are sampled within the search space and the corresponding objective function values are evaluated. Next, the values are ranked so that the better solutions have a better chance of being selected to parent another solution. After randomly selecting two parents, a crossover operation is then performed to combine the solutions of two of the selected parents and form one or more offspring. Then, with a low probability, a mutation will be introduced into the offspring produced by crossover so as to produce a solution that contained some variation relative to the parent solutions. With enough function evaluations the GA is often capable of consistently finding the global optimum of the model.

SCE was developed by Duan et al. (1992) as a global optimization method for the calibration of conceptual rainfall-runoff models. The algorithm incorporates a series of different techniques including multiple start points, controlled random search, competitive evolution, and an information sharing tool through complex shuffling (Duan et al. 1992). Multiple start points (based on the simplex search method) are used in different locations or ‘neighborhoods’ to start the search over a complex response surface. To improve the efficiency of the multi-search method, it shares information between different neighborhoods. To insure that the solution progressively evolves to a better solution, potential parent solutions are selected based on the fitness of the solution. When a new better solution is found the worst solution in the community it is replaced. Using these techniques none of the information contained in the sample is ignored or lost and the algorithm keeps searching until it converges to a solution (typically a local or global optimum). One problem with this algorithm is that it can often take more than 10,000 function evaluations to converge on the optimal solution and thus can be computationally prohibitive in a hydrologic modeling context.

DDS algorithm is described as a simple stochastic single-solution based heuristic global search algorithm developed to find a good global solution within a limited number of function evaluations (Tolson and Shoemaker, 2007). The algorithm uses a user defined number of function evaluations and determines the portion of the computational budget that can be used for a global search before it progresses to a local search around the best solution. The candidate solutions are created by perturbing the decision variables (i.e. model parameters) in the best solution by a randomly sampled value from a normal probability distribution with a mean of zero. As the search progresses, the number of decision variables perturbed decreases and the algorithm performs more one-at-a-time function evaluations that can later be extracted for sensitivity information.

DDS was applied to CLASS and MESH models by Dornes et al. (2008) to calibrate the snow cover depletion curves for Granger Basin (a subwatershed in Wolf Creek), and a weighted average of snow covered area and streamflow for Trail Valley Creek. DDS was generally able to produce good results after 100 function evaluations in this case study (Dornes et al., 2008). DDS was also used for a study conducted by Seglenieks (2009) to calibrate a WATFLOOD model set up for all of Canada. In this study DDS, was selected for its ability to calibrate a model with minimal function evaluations. Tolson and Shoemaker (2007) report that DDS required 15-20% fewer model evaluations to converge on comparable solutions identified by the SCE algorithm.

2.2.2 Multi-Objective Calibration

Multi-objective calibration is the process of optimizing multiple objective functions simultaneously. This optimization strategy is typically applied to distributed hydrological models in order to simultaneously incorporate multiple sets of measured data into the model calibration process. Due to the complexity of model processes and the number of parameters required to calibrate most distributed hydrological models, if only one calibration objective is used, then an erroneous parameter set cannot easily be identified. If the data and model resources are available, a multi-objective calibration is strongly recommended to avoid erroneous parameter sets. A single measure describing calibration performance is often inadequate to account for the simulation of all characteristics in the system that are used by a hydrologist to evaluate goodness-of-fit (Madsen, 2003).

There are multiple approaches to implement a multi-objective calibration strategy. The strategy can be as simple as using a weighted average to aggregate multiple objectives into a single objective function (utilized in this thesis). Examples of this approach in hydrologic model calibration include Madsen (2003), Zhang, et al. (2008), and Tolson and Shoemaker (2008). Another strategy that is more complex would be to apply a multi-objective optimization algorithm that searches for tradeoff solutions (Pareto curve) between the multiple objectives. The Pareto curve is a set of solutions whose objective values are not dominated by any other solutions (Deb, 2001). In a two dimensional problem the curves can be graphically illustrated to examine the set of Pareto solutions and then select one as the desired solution to the calibration problem. Visualization of the Pareto curve becomes increasingly difficult as three or more dimensions are added to the objective function.

3. Case Studies

3.1 Reynolds Creek Experimental Watershed

Reynolds Creek Experimental Watershed (illustrated in Figure 3-1) is located in the Owyhee Mountains of south western Idaho, approximately 80km west of Boise, Idaho in the United States of America. The basin was set up as a research basin by the United States Department of Agriculture (USDA) in the mid 1960’s to address the issues of water supply, seasonal snow, soil freezing, water quality, and rangeland hydrology (Slaughter et al., 2000). Extensive data collection started in 1962 and continues today.

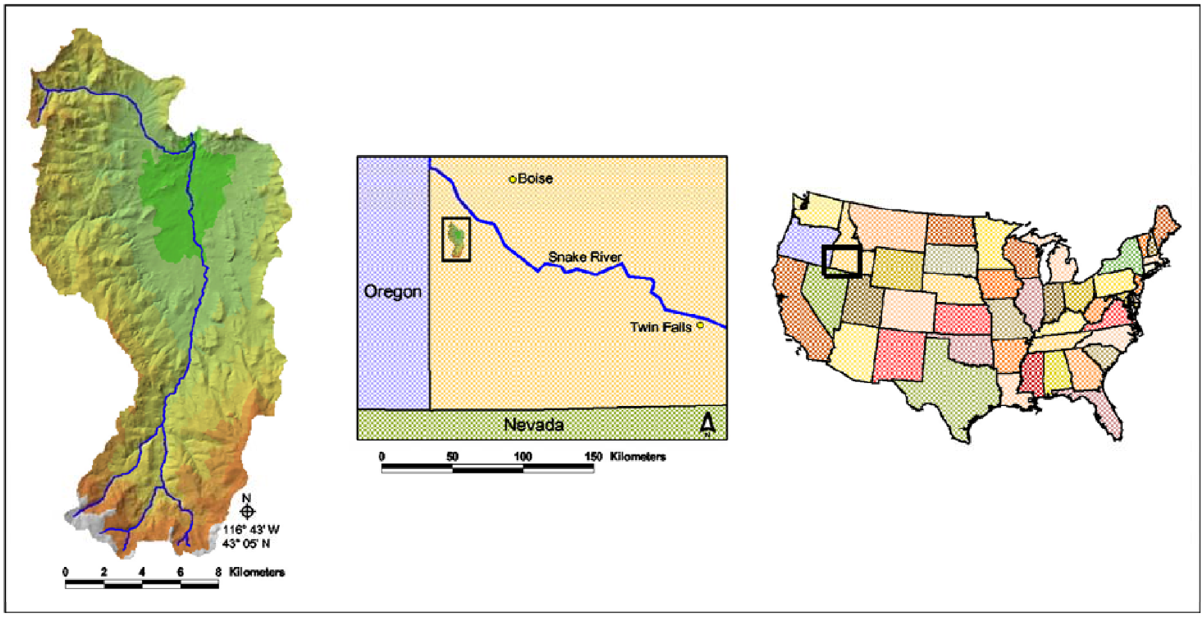

Figure 3-1: Location of Reynolds Creek Experimental Watershed from Slaughter et al. (2000)

Slaughter et al. (2000) describe Reynolds Creek as a third order subwatershed draining into the Snake River basin. The watershed elevations range from 1101 MSL to 2241 MSL (Seyfried et al., 2000a). This watershed has some very unique features in terms of vegetation, temperature, and precipitation. The vegetation of the watershed includes various species of sagebrush, greasewoods, aspen, conifers and some agricultural grazing lands (Seyfried et al., 2000a). The mean air temperature in the lower ranges of the basin is 8.9˚C and in the higher basin elevations at the headwaters the mean air temperature is 4.7˚C (Hanson et al., 2000). Precipitation in the basin also varies with elevation. The lower elevations receive about 230 mm of precipitation per year while the higher regions receive as much as 1100 mm per year, mostly in the form of snow (Hanson, 2000). The distribution of precipitation in the watershed is illustrated later in Figure 3-2.

Data for Reynolds Creek has been collected by the USDA. Extensive data preparation, quality assurance, and quality control was performed by the USDA prior to the data being released for public use.

Reynolds Creek was modelled using MESH for a study period of 10 years. The meteorological data were set up from January 1st, 1986 to September 30th, 1996. This study period was selected based on the availability of hourly meteorological data collected within the Reynolds Creek watershed. The model runs were initialized based on the approximate start of the water year. Runs were started on September 2nd, 1986.

3.1.1 Other Hydrological Modelling Studies Conducted at Reynolds Creek

Reynolds Creek is an incredibly well instrumented research basin and as a result, many other research studies have also used the Reynolds Creek watershed for hydrological modelling. This section briefly reviews some of the other Reynolds Creek modelling studies simulating either SWE or streamflow (the focus of this thesis).

Bathurst and Cooley (1996) simulated SWE and streamflow in the Reynolds Mountain East subwatershed to validate the snow melt component of the SHE hydrologic model (Abbott et al., 1986). The Reynolds Mountain East sub basin was set up with a 50 m^2 grid cell resolution to validate snow melt routines and sensitivities for a twelve-day period during an early melt event in mid February. Bathurst and Colley (1996) report that multi-objective calibration for streamflow and SWE must be considered for model calibration and that multiple measures should be applied for validation.

Zhang et al. (2008) modelled Reynolds Creek using the SWAT hydrologic model (Arnold et al., 1998) and focused on the calibration procedure for streamflow. They compare the performance of a single objective model calibration to a multi-objective calibration method using data from three streamflow stations for the objective functions. The three streamflow locations used for model calibration are Reynolds Creek Outlet, Salmon Creek, and Reynolds Creek at Tollgate. The result of Zhang et al. (2008) show that the single objective calibration and multi-objective calibration for different sites can lead to different objective function values, parameter solutions and hydrographs.

Their study failed to consider water loss at the outlet of Reynolds Creek caused by irrigation during drought years as reported by, Johnson and Smith (1979). For example, approximately 72% and 23% of flows were diverted during 1968 and 1969, respectively. These two years were used by Zhang et al. (2008) for model calibration. This would likely account for some of the differences experienced in calibration results noted by Zhang et al. (2008) between the Reynolds Creek Outlet and the Tollgate subwatershed. Since irrigation impacts appear to have influenced calibration in Zhang et al. (2008), the Reynolds Creek Outlet is not the focus of calibration in this thesis and instead, calibration focuses mainly on the Tollgate subwatershed as no irrigation occurs within the Tollgate subwatershed.

3.1.2 Meteorological Data Collection at Reynolds Creek

There are three meteorological stations set up at different elevations for the period of study used. The locations of the meteorological stations are presented in Figure 3-2. Data collected at the meteorological stations include relative humidity, ambient air temperature, dew-point temperature, shortwave solar radiation, wind speed and direction, daily Class A pan evaporation, and barometric pressure (Hanson et al., 2000). Station locations and the data collected at each station is summarized in Table 3-1. Meteorological data collection began on a daily basis between 1974 and 1977. Hourly data collection for most parameters began between 1981 and 1984 for all three meteorological stations.

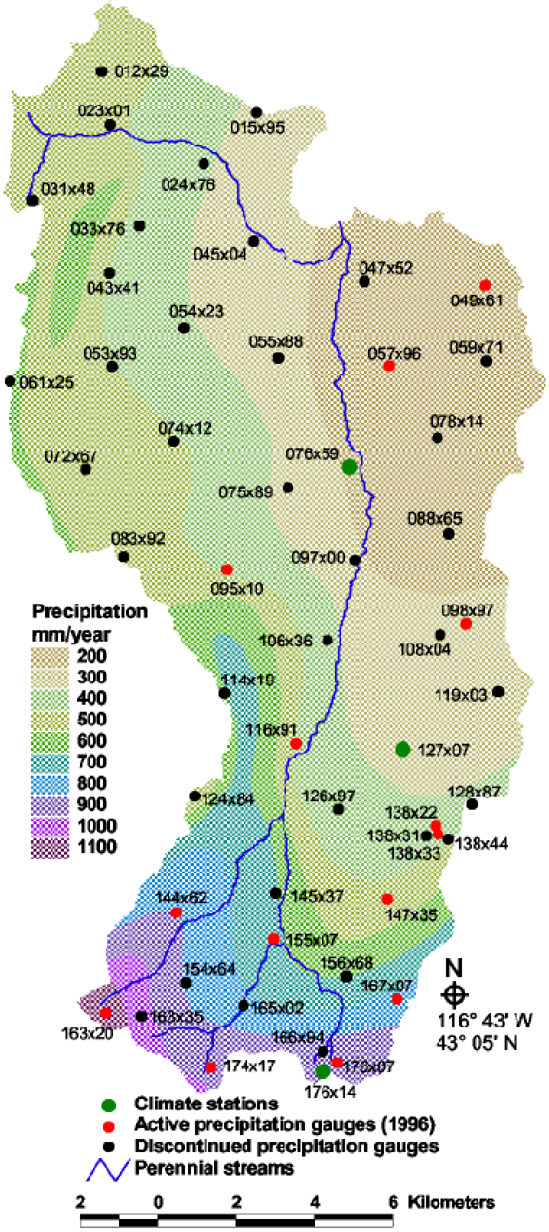

Figure 3-2: Climate and precipitation station locations and the distribution of annual precipitation at Reynolds Creek from Hanson et al. (2000)

Table 3-1: Summary of meteorological data collected at Reynolds Creek stations, taken from Hanson et al. (2000)

Station	**076x59**	**127x07**	**176x14**
Location			
Easting	520,367	521,742	519,693
Northing	4,783,418	4,776,189	4,767,923
Elevation			
GPS	1207	1652	2097
DEM	1202	1653	2097
Data Collected			
Relative Humidity	✓	✓	✓
Air Temperature	✓	✓	✓
Dew-point Temperature	✓	✓	✓
Shortwave Radiation	✓	✓	✓
Wind speed and Direction	✓	✓	✓
Pan Evaporation	✓	✓	✓
Atmospheric Pressure	✓		
Vapor Pressure	✓	✓	✓

An extensive precipitation gauge network was set up set up in conjunction with the meteorological stations. The original gauge network consisted of 83 unshielded weight-recording gauges (Hanson, 2000). Due to the undercatchment of snow in the network, the gauge network was converted to 46 pairs of unshielded and shielded gauges with orifices 3.05 m above ground (Hanson, 2000). Of these 46 dual-gages, the 17 dual-gauges that were still active in 1996 (noted in Figure 3-2) were used in the model set up of Reynolds Creek. See Hanson (2000) for the period of record for these 17 stations.

3.1.3 Subwatershed Streamflow monitoring

The streamflow monitoring network in place at Reynolds Creek contains both perennial and ephemeral streams. For this study, four perennial streamflow locations were utilized for model calibration and/or validation:

- Reynolds Creek Outlet;

- Reynolds Creek Tollgate;
- Salmon Creek; and
- Reynolds Mountain East.

The above locations were selected to have a variety of different stream flow characteristics. Figure 3-3 shows the locations of the streamflow gauges in the watershed that are monitored by the USDA and Table 3-2 lists locations, elevations and drainage areas for the streamflow locations that were used for model calibration and/or validation.

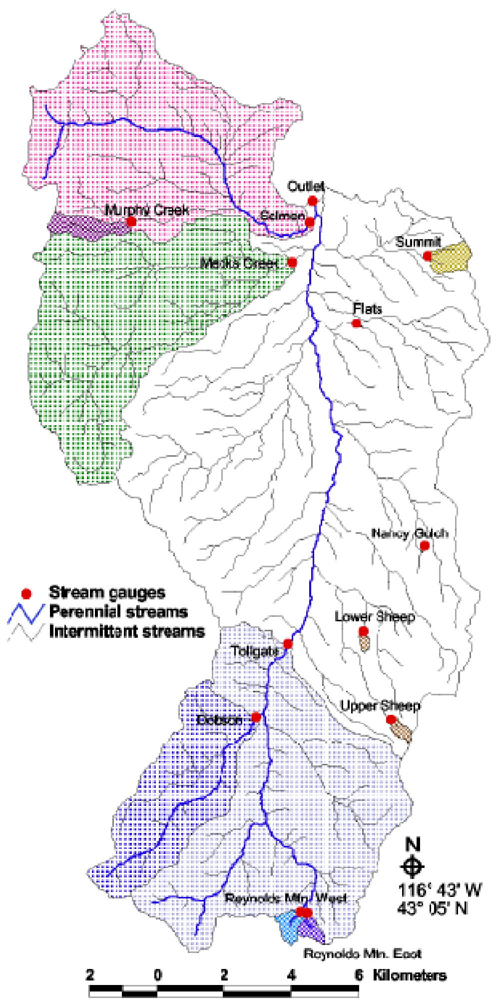

Figure 3-3: Streamflow monitoring station from Pierson et al. (2000)

Table 3-2: Summary of streamflow gauging adapted from Peirson et al. (2000)

Watershed Name	Reynolds Creek outlet	Reynolds Creek Tollgate	Salmon Creek	Reynolds Mountain East
Location				
Easting	520,111	519,393	520,015	519,746
Northing	4,789,673	4,776,495	4,788,996	4,768,494
Elevation Range (m)	1101-2241	1410-2241	1121-1918	2026-2137
Duration of Record	1963-1996	1966-1996	1964-1996	1963-1996
Drainage Area(km^2)	238.66	54.57	36.19	0.36

3.1.3.1 Reynolds Creek Outlet

Reynolds Creek Outlet has the largest drainage area of the four streamflow gauging stations utlized in this study. The outlet weir is located in a narrow canyon approximately 11 km south of the confluence at Snake River (Perison et al., 2000). Reynolds Creek is a tributary of Snake River.

3.1.3.2 Reynolds Creek Tollgate

The "Tollgate" study area is part of Reynolds Creek. Defined by the Tollgate weir, this study area encompasses the upper section of Reynolds Creek, and also contains the headwaters of the watershed (Peirson et al., 2000). These areas receive the most annual precipitation in watershed, typically in the form of snow. All the snow course sites and the snow study site are located in the Tollgate subwatershed.

3.1.3.3 Salmon Creek

The Salmon Creek watershed is characterized by steep mountainous topography, rock outcrops and shallow rocky soil (Peirson et al., 2000). This subwatershed is located in the north-west sector of Reynolds Creek and the Outlet of Salmon Creek is located less than 1 km upstream from the weir location for Reynolds Creek Outlet. There were no precipitation gauges or meteorological stations located within this subwatershed and all data used for modelling was interpolated from nearby stations.

3.1.3.4 Reynolds Mountain East

Reynolds Mountain East is the smallest continuously monitored subwatershed (0.36 km^2). It is located at the upper most head waters of Reynolds Creek. The streamflow in the basin consists predominately of the melting snow pack that typically accumulates in the area.

3.1.4 Snow Monitoring Data

When Reynolds Creek was initially set up as an experimental watershed one of the essential measurements was snow with regards to "snow deposition and melt as related to the nature of snowfall, shifting by wind, vegetation, topography, and meteorological factors" (Robins et al., 1965). Seven initial snow course sites were established in 1961, and one additional snow course was added in 1970 (Marks et al., 2000). A snow pillow was also installed on the site in 1983 to record daily SWE readings. The locations of the eight snow course sites plus one snow pillow site are presented Figure 3-4. Five of these sites (4 snow course sites plus the snow pillow site) were selected for model development purposes and are highlighted in red or green in Figure 3-4 and also summarized in Table 3-3.

Table 3-3: SWE site locations adapted from Marks et al. (2000)

USDA	Location		Elevation	
Site ID	Easting	Northing	GPS	DEM
155x54	517,892	4,770,341	1743	1733
163x20	514,134	4,769,430	2170	2166
167x07	521,613	4,769,718	2010	2009
174x26	516,719	4,767,777	2078	2072
Snow Pillow	520,055	4,768,117	2061	2058

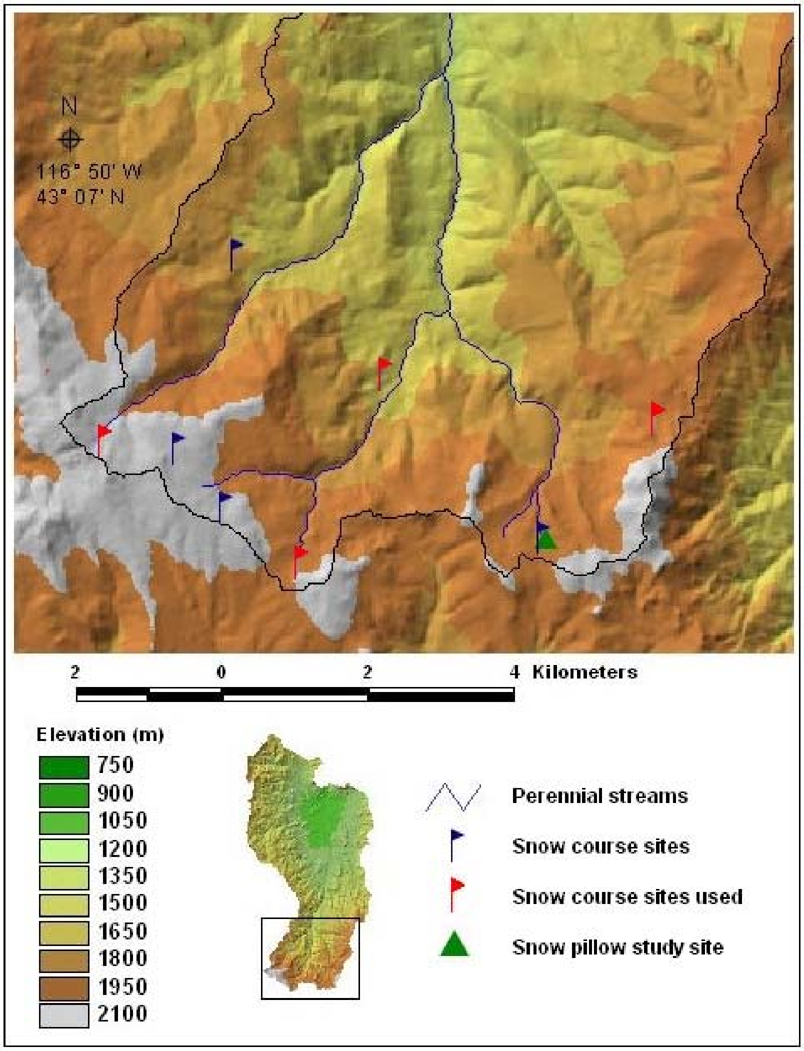

Figure 3-4: Snow monitoring locations within Reynolds Creek, adapted from Marks et al. (2000)

For a detailed description of each snow course refer to Marks et al. (2000). The sites modeled in this study are typically situated near a stand of fir or aspen trees. They are located in the extreme headwaters of the watershed, often near the ridgeline boundaries. There is also a substantial amount of variance of annual snowfall within the upper elevations of the watershed. The average annual SWE for the snow sites located within the watershed can vary from 200 mm/yr to 750 mm/yr, while the maximum SWE values vary from 450 mm/yr to 1150 mm/yr (Marks et al., 2000). Therefore, it is very important to accurately measure the precipitation distribution within the watershed, particularly within the headwaters of the watershed.

3.2 Wolf Creek Research Basin

Wolf Creek is a subwatershed of the Yukon River located approximately 15 km south of Whitehorse, Yukon at 61^{o} N Latitude (Janowicz, 1998). Wolf Creek became a research basin in 1992 to further the understanding of a sub-arctic watershed in the Yukon. The project had a broad range of objectives including to "preserve and enhance the integrity, health, biodiversity, and productivity of Arctic ecosystems for the benefit of future generations" (Janowicz, 1998). A number of different studies have been carried out at Wolf Creek, the most recent being the IP3 project that this research is supported by. Previous research projects conducted at Wolf Creek include the Arctic Environmental Strategy, the Canadian Climate programme: Global Energy and Water Cycle Experiment (GEWEX), and then the Mackenzie GEWEX Study (MAGS).

The Wolf Creek watershed, see Table 3-4, covers an area of approximately 185 km^2 over an elevation range from 800 to 2250 MSL, including a wide range of ecosystems ranging from boreal forest to alpine tundra. The three dominant ecosystems are boreal forest (spruce, pine, and aspen), subalpin taiga (shrub tundra) and alpine tundra (Francis, 1997).

Table 3-4: Streamflow stations used at Wolf Creek watershed

Watershed Name	Wolf Creek at the Alaska Highway
Location	
Easting	503,000
Northing	6,718,900
Elevation Range (m)	800-2250
Duration of Record	1996-2000
Drainage Area(km^2)	185.00

The climate at Wolf Creek watershed has a large temperature variation, and relatively low annual precipitation. The mean annual temperature is estimated at -3^{o}C, with summer and winter temperatures averages ranging from 5^{o}C to 15^{o}C and -10^{o}C to -20^{o}C, respectively. Summer highs are approximately 25^{o}C and winter lows can reach temperatures of -40^{o}C. The average annual precipitation for the area ranges from 300 to 400 mm per year (Wahl et al. 1987).

4. Model Setup

There are four main steps to setting up a watershed simulation using the MESH model. For each watershed, a grid size must be selected, meteorological forcing data has to be processed and distributed to each grid cell throughout the watershed, each grid cell must be split into grouped response units (GRUs), and the input parameters and initial conditions defined for each GRU.

The size of the grid cells will typically depend on the resolution of data that is available for each individual watershed. If detailed data are available, then a small grid could be used (e.g. 1-4 km^2). If meteorological data for the watershed is limited or the watershed is very small, the whole watershed can be modeled as a single, (potentially large) grid cell. The case studies were modelled using different spatial discretizations. For Reynolds Creek, a grid cell size of 2 km was selected. A smaller resolution was not used due to the large computational resources required to run the model. Wolf Creek was setup and run using a single 30 km grid cell that encompasses the whole watershed.

The MESH model requires seven meteorological input forcings. They are as follows:

1. shortwave radiation;
2. longwave radiation;
3. wind speed;
4. barometric pressure;
5. absolute humidity;
6. temperature; and
7. precipitation.

The MESH model uses the grouped response approach developed by Kouwen et al. (1993) for modeling subgrid variability in addition to the CLASS approach of subgrid variability that allows the GRUs to be subdivided based on vegetation into “tiles”. The GRU approach was developed for the WATFLOOD hydrological model to deal with basin heterogeneity by combining areas of similar hydrological behavior. The concept is illustrated in Figure 4-1. MESH allows users to merge the GRU and tile subgrid representation approaches by defining multiple GRUs per grid and then further subdividing the GRUs into different tiles (Pietroniro (2007) refers to this as a mosaic approach). However, for the modeling work with MESH in this thesis, only one level of subgrid representation is applied so that multiple GRUs only have one vegetation class and there is no further subdivision of the GRUs into tiles. Under this GRU strategy, MESH computes fluxes (e.g. overland flow) and tracks state or prognostic variables (e.g. SWE, soil water content) for each GRU in every grid cell.

Streamflow for each grid cell outlet is computed from the total (or area weighted average) overland flow, interflow and groundwater of all GRUs in the grid cell. Similarly, fluxes into the atmosphere for each grid cell are calculated as the area weighted averages from the GRUs. A greater number of GRUs modelled will result in an overall increase in model run time.

