Assessment of SWAT to Enable Development of Watershed Management Plans for Agricultural Dominated Systems under Data-Poor Conditions

by

Javier Mauricio Osorio Leyton

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biological Systems Engineering

Mary Leigh Wolfe, Chair
Darrell J. Bosch
Conrad D. Heatwole
Christopher W. Zobel

May 2, 2012
Blacksburg, Virginia

Keywords: Watershed management, Uncertainty analysis, Data-poor environments, Watershed modeling, Digital soil mapping

Assessment of SWAT to Enable Development of Watershed Management Plans for Agricultural Dominated Systems under Data-Poor Conditions

Javier Mauricio Osorio Leyton

Abstract

Modeling is an important tool in watershed management. In much of the world, data needed for modeling, both for model inputs and for model evaluation, are very limited or non-existent. The overall objective of this research was to enable development of watershed management plans for agricultural dominated systems under situations where data are scarce. First, uncertainty of the SWAT model's outputs due to input parameters, specifically soils and high resolution digital elevation models, which are likely to be lacking in data-poor environments, was quantified using Monte Carlo simulation. Two sources of soil parameter values (SSURGO and STATSGO) were investigated, as well as three levels of DEM resolution (10, 30, and 90 m). Uncertainty increased as the input data became coarser for individual soil parameters. The combination of SSURGO and the 30 m DEM proved to adequately balance the level of uncertainty and the quality of input datasets. Second, methods were developed to generate appropriate soils information and DEM resolution for data-poor environments. The soils map was generated based on lithology and slope class, while the soil attributes were generated by linking surface soil texture to soils characterized in the SWAT soils database. A 30 m resolution DEM was generated by resampling a 90 m DEM, the resolution that is readily available around the world, by direct projection using a cubic convolution method. The effect of the generated DEM and soils data on model predictions was evaluated in a data-rich environment. When all soil parameters were varied at the same time, predictions based on the derived soil map were comparable to the predictions based on the SSURGO map. Finally, the methodology was tested in a data-poor watershed in Bolivia. The proposed methodologies for generating input data showed how available knowledge can be employed to generate data for modeling purposes and give the opportunity to incorporate uncertainty in the decision making process in data-poor environments.

Dedication

To my mother, Hortensia Leyton

and

to my nieces: Ana, Camila and Heliana,

wishing one day they will walk this way as well.

Dedicatoria

A mi mamá, Hortensia Leyton

y

a mis sobrinas: Ana, Camila y Heliana

con la esperanza de que un día sigan este camino.

Acknowledgments

I would like to express my appreciation to my mentor, Dr. Mary Leigh Wolfe. Many things to be thankful for: her kindness, patience, friendship, and guidance throughout my program. Thanks for all the opportunities, for making possible this invaluable experience, and moreover, for the privilege to work under her direction.

I would like to acknowledge my committee members: Dr. Darrell Bosch, Dr. Conrad Heatwole, and Dr. Christopher Zobel for the assistance and the insightful guidance.

I am thankful to:

- Dr. Cully Hession for sharing his computer resources.
- Mrs. Barbara Wills and Mr. Denton Yoder for all the help, kindness and friendship.
- Mirco Peñaranda and Miguel Cabrera, for all the help during the field work.
- My colleagues at the BSE department, for the pleasure to share time with and for making me feel home away from home.
- My soccer team, the Black Knights Football Club.

This study was made possible through the support of the Department of Biological Systems Engineering at Virginia Polytechnic Institute and State University.

Table of Contents

List of figures

List of tables

Chapter I. Introduction

Agricultural production systems are complex and variable, comprising a large number of subsystems that interact dynamically over time. These complex systems pose a potential environmental risk due to the nature of the production system. Moreover, agricultural systems have high levels of uncertainty because they are influenced by biological, climatic, economic and social factors.

The type of agricultural system practiced depends on the local conditions, availability of resources and environmental limitations. Generally speaking, agricultural systems share the same production limiting factors: scarce water, soil erosion and overgrazing. Because of adverse climatic and geographic conditions and space limitations, it is a challenge in many locations to maintain reasonable agricultural production levels without overusing natural resources. Agricultural systems are important in terms of agricultural capacity, in which, unsuccessful traditional practices, have increased pressures on the soil, water, and nutrient resources, thus endangering sustainability. This condition creates the need for improved agricultural production systems that embrace sustainable use of resources and pollution control of surrounding water systems.

Understanding the agricultural system's behavior in a given location is the first step towards the development of watershed management policies and plans. However, due to the nature of complex watershed systems, the assessment of environmental impacts of natural processes and human activities requires a holistic approach that involves the use of decision support tools. Simulation models are useful tools to cope with the complexity of real systems; however, there is not a "best" model that will fit all situations. Difficulties in environmental modeling are due to natural complexity of the system, spatial heterogeneity and lack of available data. In fact, many places in the world where modeling could be useful are scarcely gauged. Measurements available in these regions are often scarce, limited, incomplete or nonexistent, poor-quality and often uncertain. In this context, a "data-poor environment" represents a place where there is not enough data to accurately characterize and evaluate the watershed under study.

Data limitation is probably the most significant challenge in watershed modeling because it restrains the use of data-intensive models and might compromise the ability to assess environmental impacts in a watershed of interest. Two different sets of data are particularly difficult to obtain: information to characterize the watershed (climate databases, soils properties,

etc.) and measured data sets for model evaluation (sediment yield, nutrient concentration, runoff volume, etc.).

Despite the lack of reliable data, modeling analysis is still required to support planning and decision-making. Consequently, research is needed to quantify the uncertainty in model predictions due to input data likely to be absent in data-poor environments and to develop methods that combine scarce data with creative alternatives and expert knowledge to improve/complete available information to apply data-intensive models in data-poor environments. The results of such research would contribute to development of a framework for use of data-intensive models under data-poor conditions to develop watershed management plans.

1.1. Objectives

The overall objective of the proposed research is to enable development of watershed management plans for agricultural dominated systems under situations where data are scarce.

The specific objectives necessary to realize this research objective are the following:

1. Characterize the uncertainty in SWAT's outputs due to input parameters likely to be lacking in data-poor environments;
2. Develop methods to gather qualitative and quantitative data that will allow modeling agricultural watershed systems under data-poor environments; and
3. Determine to what extent the proposed methodology to use SWAT with limited data will be able to represent water quality impacts of agricultural watershed systems in data-poor environments.

1.2. Research questions

The research questions to be answered are:

1. What is the capability of SWAT to generate reliable predictions in data-poor environments?
2. How can available knowledge be employed to generate data for modeling purposes?
3. What are the key input data requirements to reduce predictive uncertainty in data-poor environments?
4. Does characterization of uncertainty help to overcome the problem of not being able to validate a model in data-poor environments?

Chapter II. Literature Review

The purpose of this literature review is to provide the reader with a general overview of the application of data-intensive models in data-poor environments, with an emphasis on the Soil Water Assessment Tool (SWAT). The first part of this chapter gives a brief description of past and current work that used SWAT for watershed modeling around the world. Following this, a brief discussion of model evaluation methods is presented. Next, a review of uncertainty analysis and examples of studies that used Monte Carlo Simulation are presented. Then, spatial variation of soil properties and processes are reviewed. Finally, a brief overview of digital soil mapping is given.

2.1. Application of data-intensive models in data-poor environments

The ability of a model to simulate a watershed system depends on how well the watershed processes are represented by the model and how well the watershed system is described by model input parameters (Tripathi et al., 2006). The application of data-intensive simulation models depends upon several factors and conditions: the purpose of the study, understanding the nature of the watershed (natural complexity of the system, spatial heterogeneity and temporal variability), data limitations (quantity and quality), and computational procedures of the model (Letcher et al., 1999). Consequently, to perform watershed analysis, the right model selection depends on the purpose of the research and the availability of data.

The type and amount of data required for watershed-scale simulation vary by hydrologic model and intended application. Whether a model is data-intensive or simplistic, it is desirable to have high quality and quantity data. One must be sure to consider data requirements, alternative data sources, and proper data management in order to ensure a cost-effective system for collecting high quality data (Schafer and Hanlon, 2001). Lack of available input data means that many of the model parameters must be determined through calibration. This leads to a problem in that many best parameter sets will fit observed data, therefore, physical interpretation of the parameter will be questionable (Wheater et al., 1993; Letcher et al., 1999). Wheater et al. (1993) also added that if parameters cannot be uniquely identified, then they cannot be deterministically linked to catchment characteristics.

Different approaches to use available information for applying data-intensive distributed models in relatively data-poor environments have been investigated. Letcher et al. (1999) combined an empirical landscape-factor with simple a conceptual model to predict runoff. The landscape-factor model uses a simple empirical but landscape driven method that allows for the

dynamics of runoff generation in response to precipitation in ungauged watersheds. Elshorbagy and Ormsbee (2006) illustrated the potential use of an object-oriented simulation environment based on the concepts of system dynamics for surface water quality for management purposes, especially to cope with insufficient data for water quality modeling. They affirm that a model capable of representing complex systems in a realistic way is possible when object-oriented simulation approach is used. They conclude that object-oriented simulation based on the concepts of system dynamics is not a replacement for traditional hydrologic models but rather a feasible alternative in data-poor conditions.

Zilli Bacic et al. (2008) reported AgNPS was applied in a data-poor environment in Brazil with a limited observation dataset. The model was calibrated based on a best guess for model parameters, which was defined as initial parameter values adapted as close as possible to reality according to available data and authors' experience. Since quantitative data for model calibration and validation were lacking, this "best guess" scenario was created as an alternative to compensate in part for limited calibration data. The previous study by Zilli Bacic et al. (2008) shows that expert knowledge and information reported in the literature will allow the implementation of modeling in data-poor environments. However, the authors also conclude that even in a data-poor environment some data must be acquired.

Another way to maximize model usefulness with minimal data is to aggregate input data (FitzHugh and Mackay, 2001). However, it has been demonstrated that model predictions vary depending on the level of aggregation of input data (Jha et al., 2004; Arabi et al., 2006; Chang, 2009). Therefore, proper model use requires an understanding of how model predictions vary according to level of data aggregation and whether or not those variations can be attributed to differences in watershed characteristics (FitzHugh and Mackay, 2001; Chang, 2009). Distributed models divide the watershed into smaller units to represent heterogeneity within the watershed and model outputs are affected by geomorphologic resolution (Arabi et al., 2006). There is a direct relationship between the number and size of the homogenous units over which calculation will be performed and the detail of input requirements to run the model. More detail in the input data is required to better describe spatial variability of the watershed.

Jha et al. (2004), Rouhani et al. (2006), and Chang (2009) highlighted how spatial resolution effects resulting from watershed subdivision have strong influence on model-based predictions of long-term impacts on fate and transport of sediments and nutrients within watersheds. Reducing the size and increasing the number of sub watersheds would be expected to affect the simulation results from the entire watershed (Tripathi et al., 2006). When

analyzing the impact of input parameter aggregation on model output, FitzHugh and Mackay (2001) concluded that the relationship between outlet sediment yield and sub watershed size changes depending on the watershed characteristics, whether the watershed is source-limited or transport-limited.

On the contrary, Arabi et al. (2006) studied watershed subdivision for assessment of the impacts of BMPs on sediment and nutrient yields. They concluded that predicted reductions of sediment and nutrient yields as a result of BMP implementation were insignificant when more coarse levels of subdivision were applied. Additionally, Tripathi et al. (2006) and Jha et al. (2004) found that the level of watershed subdivision did not affect the accuracy of simulations, at least for streamflow. Also, Ghaffari (2011) concluded that decreasing the resolution above 50 m did not substantially affect the simulation of runoff but it did have an impact on simulation of sediment yield, which allows the possibility that data generalization can be made without affecting model predictions, at least to some degree and for some constituents.

Lopes and Canfield (2004) concluded that the selection of the appropriate level of watershed representation could have important theoretical and practical implications on modeling. However, it is impossible to divide a watershed into an unlimited number of sub watersheds (Chang, 2009). Mamillapalli et al. (1996) reported that there was a threshold beyond which higher resolution of data does not produce better results of predicted flow. Given that, Jha et al. (2004) and Chang (2009) recommend that watershed assessment based on modeling should include a sensitivity analysis with varying sub watershed size and number. There is not any established method for determining the optimal sub watershed/hydrologic response unit (HRU) configuration. Additional research is necessary to develop criteria that relate the watershed subdivision and the detail of data required to adequately represent its characteristics, while keeping the accuracy of model prediction of runoff, sediments and nutrients fate and transport. It is evident that poor catchment representation of important hydrological features may lead to poor performance of the model.

2.2. Soil Water Assessment Tool under data-poor conditions

SWAT is a physically based, semi-distributed, long-term continuous, watershed scale model that makes calculations on a daily time step to predict the impact of climate, land use, soil type, topographic characteristics, and land management practices upon hydrology, sediment, nutrients, pesticides, and bacteria in large ungauged watersheds (Bora and Bera, 2003; Arnold and Fohrer, 2005; Gassman et al., 2007). The model uses readily available input databases and does not require calibration; however, it is not capable of performing single-event flood routing

(Borah and Bera, 2003). It is the continuation of a long-term effort of nonpoint source pollution modeling by the USDA-Agricultural Research Service (ARS), including development of CREAMS (Knisel, 1980), GLEAMS (Leonard et al., 1987), SWRRB (Williams et al., 1985), EPIC (Williams et al., 1990), and ROTO (Arnold et al., 1995).

SWAT allows a basin to be divided into sub-watersheds and hundreds or thousands of lumped-homogeneous areas called Hydrologic Response Units (HRU), which are unique combinations of landuse – soil – slope - management (Neitsch et al., 2002) on which the computations are performed. The HRUs do not have spatial orientation (Chaubey and White, 2006; Gassman et al., 2007). The model simulates major hydrologic processes (overland and in-stream), such as water balance, sedimentation, plant dynamics, chemical fate and transport.

The SWAT model has many built-in databases, including weather, soils, crops, land use, fertilizers and pesticides (Gassman et al., 2007). This characteristic facilitates the modeling process, especially if the systems to be modeled are inside the U.S. Built-in databases can also be used as a best available reference for initial work outside the U.S. While initially developed for use mainly inside the U.S., the SWAT model has been reported to be an effective tool for assessing water resource and nonpoint-source pollution problems for a wide range of scales and environmental conditions across the globe (Gassman et al., 2007). In Europe, the SWAT model has been used in several ongoing major projects for the European Framework Directive (Arnold and Fohrer, 2005). SWAT applications have been reported in many other countries (e.g., cite a few examples); however, examples of applications in Latin America are scarce.

SWAT has reported applications in Canada (Chanasyk et al., 2006; Mapfumo et al., 2004), China (Cheng et al., 2006; Zhang et al., 2007), Australia (Watson et al., 2003; Cao et al., 2006), India (Tripathi et al., 2006; Behera and Panda, 2006), Africa (Jayakrishnan et al., 2005; Birhanu, 2007; Bouraoui et al., 2005) and Europe (van Griensven et al., 2006: Barlund et al., 2007; Bouraoui et al., 2002). Most of the studies concluded that the SWAT model has a good potential for application in hydrology/water quality assessment in countries around the world under a wide variety of watershed characteristics. Most of these reports do not mention characteristics of the input data nor how data limitation was overcome. Only a few of them recommended further testing/customizing the SWAT model for different watershed conditions. Some articles indicated that the model performance efficiency is higher when coupled with the use of high-resolution data sets. However, under different characteristics, others (Tripathi et al., 2006; Jha et al., 2004) indicated that higher resolution of spatial data does not necessarily improve the performance of SWAT.

Using SWAT and limited digital data on land use, soil and elevation, Jayakrishnan et al. (2005) modeled the hydrology of Sondu River Watershed of Kenya. The use of the model was limited by the lack of detailed digital data on landuse, soil and elevation for model input. Digital data were coarse (1 km resolution DEM) and the information of one soil type was assumed over the entire watershed (3050 km^2). Climate and streamflow information were available but data were missing for several days. No observed sediment data were available to calibrate/evaluate the model predictions. Despite the limited data and poor Nash-Sutcliffe efficiencies (NSE) of the evaluated results, Jayakrishnan et al. concluded that the application of SWAT is possible under African conditions and that better elevation data, and more detailed soil and weather data combined with detailed parameter calibration efforts, should improve the results.

Watson et al. (2003) applied SWAT to the Woady Yaloak River watershed in Australia. The model performed extremely well at predicting annual streamflows (NSE 0.75 and 0.77), but problems with groundwater and eucalyptus growth (Leaf Area Index simulation) constrained the ability to accurately model water balance. Watson et al. proposed to modify the groundwater and LAI components to adequately represent Australian conditions when there is lack of information to represent the watershed. SWAT was applied for modeling the WeruWeru watershed in Northern Tanzania (Birhanu, 2007). Results for predicted mean daily streamflow were reported exactly as observed during the water balance simulation. Authors of this study concluded that the SWAT model could be used for watershed studies in mountainous catchments in tropical regions.

SWAT sediment and nutrient simulations vary quite dramatically with the number and size of sub watersheds and also SWAT model predictions are very sensitive to HRU distribution (Arabi et al., 2006; Cho, 2010). Therefore to answer the question about how many sub watersheds to use when modeling with SWAT will depend on need to minimize data and watershed characteristics. Rouhany et al. (2006) pointed out that runoff volume predicted with SWAT has an impact due to the number of sub watersheds and HRUs. The threshold value to define the number of sub watersheds and HRUs as a function of the geomorphologic properties and hydrological behavior of watersheds still remains unanswered. Tripathi et al. (2006) reported the effect of watershed subdivision on simulated water balance components using SWAT for the Nagwan watershed in India. Their results revealed that the number and size of sub-watersheds did not appreciably affect surface runoff, but had a significant effect on the water balance components. Jha et al. (2004) reported that variation in the total number of sub watersheds had very little effect on streamflow, but the opposite result was found for sediment,

nitrate, and inorganic phosphorous. Jha et al. (2004) and Chang (2009) recommended that restricting the watershed subdivision to an optimum threshold level to reduce input preparation efforts and subsequent computational evaluation. There is still research to be done to establish how data resolution and aggregation impact SWAT prediction of different watershed sizes. HRUs are assumed to be homogeneous in hydrologic response, because are assumed with parameters representative of the entire unit, therefore the less number of those homogeneous units the less number of input parameters will be necessary to characterize the model (Jha et al., 2004; Tripathi et al., 2006). However, the size of the unit to be modeled affects the homogeneity assumption, since larger areas are more likely to have variable conditions (Jha et al., 2004; Jayakrishnan et al., 2005; Tripathi et al., 2006), therefore, data generalization could affect accuracy of model predictions.

The SWAT model has been adopted and applied worldwide in a wide range of applications and conditions (Gassman et al., 2010). However, clear limitations exist for how the model can be applied for some problems due to lack of data and/or modeling expertise and lack of suitable model algorithms. Several publications highlight the same major limiting factor in the success of SWAT simulations: lack of available high quality input data (Jacobs and Srinivasan, 2005; Krysanova et al., 2005). Also, the need for measured data collection to improve SWAT model evaluation was highlighted by several authors (e.g., Watson et al., 2003; Tripathi et al., 2006).

Krysanova et al. (2005) suggested that the following problems related to regional applications should be addressed and discussed: (a) general data needs and options to reduce data requirements, (b) choice of strategy for model validation, and (c) analysis of uncertainty related to model parametrization and input data. Other researchers (cite references) pointed out the necessity of continuing research about how different spatial and temporal input data have to be integrated in a consistent way. Therefore, the question of what effects different spatial resolutions and data aggregation of the input data have on the simulation results is still not fully answered. Consequently, there are no established scientifically sound criteria that support the use of data intensive models under data-poor conditions.

This review demonstrates one of the biggest challenges when any hydrologic modeling is performed outside North America. In the United States, national-scale digital data, such as soils and land use, are readily available for hydrologic analysis and modeling studies. The availability of detailed model input data at such scales is very limited outside the United States, especially in Africa and Latin America. Even with this limitation, detailed hydrologic/water quality

models, developed and studied widely in the United States, can be applied in other latitudes, although the need for additional model input data collection to improve model parameter calibration and simulation results must be stressed. Finally, the SWAT model has good potential for application in hydrologic/water quality studies in countries around the world as a tool to develop time and cost-efficient analyses for watershed/water resources management. SWAT as a project is currently supported by USDA. This guarantees that the model will keep being developed, with up to date documentation, large number of publications, on-line technical support, workshops and training programs, and there is an ongoing effort to promote the use of the model outside the U.S (Gassman et al., 2010). Additionally, SWAT has open source code, which is available to the public so scientists can further develop or include new research in order to improve model performance and applicability.

2.3. Model evaluation

Moriasi et al. (2007) pointed out that watershed models can save time and money because of their ability to perform long term simulation of the effects of watershed processes and management activities on water quality and water quantity. Nonetheless, in order to use model outputs for tasks ranging from regulation to research, models should be scientifically sound, robust, and defensible (U.S. EPA, 2002).

A model that accurately predicts the impact of alternative management scenarios on delivery of nutrients (nitrogen and phosphorous) to surface and ground water is an essential tool in developing watershed management plans focused on reducing nutrient delivery to surface and ground water. Before using a model for such plan development, it is important that the model be evaluated with respect to its predictive capabilities related to the delivery of nutrients to surface and ground water.

According to Moriasi et al. (2007), “model evaluation” refers to the applicable steps of sensitivity analysis, calibration, validation, uncertainty analysis and application. The general objective for model validation is to obtain a reliable modeling tool that can be used to evaluate such responses. To create such a tool, it will be necessary to first calibrate and then validate the model. It is reported that SWAT does not need calibration (Borah and Bera, 2003). However, calibration alone is not sufficient to assess the predictive capability of the model. Singh et al. (2004) recommended that before the evaluation process starts the parameter to be evaluated must be selected and tolerance limits for model must be fixed as the criteria or critical values for evaluation. General performance ratings for model evaluation are presented on table 1 (Singh et al., 2004).

Table 1. General performance ratings for recommended statistics for a monthly time step (Singh et al., 2004).

Performance Rating	RSR[1]	NSE[2]	PBIAS (%)[3]		
			Streamflow	Sediment	N, P
Very good	0.00 ≤ RSR ≤ 0.50	0.75 < NSE ≤ 1.00	PBIAS < ±10	PBIAS < ±15	PBIAS < ±25
Good	0.50 < RSR ≤ 0.60	0.65 < NSE ≤ 0.75	±10 ≤ PBIAS < ±15	±15 ≤ PBIAS < ±30	±25 ≤ PBIAS < ±40
Satisfactory	0.60 < RSR ≤ 0.70	0.50 < NSE ≤ 0.65	±15 ≤ PBIAS < ±25	±30 ≤ PBIAS < ±55	±40 ≤ PBIAS < ±70
Unsatisfactory	RSR > 0.70	NSE ≤ 0.50	PBIAS ≥ ±25	PBIAS ≥ ±55	PBIAS ≥ ±70

(1) Root Mean Square Error observation standard deviation ratio (RSR)
(2) Nash-Sutcliffe Efficiency (NSE)
(3) Percent Bias (PBIAS)

The visual comparison, which often takes the form of graphical plots of simulated and observed flows, is a necessary first step in an evaluation. It provides a general overview of model performance and provides an overall feeling for model capabilities (ASCE, 1993). A general visual agreement between observed and predicted values will indicate adequate calibration and validation over the range of values being simulated (Singh et al., 2004). When the performance of a single model is evaluated, quantitative assessment is needed, which can be met using one or more statistical goodness-of-fit criteria (ASCE, 1993). Dimensionless techniques provide a relative model assessment.

In most watershed modeling projects, model output is compared to corresponding measured data with the assumption that all error variance is contained within the predicted values and that observed values are error free (Moriasi et al., 2007). The problem for model evaluation arises when the model is applied to data-poor conditions, where there is very little or no measured data to evaluate model predictions. Very little information can be found in the literature with regard to methods and approaches for model evaluation in conditions of scarce data. One option for evaluating predictions in data-poor environments is to take field measurements. According to Sivapalan (2003) in some situations, a solution to the ungauged basin problem will be to take some (a small number) field measurements to help model evaluation. However, the question is how long a measured record is necessary to obtain an optimal model evaluation. In addition, Seibert and Beven (2009) raised the question of how many measurements might be necessary to achieve a desired and cost-effective reduction in uncertainty. With the actual evaluation procedures, large datasets for calibration and validation are needed. Monitoring plans to collect field data are labor intensive and expensive, therefore, it

is unlikely that long-term data collection in an ungauged catchment would be possible (Seibert and Beven, 2009). Therefore, it can be inferred that new approaches are needed for evaluating model predictions in data-poor conditions.

Moriasi et al. (2007) noted that, in situations when a complete measured time series does not exist, the available data may not be sufficient for analysis using the recommended statistics. In such situations, comparison of frequency distributions and/or percentiles (e.g., 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th}) may be more appropriate than the quantitative statistics guidelines. Sivaplan et al. (2003) suggested that through the evaluation of the worth of data, new insights into appropriate predictive strategies, as a function of data availability, will be developed. Some research has demonstrated that, if selected in an intelligent way, a small fraction of data points in a longer time series might contain almost all information of the entire data series (Juston et al., 2009). Small datasets are most likely to be affordable and therefore to be collected, especially in remote places where the access to technology is not always possible. According to Seibert and Beven (2009) it is encouraging, that a hydrologically intelligent choice of a small number of observations performs well relative to either regularly or randomly chosen measurement times. The intelligent choice should reflect what the model is expected to do.

Grayson et al. (2002) presented methods for the comparison of patterns to assess hydrologic characteristics of model performance. These methods include visual comparison, point-to-point comparisons, spatial relationships between errors and landscape parameters, transects, and optimal local alignment. According to Grayson et al. (2008), visual assessment is probably the most powerful comparison method; however, the disadvantage of this method is that it does not provide a quantitative measure of model performance. A different concept, a type of reference watershed, was introduced by Sivaplan et al. (2003): a comparative evaluation of the predictive performances of a variety of models, when simultaneously applied to selected “gauged” basins, can lead to “cross-fertilisation” of ideas and insights into basin responses.

Gupta et al. (2008) pointed out that there is a strong need for sophisticated approaches to model evaluation to solve the problem of prediction in ungauged basins. Gupta et al. (2008) proposed a framework for a diagnostic approach to model evaluation based on signature index matching and described how it applies to the problem of prediction in ungauged basins. Signature-based evaluation enables the Bayesian inference framework to be applied to the problem of prediction in ungauged basins, by exploiting three kinds of information about the system: the signature-based likelihood, the local prior, and the signature-based regional prior.

The method provides a strong basis for diagnostic model evaluation and a framework for reconciling environmental theory with data.

2.4. Uncertainty analysis

It is not possible for any model to represent nature's processes perfectly (Muleta and Nicklow, 2005). Thus, simplifications must be made so scientific knowledge can be applied to represent the system (Shirmohammadi et al, 2006). The nature of watershed systems and the types of simplifications may lead to many different sources of uncertainty. According to Hession and Storm (2000) and MacIntosh et al. (1994), uncertainty is everywhere and, therefore, must be incorporated into watershed–level assessment and management to enhance the decision making process. Haimes (2004) defined uncertainty as "the inability to determine the true state of affairs in a system." Hattis and Burmaster (1994) stated "uncertainty is a description of the imperfection in knowledge of the true value of a particular parameter."

Uncertainty can be either knowledge or stochastic (Helton, 1994; Hession and Storm, 2000; Walker et al., 2003). Knowledge uncertainty can be reduced in several ways, for instance, by improved instrumentation and improvements in model formulation. In contrast, stochastic variability cannot be reduced, but can be quantified (Helton, 1994; Hession and Storm, 2000; Srivastav et al., 2007). Stochastic uncertainty is due to unexplained random variability of the natural environment (Helton, 1994; MacIntosh et al., 1994; Hattis and Burmaster, 1994). Knowledge uncertainty, on the other hand, is due to incomplete understanding or oversimplification of the complex and variable real system (Helton, 1994; MacIntosh et al., 1994;Haan, 1989; Muleta and Nicklow, 2005). Additionally, much of the uncertainty on modeling predictions can be also attributed to incomplete information -quantity and quality of the input data- and to uncertainty in the estimates of the parameters used as input for the analysis (Shirmohammadi et al. 2006; Muleta and Nicklow, 2005; Srivastav et al., 2007). Knowledge uncertainty in models can be further classified as model and data uncertainty (Hession and Storm, 2000). Srivastav et al. (2007) listed the following as the primary sources of knowledge uncertainty: input data, model parameters, model structure, and measured data used during calibration.

Helton et al. (2006) and O'Hagan (2006) define uncertainty analysis as the "determination of the uncertainty in analysis results that derives from uncertainty in analysis inputs." Muleta and Nicklow (2005) define it as the "technique of determining reliability of model predictions, accounting for various sources of uncertainty." Uncertainty analysis provides a method for quantifying the reliability and confidence in model predictions (Mowrer, 2000),

expressing this variability as a distribution of possible or expected values (Chapra and Reckhow, 1979). Loague and Corwin (1996) state, “quantification of uncertainty establishes the extent to which simulated results are reliable predictions of observed truth.”

Even though various approaches for representing uncertainty exist, probabilistic analysis is the most widely used method for characterizing uncertainty. Probability distributions are used to provide information about input uncertainty (Helton, 1994). Uncertainties associated with model inputs are described by probability distributions, and the objective is to estimate the output probability distributions (O’Hagan, 2006). Sampling from each of these distributions provides a set of possible model inputs, which are used to state an output quantity. Repeating this many times, and combining the outputs generated from each repetition, provides a distribution, which represents the uncertainty present in the output.

2.5. Uncertainty analysis of SWAT

Chaubey et al. (2005) analyzed the effect of input data resolution on uncertainty of SWAT predictions for Moores Creek watershed in Arkansas by running the model for seven different DEM resolutions. The authors concluded that DEM data at a finer resolution should be incorporated in SWAT in order to minimize uncertainties in predictions. Muleta and Nicklow (2005) applied SWAT to a watershed located in southern Illinois to analyze the uncertainty due to input data upon streamflow and sediment yield using Generalized Likelihood Uncertainty Estimation (GLUE). They found fairly consistent streamflow prediction with narrow uncertainty bounds, while sediment yield prediction was found to involve a great deal of uncertainty.

Shirmohammadi et al. (2006) presented three case studies of uncertainty analyses using SWAT, with Monte Carlo, Latin-Hypercube- Monte Carlo, and GLUE approaches. They concluded that large uncertainty should be expected when analyzing runoff due to curve number. Yang et al. (2008) compared five uncertainty analysis procedures for an application of SWAT to the Chaohe Basin in China. Differences and similarities between Generalized Likelihood Uncertainty Estimation (GLUE), Parameter Solution (ParaSol), Sequential Uncertainty FItting algorithm (SUFI-2), and a Bayesian framework implemented using Markov chain Monte Carlo (MCMC) and Importance Sampling (IS) techniques were analyzed. They recommended Bayesian-based approaches because of their solid conceptual basis.

Sohrabi et al. (2003) studied Monte Carlo simulation technique connected with Latin hypercube Sampling (LHS) to analyze the uncertainty of SWAT outputs concerning nutrients and sediment losses from agricultural lands. The authors concluded that using a best possible

distribution for the input parameters to reflect the impact of soils and land use diversity may be more accurate than using average values for each input parameter. Zhang (2009) conducted calibration and uncertainty analysis for SWAT using a Genetic Algorithm (GA) and Bayesian Model Averaging (BMA). The results obtained in the two watersheds showed that this combined method can provide deterministic predictions comparable to the best calibrated model using GAs. Finally, Van Greinsven and Meixner (2006) described several uncertainty analysis tools that have been incorporated directly within the SWAT model, including a modified Shuffled Complex Evolution (SCE) algorithm called "Parameter Solutions" (ParaSol), the Sources of Uncertainty Global Assessment using Split Samples (SUNGLASSES), and the Confidence Analysis of Physical Inputs (CANOPI).

2.6. Monte Carlo Simulation (MCS)

Monte Carlo simulation (MCS) is the most popular technique for quantifying uncertainty. MCS is a computerized mathematical technique that allows people to account for risk in quantitative analysis and decision making under uncertainty (Helton, 1994; Hattis and Burmaster, 1994; Shirmohammadi et al., 2006; Muleta and Nicklow, 2005). Advantages of this methodology are presented by Srivastav et al. (2007). They affirm MCS is cost-effective, suitable for models with many parameters, large application area, simultaneous sampling of parameters, direct estimation of distributions of outputs, and simple to use and implement. Hession and Storm (2000) stated that the Monte Carlo methodology allows for partitioning the different sources of uncertainty, which can provide valuable information for the decision-making process.

Simple Monte Carlo simulation (SMCS) performs uncertainty analysis to derive a distribution of model outputs by substituting a range of sampled random inputs. The model calculates results over and over, each time using a different set of random values from the probability functions. The result is a probability distribution of possible values or confidence interval (CI) instead of a single deterministic output (Palisade, 2011). Monte Carlo simulation quantifies parameter/input uncertainty, without specific attention given to other forms of error, such as model uncertainty or decision uncertainty. The accuracy of output uncertainty estimates depends obviously on the choice of the parameter distribution and also on the number of model simulations performed (Haan, 1989).

2.7. Spatial variation of soil properties and processes

Spatial and temporal variability are attributes of all soils; soils exhibit great differences from one location to another due mainly to pedogenesis factors (Hillel, 1998). Brady and Weil (2008) noted that changes across the landscape are mainly due to five factors: climate, parent material, organisms, topography and time. Climate and parent material influence soil variability at large scales, while most small-scale soil variations involve changes in topography (where slope is critical) and the effect of microorganisms.

Soil maps are representation of the soil spatial distribution. For spatially distributed models, soils data sets have two dimensions. The first dimension is a combination of polygons representing distribution of mapping units and soil classes. The second dimension is an attribute table for profile observations of soil horizons properties. Therefore, soil properties in a map vary horizontally and also vertically.

Soil properties can be classified as static or dynamic (Hillel, 1998). Static properties are attributes of the soils regardless of process. Dynamic properties are a response of the soil to induced processes. Static properties typically exhibit normal distribution curves. Dynamic properties have greater variation and tend to exhibit skewed distributions (log-normal). Recognizing the existence of spatial variability implies acceptance of uncertainty. We can never determine a property or quantify a process exactly. All that nature allows us to do is to assess the probability that the property lies within a specifiable range of values.

2.8. Empirical soil and land knowledge systems

Local communities have roots in ancient cultures, which have strong communal linkages and organization, leadership and environmental respect (Vale et al., 2007). Stakeholders carry ancient knowledge about environmental practices; they know which actions will lead to environmental degradation; they have ancient rotational systems for crops and grazing livestock, diversify production system for food security, use ancient ingenious irrigation systems to optimize water resource (Pawluk et al., 1992; Vale et al., 2007). Most of the ancient cultures from South America have valued the soil as one of the main sources for agriculture and livestock production; therefore, developed knowledge about soils characteristics, management and conservation.

Multidisciplinary research models, combining advanced scientific knowledge with indigenous knowledge for the improvement of the quality of life of rural people, have been applied all over the world (Riley, 2001; Barrera-Bassols and Zinck, 2003; Barrios and Trejo,

2003). The understanding of indigenous approaches to soil perception, classification and appraisal is the aim of ethnopedology (Barrera-Basol and Zinck, 2003). Several authors (Barrera-Bassols and Zinck, 2003; Riley, 2001; Barrios and Trejo, 2003) reported ethnopedology deals mainly with the formalization of local soil knowledge into classification schemes and the comparison of local and technical soils classification.

In many areas of the world, locals have developed complex systems of management of soils and land resources. For example, Barrera-Bassols and Zinck (2003) cited the following classification criteria used by ethnic groups: color and texture, consistency and soil moisture, organic matter, stoniness, topography, land use and drainage, fertility, productivity, workability, structure, depth and soil temperature. General principles of indigenous soils classification, such as knowledge about land management or recognition of basic soil attributes, can be applied to join local qualitative classification with scientific quantitative soil databases. However, there is very little information when it comes to relating empirical soil and land knowledge systems of rural populations with scientific soils databases (Barrera-Bassols and Zinck, 2003). Validation of ethnopedological knowledge with modern soil science is needed in order to provide more scientifically sound basis for technological change and also to use local knowledge for modeling purposes.