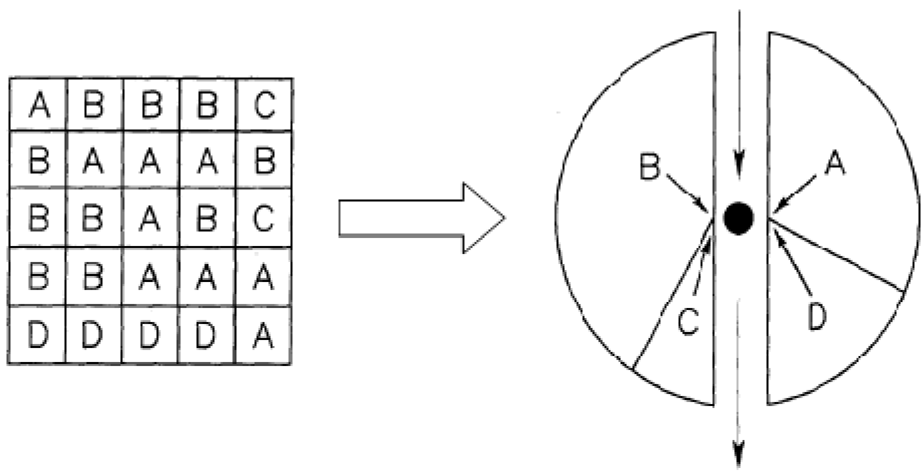

Figure 4-1: The GRU concept, figure adapted from Kouwen et al. (1993); a grid square (on the left with GRU types A through D) is generalized to four response units (on the right) where spatial locations of response units within the grid cell are not represented

There are many approaches to determine different GRUs for the model. One important factor for determining the GRUs is the parameters that are associated with each GRU. Each GRU will have the same vegetation parameters, soil parameters and hydrological parameters. Traditionally, GRUs are defined by the land cover vegetation. Land cover is typically the simplest parameter to use as information can usually be obtained from LANDSAT or LIDAR images that have been processed into GIS files. Other options for GRU categorization include soil data or topographical data such as slope and aspect. Slope and aspect were applied by Dornes et al., (2008) to a MESH modeling study of Granger Basin, a subwatershed of Wolf Creek. The study examined slope and aspect for GRU distributions to model SWE and snow-covered area. In all case studies presented in this thesis, land cover vegetation was selected to distribute the GRUs.

MESH input parameter files were generated using estimated values that were determined from default values in the CLASS or MESH documentation. Measured data were used if available. There are two parameter input files for MESH; CLASS inputs and hydrology inputs. Parameters used as part of the CLASS model are given in detail in Appendix A with their abbreviated names that are used for this thesis. The input parameters for each GRU can be generally summarized into three different types of parameters; vegetation parameters, hydrology and soil parameters, and initial conditions or initial state variable values.

Vegetation parameters (described in more detail in Verseghy (2008)) include:

- leaf area index;
- natural logarithm of roughness length for atmospheric parameters;
- visible albedo and near infrared albedo;
- standing biomass density;
- rooting depth;
- stomatal resistance; and
- coefficient governing the responses of stomates to light, vapor pressure deficit, and soil water suction.

Hydrology and soil parameters (described in more detail in Soulis (2007)) include:

- drainage index which controls soil physics for model drainage;
- drainage density representing linear streams per square area;
- valley slope;
- manning's n for overland flow surface roughness;
- lateral saturated hydraulic conductivity at the surface;
- change in lateral conductivity between layers;
- permeable depth of the soil column which limits the laterally depth that water is transported;
- percentages of sand, clay and organics in each soil layer;
- channel roughness coefficients;
- maximum allowable snow depth; and
- maximum water ponding depths (depression storage) for snow free and snow covered areas.

Initial conditions that must be specified include:

- temperatures of soil layers, canopy, snow pack and ponded water;
- fractional volume of liquid water and frozen water in each soil layer;

- depth of ponded water;
- liquid and frozen water held on the vegetation canopy;
- depth of snow present on ground;
- albedo and density of snow;
- vegetation growth index for full leaf or no leaves; and
- start and end dates for the meteorological data.

Most of the parameters that are required by MESH are not measured on site and appropriate values must be indirectly estimated. The unknown values will be estimated through extensive model calibration.

4.1 Reynolds Creek

4.1.1 Meteorological Forcing Data

The Reynolds Creek basin was modelled using MESH for a study period of almost 10 years starting on January 1st, 1986 and ending on September 30th, 1996. Model spin-up (model initialization period) took place from September 2, 1986 to December 31st, 1986 and model calibration used data from January 1st, 1987 to December 31st, 1988. All other study years were used for model validation.

Because of the dense network of gauges within Reynolds Creek (3 meteorological stations at low, mid and high range elevations and 17 dual-precipitation gauges) there was enough information collected to use a simple interpolation method such as inverse distance weighting. This method (given in the following equation) uses the distance between the points and a decay function to determine how much influence each station has on the resultant value (Dingman, 2002).

$$u(x) = \frac{\sum_{k=0}^{N} w_k(x) u_k}{\sum_{k=0}^{N} w_k(x)}, where \quad w_k(x) = \frac{1}{d(x, x_k)^p}$$

Where $u(x)$ is the interpolated value, $d(x,x_k)$ is the distance between the measured point (x) and the interpolated point (k), p was set equal to 2 for an inverse-square distance weighting, and u_k is the measured value at point k.

The meteorological forcing data were distributed using inverse distance weighting to the centre of each 2km square grid cell. Details are given for each of the meteorological forcings in the following sections.

4.1.1.1 Solar Radiation

MESH requires shortwave and longwave radiation inputs for each grid cell. Shortwave solar radiation was measured at the three meteorological stations in the watershed at an hourly time step. The measured solar radiation could not be directly distributed using inverse distance weighting while accounting for the impact of slope and aspect.

To incorporate the impacts of slope and aspect, the solar radiation algorithms of the hydrological model Raven, currently under development by Dr. Craig (2009) were applied. The solar radiation input routines in Raven were developed using equations from Physical Hydrology: Appendix D by Dingman (2002) and Yin (1997) to calculate maximum clear sky solar radiation on a sloped surface and total clear sky solar radiation. Maximum and total clear sky solar radiation required the input of slope, aspect, Julian day, year, latitude, dew point temperature and albedo. Slope and aspect were calculated using the 30 m Digital Elevation Model (DEM) and then averaged to determine the slope and aspect for each grid square. Dew point temperature was measured on site and the average albedo was estimated from Hanson (2001).

The total calculated clear sky solar radiation was compared to the observed shortwave radiation values to estimate three average daily fractional cloud cover values that could then be distributed to each grid cell using inverse distance weighting.

The input shortwave solar radiation for MESH was calculated for each grid by applying the distributed daily fractional cloud cover to the slope and aspect corrected clear sky shortwave solar radiation from Raven.

In summary, the shortwave solar radiation was calculated using the following steps:

1. Determine solar radiation inputs for each grid cell;
2. Run Raven (Craig, 2009);
3. Calculate daily fractional cloud cover;
4. Distribute fractional cloud cover using inverse distance weighting;
5. Correct calculated maximum clear sky solar radiation using fractional cloud cover.

Longwave solar radiation was not measured in the watershed during the period that the model was to be simulated and was estimated as described below. Longwave radiation was calculated using the output fractional daily cloud cover from Raven and the method described in the Simultaneous Heat and Water (SHAW) Model: Technical Documentation (Flerchinger, 2000). The inputs for longwave radiation are daily fractional cloud cover and ambient air temperature.

1. Determine the clear-sky atmospheric emissivity, (ε_a)

$$\varepsilon_a = 1 - a_\varepsilon exp[-b_\varepsilon T_a^2]$$

Where: a_ε=0.261

b_ε=7.77x10^{-4}

T_a= air temperature in ˚C

2. Determine the atmospheric emissivity (ε_{ac})

$$\varepsilon_{ac} = (1 - 0.84)\varepsilon_a + 0.84C$$

Where: C = daily fractional cloud cover

3. Determine long-wave radiation incident on the surface, (L_i)

$$L_i = \varepsilon_{ac}\sigma T_K^4$$

Where: σ= Stefan-Boltzman constant (5.6697x10^{-8})

T_K= air temperature in Kelvin

4.1.1.2 Wind Speed

Wind speed and wind direction data were collected at the three meteorological stations located throughout Reynolds Creek. The anemometers were located at 2 m above the ground climate station (Hanson et al., 2000). Data were distributed using inverse distance weighting.

4.1.1.3 Barometric Pressure

Barometric pressure was measured at one climate station, 076x59 starting in February 1987 (Hanson et al., 2000). As the MESH model is not overly sensitive to barometric pressure, the missing data were interpolated on the basis of measured data before and after the period with missing data.

4.1.1.4 Humidity

Average relative humidity or dew-point temperature was recorded at each of the three meteorological stations located in the watershed. Prior to the public distribution of the data by the USDA, relative humidity was calculated using standard conversion, and missing data were estimated using regression relationships (Hanson et al., 2000).

The measured relative humidity, barometric pressure and air temperature were applied to the methods outlined in Lowe (1976) and Brutsaert (2005) to calculate saturated vapour pressure and convert relative humidity to specific humidity.

4.1.1.5 Temperature

Temperature data were collected at the three meteorological stations located throughout Reynolds Creek. The data were distributed using a lapse rate, minimum grid cell elevation and inverse distance weighting interpolation, as described below. The minimum grid cell elevation was used rather than average or median grid cell elevation as this is the required elevation input for the MESH model.

The environmental lapse rate was calculated on an hourly basis. Linear regression was used to determine the environmental lapse rate between the temperature measured at each meteorological station and the corresponding elevation. Next, the environmental lapse rate was used to convert the hourly temperature at each gauge to mean sea level. Once the three stations were at mean sea level, inverse distance weighting was used to calculate the temperature at each grid cell for mean sea level. Finally, the calculated temperature grid was corrected for elevation using the hourly environmental lapse rate and the minimum basin elevation for each grid cell.

4.1.1.6 Precipitation

Precipitation for each gauge location was calculated by the USDA using shielded and unshielded data. The computed precipitation values distributed by the USDA for each gauge were used with the inverse distance weighting method to calculate the hourly precipitation for each grid cell.

4.1.2 Grouped Response Units

The GRUs for Reynolds Creek were distributed based on land cover vegetation data. Detailed vegetation mapping was undertaken by the USDA. A Soil Adjusted Vegetation Index (SAVI) map was defined from topographically corrected LANDSAT images (Seyfried et al., 2000a). The original highly detailed vegetation map contained 90 different vegetation mapping units and was considered too detailed for hydrological modeling purposes. The consolidated vegetation uses nine vegetation classes and is presented in Figure 4-2. As this information is also quite detailed, and would hamper the computational efficiency of the model, the data were further consolidated into six land classes for hydrological modeling purposes. The six consolidated land classes are presented in Figure 4-3 and the corresponding descriptions are given in Table 4-1.

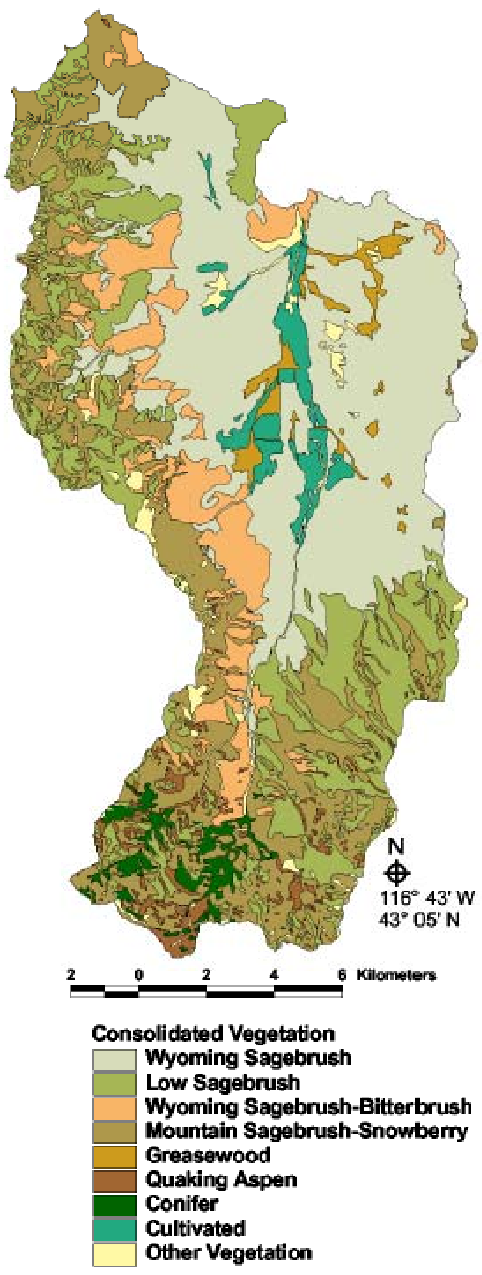

Figure 4-2: Map of consolidated vegetation taken from Seyfried et al. (2000a)

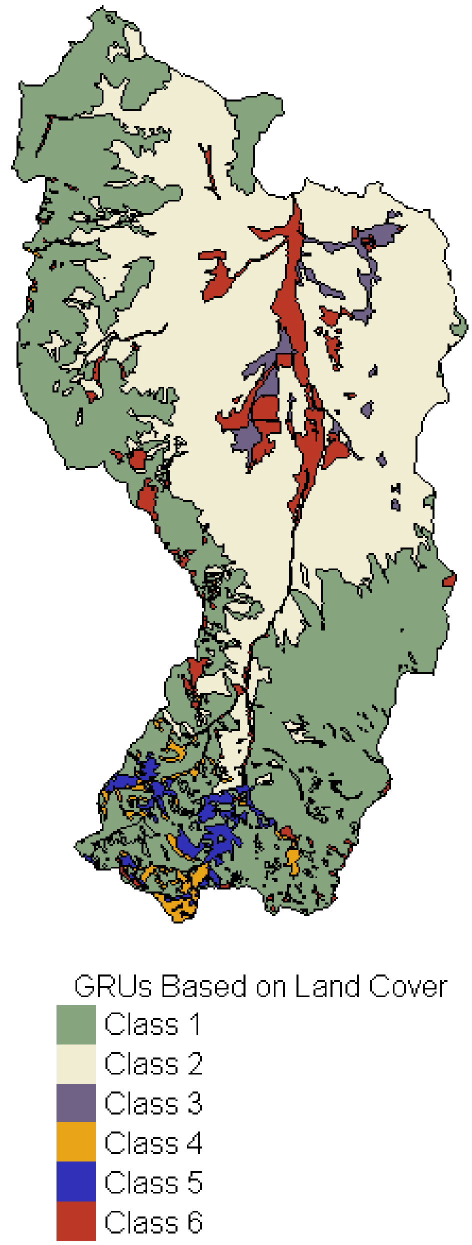

Figure 4-3: Consolidated 6 GRU land classes

Table 4-1: Land class descriptions and percent area for Reynolds Creek and Tollgate subwatershed

GRU Number	CLASS code	Area (%) Reynolds,	Tollgate	CLASS Descriptions	Vegetation Description (USDA)
1	2	48%	64%	Broad Leaf	Low Sagebrush Mountain Sagebrush-Snowberry
2	2	40%	15%	Broad Leaf	Wyoming Sagebrush Wyoming Sagebrush-Bitterbrush
3	1	2%	0%	Needle Leaf	Greasewood
4	2	2%	9%	Broad Leaf	Quaking Aspen
5	1	2%	8%	Needle Leaf	Conifer
6	3	6%	4%	Crops	Cultivated Other Vegetation

4.1.3 Parameter Estimates

The setup of any hydrological model requires the specification of initial estimates of model parameter values. With the MESH model, the modeler is responsible for specifying these values. Although experienced modelers can assign a set of initial parameter estimates without considering site specific measured data, the quality of such parameter sets is unknown. Furthermore, a poor parameter set could negatively influence the final model calibration results. Therefore, considerable effort was spent to assign reasonable, and if possible, site specific initial parameter values. The approach to estimating the initial values of all MESH model parameters are detailed in the following subsections. Copies of the MESH initial parameter input files are included in Appendix B.

4.1.3.1 Vegetation Parameters

Pairs of lysimeters were installed at two locations in the watershed, Lower Sheep Creek and Reynolds Mountain. The lysimeters at the Lower Sheep Creek site are dominated by low sagebrush and perennial bunchgrasses and forbs (Seyfried et al., 2000b). The lysimeters located at the Reynolds Mountain site are dominated by mountain sagebrush and other plants, including mountain snowberry, Idaho fescue and yarrow (Seyfried et al., 2000b). The leaf area index (LAI) was monitored several times every year at each lysimeter using the point quadrature method (Seyfried et al., 2000b). The results of this study are summarized in Table 4-2.

Table 4-2: Leaf Area Index, adapted from Seyfried et al. (2000b)

Lysimeter	Maximum Date	Maximum LAI	Minimum Date	Minimum LAI
Reynolds Mt. North	June 30	2.040	October 1	0.594
Reynolds Mt. South	July 3	1.780	October 4	0.441
Lower Sheep Creek East	May 23	1.570	September 20	0.296
Lower Sheep Creek West	May 22	1.500	September 21	0.352

Root depth values for each grouped response unit were compiled for the dominant land cover for each land class and are listed in Table 4-3 .

Table 4-3: Max root depth for the dominant plant cover

GRU	Dominant plant Cover	Max Root Depth (m)	Reference
Land Class 1	Mountain Sagebrush	1.23	Sturges et. al., 1978
Land Class 2	Wyoming Sagebrush	2.13	Sturges et. al., 1978
Land Class 3	Greasewood	18	Cooper et. al., 2006
Land Class 4	Quaking Aspen	2	Canadell et. al., 1996
Land Class 5	Conifer	1.5	Rushton, 2003
Land Class 6	Cultivated and Other	2	Neitsch et. al., 2005
Singel GRU	Sagebrush	1.23	Sturges et. al., 1978

If there was no data available then the default values from CLASS documentation were applied. Default values were assigned for the following parameters:

- natural logarithm of the roughness length;
- visible and near infrared albedo;
- standing biomass density (kg/m^2);
- minimum stomatal resistance; and
- coefficients governing the response of stomates to light, vapor pressure deficit, and soil water suction.

4.1.3.2 Hydrological and Soil Parameters

No data were available for most of the hydrological parameters. Default values were estimated using MESH documentation (Soulis, 2007) for the following parameters:

- drainage density;
- valley slope;
- manning's *n* for surface roughness;
- lateral k_{sat} at surface; and
- change in lateral conductivity.

The MESH model requires information regarding the soil characteristics of the watershed. This information is input into the model as percentages of sand, clay and organic material. These inputs are used to calculate the soil layer thermal and hydraulic properties (Verseghy, 2008).

The default soil representation in MESH of three soil layers is utilized for Reynolds Creek. The depths of the soil layers were determined using the available soils data for Reynolds Creek and is explained in further detail in Appendix C. The depths for the three soil layers (in meters) are: 0.0-0.18, 0.18-0.48, and 0.48-1.20.

Soil data for the watershed is available as soil survey maps at a 1:20,000 scale based on aerial photographs. The original soil map contained 30 soil series and 197 soil mapping units. However, the map was further consolidated for practical purposes (Seyfried et al., 2000a). The consolidated soils map is presented in Figure 4-4. As there is a large range of diversity in the watershed with respect to the soil conditions, it is very difficult to approximate the soil conditions on a land class basis.

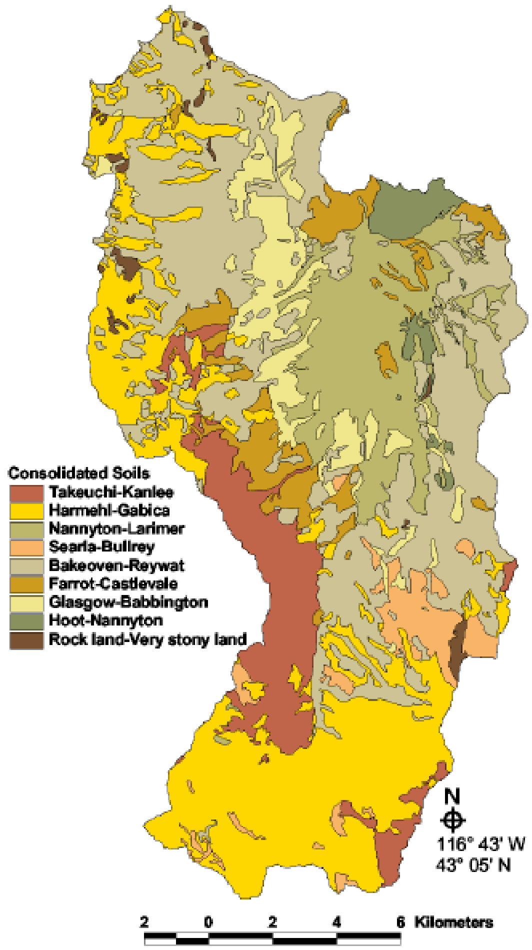

Figure 4-4: Consolidated soils for Reynolds Creek taken from Seyfried et al. (2000a)

To help better determine the approximate soil profiles for each land class, the land classes were isolated and overlaid with the soil types. The soil coverage for land classes 1 to 6 are presented in Figure 4-5 though to Figure 4-10. The corresponding soil types for each land class are listed in Table 4-4, Table 4-5, and Table 4-6.

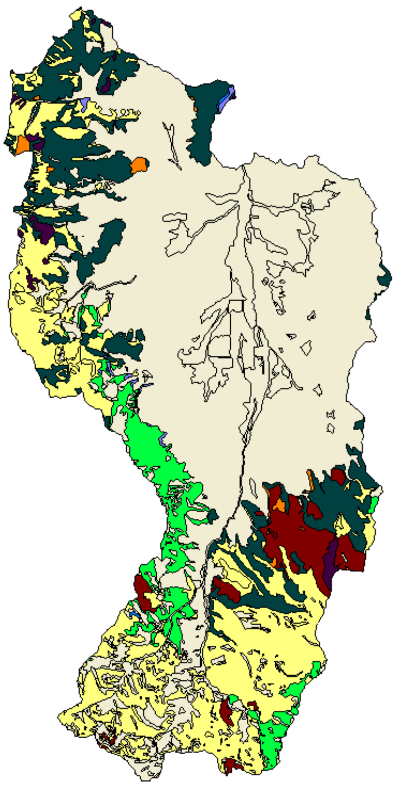

Figure 4-5: Soil types in Land Class 1

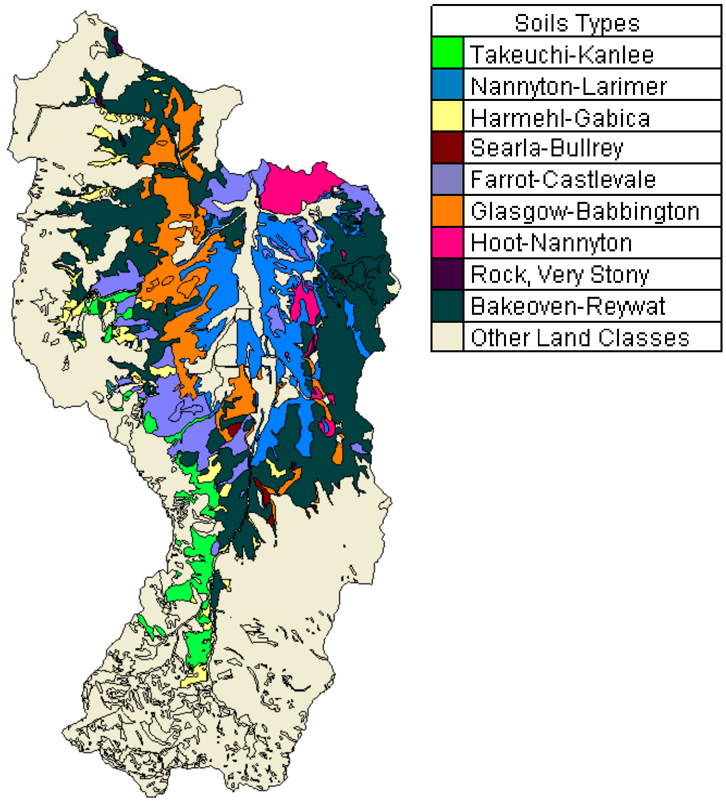

Figure 4-6: Soil types in Land Class 2

Table 4-4: Soil types in land classes 1 and 2

Land Class 1:	***Land Class 2:***
Harmehl-Gabica	Bakeoven-Reywat
Bakeoven-Reywat	Glasgow-Babbington
Takeuchi-Kanlee	Nannyton-Larimer
Searla-Bullrey	Farrot-Castlevale
Rock, Very Stony	Hoot-Nannyton
Glasgow-Babbington	Takeuchi-Kanlee
Farrot-Castlevale	Harmehl-Gabica
	Searla-Bullrey
	Rock, Very Stony

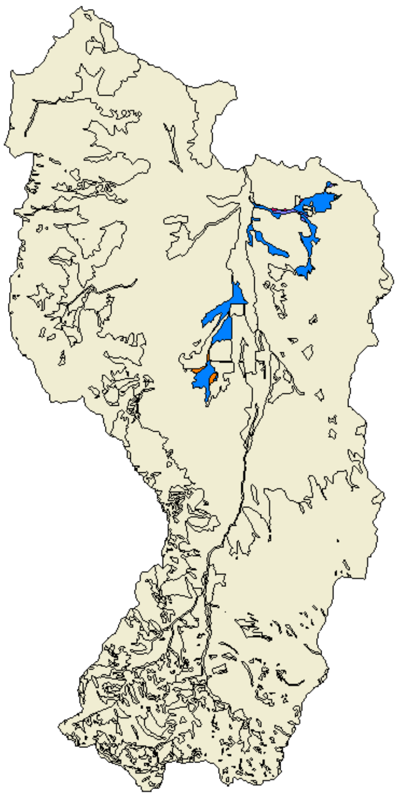

Figure 4-7: Soil types in Land Class 3

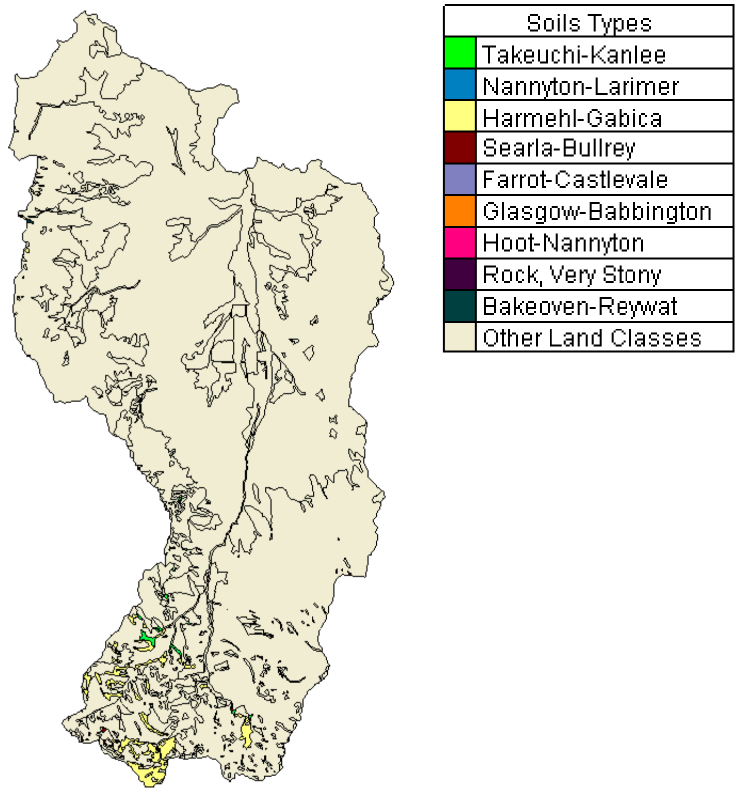

Figure 4-8: Soil types in Land Class 4

Table 4-5: Soil types in land classes 3 and 4

Land Class 3:	***Land Class 4:***
Nannyton-Larimer	Harmehl-Gabica
Glasgow-Babbington	Takeuchi-Kanlee

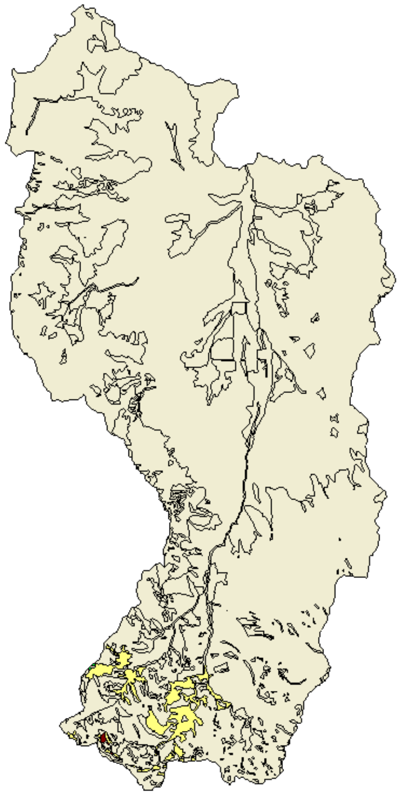

Figure 4-9: Soil types in Land Class 5

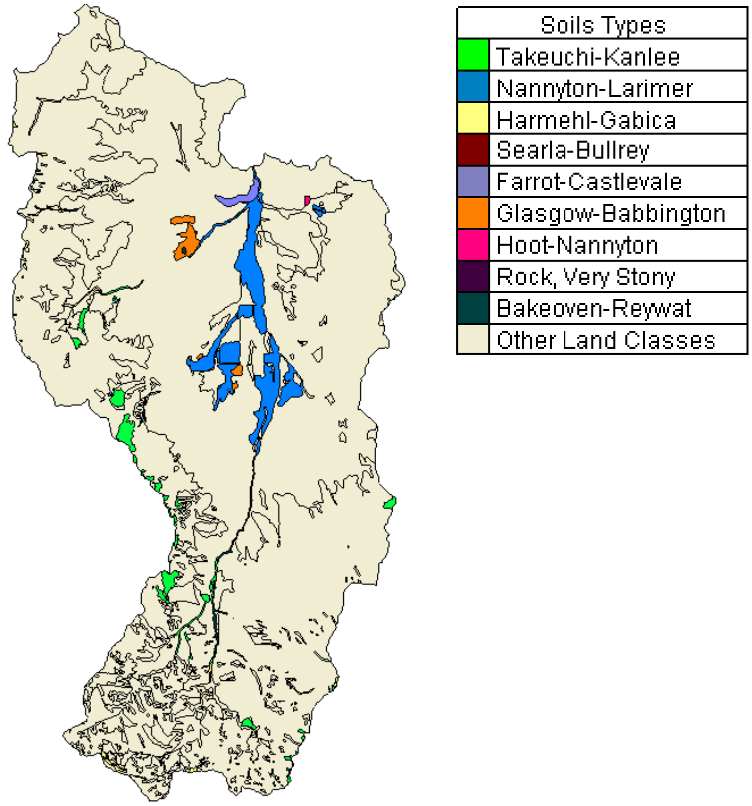

Figure 4-10: Soil types in Land Class 6

Table 4-6: Soil types in land classes 5 and 6

Land Class 5:	***Land Class 6:***
Harmehl-Gabica	Nannyton-Larimer
Searla-Bullrey	Takeuchi-Kanlee
	Glasgow-Babbington
	Farrot-Castlevale
	Hoot-Nannyton

The soil properties for the soil series in Reynolds Creek are listed in Appendix D. The majority of the soil series found in Reynolds Creek watershed are classified as a loam soil. The soil descriptions and the soil texture triangle were used to approximate the soil conditions and these are presented in Table 4-7. The permeable depth of the soil column was also estimated based on the soil descriptions. Some soil series in the watershed, particularly land classes 1, 4 and 5, are not very deep soils and quickly transition to fractured bed rock. By using a smaller permeable depth (e.g. 0.5 m) the soil column reduces the soil volume available for water storage in the individual GRU.

Table 4-7: Estimated soil conditions and permeable depth of the soil column

Land Class	Soil Layer (m)	Dominat Soil Series		Estimated Initial Conditions				
		Name	Discription	Sand (%)	Clay (%)	Organic (%)	Silt (%)	SDEPROW* (m)
Single GRU	0.00 - 0.18	Bakeoven-Reywat	Stoney Loam	40	25	5	30	1
	0.18 - 0.48		Loam	30	40	0	30	
	0.48 - 1.20		Sandy Loam	60	10	0	30	
1	0.00 - 0.18	Harmehl-Gabica	Gravelly Loam	35	30	5	30	0.5
	0.18 - 0.48		Sandy Loam	50	20	0	30	
	0.48 - 1.20		Fractured Bedrock	0	100	0	0	
2	0.00 - 0.18	Bakeoven-Reywat	Stoney Loam	40	15	5	40	1
	0.18 - 0.48		Loam	45	25	0	30	
	0.48 - 1.20		Sandy Loam	60	10	0	30	
3	0.00 - 0.18	Nannyton-Larimer	Fine Sandy Loam	65	10	5	20	1.2
	0.18 - 0.48		Sandy Loam	65	10	0	25	
	0.48 - 1.20		Sand	85	5	0	10	
4	0.00 - 0.18	Harmehl-Gabica	Gravelly Loam	40	20	10	30	0.5
	0.18 - 0.48		Clay Loam	40	30	0	30	
	0.48 - 1.20		Fractured Bedrock	0	100	0	0	
5	0.00 - 0.18	Harmehl-Gabica	Gravelly Loam	40	25	5	30	0.5
	0.18 - 0.48		Clay Loam	40	30	0	30	
	0.48 - 1.20		Fractured Bedrock	0	100	0	0	
6	0.00 - 0.18	Nannyton-Larimer	Fine Sandy Loam	65	25	5	5	1.2
	0.18 - 0.48		Sandy Loam	65	10	0	25	
	0.48 - 1.20		Sand	85	5	0	10	

*Permeable depth of the soil column

4.1.4 Initial Conditions

Like all other hydrological simulation models, MESH requires that initial conditions for model state variables (or prognostic variables) be specified. Since it is extremely difficult to accurately specify all initial state variables across the entire spatial domain of the model, the impact of inaccurate initial conditions on model predictions can be minimized by an initialization or spin-up period. By allowing the model a spin-up period before the calibration period, state variables can stabilize using observed meteorological and other data. Model performance is not assessed for the spin-up period. For example, the spin-up period would proceed the model calibration period where simulated results are compared to the measured data. The longer the spin-up, the less impact initial conditions have but the tradeoff is that the model simulation requires more computation time.

State variables in MESH were initialized on the basis of available data collected on or assumed to be representative of Sept. 2, 1986. This date was selected based on the available state variable data as well as the desire to minimize the necessary spin-up time of the model and thus maximize computational efficiency. Based on the calibration period starting on January 1, 1987 (as later discussed in Chapter 5) a spin-up period of 4 months was used. Measured soil temperature, soil moisture, and ambient air temperature data from multiple locations across the basin was used to specify many of the initial conditions. Using this late summer start date made it easier to identify initial conditions such as the snow pack density (assigned 0) and vegetation growth index (assigned to 1 as plants are still assumed to have all their leaves). As described in Section 4.1.3, soils in the MESH model are represented with three layers and each of these requires initial temperatures and soil moisture levels. The following subsections detail how measured data were translated into initial conditions for September 2, 1986.

4.1.4.1 Soil Temperature

The soil temperature was measured on an hourly or weekly basis at five stations throughout the watershed. On September 2, 1986 data were only available for two stations in the watershed, Probe 76x59 and Probe 127x07. Temperature values were measured at various depths for each station, as illustrated in Figure 4-11.

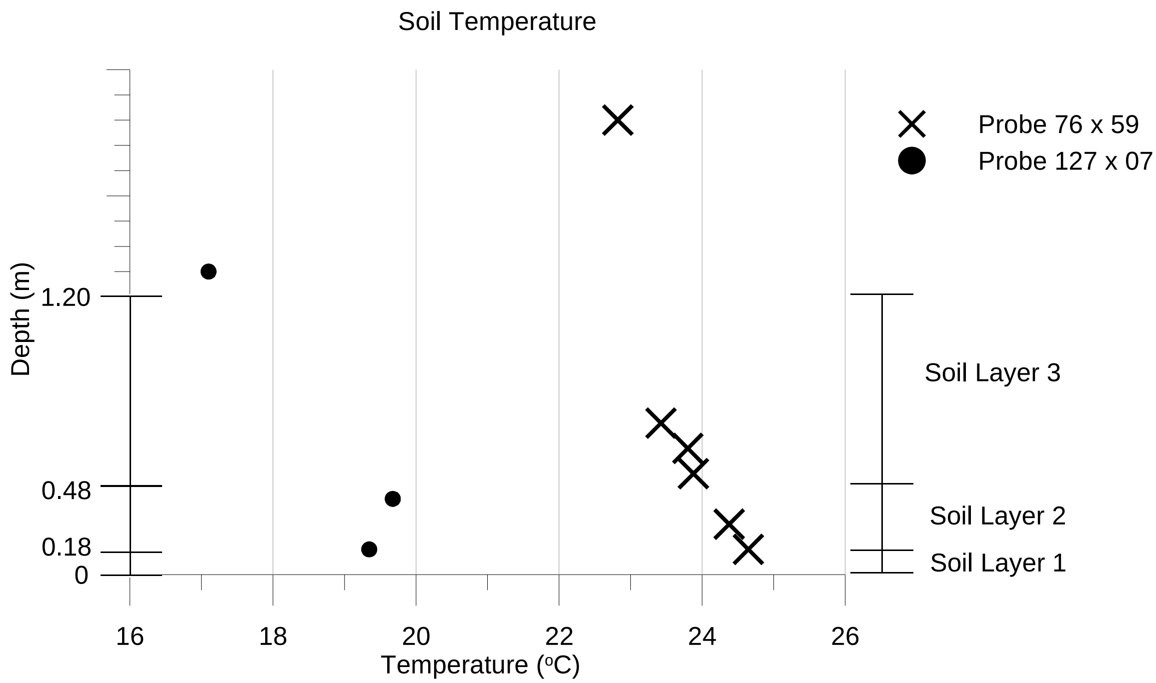

Figure 4-11: Soil temperature at Reynolds Creek, September 2, 1986

Based on the observed soil temperatures at each location and depth, the approximate initial soil temperature was estimated for each GRU based on location within the basin and depth. The single GRU model configuration initial temperatures were assigned based on the average values throughout the basin. The resulting initial soil temperatures are summarized in Table 4-8.

Table 4-8: Initial soil temperature for Reynolds Creek, September 2, 1986

Initial Conditions - Soil Temperature							
Depth	1 GRU	6 GRU					
	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
0.00 - 0.18	22.00	19.35	24.65	24.65	19.35	19.35	24.65
0.18 - 0.48	22.64	19.68	24.38	24.38	19.68	19.68	24.38
0.48 - 1.20	21.44	17.10	23.61	23.61	17.10	17.10	23.61

4.1.4.2 Soil Moisture

The soil moisture was measured with a neutron probe at various depths and locations though out the watershed several times though out the course of the year. On September 2, 1986 data were available from seventeen stations in the watershed. Soil moisture values are graphically represented with respect to depth for each grouped location in Figure 4-12.

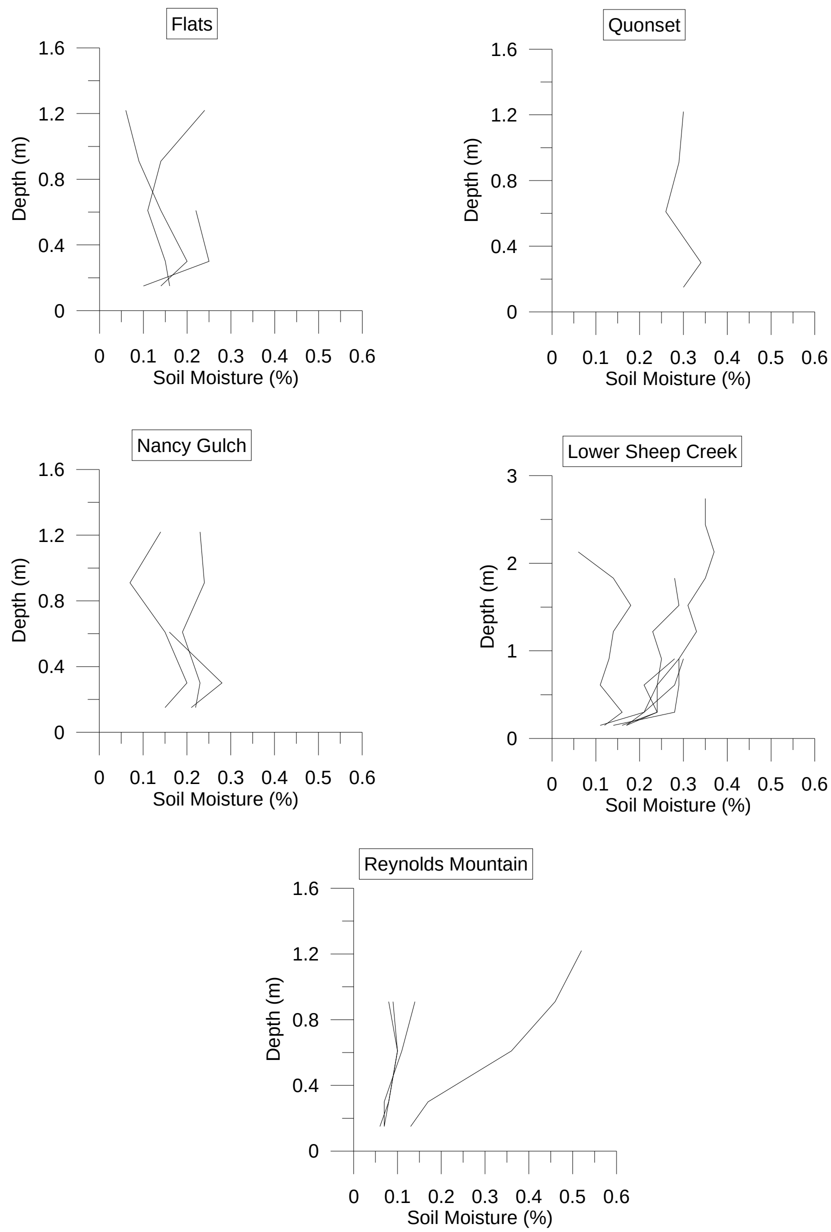

Figure 4-12: Measured soil moisture for sites at Reynolds Creek, September 2, 1986

Soil moisture was measured at the following depth intervals in meters: 0.15, 0.30, 0.61, 0.91, 1.22, 1.52, 1.83, 2.13, 2.44, and 2.74. Some assumptions had to be made to fit the soil moisture data into the three pre-defined soil depths.

1. The values at 0.15 are appropriate for the soil layer 1(0.0-0.18);
2. the values at 0.30 are appropriate for the soil layer 2 (0.18-0.48); and
3. the values at 0.61, 0.91 and 1.22 are appropriate for soil layer 3 (0.48 to 1.2 m).

Another important factor to be aware of is that the land classes were assigned based on a consolidated vegetative cover and this does not account for the spatial variability of the soil types. As a result the soil conditions are assumed the same though out each land class. The initial conditions were therefore determined based on geographic proximity and samples collected in the land class.

Two sites are located within Land class 1, Reynolds Mountain and Lower Sheep Creek. These sites have a total of ten neutron probe monitoring locations. One probe located in Reynolds Mountain, 176006, was not included because the observed values are much higher than other measurements taken in the area. This is not unexpected as there are many ground water springs located in the Reynolds Mountain and headwater areas of the watershed; however this measurement would not properly represent the soil moisture in the watershed. Land class 2 has two sites, Quonset and Nancy Gulch, located within the land class for a total of four neutron probe monitoring locations. Land class 6 has one site, Flats, located within the land class for a total of three neutron probe monitoring locations. There are no neutron probes located within land classes 3, 4 and 5. Land class 3 is presumed to be similar to land class 6 based on proximity. Land classes 4 and 5 uses the Reynolds Mountain site. The approximated initial soil moisture conditions for each land class are listed in Table 4-9.

Table 4-9: Estimated initial soil moisture for September 2, 1986

Initial Conditions - Soil Moisture							
Depth	Single GRU	6 GRU					
		Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
0.00 - 0.18	0.15	0.12	0.22	0.13	0.07	0.07	0.13
0.18 - 0.48	0.20	0.17	0.26	0.20	0.08	0.08	0.20
0.48 - 1.20	0.19	0.20	0.20	0.14	0.10	0.10	0.14

4.1.4.3 Other Initial Conditions

State variables not mentioned above were approximated based on the meteorological conditions for model initialization. Values for these parameters are given in Table 4-10.

Table 4-10: Parameters initialized based on meteorological values

Parameter	Value	Comments
The canopy temperature	15.1 °C	Average temperature for September 2nd, 1986
The temperature of the snow pack	15.1 °C	Average temperature for September 2nd, 1986
The temperature of ponded water	10 °C	Approximated based on average temperature
Depth of ponded water	0	No precipitation since August 23rd, 1986
Liquid water held on the vegetation canopy	0	No precipitation since August 23rd, 1986
Frozen water held on the vegetation canopy	0	No snow pack
Snow mass present on the ground	0	No snow pack
Albedo of the snow	0	No snow pack
Density of the snow	0	No snow pack
Vegetation growth index	1	Full leaf out

4.2 Wolf Creek

4.2.1 Meteorological Forcing Data

Meteorological forcing data for the Wolf Creek model was set up by Jonathan Bastien as part of his Masters of Applied Science work at the University of Waterloo. A brief description of the model setup, summarized from Bastien (2004), is provided in the following paragraph.

Meteorological data were collected at Wolf Creek from 1996 to the end of 2000. Five meteorological stations are present within the Wolf Creek watershed. All parameters necessary for the MESH model were measured except for net or incoming longwave radiation. Longwave radiation was calculated with the Cold Region Hydrological Model (CRHM) (Pomeroy et al., 2007) using shortwave radiation data and other measured meteorological data. Measures were taken by Bastien (detailed in Bastien (2004)) to insure that the meteorological data input into the model, including the longwave radiation, were reasonable.

4.2.2 Land Cover Data

Land cover data were classified using LANDSAT images. Bastien (2004) used eight primary land classes for the initial model setup. The land classes were tundra, buck brush (low-medium density alders and small shrubs with and understory of grasses and mosses), aspen forest, black spruce forest, mixed black spruce and buck brush forest, rock outcrop, permanent snowdrift, and water. In an effort to improve the computational efficiency and model calibration, the land classes were further consolidated in this research into four land cover types given in Table 4-11. A single land class model was also set up to be calibrated.

Table 4-11: Land classes used for Wolf Creek

GRU Number	CLASS code	Area (%)	CLASS Descriptions	Vegetation Description
1	4	15%	Tundra	Tundra
2	2	36%	Deciduous Broad Leaf	Buck Brush Aspen Forest
3	1	42%	Evergreen Needle Leaf	Black Spruce
4	5	7%	Rock/Water	Rock outcrop Permanent Snowdrift Water

4.2.3 Parameter Estimates

The initial parameter sets and initial conditions applied to this model of Wolf Creek were based on the final parameter sets defined by Bastien (2004). The model spin-up period for Wolf Creek was one year and is longer than the Reynolds Creek spin-up because initial condition data for Wolf Creek is more uncertain than Reynolds Creek. The initial parameter sets and initial conditions used for the single GRU model and the 4 GRU model are given in Appendix B.

5. Model Calibration

Model calibration is the perturbation of model parameters within reasonable ranges to improve the agreement between simulated model predictions and measured data for the system being modelled. Calibration may involve manual methods such as trial and error, or automatic calibration procedures that use an optimization algorithm. The work completed for this thesis uses an automatic calibration procedure. The MESH model calibration methodology for the Reynolds Creek and Wolf Creek case studies are described in the following sections.

Reynolds Creek was set up in two model configurations. The purpose of this was to determine how the model performs in calibration and validation with multiple GRUs versus a single generalized GRU. The configurations for Reynolds Creek are:

1. 2 km grid cells and 1 GRU
2. 2 km grid cells and 6 GRUs

Wolf Creek was used as a simplified case study and setup with two model configurations similar to Reynolds Creek. The configurations for Wolf Creek are:

1. 30 km grid cells and 1 GRU
2. 30 km grid cells and 4 GRUs

All MESH model simulations for Reynolds Creek and Wolf Creek were performed on SHARCnet (Shared Hierarchical Academic Research Computing Network), a parallel computing facility. This was done to decrease total computational time by taking advantage of the thousands of processors available through SHARCnet. The executable files and binary files for the modes were recompiled on the SHARCnet system (a LINUX operating system).

As part of this thesis work, a set of parameter ranges were established for MESH model calibration. The parameter ranges were assigned based on the default values defined in Verseghy (2008) and experience calibrating the MESH model. Parameter ranges will have an impact the results of automatic calibration procedure; this is further discussed in Section 6.3. A previously proposed set of parameter ranges and a new refined set of parameter ranges are given in Appendix E.

Model calibration was focused on improving the agreement between simulated and measured streamflow and/or SWE time series data. The objective function selected to quantify the quality of agreement between simulated and measured SWE and streamflow was the Nash–Sutcliffe efficiency measure (Section 2.2). It was selected because it is common measure of model performance in hydrology that can be applied to both streamflow and SWE time series. The Nash–Sutcliffe measure is a sum of squares based efficiency measure that is to be maximized with an optimal value of 1.0 and negative values indicate very poor model predictive quality. It can generally measure simulation quality in terms of the shape and volume of the hydrograph; however it places large emphasis on peak events.

All model automatic calibrations experiments were performed using the DDS optimization algorithm (Tolson and Shoemaker, 2007). DDS is well suited for optimization problems with a large number of calibration parameters, such as a distributed watershed model. DDS was designed specifically for the automatic calibration problem and the algorithm is able to rapidly converge to a good calibration solution and easily avoids poor local optima (Tolson and Shoemaker, 2007).

5.1 Reynolds Creek Model Calibration

5.1.1 Monitoring Data

Reynolds Creek is described by a very detailed measured data set containing different long-term time series that each could be used to assess model prediction quality. Two of these types of time series data, streamflow and SWE, were selected as the basis for model evaluation in this case study as they are important indicators of model performance. The details of these objectives are summarized in the following sections.

5.1.1.1 Streamflow

Four streamflow gauge locations (Reynolds Creek at Tollgate, Reynolds Mountain East, Reynolds Creek Outlet and Salmon Creek) were utilized to compare simulated and measured streamflow. Figure 3-3 shows the locations of these gauges. For the purpose of model calibration, only two streamflow locations were used; Reynolds Creek at the Tollgate weir and Reynolds Mountain East. These sites were selected for calibration because they were not impacted streamflow diversions for flood irrigation. During the melt period and summer months, streamflow from Reynolds Creek was often used for crop irrigation of fields located below the Tollgate weir (Perison et al., 2001). Within the Reynolds Creek watershed there are approximately 690 ha of pasture, hay, and grain crops that received water through flood irrigation (Johnson and Smith, 1979). The Salmon Creek subwatershed also contains approximately 36 ha of flood irrigated pasture land (Perison et al., 2000). The amount of water diverted each year from Reynolds Creek is not directly measured. A study performed by

Johnson and Smith (1979) suggests that on average 24% of the total flow is diverted between the months of March and September. During a dry year, or a low flow year, it was estimated that up to 74% of the total flow could be diverted between the months of March and September. As a result greater flow volumes can occur at the Tollgate weir (approximate drainage area of 54.4 km^2) then at the outlet of Reynolds Creek (approximate drainage area of 238.6 km^2) during the low flow years. This diversion of flows will cause some problems if the model was calibrated to flows at the Reynolds Creek Outlet since these significant diversion volumes are unknown. The Reynolds Creek Outlet and Salmon Creek sites were incorporated during post-calibration prediction quality assessment (or the model validation period).

5.1.1.2 Snow Water Equivalent

SWE was calibrated using data from four snow course locations and one snow pillow (previously presented in Figure 3-4). Data for each snow course site was measured approximately every two weeks during the snow season, while daily SWE measurements were available for the snow pillow site.

The four snow course sites were selected based on their distribution within the watershed and their proximity to precipitation gauges. Due to the extreme mountain topography, precipitation can vary significantly within a few kilometres in the higher elevations. This is illustrated in Figure 5-1 for the 1986-1987 and 1987-1988 snow seasons. In the higher elevations of the watershed the annual precipitation totals that can vary from 700 mm to 1100 mm. Therefore, to accurately simulate SWE monitored at point locations, it is very important that the precipitation inputs are accurate. All four snow course sites selected contain a precipitation gauge within 2 km of the snow course. The snow pillow is located within 1km of a precipitation gauge and a meteorological station.

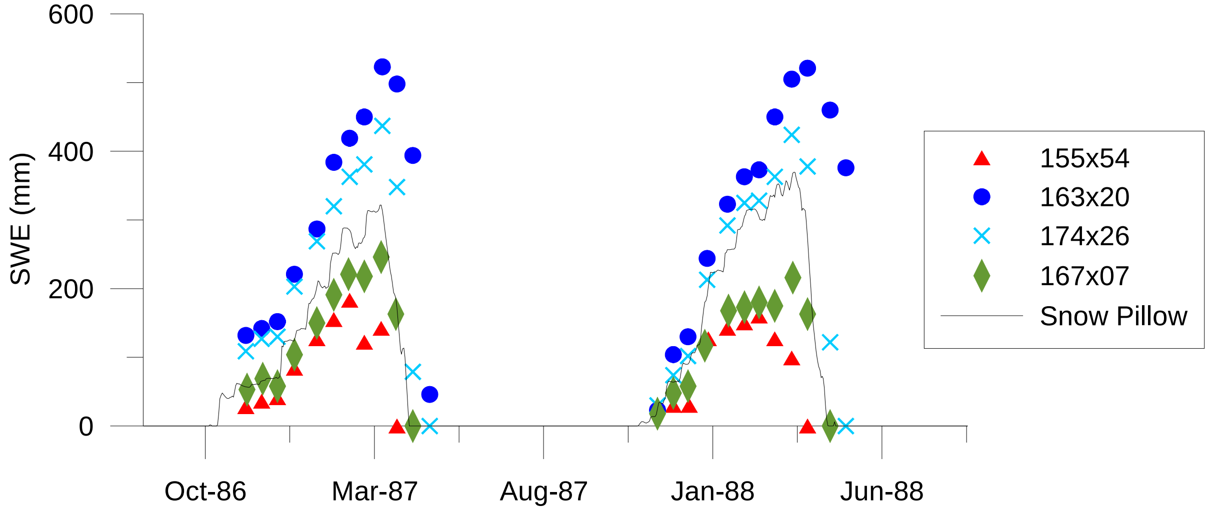

Figure 5-1: Observed SWE for 1987 and 1988 snow seasons

When snow pillows were first introduced in the early 1960's, they were generally considered quite unreliable (Marks et al., 2000). As a result, the snow pillow site was established in conjunction with a snow course site (176x07) so that the snow course data could be used to validate the snow pillow measurements. A comparison of the snow pillow data and the corresponding snow course data for the calibration years is presented in Figure 5-2 and shows that the two SWE measurement strategies are consistent for the calibration years. Therefore, the daily snow pillow data were utilized for model calibration instead of the snow course data at site 176x07.

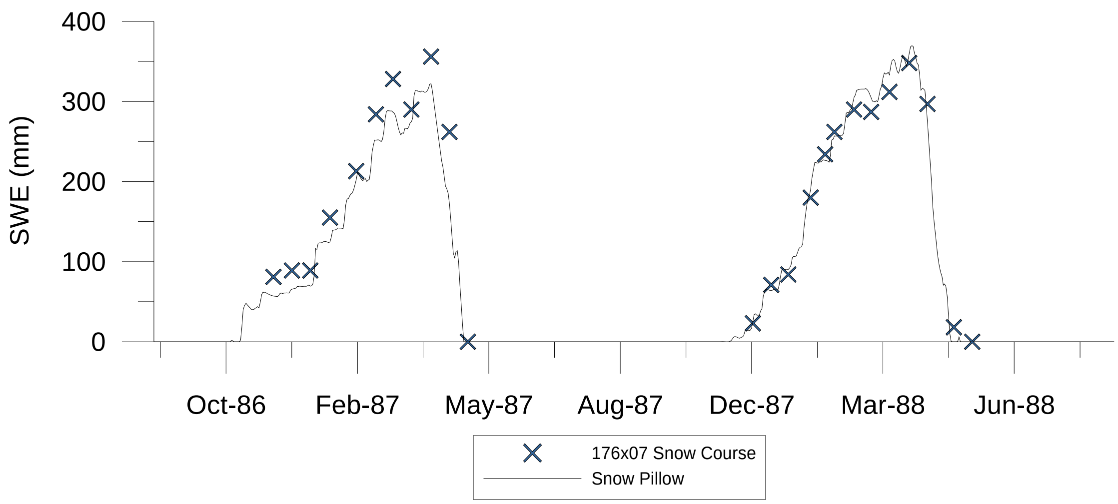

Figure 5-2: Snow pillow data and snow course data for 1987 and 1988

5.1.2 Spin-up and Calibration Periods

The simulation period for Reynolds Creek begins on September 2nd, 1986 and runs to December 31st, 1988. Because the model was initialized with measured data, a relatively short spin-up period was used to reduce the computational time of the model. Model spin-up took place from September 2nd, 1986 to December 31st, 1986. Model calibration took place during the 1987 and 1988 calendar years. The model validation period (post-calibration performance) covers the 1989 to 1996 calendar years.