2.9. Digital soil mapping (DSM)

Soil is a fundamental natural resource of agricultural production (Ge et al., 2011; Scull et al., 2003; Aksoy et al., 2009). Soil properties are a significant input into models because they are involved in a large number of processes such as evapotranspiration, photosynthesis, and biogeochemical cycling (Palacios-Orueta and Ustin, 1998). In the United States, soil survey work began in 1896 by USDA with the aim of providing soil maps to aid in landuse decision making (Scull et al., 2003). Traditional soil survey concepts are based on qualitative recognition of soil properties in relation to landscape characteristics (Scull et al., 2003; Carre and Girand, 2002). Traditional soil survey consists of three steps (Cook et al., 1996): direct observation of ancillary data (aerial photography) and soil profile characteristics, observation of soil attributes to infer soil variation, and prediction of soil variation at unobserved sites. This process is then transformed into a cartographic model by drawing map unit boundaries on aerial photographs (photographic scale determines resolution of the soil map).

Spatially distributed soil information in many places of the world is scarce, incomplete or nonexistent. The main reason for lack of spatial soil data is that conventional soil survey methods are time consuming, labor intensive and expensive (Moore et al., 1993; Lagacherie

and McBratney, 2007; Saunders and Boettinger, 2007; Aksoy et al., 2009). Therefore, there is a need for methodologies that can help to fulfill this absence of one of the most important input datasets for modeling.

The use of digital elevation model (DEM) and Landsat Thematic Mapper (TM) imagery for soil mapping has become popular (McBratneya et al., 2003). Technological advances in geographic information science and remote sensing have created potential for improving the way that soil maps are produced (Carre and Girand, 2002; Aksoy et al., 2009; Leon et al., 2003). Along with technology, remotely sensed data have also become freely available, which made possible the use of alternative approaches for soil mapping. Previous studies have demonstrated the usefulness of digital data sources, like DEM and satellite data, for creating maps of large areas (McBratneya et al., 2003; Dobos and Montanarella, 2007). DEMs are used to derive terrain attributes that make terrain analysis feasible for predictive soils mapping (Aksoy, 2009). Remote sensing has made it possible to gather data about soils (Leon et al., 2003; Batchily et al., 2003). Banmgardner et al. (1985) provided extensive reviews of soil properties that affect reflectance and, therefore, the remote sensing of soils. Soils reflectance is affected by soil properties like particle size distribution, soil structure, surface roughness, moisture content, organic matter content, soil salinity and mineral content (Frazier, 1989; Banmgardner et al., 1985).

The soil cover is not simply randomly distributed (Jenny, 1941). Soil profile character is a function of climate, organisms, relief, parent material and time, implying that if the spatial distribution of the soil-forming factors is known, soil character may be inferred (Jenny, 1941; Scull et al., 2003). Therefore, it might be possible to make predictions about the spatial distribution of the soils in a particular area if some of the soil-forming factors can be defined by remote sensing technology and numerical modeling.

This prediction process, based on numerical models and remotely sensed data, has been defined by different authors in several ways. Scull et al. (2003) defined “predictive soil survey” as the process that takes information on what it is more-or-less known with a given level of uncertainty and infers unknown data with also a given level of accuracy. The same concept was named “soil inference systems” by Lagacherie and McBratney (2007). In more recent years, the term Digital Soil Mapping (DSM) has become the most used. DSM is defined by Lagacherie and McBratney (2007) as “the creation of spatial soil information systems by numerical models inferring the spatial and temporal variations of soil types and soil properties from soil observation and knowledge and from related environmental variables.”

Several authors (Carre and Girand, 2002; Aksoy et al., 2009; Scull et al., 2003; Lagacherie and McBratney, 2007) have presented different methodologies for mapping soil types using DEMs and satellite imagery with the aim of collecting and implementing the scientific results of studies to develop a quantitative methodology for soils mapping and supporting soils database creation for modeling purposes. The main hypothesis underlying these methods is that soil types result from environmental factors (Carre and Girand, 2002; Lagacherie and McBratney, 2007). Soil map unit delineation is based on two primary phenomena: terrain and lithology (Acron, 1965; Moore et al., 1993; Batchily et al., 2003). The soils mapping approach is based on polygons representing homogeneous physiographic and lithologic units of the earth's surface. Soil information is assigned to the polygons as soil associations, such that each soil represents a unique combination of terrain and soil characteristics (Acron, 1965; Moore et al., 1993; Dobos and Montanarella, 2007).

Lithology has been used to identify the distribution of soil-forming materials. Additionally, the use of remotely sensed images for lithological mapping is very well known (Li and Li, 2001). Several studies reported lithologic classification based on remote sensing (Alberti et al., 1993; Mather et al., 1998; Schetselaar et al., 2000; Ricchetti, 2000; Gad and Kusky, 2006). Alberti et al. (1993) and Boettinger et al. (2007) used combination of Landsat spectral bands can to represent geological and environmental covariates of parent material and soil. Color ratio composites are derived by dividing one band by another (Bennett et al, 1993; Mshiu, 2011). Band ratio images (5/3, 5/1, 7/5) and (7/5, 5/4, 3/1) were proved to be very effective in the lithological discrimination (Bennett et al, 1993; Gad and Kusky, 2006; Mshiu, 2011). Sultan et al. (1987) proved that band ratio image composition (5/7, 5/1, 5/4*3/4) is the most suitable for mapping lithological formations in arid zones. A detailed list of band ratio combinations with lithological mapping purposes can be found at Crippen (1989).

Chapter III. Methods

In this chapter, a summary of the research methods used in this study is followed by a detailed description of the methods used to achieve each objective. In order to characterize the uncertainty in SWAT's outputs due to input parameters likely to be lacking in data-poor environments, the Soil Water Assessment Tool (SWAT) was run for a data-rich environment to determine which input data are essential and which inputs are not, so SWAT can be run without affecting its capacity to make accurate predictions. Sensitivity of model predictions to soil parameters was approximated using the relative sensitivity (Sr) to determine parameters that most influenced predicted streamflow, sediment yield, and nutrient model outputs. The sensitive soil parameters then were included in the uncertainty analysis. Simple Monte Carlo simulation (SMCS) methodology was used to quantify the uncertainty in model outputs induced by uncertainty in inputs that are likely to be missing in data-poor conditions. Additionally, with the aim of defining whether or not coarse resolutions have a combined effect on soil parameters, two sets of input data were also sequentially changed: DEM resolution (10, 30 and 90 m) and the source of the soils database (SSURGO and STATSGO). Output distributions are presented as modified Tukey box plots providing a range and distribution of uncertainty as the 10^{th} and 90^{th} percentiles.

The results of the sensitivity and uncertainty analyses helped identify the qualitative and quantitative data for which estimation methods needed to be developed to allow modeling agricultural watershed systems under data-poor environments. Methods were required for DEM and soils map generation. Direct projection using cubic convolution as a resampling technique was applied to improve the resolution of freely available 90 m DEM. To generate a soil data set, map units were defined by overlaying a slope class map with a lithologic raster to represent the two primary soil forming factors: terrain and parent material. The methodology to generate the soils map attribute table is based on associating textural information of the top layer with a complex data base using a simple lookup procedure. The generated DEM and soils dataset were sequentially replaced in a data-rich environment to assess the impact of the generated datasets on model predictions, and uncertainty of those predictions, compared to those based on measured data in the data-rich environment. The comparison included six scenarios resulting from the combination of two sources of 30 m DEMs (USGS and resampled CGIAR-CSI) and three sources of soil databases (SSURGO, STATSGO, and SATSLP (generated in this study)).

The developed data generation methodology was then tested in a real data-poor environment to determine to what extent the proposed methodology to use SWAT with limited data will be able to represent water quality impacts of agricultural watershed systems in data-poor environments. The study site is the Huanquisco River watershed in Bolivia. Topographic information to run SWAT for a 10-year period (2002 - 2011) was derived from the original USGS 90 m resolution DEM resampled to a 30 m grid by direct projection using cubic convolution as a resampling technique. Results of a field soil survey were used to create a geospatial database of sampling points and textural proportions to complete the information required to create the soil map and associated attribute table. Management operations required to characterize the watershed were set based on the results of a questionnaire applied in situ. One more time, the same parameters and procedure were applied to quantify uncertainty of predictions. Finally, the model was run for a second time with a daily time step to perform a comparison between predicted and observed streamflow values.

3.1. Objective 1 - Characterize the uncertainty in SWAT's outputs due to input parameters likely to be lacking in data-poor environments.

3.1.1. Model and input data

For this study, the selected model is the Soil Water Assessment Tool (SWAT) developed by Arnold and other scientists at the USDA-Agricultural Research Service (Arnold et al., 1998). The aim was to determine which input data are essential and which inputs are not, so SWAT can be run without affecting its capacity to make accurate predictions. All input files were prepared with the model version for ArcGIS, ArcSWAT 2.3.4. Interface for SWAT2005 (Olivera et al., 2006). SWAT was run first for a data-rich environment, specifically, two watersheds from the continental U.S., a place where there is enough available information to comprehensively characterize a watershed system and evaluate the model. Detailed information to run the model and measured data to evaluate SWAT outputs were collected from the ARS Watershed Network (Neff, 1965).

Little River Experimental Watershed (LREW) is located in the headwaters of the Suwannee River Basin, Georgia. The watershed area is 334 km^2, with landuse classified as 50% forest, 41% mixed agricultural, 7% urban, and 2% water. Agricultural cropping rotations are variable, livestock production is scarce, and there are no permitted point source discharges in the watershed (Bosch et al., 2007). The LREW was instrumented with recording rain gauges, flumes and weirs to measure rainfall, streamflow and water chemistry since the late 1960s and has been in continuous operation since that time (Seyfried et al., 2001). Databases for climate,

streamflow, and water quality were downloaded from the anonymous ftp site (ftp://www.tiftoarss.org/) maintained by the USDA-ARS, Southeast Watershed Research Laboratory (SEWRL).

The Reynolds Creek Experimental Watershed (RCEW) comprises 239 km^2 located in the Owyhee Mountains of southwest Idaho. Reynolds Creek is a north flowing tributary of the Snake River. Elevation ranges from 1101 m to 2241 m. The area has a wide diversity in local climate, geology, soils, and vegetation. Precipitation varies from 230 mm at the lower elevations to more than 1100 mm in the higher regions. Plant communities at lower elevations are typical of the Great Basin Desert, while forest and alpine plant communities may be found in the higher sectors (Slaughter et al., 2001).

Digital elevation models (DEMs) for both watersheds were acquired from USGS-Seamless Data Warehouse (http://seamless.usgs.gov/) at resolutions of 1 arc-second (30 m) and 1/3 arc-second (10 m). Because the 3 arc-second DEMs were not available at the Seamless Data Warehouse, the Shuttle Radar Topography Mission (SRTM) 90 m Digital Elevation Database v4.1 (Jarvis et al., 2008) was used as the 90 m DEM source. The SRTM DEMs were downloaded from the GeoPortal of the Consultative Group on International Agricultural Research - Consortium for Spatial Information (CGIAR-CSI) (http://srtm.csi.cgiar.org). For both watersheds, landuse maps and related information were obtained from the USGS Land Cover Institute (LCI) (Fry et al., 2009) available at the USGS-Seamless Data Warehouse (http://seamless.usgs.gov/) with a resolution of 1 arc-second (30 m).

Detailed soils maps were requested from the Natural Resources Conservation Service (NRCS) - Soil Survey Geographic (SSURGO) Database (Soil Survey Staff NRCS-USDA, 2011) (Available at: http://soildatamart.nrcs.usda.gov). For the LREW, the soils map was built with portions of Turner and Worth counties; for the RCEW, the map is contained inside Owyhee county. Additionally, soils maps from the State Soil Geographic (STATSGO) Data Base were downloaded for both watersheds from the Water Resources National Spatial Datasets (NSDI) Node website (Soil Survey Staff NRCS-USDA, 2011) (Available at: http://water.usgs.gov/GIS/metadata/usgswrd/XML/ussoils.xml#stdorder). For the LREW, according to Sullivan et al. (2007), a majority of the soil classifies as Tifton loamy sand (36%) followed by Alapaha loamy sand (12%) and Kinston and Osier fine sandy loam (8%). In the case of the RCEW, soils derived from granitec and volcanic rocks and lake sediments are present on the watershed and range from shallow, desertic soils at lower elevations to deep,

commonly moist soils at the higher elevations, which are dominated by forests (Slaughter et al., 2001).

Measured data from monitoring stations for streamflow were obtained from USDA-ARS SEWRL. Long-term streamflow data are available on the SEWRL anonymous ftp site (ftp://www.tiftonars.org/) for up to eight flow measurement sites within LREW. However, for this study, only the flow control installed at Station B was used. Station B is the flow measurement installation located at the outlet of the LREW and it is instrumented to measure streamflow for the entire drainage area (Bosch and Sheridan, 2007).

3.1.2. Define periods of time for calibration and validation for Little River EW

The time period for the model evaluation was defined based on the availability of climatic and measured data for the different constituents. Streamflow and nutrients data for the LREW are available for the time period from October 10, 1978 to December 31, 1985. Complete records for streamflow, sediments and nutrients are available from January 1, 1986 to December 31, 2003. Climate data comprises 41 years from January 1, 1969 to December 31, 2009. Figures 1 and 2 show the annual precipitation for LREW from 1978 to 2003 and for Reynolds Creek EW from 1968 to 1996.

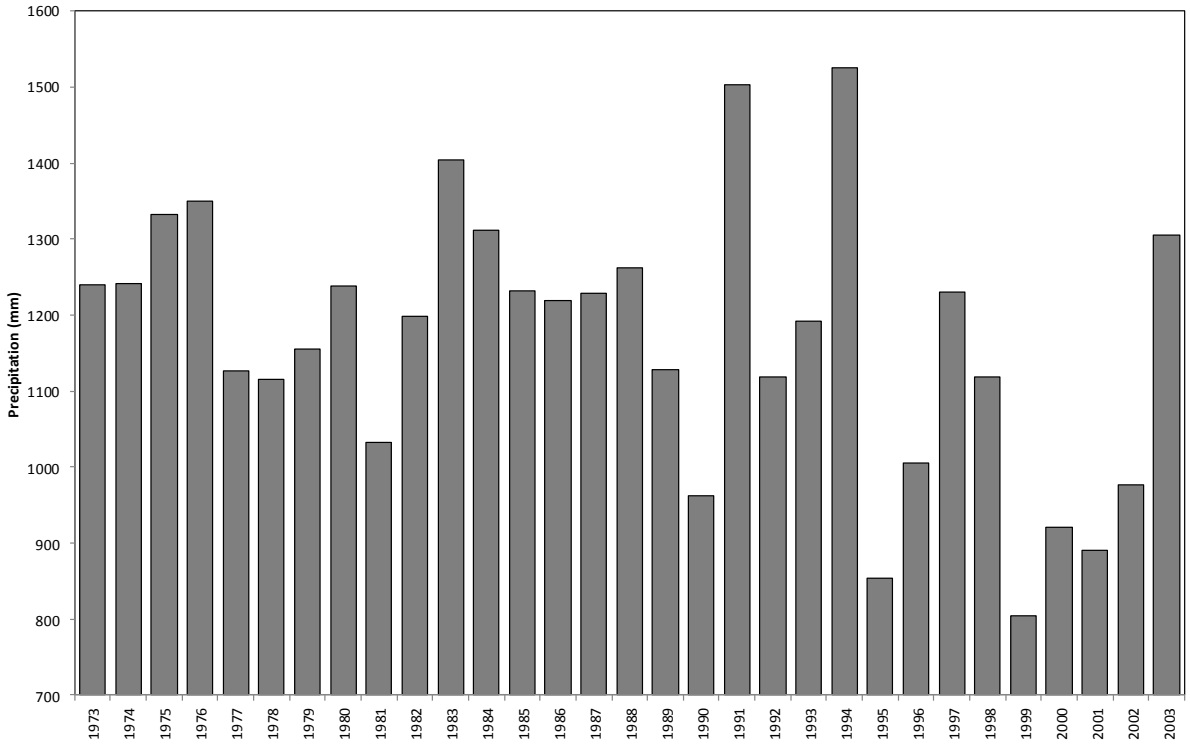


Figure 1. Annual precipitation for Little River Experimental Watershed from 1978 to 2003.

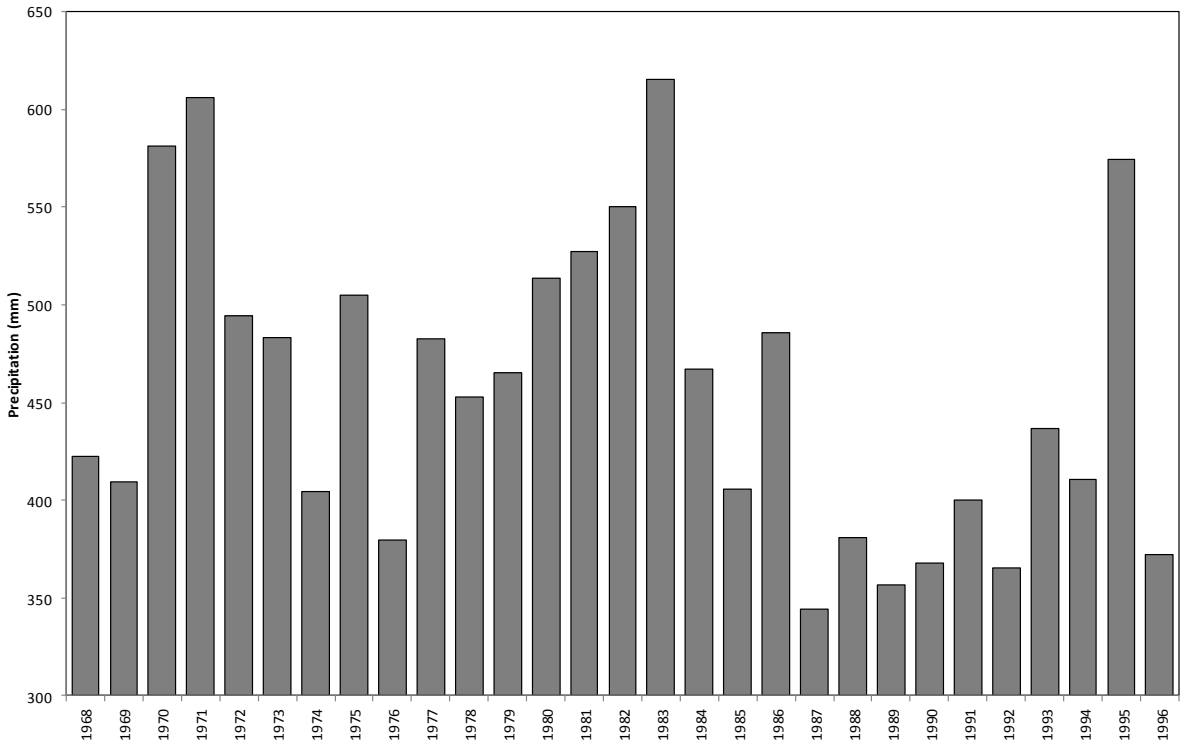


Figure 2. Annual precipitation for Reynolds Creek Experimental Watershed from 1968 to 1996.

In order to determine the time periods for calibration and validation, the average annual precipitation was calculated for each watershed, 1119.2 mm for LREW for the period 1978 – 2003 and 457.4 mm for RCEW from 1968 - 1996. Then, differences between the average annual amount and each year were calculated and plotted (figs. 3 and 4) to determine years that were drier or wetter than the average for the period.

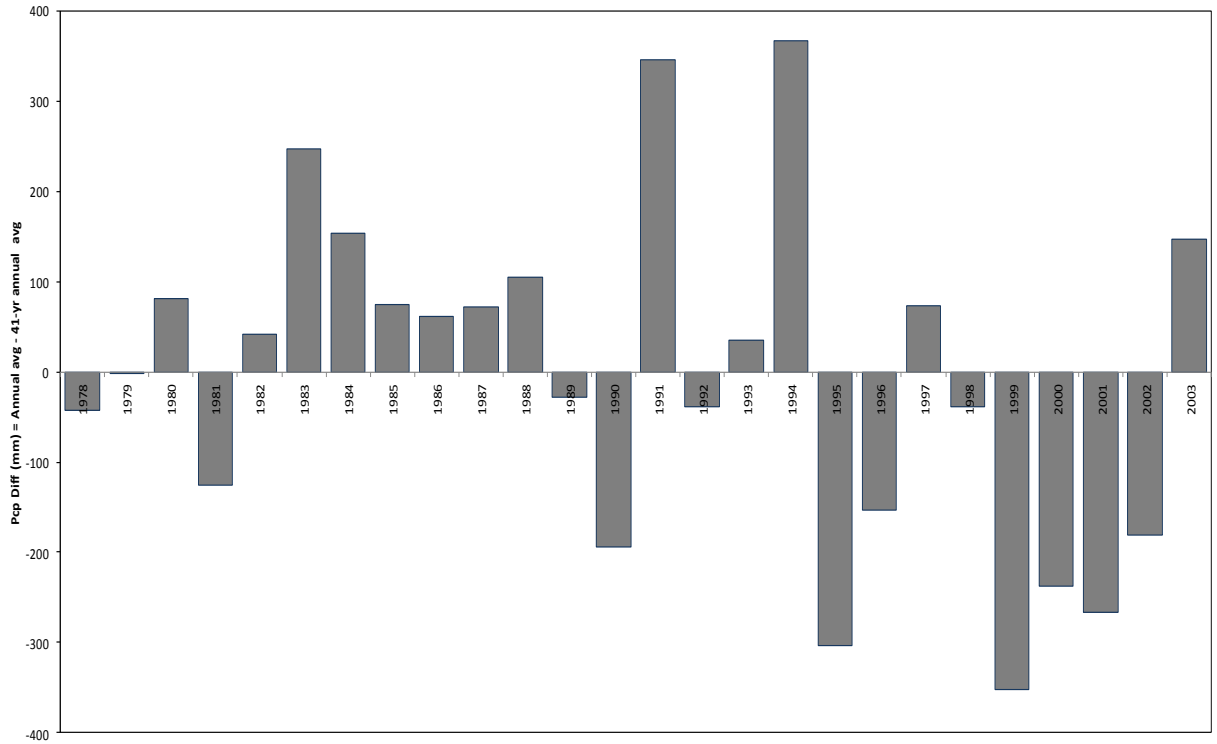


Figure 3. Difference between average annual precipitation and annual precipitation for each year from 1978 to 2003 and the selected calibration and validation periods for this study for the Little River Experimental Watershed.

Two periods of six years in length that each includes three drier years and three wetter years were selected for each watershed. The calibration period for LREW was selected as January 1, 1987 to December 31, 1992 and the validation period from January 1, 1993 to December 31, 1998 (fig. 3). Additionally, January 1, 1986 to December 31, 1986 was used to initialize the model. For RCEW (fig. 4), the calibration period goes from January 1, 1973 to December 31, 1978 and the validation period from January 1, 1979 to December 31, 1984. Additionally, 1978 was used to initialize the model. The initialization year of each simulation was discarded because the arbitrary initial conditions affect the simulation results during this time period.

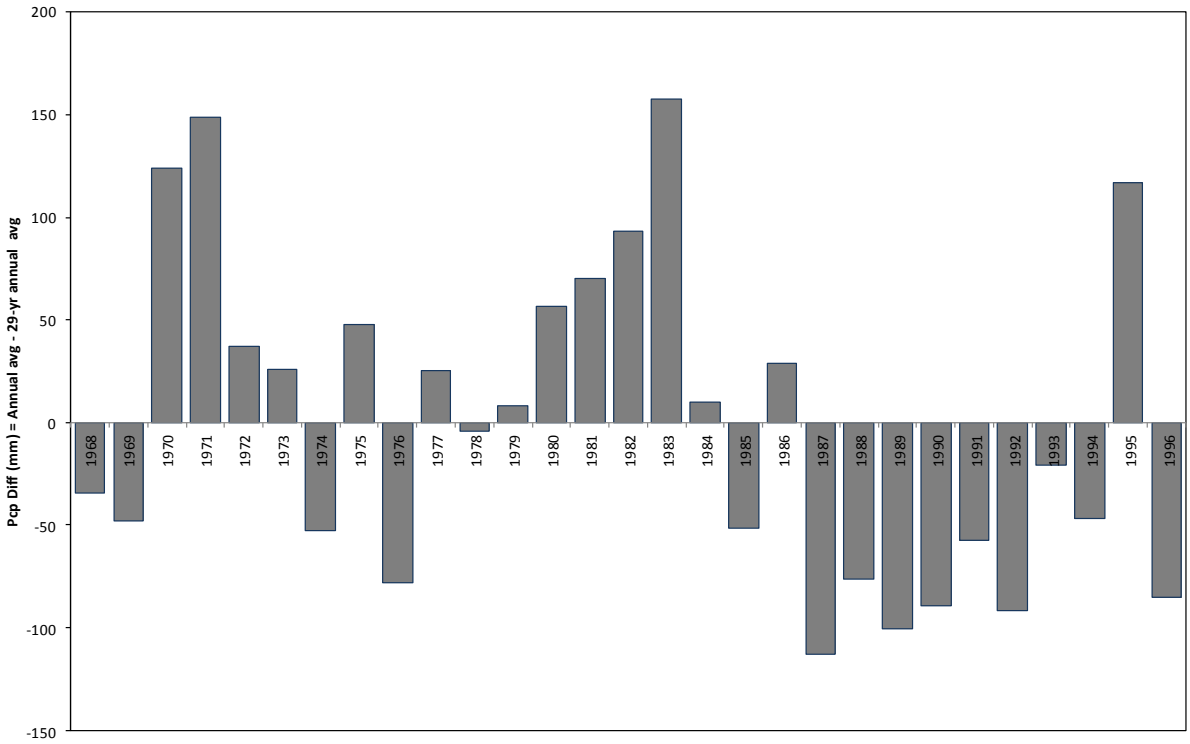


Figure 4. Difference between average annual precipitation and annual precipitation for each year from 1968 to 1996 and the selected calibration and validation periods for this study for the Reynolds Creek Experimental Watershed.

The following input datasets for SWAT were prepared for each watershed:

- Weather gauge locations (Wgnstations.dbf)
- Precipitation gauge location (pcpstations.dbf)
- Temperature gauge location (tmpstations.dbf)
- Precipitation data table (pcptiffton.dbf)
- Temperature data table (tmptiffton.dbf)

In order to avoid creating individual files for each evaluation period, precipitation and temperature files were prepared for the full length of the available dataset: from January 1, 1978 to December 31, 2003 for LREW and from January 1, 1968 to December 31, 1996 for RCEW. However, SWAT was set to be run from January 1, 1986 to December 31, 1999 for Little River and from January 1, 1973 to December 31, 1984 for Reynolds Creek. The evaluation periods corresponding to calibration and validation were set once the databases were exported to MS-Excel.

Table 2 includes a description of the SWAT state variables used in all analyses in this study. Because of the type of available measured information for evaluation, organic and mineral phosphorous were added together as total phosphorous. Similarly, organic, nitrate, ammonium, and mineral nitrogen were aggregated as total nitrogen.

Table 2. Description of SWAT output variables used in the analysis.

Variable	Code	Description
Flow-related	FLOW_OUT	Average daily streamflow at the outlet of watershed, includes channel losses (m^3 s^{-1})
Sediment-related	SED_OUT	Sediment transported with water at the outlet of watershed, includes channel losses (t)
Nutrient-related	ORGN_OUT	Organic nitrogen transported with water at the outlet of watershed (kg N)
	ORGP_OUT	Organic phosphorus transported with water at the watershed outlet (kg P)
	NO3_OUT	Nitrate transported with water at the watershed outlet (kg N)
	NH4_OUT	Ammonium transported with water at the watershed outlet (kg N)
	MINP_OUT	Mineral phosphorus transported with water at the watershed outlet (kg P)
	MINN_OUT	Mineral nitrogen transported with water at the watershed outlet (kg N)

3.1.3. Manual calibration procedure

A modification of the procedure for manual calibration in watershed models described by Santhi et al. (2001) was employed. The calibration sequence started with streamflow, followed by sediment, and, finally, total nitrogen and phosphorous. For each of the constituents, the flow chart (fig. 5) shows the model parameters and the order in which they were modified in a sequential manner. The calibration was done with one parameter at a time within each constituent. The first parameter was adjusted, the model was run and the outputs evaluated. If the criteria were not met, the second parameter was adjusted and the outputs evaluated again. If the criteria were not met after modifying the last parameter, a new cycle was started modifying the first parameter in the list. When the modification in a parameter did not improve the predictions, the parameter was reset to its last value and the next parameter in the list was modified. The calibration procedure stopped as soon as the criteria were met.

Once the model was successfully calibrated for the first output, the resulting set of parameters was used as the starting point for the next output. This process was performed until the model was calibrated for all outputs of interest. The calibration was performed first for

annual averages; then the set of parameters was tested with monthly values. Depending on the results of the evaluation, the model parameters were fine-tuned for monthly averages.

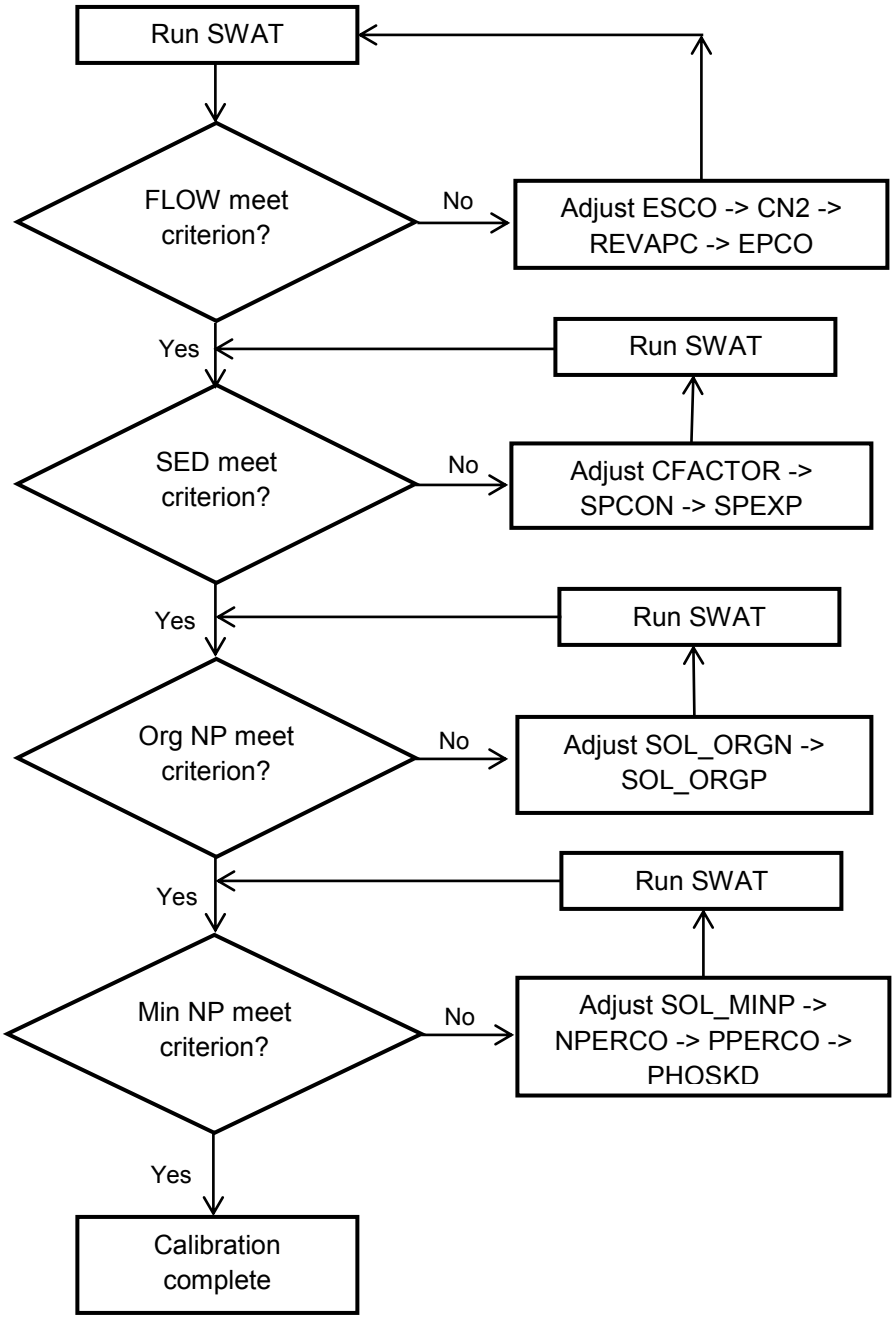


Figure 5. General calibration procedure for flow, sediment, and nutrients (based on Santhi et al., 2001).

For each change in the parameter value, the model was run, the output tabulated and the R-code for statistics was run. An R-code was programmed to calculate descriptive statistics of observed and simulated values and to calculate the coefficient of determination (R^2). Additionally, the following goodness-of-fit functions for numerical comparison of simulated and observed time series were included in the R-code from the HydroGOF package (Zambrano, 2010): Nash-Sutcliffe Efficiency (NSE), Percent Bias (PBIAS) and Root Mean Square Error observation standard deviation ratio (RSR).

3.1.4. Calibration criteria

The calibration was performed to maximize the coefficient of determination (R^2) and the Nash-Sutcliffe Efficiency (NSE) value and to minimize Percent Bias (PBIAS) and RMSE – observation standard deviation ratio (RSR) according to the criteria presented by Singh et al. (2004) (table 1). Two different ranges of evaluation criteria were established for the evaluation of annual and monthly output values, with the range for annual values being more rigorous. The intervals of the “Satisfactory” category were used for the monthly time step evaluation and the intervals for the “Good” category were employed for the annual time step. Based on Santhi et al. (2001), the criteria for the coefficient of determination (R^2) was set to 0.65 and 0.60 for annual and monthly output values, respectively.

The range of values for each of the parameters was initially fixed based on reported limits in the SWAT user’s manual (table 3) (Neitsch et al., 2005); these limits represent the minimum and maximum values that the model sets by default for each parameter. However, soil evaporation compensation factor (ESCO) and plant uptake compensation factor (EPCO) were additionally constrained following SWAT recommendations, which are recommended values for a particular parameter based on previous experiences (Neitsch et al., 2005). Table 3 additionally shows the initial values which were used to run the baseline. These values were set based on the reported default values in the SWAT user’s manual.

Table 3. SWAT calibration parameters range and initial values.

Variable name	Processes	Description	Model range	Initial value
CN2	Flow	Curve number	0 to 100	Varies as f(landuse)
GW_REVAP	Flow	Ground water revap coefficient	0.00 to 1.00	0.02
ESCO	Flow	Soil evaporation compensation factor	0.00 to 1.00	0.95
EPCO	Flow	Plant uptake compensation factor	0.00 to 1.00	1
SOL_AWC	Flow	Soil available water	0.00 to 1.01	Varies as f(soil)
C FACTOR	Sediment	Cover or management factor	0.003 to 0.45	Varies as f(landuse)
SPCON	Sediment	Linear factor for channel sediment routing	0.0001 to 0.01	0.0008
SPEXP	Sediment	Exponential factor for channel sediment routing	1.0 to 1.5	1
SOL_ORGN	Organic N	Initial organic N concentration in the upper soil layer for a particular land use		Varies as f(landuse)
SOL_ORGP	Organic P	Initial organic P concentration in the upper soil layer for a particular land use		Varies as f(landuse)
NPERCO	Mineral N	N percolation coefficient	0.2 to 0.6	0.05
SOL_MINP	Mineral P	Initial mineral P concentration in the upper soil layer for a particular land use		Varies as f(landuse)
PPERCO	Mineral P	P percolation coefficient	10.0 to 17.5	10
PHOSKD	Mineral P	P soil partition coefficient	100 to 175	100

3.1.5. Sensitivity analysis

Sensitivity is expressed by a dimensionless index, which was calculated as the ratio between the relative change of model output and the relative change of a parameter. Sensitivity was approximated using the relative sensitivity (S_r), with a conventional variation by a fixed

percentage of the initial parameter value to determine parameters that most influenced predicted streamflow, sediment yield, and nutrient model outputs.

Sensitivity analysis was performed with the purpose of determining which parameters have an impact on the model predictions (Shirmohammadi et al., 2006) and, therefore, will be important to measure in the field or find alternative ways to estimate their values. Parameters for the sensitivity analysis were selected after studying all the possible input parameters required to create a SWAT project. All the soil parameters were included for this analysis because it is well known that soils information is likely to be missing in data-poor conditions and also that the most sensitive parameters for hydrology and water quality are the physical soil properties such as bulk density and available water capacity (Lenhart et al., 2002). Values were changed in an efficient way to cover the whole range in a few steps, keeping their actual value as a base line. The values for parameters were automatically varied one at a time while holding the others fixed; the following variations were used: ±10%, 25% and 50%.

Natural basins show great variability and complexity with respect to soil characteristics and distribution. Physical soil properties vary in two ways: horizontally, spatial distribution of soil types across the watershed area (soil type), and vertically, differences among soil horizons within each soil type. If property values are changed in only one layer and one soil, values of other layers of the same soil or of other soils may be violated. To avoid this, the parameter was changed simultaneously for all layers of all soils at the same time and in the same proportion. Table 4 shows the seven soil parameters and the maximum and minimum values.

Table 4. Soil parameters with their minimum and maximum values included in the sensitivity analysis.

Parameter Name	SWAT Id	Min	Max
Bulk density (g cm-3)	SOL_BD	0.00	1.73
Available water capacity (mm H2O mm soil-1)	SOL_AWC	0.00	0.70
Saturated hydraulic conductivity (mm hr-1)	SOL_K	0.00	600.00
Organic carbon content (%)	SOL_CBN	0.00	2.03
Soil albedo (Moist)	SOL_ALB	0.00	0.44
USLE soil erodibility (0.013 t m2h)/(m3 t cm)	USLE_K	0.00	0.37
Salinity (%)	SOL_EC	0.00	100.00
Porosity fraction from which anions are excluded	ANION_EXCL	0.00	0.50

The calibrated parameter set for streamflow, along with the same available data set of Little River EW used for the model evaluation, was used to perform automated sensitivity analysis. The model was run for the period from January 1, 1986 to December 31, 1992. Also, the first year of simulations was used to initialize the model, so model outputs from January 1, 1986 to December 31, 1986 were discarded for the analysis to follow.

3.1.6. Uncertainty analysis procedure

Simple Monte Carlo simulation (SMCS) methodology was used to quantify the uncertainty in model outputs induced by uncertainty in inputs that are likely to be missing or incomplete in data-poor conditions. SWAT-MonteCarlo (SWATmc) was developed for this study as a macro using the embedded Visual Basic for Applications (VBA) in Microsoft Excel. The core routines for Monte Carlo simulation were programmed in VBA and a spreadsheet was used as a data input interface and to store data generated by the code for each model run.

SWATmc is a semi-automated modeling procedure that performs one-phase Monte Carlo simulations for uncertainty analysis. SWATmc calculates results over and over by substituting a set of sampled random parameters from probability functions. The result is a probability distribution of output values and a confidence interval. The procedure is general, iterative, and has no inherent limitations in terms of the number of parameters that can be

considered simultaneously. The tool lumps all sources of uncertainties. It is simple and very convenient to use, does not require previous calibration of the model, and can be applied for any modeling work that requires uncertainty analysis with SWAT.

SWATmc works in the following way. Step 1, the Arc SWAT interface is used for the setup and parameterization of the model and to transcribe all the model input and parameter databases in text format. Step 2, the user defines values to characterize the distributions of input parameter values. Step 3, the code creates a set of random values for each parameter with one of the following predetermined probability distributions: uniform, normal, triangular, exponential or lognormal. Step 4, SW_Edit.exe replaces the parameter value in SWAT input files. Step 5, SWAT2005.exe is called to run in an MS-DOS mode. Step 6, the code calls the SW_Extract.exe program to extract the required outputs from SWAT's reach output file and store them in a txt-file. Steps 4 to 6 are repeated the number of iterations specified by the user. In this case, it was fixed at 1000 times assuming that only one parameter was uncertain.

Due to the computational characteristics of SWATmc, the tool calculates a constant adjustment factor to deal with changes in multiple dimensions. Therefore, variations in parameter values for the Monte Carlo analysis are made based on a constant adjustment factor, in such manner that the soil properties vary vertically and horizontally keeping the varying proportion constant. Thus, SWATmc was set according to the following scheme (fig. 6):

1. For each uncertain parameter, the following inputs are required to run SWATmc: number of uncertain parameters, number of iterations, uncertain parameter's code name and the parameters required to characterize its probability distribution (e.g., mean and variance for log-normal).
2. The mean and variance for each parameter were calculated based on the soil map's attribute table. The soil parameter's values were grouped in one pool of data; this pool lumps data from all soil horizons. The assumption is that the attribute table contains representative - deterministic data of the real soil system in the study area.
3. A random value between 0 and 1 is generated using the Marse - Roberts code (Marse and Roberts, 1983). According to the authors, the algorithm produces uniformly distributed numbers equally likely to assume any value between 0 and 1. Because it has a long cycle length, the procedure guarantees that more than 1013 numbers will be generated before the repetition begins.

4. A randomized parameter is calculated with Excel's log-normal function (LOGNORM_INV), which is a function of random number (generated with the Marse - Roberts code), mean and variance calculated to characterize the distribution.
5. The relative difference between random value and the mean is computed, i.e., (random value – mean)/mean.
6. An adjustment factor is then calculated as (1 + the relative difference).
7. The code includes a max/min control to avoid values that go beyond the values that the model sets for each parameter reported in the SWAT user's manual (Neitsch et al., 2005).
8. The adjustment factor is then stored in a text file and the loop goes back to step 3 as many times as the specified number of iterations.
9. The initial parameter values for each soil and each horizon are multiplied by the calculated adjustment factor.
10. SW_Edit.exe replaces the adjusted parameter value in SWAT input files.
11. SWAT2005 is run and the required outputs for the selected constituents are extracted by SW_extract.exe for further analysis.
12. Steps 9 to 11 are repeated as many times as the specified number of iterations.

When the last SWAT run is done, the tool reads all the output files and calculates the following statistics: mean, standard error, median, standard deviation, variance, confidence interval, skewness, and kurtosis. The tool then imports all the generated values to an Excel sheet for table creation and plotting modified Tukey box plots in which uncertainty is represented by the 90^{th} and 10^{th} percentiles.

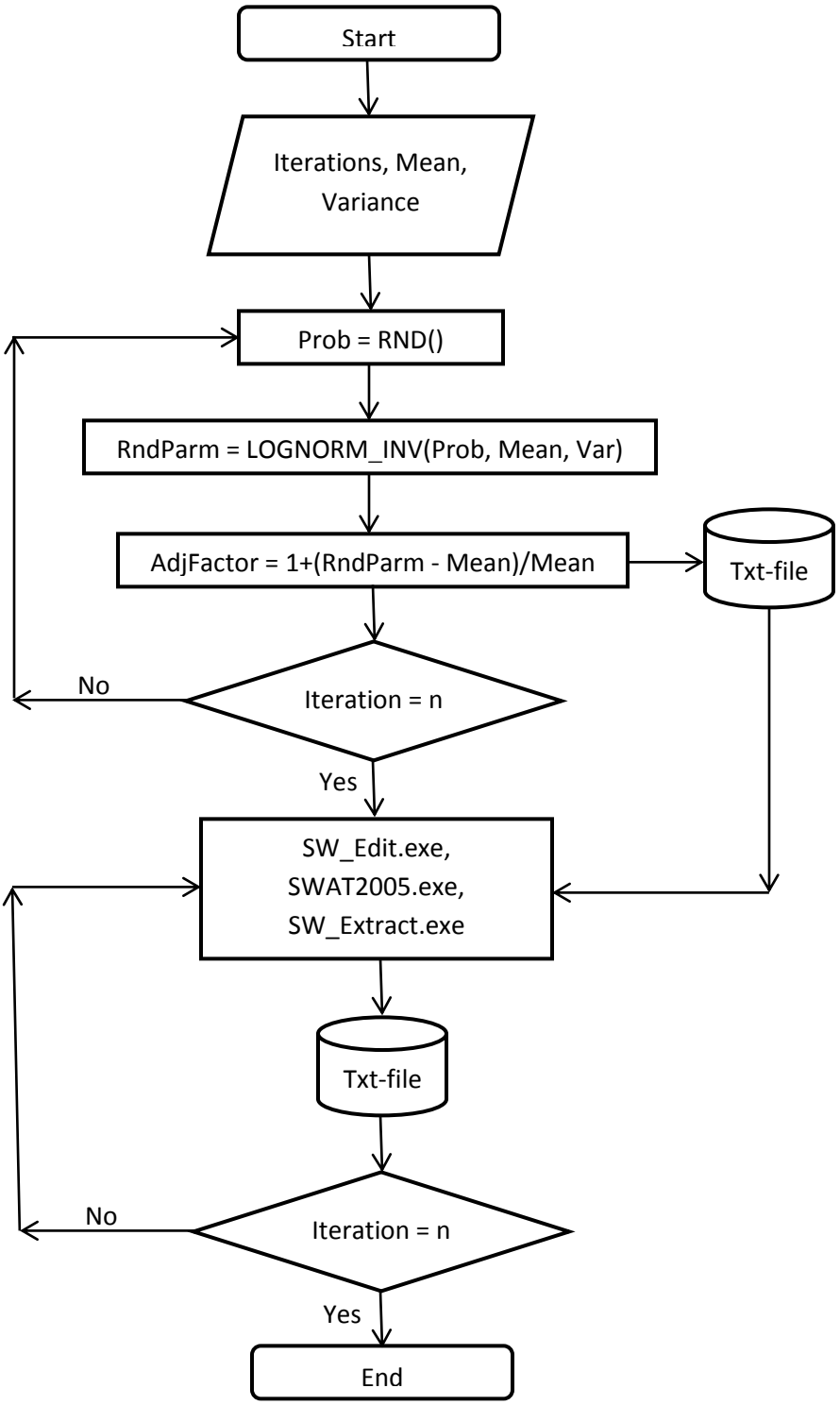


Figure 6. Soil parameter randomization and uncertainty procedure.

The uncertainty analysis was performed for both experimental watersheds, Little River in Georgia and Reynolds Creek in Idaho. The selection of the parameters for uncertainty analysis was made based on the results of the sensitivity analysis. Only the soil parameters categorized as at least slightly sensitive were included (SOL_AWC, SOL_BD, USLE_K, SOL_CBN, SOL_ALB). The output distributions are presented as modified Tukey box plots providing a range and distribution of uncertainty of SWAT outputs due to selected soil parameters and also were used to establish a 10^{th} - 90^{th} percentile range (CI) for the output. Additionally, with the aim of defining whether or not coarse resolutions have a combined effect on soil parameters, two sets of input data were also sequentially changed: DEM resolution (10, 30 and 90 m) and the source of the soils database (SSURGO and STATSGO).

3.2. Objective 2 - Develop methods to gather qualitative and quantitative data that will allow modeling agricultural watershed systems under data-poor environments

3.2.1. Data creation

Elevation models (DEMs), soil datasets and climate datasets are the most important input data for SWAT. In data-poor environments, these same inputs are typically the ones that are missing or do not have enough resolution for modeling purposes. The following methodology describes an attempt to improve the resolution of freely available DEMs and to generate appropriate soil data sets for running SWAT. The following sections describe the procedure for digital soil mapping applied to a data-rich environment; once the method was tested, it was applied to a data-poor environment.

3.2.1.1. Resampling of DEM

The CGIAR-CSI distributes processed SRTM digital elevation data as 3-arcsecond DEMs (approximately 90 m resolution). The CGIAR-CSI elevation models are derived from the original USGS DEMs, which were averaged to 90 m from the original 1 arc second data (approximately 30 m). The 90 m-USGS-DEMs have been processed at CGIAR-CSI to fill in no-data voids. The data processing involved the production of vector contours and points and the re-interpolation of the derived contours back into a raster DEM. The interpolated DEM values were then used to fill in the original no-data holes within the SRTM data (Jarvis et al., 2008).

For this part of the study, the 90 m resolution SRTM DEM obtained from CGIAR for the Little River watershed was resampled to a 30 m grid (UTM – Zone 17 North) by direct projection using cubic convolution as a resampling technique. Cubic convolution is a technique for

resampling raster data in which the weighted average of the nearest 16 cells is used to calculate the new cell value (Keys, 1981). The decision to use a resampling technique instead of more complex interpolation methods was made based on the work performed by Her and Heatwole (2008). The authors proved that direct projection of the SRTM source decimal degree data to a 30 m grid using cubic convolution is comparable to other interpolation methods and is much simpler to implement than more complicated interpolation procedures.

3.2.1.2. Generation of soils map and attribute table

Soils data are the most critical input for the model and constitute the most challenging dataset to gather. For spatially distributed models, soils data sets have two dimensions. The first dimension is a representation of the soil spatial distribution, a soils map, which is a combination of polygons representing mapping units and soil classes. The second dimension is an attribute table of soil profile observations, which is the representation of soil properties. Frequently, information about both dimensions is missing or incomplete.

Soils map and attribute table generation depends on the available information for the study area. To overcome incomplete or lack of soils data under data-poor conditions, a framework was developed to generate missing or incomplete information for modeling. GIS and remote sensing analysis have been applied successfully to a wide variety of situations lacking sufficient information to create base soil survey maps using Landsat satellite imagery and DEM data in a GIS framework. Soils data map generation takes two steps: soil unit determination (soils map definition) and soil properties definition (attribute table definition).

a) Map generation

The main idea is to generate a soils map by overlaying a slope class map and color composite images. The technique explained here is based on freely available information, such as LANDSAT imagery and SRTM digital elevation data.

Map units or soil classes were defined based on a GIS and remote sensing procedure that uses the following general soil prediction model (eq. 1) proposed by McBratney et al. (2003):

$$S_{(x,y)} = f(Q) + \varepsilon \quad (1)$$

Where: Q is predictor variable(s), and

S(x,y) is the soil class or attribute at some spatial location x, y.

The soil profile character is a function of climate, organisms, relief, parent material and time (Jenny, 1941; Scull et al., 2003). All these factors are represented in the equation by Q, as empirical quantitative descriptors of relationships between soil and other spatially referenced factors. All these soil forming factors could be used as soil prediction functions.

The mapping procedure uses terrain (slope) and parent material characteristics (lithology) to delineate the mapping units. Only two factors, topography (r) and parent material (p), were considered because several authors (Acron, 1965; Moore et al., 1993; Batchily et al., 2003) consider them as the two primary soil forming factors and they have been used successfully to predict different soil classes (eq. 2). The model can then be represented as:

$$S_{(x,y)} = f(r, p) + \varepsilon \quad (2)$$

Topographic data derived from DEM and band combination and ratios of Landsat remotely sensed spectral data were selected to represent soil-forming factors. Basically, the methodology implies grouping and classification of slope to make a broad subdivision of the map area. Further subdivision of the soil units was made according to lithology (parent material) to complete the delineation procedure.

Slope phases were used to divide soil series. The slope class map (r-factor) was generated from the DEM with the ArcGIS Spatial Analyst tool. The first step was to classify slope derived from a DEM following the Soil Survey Division Staff (1993) slope classification. This classification is based on the guide and terminology for various slope classes defined in terms of gradient and complexity. Six classes were formed with slope intervals; table 5 shows the definitions of slope classes by slope gradient limits.

Table 5. Slope gradient classes (Soil Survey Division Staff, 1993).

Class	Description	%
1	Nearly level	0 – 1.0
2	Gently sloping	1.0 - 5.0
3	Strongly sloping	5.0 – 15.0
4	Moderately steep	15.0 – 30.0
5	Steep	30.0 – 60.0
6	Very steep	> 60.0

The p-factor (parent material) information was obtained from processing a satellite image. Lithological mapping was performed manually in ERDAS Image 9.1 software, using as base the Landsat 4-5 Thematic Mapper (TM) imagery. TM data collected in September 2011 were used in this part of the study. Little River EW is included in Landsat TM data scene number LT50180382011255GNC01 (path/row = 18/38), consisting of seven bands. The image was downloaded with the online USGS Global Visualization Viewer from Earth Resources Observation and Science Center (EROS), available at http://glovis.usgs.gov/. It has 0 % cloud cover. The data were obtained as final product processed tapes (Quality 9) and were corrected to the corresponding UTM coordinate system. No other corrections were made.

Unsupervised image processing was applied for lithological discrimination of Little River EW. The result of the digital image processing of Landsat satellite data is a false color composite image. This image was constructed by overlaying combinations of the LANDSAT TM band ratios (5/7, 5/4 and 5/4*4/1). Resulting files were reclassified in 12 classes and the image was converted to a raster format to facilitate further processing. Finally, soils map units were defined by overlaying slope class map with the lithologic raster. The combination of lithology and slope class was then reclassified into discrete units using the ArcGIS's spatial analyst tool.

b) Attribute table generation

A large array of agriculturally-important soil properties (including texture, organic and inorganic carbon content, macro- and micro-nutrients, moisture content, cation exchange capacity, electrical conductivity, pH, and iron) can be quantified with remote sensing to varying extents (Ge et al., 2011). However, a major limitation of conventional satellite remote sensing is that it is typically only capable of collecting information relevant to the soil surface. For modeling purposes, the soil attribute table must include information by soil layers or horizons.

The methodology to generate the attribute table is based on associating textural information of the top layer with a complex data base using a simple lookup procedure. The SWAT model has an extensive soils database for the continental U.S., which follows the STATSGO format. Each entry in the soils database is identified by MUID (Map Unit Identifier) and an S5ID (Soil Interpretations Record Number). This database has all the necessary soil physical and chemical parameters by soil horizon to run the SWAT model.

The soils database in SWAT comprises an individual table for each state, therefore, in order to synthesize and group soils by textural class, the information for all 50 states was joined all together in one table and the repeated S5ID were deleted, creating a soils database with

18,240 unique S5ID records grouped in 12 subsets based on USDA's 12 textural classes. Each class ranges in a size distribution that is consistent with a rather narrow range in soil behavior; each class is named to identify the size separate or separates having the dominant impact on properties.

The textural proportion expressed as a percentage of sand (S), silt (Y) and clay (C) was used to define soil textural class. The matching process is performed by a set of logical rules or a collection of "if-then" statements. The lookup process first identifies the textural class of the observed soil, sets its proportions in one of the 12 textural classes in the database, and then tries to match the observed proportions with the ones in the soils database. If the observed proportions for the three soil classes are found in the subset of the database, then the S5ID code for that soil class is recorded and the lookup process starts over for the next soil unit. If more than one database record matches the observed proportions, a process of successively partitioning data into homogeneous subsets is started, followed by a random selection from the pool of matching database entries using a uniform probability function where the minimum and maximum values are set to the interval of the matching entries.

When the observed textural distribution does not match perfectly the textural distribution of any of the database records, the lookup process starts searching again with the type of soil particle (sand, silt or clay) that makes up the highest percentage of the sample used to describe the soil texture class. Once this is matched, it tries to match the second particle size proportion and then approximates the third particle size. When none of the three fractions is dominant, the textural class is loam and the searching sequence is sand, clay and silt. If only one soil entry in the database is found with this process, its S5ID code is recorded; a random selection is made if there is more than one matching record. When none of the proportions match the database, to avoid exact percentages, the silt, sand and clay proportion values are modified in ±10% in a successive fashion, starting with the dominant particle size. This process is repeated with 10% increments until at least the dominant particle size proportion matches the database. When the process has been repeated for all soil units in the map, a text file matching the different soil map units with an S5ID code from the database is created. This file, with Value and S5ID headers, is required by the SWAT interface to further associate each soil unit in the map with the model's soils database using the S5ID as a linkage code.

For Little River, the following procedure was used to mimic the field soil sampling that takes place in a real data-poor environment. A grid of points with 1 km interval was created in ArcGIS using the Fishnet function. This point layer was overlaid with the SSURGO map to pass

over the textural proportions of the top soil from the map attribute table to the grid of points. The resulting grid of textural proportions was then overlaid with the soils map generated from satellite imagery and slope classification (SATSLP). As a result, each pixel in SATSLP map gets a texture proportion; therefore, its attribute table contains the number of pixels of similar texture for each soil map unit. The calculated mode for each soil unit was selected as the final texture proportion to be matched with the SWAT soils database.

3.2.2. Validation of the map generation procedure

Since the analysis for the data-rich environment was complete, input likely to be absent in data-poor conditions were sequentially replaced by alternative datasets. Two sequential changes, DEM source and soils dataset sources, were made to compare with the results obtained in the data rich analysis (objective 1). The original 30 m USGS DEM was compared with the 30 m resampled CGIAR-CSI DEM using both soils sources: SSURGO and STATSGO. In addition, the SSURGO and STATSGO databases used to model LREW were replaced with the generated soil map and database (SATSLP). Therefore, the comparison includes six scenarios as a result of the combination of two sources of 30 m DEMs (USGS and resampled CGIAR-CSI) and three sources of soil databases (SSURGO, STATSGO, and SATSLP). The results of this comparison analysis quantify the difference between the downloaded data and the equivalent generated data.

The same set of parameters for uncertainty analysis were included here (SOL_AWC, SOL_BD, USLE_K, SOL_CBN, SOL_ALB). Then, the model was run one thousand times for each of the individual parameters and all of them at the same time. SWAT was set for a 13-year period, January 1, 1986 to December 31, 1998 (1986 was used to initialize the model). The required outputs for the selected constituents were extracted for further analysis and the following statistics were calculated: mean, standard error, median, standard deviation, variance, confidence interval, skewness, and kurtosis. Also, output distributions were presented as modified Tukey box plots providing a range and distribution of uncertainty as the 10^{th} and 90^{th} percentiles.

3.3. Objective 3 - Determine to what extent the proposed methodology to use SWAT with limited data will be able to represent water quality impacts of agricultural watershed systems in data-poor environments.

3.1.1. Watershed descriptions and SWAT input data

The Huanquisco River watershed (fig. 7) is a 124 km^2 mountainous river basin, which is located on the eastern side of Lake Titicaca in Bolivia. The area has 793 inhabitants (242 families) from six communities (Chikchaya, Cohani, Canta, Calahuncane, Quesuni, and Chojñapata). The basic characteristics of the watershed are introduced as follows.

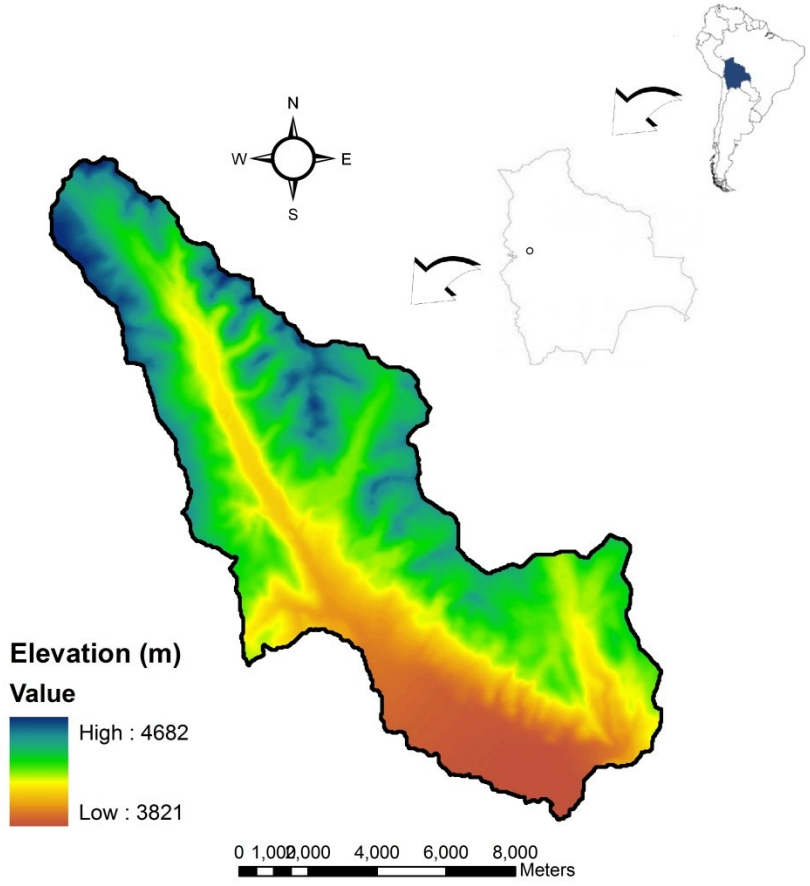
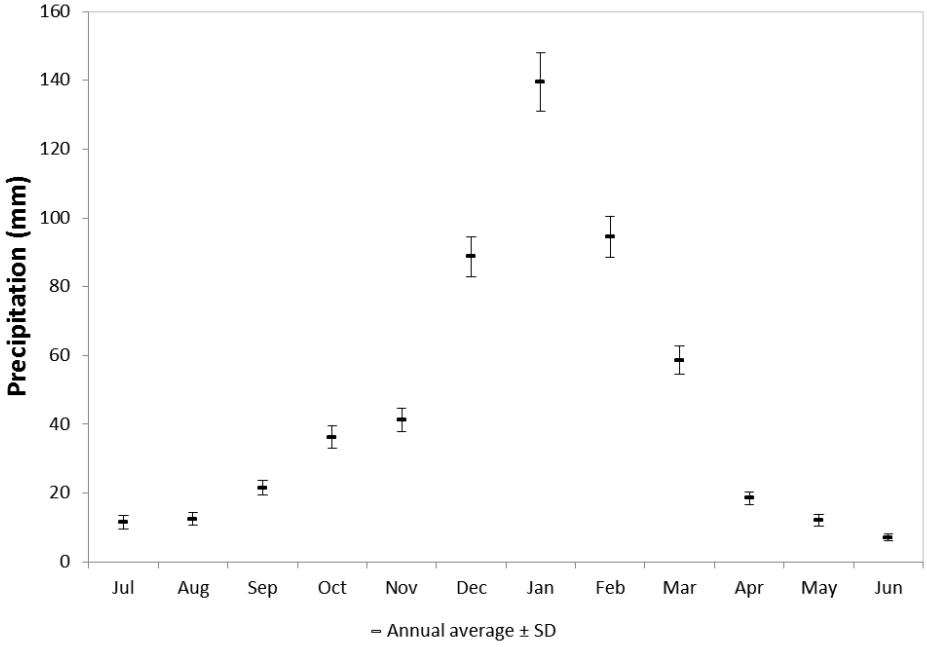


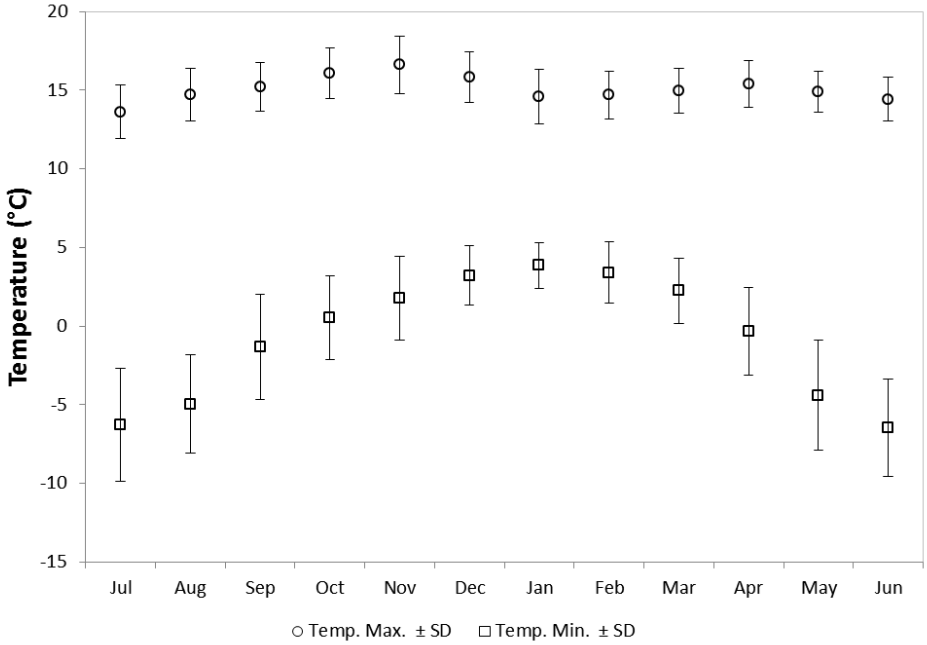
Figure 7. Elevation map of the Huanquisco River Watershed, Ancoraimes - Bolivia.

The average elevation is about 4222 m. The minimum and maximum elevations determined from the DEM data range between 3820 and 4680 m. Weather in this region was characterized based on data collected from the Meteorology and Hydrology National Service (SENAMHI) for the time period of 2001 to 2011 (accessed at: http://www.senamhi.gob.bo). Average annual rainfall for the study period was about 480 mm with 80% falling between

October and March (fig. 8a). The average annual temperature for the Huanquisco watershed is 14.0 °C. In winter, the average minimum temperature is -6.4 °C, while in summer the average maximum temperature is 16.6 °C (fig. 8b).



(a)



(b)

Figure 8. Average monthly precipitation (mm) (a) and average maximum and minimum temperature (°C) (b) for the Huanquisco river watershed in Bolivia (Meteorology and Hydrology National Service, 2012).

The watershed is part of the Lake Titicaca basin, which is characterized by intensive agriculture in relatively small fields in upland areas and riparian forests along stream channels. Major land uses (fig. 9) within the watershed are pasture (55%), agricultural (16%), and bare or rocky land (28%). Other land use/land cover (Andean wetland and forest) accounts for the remaining 1% of the area. Crop production is dominated by potatoes (42%), onions (20%), and barley (20%); the remaining cropping area is green and fava beans, wheat and quinoa. Animal production is in the form of South American camelids and sheep flocks located in the northeast side of the watershed.

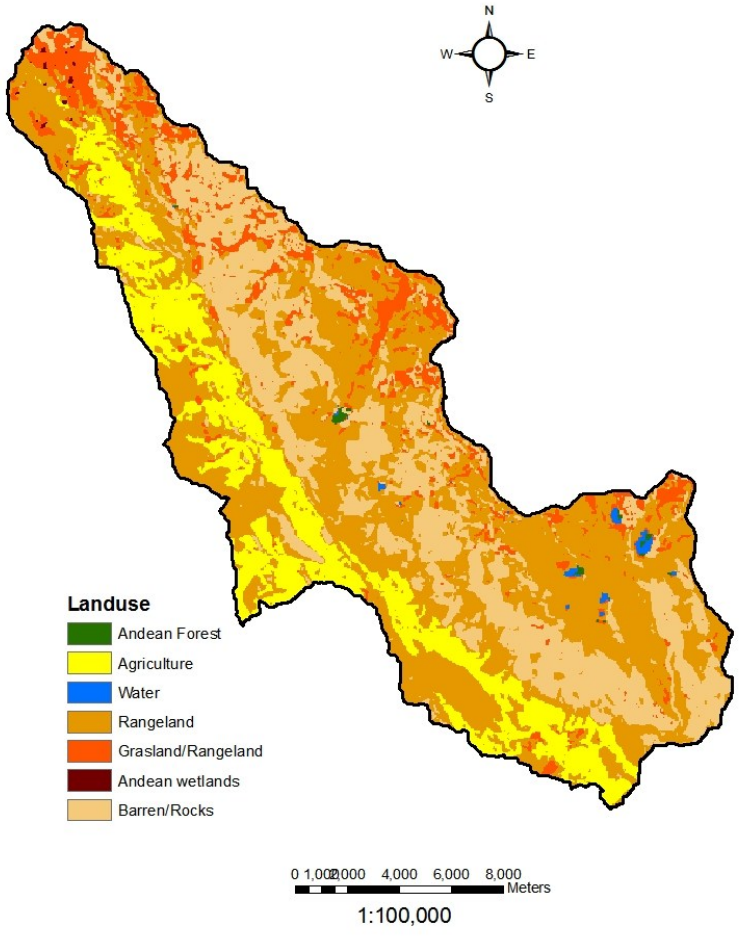


Figure 9. Landuse map for the Huanquisco river watershed (Penaranda et al., 2011).

3.1.2. SWAT input data setup

Land use, soil, and topography are the required spatially distributed input data for SWAT to create Hydrologic Response Units (HRUs), the smallest units used to simulate all of the hydrologic processes. Topographic information is provided in the form of digital elevation model (DEM). For the Huanquisco watershed, the original USGS 90 m resolution SRTM DEM acquired from USGS-Seamless Data Warehouse (http://seamless.usgs.gov/) was resampled to a 30 m grid (UTM – Zone 17 South) by direct projection using cubic convolution as a resampling technique. The landuse map was obtained from University of San Andres (Penaranda et al., 2011). The soil map units and associated attribute data were generated following the procedure presented in section 3.2.1.2.

Two other key sets of inputs required for simulating with SWAT are climate and management data. The daily climate inputs consist of precipitation, maximum and minimum temperature, solar radiation, wind speed, and relative humidity. The daily information collected from the Meteorology and Hydrology National Service (SENAMHI) for the time period of 2001 to 2011 was synthesized to monthly climate statistics (average and standard deviation) as required climate input data for the built-in weather generator (table 6).

Table 6. Climatic parameters for the Huanquisco River watershed. Monthly average and standard deviation for the 2001 to 2011 time period.

Month	Temp. Max. (± SD) (°C)	Temp. Min. (± SD) (°C)	Wind Velocity (± SD) (m/s)	Solar Radiation (± SD) (MJ/m^2)	Re. Hum. (± SD) (%)	Precipitation (± SD) (mm)
July	13.6 (1.7)	-6.3 (3.5)	1.2 (0.9)	19.3 (2.4)	0.1 (0.09)	11.5 (2.0)
August	14.7 (1.6)	-5.0 (3.1)	1.4 (0.7)	21.8 (2.5)	0.6 (0.09)	12.5 (1.8)
September	15.2 (1.5)	-1.3 (3.3)	1.6 (0.7)	22.8 (2.9)	0.6 (0.08)	21.5 (2.1)
October	16.1 (1.5)	0.6 (2.6)	1.4 (0.6)	24.3 (2.8)	0.6 (0.07)	36.3 (3.1)
November	16.6 (1.8)	1.8 (2.6)	1.4 (0.6)	24.8 (3.2)	0.6 (0.07)	41.3 (3.3)
December	15.8 (1.6)	3.2 (1.8)	1.3 (0.5)	23.2 (2.6)	0.7 (0.06)	88.7 (5.8)
January	14.6 (1.7)	3.9 (1.4)	1.5 (0.7)	21.3 (2.5)	0.7 (0.05)	139.5 (8.6)
February	14.7 (1.5)	3.4 (1.9)	1.6 (0.7)	21.3 (2.5)	0.7 (0.05)	94.5 (5.9)
March	15.0 (1.4)	2.2 (2.0)	1.5 (0.8)	20.9 (2.3)	0.7 (0.04)	58.7 (4.1)
April	15.4 (1.4)	-0.3 (2.7)	1.4 (0.7)	20.5 (2.2)	0.7 (0.06)	18.6 (1.8)
May	14.9 (1.3)	-4.4 (3.5)	1.2 (0.6)	19.8 (2.0)	0.6 (0.11)	12.2 (1.6)
June	14.4 (1.3)	-6.4 (3.0)	1.1 (0.5)	19.1 (1.8)	0.5 (0.09)	7.2 (1.0)

With the assistance of two local technicians, a survey questionnaire was used to interview 15 farmers of the Huanquisco River watershed. The local technicians spoke native language of Aymara and were well known in the area. This increased farmer's confidence in sharing their knowledge and completing the questionnaire. The questionnaire had questions to collect information in the following categories: crop production system (including fertilization and pesticide management); soils classification and management; grassland and livestock management; and knowledge about erosion and runoff processes. Additionally, the questions

were intended to obtain information regarding how the farmers deal with adverse weather events and how their decision making process works.

The information from the questionnaire was used to characterize the watershed and the management operations required by ArcSWAT, such as, planting, harvesting, and fertilizer applications for the agricultural HRUs. Other key options that were selected for these simulations included (1) the Runoff Curve Number (CN) method for estimating surface runoff from precipitation, (2) the Penman-Monteith method for estimating potential evapotranspiration (ET), and (3) the variable storage method to simulate channel water routing. All other model parameter values were left as the default values.

The model was run for a 10-year period, January 1, 2002 to December 31, 2011 selected based on the availability of climatic data. Additionally, January 1, 2001 to December 31, 2001 was used to initialize the model. The same parameters included for uncertainty analysis for Little River EW were used for Huanquisco watershed: SOL_AWC, SOL_BD, USLE_K, SOL_CBN, SOL_ALB. The output distributions were presented as modified Tukey box plots providing a range and distribution of uncertainty of SWAT outputs due to selected soil parameters and also were used to establish the 10^{th} and 90^{th} percentiles for the output.

The model was run for a second time with a daily time step for the time period 2008 and 2009 to perform a comparison between predicted and observed streamflow values. Streamflow data for Huanquisco RW is available for the time period from December 31, 2009 to May 31, 2009. The information was collected through a monitoring program sponsored by SANREM-CRSP (USAID) (Penaranda et al., 2011).