5.1.3 Calibrated Parameters

5.1.3.1 Streamflow Parameters

Streamflow is an aggregated hydrological response and therefore, model parameters impacting any hydrologic process can impact simulated streamflow. For each GRU the vegetation, hydrology and soil parameters as discussed in Chapter 4 were calibrated independently. This means that when the same parameter was calibrated in multiple GRUs, the value for each GRU could be varied without simultaneously modifying values in other GRUs.

For the single GRU configuration of Reynolds Creek, thirty-one vegetation, hydrology and soil parameters were calibrated. For the 6 GRU configuration, sixty-two vegetation, hydrology and soil parameters where calibrated in the two dominant GRUs. This was done to reduce the overall number of calibration parameters. The GRUs calibrated, GRU 1 and GRU 2, contain approximately 88% of the land area for Reynolds Creek watershed and 80% of the land area for Tollgate subwatershed. Both GRUs are dominated by different species of sagebrush.

5.1.3.2 Snow Water Equivalent

The SWE measurement locations chosen for calibration were mapped to the appropriate GRU and grid cell model output to compare simulated results to measured data. The GRU type assigned to each SWE location was determined by overlaying the GRU map with snow survey locations. The results of this process are summarized in Table 5-1.

Table 5-1: Snow survey locations and GRUs

Site	Location		GRU	GRU Description
	Easting	Northing		
155x54	517892	4770341	5	Conifer
163x20	514041	4769438	1	Low Sagebrush
167x07	521613	4769718	1	Low Sagebrush
174x26	516719	4767777	4	Quaking Aspen
Snow Pillow	520055	4768117	4	Quaking Aspen

Not all model parameters impact SWE simulation results in MESH. For calibrating SWE data, only parameters that affect SWE, as given in Dornes et al. (2008) and summarized in Table 5-2 were applied. Based on the locations of the snow survey sites, three GRUs (GRU 1, GRU 4 and GRU 5) were optimized. A total of thirty-nine parameters were optimized for calibration runs focused only on SWE. The parameters optimized consist of twelve vegetation parameters that impact the energy balance and one hydrological parameter that limits snow depth.

Table 5-2: Calibrated SWE parameters

SWE Parameters	
LAMXROW	Maximum leaf area index
LAMNROW	Minimum leaf area index
LNZ0ROW	Natural logarithm of the roughness length
ALICROW	Near infrared albedo
ALVCROW	Visible albedo
CMASROW	Standing biomass density ($kg \cdot m^{-2}$)
RSMNROW	Minimum stomatal resistance
QA50ROW	Coefficient governing the response of stomates to light
VPDAROW	Coefficient governing the response of stomatal resistance to vapour pressure deficit
VPDBROW	Coefficient governing the response of stomatal resistance to vapour pressure deficit
PSGAROW	Coefficient governing the response of stomatal resistance to soil water suction
PSGBROW	Coefficient governing the response of stomatal resistance to soil water suction
ZSNLROW	Limiting snow depth

5.1.4 Calibration Strategies

For both sets of selected calibration data, SWE and streamflow, multiple sites were used for model calibration. To optimize these multi-site calibrations simultaneously, three different calibration strategies were applied:

- single objective streamflow, with two streamflow monitoring locations;
- single objective SWE, with five SWE monitoring locations; and
- single objective streamflow and SWE (above two combined).

To calibrate a multi-site streamflow problem using a single objective function, a measured volume weighted average of the two Nash-Sutcliffe values was used. The volume-weighted average was used instead of a regular average to account for the differences in the magnitude of flows due to size differences of the watersheds for the two sites. By using the weighted average for these two sites approximately 99% of the objective function value is based on the Tollgate site (mean annual streamflow 0.424 m^3/s) and 1% of the objective function value is based on the Reynolds Mountain East site (mean annual streamflow 0.00671 m^3/s) for the calibration period.

To calibrate the multi-site SWE values, the Nash-Sutcliffe value was calculated for each of the five sites. As each SWE site represented a point measurement (rather than an aggregated response like streamflow) all five sites were given equal importance. Thus, the average of the five Nash-Sutcliffe values was defined as the objective function for SWE.

To calibrate both objective functions (streamflow and SWE) simultaneously, the objective function was defined as the average of the overall streamflow and overall SWE objective functions. The goal of calibrating both objectives equally was to try and improve the overall accuracy of the calibrated parameter set. By focusing on multiple objectives if is less likely that the algorithm will converge on an incorrect parameter set.

These three calibration strategies are methods a hydrologist might utilize to come up with a single calibrated model parameter set (depending on their modeling objectives). The differences in calibrated model prediction quality achieved under each objective are discussed in Chapter 6.

5.2 Wolf Creek

5.2.1 Monitoring Data

One streamflow gauging station, Wolf Creek at the Alaska Highway, was used for the calibration of Wolf Creek. No other objective functions, such as SWE, were applied for this study of Wolf Creek.

5.2.2 Spin-up and Calibration Periods

The streamflow data for Wolf Creek was divided into a calibration period and validation period. The calibration took place during the 1999 and 2000 calendar years. The 1998 calendar year was used as a one year spin-up period. A longer spin-up period was used for Wolf Creek because no information regarding the initial soil moisture and temperature was available. Model validation was performed with the 1997 and 1998 calendar years with 1996 as a one year spin-up period. The 1997 calendar year was selected for model validation and not model calibration because of a break in an ice dam during the melt period that caused abnormally large flows (Soulis, 2009). The observed flows for Wolf Creek during the model simulation period are presented in Figure 5-3.

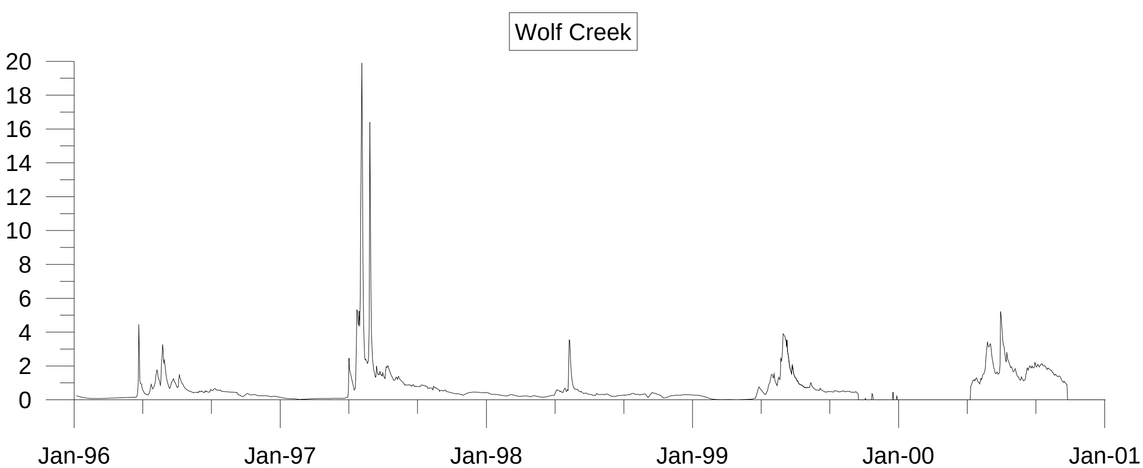

Figure 5-3: Hydrograph for Wolf Creek (1996-2000)

5.2.3 Calibrated parameters

As with Reynolds Creek vegetation, hydrology and soil parameters were calibrated for streamflow at Wolf Creek. For the single GRU configuration of Wolf Creek, 31 vegetation, hydrology and soil parameters were calibrated. For the 4 GRU configuration of Wolf Creek, 62 vegetation, hydrology and soil parameters were calibrated. Only parameters in two of the four GRUs were calibrated to reduce the overall number of parameters calibrated. The GRUs calibrated, GRU 2 and GRU 3, were selected for calibration as they cover approximately 78% of the land area at Wolf Creek.

6. Model Calibration Experiments

As the MESH model is still in the development phase, a critical component of the model development is a thorough analysis of the model setup alternatives, model performance and the limitations of the model. In practice, after all model setup assumptions are determined and the calibration strategy is selected, then the calibration procedure would be completed once to identify a calibrated parameter set. The model calibration experiments outlined in this chapter are designed to assess the impact that different model set-up assumptions and different calibration objectives have on the model predictions.

The design of each of the experiments is impacted by a few factors. First, each of the automatic calibration problems being solved is a difficult, nonlinear, multi-modal, black-box type of optimization that requires a global optimization method. DDS is a stochastic global optimization algorithm and thus the final answer provided by DDS is not guaranteed to be optimal and can change based on the random numbers utilized throughout the search. Secondly, these automatic calibration problems are computationally expensive. In practice, the calibration of a model like MESH would be conducted under a limited computational budget where the modeler dictates the optimization algorithm duration, instead of waiting for the algorithm to stop when it has found the local optimum. Each of these factors suggests the identification of reliable differences between calibration results under different calibration assumptions or objectives must be based on calibration experiments with multiple samples. Here, multiple samples with different random seeds of the same calibration experiment are referred to as optimization trials. For most calibration problems five trials were conducted.

Many hydrological models have an extensive range of inputs that include meteorological data, topography, land use data, hydrological gauging data, and parameter inputs. Reynolds Creek has had a substantial amount of data collected at the site and is an excellent case study for calibration experiments requiring extensive, high quality data sets. Calibration experiments that did not require as much data were performed using the Wolf Creek model. Although Wolf Creek is also a research watershed not all of the existing data were used for these experiments, Wolf Creek is being used as a simplified case study in this thesis.

Each of the following subsections (6.1 to 6.6) presents an experiment or calibration strategy. Each subsection first introduces a method and purpose and then includes the corresponding results and discussion. Each set of model experiments was designed to assess the performance of the MESH model, provide some feedback on the best approach for model setup and compare model calibration strategies. Where applicable, the experiment results are compared to the baseline calibration, results in section 6.1 which were expected to produce the best calibration results.

The baseline experiments were selected as the basis for model validation, which was only performed using the best calibration results (parameter sets). Model validation with respect to SWE simulation results are described in Section 6.6.

6.1 Initial Model Calibration

6.1.1 Calibration Strategy

A series of model calibration runs were initially performed to be used as the "baseline" experiment. Other calibration configurations are compared to this baseline experiment. These calibrations were considered the baseline results because they were designed with the goal of achieving the best possible single objective model performance for each model spatial configuration considered. For example, preliminary model runs indicated that refined model parameter ranges (as defined in Appendix E) along with estimated set parameter defined in Appendix B (rather than a randomly chosen set of values). The baseline model focused on streamflow calibration only and the details of the calibration procedure are described in Section 5. The model calibration experments and configurations for Wolf Creek and Reynolds Creek are summarized in Table 6-1 and Table 6-2, respectively.

Table 6-1: Initial model calibration for Wolf Creek

Configuration	30km 1 GRU	30km 4 GRU
GRUs Calibrated	Single GRU	GRU 2 and 3
No. Parameters Calibrated	31	62
Parameter Ranges	Refined	Refined
Initial Solution	Estimated Parameters	Estimated Parameters
Objective Calibrated	Streamflow at Outlet	Streamflow at Outlet
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000, 10000	1000, 5000
Trials	5	5

Table 6-2: Initial model calibration for Reynolds Creek

Configuration	2km 1 GRU	2km 6 GRU
GRUs Calibrated	Single GRU	GRU 1 and 2
No. Parameters Calibrated	31	62
Parameter Ranges	Refined	Refined
Initial Solution	Estimated Parameters	Estimated Parameters
Objective Calibrated	Streamflow at Tollgate	Streamflow at Tollgate
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000
Trials	5	5

The baseline experiments for Wolf Creek were run with 1,000 and 10,000 objective function evaluations for the single GRU configuration and 1,000 and 5,000 objective function evaluations for the 4 GRU configuration. The purpose of this approach was to determine how the computational budget impacts the overall results of the calibration.

For both watersheds, two different GRU configurations were used to determine if increasing the detail of the GRUs will improve model performance.

When the baseline 2 km 6 GRU configuration of Reynolds Creek was set up on the UNIX platform, some inconsistencies in the model developed, likely because of the FORTRAN compiler used on the UNIX platform. The 2km 6 GRU model configuration was terminating in error during approximately 40% of the model calibration runs performed during an initial trial. The error was the result of a numerical error in the water balance. This indicated that the model was having difficulty possibly due to the increased complexity of multiple grid, 6 GRUs configuration. As this was the first time the MESH model was tested and run at this discretization on the Reynolds Creek case study, some simplifications were made to the soil parameters to temporarily resolve the water balance issues. The soils in the uncalibrated land classes were modified so that GRUs 3 to 6 had the same values for soil parameters (despite their differences in the measured data). This reduced the model crash rate to less than 10% and thus enabled model calibration to proceed. This issues requires a more in depth investigation by model developers.

6.1.2 Results and Discussion

6.1.2.1 Model Calibration

The results, presented inTable 6-3 and Table 6-4 are set up to include a brief snap shot of each set of model calibration results for Wolf Creek and Reynolds Creek, respectively. Each trial had an initial Nash-Sutcliffe value associated with it and over the course of automatic model calibration this value improved. The calibration results are summarized in terms of best and worst Nash-Sutcliffe values (to show the range of each of the trials) and then the average Nash-Sutcliffe calibration result for all of the trials.

The calibration results for the baseline Wolf Creek models are summarized inTable 6-3. The Wolf Creek single GRU configuration was able to produce good results after 10,000 function evaluations (mean Nash-Sutcliffe value, 0.76), but could not consistently produce good results after 1,000 function evaluations (mean Nash-Sutcliffe value, 0.66 but values as low as 0.50). The results for the 4 GRU configuration of Wolf Creek did improve from the initial solution with calibration, however after 1,000 function evaluations the model was not able to produce good results (mean Nash-Sutcliffe value, 0.31). After 5,000 function evaluations of the 4 GRU configuration of Wolf Creek had a mean Nash-Sutcliffe value of 0.42, a very poor result. It was expected that the 4 GRU spatially distributed model should be able to perform as well if not better then the lumped single GRU model configuration. In an attempt to resolve this issue the 4 GRU configuration was re-calibrated with all four GRUs being calibrated instead of two GRUs. There was no improvement in model performance when four GRUs were calibrated and therefore no results were reported. As the model was calibrated with a sufficient number of function evaluations to produce equal or better results than the single GRU configuration, it is possible that the problem is a result of the GRU delineation. Future work should be conducted to further investigate this issue.

Table 6-3: Initial calibration results for various Wolf Creek model configurations

Watershed	Wolf Creek	Wolf Creek	Wolf Creek	Wolf Creek
Configuration	30km 1 GRU	30km 1 GRU	30km 4 GRU	30km 4 GRU
Function Evaluation	1000	10000	1000	5000
Trials	5	5	5	5
Initial NS Value	-3.66	-3.66	-5.56	-5.56
Calibrated NS Value				
Worst	0.50	0.72	0.09	0.37
Best	0.72	0.82	0.40	0.46
Average	0.66	0.76	0.31	0.42

The calibration runs for Wolf Creek were performed using different numbers of function evaluations (1,000, 5,000 and 10,000) to determine the benefits of using a greater number of objective function evaluations. The results for Wolf Creek with 10,000 and 5,000 function evaluations were considerably better than the results for Wolf Creek with only 1,000 function evaluations. Therefore, findings in the rest of the thesis may be limited by the fact that with a limited computational budget (i.e. 1000 or less function evaluations) the calibration solutions returned by DDS could be significantly worse than the true but unknown optimal solutions.

Wolf Creek is less computationally intensive problem compared to Reynolds Creek and was therefore tested with larger number of function evaluations. This same procedure was not able to be repeated with the Reynolds Creek case study due increased computational time required. The computational time required to run the 1 year spin-up period and 2 year calibration period for the single GRU configuration of Wolf Creek takes 12 seconds per function evaluation (or one model simulation) and 29 seconds for the 4 GRU configuration. In comparison, to run the Reynolds Creek model for the 4 month spin-up and 2 year calibration period takes 75 seconds for the single GRU configuration and 175 seconds for the 6 GRU configuration. The 6 GRU configuration with SWE is the most computationally expensive set up and takes 198 seconds per function evaluation.

The calibration results for the baseline Reynolds Creek models are summarized in Table 6-4. The baseline set of model calibration yielded very good results for the single GRU (mean Nash-Sutcliffe value, 0.75) and 6 GRU (mean Nash-Sutcliffe value, 0.78) configurations of Reynolds Creek. Both model configurations performed quite well during the model calibration periods. The 6 GRU version had slightly better results than the single GRU version during the calibration runs

(Nash–Sutcliffe of 0.82 for the single GRU and 0.85 for 6 GRU configuration). However, the 6 GRU configuration generally had more stability issues resulting in the model frequently terminating in error before the end of the calibration period.

Table 6-4: Initial calibration results for Reynolds Creek

Watershed	Reynolds Creek	Reynolds Creek
Configuration	2km 1 GRU	2km 6 GRU
Objective Calibrated	Streamflow	Streamflow
Function Evaluations	1000	1000
Trials	6	6
Initial NS Value	0.08	0.20
Calibrated NS Value		
Worst	0.71	0.72
Best	0.82	0.85
Average	0.75	0.78

In the Reynolds Creek case study the two dominant and calibrated GRUs were both various species of sagebrush. The single GRU configuration of Reynolds Creek may be performing very similarly to the 6 GRU configurations due to the large percentage of land area covered by sagebrush. In other words, achieving the benefit of increased spatial detail is only possible under careful and representative GRU descriptions.

The results for the 6 GRU configuration of Reynolds Creek had slightly better results than the single GRU configuration, however there is no conclusive evidence to say that one configuration performed better than the other. It is encouraging that both model configurations were able to achieve comparable results with the same computational budget but twice as many parameters were calibrated in the 6 GRU configuration. Only 31 parameters were calibrated for the single GRU configuration and 62 parameters were calibrated for the 6 GRU configuration.

6.1.2.2 Reynolds Creek Streamflow Validation

For the single GRU configuration of Reynolds Creek, validation runs were performed for Tollgate, Reynolds Mountain East, Reynolds Creek Outlet and Salmon Creek. The model was unable to run continuously from September 2nd, 1986 until September 30th, 1996 and terminated due to an error in September 1993 when only running the model for Tollgate and Reynolds Mountain East. When the model was run with Salmon Creek and Reynolds Creek Outlet an error occurred in May, 1989 and the model was re-run starting on September 2, 1990. The model again terminated with an error in September, 1993. The hydrographs for the calibration and/or validation of Reynolds Mountain East, Tollgate, Salmon Creek, and Reynolds Creek Outlet for the single GRU model configuration are presented in Figure 6-1 to Figure 6-4.

Model performance statistics (Nash-Sutcliffe (NS), absolute percent bias (APB), and coefficient of determination (R^2)) were calculated for the calibration and validation periods for the single GRU configuration of the Reynolds Creek model at Tollgate and Reynolds Mountain East subwatersheds. Statistics were not calculated for Salmon Creek and Reynolds Creek Outlet due to model stability issues. The statistical results are presented in Table 6-5 and will be discussed with the corresponding hydrographs.

Table 6-5: Summary of annual statistics for Reynolds Creek, 2km 1 GRU model configuration

	Reynolds Creek Watershed					
	Tollgate			Reynolds Mountain East		
	NS	APB	R^2	NS	APB	R^2
1987	0.79	21%	0.90	0.60	51%	0.90
1988	0.86	21%	0.94	0.54	54%	0.86
1989	0.73	4%	0.92	0.75	26%	0.88
1990	0.84	4%	0.94	0.72	29%	0.86
1991	0.55	22%	0.84	0.61	24%	0.80
1992	0.57	3%	0.92	0.65	23%	0.83
1993	0.26	16%	0.86	0.85	11%	0.94
1987-1993	0.64	4%	0.89	0.81	20%	0.90

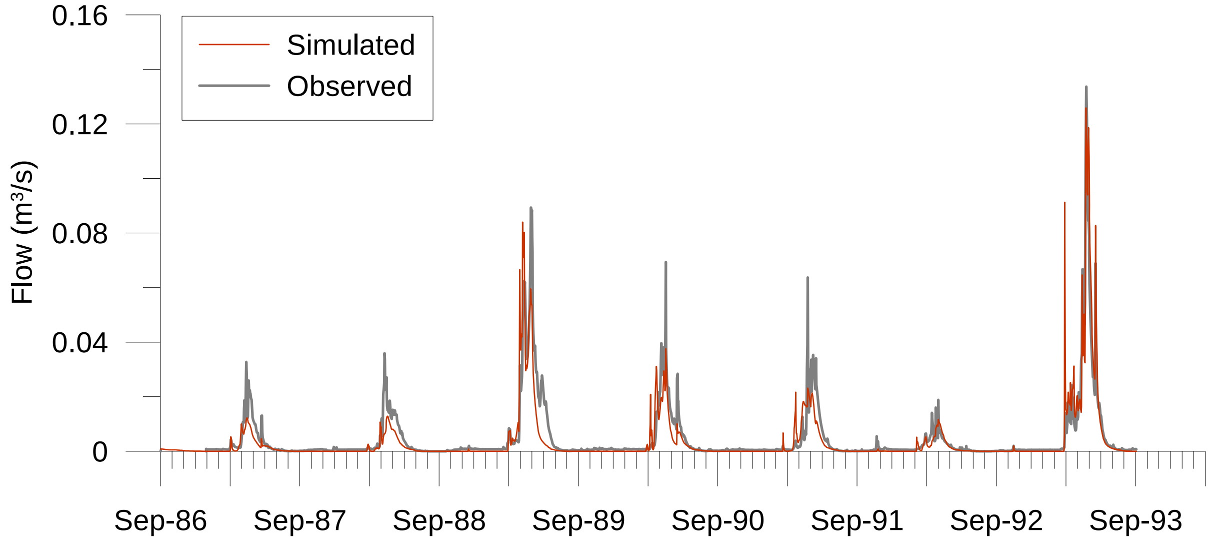

Figure 6-1: Calibration (1987-1988) and validation (1989-1993) for Reynolds Mountain East, 2km 1 GRU

Reynolds Mountain produced very good results given the low flow regime of the subwatershed. The calibration period (1987-1988) consisted of two low flow years. Both calibration years were able to closely match the shape of the hydrograph (R^2 = 0.90 and 0.86) however the overall discharge was underestimated (APB = 51% and 54%). Even though the calibration period used low flow years, in the validation period the high flow years (1989 and 1993) were not under estimated and produced some of the better results (NS = 0.75 and 0.85, respectively). The other validation years (1990, 1991, 1992) had very good timing and shape however the simulated results did not match the peaks of the hydrograph.

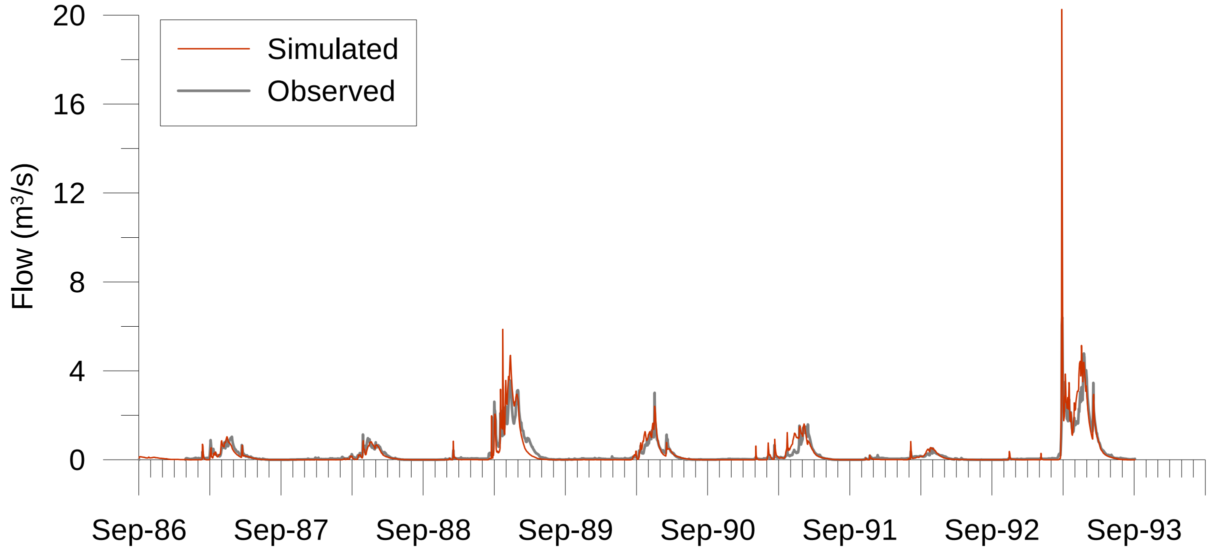

Figure 6-2: Calibration (1987-1988) and validation (1989-1993) for Tollgate, 2km 1 GRU

Reynolds Creek at Tollgate produced very good results with respect to the shape, timing and peaks of the hydrograph. The calibration period (1987-1988) consisted of two low flow years. Both calibration years were able to closely match the shape of the hydrograph (R^2 = 0.90 and 0.94) and the overall discharge (APB = 21% and 21%) resulting in very good results for the calibration period (NS = 0.79 and 0.86). The validation years (1989 to 1993) were generally able to capture timing for the melt period of the hydrograph; however the volume and peaks did not match as well as the calibration period. One year that particularly stands out as a poor model validation year was 1993, where the peak simulated flows were approximately 10 m^3/s greater than the observed flow rate. The MESH model greatly over estimated the spring melt that occurred in early March and this resulted in a very poor Nash–Sutcliffe value (NS = 0.26).

The results for Reynolds Mountain East subwatershed were generally not as good as Tollgate. However, the weighted average of the total flow volume was used for model calibration put little emphasis into achieving the best calibration for Reynolds Mountain East. The overall area of the site is only 0.38 km^2 and peak flows range from 0.04 m^3/s for a low flow year to 0.14 m^3/s for a high flow year. For this scale, the overall results for this small study subwatershed are quite good.

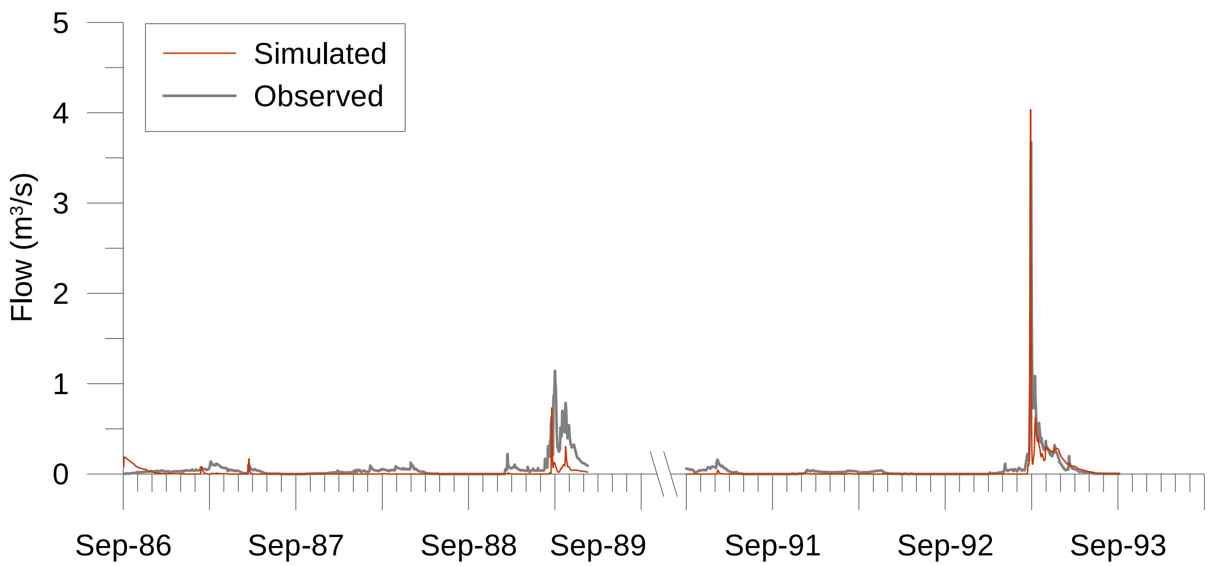

Figure 6-3: Model validation for Salmon Creek, 2km 1 GRU

The validation runs for Salmon Creek did not capture the low flow events and under estimated the higher flow events. This could possibly be a result of the interpolation used to distribute the precipitation forcings. The precipitation distribution for the watershed (previously presented in Figure 3-2) shows that the Salmon Creek subwatershed can receive between 500-200 mm/year, however geographically the nearest precipitation gauges only receive 200 mm/year. During the higher flow years (1989 and 1993), the overall shape and timing of the hydrograph was quite good however the flows were still drastically underestimated.