3.1.3. Field soil sampling and sample processing

Local measurements to define texture of the top soil layer were taken easily and inexpensively. The field soil survey was used to create a geospatial database of sampling points and textural proportions.

A 1 km by 1 km grid was created to guide the soil sampling in order to reduce the sampling bias and to obtain a homogeneous distribution of the sampling. However, the final sampling points were decided in the field based on accessibility of location. A total of 140 soil samples were collected. With a sampling tube, five subsamples inside each grid were taken and combined for one sample attributed to the geo-referenced point. For each grid, the soil subsamples were taken following a sigma pattern (∑) to better characterize the variability within

the grid. Soil sampling depth was set to 12 to 20 cm for each subsample. The subsamples were mixed together in a plastic bag. Each bag was identified with a sequential number and recorded waypoint from a GPS unit.

Soil samples were processed to determine their texture following the particle separation by suspension method procedure described in detail in Appendix A. Once the proportions were determined, the soil texture class was defined by triangulating the formation in the soil texture triangle (Appendix B). With the texture class and textural proportion of sand (S), silt (Y) and clay (C) for each of the nine soil units defined in the map, a random selection from the pool of homogeneous subsets in the SWAT database was performed using a uniform probability function where the minimum and maximum values were set to the interval of the matching entries. The result of this process is a table of map unit IDs and their corresponding S5ID codes.

Chapter IV. Results

4.1. Objective 1 - Characterize the uncertainty in SWAT's outputs due to input parameters likely to be lacking in data-poor environments.

4.1.1. Manual calibration of streamflow

In order to set the baseline dataset, all parameters of interest were manually set to default values reported in the SWAT user's manual. The evaluation process started with the calibration of SWAT for streamflow. The baseline was run, and only the R^2 criterion was met (first row in table 7). Adjusting CN2, GW_REVAP and EPCO did not cause enough change in the predictions to make the statistics meet the criteria (table 7), so their values were maintained at their initial values. SWAT uses the curve number (CN2) to estimate surface runoff and initial abstractions. The Curve Number is a dimensionless parameter which is related to soil type, land use, hydrologic condition, and antecedent soil moisture condition (SCS, 1986). Curve Numbers are derived from experimental data using daily rainfall and runoff data from monitoring experiments. Therefore, comprehensive information is required to adjust CN2. Because there was not enough information to make further adjustments to this parameter, no additional changes were made to CN2. On the other hand, the soil evaporation compensation factor (ESCO) adjusts the depth of the soil profile from which SWAT meets soil evaporative demand. As ESCO increases, the depth to which soil evaporative demand can be met decreases (Neitsch et al., 2005). This limits soil evaporation and reduces the simulated value for evapotranspiration. ESCO is reported as a sensitive parameter in the literature (Feyereisen et al., 2007); however, it is not a physically based parameter. Therefore, it made more sense to calibrate streamflow by further adjustment of this factor. Only adjustments to ESCO made a positive impact on SWAT predictions of annual streamflow. When ESCO was set to 0.4 (fifth row in table 7), statistics were produced that met the calibration criteria.

Table 7. Statistics for annual calibration process for streamflow simulated with SWAT for Little River Experimental Watershed (1987-1992).

Description	Statistics				Parameter Values
	R^2 [(1)] ≥ 0.65	NSE[(2)] ≥ 0.65	PBIAS[(3)] ≤ ±15	RSR[(4)] ≤ 0.60	
Baseline	0.77	-1.53	-40.20	1.45	CN2=f(landuse); GW_REVAP=0.02; ESCO=0.95; EPCO=1.0 (model default values)
Change: Adjusted CN2	0.53	-1.25	-38.60	1.37	CN2=baseline f(landuse) +10%; GW_REVAP=0.02; ESCO=0.95; EPCO=1.0
Change: Adjusted ESCO	0.70	0.57	-10.90	0.59	CN2=baseline f(landuse); GW_REVAP=0.02; ESCO=0.5; EPCO=1.0
Calibrated	0.86	0.76	-9.90	0.45	ESCO=0.4; CN2=baseline f(landuse); GW_REVAP=0.02; EPCO=1.0

(1) Coefficient of determination (R^2)
(2) Root Mean Square Error observation standard deviation ratio (RSR)
(3) Nash-Sutcliffe Efficiency (NSE)
(4) Percent Bias (PBIAS)

With the calibrated model parameters for annual values as the starting point, the model was run for the same period of time (1987-1992) to evaluate monthly values. Compared with the annual values, as expected, the statistics showed a slight decrease for R^2 and NSE and a slight increase in RSR. On the other hand, a slight improvement was seen in the PBIAS value. However, all statistics met the criteria set for monthly values without any further adjustment (table 8).

Table 8. Statistics for monthly calibration process for streamflow simulated with SWAT for Little River Experimental Watershed (1987-1992).

Description	Statistics				Parameter Values
	R^2 [(1)] ≥ 0.65	NSE[(2)] ≥ 0.65	PBIAS[(3)] ≤ ±15	RSR[(4)] ≤ 0.60	
1st Try	0.77	0.56	-2.80	0.65	CN2=-10%; GW_REVAP=0.02; ESCO=0.4; EPCO=1.0
Calibrated	0.81	0.73	-9.70	0.51	ESCO=0.4; CN2=0%; GW_REVAP=0.02; EPCO=1.0

(1) Coefficient of determination (R^2)
(2) Root Mean Square Error observation standard deviation ratio (RSR)
(3) Nash-Sutcliffe Efficiency (NSE)
(4) Percent Bias (PBIAS)

With the set of parameters obtained from the calibration for the time interval 1987 to 1992, the model was run for the 1993 to 1999 time interval. The results of the validation process for annual and monthly predictions are shown in tables 9 and 10, respectively. The model predictions met the criteria for annual values. On the other hand, statistics for the evaluation of the predictions of monthly values showed that while the criteria were met for R^2 and PBIAS, NSE and RSR fell slightly short of the criteria.

Table 9. Statistics for annual validation process for streamflow simulated with SWAT for Little River Experimental Watershed (1993-1998).

Description	Statistics				Parameter Values
	R^2 [(1)] ≥ 0.65	NSE[(2)] ≥ 0.65	PBIAS[(3)] ≤ ±15	RSR[(4)] ≤ 0.60	
Validated	0.78	0.68	9.90	0.52	ESCO=0.4; CN2=0%; GW_REVAP=0.02; EPCO=1.0

(1) Coefficient of determination (R^2)
(2) Root Mean Square Error observation standard deviation ratio (RSR)
(3) Nash-Sutcliffe Efficiency (NSE)
(4) Percent Bias (PBIAS)

Table 10. Statistics for monthly validation process for streamflow simulated with SWAT for Little River Experimental Watershed (1993-1998).

Description	Statistics				Parameter Values
	R^2 [(1)] ≥ 0.65	NSE[(2)] ≥ 0.65	PBIAS[(3)] ≤ ±15	RSR[(4)] ≤ 0.60	
Validated	0.60	0.37	10.20	0.79	ESCO=0.4; CN2=0%; GW_REVAP=0.02; EPCO=1.0

(1) Coefficient of determination (R^2)
(2) Root Mean Square Error observation standard deviation ratio (RSR)
(3) Nash-Sutcliffe Efficiency (NSE)
(4) Percent Bias (PBIAS)

Additionally, to observe a general visual agreement between observed and simulated data, the results of the annual and monthly values of streamflow were plotted in figures 10 and 11, respectively. Scatter plots with a one-to-one line were prepared for monthly values. Figure 12 show a generally even distribution for over predicted and under predicted values for the calibrated values. A similar plot for the validation (fig. 13) shows a couple outliers, which are the reason for the NSE and RSR statistics not meeting the criteria.

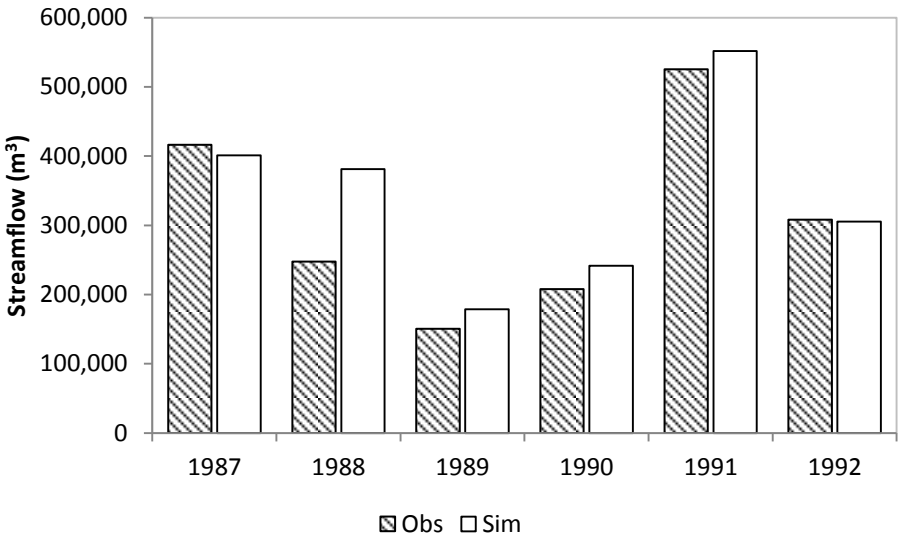


Figure 10. Observed and simulated annual streamflow (m^3) for the calibration process for Little River Experimental Watershed (1987-1992).

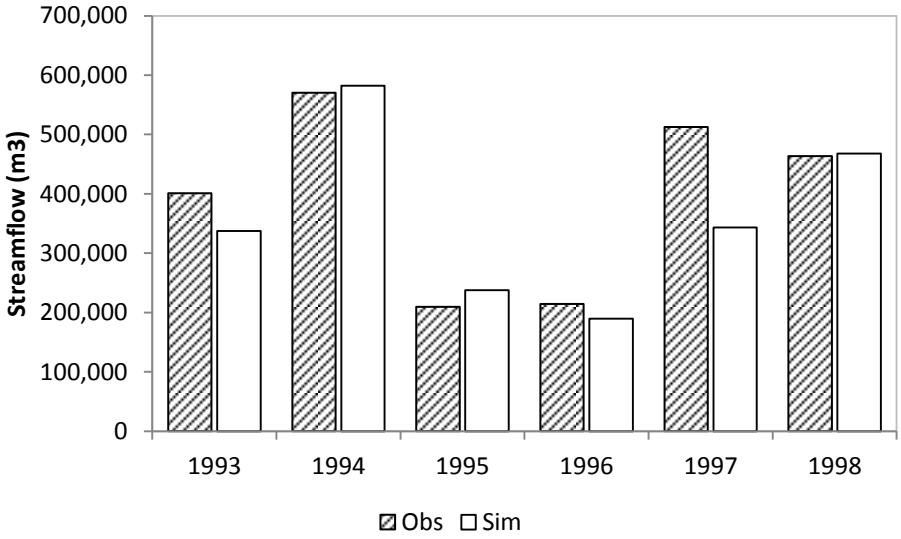


Figure 11. Observed and simulated annual streamflow (m^3) for the validation process for Little River Experimental Watershed (1993-1998).

Scatter plots with a one-to-one line were prepared for monthly values. Figure 12 shows a generally even distribution for over predicted and under predicted values.

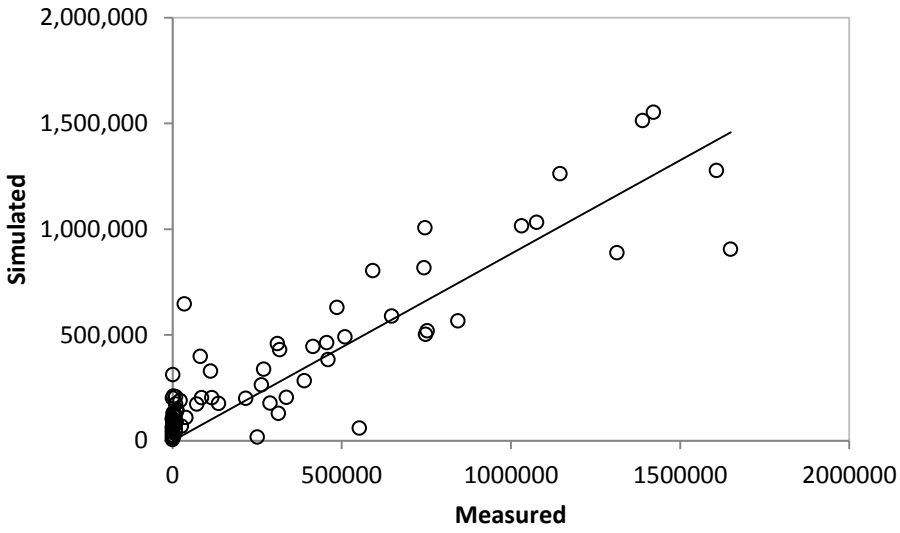


Figure 12. Observed and simulated monthly streamflow (m^3) for the calibration for Little River Experimental Watershed (1987-1992).

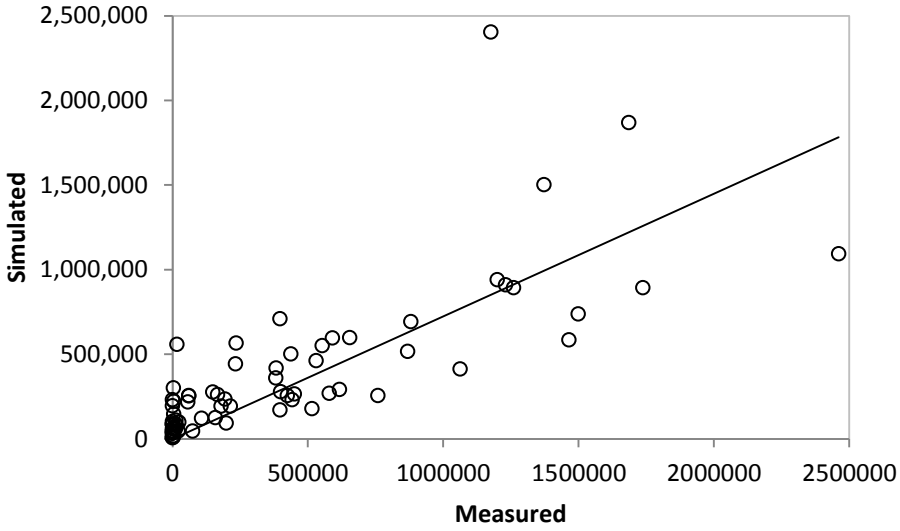


Figure 13. Observed and simulated monthly streamflow (m^3) and for the validation process for Little River Experimental Watershed (1993-1998).

Calibration of the SWAT model for sediments and nutrients for Little River Experimental Watershed was not completed successfully. In the case of sediments, the measured data did not match the evaluation period defined for streamflow (1986-1999). Through a personal communication, researchers of the Southeast Watershed Research Laboratory (SEWRL) reported inconsistency in the measured datasets for sediments and nutrients for LREW. An unsuccessful attempt to calibrate SWAT for sediments and nutrients was also performed at SEWRL. The available measured sediment data corresponds to years 2000 to 2003, while the

predicted values correspond to years 1986 to 1998. The period of time 2000 to 2003 was not included in the evaluation period because it did not provide the appropriate balance between wet and dry years, which was the initial criteria to define the evaluation time. hree out of four years in the 2000-2003 period were drier than the average. Therefore, due to the mismatch in the periods of available measured data and the predicted interval, sediment yield was not calibrated.

In the case of nutrients, measured total nitrogen and phosphorous are available for a longer period of time, however, the trend of the measured data made manual calibration impossible. Specifically, reported measured nutrient loads do not match the trend of precipitation; in addition to that, the differences in magnitude between observed and predicted values were enormous. Additionally, since SWAT nutrient predictions are driven by precipitation, the distribution of the predicted values is completely different than the distribution showed by the measured dataset. There is not a regular pattern between the measured and predicted datasets. As observed (fig. 14), some years are over-predicted and others are under-predicted. Moreover, cases in which the measured phosphorous have similar values, simulated values have different responses; one year is over-predicted and the other is under-predicted. A close look was taken to the measured and predicted relationships taking into account precipitation and nutrient loads; the result did not provide more light to solve the problem. Because of this particular behavior of the observed data (fig. 14), manual calibration was unsuccessful because it adjusts either the over predicted values or the under-predicted ones and vice versa.

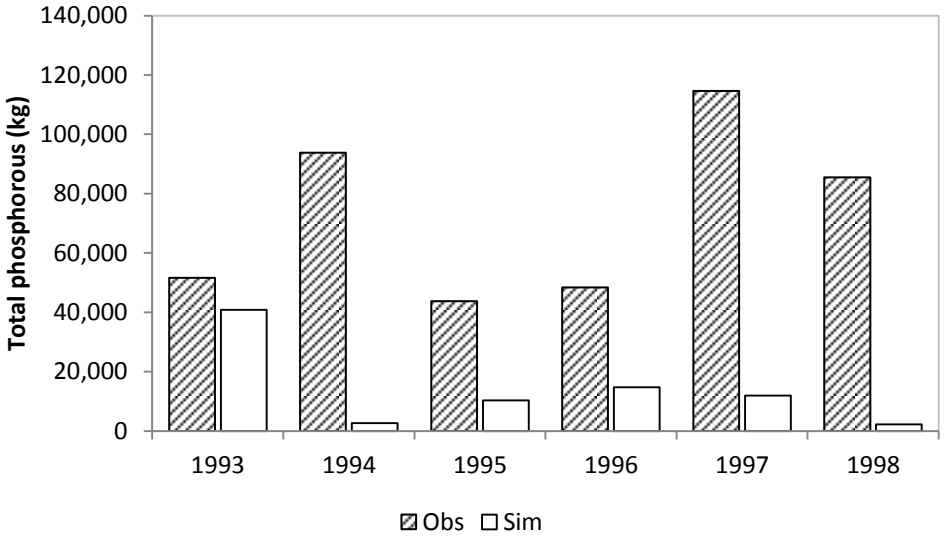


Figure 14. Observed and simulated annual total phosphorous (kg) for the calibration period for Little River Experimental Watershed (1987-1992).

When manual calibration managed to adjust over-predicted values, the other set was completely under-predicted. Consequently, the computed statistics are completely off of the required criteria (table 11).

Table 11. Statistics for annual calibration process for total phosphorous simulated with SWAT for Little River Experimental Watershed (1987-1992).

Description	Statistics				Parameter Values
	R^2 [(1)] ≥ 0.65	NSE[(2)] ≥ 0.65	PBIAS[(3)] ≤ ±15	RSR[(4)] ≤ 0.60	
Baseline	0.04	-2.00	-74.40	1.58	ESCO=0.4; CN2=f(landuse); GW_REVAP=0.02; EPCO=1.0; SOL_ORGP=5; PPERCO = 10
Change: Sol_orgP	0.06	-1.88	-69.40	1.55	ESCO=0.4; CN2=f(landuse); GW_REVAP=0.02; EPCO=1.0; SOL_ORGP=1; PPERCO = 10
Change: Sol_orgP	0.06	-1.87	-68.60	1.55	ESCO=0.4; CN2=f(landuse); GW_REVAP=0.02; EPCO=1.0; SOL_ORGP=0.1; PPERCO = 10
Change: Sol_orgP + PPerco	0.25	-96.45	107.80	9.01	ESCO=0.4; CN2=f(landuse); GW_REVAP=0.02; EPCO=1.0; SOL_ORGP=0.1; ; PPERCO = 15
Change: Sol_orgP + PPerco	0.03	-1.80	-40.60	1.53	ESCO=0.4; CN2=f(landuse); GW_REVAP=0.02; EPCO=1.0; SOL_ORGP=0.1; PPERCO = 17.5

[(1)] Coefficient of determination (R^2)
[(2)] Root Mean Square Error observation standard deviation ratio (RSR)
[(3)] Nash-Sutcliffe Efficiency (NSE)
[(4)] Percent Bias (PBIAS)

In this part of the study, manual calibration was performed to find a set of model parameters that can successfully approximate observed values. The set of calibrated parameters could have been used for other watersheds with similar characteristics because similar watersheds will show a similar hydrological behavior and thus can be modeled using the same set of model parameters. However, several drawbacks arise due to the non-uniqueness of the model parameters; consequently based on the results obtained in this study, generalization of calibrated model parameters could not be proved appropriate.

4.1.2. Sensitivity analysis

The analysis was based in this statement: the greater the relative sensitivity (Sr), the more sensitive a model output variable was to that particular parameter. Details of the sensitivity analysis results are given in the following sections. In summary, streamflow, sediment yield, total nitrogen and total phosphorous were found to be most sensitive to bulk density (SOL_BD), available water capacity (SOL_AWC), organic carbon content (SOL_CBN), soil albedo

(SOL_ALB), saturated hydraulic conductivity (SOL_K), and USLE soil erodibility (USLE_K), based on detailed sensitivity analyses. On the other hand, streamflow, sediment yield, total nitrogen and total phosphorous were not sensitive to porosity fraction from which anions are excluded (ANION_EXCL) and salinity (SOL_EC).

4.1.2.1. Sensitivity of the model predictions to bulk density

The sensitivity of streamflow to bulk density seems to be parabolic; with streamflow predictions increasing as the percentage change from the baseline is greater, whether positive or negative. There is a direct relationship between bulk density and sediment yield. When there are more voids or the pores are bigger, the capacity of the soil to stay together is less, so the energy of the rain will detach the particles and they will be transported. Bulk density has an inverse relationship with total nitrogen and total phosphorous. It seems that when the pores are smaller they can hold nutrients better than when the voids are bigger. Table 12 shows how changes in bulk density impacted different outputs. Streamflow is slightly to moderately sensitive, sediment yield is moderately sensitive to sensitive, N is moderately sensitive and P moderately sensitive to sensitive.

Table 12. Results of sensitivity analysis of SWAT outputs with changes of bulk density (SOLBD) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 0.87	-50	3.84	0.08[2]	-26.57	0.53[3]	36.64	0.73[3]	33.09	0.66[3]
0 - 1.29	-25	2.58	0.10[3]	-22.80	0.91[3]	17.67	0.71[3]	24.55	0.98[3]
0 - 1.56	-10	0.57	0.06[2]	-10.73	1.07[4]	1.46	0.15[3]	10.67	1.07[4]
0 - 1.73	- Baseline -								
0 - 1.90	10	.095	0.10[2]	2.30	0.23[3]	-2.00	0.20[3]	-3.85	0.39[3]
0 - 2.16	25	8.22	0.33[3]	12.91	0.52[3]	-12.33	0.49[3]	-10.84	0.43[3]
0 - 2.59	50	12.65	0.25[3]	10.95	0.22[3]	-23.51	0.47[3]	-22.25	0.44[3]

Note: (1) Insensitive: S_r <|0.01|
(2) Slightly sensitive: |0.01|≤ S_r <|0.1|
(3) Moderately sensitive: |0.1|≤ S_r <|1.0|
(4) Sensitive: |1.0|≤ S_r <|2.0|
(5) Extremely sensitive: S_r ≥|2.0|

4.1.2.2. Sensitivity of the model predictions to available water capacity

Available water capacity has an inverse relationship with streamflow, sediment yield, total nitrogen and total phosphorous. Changes in the AWC value change streamflow and sediment yield almost linearly. Consequently, values of dissolved and attached nutrients are

also changed in the same pattern. As shown in table 13, all of the outputs are moderately sensitive to AWC.

Table 13. Results of sensitivity analysis of SWAT outputs with changes of available water capacity (SOL_AWC) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 0.35	-50	26.20	0.52[3]	29.28	0.59[3]	36.64	0.73[3]	9.27	0.19[3]
0 - 0.53	-25	11.44	0.46[3]	14.32	0.57[3]	17.67	0.71[3]	7.45	0.30[3]
0 - 0.63	-10	4.56	0.46[3]	6.67	0.67[3]	1.46	0.15[3]	2.12	0.21[3]
0 - 0.70	---------- Baseline ----------								
0 - 0.77	10	-4.08	0.41[3]	-5.11	0.51[3]	-2.00	0.20[3]	-2.24	0.22[3]
0 - 0.86	25	-7.75	0.31[3]	-9.13	0.37[3]	-12.33	0.49[3]	-3.78	0.15[3]
0 - 1.00	50	-14.63	0.29[3]	-19.21	0.38[3]	-23.51	0.47[3]	-12.07	0.24[3]

Note: (1) Insensitive: S_r <|0.01|
(2) Slightly sensitive: |0.01|≤ S_r <|0.1|
(3) Moderately sensitive: |0.1|≤ S_r <|1.0|
(4) Sensitive: |1.0|≤ S_r <|2.0|
(5) Extremely sensitive: S_r ≥|2.0|

4.1.2.3. Sensitivity of the model predictions to organic carbon content

The sensitivity pattern of the outputs to organic carbon content (SOL_CBN) is presented in table 14. Streamflow varies linearly with organic carbon content and is slightly sensitive to carbon content. According to this result, the more organic carbon content in the soil, the more streamflow is generated. Carbon content is commonly used to characterize the amount of organic matter in soils. Soil organic matter content and composition affect both soil structure and adsorption properties; therefore, water retention may be affected by changes in soil organic matter (Rawlsa et al., 2003). Therefore, the result obtained for streamflow did not make sense, because in theory, water retention in the soil increases with more organic matter content so less runoff should be observed.

There is no significant influence of carbon content on sediment yield at levels of SOL_CBN below 2.03. Sediment yield is slightly sensitive to carbon content. For both nutrients, N and P, organic carbon content has a direct relationship, with increased yield of nitrogen and phosphorous as the carbon content is increased. Total N and P are moderately sensitive to changes in carbon content.

Table 14. Results of sensitivity analysis of SWAT outputs with changes of organic carbon content (SOL_CBN) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 1.015	-50	-0.73	0.01[2]	-0.27	0.01[1]	-6.75	0.14[3]	-8.14	0.16[3]
0 - 1.523	-25	-0.36	0.01[2]	-0.13	0.01[1]	-3.48	0.14[3]	-4.16	0.17[3]
0 - 1.827	-10	-0.11	0.01[2]	-0.03	0.00[1]	-1.23	0.12[3]	-1.47	0.15[3]
0 - 2.030	- - - - - - Baseline - - - - - -								
0 - 2.233	10	0.16	0.02[2]	-0.24	0.02[2]	0.80	0.08[2]	1.39	0.14[3]
0 - 2.538	25	0.34	0.01[2]	-0.40	0.02[2]	2.53	0.10[3]	3.77	0.15[3]
0 - 3.045	50	0.62	0.01[2]	-0.30	0.01[1]	6.02	0.12[3]	7.90	0.16[3]

Note: (1) Insensitive: S_r <|0.01|
(2) Slightly sensitive: |0.01|≤ S_r <|0.1|
(3) Moderately sensitive: |0.1|≤ S_r <|1.0|
(4) Sensitive: |1.0|≤ S_r <|2.0|
(5) Extremely sensitive: S_r ≥|2.0|

4.1.2.4. Sensitivity of the model predictions to USLE soil erodibility

There is no impact of the USLE K-Factor on streamflow. As expected, there is a direct relationship between soil saturated hydraulic conductivity and sediment yield, total nitrogen and total phosphorous. As the value of USLE_K increases, the sediment yield increases, consequently, there are more attached nutrients transported along with the eroded particles. The sensitivity of sediment yield and nutrients to USLE_K ranges from slight to moderate (table 15).

Table 15. Results of sensitivity analysis of SWAT outputs with changes of USLE soil erodibility (USLE_K) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 0.19	-50	-0.04	0.001[1]	-11.87	0.24[3]	-7.51	0.15[3]	-28.32	0.57[3]
0 - 0.28	-25	-0.02	0.001[1]	-4.22	0.17[3]	-3.61	0.14[3]	-14.16	0.57[3]
0 - 0.33	-10	0.00	0.000[1]	-0.91	0.09[2]	-0.83	0.08[2]	-3.31	0.33[3]
0 - 0.37	- - - - - - Baseline - - - - - -								
0 - 0.41	10	0.01	0.001[1]	1.35	0.13[3]	1.48	0.15[3]	5.91	0.59[3]
0 - 0.46	25	0.01	0.000[1]	2.62	0.10[3]	2.87	0.11[3]	11.46	0.46[3]
0 - 0.55	50	0.02	0.000[1]	4.40	0.09[2]	5.60	0.11[3]	21.17	0.42[3]

Note: (1) Insensitive: S_r <|0.01|
(2) Slightly sensitive: |0.01|≤ S_r <|0.1|
(3) Moderately sensitive: |0.1|≤ S_r <|1.0|
(4) Sensitive: |1.0|≤ S_r <|2.0|
(5) Extremely sensitive: S_r ≥|2.0|

4.1.2.5. Sensitivity of the model predictions to albedo

Streamflow and sediment yield were slightly sensitive to albedo (table 16). Total phosphorous was slightly sensitivity for levels of SOL_ALB greater than 0.33 and insensitive for lower values. Total N was insensitive to changes in albedo.

Table 16. Results of sensitivity analysis of SWAT outputs with changes of albedo (SOL_ALB) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 0.22	-50	-0.60	0.01[2]	-1.02	0.02[2]	-0.47	0.009[1]	-0.47	0.009[1]
0 - 0.33	-25	-0.30	0.01[2]	-0.50	0.02[2]	-0.20	0.008[1]	-0.28	0.01[2]
0 - 0.39	10	-0.13	0.01[2]	-0.24	0.02[2]	-0.07	0.007[1]	-0.15	0.01[2]
0 - 0.44	---------- Baseline ----------								
0 - 0.48	10	0.14	0.01[2]	0.30	0.03[2]	0.08	0.008[1]	0.17	0.02[2]
0 - 0.55	25	0.36	0.01[2]	0.73	0.03[2]	0.16	0.006[1]	0.42	0.02[2]
0 - 0.66	50	0.72	0.01[2]	1.35	0.03[2]	0.35	0.007[1]	0.93	0.02[2]

Note: (1) Insensitive: S_r <|0.01|
(2) Slightly sensitive: |0.01|≤ S_r <|0.1|
(3) Moderately sensitive: |0.1|≤ S_r <|1.0|
(4) Sensitive: |1.0|≤ S_r <|2.0|
(5) Extremely sensitive: S_r ≥|2.0|

4.1.2.6. Sensitivity of the model predictions to saturated hydraulic conductivity

Streamflow, sediment yield, and both nutrients were slightly sensitive to hydraulic conductivity (table 17). There is a direct relationship between saturated hydraulic conductivity with both streamflow and sediment yield; therefore, it is expected that dissolved and attached nutrient values should also increase. However, based on the results of this part of the research, both nutrients show an inverse relationship with saturated hydraulic conductivity.

Table 17. Results of sensitivity analysis of SWAT outputs with changes of hydraulic conductivity (SOL_K) for Little River Experimental Watershed data set.

Input Parameter	Δ Input Parameter (%)	Streamflow		Sediment		Total N		Total P	
		Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r	Δ Output (%)	S_r
0 - 300	-50	-0.97	0.02[2]	-1.16	0.02[2]	1.13	0.02[2]	0.55	0.01[2]
0 - 450	-25	-0.54	0.02[2]	-0.66	0.03[2]	1.70	0.07[2]	0.64	0.03[2]
0 - 540	-10	-0.21	0.02[2]	-0.27	0.03[2]	0.68	0.07[2]	0.33	0.03[2]
0 - 600	----------	----------	----------	----------	Baseline	----------	----------	----------	----------
0 - 660	10	0.21	0.02[2]	0.30	0.03[2]	-0.78	0.08[2]	-0.18	0.02[2]
0 - 750	25	0.53	0.02[2]	0.84	0.03[2]	-1.98	0.08[2]	-0.55	0.02[2]
0 - 900	50	1.05	0.02[2]	2.45	0.05[2]	-2.77	0.08[2]	-0.72	0.01[2]

Note: (1) Insensitive: $S_r < |0.01|$
(2) Slightly sensitive: $|0.01| \leq S_r < |0.1|$
(3) Moderately sensitive: $|0.1| \leq S_r < |1.0|$
(4) Sensitive: $|1.0| \leq S_r < |2.0|$
(5) Extremely sensitive: $S_r \geq |2.0|$

4.1.3. Results of uncertainty analysis

The results of assessing uncertainty of SWAT outputs due to parameters likely to be absent in data-poor environments are presented in this section. The figures and tables present the following results. Figures 15 to 20 show results of uncertainty analyses of available water capacity, soil bulk density, USLE erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo on predicted mean annual streamflow, sediment yield, total nitrogen and total phosphorous for the Little River Experimental Watershed for 1987-1998. Each box plot set illustrates the differences due to changes in DEM resolution (10, 30 and 90 m) and to the source of the soils database (SSURGO and STATSGO). For the individual soil parameters, the analysis was performed by varying one parameter at a time, assuming that the parameter analyzed was the only uncertain parameter and that all other model parameters were known values and held constant. Additionally, figure 21 shows the results of varying all the parameters at the same time, assuming that all soil parameters included in this part of the analysis were uncertain. In the same way, figures 22 to 27 show parallel results of uncertainty analyses of available water capacity, soil bulk density, USLE erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo on predicted mean annual streamflow, sediment yield, total nitrogen and total phosphorous for Reynolds Creek Experimental Watershed for 1973-1984, and figure 28 shows the results of varying all the parameters at the same time.

Tables 18 and 19, for Little River and Reynolds Creek, respectively, show F-test results comparing the effect of DEM resolution, soils database and their combined effect on SWAT

outputs when the selected soil parameters were varied. For Little River (table 18), the individual effect of soils database source caused significant differences ($p < 0.001$) in the predictions of streamflow, sediment yield, total N and total P for all parameters. The DEM resolution led to significant differences in the predictions of sediment yield and total P for all soil parameters, while the differences for streamflow and total nitrogen for some soil parameters were not significant. DEM resolution did not lead to a significant difference in predicted average annual streamflow when soil available water content was varied. Bulk density, organic carbon content, saturated hydraulic conductivity or all of the soil parameters combined did not make a significant impact in the predictions of nitrogen.

The effect of the interaction between soils database and DEM resolution was different for each output. Significant differences were always seen for soil albedo and saturated hydraulic conductivity regardless of output. In contrast, when all the parameters were varied at the same time, there were no significant differences in annual average output as a function of soils database or DEM resolution. The interaction of soil database and DEM resolution had no impact on streamflow prediction as a function of uncertainty in available water content ($p = 0.7483$), while all other soil parameters showed differences due to the interaction. Sediment yield prediction was impacted by available water content, bulk density and the USLE erodibility factor. Total nitrogen showed no differences due to bulk density, while all other outputs showed significant differences due to the effect of the interaction. Total phosphorous showed no impact of bulk density, organic carbon content and USLE erodibility factor. For Reynolds Creek, the individual effect of soils database source (table 19), DEM resolution and the interaction were significant ($p < 0.05$) for all predictions of streamflow, sediment yield, total nitrogen and total phosphorous for all parameters. The effect of uncertainty of all soil parameters was also significant for predictions of all outputs.

Based on these results, characteristics of the watershed, such as size and topography, can have considerable influence on the predictions regardless of which soil parameter is analyzed. Under the conditions of this study, soil database had a different impact over model predictions for both watersheds. Depending on the watershed, a change in soils database source may come with a significant ($p < 0.001$) change in model predictions. Thus, soils data for data-poor environments must be generated carefully.

Changes in DEM resolution did not impact SWAT outputs in the same way for both experimental watersheds (LREW and RCEW). For LREW, a change in DEM resolution did not always changed the predictions, while for RCEW, the model responses were impacted by

changes in DEM resolution. Watershed size and topography are different; LREW has an area of 334 km^2 while RCEW has 239 km^2. In Little River, elevation goes from 81 to 148 m while in Reynolds Creek goes from 1108 to 2245 m. Therefore, it could be inferred that impacts of DEM resolution over SWAT predictions are highly influenced by the size and characteristics of the watershed. The soil parameter included in the analysis can cause differences in model predictions. The combination of soils database and DEM can also have a different impact on model predictions.

Tables 20 to 26 for Little River and 27 to 33 for Reynolds Creek show relative differences between the 10^{th} and 90^{th} percentile range of each soil-DEM resolution combination with respect to the 10^{th} and 90^{th} percentile range of the combination of SSURGO and the 10 m DEM (baseline). The relative differences are expressed as a percentage, positive values indicating the interval increased relative to the baseline; therefore, the uncertainty is higher. If the values are negative, the confidence interval was reduced and the uncertainty was also reduced. The individual effect of soil dataset had a bigger widening impact than the DEM resolution did. The combined effect of soils database and DEM resolution had an increasing effect as the combination became coarser (STATSGO and 90 m DEM).

The 10^{th} - 90^{th} percentile range quantifies uncertainty of the model predictions. Ideally, this range should be as narrow as possible; the smaller the value is (tending to zero), the less uncertain the predicted output is. For Little River experimental Watershed, regardless of the resolution of the DEM, the streamflow mean and median values for SSURGO and STATSGO databases remained at almost the same level, approximately 6.1 and 5.5 m^3s^{-1} (fig. 15 to 20), respectively, for all individual soil parameters and also for all parameters varied at the same time (fig. 21). However, the corresponding percentile values varied in several different ways. For available water capacity (table 20), bulk density (table 21), organic carbon content (table 22), and USLE soil erodibility (table 23), the 10^{th} – 90^{th} percentile ranges were wider for STATSGO than for SSURGO combinations. For albedo and saturated hydraulic conductivity (tables 22 and 24), the 10^{th} – 90^{th} percentile ranges of streamflow were narrower for STATSGO than for SSURGO. For Little River, when all the soil parameters were varied at the same time (table 26), SSURGO and 10 m resolution DEM had a wider percentile range than the combinations of SSURGO with 30 and 90 m resolution DEM. The previous result was not expected, because the SSURGO/10 m DEM represents the highest quality data for soils and elevation. This seems to be an effect of DEM resolution, because it is well known that coarser DEM resolutions will generate lower prediction values for streamflow. On the other hand, table 27 shows that the

combinations with STATSGO produced wider 10^{th} – 90^{th} percentile range than any SSURGO combination.

Regardless of the DEM resolution and whether the parameters were run individually or all at the same time, the predicted mean and median sediment yield values were around 17,600 t and 14,000 t for SSURGO and STATSGO, respectively (fig. 15 to 20). The 10^{th} – 90^{th} percentile ranges for sediment yield where wider for the following parameters: available water capacity (table 21), bulk density (table 22), and organic carbon content (table 23). The opposite behavior was observed for albedo (table 24), USLE soil erodibility (table 25), and saturated hydraulic conductivity (table 26), combinations that included SSURGO had wider percentile ranges than the combinations that included STATSGO. When all the soil parameters were varied at the same time the 10^{th} – 90^{th} percentile ranges were wider for the combinations that included SSURGO than the corresponding STATSGO combinations. Unexpectedly, the combination of STATSGO and the 90 m DEM produced the narrowest 10^{th} – 90^{th} percentile range.

For nitrogen, the impact on predicted mean and median values was different for each soil parameter. For available water capacity (fig. 22), organic carbon content (fig. 25), albedo (fig. 26), and saturated hydraulic conductivity (fig. 27), the mean and median increased as the DEM resolution became coarser. On the other hand, the mean and median tended to stay the same for bulk density (fig. 23) and tended to decrease due to USLE soil erodibility (fig. 24). Regardless of the DEM resolution, for all the soil parameters except albedo and saturated hydraulic conductivity, the 10^{th} – 90^{th} percentile ranges with all combinations of STATSGO were wider than those that included SSURGO. For albedo and saturated hydraulic conductivity, regardless the soil database, the 10^{th} – 90^{th} percentile ranges became narrower with coarser DEM. When all soil parameters were varied at the same time, the mean and median values for annual total nitrogen remained at 1,390 t for all combinations of soil dataset (SSURGO and STATSGO) and DEM resolution (10, 30, 90 m) and the 10^{th} – 90^{th} percentile ranges were wider for all the combinations of STATSGO than the combinations with SSURGO.

For total phosphorous, whether SSURGO or STATSGO was used, all individual soils parameters showed mean and median values tending to be lower as the DEM resolution became coarser. Mean and median values from the combinations including STATSGO were higher than their corresponding soil-DEM combinations that included SSURGO. The behavior described in the previous paragraph for nitrogen was repeated when all the parameters were varied at the same time. For all the soil parameters except albedo, the 10^{th} – 90^{th} percentile

ranges were again wider for DEM resolution with STATSGO than the ranges resulting when using SSURGO. For albedo (table 23), independently of the soils source, the 10^{th} – 90^{th} percentile ranges got wider as the DEM resolution got coarser. However, combinations that included STATSGO were narrower than the combinations with SSURGO. When all the parameters were varied at the same time, the 10^{th} – 90^{th} percentile ranges were wider as the input data became coarser. Both, SSURGO and STATSGO combinations had wider 10^{th} – 90^{th} percentile ranges as the DEM tended to 90 m resolution. Additionally, the 10^{th} – 90^{th} percentile ranges was wider with all the combinations than included STATSGO.

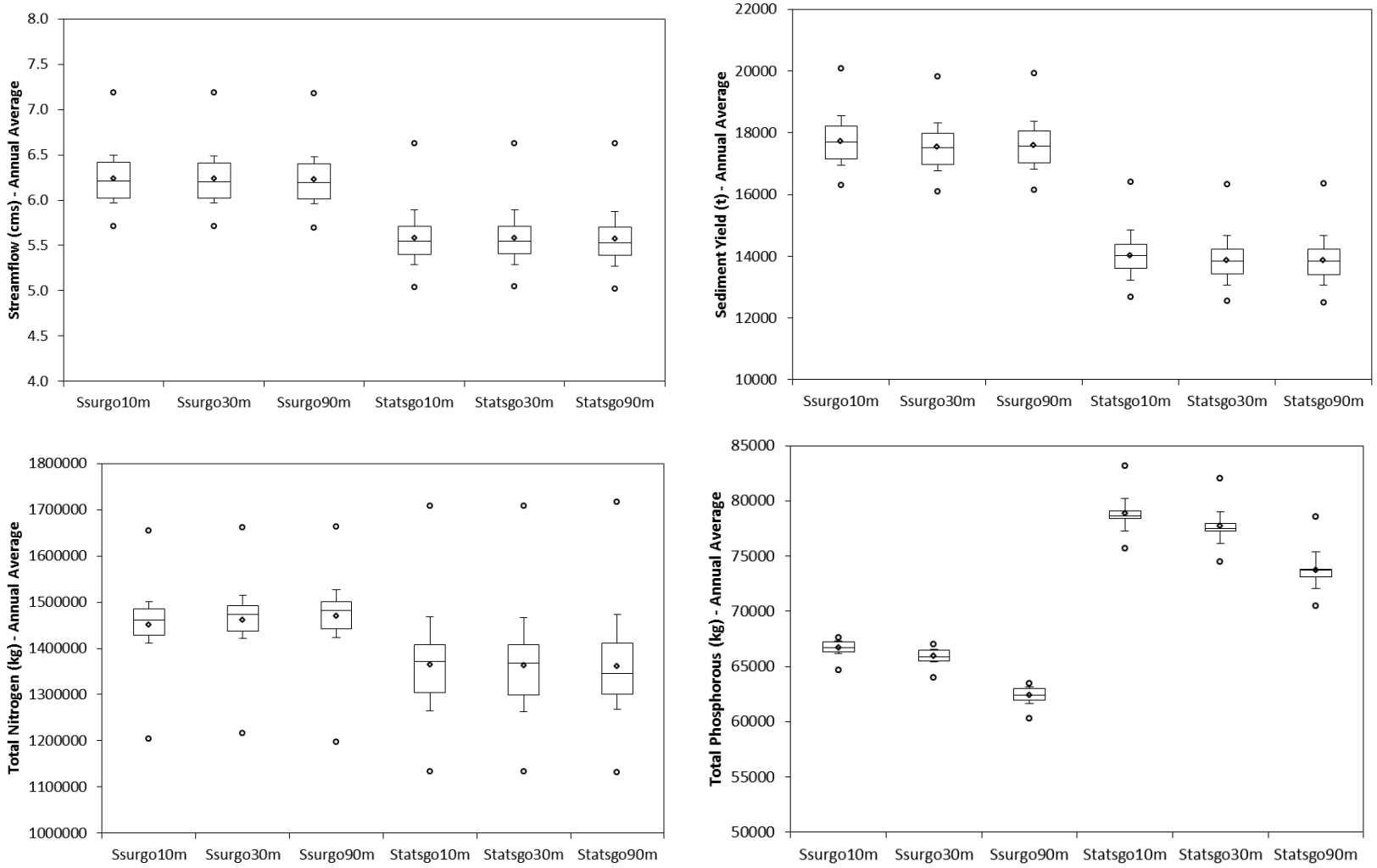


Figure 15. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to available water capacity (SOL_AWC) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

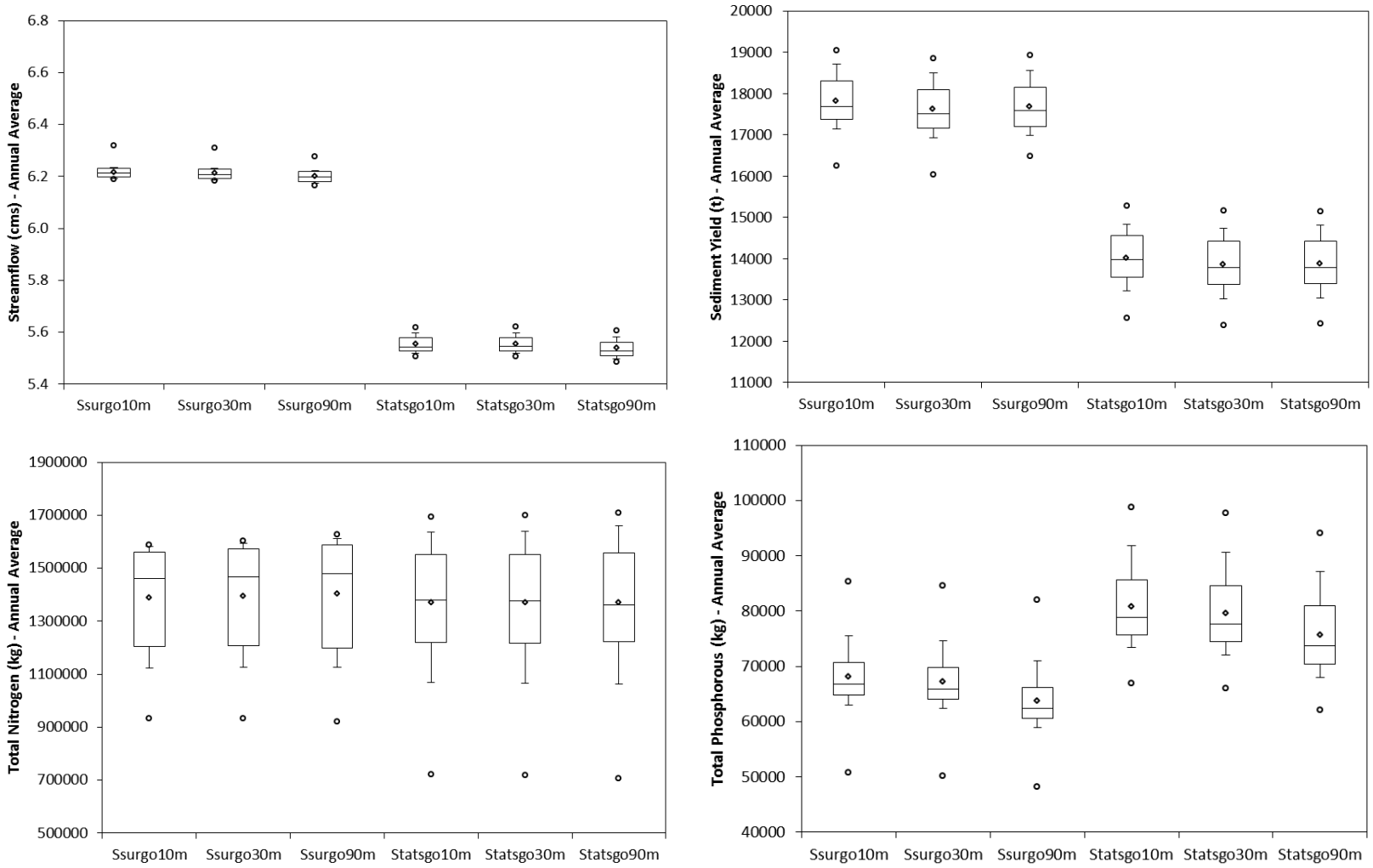


Figure 16. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to soil bulk density (SOL_BD) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

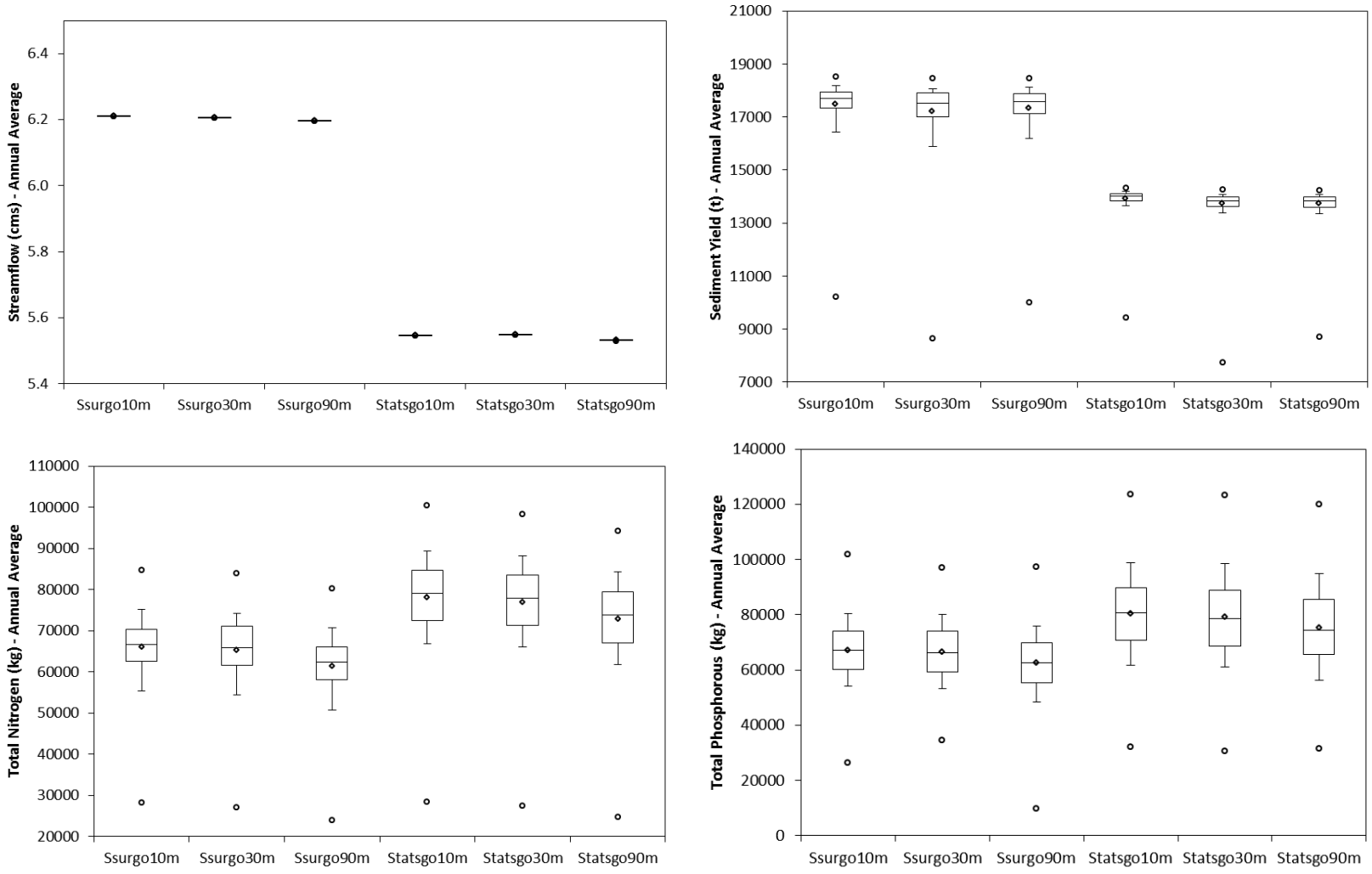


Figure 17. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to USLE equation erodibility factor (USLE_K) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

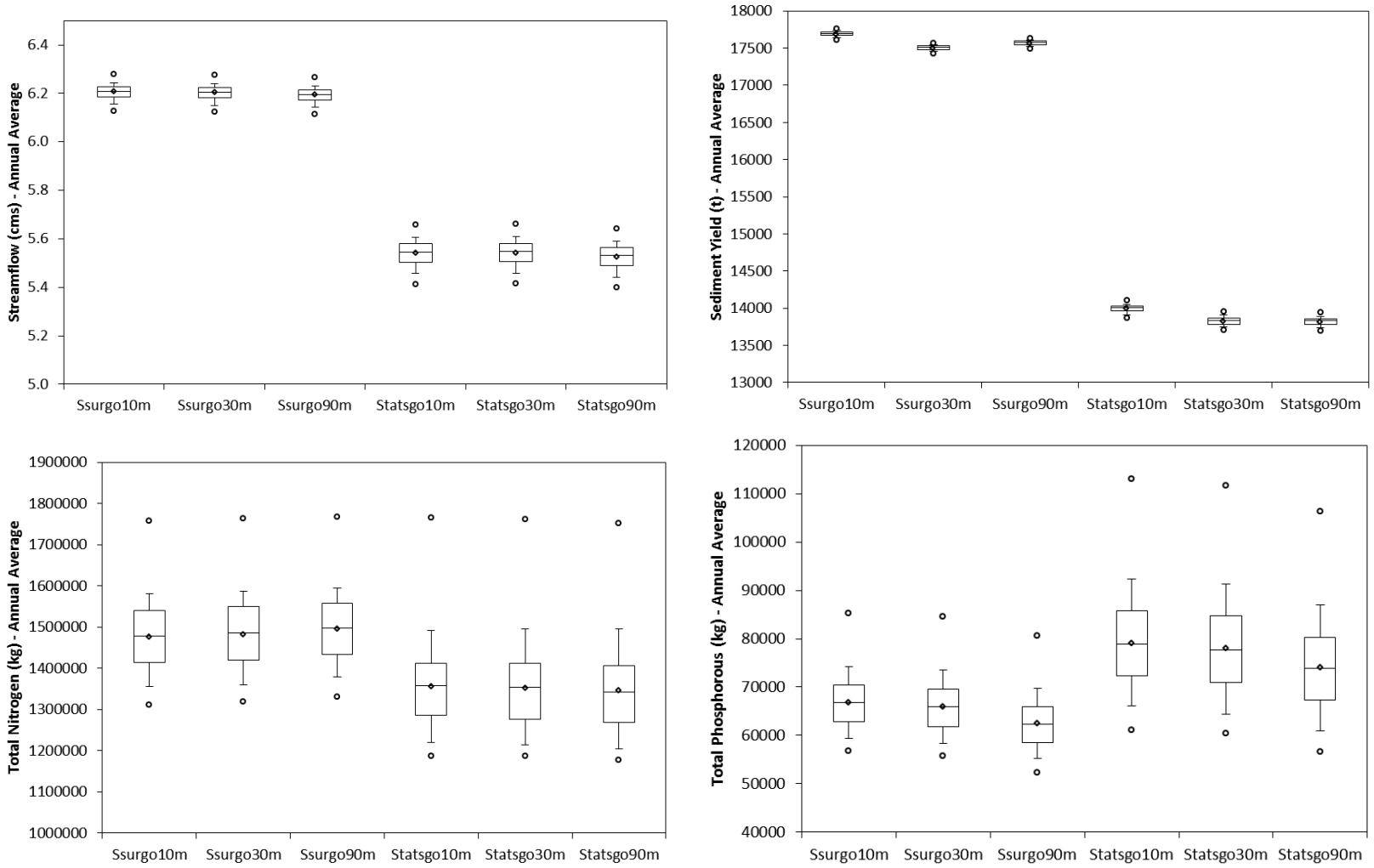


Figure 18. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to organic carbon content (SOL_CBN) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

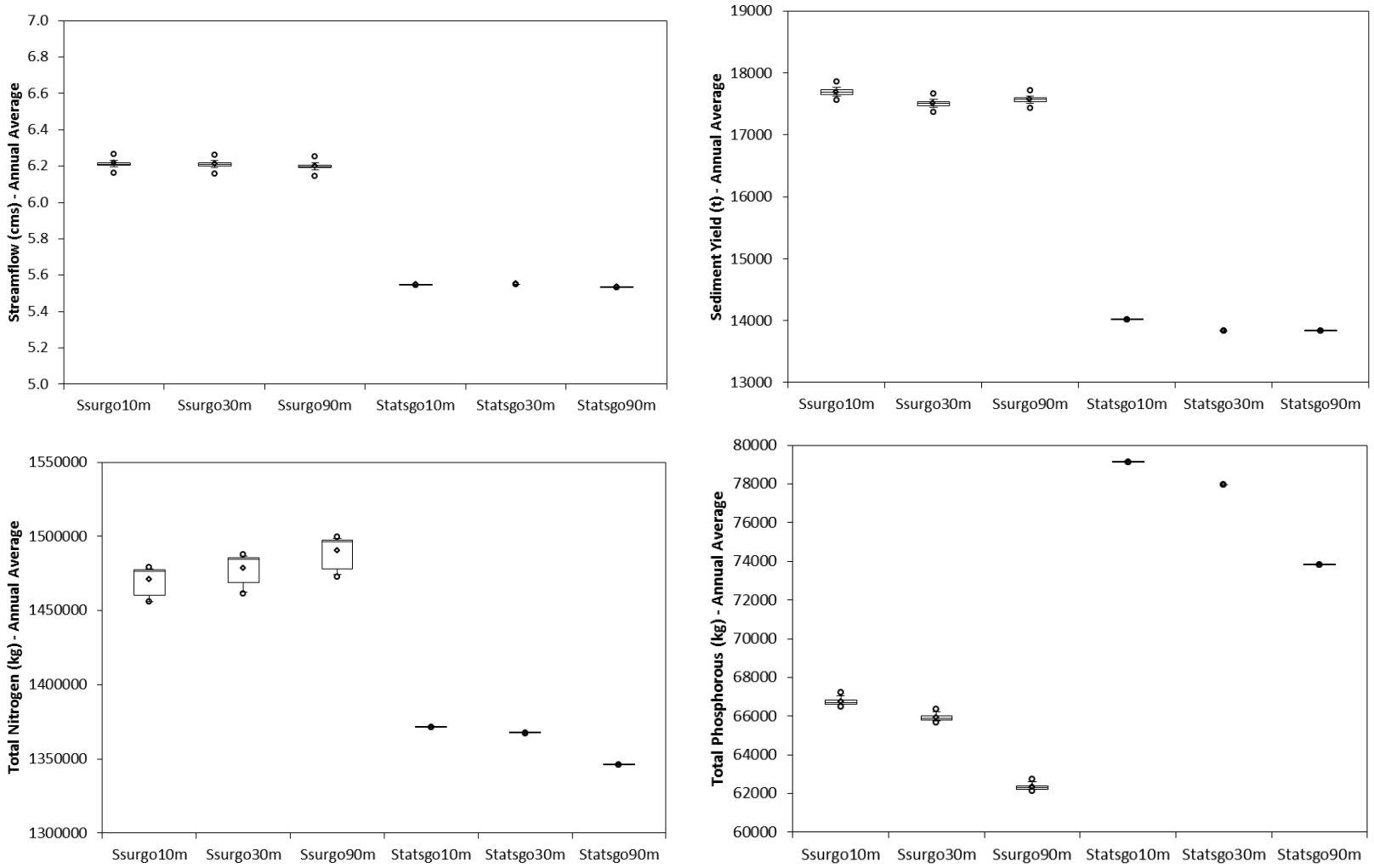


Figure 19. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to soil albedo (SOL_ALB) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

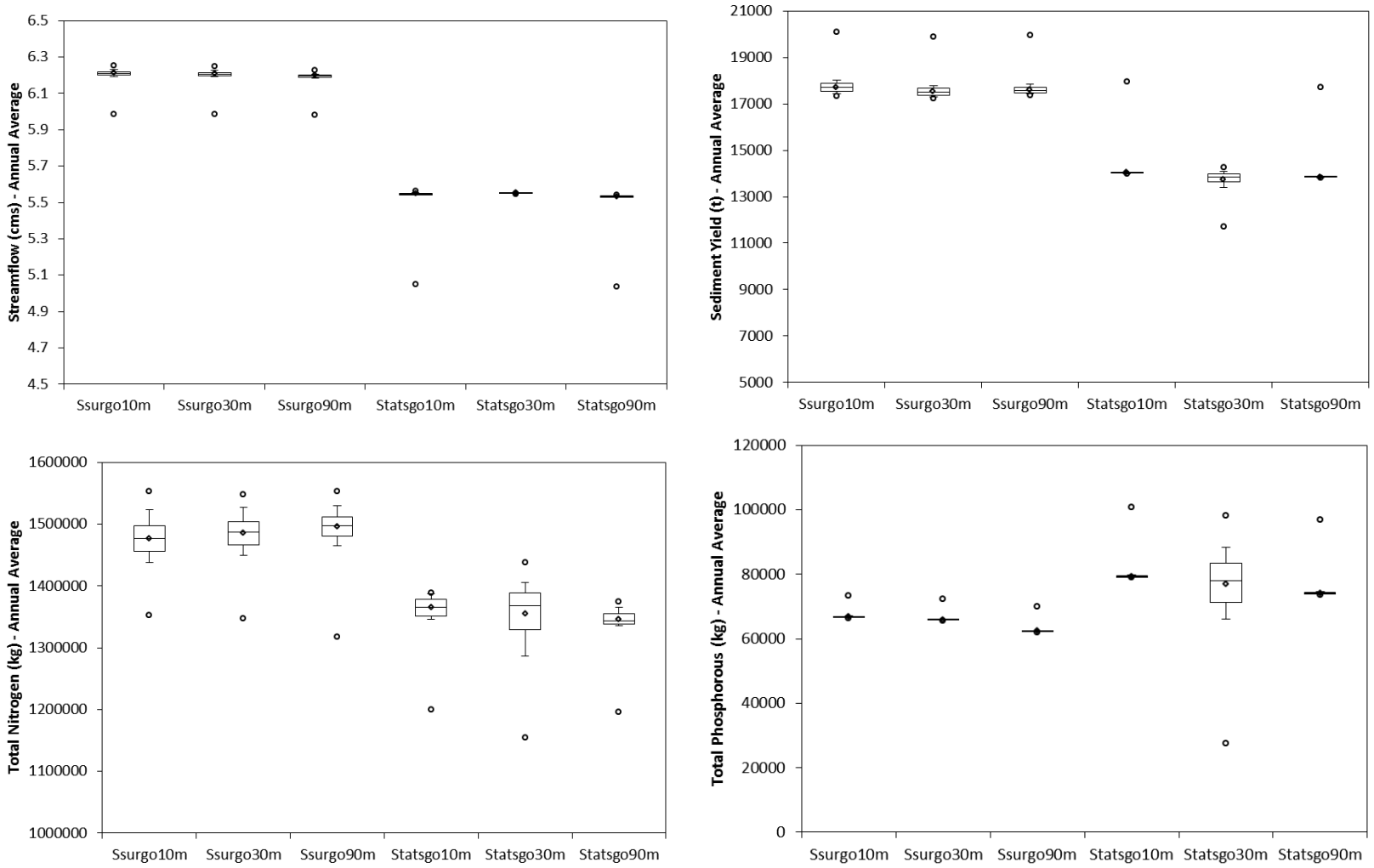


Figure 20. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to saturated hydraulic conductivity (SOL_K) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

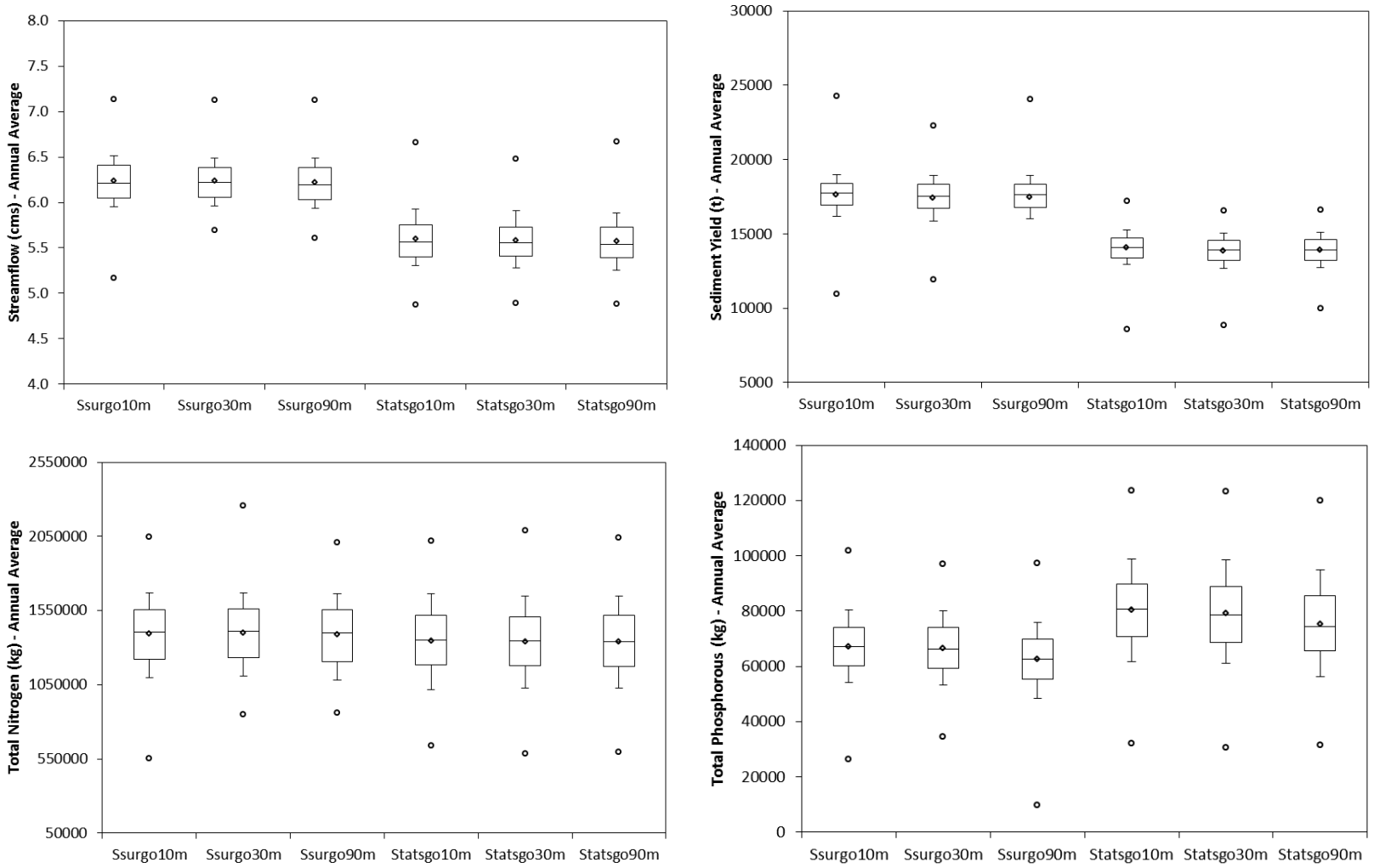


Figure 21. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to all soils factors for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

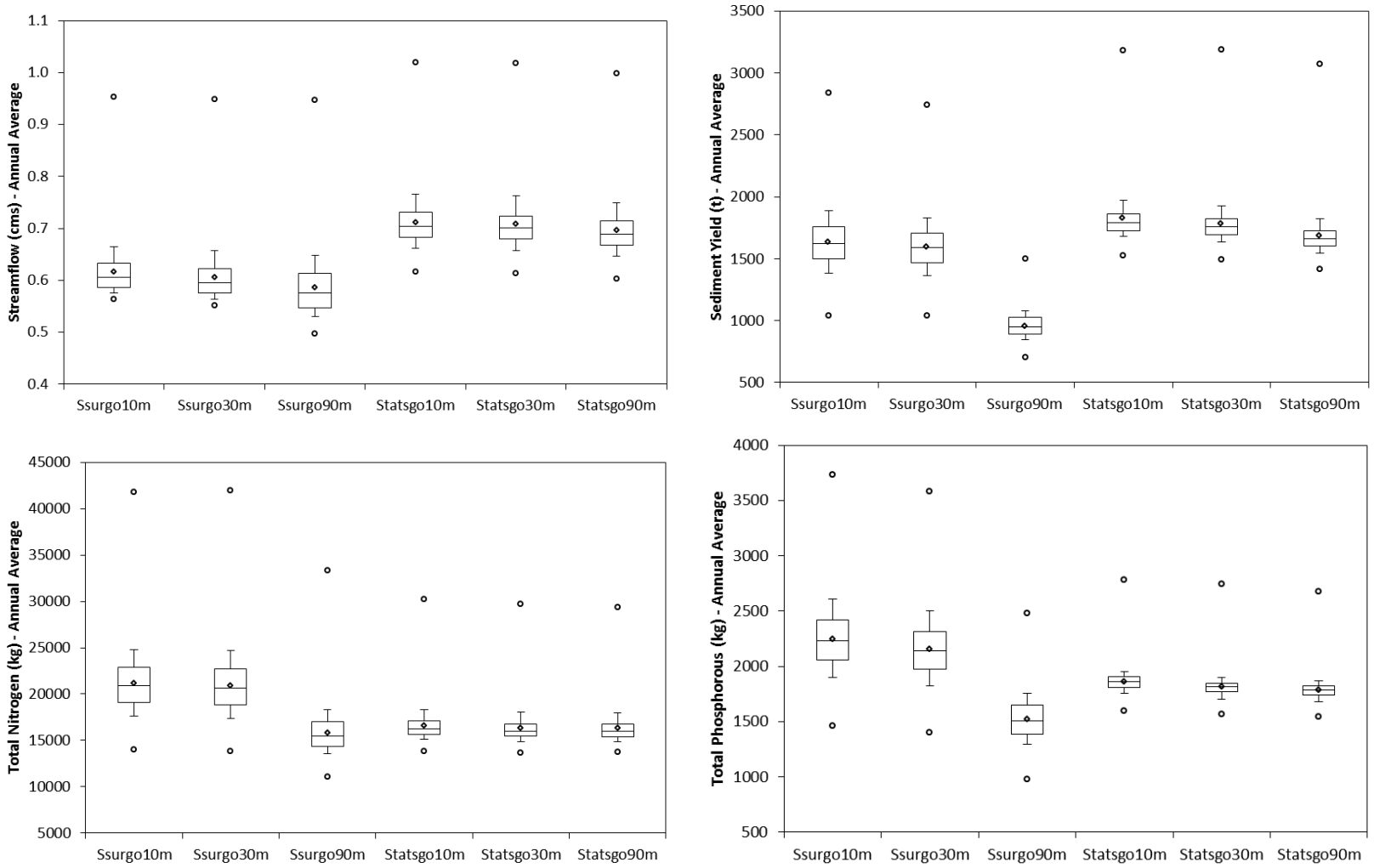


Figure 22. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to available water capacity (SOL_AWC) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