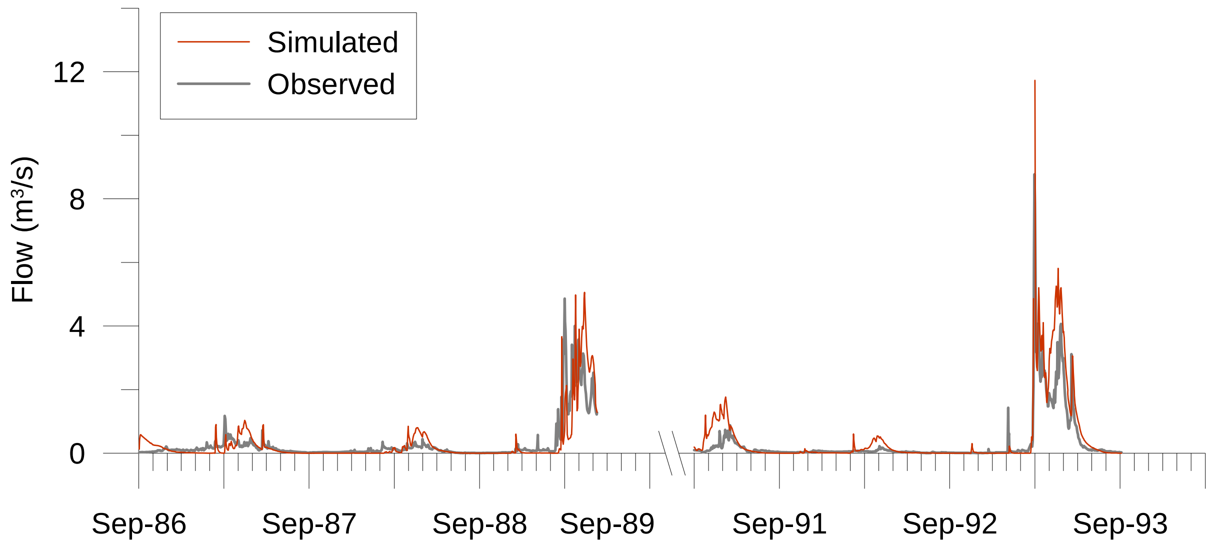

Figure 6-4: Model validation for Reynolds Creek Outlet, 2km 1 GRU

The simulated flows for Reynolds Creek Outlet typically over-predicted the observed flow values, particularly during the low flow years. This is likely a result of withdraw of streamflow for flood irrigation, as documented by Johnson and Smith (1979). The timing errors observed at Reynolds Creek Outlet are likely a result of one of the hydrological parameters that was not calibrated for the outlet as we do not see the same errors at the Tollgate weir.

The validation performed on the single GRU model configuration for Reynolds Creek Outlet showed that the model produces reasonably good results using a single GRU model configuration for this watershed. In some instances there was a slight deviation of fit or timing but the overall model validation generally showed positive results.

For the 6 GRU configuration of Reynolds Creek, validation runs were performed for Tollgate, and Reynolds Mountain East only. Reynolds Creek and Salmon Creek were not used for model validation due to model stability issues. The 6 GRU configuration of MESH was unable to run continuously from September 2nd, 1986 until September 30th, 1996 and terminated due to an error in early 1989. In an attempt to obtain data for model validation the model was restarted in September 1991 but terminated with an error in May 1993 and was again restarted for September 1993 and terminated with an error in March 1995. The hydrographs for Reynolds Mountain East and Tollgate for the 6 GRU model configuration are presented in Figure 6-5 and Figure 6-6, respectively.

Model performance statistics were calculated for the calibration and validation results for the 6 GRU model configuration for the Tollgate and Reynolds Mountain subwatersheds. The hydrological statistics for the 6 GRU model configuration showed similarities to the single GRU model configuration. The statistics were only calculated for years that did not include a spin-up period. Overall statistics for the entire period (1987-1995) were not calculated because the model was unable to run continuously. The statistical results are presented in Table 6-6.

Table 6-6: Summary of annual statistics for Reynolds Creek, 2km 6 GRU model configuration

	Reynolds Creek Watershed					
	Tollgate			Reynolds Mountain East		
	NS	APB	R^2	NS	APB	R^2
1987	0.81	11%	0.92	0.51	57%	0.86
1988	0.90	1%	0.95	0.59	51%	0.88
1989	0.68	32%	0.86	0.07	77%	0.82
1992	0.65	3%	0.92	0.43	52%	0.73
1994	0.82	10%	0.93	0.66	43%	0.87
1995	0.20	6%	0.64	0.20	55%	0.63

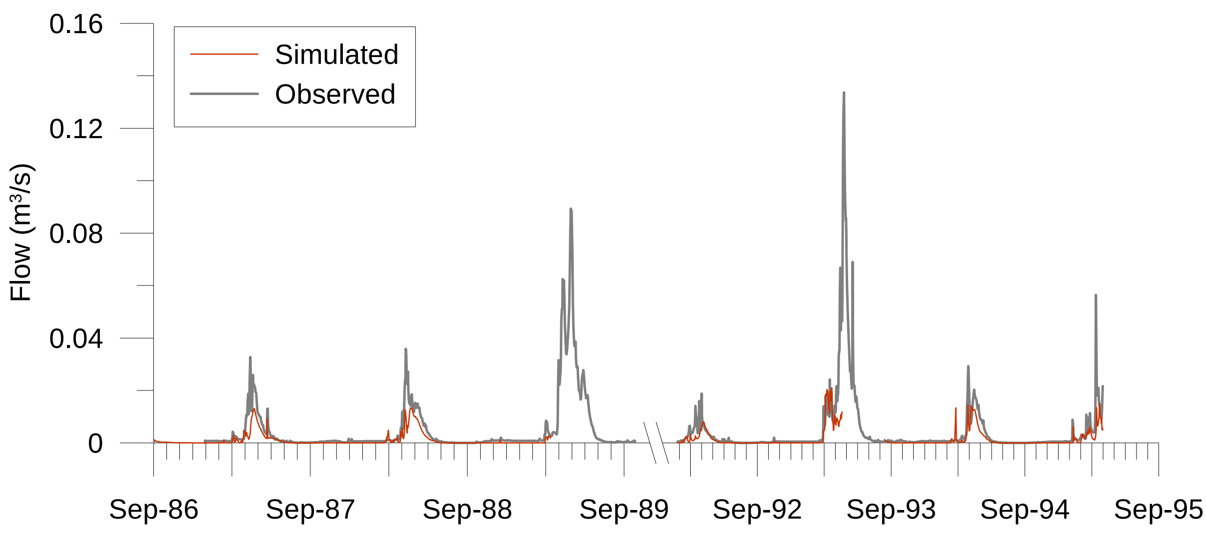

Figure 6-5: Calibration (1987-1988) and validation (1989-1995) for Reynolds Mountain East, 2km 6 GRU

Reynolds Mountain East site generally underestimated the discharge for the subwatershed. The calibration period (1987-1988) missed the timing of the peak for both calibration years and underestimated the flow (APB = 57% and 51%). Both calibration years were able to generally match the shape of the hydrograph (R^2 = 0.86 and 0.88). The other validation years (1992, 1994 and 1995) had some timing issues in 1992, but were generally good for 1994 and 1995. All other validation years missed the discharge peaks and under estimated the total flow volume.

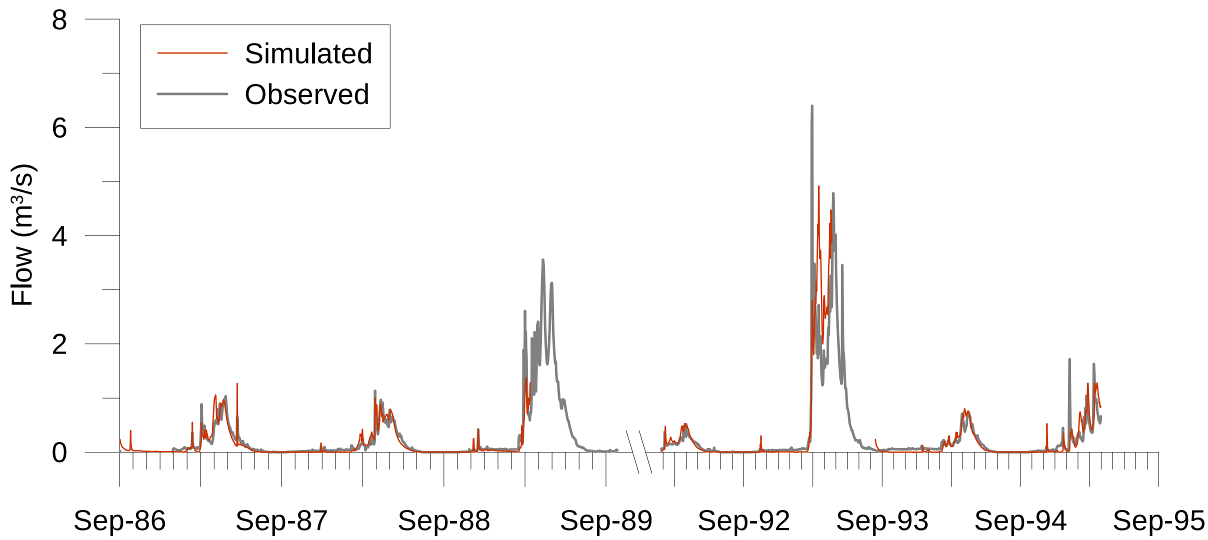

Figure 6-6: Calibration (1987-1988) and validation (1989-1995) for Tollgate, 2km 6 GRU

Reynolds Creek at Tollgate produced very good results with respect to the shape, timing and peaks of the hydrograph. The calibration period (1987-1988) consisted of two low flow years. Both calibration years were able to closely match the shape of the hydrograph (R^2 = 0.92 and 0.95) and the overall discharge (APB = 11% and 1%) resulting in very good results for the calibration period (NS = 0.81 and 0.90). The validation years (1992, 1994 and 1995) were generally able to capture timing for the melt period of the hydrograph for the lower flow years. The higher flow years (1989 and 1993) resulted in the model terminating in error. However the simulated peaks for the 1993 melt event did not over estimate the melt period streamflow like the single GRU configuration.

The difficulty encountered while trying to run the validation period for the model shows that there are major stability issues inherent in using the distributed version of the model. These errors include water balance errors, temperature solver errors, and energy balance errors. It is not clear why these stability errors are more prevalent on the 6 GRU model configuration of Reynolds Creek.

6.1.2.3 Wolf Creek Streamflow Validation

For the Wolf Creek case study model calibration (1999 and 2000) and validation (1997 and 1998) are presented in Figure 6-7 and Figure 6-8 for the single GRU and 4 GRU configurations, respectively. The Nash–Sutcliffe values (NS), absolute percent bias (APB) and correlation coefficients (R^2) were calculated for both model configurations and are presented in Table 6-7.

Table 6-7: Summary of annual statistics for Wolf Creek

	Wolf Creek Watershed					
	1 GRU			4 GRU		
	NS	APB	R^2	NS	APB	R^2
1997	0.12	17%	0.40	0.22	10%	0.49
1998	-0.59	68%	0.25	-1.54	43%	0.05
1999	0.77	6%	0.94	0.18	28%	0.82
2000	0.45	7%	0.74	0.13	3%	0.57
1997-2000	0.32	12%	0.59	0.26	5%	0.57

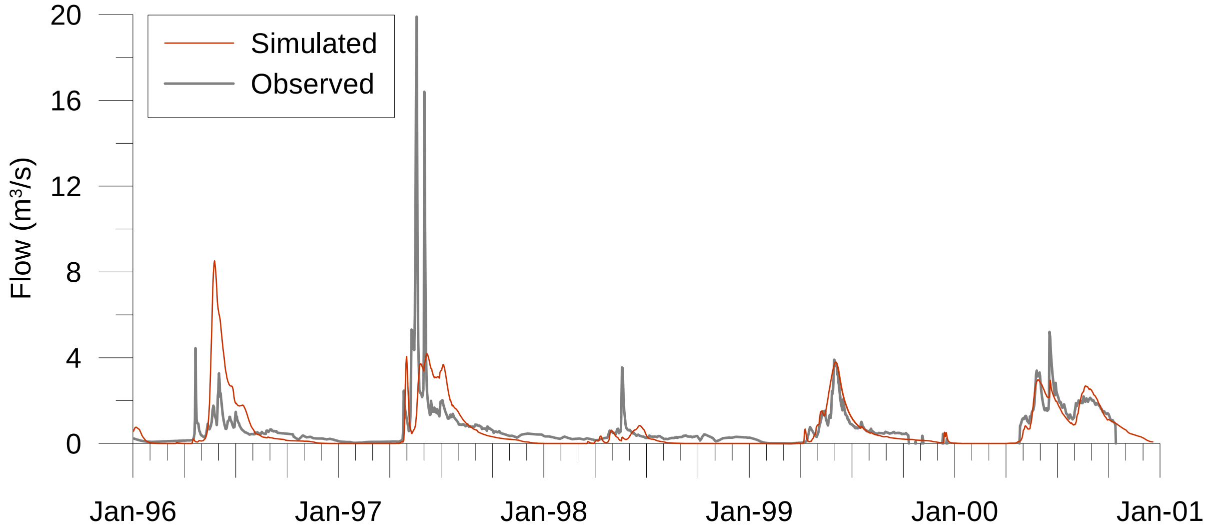

Figure 6-7: Calibration (1999-2000) and validation (1997-1998) for Wolf Creek, 30km 1 GRU

While the single GRU configuration did rather well for the model calibration, the validation results were quite poor. For the calibration period (1999-2000) the model was able to reproduce the timing, shape ($R^2 = 0.94$ and 0.74) and volume (APB = 6% and 7%) of the peaks quite well. In the first validation year, 1997, the model was unable to simulate the shape or volume. However the large peaks in the observed data were caused by ice jam coupled with rain fall events (Soulis, 2009). The second validation year, 1998, was a low flow year. The model was unable to simulate the peak flows, the volume or the shape of the event. While the calibration results were promising, the model could not reproduce the results for validation period showing that this model configuration and/or the calibrated parameter set are not a robust predictor of streamflow for Wolf Creek.

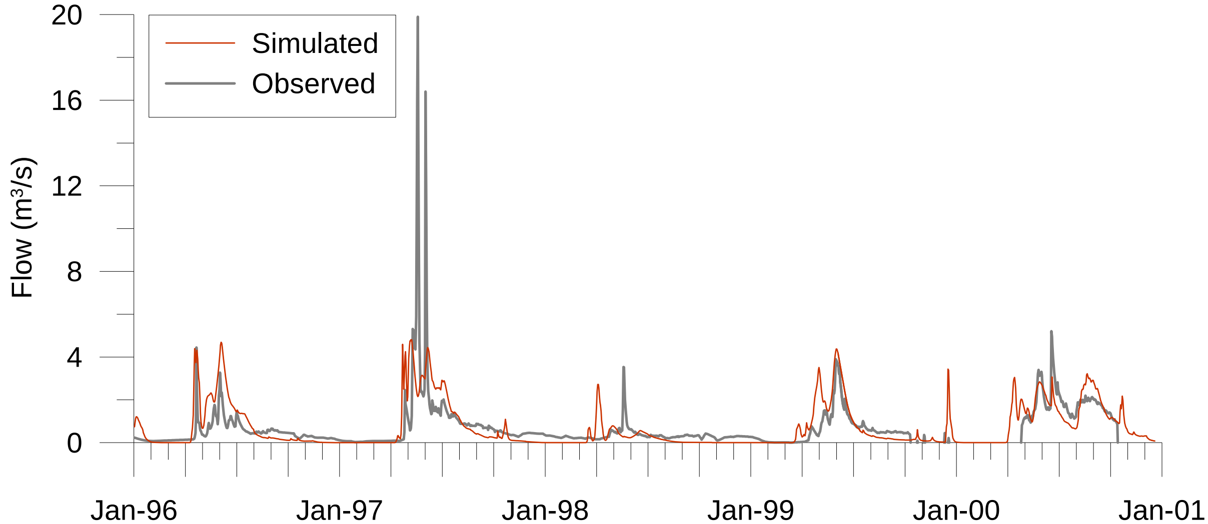

Figure 6-8: Calibration (1999-2000) and validation (1997-1998) for Wolf Creek, 30km 4 GRU

Overall, both the calibration and validation for the 4 GRU configuration of Wolf Creek were poor. Although the model was able to simulate the peak and the shape for the 1999 melt event, it erroneously simulated two additional peaks (one in May and one in December) that did not occur in the measured data. The second calibration year was generally able to match the shape, but under or over estimated the peak flows. The first validation year, 1997, the model was generally able to simulate the shape of the hydrograph, with the exception of the ice jam events that occurred in April, 1997. For second validation year, 1998, the model predicted the peak approximately two months before the melt actually occurred. Given the poor calibration results, the validation results were not expected to be much better. The statistics calculated for these results confirm the poor performance of this model configuration. The daily Nash-Sutcliffe values calculated annually ranged from -1.54 to 0.22.

6.2 Initial Parameter Set

6.2.1 Calibration Strategy

The initial parameter values used for the calibration of the MESH model (discussed in Section 4.1.3) were developed based on physical values where possible, particularly for Reynolds Creek. Coefficients and other parameters that could not be measured were assigned values based on recommend values from CLASS or MESH model documentation. This process of determining good quality initial parameter values can be a difficult part of the model set up and typically requires a strong working knowledge of the model.

As an alternative to the time consuming process of determining a physically based initial parameter set as described in Sections 4.1.3 and 4.1.4, a randomly generated set of parameter values confined by appropriate ranges could also be used to initialize the automatic calibration algorithm. This is a common approach when global optimization algorithms are used for automatic calibration (Tolson and Shoemaker, 2007; Duan 1992) and thus a potentially time saving alternative.

To determine how the initial set of parameter impacts the overall calibration of the model, a set of model runs were performed using randomly generated solutions. The calibration experiments were performed for Wolf Creek and Reynolds Creek and are summarized in Table 6-8. These results are contrasted with the baseline model results which utilized the time-intensive, data-based initial parameter set.

Wolf Creek was included as part of this case to determine if an increase in function evaluations (10,000) is able to achieve equally good calibration results. This experiment was designed to see if it was possible to simply optimize for long enough to overcome initial parameter difficulties. Although orders of magnitude more function evaluations are likely necessary to overcome the poor calibration performance when using a randomly generated initial parameter sets, for a model like MESH, 10,000 was deemed representative of the maximum number that could be conducted in most practical calibration problems.

Table 6-8: Model runs performed with a randomly generated initial parameter sets

Watershed	Reynolds Creek	Wolf Creek	
Configuration	2km 6 GRU	30km 1 GRU	30km 4 GRU
GRUs Calibrated	GRU 1 and 2	Single GRU	GRU 2 and 4
No. Parameters Calibrated	62	31	62
Parameter Ranges	Refined	Refined	Refined
Initial Solution	Random	Random	Random
Objective Calibrated	Streamflow at Tollgate	Streamflow at Outlet	Streamflow at Outlet
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000, 10000	1000, 5000
Trials	5	5, 5	5,5

6.2.2 Results and Discussion

To determine how the initial set of parameter impacts the overall calibration of the model, a set of model runs was performed using randomly generated solutions. The results of these calibration runs for Wolf Creek and Reynolds Creek are summarized in Table 6-9, and graphically in Figure 6-9.

Table 6-9: A comparison of Nash–Sutcliffe values for an initial parameter (calibration strategy 6.1) set and a random parameter set (calibration strategy 6.2)

Watershed	Wolf Creek		Wolf Creek		Wolf Creek		Wolf Creek		Reynolds Creek	
Configuration	30km 1 GRU		30km 1 GRU		30km 4 GRU		30km 4 GRU		2km 6 GRU	
Evaluations	1000		10000		1000		5000		1000	
Trials	5		5		5		5		5	
Calibration Strategy	6.1	6.2	6.1	6.2	6.1	6.2	6.1	6.2	6.1	6.2
Average Initial NS	-3.66	-37.52	-3.66	-29.19	-5.56	-14.40	-5.56	-10.36	0.20	-47.00
Calibrated NS Value										
Worst	0.50	0.19	0.72	0.40	0.09	-0.58	0.37	0.37	0.72	-0.40
Best	0.72	0.38	0.82	0.52	0.40	0.22	0.46	0.39	0.85	0.31
Average	0.66	0.27	0.76	0.46	0.31	-0.10	0.42	0.38	0.78	0.03

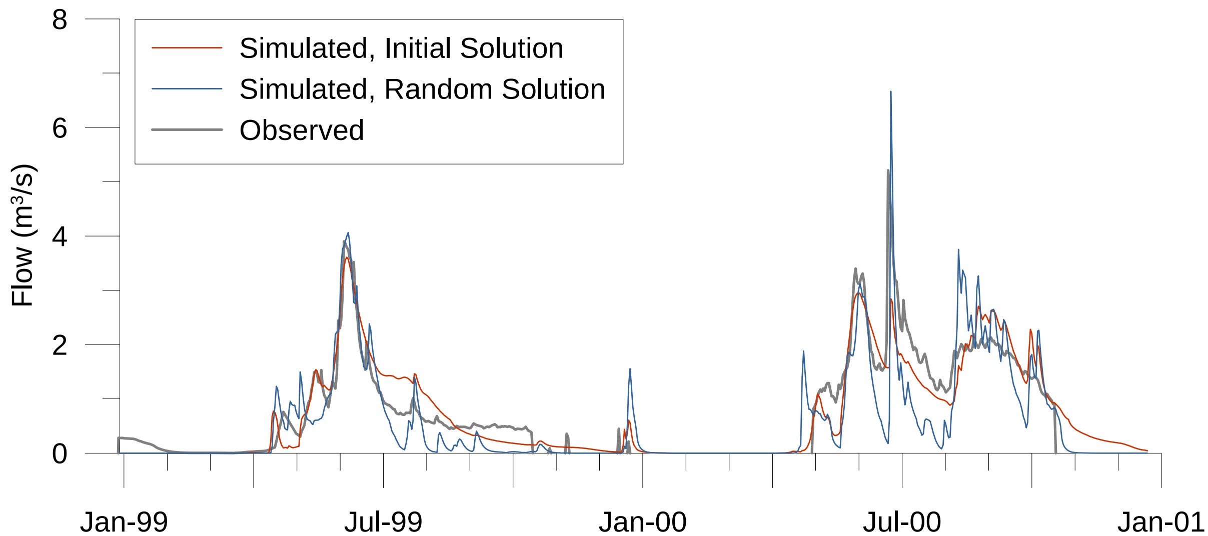

Figure 6-9: Calibration results for Wolf Creek (single GRU, 10,000 function evaluations) using an initial solution and random values

On average the best solutions obtained using random initial parameter sets were an order of magnitude worse than the baseline result. Even after 10,000 model evaluations and multiple trials the model still performs better with the initial estimated parameter set.

The results from this experiment show that using random initial parameter set will greatly degrade the search algorithm performance. In contrast, starting the search algorithm form a carefully estimated initial set of parameters (preferably, based on available data) is able to produce better calibration results.

6.3 Parameter Range Specification

6.3.1 Calibration Strategy

Parameter ranges specified for calibration parameters can have a large impact on the quality of the calibration results. If the parameter range is too large then the calibration algorithm will have a larger parameter space to search, making it more difficult to find good quality parameter sets. With a larger parameter space it is also more likely to encounter a parameter set that will cause the model to terminate in error if the feasible bounds of the parameters are exceeded. This is particularly a problem with MESH due to the instabilities of the distributed model configuration. One goal of this thesis work was to find a set of realistic ranges for calibrating the MESH model and to demonstrate that extremely wide parameter ranges can negatively impact model calibration performance with DDS. The refined parameter ranges were developed using default parameter values and recommendations from MESH model developers. These parameter ranges were used to generate the baseline results. The initial set of parameter rages, used in early calibration experiments were based on a wide set of parameter ranges loosely based on default values. Both the initial set of parameter ranges and refined set of parameter ranges are listed in Appendix E. A summary of the model experiments is presented in Table 6-10.

Table 6-10: Model runs performed with an unrefined set of parameter ranges

Watershed	Reynolds Creek	Wolf Creek
Configuration	2km 6 GRU	30km 1 GRU
GRUs Calibrated	GRU 1 and 2	Single GRU
No. Parameters Calibrated	62	31
Parameter Ranges	Initial Range	Initial Range
Initial Solution	Estimated Parameters	Estimated Parameters
Objective Calibrated	Streamflow at Tollgate	Streamflow at Outlet
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000
Trials	5	5

6.3.2 Results and Discussion

To determine if parameter ranges used for model calibration parameters have an impact on the quality of calibration results, the model was calibrated with an initial set of parameter ranges. The results are compared to the baseline results, calibrated using a refined set of parameter ranges. The experiment results are presented Table 6-11.

Table 6-11: A comparison of Nash–Sutcliffe values for initial parameter ranges (calibration strategy 6.3) and a refined parameter ranges (calibration strategy 6.1)

Watershed	Wolf Creek		Reynolds Creek	
Configuration	30km 1 GRU		2km 6 GRU	
Evaluations	1000		1000	
Trials	5		5	
Calibration Strategy	6.1	6.3	6.1	6.3
Model Crashes (%)	< 1%	< 1%	6%	40%
Calibrated NS Value				
Worst	0.50	0.07	0.72	0.62
Best	0.72	0.60	0.85	0.76
Average	0.66	0.44	0.78	0.71

The results in Table 6-11 show that the refined parameter ranges produce improved results with the same number of function evaluations reducing the model search space. With regard to model stability, there is little impact on the lumped model results (Wolf Creek) however there is a significant improvement for the distributed model results (Reynolds Creek). With the initial parameter ranges, approximately 40% of the model evaluations terminated in error; this will naturally impact the calibration results. By using the refined parameter ranges model stability and calibration results were greatly improved.

6.4 Soil Constraints

6.4.1 Calibration Strategy

The MESH model uses percentages of sand, clay and organic as an input to calculate the hydraulic conductivity, porosity and other thermal properties for each soil layer (Verseghy, 2008). For example, modelers using the current version of MESH can only modify the conductivity and porosity of soils indirectly by modifying the soil component percentages. The purpose of this experiment is to determine how calibrating the percentage of sand, clay, and organic contribute to the overall model performance compared to assigning soil parameters based purely on soil database values (field data).

In the first experiment, the single GRU configuration of Reynolds Creek was calibrated for streamflow the same way as the baseline calibration (see section 6.1) except that the soil parameters were fixed (not calibrated). This experiment was designed to determine how calibration is impacted by not calibrating the soil composition and using the initial soil survey estimates. In the second experiment, the percentage of sand and clay in each layer were calibrated and the organic content in the soil layers was confined to less than 10% (compared to 100% for the baseline calibration runs) based on the available data for organic soil content in Reynolds Creek soils. The sand and clay composition percentages were not constrained as they have a much larger range of appropriate values. The second experiment was designed to assess the impact of realistic constraints on the organic content percentages of the soil layers. A summary of the calibration runs for Reynolds Creek is given in Table 6-12.

Table 6-12: Model runs performed with limited calibration of the organic soil layer

Watershed	Reynolds Creek	Reynolds Creek
Configuration	2km 1 GRU	2km 1 GRU
GRUs Calibrated	GRU 1	GRU 1
No. Parameters Calibrated	31[a]	22[b]
Parameter Ranges	Refined with the organic soils set to <10%	Refined
Initial Solution	Estimated Parameters	Estimated Parameters
Objective Calibrated	Streamflow at Tollgate	Streamflow at Tollgate
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000
Trials	5	5

[a] Same parameter set and ranges as in Experiment 6.1 except for organic soils.

[b] Reduced parameter set due to fixing the nine total soil composition parameters.

6.4.2 Results and Discussion

Reynolds Creek was examined using two soil calibration strategies; first the soil parameters were fixed (not calibrated) using estimated soils data, and second the soils were calibrated but confined to stay within realistic ranges (organic less than 10%). The results are summarized in Table 6-13.

Table 6-13: Comparison of Nash–Sutcliffe efficiencies for uncalibrated and constrained soil parameters

Watershed	Reynolds Creek	Reynolds Creek	Reynolds Creek
Configuration	2km 1 GRU	2km 1 GRU	2km 1 GRU
Evaluations	1000	1000	1000
Trials	5	5	5
Soils	Unconstrained and Calibrated	Constrained and Calibrated	Uncalibrated
No. Parameters	31	31	22
Calibration Strategy	6.1	6.4	6.4
Initial NS Value	0.08	0.08	0.08
Calibrated NS Value			
Worst	0.71	0.65	0.54
Best	0.82	0.77	0.71
Average	0.75	0.73	0.59

Results in Table 6-13 clearly show that by calibrating all soil parameters the model prediction quality is improved. This is expected since calibrating with more parameters means that there is more flexibility to adjust the model simulation results towards measured calibration data. Unfortunately, as shown in Table 6-14, the unconfined calibration of the soil parameters can result in a large and unrealistic configuration of the soil parameters. Often the organic layer nearly reaches100% when site surveys suggest that there is little to no organic matter in the soils. While the best prediction quality is attained at the expense of unrealistic values, the constrained parameter ranges only caused a small decrease in the quality of the calibration results.

Table 6-14: Summary of the calibrated unconstrained organic content in each soil layer for calibration experiment 6.1 in comparison to the soil survey data

	Layer 1	Layer 2	Layer 3
Approximate Organic Content (%) from soil survey	0-10	0-5	0
Max Organic Content (%)	82.43	72.36	95.47
Minimum Organic Matter (%)	11.89	16.99	59.16
Average Organic Matter (%)	44.25	44.14	79.81

The results in Table 6-13 do show that some level of calibration for the soil parameters is required as fixing soil parameters to initial soil properties estimated from soil survey data yielded substantial decrease in the results (average Nash-Sutcliffe is reduced from 0.75 to 0.59). Therefore, even with site specific soil composition data, calibrating soil parameter is recommended. However it is important to consider the ranges of the site specific soil conditions for model calibration.

6.5 Effects of Solar Radiation Estimation

6.5.1 Calibration Strategy

Solar radiation forcings can be very difficult to obtain or estimate for the purposes of hydrological modeling. The MESH model simulates the energy balance of the land surface and therefore solar radiation data will have an impact on simulated results.

The short wave solar radiation was calculated using two methods. The first is described in Chapter 4 Model Setup section 4.1.1.1. This method used the available topographic data to find the average slope and aspect in each grid cell. Dew point temperature and average albedo are used to determine the maximum clear sky solar radiation. This was then used with the calculated fractional daily cloud cover (extrapolated from the measured solar radiation data) to calculate the corrected solar radiation forcings.

Another possible and much simpler, method to determine the shortwave solar radiation forcings would be to apply an inverse distance weighting scheme to the three sites within the watershed where short wave solar radiation was measured.

The purpose of this experiment is to see if the slope and aspect corrections applied to the solar radiation data (the baseline models in section 6.1) have a notable impact of the performance of the model.

Models using each solar radiation schemes were calibrated for the distributed Reynolds Creek case study for both single GRU and 6 GRUs model configurations. Single objective calibration was performed for streamflow only. The model runs are summarized in Table 6-15.

Table 6-15: Model calibrations performed using an inverse distance weighting method to distribute solar radiation

Watershed	**Reynolds Creek**	
Configuration	2km 1 GRU	2km 6 GRU
GRUs Calibrated	Single GRU	GRU 1 and 2
No. Parameters Calibrated	31	62
Parameter Ranges	Refined	Refined
Initial Solution	Estimated Parameters	Estimated Parameters
Objective Calibrated	Streamflow at Tollgate	Streamflow at Tollgate
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000
Trials	5	5

6.5.2 Results and Discussion

The two solar radiation estimation strategy (calibration strategy 6.1 uses solar forcings calculated with slope and aspect while the calibration strategy 6.5 uses inverse distance weighting) are compared in Table 6-16 and graphically for one model configuration in Figure 6-10. Model performance was only assessed based on streamflow prediction quality.

Table 6-16: A comparison of daily Nash–Sutcliffe values for two solar radiation estimation techniques

Watershed	Reynolds Creek		Reynolds Creek	
Configuration	2km 1 GRU		2km 6 GRU	
Evaluations	1000		1000	
Trials	5		5	
Calibration Strategy	6.1	6.5	6.1	6.5
Initial NS Value	0.08	0.14	0.20	0.48
Calibrated NS Value				
Worst	0.71	0.76	0.72	0.71
Best	0.82	0.84	0.85	0.84
Average	0.75	0.82	0.78	0.81

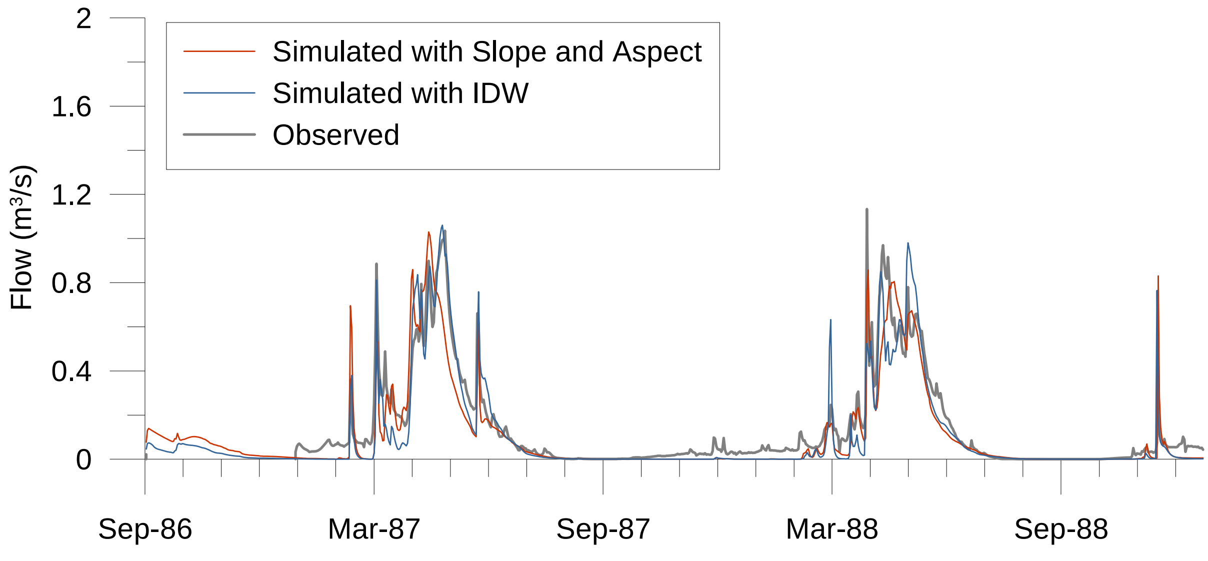

Figure 6-10: Calibration results for Reynolds Creek (single GRU, 1,000 trials) using two different solar radiation forcing schemes

The same initial parameter set was used for both single GRU configurations. However the initial parameter set for the 6 GRU configuration caused the solar inverse distance weighting model configuration to terminate in error. Minor changes were made to the initial parameter set to calibrate this model configuration. Slight changes were made to the initial permeable soil depth for the inverse distance weighting radiation calibration experiments in order to avoid the model terminating in error. The initial Nash-Sutcliffe values for both model configurations were better using the inverse distance weighting solar radiation forcings. The average calibration results for the inverse distance weighting method produced better results than the slope and aspect distributed solar forcings. For example, results in Table 6-16 show the average Nash-Sutcliffe values to increase from 0.75 to 0.82 under strategy 6.5 for the single GRU configuration. For the 6 GRU configuration, the average Nash-Sutcliffe values increased from 0.78 to 0.81.

Results demonstrate that the MESH streamflow predictions appear to be sensitive to the solar radiation inputs. Further investigation is necessary to determine the best method of solar radiation estimation method. The impact of solar radiation is not limited to streamflow and experiments that calibrate SWE (as well as streamflow) as the objective function should also be performed.

6.6 SWE as a Calibration Objective

6.6.1 Calibration Strategy

The baseline experiments (Section 6.1) only considers streamflow as the calibration objective. Reynolds Creek has many additional time series available that could be used for model calibration. For this set of experiments, the impact of calibrating to SWE data instead of, or in addition to, streamflow data are evaluated.

SWE was first calibrated by defining the objective function based on SWE prediction quality at five locations (see Section 5.1.4 for further calibration strategy details). Only thirty-nine parameters in three GRUs were calibrated as the other parameters that were calibrated in the baseline models only impact streamflow and have little to no impact on SWE. The second model experiment calibrated both streamflow and SWE, assuming equal importance for both objectives. This model calibration strategy was also outlined in Section 5.1.4. The quality of the simulated streamflow and SWE results will be compared between the baseline (streamflow objective) and these two different calibration objectives. Model runs are summarized in Table 6-17.

Table 6-17: Model calibrations performed to compare SWE and streamflow calibration objectives.

Watershed	**Reynolds Creek**	
Configuration	2km 6 GRU	2km 6 GRU
GRUs Calibrated	GRU 1, 4, 5	GRU 1, 2, 4, 5
No. Parameters Calibrated	39 (13, 13, 13)	88 (31, 31, 13, 13)
Parameter Ranges	Refined	Refined
Initial Solution	Estimated Parameters	Estimated Parameters
Objective(s) Calibrated	SWE at 5 locations	Streamflow at 2 locations SWE at 5 locations
Objective Function	Nash-Sutcliffe	Nash-Sutcliffe
Function Evaluations	1000	1000
Trials	5	5

6.6.2 Results and Discussion

6.6.2.1 Single objective SWE

The calibration results presented in Table 6-18 summarize the average Nash-Sutcliffe value for five SWE sites over the calibration period. Overall, calibration resulted in a substantial improvement of the initial simulated SWE results, however the aggregated SWE calibration results are not as good as the flow weighted average streamflow results. The SWE simulation results are poor for two site locations (Nash-Sutcliffe, -0.39 and 0.02), but are very good for the three other sites (Nash-Sutcliffe, 0.69, 0.77 and 0.96).

Table 6-18: Nash–Sutcliffe values for the calibration results of aggregated SWE

Watershed	Reynolds Creek
Configuration	2km 6 GRU
Objective Calibrated	SWE
Function Evaluations	1000
Trials	4
Initial NS Value	-0.72
Calibrated NS Value	
Worst	0.16
Best	0.45
Average	0.31

Like other calibration experiments using the 2 km 6 GRU model configuration, substantial model stability issues were encountered when the calibrated parameter sets were used to simulate the model validation period. When model validation was attempted, the model terminated in error in June 1991. In order to generate further validation results, the model was restarted in September 1991 and ran until early April, 1993 when it terminated again in error. Although the validation results are not the result of a continuous simulation, the model terminated in error during the summer months when snow is not present. The results are therefore thought to be representative of what would have been simulated in the absence of model stability errors. The calibration and validation results for each SWE site are presented in Figure 6-11 though to Figure 6-15.

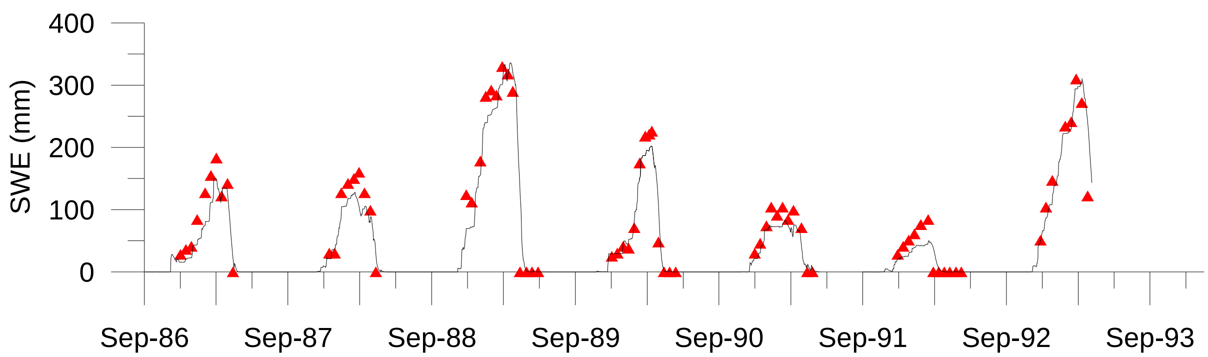

Figure 6-11: Calibration (1987-1988) and validation (1989-1993) for SWE, at site 155x54 where the solid line is simulated SWE and points are measured SWE

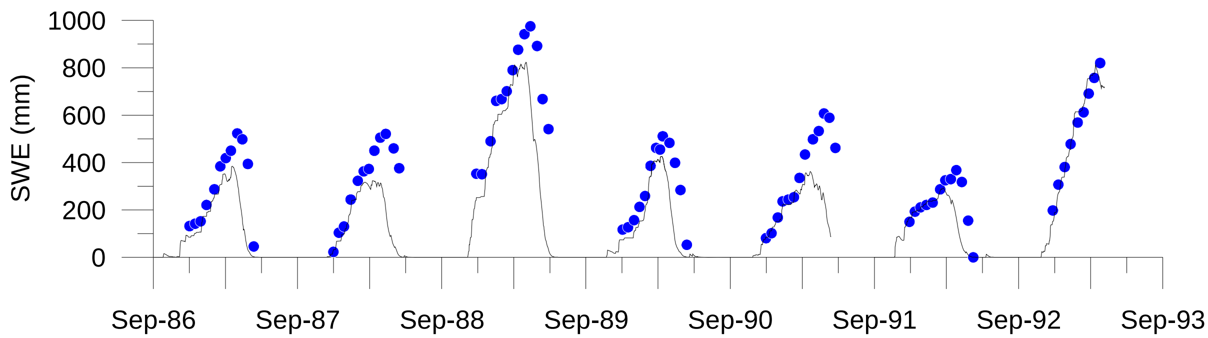

Figure 6-12: Calibration (1987-1988) and validation (1989-1993) for SWE, at site 163x20 where the solid line is simulated SWE and points are measured SWE

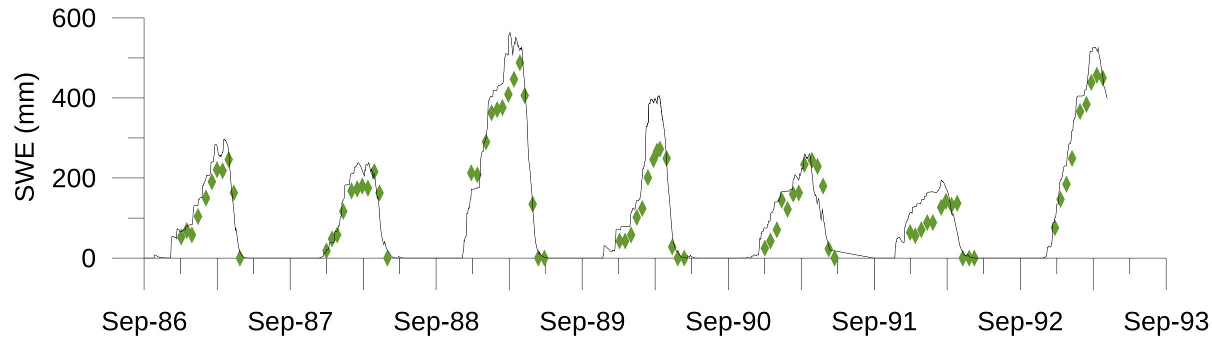

Figure 6-13: Calibration (1987-1988) and validation (1989-1993) for SWE, at site 167x07 where the solid line is simulated SWE and points are measured SWE

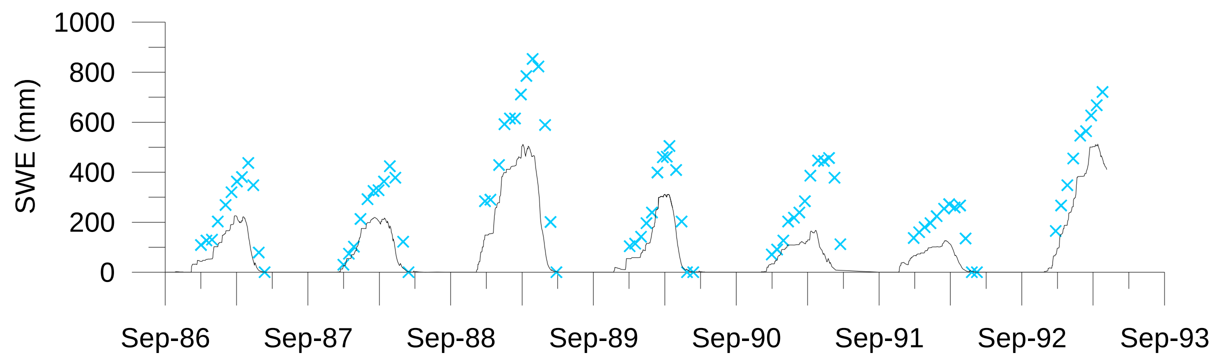

Figure 6-14: Calibration (1987-1988) and validation (1989-1993) for SWE, at site 174x26 where the solid line is simulated SWE and points are measured SWE

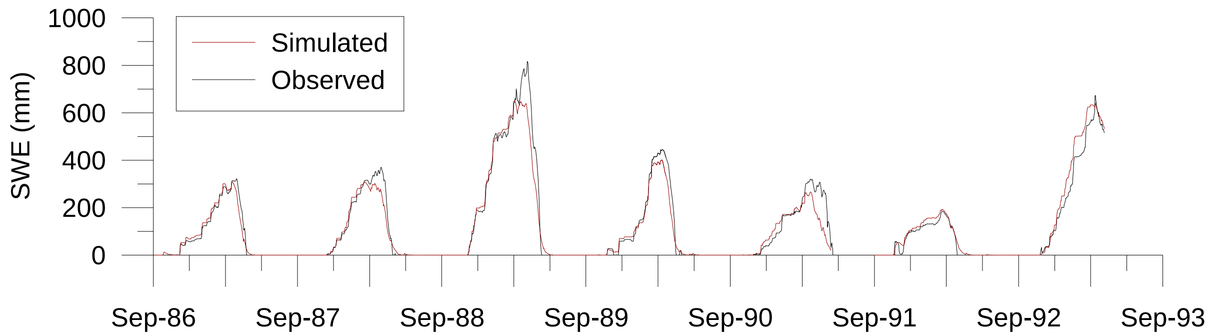

Figure 6-15: Calibration (1987-1988) and validation (1989-1993) for SWE, at the snow pillow site

The overall results for the SWE calibration were quite good, however two sites, 174x26 and 163x20, typically underestimated SWE. Because of the close proximity of the snow course sites to precipitation gauges, the input precipitation for each model grid was presumably accurate but the MESH model is unable to account for the transfer of snow between different grid cells. As a result the model does not allow for snow processes such as blowing snow or snow drifts that frequently occur at various locations. Both snow course sites that did not perform as well, 174x26 and 163x20, were documented to be shaded areas of high deposition on a north east slope (Marks et al., 2000). Therefore, it is not expected that the overall SWE calibration results will improve unless additional snow transportation processes are incorporated into the MESH model.

Various model performance statistics for the calibration and validation periods were computed for the SWE results and given in Table 6-19.

Table 6-19: Overall statistics for SWE calibration sites

Reynolds Creek Watershed - SWE							
		Calibration			Validation		
	GRU	NS	APB	R^2	NS	APB	R^2
155x54	5	0.69	22%	0.98	0.93	3%	0.97
163x20	1	-0.39	40%	0.92	0.40	25%	0.96
167x07	1	0.77	12%	0.83	0.80	21%	0.88
174x26	4	0.02	51%	0.60	0.09	49%	0.83
Snow Pillow Site	4	0.96	2%	0.93	0.94	3%	0.95

The statistics confirmed very poor model simulation for sites 163x20 and 174x26. Both sites also had another site located within the same GRU type being calibrated that did quite well. This suggests that the poor results for these two sites are because of the local conditions that exist at the site location that are not simulated.

6.6.2.2 Combined SWE and Streamflow Objective

By averaging two objectives and calibrating them simultaneously the calibration produces a single parameter set. Figure 6-16, compares the best results of the single objective SWE and streamflow calibrations with the best result for the simultaneous calibration of streamflow and SWE. The number of parameters calibrated for single objective streamflow, single objective SWE and multi-objective optimizations varied for each case. As a result, it is difficult to make a direct comparison of the different optimization strategies. The single objective calibration of SWE was calibrated using 39 parameters in three GRUs. Only parameters that influence SWE were calibrated, other parameters that are calibrated for streamflow were not calibrated for SWE. The single objective streamflow was calibrated using 62 parameters in two GRUs. Finally, the multi-objective optimizations calibrated 88 parameters in four GRUs. This suggests that the multi-objective calibration has the largest search space, but also a much better chance of producing better overall results.

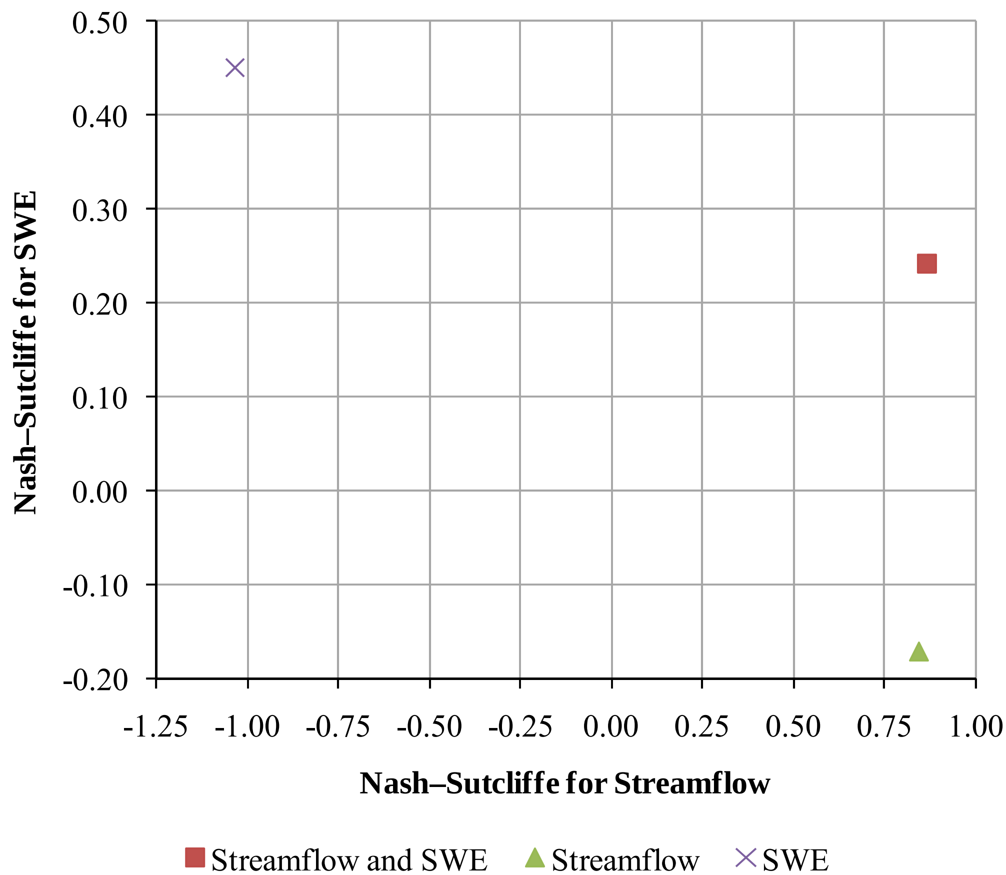

Figure 6-16: Results of multi-objective calibration strategies with single objective end points

Calibrating streamflow individually yields poor results for SWE, and vice versa. By calibrating SWE and streamflow simultaneously a good balance was achieved by using the average to the two objective functions. For example, streamflow obtained a best Nash-Sutcliffe result of 0.84 as a single objective calibration and 0.87 as a multi-objective calibration. For SWE the best single objective Nash-Sutcliffe result was 0.45 and the best multi-objective Nash-Sutcliffe value was 0.24. While the multi-objective value is not as good as the best SWE calibration result, it does fall within the range of values achieved by the SWE calibration (0.16-0.45). This overall improvement in the calibration provides a better solution for both objectives, and was identified using the same number of function evaluations while calibrating a larger number of parameters.

7. Conclusions and Recommendations

Reynolds Creek was setup as a distributed model with a 2km grid cell size. This case study can potential be used for future hydrological modeling and calibration research. The long term data set and large number of measured state variables makes it an ideal case study for model testing. More data has been collected for Reynolds Creek since the initial public release of the data in 2000. This additional data set could also be incorporated into the current Reynolds Creek model setup.

The MESH model was calibrated and validated for distributed model of Reynolds Creek and the lumped model of Wolf Creek. The overall model performance of Wolf Creek was fair and more effort in the overall model calibration and setup would have likely produced better model results.

Calibration and validation of Reynolds Creek produced very good model results for streamflow predictions. However, model stability and computational efficiency hampered calibration of the model. The streamflow results for the 6 GRU watershed configuration showed only minor improvements over the single GRU model configuration. This may be partly because of the two dominant land classes present within the watershed were similar (both different species of sagebrush) and occupied approximately 88% of the total land area. The 6 GRU model configuration experienced issues with model instabilities and was highly sensitive to parameter calibration ranges. During model validation for both the single GRU and 6 GRU model configurations, the model terminated early with an error thus decreasing the length of the planned model validation period and generally made model validation very difficult.

In addition to streamflow, Reynolds Creek was also calibrated to SWE. Results were very good for three snow sites and poor for two snow sites, where the MESH model does not simulate some of the snow processes such as relocation that is known to occur at these sites. Further model validation for SWE could be performed using additional sites located though out the basin. Other measures of performance, such as snow depth and snow density, could also be used for model validation.

When streamflow or SWE are optimized as a single objective problem, the resulting solution will only provide good results for the optimized objective. If both objectives are optimized simultaneously with a weighted average approach, the results can yield good results for both objectives without an increase in computational budget. The weighted average approach yields a preferable solution then calibrating to streamflow alone as the SWE results are improved without degrading the quality of the streamflow results.

The Wolf Creek case study was used to demonstrate how the number of model function evaluations can impact the end calibration result. A larger computational budget typically produced better results. For the Wolf Creek case study, the model did not perform very well with a computational budget of 1,000 function evaluations and significant improvements were observed with a larger computational budget of 10,000 function evaluations. The Reynolds Creek case study was not tested on more than 1,000 function evaluations due to the computational time required by the model; however after 1,000 function evaluations the model was able produce good quality results for Reynolds Creek. Depending on the case study, a budget of 1,000 model evaluations for calibration may not be adequate to achieve a good calibration result.