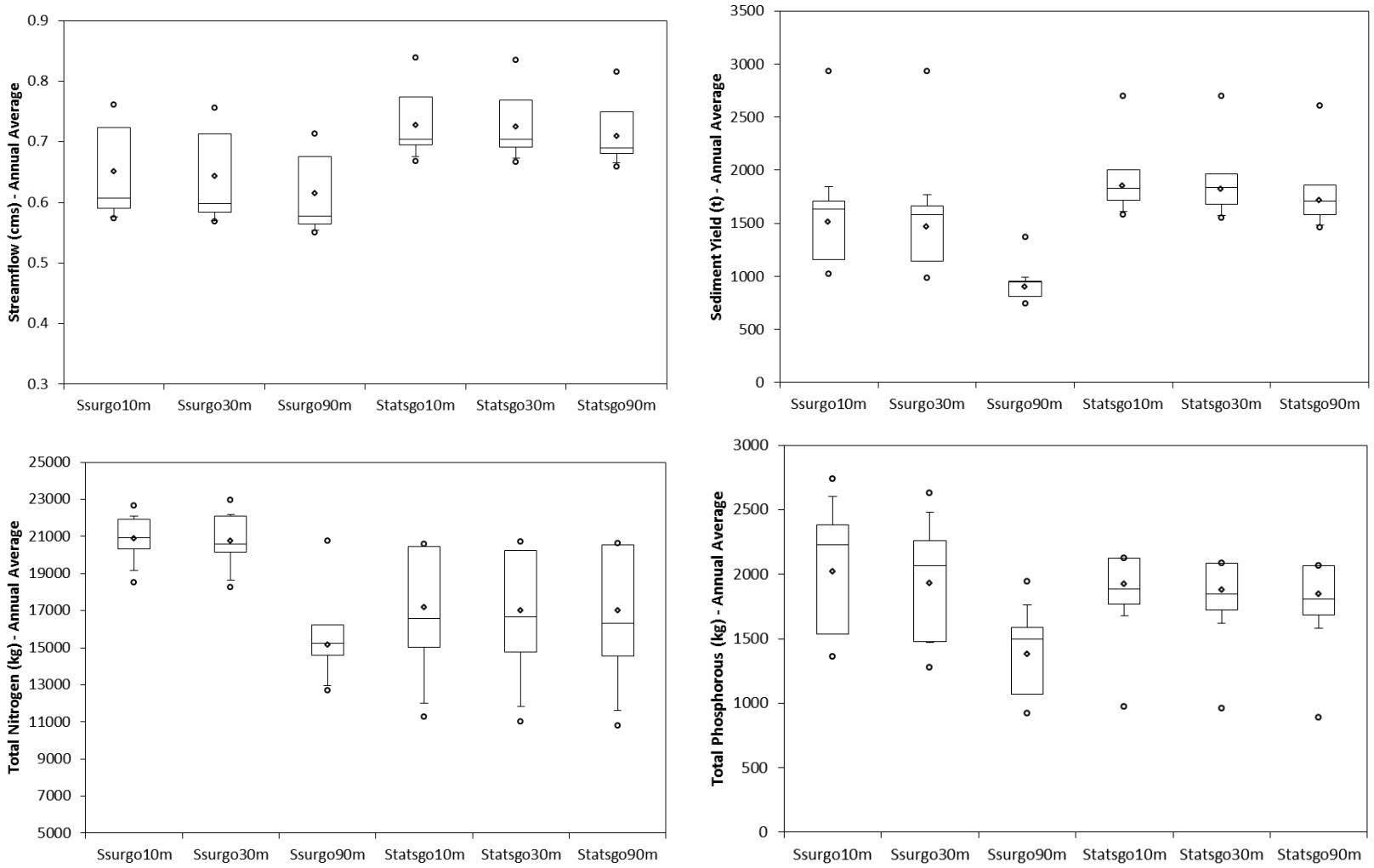


Figure 23. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to soil bulk density (SOL_BD) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median

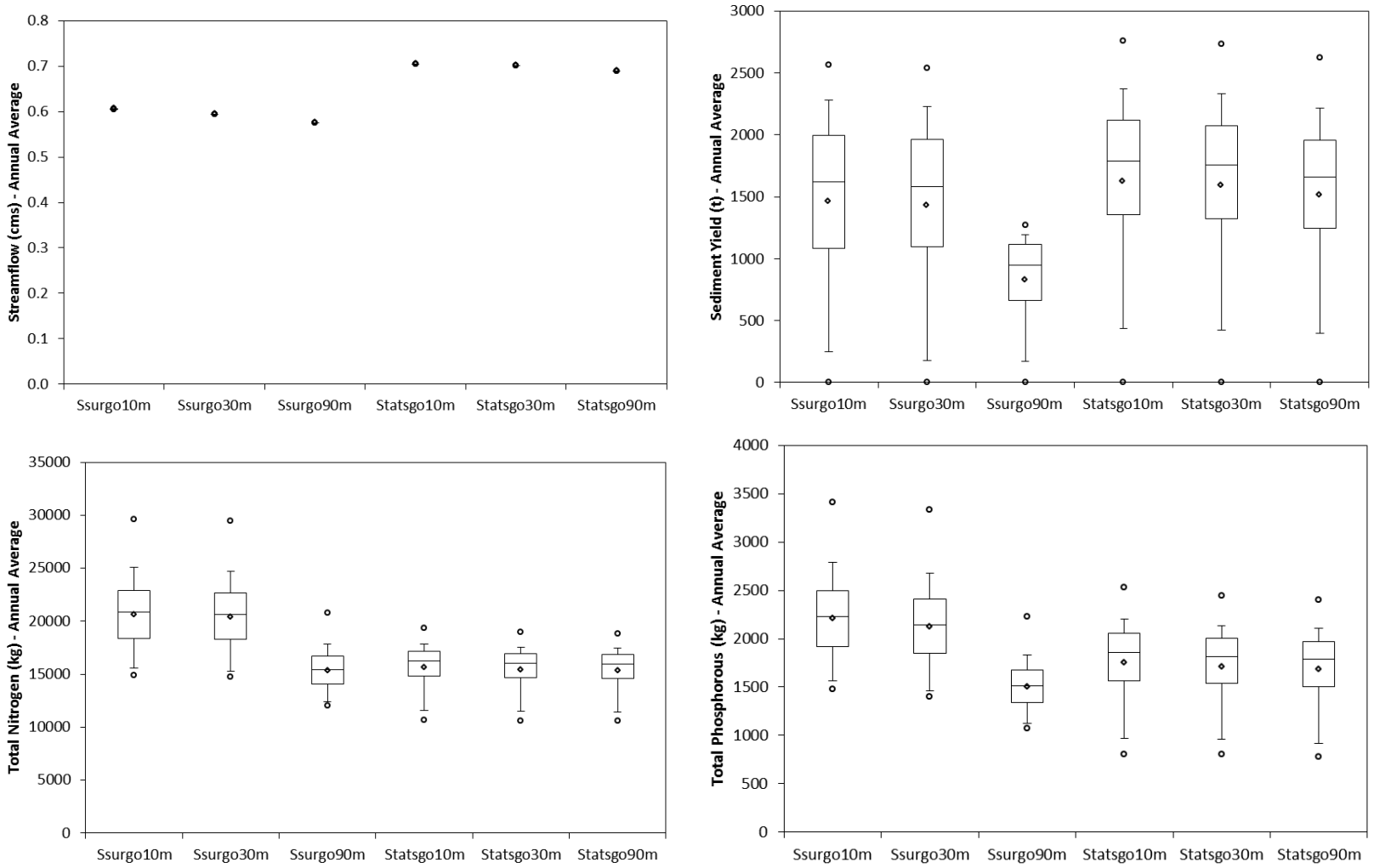


Figure 24. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to USLE equation erodibility factor (USLE_K) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

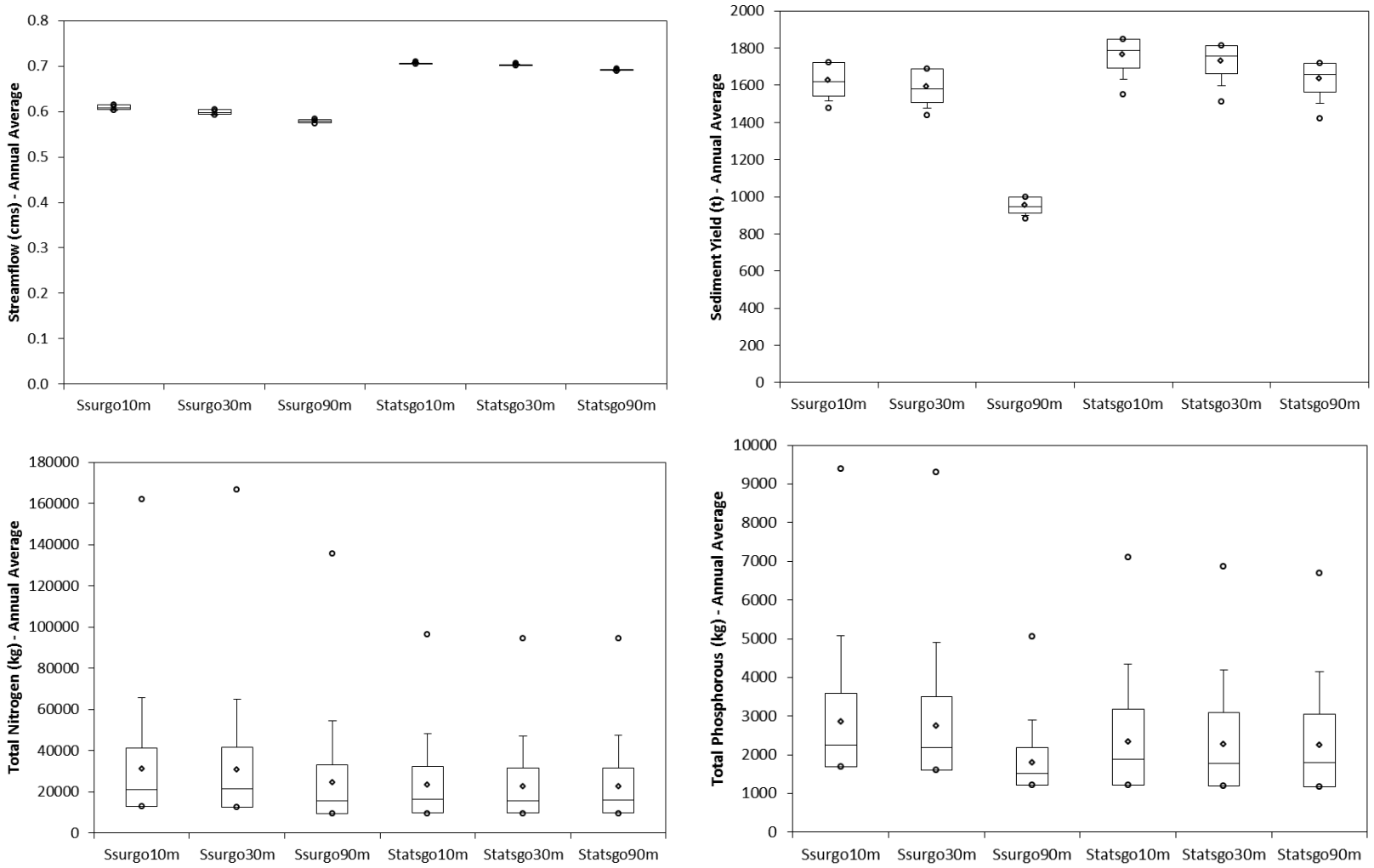


Figure 25. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to organic carbon content (SOL_CBN) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

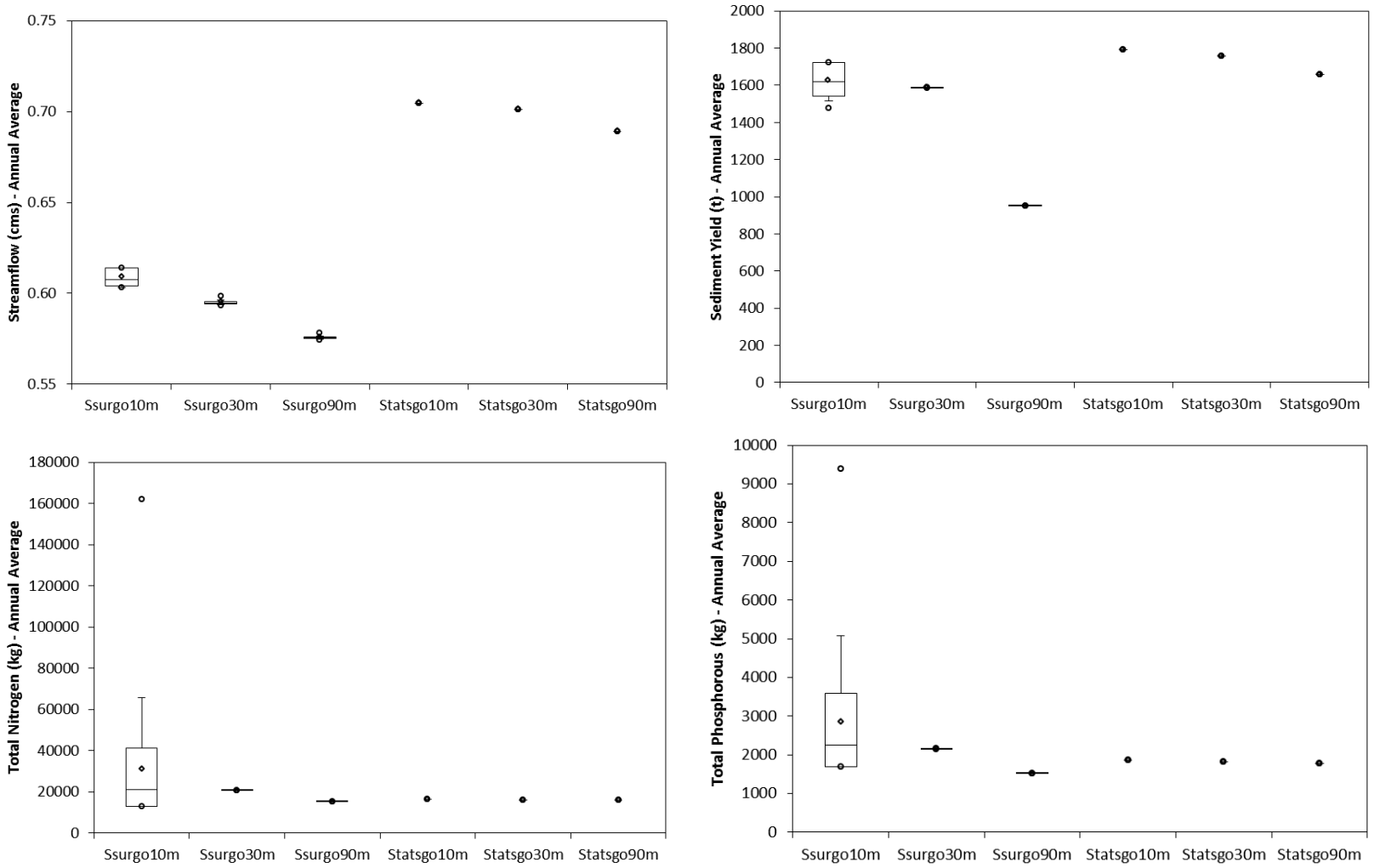


Figure 26. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to soil albedo (SOL_ALB) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

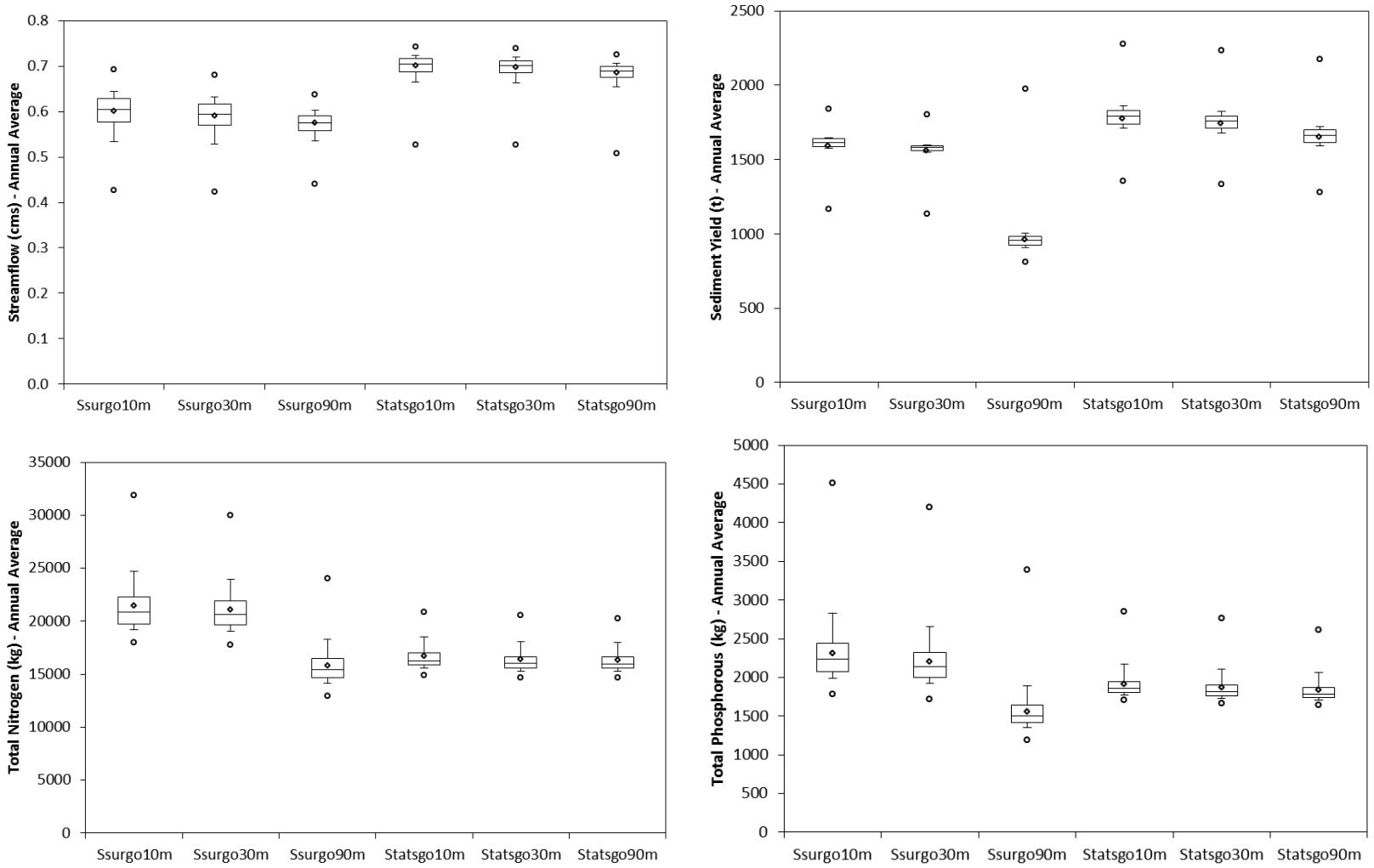


Figure 27. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to soil saturated hydraulic conductivity (SOL_K) for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

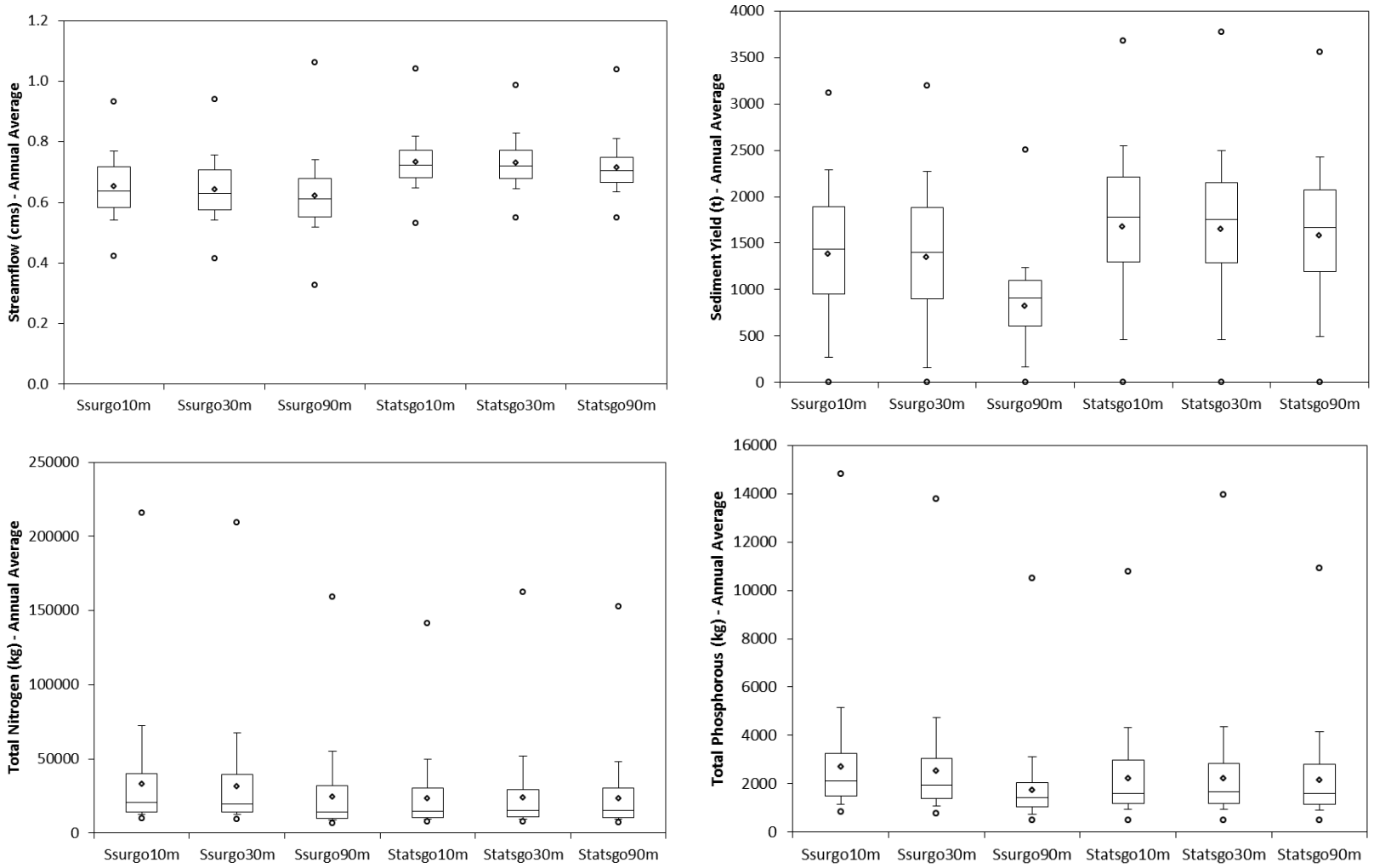


Figure 28. Modified Tukey box plots representing the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) on the uncertainty of SWAT outputs due to all soils factors for Reynolds Creek Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

Table 18. F-Test results to compare the effect of DEM resolution, soils database and the combined effect on SWAT average annual outputs due to soil parameters for Little River Experimental Watershed. Specific tests are Soils DB: average annual output (SSURGO) = average annual output (STATSGO); DEM: average annual output (10 m) = average annual output (30 m) = average annual output (90 m); and Soils*DB: average annual output (SSURGO x 10 m) = average annual output (SSURGO x 30 m) = average annual output (SSURGO x 90 m) = average annual output (STATSGO x 10 m) = average annual output (STATSGO x 30 m) = average annual output (STATSGO x 90 m).

	Streamflow			Sediment Yield			Total Nitrogen			Total Phosphorous		
	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM
SOL_ALB	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_AWC	< 0.0001**	0.1973^{NS}	0.7483^{NS}	< 0.0001**	< 0.0001**	0.2696^{NS}	< 0.0001**	0.0028**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_BD	< 0.0001**	< 0.0001**	0.0003**	< 0.0001**	< 0.0001**	0.2414^{NS}	< 0.0001**	0.4555^{NS}	0.5201^{NS}	< 0.0001**	< 0.0001**	0.1040^{NS}
SOL_CBN	< 0.0001**	< 0.0001**	0.0183*	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	0.2998^{NS}	< 0.0001**	< 0.0001**	< 0.0001**	0.3414^{NS}
SOL_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	0.5102^{NS}	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
USLE_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	0.0460^{NS}	< 0.0001**	0.0091**	< 0.0001**	< 0.0001**	< 0.0001**	0.5894^{NS}
SOL_ALL	< 0.0001**	0.0045**	0.5739^{NS}	< 0.0001**	< 0.0001**	0.8649^{NS}	< 0.0001**	0.6739^{NS}	0.7425^{NS}	< 0.0001**	< 0.0001**	0.6787^{NS}

*	$p \leq 0.05$
**	$p \leq 0.01$
NS	Not Significant
SOL_ALB	Soil albedo
SOL_AWC	Available water capacity
SOL_BD	Bulk density
SOL_CBN	Organic carbon content
SOL_K	Saturated hydraulic conductivity
USLE_K USLE	USLE soil erodibility
SOL_ALL	All of the above

Table 19. F-Test results to compare the effect of DEM resolution, soils database and the combined effect of SWAT average annual outputs due to soil parameters for Reynolds Creek Experimental Watershed. Specific tests are Soils DB: average annual output (SSURGO) = average annual output (STATSGO); DEM: average annual output (10 m) = average annual output (30 m) = average annual output (90 m); and Soils*DB: average annual output (SSURGO x 10 m) = average annual output (SSURGO x 30 m) = average annual output (SSURGO x 90 m) = average annual output (STATSGO x 10 m) = average annual output (STATSGO x 30 m) = average annual output (STATSGO x 90 m).

	Streamflow			Sediment Yield			Total Nitrogen			Total Phosphorous		
	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM
SOL_ALB	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_AWC	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_BD	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_CBN	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
USLE_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_ALL	< 0.0001**	< 0.0001**	< 0.0140*	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	0.0019**	< 0.0001**	< 0.0001**

* $p \leq 0.05$
** $p \leq 0.01$
SOL_ALB Soil albedo
SOL_AWC Available water capacity
SOL_BD Bulk density
SOL_CBN Organic carbon content
SOL_K Saturated hydraulic conductivity
USLE_K USLE soil erodibility
SOL_ALL All of the above

Table 20. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for available water capacity (SOL_AWC), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-0.31	0.24	16.11	16.32	17.15	-3.18	-2.43	0.67	1.00	1.48	3.69	14.51	127.55	127.69	129.62	2.65	38.62	159.42	158.05	198.20
Ssurgo 30 m DEM	-	0.55	16.47	16.68	17.51	-	0.77	3.97	4.32	4.81	-	10.44	119.45	119.58	121.44	-	35.05	152.73	151.39	190.51
Ssurgo 90 m DEM		-	15.83	16.04	16.87		-	3.18	3.52	4.01		-	98.71	98.83	100.51		-	87.15	86.16	115.12
Statsgo 10 m DEM			-	0.18	0.89			-	0.34	0.81			-	0.06	0.91			-	-0.53	14.95
Statsgo 30 m DEM				-	0.71				-	0.47				-	0.85				-	15.56

Table 21. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for to soil bulk density (SOL_BD), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	9.05	17.73	90.14	94.16	98.19	0.60	0.88	4.10	9.89	13.02	2.49	6.25	23.77	25.30	30.54	-2.44	-4.94	46.61	47.13	52.97
Ssurgo 30 m DEM	-	7.95	74.35	78.04	81.73	-	0.27	3.47	9.23	12.35	-	3.67	20.77	22.26	27.37	-	-2.56	50.27	50.81	56.80
Ssurgo 90 m DEM		-	61.51	64.93	68.35		-	3.19	8.93	12.04		-	16.50	17.93	22.86		-	54.23	54.78	60.92
Statsgo 10 m DEM			-	2.12	4.23			-	5.56	8.57			-	1.23	5.47			-	0.36	4.34
Statsgo 30 m DEM				-	2.07				-	2.85				-	4.18				-	3.97

Table 22. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for organic carbon content (SOL_CBN), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	2.93	-1.84	66.82	72.80	68.51	0.66	-7.53	49.83	82.78	73.57	1.17	-4.78	20.18	23.93	28.52	2.60	-1.85	78.22	82.24	76.43
Ssurgo 30 m DEM	-	-4.63	62.07	67.89	63.72	-	-8.13	48.85	81.59	72.43	-	-5.88	18.79	22.50	27.04	-	-4.34	73.70	77.62	71.96
Ssurgo 90 m DEM		-	69.95	76.05	71.67		-	62.02	97.66	87.69		-	26.21	30.16	34.98		-	81.57	85.67	79.75
Statsgo 10 m DEM			-	3.59	1.01			-	21.99	15.84			-	3.13	6.94			-	2.25	-1.00
Statsgo 30 m DEM				-	-2.48				-	-5.04				-	3.70				-	-3.18

Table 23. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for soil albedo (SOL_ALB), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	10.13	15.19	-98.78	-99.76	-98.78	1.02	-12.25	-98.63	-99.25	-98.75	10.31	8.85	-99.07	-99.53	-99.69	0.57	1.40	-7.29	-7.54	-8.98
Ssurgo 30 m DEM	-	4.60	-98.89	-99.78	-98.89	-	-13.13	-98.64	-99.26	-98.77	-	-1.33	-99.16	-99.57	-99.72	-	0.83	-7.81	-8.07	-9.50
Ssurgo 90 m DEM		-	-98.94	-99.79	-98.94		-	-98.44	-99.15	-98.58		-	-99.15	-99.57	-99.71		-	-8.57	-8.82	-10.24
Statsgo 10 m DEM			-	-80.00	0.00			-	-45.45	-9.09			-	-48.97	-66.26			-	-0.28	-1.83
Statsgo 30 m DEM				-	400.00				-	66.67				-	-33.87				-	-1.56

Table 24. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for USLE equation erodibility factor (USLE_K), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-40.0	-20.0	60.00	20.00	80.00	24.24	10.34	-68.46	-60.11	-59.15	0.06	1.61	14.18	12.58	14.07	3.49	4.66	41.81	43.40	48.49
Ssurgo 30 m DEM	-	33.33	166.67	100.00	200.00	-	-11.19	-74.61	-67.89	-67.12	-	1.55	14.11	12.50	14.00	-	1.13	37.02	38.56	43.48
Ssurgo 90 m DEM		-	100.00	50.00	125.00		-	-71.41	-63.85	-62.97		-	12.37	10.79	12.26		-	35.49	37.01	41.87
Statsgo 10 m DEM			-	-25.00	12.50			-	26.46	29.52			-	-1.41	-0.10			-	1.12	4.71
Statsgo 30 m DEM				-	50.00				-	2.42				-	1.33				-	3.55

Table 25. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for saturated hydraulic conductivity (SOL_K), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-8.22	-41.5	-67.7	-98.6	-78.00	-19.1	-25.9	-87.6	25.32	-90.4	-9.03	-23.9	-51.2	39.00	-65.1	-24.88	-61.64	38.82	4523.30	79.73
Ssurgo 30 m DEM	-	-36.2	-64.9	-98.5	-76.03	-	-8.48	-84.7	54.97	-88.1	-	-16.4	-46.37	52.79	-61.68	-	-48.93	84.79	6054.15	139.24
Ssurgo 90 m DEM		-	-44.9	-97.6	-62.38		-	-83.3	69.32	-87.0		-	-35.81	82.88	-54.14		-	261.87	11951.39	368.50
Statsgo 10 m DEM			-	-95.7	-31.72			-	914.8	-22.2			-	184.88	-28.56			-	3230.31	29.47
Statsgo 30 m DEM				-	1518.3				-	-92.3				-	-74.92				-	-96.11

Table 26. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for all soil factors at the same time, Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-6.98	-2.77	11.50	12.30	12.16	10.19	4.38	-17.40	-14.63	-15.19	-2.34	1.89	12.63	8.66	8.04	3.49	4.66	41.81	43.40	48.49
Ssurgo 30 m DEM	-	4.52	19.86	20.73	20.58	-	-5.27	-25.05	-22.53	-23.03	-	4.32	15.32	11.26	10.62	-	1.13	37.02	38.56	43.48
Ssurgo 90 m DEM		-	14.68	15.50	15.36		-	-20.87	-18.22	-18.75		-	10.54	6.65	6.04		-	35.49	37.01	41.87
Statsgo 10 m DEM			-	0.72	0.60			-	3.35	2.68			-	-3.52	-4.08			-	1.12	4.71
Statsgo 30 m DEM				-	-0.12				-	-0.65				-	-0.57				-	3.55

Table 27. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for available water capacity (SOL_AWC), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	3.68	32.34	16.63	18.96	16.26	**465**	**232**	**292**	**290**	**277**	1.50	-33.48	-55.28	-54.70	-56.54	-4.43	-35.03	-72.26	-72.59	-74.08
Ssurgo 30 m DEM	-	27.65	12.49	14.73	12.13	-7.97	-54.01	-42.07	-42.62	-45.17	-	-34.47	-55.94	-55.37	-57.19	-	-32.02	-70.97	-71.32	-72.88
Ssurgo 90 m DEM		-	-11.87	-10.11	-12.16	-	-50.03	-37.05	-37.65	-40.42		-	-32.77	-31.90	-34.67		-	-57.30	-57.81	-60.10
Statsgo 10 m DEM			-	2.00	-0.32		-	25.96	24.78	19.22			-	1.29	-2.82			-	-1.19	-6.55
Statsgo 30 m DEM				-	-2.27			-	-0.94	-5.35				-	-4.06				-	-5.43

Table 28. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for to soil bulk density (SOL_BD), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-3.00	-17.6	-33.36	-35.73	-42.89	-8.38	-72.82	-42.40	-42.31	-44.79	21.85	12.77	189.91	187.54	204.02	-5.99	-35.27	-58.08	-56.30	-54.78
Ssurgo 30 m DEM	-	-15.1	-31.30	-33.75	-41.13	-	-70.33	-37.13	-37.04	-39.74	-	-7.45	137.93	135.99	149.51	-	-31.15	-55.41	-53.52	-51.90
Ssurgo 90 m DEM		-	-19.09	-21.97	-30.66		-	111.91	112.23	103.12		-	157.07	154.97	169.59		-	-35.23	-32.48	-30.13
Statsgo 10 m DEM			-	-3.56	-14.31			-	0.15	-4.15			-	-0.82	4.87			-	4.25	7.88
Statsgo 30 m DEM				-	-11.14				-	-4.29				-	5.73				-	3.48

Table 29. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for USLE equation erodibility factor (USLE_K), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM						0.87	-49.54	-4.86	-5.87	-10.61	-1.13	-42.74	-34.11	-36.29	-36.50	-0.60	-42.55	0.61	-3.53	-3.00
Ssurgo 30 m DEM						-	-49.97	-5.69	-6.69	-11.38	-	-42.09	-33.36	-35.56	-35.78	-	-42.20	1.21	-2.95	-2.42
Ssurgo 90 m DEM							-	88.52	86.52	77.14		-	15.07	11.26	10.90		-	75.12	67.91	68.85
Statsgo 10 m DEM								-	-1.06	-6.04			-	-3.31	-3.63			-	-4.12	-3.58
Statsgo 30 m DEM									-	-5.03				-	-0.33				-	0.56

Table 30. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for organic carbon content (SOL_CBN), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	10.85	-24.0	-73.64	-74.34	-72.87	1.58	-50.65	5.59	4.11	3.46	-1.26	-15.22	-26.95	-29.48	-28.72	-2.59	-49.70	-7.44	-11.16	-11.91
Ssurgo 30 m DEM	-	-31.5	-76.22	-76.85	-75.52	-	-51.42	3.95	2.48	1.85	-	-14.13	-26.02	-28.58	-27.81	-	-48.36	-4.99	-8.80	-9.57
Ssurgo 90 m DEM		-	-65.31	-66.22	-64.29		-	113.95	110.94	109.63		-	-13.84	-16.82	-15.92		-	84.00	76.62	75.13
Statsgo 10 m DEM			-	-2.65	2.94			-	-1.41	-2.02			-	-3.46	-2.42			-	-4.01	-4.82
Statsgo 30 m DEM				-	5.74				-	-0.62				-	1.08				-	-0.84

Table 31. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for soil albedo (SOL_ALB), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-78.2	-84.5	-100.0	-100.0	-100.0	-98.8	-99.7	-100.0	-100.0	-100.0	-99.8	-99.9	-100.0	-100.0	-100.0	-99.65	-99.98	-100.0	-100.0	-100.0
Ssurgo 30 m DEM	-	-28.5	-100.0	-100.0	-100.0	-	-76.2	-100.0	-100.0	-100.0	-	-80.7	-100.0	-100.0	-100.0	-	-93.76	-100.0	-100.0	-100.0
Ssurgo 90 m DEM		-	-100.0	-100.0	-100.0		-	-100.0	-100.0	-100.0		-	-100.0	-100.0	-100.0		-	-100.0	-100.0	-100.0
Statsgo 10 m DEM																				
Statsgo 30 m DEM																				

Table 32. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for saturated hydraulic conductivity (SOL_K), Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-5.55	-38.3	-46.28	-47.56	-52.36	-31.8	29.24	105.08	95.76	72.92	-10.90	-24.49	-46.79	-49.80	-51.54	**733**	**547**	**397**	**380**	**355**
Ssurgo 30 m DEM	-	-34.6	-43.12	-44.48	-49.56	-	89.74	201.09	187.41	153.88	-	-15.26	-40.28	-43.66	-45.62	-12.4	-34.55	-52.59	-54.61	-57.53
Ssurgo 90 m DEM		-	-12.93	-15.01	-22.78		-	58.68	51.47	33.80		-	-29.53	-33.52	-35.82	-	-25.29	-45.88	-48.19	-51.52
Statsgo 10 m DEM			-	-2.39	-11.32			-	-4.55	-15.68			-	-5.65	-8.93		-	-27.57	-30.65	-35.12
Statsgo 30 m DEM				-	-9.14				-	-11.67				-	-3.47			-	-4.26	-10.42

Table 33. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (10, 30, and 90 m) and soils database (STASTSGO and SSURGO) for all soil factors at the same time, Reynolds Creek Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM	Ssurgo 30 m DEM	Ssurgo 90 m DEM	Statsgo 10 m DEM	Statsgo 30 m DEM	Statsgo 90 m DEM
Ssurgo 10 m DEM	-5.28	-0.89	-24.28	-18.89	-22.10	5.01	-46.73	3.99	1.39	-4.11	-6.97	-21.60	-32.52	-28.73	-34.99	-8.23	-40.20	-15.85	-15.16	-19.28
Ssurgo 30 m DEM	-	4.63	-20.06	-14.37	-17.76	-	-49.27	-0.97	-3.45	-8.69	-	-15.73	-27.47	-23.40	-30.13	-	-34.83	-8.31	-7.55	-12.04
Ssurgo 90 m DEM		-	-23.60	-18.16	-21.40		-	95.20	90.32	80.00		-	-13.93	-9.10	-17.08		-	40.71	41.86	34.97
Statsgo 10 m DEM			-	7.12	2.88			-	-2.50	-7.79			-	5.61	-3.67			-	0.82	-4.08
Statsgo 30 m DEM				-	-3.96				-	-5.42				-	-8.78				-	-4.86

Results of this part of the study showed that, while the model predictions did not change or decrease as the input data became coarser, the uncertainty increased. Therefore, research and additional effort should be placed on improving available information in data-poor conditions to reduce the uncertainty in model predictions.