Calibration of the MESH model using DDS was shown, for the two case studies, to produce better results with an initial parameter set estimated from default and observed values as opposed to a random initial parameter set. On average, the random initial parameter sets produced an initial Nash–Sutcliffe value that was an order of magnitude worse than the estimated initial parameter set. Even after 10,000 function evaluations for the Wolf Creek case study, the random initial solutions produced significantly worse results.

The parameter ranges used for calibration of the MESH model will impact model performance with regards to overall calibration and model stability issues with the distributed model configuration. Typically, the refined parameter ranges (presented in Appendix E) will improve model calibration performance by reducing the size of the large parameter search space.

By calibrating the soil parameters for a watershed, the model is able to produce notably better calibration results. While calibration is necessary, the soil composition should be confined with ranges that reflect the observed soil conditions. If the soil parameters ranges are not confined to observed conditions then the resulting calibrated soil can result in unrealistic soil conditions.

Solar distribution and estimation of the radiation forcings will have an impact on the model prediction quality for streamflow. Two different solar radiation estimation schemes (slope and aspect compared to inverse distance weighting) were tested for their impact on streamflow calibration. Results showed the inverse distance weighting scheme produced slightly better results. Further calibration and validation experiments should be performed using SWE to better determine if one approach is clearly better than the other in terms of overall model prediction quality.

Recommendations on future calibration experiments with MESH include the following:

- By reducing number parameters calibrated, the efficiency of the calibration should improve. This thesis work focused to attempting to achieve the best possible calibration results and therefore focused on a large number of parameters. Further calibration experiments should be conducted using the sensitive parameters for calibration.

- Alternate GRU definitions could be examined for Reynolds Creek. Soil or hydrological parameters could be used to setup the model as the aggregation of the vegetation parameters (done with the single GRU configuration) did not negatively impact model calibration.
- Additional multi-objective calibration strategies should be implemented and compared to the weighted average approach applied in this thesis.

Recommendations for future MESH model development include the following:

- Further testing of new versions of MESH should include testing on a highly detailed, long duration (e.g. 10 years) distributed case study such as Reynolds Creek to make sure that numerical problems in the model have the highest chance of being identified. Using simplified, short duration watersheds to test model stability and performance will not identify many of the model problems that were encountered with Reynolds Creek.
- The input parameter sets for MESH require the setup of each GRU to have the same hydrological, vegetative, soils and prognostic variables. This forces the user to make assumptions that regarding the parameterization of the model that may or may not be true. By partitioning the parameter file so that these inputs are no longer linked the usability of the MESH model could greatly be improved.
- The user should be able to specify directly the values of soil parameters like porosity and hydraulic conductivity instead of indirect specification through the component soil percentages. Although results do not show this will definitely improve model calibration performance, it will make the calibration process more straightforward.

References

Abbott, M.B, Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., (1986). An introduction to the European Hydrological System – System Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modeling system. *Journal of Hydrology*, 87:45-59.

Arnold, J. G., Srinivasan, R., Muttiah, R. R., Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: Model development. *Journal of American Water Resources Association*, 34: 73– 89.

Bastien, Jonathan. Testing the WATCLASS 3.0 Distributed Hydrological Model on Two Canadian Watersheds: Data Preparation, Model Calibration and Performance Assessment. MASc Thesis, University of Waterloo, Waterloo, ON, 2004.

Bathurst, J.C. and Cooley, K.R., (1996). Use of the SHE hydrological modeling system to investigate basin response to snow melt at Reynolds Creek, Idaho. *Journal of Hydrology,* 175(1996) 181-211.

Beven, K J and Freer, J, (2001). A Dynamic TOPMODEL. *Hydrological Processes*, 15(10), 1993-2011.

Brutsaert, Wilfried. Hydrology: an Introduction. Cambridge University Press, 2005.

Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D., (1996). Maximum rooting depth of vegetation types at the global scale. *Oecologia,* 108:583-595.

Cooper, D.J., Sanderson, J.S., Stannard, I.D., Groeneveld, D.P., (2006). Effects of long-term water table drawdown on evapotranspiration and vegetation in an arid region phreatophyte community. *Journal of Hydrology,* 325:21-34.

Craig, James. "Raven: Solar Radiation Generator" Personal communication. February 23rd, 2009.

Davison, B., Pohl, S., Dornes, P., Marsh, P., Pietroniro, A., MacKay, M., (2006). Characterizing Snowmelt Variability in a Land-Surface-Hydrologic Model. *Atmosphere-Ocean,* 44(3) 2006, 271-287.

Deb, Kalyanmoy. Multi-objective optimization using evolutionary algorithms. John Wiley and Sons, 2001.

DHI (1999) "MIKE SHE WQ—User Manual". DHI Water & Environment, Hørsholm, Denmark.

Dingman, S. Lawernce. Physical Hydrology 2nd Ed. New Jersey: Prentice-Hall, 2002.

Dornes, P., Tolson, B., Davison, B., Pietroniro, A., Pomeroy, J., Marsh, P., (2008). Regionalization of land surface hydrological model parameters in subarctic and arctic environments. *Physics and Chemistry of the Earth*, doi:10.1016/j.pce.2008.07.007.

Duan, Q., Sorooshian, S., Gupta, V., (1992). Effective and Efficient Global optimization for Conceptual Rainfall-Runoff Models. *Water Resources Research,* 28(4), 1015-1031.

Flerchinger, G. N. (2000). "The Simultaneous Heat and Water (SHAW) Model: Technical Documentation." Northwest Watershed Research Center. USDA Agricultural Research Service.

Francis, S. (1997). "Data Integration and Ecological Stratification of Wolf Creek Watershed, South-Central Yukon". Prepared for Indian & Northern Affairs Canada and Agriculture Canada. Applied Ecosystem Management Ltd.

Graham, L.P., and Bergstrom, S. (2000). Land surface modeling in hydrology and meteorology-lessons learned from the Baltic Basin. *Hydrology and Earth System Sciences*, 4(1), 13-22.

Gupta, H. V., Sorooshian, S., Yapo, P. O., (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34, 751– 763.

Hanson, C. L., Marks, D., Van Vactor, S. S., (2000). "Climate Monitoring at the Reynolds Creek Experimental Watershed, Idaho, USA." ARS Technical Bulletin NWRC-2000-6.

Hanson, C. L. (2000). "Precipitation Monitoring at the Reynolds Creek Experimental Watershed, Idaho, USA." ARS Technical Bulletin NWRC-2000-4.

Hanson, C. L. (2001). Clear-Sky Albedo Measured at Seven Rangeland Sites in Southwest Idaho. *Journal of Hydraulic Engineering,* vol.6 issue 6:532-534

IP3 Network, "IP3 Network: Improved Processes and parameterization for Prediction in Cold Regions", 16 April 2008. Last Accessed 27 July 2009. http://www.usask.ca/ip3/

Janowicz, J.R. (1998). "Wolf Creek Research Basin – Overview. Proceedings from the Wolf Creek Research Basin: Hydrology, Ecology, Environment workshop", Whitehorse, Yukon, March 1998, 121-130.

Johnson, C., and Smith, J., (1979). Reducing Stream Sediment Loads by Irrigation Diversions. *Transactions of the American Society Agriculture Engineers,* 22, 573-576, 1979.

Legates, D.R., and McCabe, G.J. Jr., (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research, 35:233-241.*

Lemmem, D.S., Warren, F.J., Lacroix, J., Bush, E. editors (2008). "From Impacts to Adaptation: Canada in a Changing Climate 2007"; Government of Canada, Ottawa, ON, 448p.

Lowe, P.R. (1976). An Approximating Polynomial for the Computation of Saturation Vapor Pressure. *Journal of Applied Meteorology,* vol. 16:100-103.

Kouwen, N., Soulis, E.D., Pietroniro, A., Donald, J., Harrington, R.A., (1993). Grouped Response Units for Distributed Hydrologic Modeling. *Journal of Water Resources Planning and Management,* 119, 289-305.

Madsen, H., (2000). Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology,* 235(2000) 276 – 288.

Madsen, H., (2003). Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives. *Advances in Water Resource,* 26:205 – 216.

Marks, D., Cooley, K.R., Rovertson, D.C., Winstral, A., (2000). "Snow Monitoring at the Reynolds Creek Experimental Watershed, Idaho, USA." ARS Technical Bulletin NWRC-2000-5.

Nash J.E., and Sutcliffe J.V., (1970). River flow forecasting though conceptual models. Part I: a discussion of principles. *Journal of Hydrology,* 10:282-290.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J. R., (2005). "Soil and Water Assessment Tool Theoretical Documentation". Agriculture Research Service, U.S. department of Agriculture, Temple, Texas.

Pierson, F.B., Slaughter, C.W., Cram, Z.K., (2000). "Monitoring Discharge and Suspended Sediment, Reynolds Creek Experimental Watershed, Idaho, USA." ARS Technical Bulletin NWRC-2000-8.

Pierson, F.B., Slaughter, C.W., Cram, Z.K., (2001). Long-term stream discharge and suspended-sediment database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research*, vol. 37 no. 11:2857-2861.

Pietronio, A., Fortin, V., Kouwen, N., Neal, C., Turcotte, R., Davison, B., Verseghy, D., Soulis, E.D., Caldwell, R., Evora, N., Pellerin, P., (2007). Development of the MESH modeling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale. *Hydrology and Earth System Sciences,* 11: 1279-1294.

Pomeroy, J.W., Gray, D.M, Brown, T., Hedstrom, N.R, Quinton, W.L., Granger, R.J., Carey, S.K., (2007). The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes*, Vol 21 (19) 2650-2667.

Robins, J. S., Kelly, L. L., Hamon, W. R., (1965). Reynolds Creek in southwest Idaho: An outdoor hydrologic laboratory. *Water Resources Research*, 1, 407-413, 1965.

Rushton, K.R. Groundwater Hydrology: Conceptual and Computational Models. John Wiley and Sons, 2003.

Seglenieks, Frank. Creation of a gridded time series of hydrological variables for Canada. PhD Thesis, University of Waterloo, Waterloo, On, 2009.

Seyfried, M.S, Harris, C. R., Marks, D., Jacobs, B., (2000a). "A Geographic Database for Watershed Research, Reynolds Creek Experimental Watershed, Idaho, USA." ARS Technical Bulletin NWRC-2000-3.

Seyfried, M.S, Flerchinger, G. N., Hanson, C. L., Murdock, M.D., Van Vactor, S.S., (2000b). "Long-Term Soil Microclimate Database: Reynolds Creek Experimental Watershed, Idaho, USA ." ARS Technical Bulletin NWRC-2000-7.

Singh, V. P., and Woolhiser, D. A., (2002). Mathematical modeling of watershed hydrology. *Journal Hydrologic Engineering*, 7, 270– 292.

Slaughter, C.W., Marks, D., Flerchinger, G.N, Van Vactor, S. S., Burgess, M. (2000). "Research Data collection at the Reynolds Creek experimental Watershed, Idaho, USA". ARS Technical Bulletin NWRC-2000-2.

Soulis, E.D. "MEC/MESH Hydraulic Parameters" Personal document. June 5th, 2007.

Soulis, E.D. "Wolf Creek" Personal communication. July 23rd, 2009.

Soulis, E.D., and Seglenieks, F.R., (2007). The MAGS Integrated Modeling System. Cold Regions Atmospheric and Hydrologic Studies: the Mackenzie GEWEX Experience. Vol. II: Hydrologic Processes, Ed. M.K. Woo, Springer-Verlag. 445-474.

Soulis, E.D., Senlgrove, K.R., Kouwen, N., Seglenieks, F.R., Verseghy, D.L, (2000). Towards Closing the Vertical Water Balance in Canadian Atmospheric Models: Coupling of the Land Surface Scheme CLASS with the Distributed Hydrological Model WATFLOOD. *Atmosphere-Ocean*, 38 (1) 2000, 251-269.

Sturges, D.L., and Trlica, M.J., (1978). Root weights and Carbohydrate Reserves of Big Sagebrush. *Ecology,* 59(6):1282-1285.

Tolson, B. A., and Shoemaker, C. A., (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43, W01413, doi:10.1029/2005WR004723.

Tolson, B. A., and Shoemaker C. A., (2008). Efficient prediction uncertainty approximation in the calibration of environmental simulation models. *Water Resources Research*, 44, W04411, doi:10.1029/2007WR005869.

Verseghy, D., (1991). CLASS – a Canadian land surface scheme for GCMs, I. Soil model. *International Journal of Climatology, 11, 111-133.*

Verseghy, D. L., McFarlane, N.A., Lazare, M., (1993). CLASS – a Canadian land surface scheme for GCMs, II. Vegetation model and coupled runs, *International Journal of Climatology, 13, 347-370.*

Verseghy, D. (2008). "The Canadian Land Surface Scheme Technical Documentation – Version 3.4." Climate Research Division, Science and Technology Branch. Environment Canada.

Vrugt, J. A., Gupta, H.V., Bastidas, L.A., Bouten, W., Sorooshian, S., (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resources Research,* 39(8):1214.

Wahl, H.E. Fraser, D.B., Harvey, R.C., Maxwell, J.B., (1987). "Climate of the Yukon". Environment Canada, Atmospheric Environment Services, Climatological Studies Number 40, Ottawa.

Wang, Q.J., (1991). The Genetic Algorithm and Its application to Calibrating Conceptual Rainfall-Runoff Models. *Water Resources Research*, 27(9)2467-2471.

Yapo, P. O., Gupta, H. V., Sorooshian, S., (1998). Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, 204(1-4) 83-97.

Yin, X., (1997). Optical Air Mass: Daily Integration and its Applications, *Meteorology and Atmospheric Physics*, 63, 227-233.

Zhang, X., Srivivasan, R., Van Liew, M., (2008). Multi-Site Calibration of the SWAT Model for Hydrologic Modeling. *Transactions of the ASABE,* Vol. 51(6):2039-2049.

Appendix A

“CLASS 3.0 template used to create the class.ini initialization file used by WATCLASS” Created by D. Verseghy

Variables Required in Initialization File in CLASS 3.0

Line 1:
TITLE1 – TITLE6, FORMAT(2X, 6A4), 24 character title for your model run

Line 2:
NAME1 – NAME6, FORMAT(2X, 6A4), 24 characters for the name of the researcher

Line 3:
PLACE1 – PLACE6, FORMAT(2X, 6A4), 24 characters for the place name

Line 4:
DEGLAT, Latitude of your site or grid-cell in degrees
DEGLON, Longitude of your site or grid-cell in degrees measured east from 0°
ZRFMGRD(1), Reference height (measurement height) for momentum (wind speed)
ZRFHGRD(1), Reference height (measurement height) for heat (temperature and humidity)
ZBLDGRD(1), The blending height for aggregating surface roughness (50 m is reasonable)
GCGRD(1), Set GCGRD to -1
ILW, Set ILW to 1 if incoming longwave radiation is provided, and to –1 if net longwave radiation is provided.
NLTEST, The number of grid-cells being run, usually set to 1
NMTEST, The number of mosaic tiles being used, set to 1 for single site tests

FORMAT(5F10.2,F7.1,3I5)

Line 5:
(FCANROW(I,M,J),J=1,ICAN+1), Fraction of the grid-cell occupied by 1. needleleaf trees, 2. broadleaf trees, 3. crops, 4. grass, and 5. urban areas
(LAMXROW(I,M,J),J=1,ICAN), Maximum leaf area index for 1. needleleaf trees, 2. broadleaf trees, 3. crops, and 4. grass

FORMAT(9F8.3)

Line 6:
(LNZ0ROW(I,M,J),J=1,ICAN+1), Natural logarithm of the roughness length for 1. needleleaf trees, 2. broadleaf trees, 3. crops, 4. grass, and 5. urban areas
(LAMNROW(I,M,J),J=1,ICAN), Minimum leaf area index for 1. needleleaf trees, 2. broadleaf trees, 3. crops, and 4. grass

FORMAT(9F8.3)

Line 7:
(ALVCROW(I,M,J),J=1,ICAN+1), Visible albedo for 1. needleleaf trees, 2. broadleaf trees, 3. crops, 4. grass, and 5. urban areas
(CMASROW(I,M,J),J=1,ICAN), Standing biomass density ($kg \cdot m^{-2}$) for 1. needleleaf trees, 2. broadleaf trees, 3. crops, and 4. grass

FORMAT(9F8.3)

Line 8:
(ALICROW(I,M,J),J=1,ICAN+1), Near infrared albedo for 1. needleleaf trees, 2. broadleaf trees, 3. crops, 4. grass, and 5. urban areas
(ROOTROW(I,M,J),J=1,ICAN), Rooting depth for 1. needleleaf trees, 2. broadleaf trees, 3. crops, and 4. grass

FORMAT(9F8.3)

Line 9:
(RSMNROW(I,M,J),J=1,ICAN), Minimum stomatal resistance for 1. needleleaf trees, 2. broadleaf trees, 3. crops, and 4. grass. We can suggest values if you are not sure what to use.
(QA50ROW(I,M,J),J=1,ICAN), Coefficient governing the response of stomates to light. It is the value of visible radiation ($W \cdot m^{-2}$) at which stomatal resistance is twice the minimum value. We have found a value of 30 – 50 works for a variety of vegetation, but use your own values if you have them.

FORMAT(4F8.3,8X,4F8.3)

Line 10:
(VPDAROW(I,M,J),J=1,ICAN), (VPDBROW(I,M,J),J=1,ICAN), These are coefficients governing the response of stomatal resistance to vapour pressure deficit. Values corresponding to the function employed in previous versions of CLASS are 0.5 for VPDAROW and 1.0 for VPDBROW. We have found other values for use with selected vegetation types.

FORMAT(4F8.3,8X,4F8.3)

Line 11:
(PSGAROW(I,M,J),J=1,ICAN),(PSGBROW(I,M,J),J=1,ICAN), These are coefficients governing the response of stomatal resistance to soil water suction. We have few data sets to test this function, but suggest values of 100 for PSGAROW and 5 for PSGBROW.

FORMAT(4F8.3,8X,4F8.3)

Line 12:
DRNROW(I,M), A drainage index, set to 1.0 to allow the soil physics to model drainage, and to a value between 0 and 1.0 to simulate impeded drainage. The calculated drainage is multiplied by this value.
SDEPROW(I,M), The permeable depth of the soil column. Set to ≤ 4.1 m.
FAREROW(I,M), When running a mosaic, the fractional area that this tile represents in a grid-cell.

FORMAT(9F8.3) Note: only the first 3 variables in the format statement are used.

Line 13:
XSLPROW(I,M), GRKFROW(I,M), WFSFROW(I,M), WFCIROW(I,M), These are WATFLOOD parameters. If you are not running WATFLOOD algorithms (which you need not do if you are running over a flat surface), these are not used. (In that case the switch IWF in RUNCLASS should be kept at a value of 0.)
MIDROW(I,M), This is a mosaic tile identifier, which has a value of 1 for land.

FORMAT(4E8.1,I8)

Line 14:
(SANDROW(I,M,J),J=1,3), The percentage sand content of soil layers 1-3. –4 denotes an ice sheet, -3 denotes impermeable rock, and –2 denotes organic soil.

FORMAT(3F10.1)

Line 15:
(CLAYROW(I,M,J),J=1,3), The percentage clay content of soil layers 1-3.

FORMAT(3F10.1)

Line 16:
(ORGMROW(I,M,J),J=1,3), The percentage organic matter in a mineral soil. If the soil is an organic soil, 1 denotes fibric peat, 2 denotes hemic peat and 3 denotes sapric peat.

FORMAT(3F10.1)

Line 17:
(TBARROW(I,M,J),J=1,IGND), The temperature (°C) of soil layers 1-3
TCANROW(I,M), The canopy temperature
TSNOROW(I,M), The temperature of the snowpack
TPNDROW(I,M), The temperature of ponded water on the surface

FORMAT(6F10.2)

Line 18:
(THLQROW(I,M,J),J=1,IGND), Fractional volume of liquid water in soil layers 1-3
(THICROW(I,M,J),J=1,IGND), Fractional volume of frozen water in soil layers 1-3
ZPNDROW(I,M), Depth of water ponded on the surface

FORMAT(7F10.3)

Line 19:
RCANROW(I,M), Liquid water held on the vegetation canopy ($kg{\cdot}m^{-2}$)
SCANROW(I,M), Frozen water held on the vegetation canopy ($kg{\cdot}m^{-2}$)
SNOROW(I,M), Snow mass present on the ground ($kg{\cdot}m^{-2}$)
ALBSROW(I,M), Albedo of the snow
RHOSROW(I,M), Density of the snow
GROROW(I,M), Vegetation growth index. Set to 0 before leaf-out, 1.0 when in full leaf, and estimate in between.

FORMAT(2F10.4,F10.2,F10.3,F10.4,F10.3)

Line 20:
JOUT1, Day of year on which half-hourly output begins
JOUT2, Day of year on which half-hourly output ends
JAV1, Day of year on which daily average output begins
JAV2, Day of year on which daily average output ends

FORMAT (4I10)

Line 21:
KOUT1, Year in which half-hourly output begins
KOUT2, Year in which half-hourly output ends
KAV1, Year in which daily average output begins
KAV2, Year in which daily average output ends

FORMAT (4I10)

Appendix B

Initial Parameter Input Files

Reynolds Creek 1 GRU

MESH_parameters_CLASS.ini

```
  IP3 Reynolds Creek                                                        01
  A. MacLean                                                                02
  U of Waterloo                                                             03
       49.30    116.60     15.00     15.00     50.00   -1.0    1    1    1   04
     0.000   1.000   0.000   0.000   0.000   0.000   2.040   0.000   0.000   05
     0.000   0.540   0.000   0.000   0.000   0.000   0.296   0.000   0.000   06
     0.000   0.030   0.000   0.000   0.000   0.000  50.000   0.000   0.000   07
     0.000   0.230   0.000   0.000   0.000   0.000   1.230   0.000   0.000   08
     0.000 125.000   0.000   0.000           0.000  40.000   0.000   0.000   09
     0.000   0.500   0.000   0.000           0.000   0.600   0.000   0.000   10
     0.000 100.000   0.000   0.000           0.000   5.000   0.000   0.000   11
     0.500   1.000   1.000   0.001                                           12
     0.020   0.100   0.100   0.100       1                                   13
       40.00     30.00     60.00                                            14
       25.00     40.00     10.00                                            15
        5.00      0.00      0.00                                            16
       22.00     22.03     22.05    15.100    15.000     10.00              17
       0.150     0.200     0.190     0.000     0.000     0.000     0.000    18
      0.0000    0.0000      0.00     0.000     0.000     1.000              19
           1         1         1         1                                  20
        1986      2100      1986      2100                                  21
           0         0         1      1986                                  22
123456789*123456789*123456789*123456789*123456789*123456789*123456789*123
```

MESH_parameters_hydrology.ini

```
1.1.a04: MESH Hydrology Parameters input file
##### Option Flags #####
----#
    2 # Number of option flags
  0.0 #1 [reserved #1]
  0.0 #2 [reserved #2]
##### Channel River Roughness Factors (WF_R2) #####
-----#-----#-----#-----#-----#
 1.872 0.150 0.150 0.150 0.150
##### GRU Independent Hydrologic Parameters #####
-------#
       2 # Number of GRU independent hydrologic parameters
     0.0 #1 [reserved #1]
     0.0 #2 [reserved #2]
##### GRU Dependent Hydrologic Parameters #####
-------#
       1 #Number of GRUs (must match number in mesh_parameters_class.ini file)
       3 #Number of GRU dependent hydrologic parameters
---------#
      0.90
      0.78
      0.75
```

Reynolds Creek 6 GRU

MESH__parameter_CLASS.ini

```
IP3 Reynolds Creek                                                              01
A. MacLean                                                                      02
U of Waterloo                                                                   03
     49.30    116.60     15.00     15.00     50.00   -1.0     1    1    6       04
   0.000   1.000   0.000   0.000   0.000   0.000   2.040   0.000   0.000        05
   0.000   0.540   0.000   0.000   0.000   0.000   0.441   0.000   0.000        06
   0.000   0.030   0.000   0.000   0.000   0.000  50.000   0.000   0.000        07
   0.000   0.230   0.000   0.000   0.000   0.000   1.230   0.000   0.000        08
   0.000 125.000   0.000   0.000           0.000  40.000   0.000   0.000        09
   0.000   0.500   0.000   0.000           0.000   0.600   0.000   0.000        10
   0.000 100.000   0.000   0.000           0.000   5.000   0.000   0.000        11
   0.500   0.500   1.000   0.001                                                12
   0.020   0.100   0.100   0.100       1                                        13
     35.00     50.00     10.00                                                  14
     30.00     20.00     80.00                                                  15
      5.00      0.00      0.00                                                  16
     19.35     19.68     17.10    15.100    15.000     10.00                    17
     0.120     0.170     0.200     0.000     0.000     0.000     0.000          18
    0.0000    0.0000      0.00     0.000     0.000     1.000                    19
   0.000   1.000   0.000   0.000   0.000   0.000   1.570   0.000   0.000        05
   0.000   0.540   0.000   0.000   0.000   0.000   0.296   0.000   0.000        06
   0.000   0.030   0.000   0.000   0.000   0.000  50.000   0.000   0.000        07
   0.000   0.230   0.000   0.000   0.000   0.000   1.230   0.000   0.000        08
   0.000 125.000   0.000   0.000           0.000  40.000   0.000   0.000        09
   0.000   0.500   0.000   0.000           0.000   0.600   0.000   0.000        10
   0.000 100.000   0.000   0.000           0.000   5.000   0.000   0.000        11
   0.500   1.000   1.000   0.001                                                12
   0.020   0.100   0.100   0.100       1                                        13
     40.00     45.00     60.00                                                  14
     15.00     25.00     10.00                                                  15
      5.00      0.00      0.00                                                  16
     24.15     23.88     23.61    15.100    15.000     10.00                    17
     0.220     0.260     0.200     0.000     0.000     0.000     0.000          18
    0.0000    0.0000      0.00     0.000     0.000     1.000                    19
   1.000   0.000   0.000   0.000   0.000   2.000   0.000   0.000   0.000        05
   0.000   0.000   0.000   0.000   0.000   0.500   0.000   0.000   0.000        06
   0.030   0.000   0.000   0.000   0.000  15.000   0.000   0.000   0.000        07
   0.190   0.000   0.000   0.000   0.000  18.000   0.000   0.000   0.000        08
 200.000   0.000   0.000   0.000          30.000   0.000   0.000   0.000        09
   0.650   0.000   0.000   0.000           1.050   0.000   0.000   0.000        10
 100.000   0.000   0.000   0.000           5.000   0.000   0.000   0.000        11
   0.500   1.200   1.000   0.001                                                12
   0.020   0.100   0.100   0.100       1                                        13
     65.00     65.00     85.00                                                  14
     10.00     10.00      5.00                                                  15
      5.00      0.00      0.00                                                  16
     24.15     23.88     23.61    15.100    15.000     10.00                    17
     0.130     0.200     0.140     0.000     0.000     0.000     0.000          18
    0.0000    0.0000      0.00     0.000     0.000     1.000                    19
   0.000   1.000   0.000   0.000   0.000   0.000   6.000   0.000   0.000        05
   0.000   0.300   0.000   0.000   0.000   0.000   0.500   0.000   0.000        06
   0.000   0.050   0.000   0.000   0.000   0.000  20.000   0.000   0.000        07
   0.000   0.290   0.000   0.000   0.000   0.000   1.230   0.000   0.000        08
   0.000 125.000   0.000   0.000           0.000  40.000   0.000   0.000        09
   0.000   0.500   0.000   0.000           0.000   0.600   0.000   0.000        10
   0.000 100.000   0.000   0.000           0.000   5.000   0.000   0.000        11
   0.500   0.500   1.000   0.001                                                12
   0.020   0.100   0.100   0.100       1                                        13
     40.00     40.00     10.00                                                  14
     20.00     30.00     80.00                                                  15
     10.00      0.00      0.00                                                  16
     19.35     19.68     17.10    15.100    15.000     10.00                    17
```

```
     0.070     0.080     0.100     0.000     0.000     0.000     0.000    18
    0.0000    0.0000      0.00     0.000     0.000     1.000              19
   1.000   0.000   0.000   0.000   0.000   2.000   0.000   0.000   0.000  05
   0.176   0.000   0.000   0.000   0.000   1.600   0.000   0.000   0.000  06
   0.030   0.000   0.000   0.000   0.000  25.000   0.000   0.000   0.000  07
   0.190   0.000   0.000   0.000   0.000   1.500   0.000   0.000   0.000  08
 200.000   0.000   0.000   0.000          30.000   0.000   0.000   0.000  09
   0.650   0.000   0.000   0.000           1.050   0.000   0.000   0.000  10
 100.000   0.000   0.000   0.000           5.000   0.000   0.000   0.000  11
   0.500   0.500   1.000   0.001                                          12
   0.020   0.100   0.100   0.100       1                                  13
     40.00     40.00     10.00                                            14
     25.00     30.00     80.00                                            15
      5.00      0.00      0.00                                            16
     19.35     19.68     17.10    15.100    15.000     10.00              17
     0.070     0.080     0.100     0.000     0.000     0.000     0.000    18
    0.0000    0.0000      0.00     0.000     0.000     1.000              19
   0.000   0.000   1.000   0.000   0.000   0.000   0.000   4.000   0.000  05
   0.000   0.000  -1.096   0.000   0.000   0.000   0.000   0.000   0.000  06
   0.000   0.000   0.060   0.000   0.000   0.000   0.000   2.000   0.000  07
   0.000   0.000   0.360   0.000   0.000   0.000   0.000   2.000   0.000  08
   0.000   0.000  85.000   0.000           0.000   0.000  30.000   0.000  09
   0.000   0.000   0.500   0.000           0.000   0.000   1.000   0.000  10
   0.000   0.000 100.000   0.000           0.000   0.000   5.000   0.000  11
   0.500   1.200   1.000   0.001                                          12
   0.020   0.100   0.100   0.100       1                                  13
     65.00     65.00     85.00                                            14
     25.00     10.00      5.00                                            15
      5.00      0.00      0.00                                            16
     24.51     23.88     23.61    15.100    15.000     10.00              17
     0.130     0.200     0.140     0.000     0.000     0.000     0.000    18
    0.0000    0.0000      0.00     0.000     0.000     1.000              19
         1         1         1         1         1         1              20
      1900      2100      1900      2100      1900      2100              21
         0         0         1      1986                                  22
123456789*123456789*123456789*123456789*123456789*123456789*123456789*123
```

MESH_parameters_hydrology.ini

```
1.1.a04:MESH Hydrology Parameters input file
##### Option Flags #####
----#
    2 # Number of option flags
    0 #1 [reserved]
    0 #2 [reserved]
##### River roughness factor (WF_R2) (5 classes maximum) #####
-----#-----#-----#-----#-----#
 0.237 0.000 0.000 0.000 0.000
##### GRU class independent hydrologic parameters #####
-------#
       2 # Number of GRU independent hydrologic parameters
   0.000 #1 [reserved]
   0.000 #2 [reserved]
##### GRU class dependent hydrologic parameters #####
-------#
       6 #Number of GRUs (must match number in mesh_parameters_class.ini file)
       3 #Number of GRU dependent hydrologic parameters
---------#---------#---------#---------#---------#---------#---------#---------#
      0.98      0.98      0.98      0.98      0.98      0.98
      0.90      0.90      0.90      0.90      0.90      0.90
      0.19      0.19      0.19      0.19      0.19      0.19
```

Wolf Creek 1 GRU

MESH__parameter_CLASS.ini

```
IP3 Project, Wolf Creek                                                      01
A. MacLean                                                                   02
U of Waterloo                                                                03
     61.00    135.5      30.00      30.00      30.00   -1.0     1    1    1  04
   1.000   0.000   0.000   0.000   0.000   3.000   0.000   0.000   0.000  05
   0.405   0.000   0.000   0.000   0.000   0.500   0.000   0.000   0.000  06
   0.030   0.000   0.000   0.000   0.000  10.000   0.000   0.000   0.000  07
   0.120   0.000   0.000   0.000   0.000   1.000   0.000   0.000   0.000  08
 200.000   0.000   0.000   0.000          30.000   0.000   0.000   0.000  09
   0.650   0.000   0.000   0.000           1.050   0.000   0.000   0.000  10
 100.000   0.000   0.000   0.000           5.000   0.000   0.000   0.000  11
   0.080   1.000   1.000   0.001                                          12
   0.020   0.100   0.100   0.100       1                                  13
     70.00     65.00     75.00                                              14
     20.00     35.00     20.00                                              15
      0.00      0.00      0.00                                              16
    -6.60     -5.000     0.00     -18.500    -18.500    -18.500              17
     0.050     0.050     0.200      0.100      0.100      0.000      0.000   18
    0.0000    0.0000     38.00      0.500    100.000      0.000              19
         1         1         1          1          0          0              20
      1900      2100      1900       2100          0          0              21
         0         0         1       1996                                    22
123456789*123456789*123456789*123456789*123456789*123456789*123456789*123
```

MESH_parameters_hydrology.ini

```
1.1.a04:MESH Hydrology Parameters input file
##### Option Flags #####
----#
    2 # Number of option flags
    0 #1 [reserved]
    0 #2 [reserved]
##### River roughness factor (WF_R2) (5 classes maximum) #####
-----#-----#-----#-----#-----#
 5.761 0.000 0.000 0.000 0.000
##### GRU class independent hydrologic parameters #####
-------#
       2 # Number of GRU independent hydrologic parameters
   0.000 #1 [reserved]
   0.000 #2 [reserved]
##### GRU class dependent hydrologic parameters #####
-------#
       1 #Number of GRUs (must match number in mesh_parameters_class.ini file)
       3 #Number of GRU dependent hydrologic parameters
---------#---------#---------#---------#---------#---------#---------#---------#
      0.73
      0.12
      0.26
```

Wolf Creek 4 GRU

MESH_parameter_CLASS.ini

```
IP3 Project, Wolf Creek                                                          01
A. MacLean                                                                       02
U of Waterloo                                                                    03
     61.00    135.5      30.00     30.00     30.00    -1.0     1    1    4     04
   0.000   0.000   0.000   1.000   0.000   0.000   0.000   0.000   1.500     05
   0.000   0.000   0.000  -3.000   0.000   0.000   0.000   0.000   0.100     06
   0.000   0.000   0.000   0.050   0.000   0.000   0.000   0.000   0.200     07
   0.000   0.000   0.000   0.290   0.000   0.000   0.000   0.000   0.100     08
   0.000   0.000   0.000 250.000           0.000   0.000   0.000  50.000     09
   0.000   0.000   0.000   0.620           0.000   0.000   0.000   0.400     10
   0.000   0.000   0.000 100.000           0.000   0.000   0.000   5.000     11
   0.030   1.000   1.000   0.001                                             12
   0.020   0.100   0.100   0.100       1                                     13
     55.00     55.00     55.00                                               14
     20.00     20.00     20.00                                               15
      0.00      0.00      0.00                                               16
    -11.00    -10.00     -9.00     -8.00     -8.00     -8.00                 17
     0.075     0.075     0.075     0.150     0.150     0.200     0.000       18
    0.0000    0.0000     38.00     0.500   100.000     0.000                 19
   0.000   1.000   0.000   0.000   0.000   0.000   2.300   0.000   0.000     05
   0.000   0.405   0.000   0.000   0.000   0.000   0.100   0.000   0.000     06
   0.000   0.050   0.000   0.000   0.000   0.000  10.000   0.000   0.000     07
   0.000   0.290   0.000   0.000   0.000   0.000   2.000   0.000   0.000     08
   0.000 200.000   0.000   0.000           0.000  50.000   0.000   0.000     09
   0.000   0.500   0.000   0.000           0.000   0.600   0.000   0.000     10
   0.000  55.000   0.000   0.000           0.000   1.000   0.000   0.000     11
   0.500   1.000   1.000   0.001                                             12
   0.020   0.100   0.100   0.100       1                                     13
     65.00     65.00     65.00                                               14
     20.00     20.00     20.00                                               15
     10.00      2.00      5.00                                               16
     -6.60    -5.000     0.00    -18.500   -18.500   -18.500                 17
     0.050     0.050     0.200     0.100     0.100     0.000     0.000       18
    0.0000    0.0000     38.00     0.500   100.000     0.000                 19
   1.000   0.000   0.000   0.000   0.000   3.000   0.000   0.000   0.000     05
   0.405   0.000   0.000   0.000   0.000   0.500   0.000   0.000   0.000     06
   0.030   0.000   0.000   0.000   0.000  10.000   0.000   0.000   0.000     07
   0.120   0.000   0.000   0.000   0.000   1.000   0.000   0.000   0.000     08
 200.000   0.000   0.000   0.000          30.000   0.000   0.000   0.000     09
   0.650   0.000   0.000   0.000           1.050   0.000   0.000   0.000     10
 100.000   0.000   0.000   0.000           5.000   0.000   0.000   0.000     11
   0.080   1.000   1.000   0.001                                             12
   0.020   0.100   0.100   0.100       1                                     13
     70.00     65.00     75.00                                               14
     20.00     35.00     20.00                                               15
      0.00      0.00      0.00                                               16
     -6.60    -5.000     0.00    -18.500   -18.500   -18.500                 17
     0.050     0.050     0.200     0.100     0.100     0.000     0.000       18
    0.0000    0.0000     38.00     0.500   100.000     0.000                 19
```

```
   0.000   0.000   0.000   1.000   0.000   0.000   0.000   0.000   1.100  05
   0.000   0.000   0.000  -3.000   0.000   0.000   0.000   0.000   0.100  06
   0.000   0.000   0.000   0.050   0.000   0.000   0.000   0.000   0.200  07
   0.000   0.000   0.000   0.290   0.000   0.000   0.000   0.000   0.100  08
   0.000   0.000   0.000 200.000           0.000   0.000   0.000  40.000  09
   0.000   0.000   0.000   0.620           0.000   0.000   0.000   0.400  10
   0.000   0.000   0.000 100.000           0.000   0.000   0.000   5.000  11
   0.500   1.000   1.000   0.001                                          12
   0.020   0.100   0.100   0.100       1                                  13
     55.00     55.00     55.00                                            14
     20.00     20.00     20.00                                            15
      0.00      0.00      0.00                                            16
    -11.00    -10.00     -9.00    -8.00     -8.00     -8.00               17
     0.100     0.100     0.100     0.200     0.200     0.200     0.000    18
    0.0000    0.0000     38.00     0.500   100.000     0.000              19
         1         1         1         1         0         0              20
      1900      2100      1900      2100         0         0              21
         0         0         1      1996                                  22
123456789*123456789*123456789*123456789*123456789*123456789*123456789*123
```

MESH_parameters_hydrology.ini

```
1.1.a04:MESH Hydrology Parameters input file
##### Option Flags #####
----#
    2 # Number of option flags
    0 #1 [reserved]
    0 #2 [reserved]
##### River roughness factor (WF_R2) (5 classes maximum) #####
-----#-----#-----#-----#-----#
 5.761 0.000 0.000 0.000 0.000
##### GRU class independent hydrologic parameters #####
-------#
       2 # Number of GRU independent hydrologic parameters
   0.000 #1 [reserved]
   0.000 #2 [reserved]
##### GRU class dependent hydrologic parameters #####
-------#
       4 #Number of GRUs (must match number in mesh_parameters_class.ini file)
       3 #Number of GRU dependent hydrologic parameters
---------#---------#---------#---------#---------#---------#---------#---------#
      0.73      0.91      0.94      0.24
      0.12      0.53      0.85      0.37
      0.26      0.01      0.03      0.04
```

Appendix C

MESH Input Soil Depths for Reynolds Creek

The soil layer depths for Reynolds Creek were determined based on the average depths of the consolidated list of soil series found in Seyfried et al. (2000) for soil layers one and two. Soil layer three was based on the depth of the deepest soil layer. For land classes where the soil column is not as deep, the depth of the soil column can be reduced with the SDEPOW parameter (permeable depth of soil column). The depths of each soil layer are graphically presented in Figure C - 1and the depths of each soil layer input into MESH and the cumulative depths for the soil layer are given in Table C - 1.

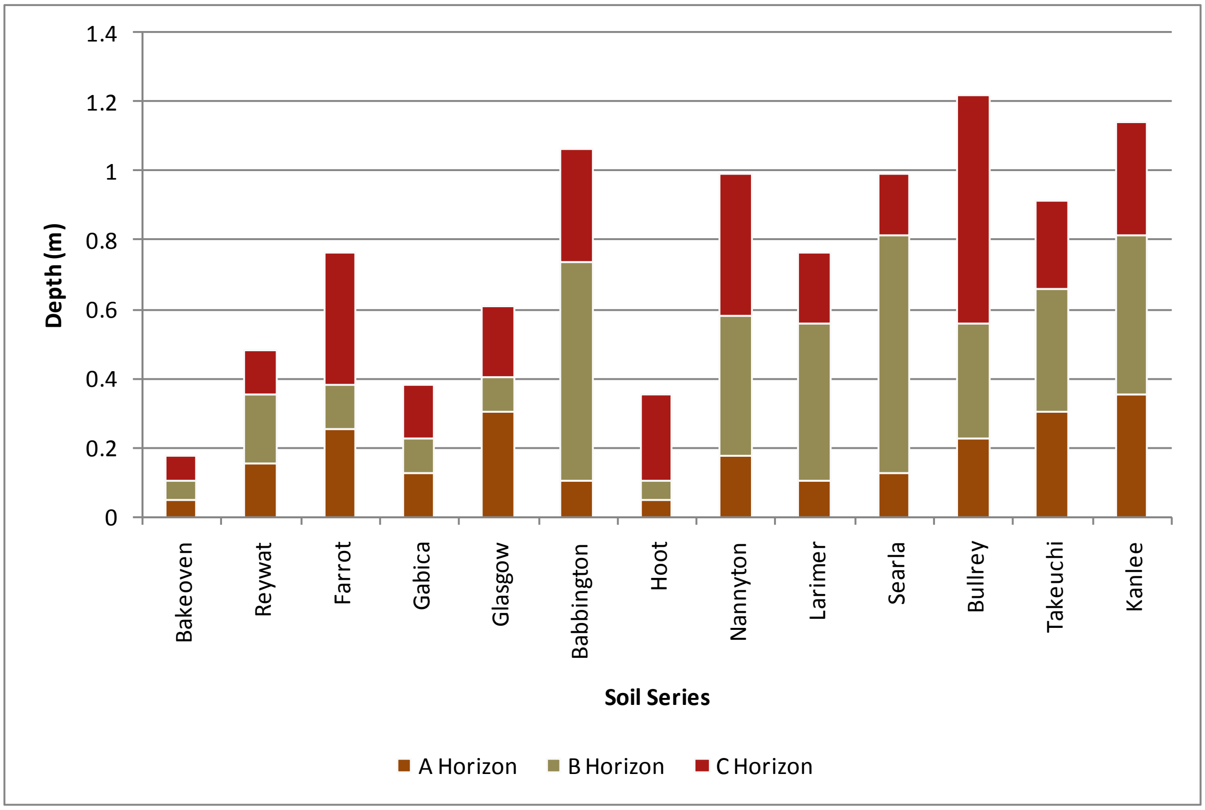

Figure C - 1: Soil Depths

Table C - 1: MESH Soil depths for Reynolds Creek

	Depth (m)	Cumulative Depth (m)
Soil Layer 1	0.18	0.00 - 0.18
Soil Layer 2	0.30	0.18 - 0.48
Soil Layer 3	0.72	0.48 - 1.20

Appendix D

Soil Descriptions for Reynolds Creek

Soil Descriptions: Taken from National Cooperative Soil Survey

	Horizon	Depth (in)		Description
Bakeoven	*very cobbly loam-rangeland*			
Source:	http://www2.ftw.nrcs.usda.gov/osd/dat/B/BAKEOVEN.html			
	A	0	2	very cobbly loam, 40 percent rock fragments
	Bw1	2	4	very gravely heavy loam, 60 percent rock fragments
	Bw2	4	7	very gravely clay loam, 60 percent rock fragments
	2R	7		basalt
Reywat	*Reywat very stony loam--rangeland.*			
Source:	http://www2.ftw.nrcs.usda.gov/osd/dat/R/REYWAT.html			
	A1	0	3	very stony loam, 20 percent gravel, 15 percent cobbles, and 20 percent stones;
	A2	3	6	very stony loam, 20 percent gravel, 15 percent cobbles, and 20 percent stones;
	Bt1	6	14	very gravely clay loam, 40 percent gravel and some cobbles;
	Bt2	14	19	very gravely clay loam, 40 percent gravel and some cobbles;
	R	19		basalt;
Farrot	*Farrot coarse sandy loam - rangeland;*			
Source:	http://www2.ftw.nrcs.usda.gov/osd/dat/F/FARROT.html			
	A1	0	4	coarse sandy loam,
	A2	4	10	coarse sandy loam,
	Bt1	10	15	sandy clay loam, 5 percent fine pebbles;
	Bt2	15	24	sandy clay loam, 10 percent fine pebbles;
	C	24	30	very gravely coarse loam, 40 percent pebbles;
	R	30		slightly weathered quartz diorite
Gabica	*Gabica gravely loam--rangeland*			
Source:	http://www2.ftw.nrcs.usda.gov/osd/dat/G/GABICA.html			
	A	0	5	gravely loam, 15 percent pebbles;
	Bt1	5	9	very gravely loam, 50 percent pebbles and 5 percent cobbles;
	Bt2	9	15	very gravely clay loam, 45 percent pebbles and 5 percent cobbles;
	R	15		fractured bedrock

Glasgow *Glasgow silt loam*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/G/GLASGOW.html

Ap	0	6	silt loam
A	6	12	silt loam
2Bt	12	16	clay
2Btk	16	24	clay loam
3R	24		fractured volcanic tuff.

Babbington *Babbington loam*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/B/BABBINGTON.html

A	0	4	Loam, 5 percent gravel
Bt1	4	13	Clay Loam
Bt2	13	29	Clay Loam, 27 to 35 percent clay
Btk	29	42	Loam, 24 to 32 percent clay, 0 to 10 percent, mainly gravel
2C	42	50	Loamy Sand, 6 to 60 percent gravel

Hoot *Hoot very cobbly loam, rangeland.*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/H/HOOT.html

A1	0	2	very cobbly loam, 20 percent pebbles, 20 percent cobbles and stones;
A2	2	4	very gravely loam, 40 percent pebbles, 10 percent cobbles;
Bt	4	14	extremely gravely clay loam, 50 percent pebbles, 15 percent cobbles;
R	14		andesite.

Nannyton *Nannyton fine gravely sandy loam*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/N/NANNYTON.html

A1	0	2	fine gravely sandy loam, 20 percent gravel
A2	2	7	loam, 5 percent gravel;
Bt	7	14	clay loam, 5 percent gravel;
Bk1	14	17	loam, 5 percent gravel;
Bk2	17	23	fine sandy loam, 5 percent gravel;
2Bk3	23	27	gravely coarse sandy loam, 25 percent gravel;
2Bk4	27	39	gravely loamy coarse sand, 30 percent gravel;
3Bk5	39	60	fine sand and fine gravel,

Larimer *Larimer fine sandy loam - grassland.*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/L/LARIMER.html

A	0	4	fine sandy loam, 5 percent gravel;
B	4	7	light loam, 5 percent gravel;
Bt	7	18	heavy loam, 5 percent gravel
Bk	18	22	loam, 5 percent gravel;
Ck1	22	30	gravely sandy loam, 20 percent gravel;
2Ck2	30	60	relatively clean gravel, cobbles, and sand, 80 % gravel

Searla *Searla very gravely loam, rangeland.*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/S/SEARLA.html

A	0	5	gravely loam, 10 percent cobbles and 20 percent gravel;
Bt1	5	12	very gravely clay loam, 10 percent cobbles and 30 percent gravel;
Bt2	12	19	very gravely sandy clay loam, 5 percent cobbles and 35 percent gravel;
Btk	19	32	very gravely sandy clay loam, 20 percent cobbles and 25 percent gravel;
Bk1	32	39	very gravely sandy clay loam, 25 percent cobbles and 30 percent gravel;
Bk2	39	60	very gravely sandy loam, 25 percent cobbles and 30 percent gravel;

Bullrey *Bullrey very gravely loam, rangeland*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/B/BULLREY.html

A1	0	4	very gravely loam
A2	4	9	very gravely loam, 45 percent coarse pebbles
Bw1	9	14	very gravely loam, 50 percent coarse pebbles
Bw2	14	22	very gravely loam, 55 percent coarse pebbles
C1	22	26	very gravely loam, 35 percent coarse pebbles
C2	26	48	gravely sandy loam, 20 percent pebbles
C3	48	60	extremely gravely sandy loam, 60 to 70 percent gravel

Takeuchi *Takeuchi fine gravely coarse sandy loam*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/T/TAKEUCHI.html

A1	0	4	fine gravely coarse sandy loam, 15 percent gravel;
A2	4	12	fine gravely coarse sandy loam, 15 percent gravel;
Bt	12	18	fine gravely coarse sandy loam, 20 percent fine gravel;
BC	18	26	fine gravely coarse sandy loam, 20 percent fine gravel;
Cr	26	36	moderately weathered granite or quartz monzonite
R	36		somewhat fractured and slightly weathered granite or quartz monzonite

Kanlee *Kanlee fine gravely coarse sandy loam*

Source: http://www2.ftw.nrcs.usda.gov/osd/dat/K/KANLEE.html

A1	0	10	fine gravely coarse sandy loam, 15 percent gravel;
AB	10	14	fine gravely coarse sandy loam, 15 percent gravel;
Bt1	14	17	fine gravely sandy clay loam, 15 percent gravel;
Bt2	17	24	fine gravely sandy clay loam, 15 percent gravel;
Bt3	24	32	fine gravely sandy clay loam, 30 percent gravel;
Cr	32	45	weathered granite

Appendix E

Parameter Ranges

Parameters		Initial Range		Refined Range	
Short Name	**Definition**	**Min**	**Max**	**Min**	**Max**
River Roughness	River roughness coefficient	0.01	6	0.01	6
DRNROW	A drainage index, set between 0 and 1	0	1	0	1
SDEPROW	Permeable depth of the soil column.	0	4.1	0.48	1.2
FAREROW	Fractional area a tile represents in a grid cell	0	5	0	1
DDENROW	Drainage density	0.0001	1000	0.0001	10
XSLPROW	Valley slope	0.0001	1	0.01	0.5
GRKFROW	Change in lateral conductivity at depth H_0	0.0001	1	0.0001	100
WFSROW	Coefficient for surface roughness	0.0001	100	0.0001	0.2
WFCIROW	lateral k_{sat} at surface	0.0001	100	0.0001	100
% Sand	Percent sand in layers 1,2 and 3	0	100	0	100
% Clay	Percent clay in layers 1,2 and 3	0	100	0	100
% Organic	Percent organic in layers 1,2 and 3	0	100	0	100
ZSNLROW	limiting snow depth	0.01	1	0.1	1
ZPLSROW	maximum water ponding depths for snow cover areas	0.01	1	0.01	0.1
ZPLGROW	maximum water ponding depths for snow free areas	0.01	1	0.01	0.1

Parameters		Initial Range		Refined Range	
Short Name	Definition	Min	Max	Min	Max
LNZ0ROW	Natural logarithm of the roughness length for the vegetation.	-10	100	-3	0.6
ALVCROW	Visible albedo for the vegetation	0	100	0.03	0.1
ALICROW	Near infrared albedo for the vegetation	0	100	0	0.5
RSMNROW	Minimum stomatal resistance for the vegetation	0	1000	50	300
VPDAROW	Coefficient of stomatal resistance to vapour pressure deficit for the vegetation	0	100	0.1	1
VPDBROW	Coefficient of stomatal resistance to vapour pressure deficit for the vegetation	0	100	0.1	2
PSGAROW	Coefficient of stomatal resistance to soil water suction for the vegetation	0	1000	50	200
PSGBROW	Coefficient of stomatal resistance to soil water suction for the vegetation	0	1000	1	10
QA50ROW	Coefficient of stomates to light for the vegetation.	0	1000	10	50
LAMXROW	Maximum leaf area index for the vegetation	0	100	1.5	10
LAMNROW	Minimum leaf area index for the vegetation	0	100	0	10
CMASROW	Standing biomass density ($kg \cdot m^{-2}$) for the vegetation	0	100	0	50
ROOTROW	Rooting depth for the vegetation	0	100	0.15	1

Appendix F
Sensitivity Analysis for Streamflow Calibration

With so many parameters in a hydrological model, it is important that calibration be supported by statistical techniques such as sensitivity analysis, identifiability analysis and auto-calibration (van Grievsven et al., 2002). Sensitivity analysis identifies parameters that do or do not have a significant influence on model simulations of real world observations for specific catchments (van Grievsven et al. 2006). Parameter sensitivity becomes particularly important when dealing with highly distributed models because of the large amount of computational time required to run the model and the large number of parameters being calibrated.

A straight forward method of measuring parameter sensitivity is through OAT (One-factor At-a-Time) design. By examining the change in the model objective function and the change in the model parameter (typically normalized by the parameter range) we can assign a set sensitivity for that parameter relative to the other parameters in the model. By post processing the results of a DDS model calibration run, the relative sensitivity of different parameters can be determined. This is because, at the end of the DDS search, most new solutions it evaluates differ from the best solution found so far by only a single parameter and thus OAT sensitivities can be computed. Generally with this type of sensitivity analysis, the sensitivity of each parameter will change depending on the parameters space and model objective function.

A sensitivity analysis was performed on each case study using one-at-time objective function evaluations for each parameter. The sensitivity was determined using the following equation.

$$SAi = \frac{|\Delta NS|}{|\Delta Xi| \div (X_{\mathrm{i\,max}} - X_{\mathrm{i\,min}})}$$

where Xi is the OAT parameter modified, and the absolute change in the objective function (Nash-Sutcliffe, ΔNS) is divided by the normalized change in the Xi parameter. The nature of this type of sensitivity analysis is highly variable depending on the location of the optimization algorithm within the parameter space. Therefore, for each parameter, the SAi measures available for all OAT DDS parameter perturbations are averaged together. As a result, different parameters will be sensitive under different conditions and therefore, the results will never be absolute. When unlimited computational efficiency is available the best results would likely be achieved with all the parameters being calibrated. As that is not often an option, a general list of parameter sensitivities was developed as a by-product of the model calibrations performed as part of this thesis work.

To address the soil parameters (percentages of sand, clay, and organic) in each layer, the sensitivities for each layer were treated as one cumulative value. This was necessary in that while individually the sand, clay, and organic may not have equal sensitivity results, all three parameters need to be calibrated as one soil layer.

The outputs of model calibration were post processed to calculate the average sensitivity for each parameter. Each calibration run produced slightly different sensitivity results, so the each parameter set produced was ranked and then averaged based on rank. By ranking the parameter sets this effectively normalizes the range in each sensitivity result while maintaining all of the sensitivity information. The ranked results of each sensitivity trial are summarized in Table F-1.

Table F-1: Sensitivity ranking information with parameters ranked from most sensitive (1) to least sensitive (25)

Average Rank	Parameter
1	DDENROW
2	XSLPROW
3	WFCIROW
4	LNZ0ROW
5	Soil Layer 3
6	ZSNLROW
7	Soil Layer 1
8	Soil Layer 2
9	SDEPROW
10	VPDAROW
11	LAMXROW
12	RSMNROW
13	QA50ROW
14	GRKFROW
15	LAMNROW
16	CMASROW
17	VPDBROW
18	ALICROW
19	ALVCROW
20	ROOTROW
21	PSGAROW
22	PSGBROW
23	ZPLSROW
24	WFSROW
25	ZPLGROW

References:

van Griensven, A., Francos, A., Bauwens, W. (2002). Sensitivity analysis and auto-calibration of an integral dynamic model for river water quality. *Water Science and Technology*, Vol45 No. 9:325-332.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R. (2006). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, 324(2006) 10-23.