The combination of SSURGO and the 30 m DEM proved to adequately balance the level of uncertainty and the quality of the input data sets. Additionally, going from 10 m DEM to 30 m did not increase uncertainty. Consequently, the next part of the research focused on developing alternatives to downscale 90 m DEMs to 30 m and to generate an equivalent soil map to SSURGO. Two reasons support this decision. The SSURGO soils database is available and free of charge for the continental US, Hawaii, Alaska, Puerto Rico and the Virgin Islands (Soil Survey Staff NRCS-USDA, 2011). Equivalent soils information can be found for Canada (National Soil Data Base – NSDB: http://sis.agr.gc.ca/cansis/nsdb/), most of Eastern Europe (European Soil Data Centre – ESDB: http://eusoils.jrc.ec.europa.eu/) and Australia (Australian Soil Resource Information System – ASRIS: http://www.asris.csiro.au/); however, most places in Central and South America, Africa and Asia do not have the same detailed soils information or it is not freely available. In the second place, the effect of using different resolution DEMs on the results from hydrological modeling has been discussed in the literature review of this document. Digital elevation models at 90 m resolution are freely available for most of the world; however, 90 m resolution is considered coarse and not recommended for use in hydrologic analysis. On the other hand, it is reported that there is a threshold beyond which higher resolution of data does not produce better predictions and that the extra cost and labor to obtain DEMs with 10 m resolution is not justified to obtain more accurate prediction (Mamillapalli et al., 1996). Because of that, a high resolution source, such as 10 m, may not be the best alternative for data-poor environments. In addition, this study showed that there was not an increase in uncertainty going from 10 m DEM to 30 m. Therefore, a 30 m grid would give a better representation of the surface in data-poor conditions than 10 and 90 m DEMs.

4.2. Objective 2 - Develop methods to gather qualitative and quantitative data that will allow modeling agricultural watershed systems under data-poor environments

4.2.1. DEM comparison

A comparison between USGS DEM 1 arc second data (30 m) and the 30 m DEM from CGIAR-CSI resampled from the 3-arcsecond DEM (approximately 90 m) shows differences for

different parameters (table 34). Minimum elevation is slightly different, but the maximum elevation has a 9-m difference, with the resampled DEM including the highest elevation (157.0 m). The average value is also different, with the resampled DEM again having the higher average (4.2 m higher). Standard deviations are almost the same, with the USGS DEM having the slightly higher value in this case. Areas are different in magnitude; the area derived with the CGIAR DEM is 11.4 km^2 larger, mainly because additional areas were added in the lower left and upper right portions of the watershed (fig. 29). However, smaller differences can be seen all along the watershed boundary. Minor differences can be seen along the rest of the watershed boundary. Additionally, differences in the texture of both models can be seen. Figure 30a shows that the DEM downloaded from USGS has a smooth appearance, while the CGIAR DEM (fig. 30b) looks rough with a different texture.

Table 34. DEM resolution comparison (USGS 30 m and CGIAR 30 m resampled) for Little River Experimental Watershed.

Parameter	USGS 30 m DEM	CGIAR 30 m DEM
Minimum (m)	81.9	82.0
Maximum (m)	148.0	157.0
Mean (m)	109.9	114.1
Standard Deviation	12.8	12.5
Area (km^2)	332.5	343.9

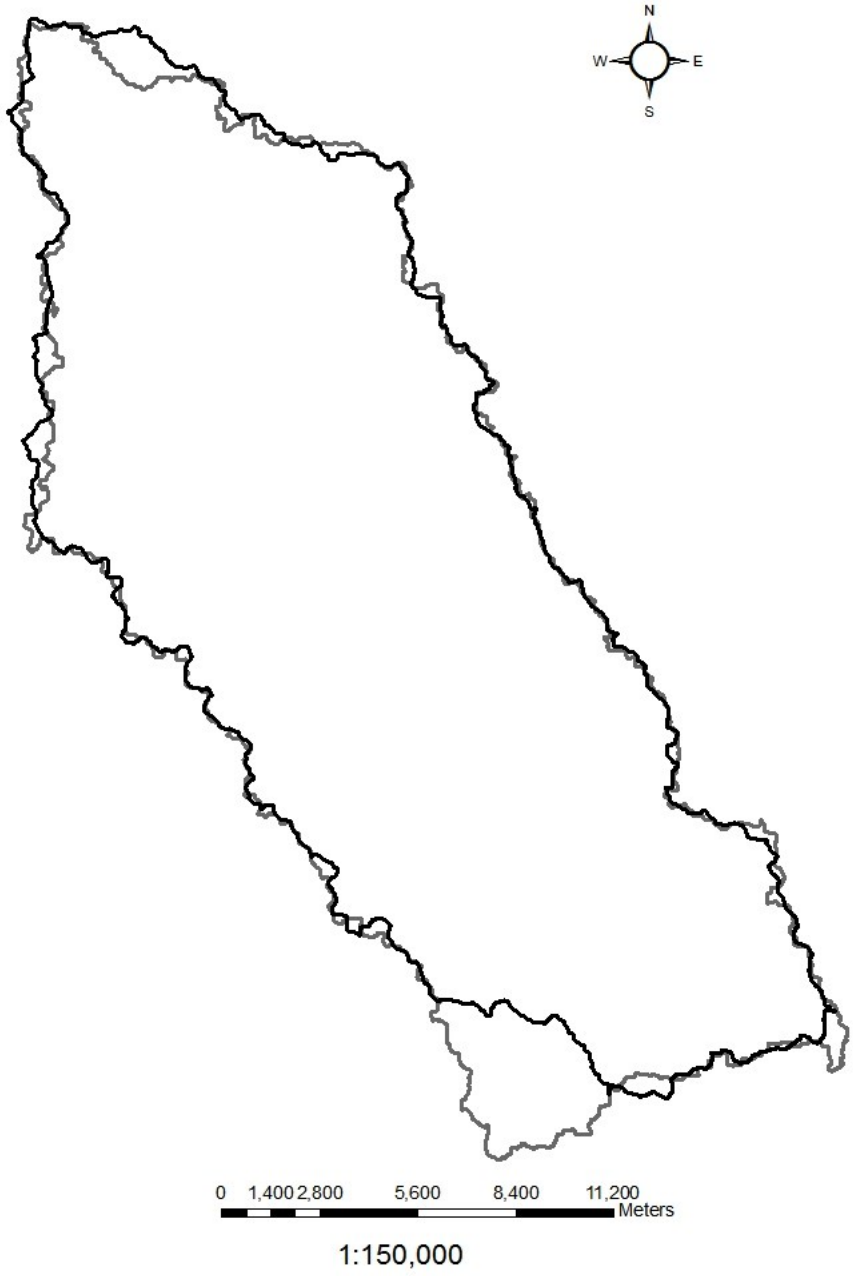
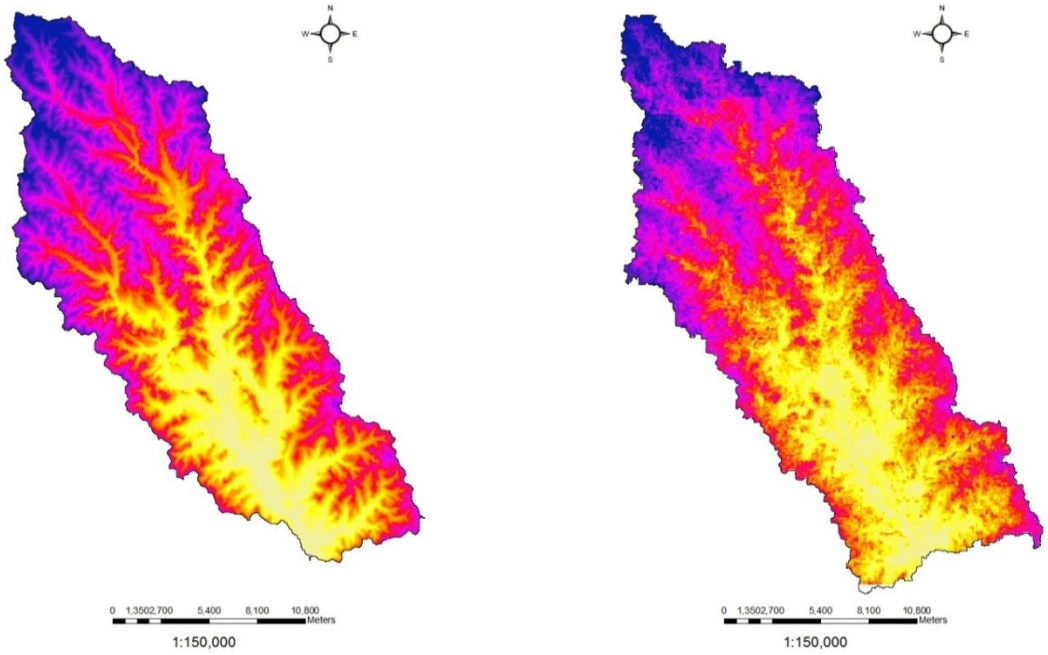


Figure 29. Watershed boundaries generated from two different digital elevation models for Little River Experimental Watershed. Black solid line is the USGS DEM 1 arc second data (30 m) and grey solid line the 30 m CGIAR-CSI DEM resampled from a 3-arcsecond DEM (approximately 90 m).



(a) (b)

Figure 30. Digital elevation models for Little River Experimental Watershed: (a) the USGS DEM 1 arc second data (30 m) (b) the 30 m CGIAR-CSI DEM resampled from a 3-arcsecond DEM (approximately 90 m).

The differences in the DEM parameters made an impact on several hydrologic properties of Little River watershed. The number of sub watersheds derived by SWAT's algorithm (figure 31) was different; the USGS 30 m DEM had 27 sub watersheds, while the resampled CGIAR 30 m DEM had 35 sub watersheds.

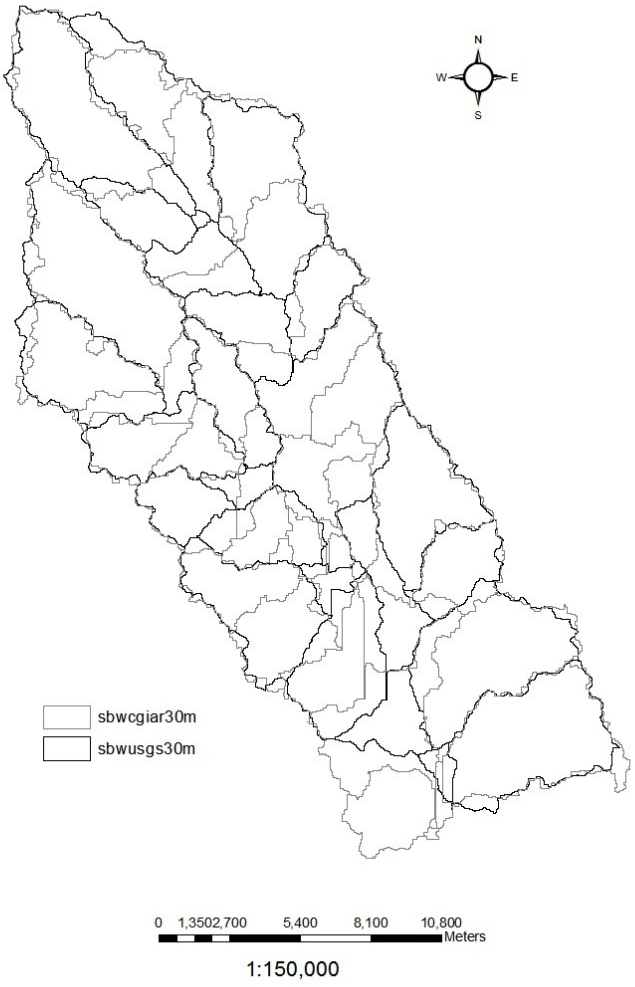


Figure 31. Sub watershed delineation for Little River Experimental Watershed. USGS DEM 1 arc second data (30 m) in black bold line and the 30 m CGIAR-CSI DEM resampled in gray dotted line.

The differences in sub watershed definition are caused by differences in the stream network delineation (fig. 32). To match the USGS watershed boundary, the outlet at the CGIAR watershed had to be displaced a few cells down and towards the left, which generated an additional small stream right above the outlet that generates the additional area in the watershed. Otherwise, when the outlet was set at the same point, the boundary generated with the CGIAR's DEM ignored a large portion of the bottom right side of the watershed. Generally, the CGIAR derived stream network is slightly displaced; the nodes where two streams become one are not the same in both networks. The lower part of the network shows the biggest differences. The displacement is approximately 1 km in the worst case.

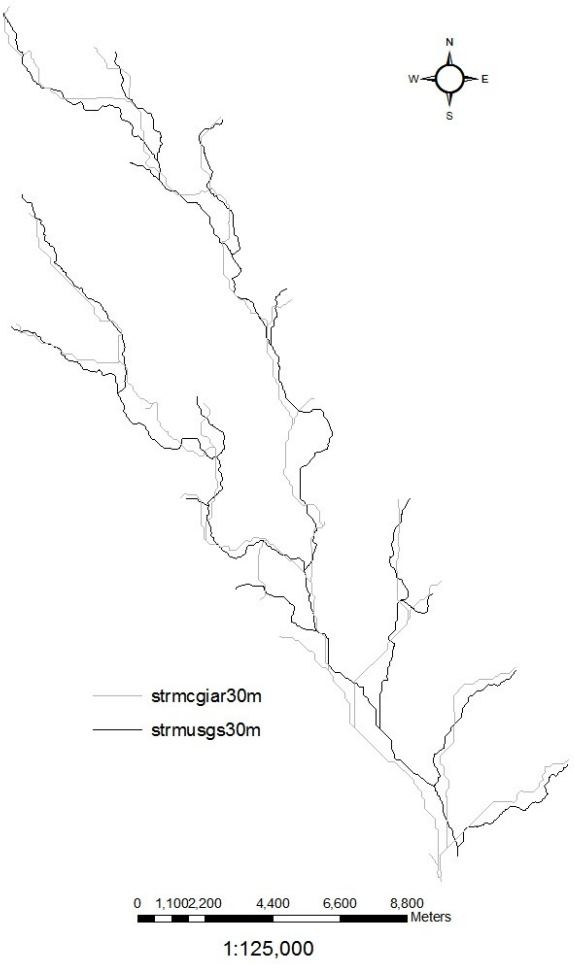


Figure 32. Stream network delineation for Little River Experimental Watershed. USGS DEM 1 arc second data (30 m) in black bold line and the 30 m CGIAR-CSI DEM resampled in gray dotted line.

4.2.2. Soils map comparison

The soils maps for the Little River watershed are shown in figure 33. All three maps show a zone in the mid-section of the watershed, similar to the brown area in the STATSGO map, that falls mostly along the stream channel network. Significant differences can be seen in the number of soil types for each map. With STATSGO, the watershed area is divided into only five soil units, with one soil type covering 77.8 % of the total watershed. With SSURGO, the watershed is divided into 93 different soil units, and the largest area covered by just one soil unit is 21.2 % of the watershed area. The soil map generated by overlaying lithology and slope class divides the area into 18 soil units; the most dominant soil type covers 55.1 % of the watershed area.

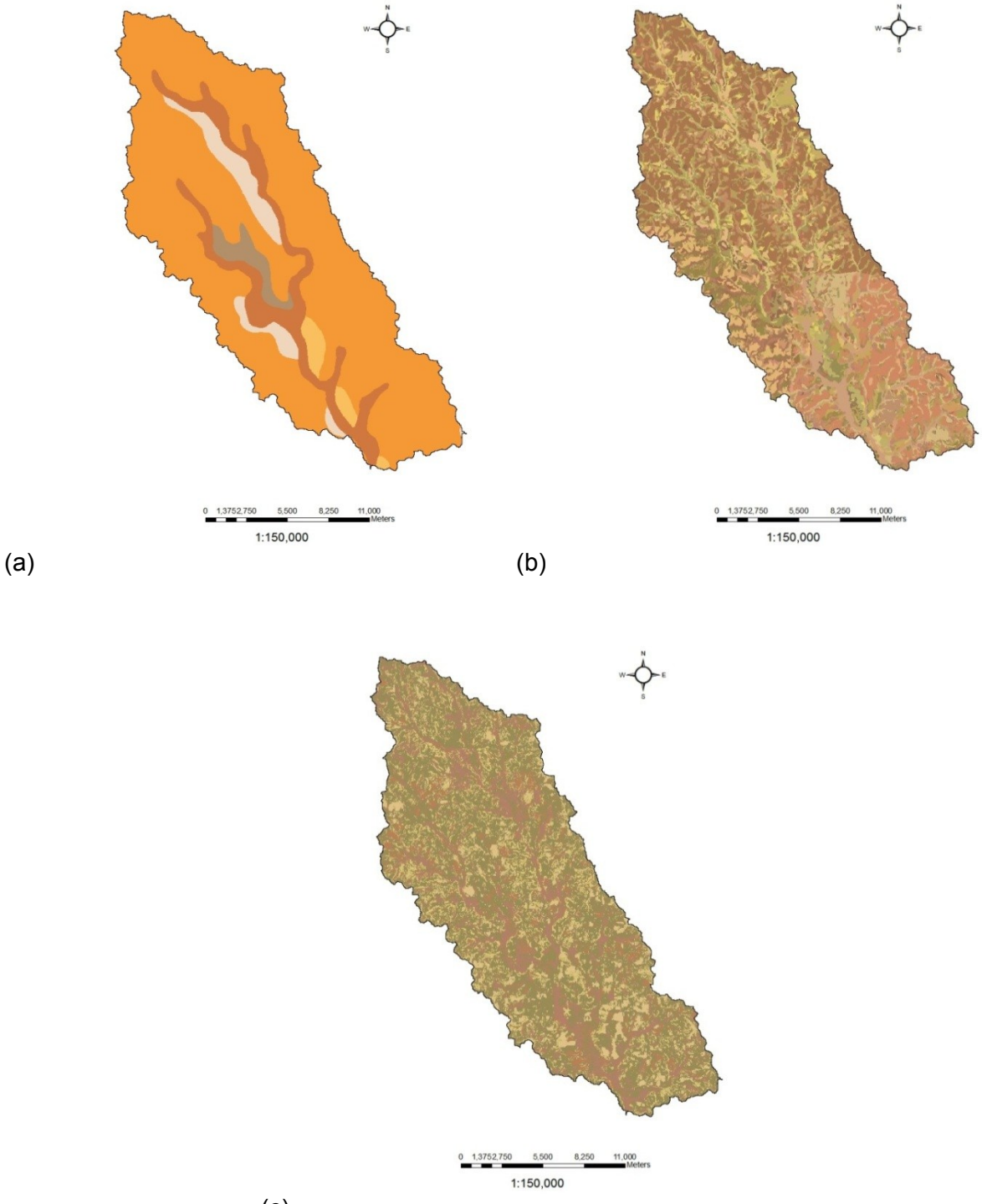


Figure 33. Soils map for Little River Experimental Watershed: (a) STATSGO, (b) SSURGO and (c) derived from overlaying lithology and slope classes.

4.2.3. Uncertainty analysis

This section discusses the results of assessing uncertainty of SWAT outputs due to soil parameters accounting for the variation in DEM source and also the soil data base. Table 35 shows F-Test results to compare the effect of DEM resolution, soils database and their combined effect on SWAT outputs when the selected soil parameters were changed. The individual effect of soils database source, individual effect of DEM resolution and the interaction of soils database and DEM caused significant differences ($p < 0.001$) in the predictions of streamflow, sediment yield, total nitrogen and total phosphorous for all soils parameters varied individually and all at once. The effects of uncertainty analysis of soil parameters (available water capacity, soil bulk density, USLE equation erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo) on predicted mean annual streamflow, sediment yield, total nitrogen and total phosphorous for the LREW (1987-1998) are shown in figures 34 to 39. Figure 40 shows effects of uncertainty analysis when all soil parameters vary at the same time. Each box plot set illustrates the differences due to changes in DEM resolution (USGS DEM and CGIAR-CSI DEM) and to the source of the soils database (SSURGO, STATSGO and SATSLP).

For all the individual soil parameters (available water capacity, soil bulk density, USLE equation erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo) and also all the parameters at the same time, the combinations of SATSLP and resampled DEM showed a narrower 10^{th} - 90^{th} percentile range than the combination of SSURGO and 30 m USGS-DEM (tables 36 – 42). In three cases, the 10^{th} - 90^{th} percentile range was narrower for the SSURGO and 30 m USGS-DEM: streamflow and sediment yield for bulk density and streamflow for USLE soil erodibility (tables 40 and 41). With the exception of available water capacity, all the other individual soil parameters (soil bulk density, USLE equation erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo) and also when all the parameters were varied at the same time, the streamflow and sediment yield predictions based on the derived soils map (mean and median = ~6.1 m^3s^{-1}) were comparable to the predictions of the combination of SSURGO-30 USGS DEM (mean and median = ~6.2 m^3s^{-1}). In addition, regardless of the soil parameter, the predictions made with the derived soils map and the resampled DEM always had the lowest mean for total nitrogen and phosphorous compared to the combinations of SSURGO and STATSGO with the 30 m USGS-DEM.

For streamflow, the combination of STATSGO and the resampled 30 m DEM produced the lowest mean and median values for all the individual soil parameters (~4.0 m^3s^{-1}; tables 36 to 42). For all of the individual soil parameters except available water content, the mean and median for the combinations of SSURGO and SATSLP with the USGS DEM were in the same range of magnitude (~6.2 m^3s^{-1}; fig.34 - 39). The combination of SSURGO and SATSLP with the resampled DEM (fig.34 - 39) had the same behavior between each other, with mean and median values of ~4.4 m^3s^{-1}. Differences in mean and median values for sediment yield and total nutrients (nitrogen and phosphorous) were more obvious than for streamflow; however, as with streamflow, the combination of STATSGO and resampled 30 m DEM provided the lowest values. The behavior of the combinations SSURGO and SATSLP with either DEM (USGS or Resampled CGIAR-CSI) showed the same tendency as that observed with streamflow, but the values for SSURGO and SATSLP were lower than when the resampled DEM was used instead of the USGS DEM.

When all of the soil parameters were varied at the same time, combinations that included the 30 m USGS DEM produced the highest values for median and mean for all four outputs. While the 10^{th} - 90^{th} percentile ranges were the same order of magnitude, the ranges for the SATSLP combination were narrower than for the SSURGO or STATSGO combinations for all outputs. Also from figure 40, regardless of the soils database source, the statistical parameters were lower whenever the resampled DEM was involved. Additionally, for streamflow and sediment yield, the combination of STATSGO and the resampled DEM produced the lowest statistical values in comparison to any other combination of soils database and DEM. For total nitrogen, the combination SATSLP and the resampled DEM produced the lowest predictions. For total phosphorous, SATSLP and the resampled DEM produced the lowest predictions only for available water content, USLE soil erodibility and organic carbon content.

Based on the results of this part of the study, it can be concluded that the coarser the DEM is, the lower the predictions are for streamflow, sediment yield, and both nutrients. Additionally, the generated soils map (SATSLP) provided the same responses as when SSURGO was involved. SATSLP statistical parameters were similar to the ones obtained with SSURGO for both DEM sources for all constituents. The fact that the SATLSP provided consistently similar results as SSURGO is considered as a promising result towards the use of SWAT in data-poor environments, especially due to the fact that the SSURGO database was considered as the highest quality source for soils.

Table 35. F-Test results to compare the effect of DEM resolution, soils database and the combined effect of SWAT average annual outputs due to soil parameters for Little River Experimental Watershed. Specific tests are Soils DB: average annual output (SSURGO) = average annual output (STATSGO) = average annual output (SATSLP); DEM: average annual output (30 m USGS) = average annual output (30 m CGIAR-SCI); and Soils*DB: average annual output (SSURGO x 30 m USGS) = average annual output (SSURGO x 30 m CGIAR-CSI) = average annual output (STATSGO x 30 m USGS) = average annual output (STATSGO x 30 m CGIAR-CSI) = average annual output (SATSLP x 30 m USGS) = average annual output (SATSLP x 30 m CGIAR-CSI).

	Streamflow			Sediment Yield			Total Nitrogen			Total Phosphorous		
	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM	Soils DB	DEM	Soils*DEM
SOL_ALB	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_AWC	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	0.0028**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_BD	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_CBN	< 0.0001**	< 0.0001**	< 0.0001***	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
USLE_K	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**
SOL_ALL	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**	< 0.0001**

**	$p \leq 0.01$
SOL_ALB	Soil albedo
SOL_AWC	Available water capacity
SOL_BD	Bulk density
SOL_CBN	Organic carbon content
SOL_K	Saturated hydraulic conductivity
USLE_K USLE	USLE soil erodibility
SOL_ALL	All of the above

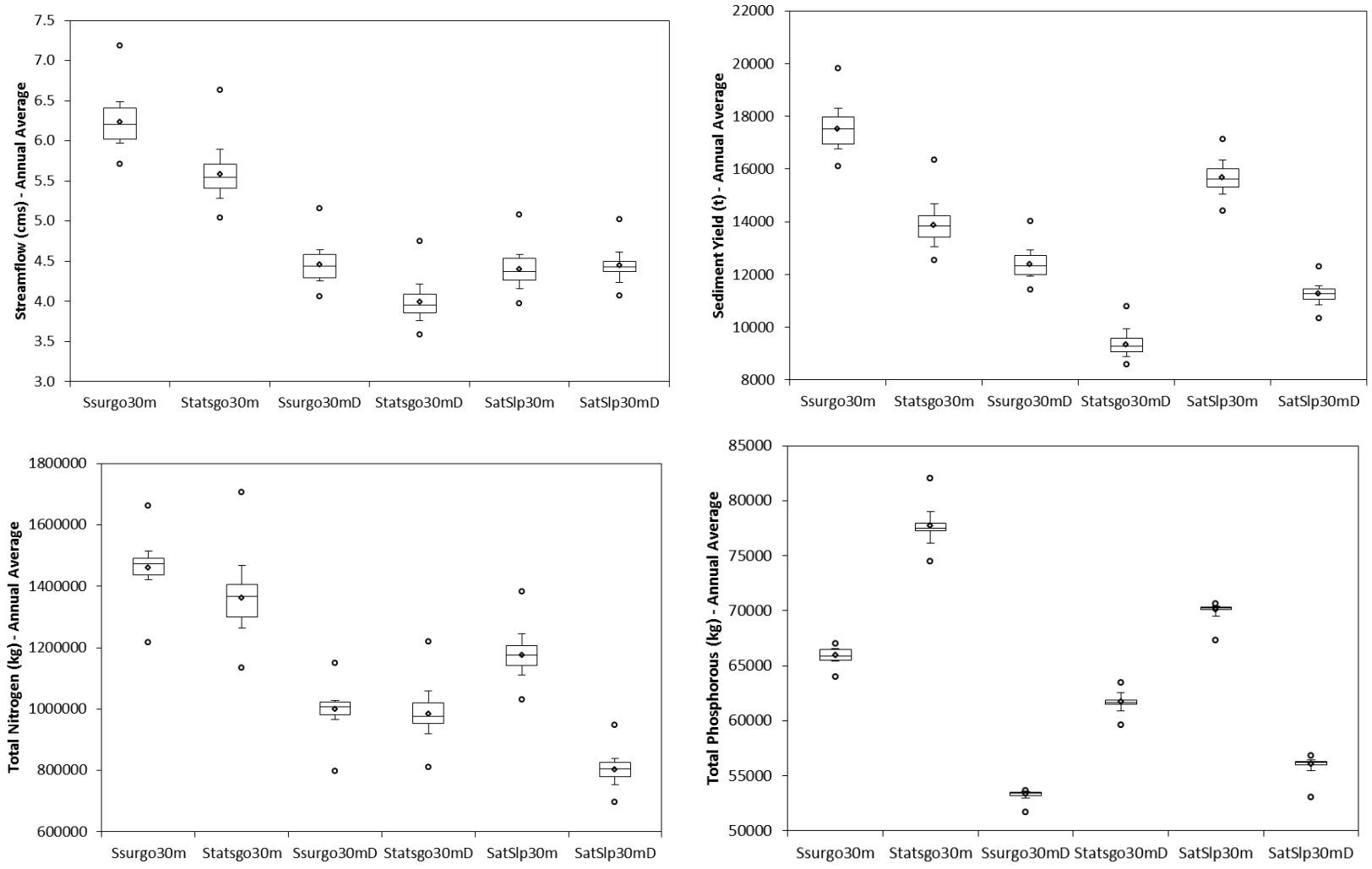


Figure 34. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to available water capacity (SOL_AWC) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

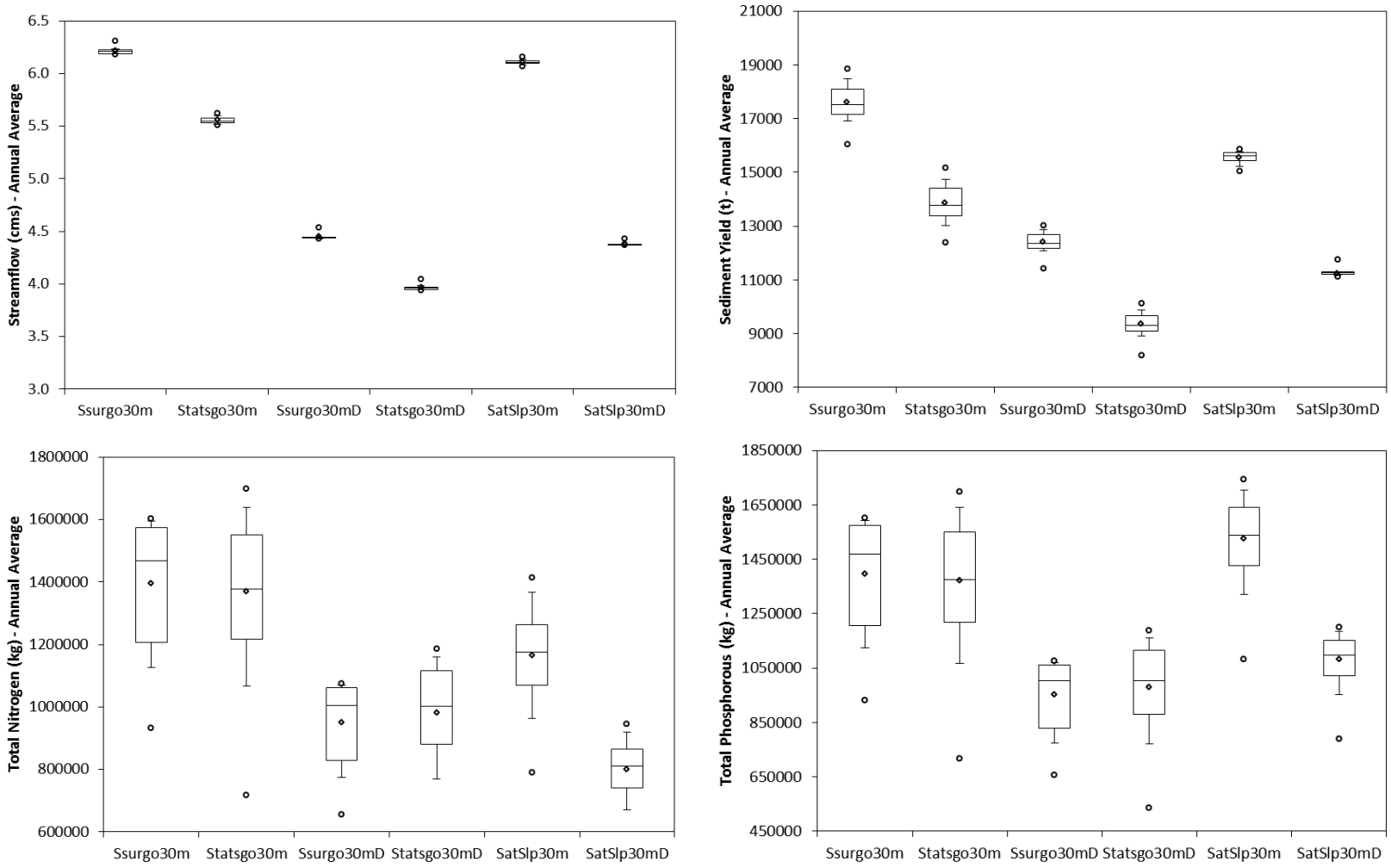


Figure 35. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to soil bulk density (SOL_BD) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

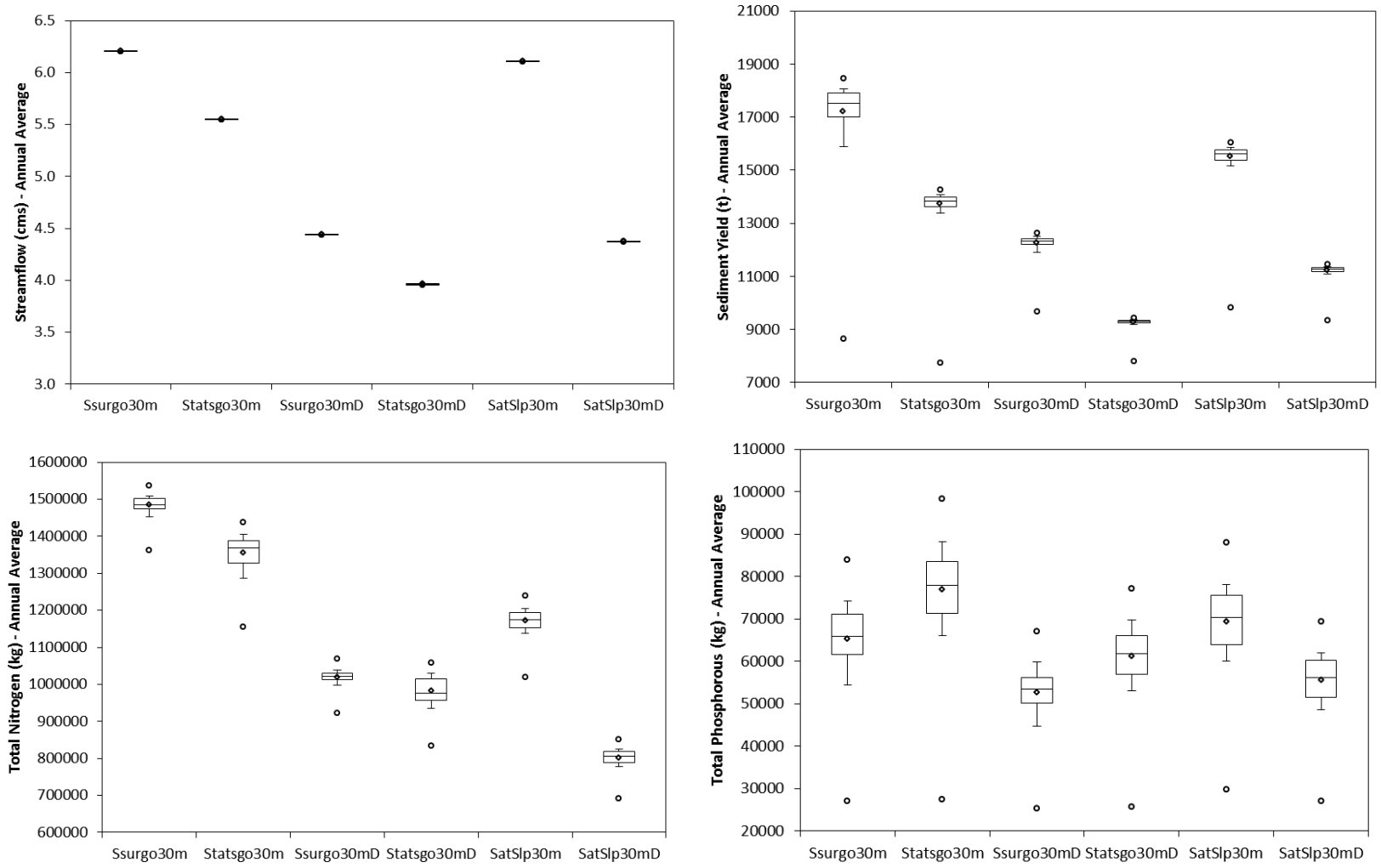


Figure 36. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to USLE equation erodibility factor (USLE_K) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

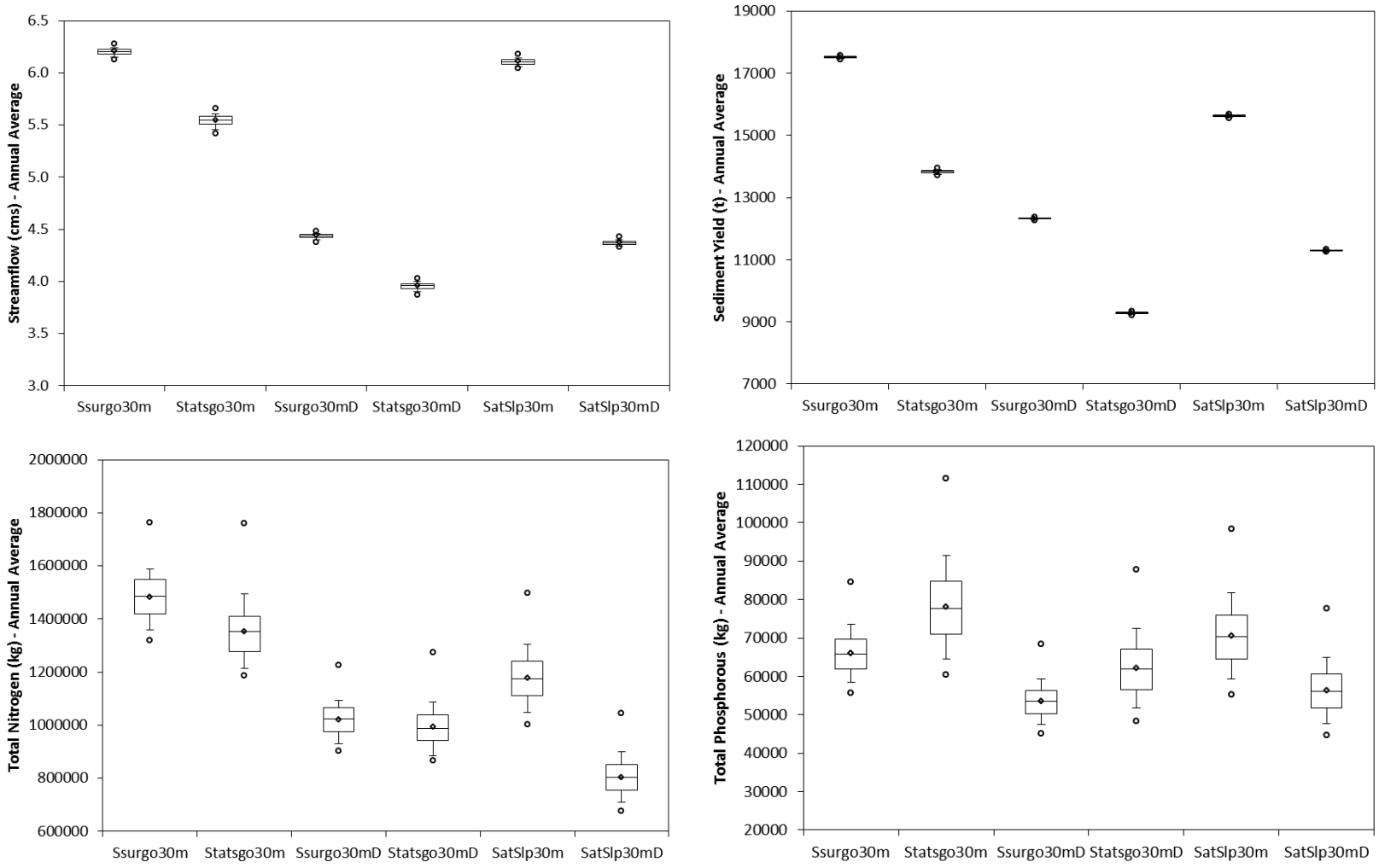


Figure 37. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to organic carbon content (SOL_CBN) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

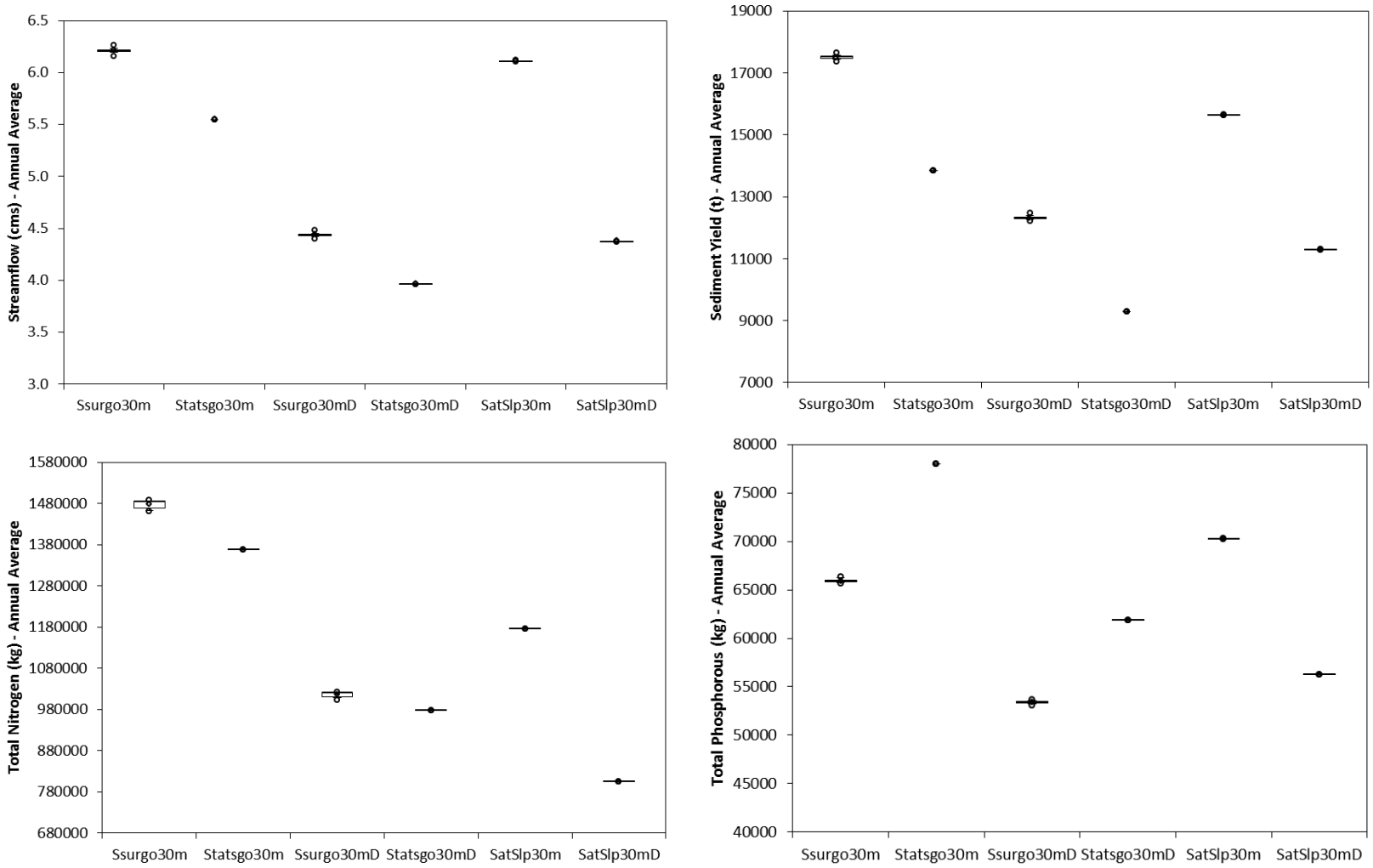


Figure 38. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to soil albedo (SOL_ALB) for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

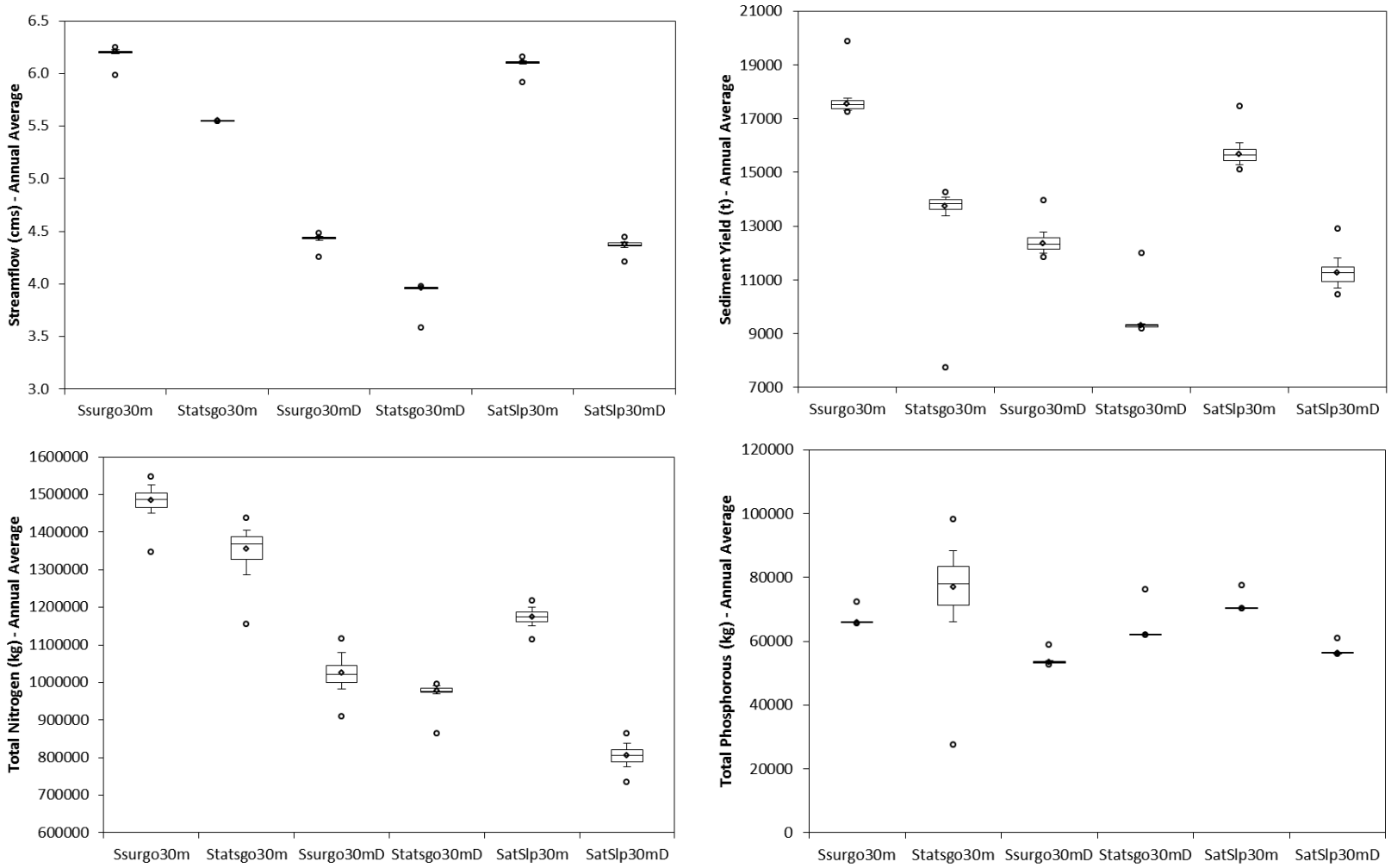


Figure 39. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to saturated hydraulic conductivity (SOL_K) for Little Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

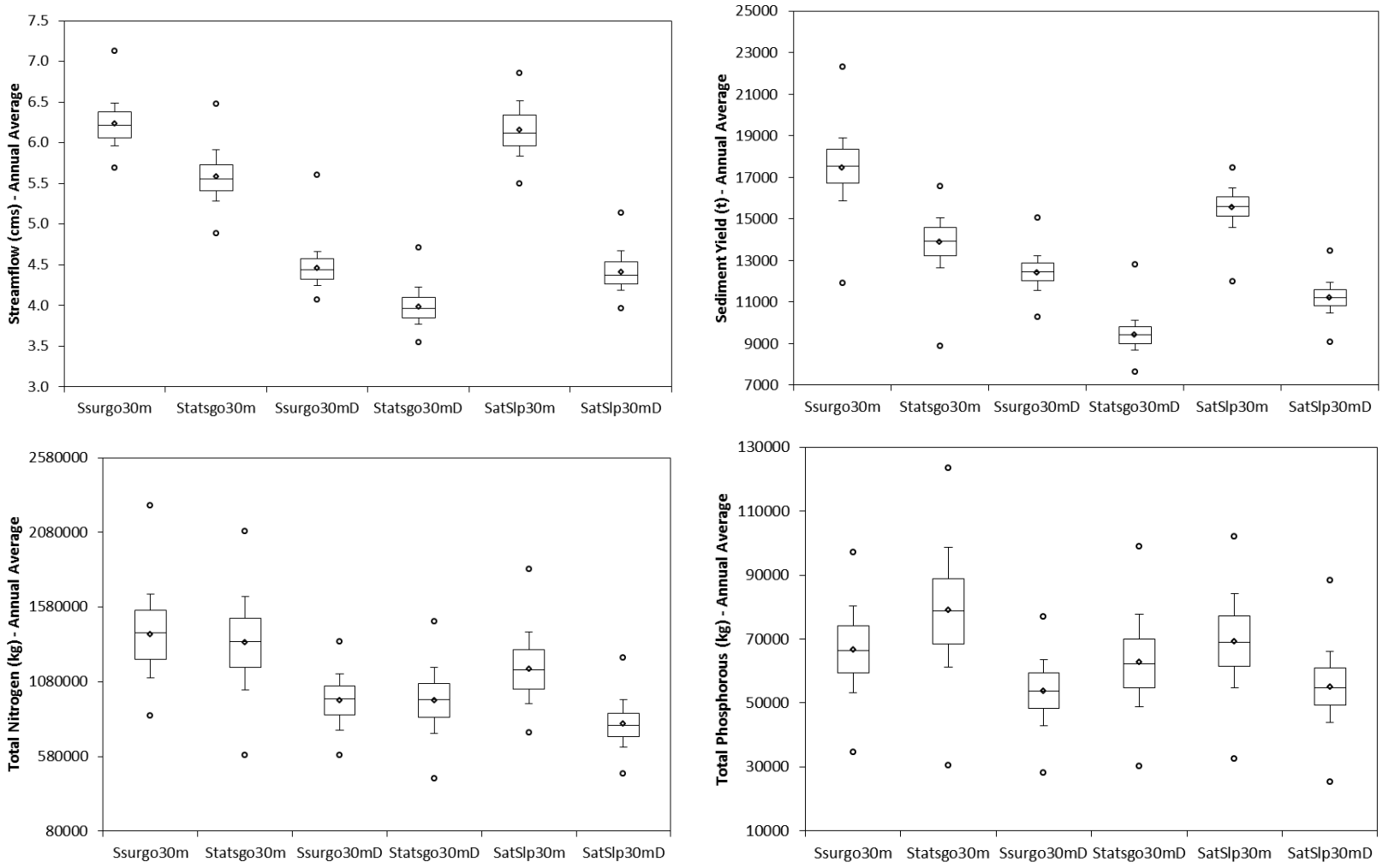


Figure 40. Modified Tukey box plots representing the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) on the uncertainty of SWAT outputs due to all soils factors for Little River Experimental Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

Table 36. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for available water capacity (SOL_AWC), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	20.73	-20.8	-12.93	29.95	-8.07	-22.53	-45.74	-52.73	-37.65	-52.70	11.26	-32.55	-21.10	-13.76	-43.56	-7.56	-40.39	-34.34	-18.75	-45.05
Statsgo 30 m DEM	-	-34.4	-27.88	7.64	-23.8	-	-29.96	-38.98	-19.52	-38.94	-	-39.38	-29.09	-22.49	-49.27	-	-35.52	-28.97	-12.11	-40.56
Ssurgo 90 m D DEM		-	10.06	64.27	16.21		-	-12.88	14.90	-12.82		-	16.98	27.85	-16.31		-	10.15	36.29	-7.82
Statsgo 10 m D DEM			-	49.25	5.59			-	31.89	0.06			-	9.30	-28.46			-	23.73	-16.32
SatSlp 30 m DEM				-	-29.2				-	-24.13				-	-34.55				-	-32.37

Table 37. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for to soil bulk density (SOL_BD), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	78.04	-65.8	-15.87	12.03	-74.8	9.23	-49.80	-37.70	-64.37	-92.68	22.26	-36.89	-16.86	-13.68	-46.91	5.97	-32.92	-25.97	-11.75	-40.97
Statsgo 30 m DEM	-	-80.8	-52.75	-37.08	-85.9	-	-54.04	-42.97	-67.38	-93.29	-	-48.38	-32.00	-29.39	-56.58	-	-36.70	-30.14	-16.72	-44.30
Ssurgo 90 m D DEM		-	146.49	228.22	-26.4		-	24.09	-29.04	-85.41		-	31.73	36.77	-15.89		-	10.36	31.56	-12.00
Statsgo 10 m D DEM			-	33.16	-70.1			-	-42.81	-88.24			-	3.83	-36.15			-	19.21	-20.27
SatSlp 30 m DEM				-	-77.5				-	-79.44				-	-38.50				-	-33.11

Table 38. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for organic carbon content (SOL_CBN), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	67.89	-28.2	7.04	-9.76	-39.1	81.59	-32.60	-13.64	-8.38	-41.87	22.50	-28.84	-11.27	12.35	-16.70	-0.17	-30.44	-27.74	-15.10	-40.83
Statsgo 30 m DEM	-	-57.2	-36.24	-46.2	-63.7	-	-62.88	-52.44	-49.55	-67.99	-	-41.91	-27.57	-8.29	-32.00	-	-30.32	-27.61	-14.95	-40.72
Ssurgo 90 m D DEM		-	49.20	25.78	-15.1		-	28.13	35.93	-13.75		-	24.69	57.88	17.06		-	3.89	22.06	-14.93
Statsgo 10 m D DEM			-	-15.7	-43.1			-	6.08	-32.69			-	26.61	-6.12			-	17.49	-18.11
SatSlp 30 m DEM				-	-32.5				-	-36.55				-	-25.86				-	-30.30

Table 39. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for soil albedo (SOL_ALB), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	-99.7	-13.24	-99.7	-82.76	-91.82	-99.2	-18.21	-99.9	-90.69	-92.54	-99.5	-47.46	-99.6	-98.60	-99.36	-8.07	-31.2	-34.3	-20.9	-45.8
Statsgo 30 m DEM	-	39150	0.00	7700	3600.0	-	10955	-91.67	1158.33	908.33	-	12178	-23.39	227.4	49.19	-	-25.2	-28.5	-14.0	-41.1
Ssurgo 90 m D DEM		-	-99.7	-80.13	-90.57		-	-99.9	-88.62	-90.88		-	-99.3	-97.33	-98.7		-	-4.47	14.95	-21.2
Statsgo 10 m D DEM			-	7700.0	3600.0			-	15000.	12000.			-	327.3	94.74			-	20.33	-17.5
SatSlp 30 m DEM				-	-52.56				-	-19.87				-	-54.4				-	-31.5

Table 40. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for USLE equation erodibility factor (USLE_K), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	100.0	-33.3	400.0	500.0	433.3	-67.89	-72.50	-91.85	-68.25	-87.05	110.78	-24.65	69.18	18.57	-13.26	-6.41	-30.45	-31.18	-19.35	-44.59
Statsgo 30 m DEM	-	-66.6	150.0	200.0	166.6	-	-14.35	-74.61	-1.12	-59.66	-	-64.25	-19.74	-43.75	-58.85	-	-25.69	-26.47	-13.83	-40.79
Ssurgo 90 m D DEM		-	650.0	800.0	700.0		-	-70.35	15.45	-52.89		-	124.52	57.35	15.10		-	-1.05	15.96	-20.33
Statsgo 10 m D DEM			-	20.00	6.67			-	289.44	58.90			-	-29.92	-48.73			-	17.19	-19.48
SatSlp 30 m DEM				-	-11.1				-	-59.20				-	-26.85				-	-31.29

Table 41. Relative differences (%) between 10th – 90th percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for to soil bulk density (SOL_BD), Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	-98.5	17.30	-55.22	-4.42	57.52	54.97	73.7	-72.92	81.92	141.6	52.79	24.66	-71.1	-36.16	-19.36	-7.17	-27.89	-35.77	-21.39	-44.16
Statsgo 30 m DEM	-	7820.0	2923.3	6353.3	10535.0	-	12.1	-82.52	17.39	55.94	-	-18.4	-81.1	-58.22	-47.22	-	-22.32	-30.81	-15.31	-39.85
Ssurgo 90 m D DEM		-	-61.83	-18.52	34.28		-	-84.4	4.70	39.08		-	-76.88	-48.79	-35.31		-	-10.92	9.02	-22.57
Statsgo 10 m D DEM			-	113.4	251.76			-	571.6	792.2			-	121.5	179.8			-	22.39	-13.07
SatSlp 30 m DEM				-	64.80				-	32.8				-	26.31				-	-28.97

Table 42. Relative differences (%) between 10^{th} – 90^{th} percentiles due to the effect of DEM resolution (30 m USGS, and 30 m CGIAR) and soils database (STASTSGO, SSURGO and SATSLP) for all soil factors at the same time, Little River Experimental Watershed.

Baseline Level	Streamflow					Sediment Yield					Total Nitrogen					Total Phosphorous				
	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 mD DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM	Statsgo 30 m DEM	Ssurgo 30 m D DEM	Statsgo 30 m D DEM	StatSlp 30 m DEM	SatSlp 30 m D DEM
Ssurgo 10 m DEM	20.73	-20.89	-12.93	29.95	-8.07	-22.53	-45.74	-52.73	-37.65	-52.70	11.26	-32.55	-21.10	-13.76	-43.56	-7.56	-40.39	-34.34	-18.75	-45.05
Statsgo 30 m DEM	-	-34.4	-27.88	7.64	-23.8	-	-29.96	-38.98	-19.52	-38.94	-	-39.38	-29.09	-22.49	-49.27	-	-35.52	-28.97	-12.11	-40.56
Ssurgo 90 m D DEM		-	10.06	64.27	16.21		-	-12.88	14.90	-12.82		-	16.98	27.85	-16.31		-	10.15	36.29	-7.82
Statsgo 10 m D DEM			-	49.25	5.59			-	31.89	0.06			-	9.30	-28.46			-	23.73	-16.32
SatSlp 30 m DEM				-	-29.2				-	-24.13				-	-34.55				-	-32.37

Results from this part of the study suggest that soil mapping and downscaling of coarser DEM can be done through the use of Landsat and SRTM data as an alternative approach that proved to be simple for watershed modeling purposes in data-poor conditions. The methodologies to overcome insufficient data to characterize the watershed represent a viable option to enable data-intensive models for conditions in which data are scarce, incomplete or nonexistent. The methods presented here are based on freely available data and procedures that are easy to reproduce. However, the methodologies require the use of specialized software to process remotely sensed data and GIS information, which can be costly to acquire in developing countries. There are open source alternatives to ERDAS-Imagine and ArcGIS, such as GRASS (Geographic Resources Analysis Support System). According to the software web page (http://grass.fbk.eu/), GRASS can be used for satellite image processing and geospatial data management and analysis; therefore, it can be applied to replicate the methodologies presented here; however GRASS software procedures and tools were not used in this study.

4.3. Objective 3 - Determine to what extent the proposed methodology to use SWAT with limited data will be able to represent water quality impacts of agricultural watershed systems in data-poor environments.

4.3.1. DEM analysis

The digital elevation model used for the Huanquisco watershed (figure 41) was resampled following the procedure described in the methodology. The original source USGS DEM 3 arc second data (approximately 90 m) was downloaded from the Seamless USGS warehouse and resampled by direct projection using cubic convolution to 30 m resolution. The map kept its characteristics; it does not show a rough texture as the resampled DEM from the CGIAR-CSI. The minimum elevation is 3821 m, the maximum is 4686 m and the average is 4222 m (standard deviation is 197.4 m).

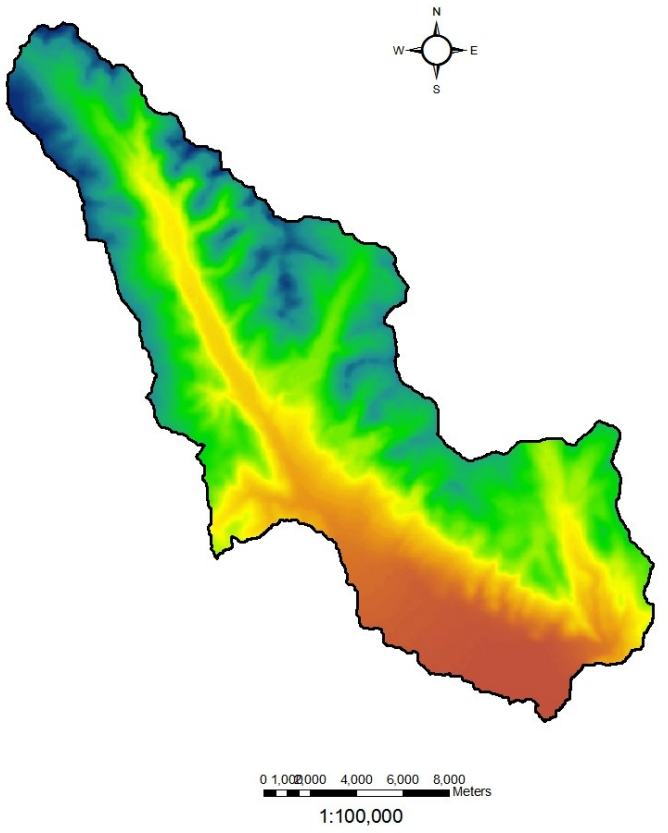
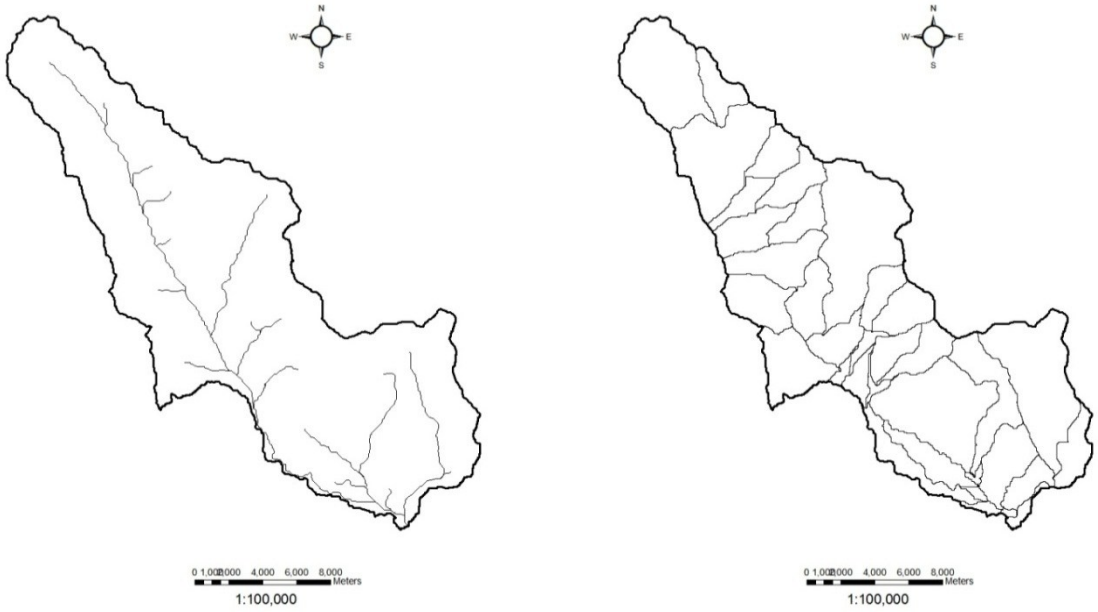


Figure 41. Digital elevation model for the Huanquisco River Watershed (30 m resolution) resampled by direct projection using cubic convolution from USGS DEM 3 arc second.

A total of 39 sub watersheds were defined from the DEM by SWAT (fig. 42). The smallest sub watershed area is 14.8 m^2 and the largest one 1394.9 m^2. The stream network was defined and is composed of a total of 148 stream segments.



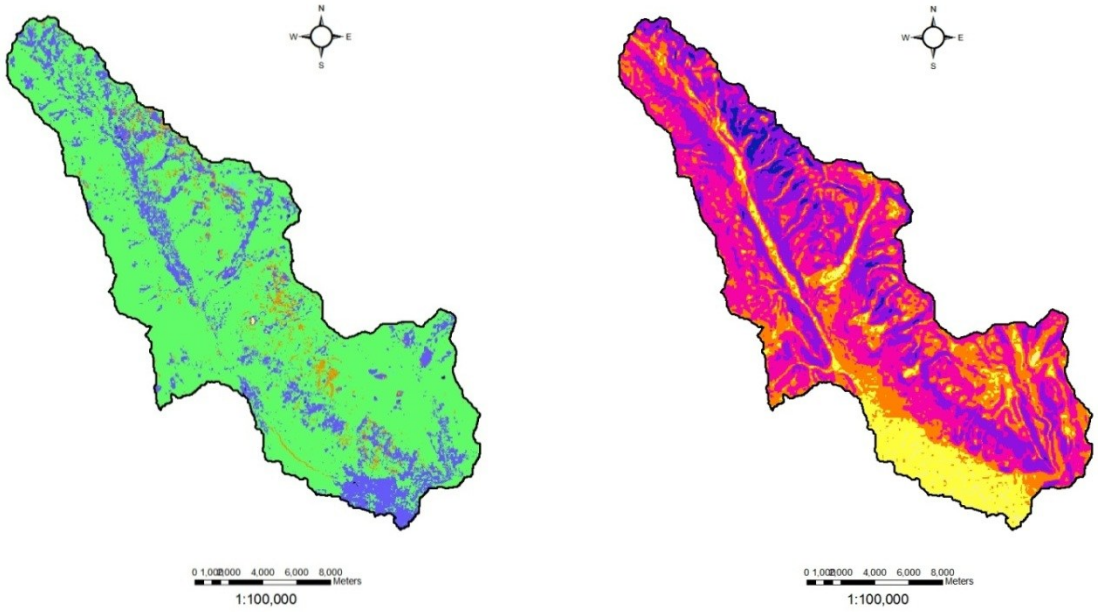
(a) (b)

Figure 42. Stream network (a) and sub watersheds (b) delineated by SWAT based on the 30 m USGS DEM for the Huanquisco River Watershed.

4.3.2. Soils map analysis

An approach for predicting and mapping soil units and relevant soil properties using advanced digital soil mapping methodologies was tested. The methodology has been proposed for soil class mapping for data-poor environments where detailed soil observations are nonexistent and costly to obtain. It comprises two steps: soil unit delineation (soils map) and soil properties definition (map attribute table).

It was observed during the process of creating the soils map that Landsat TM band ratios maximized the discrimination of individual rock types and proved to be very effective in the discrimination of lithological units (figure 43a). Lithological units were relatively easy to group based on their reflectance values. Nine classes were clearly defined with the unsupervised classification method. Slope phases were used to divide the area by slope gradient limits (figure 43b). First, using ArcGIS functions, a slope class map was derived from the DEM. Secondly, the slope map was reclassified in 6 classes. Classification based on slope gradient is very important for soils classification; usually steep slopes have shallow soil types while flat areas tend to have deeper soils.



(a) (b)

Figure 43. Lithologic map constructed from thematic mapper data and slope class raster for the Huanquisco River Watershed.

The lithologic and classified slope maps were used as base layers for creating the final soils map. The base maps were combined by simple overlaying; at the end of the reclassification process, the resulting raster was then classified in nine different classes representing the same number of soil units (figure 44).

Band ratio image (5/7, 5/1, 5/4 * 3/4), slope and unsupervised classification proved to be useful for detailed mapping in data-poor environments. This map was created without the need of detailed soils field data survey. However, the resulting map could not be compared with any detailed geologic map, because it was not possible to find a detailed map for the study zone.

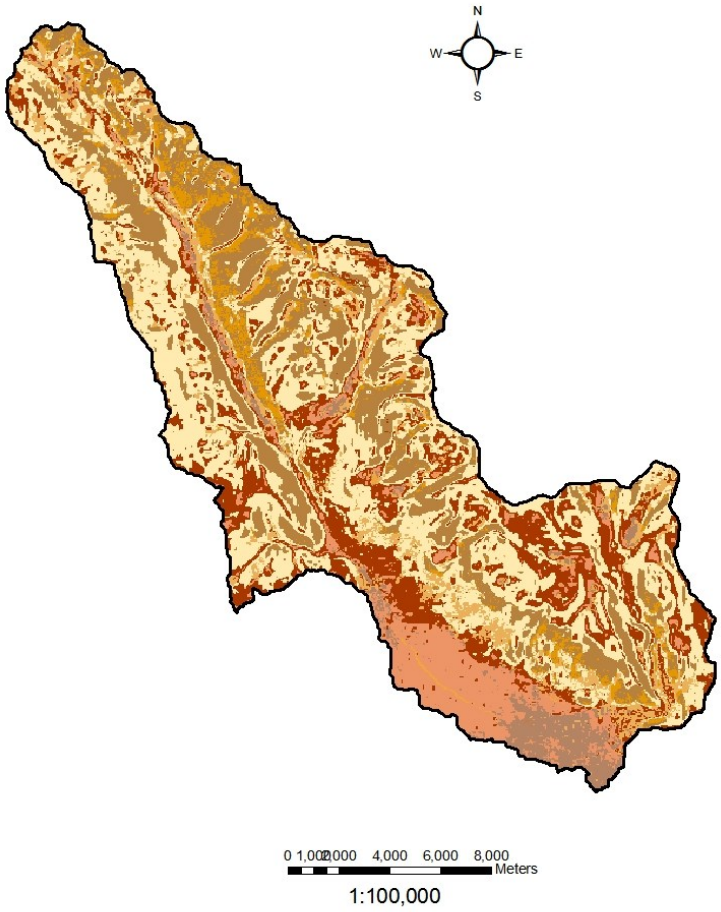


Figure 44. Final soil map for the Huanquisco River Watershed.

Field work was conducted in Ancoraimes, Bolivia for soil sampling to create a geospatial database of sampling points and textural proportions of the top soil. From the field work, a total of 140 soil samples were collected (fig. 45). It is known that the more points, the more accurate the map, but the sample density was limited by time and resources. The textural analysis performed on each of these samples, soil particle separation by suspension, was proven to be simple, easy to be reproduce and inexpensive. However, it is time consuming, since it requires a long period for shaking the solution of soil, water and the chemical compound (20 min for each sample). The readings were easy to obtain for most of the samples; the depth level of soil layers corresponding to sand and silt were measured without any problem; however, some problem arises when the last soil layer to be measured (clay) is extremely thin.

Figure 45. Sampling points for the textural analysis of the Huanquisco River Watershed.

The study was completed using LANDSAT TM and SRTM data to delineate soil units in the Huanquisco watershed. The attribute table was completed by a simple look up process (Appendix C). The method is simple and easily reproducible; the results depend on field sampling to determine texture of the top soil. Future work could improve the soils mapping concept presented here by including pedotransfer functions to predict additional soil parameters for the matching process in order to approximate the soil selection in a better way.

4.3.3. Quantification of uncertainty for the Huanquisco River watershed

Uncertainty was estimated for the following soil parameters likely to be absent in data-poor environments: available water capacity, soil bulk density, USLE equation erodibility factor, organic carbon content, saturated hydraulic conductivity and soil albedo. The same type of

uncertainty analysis performed for the data-rich scenarios was applied here. The analysis was performed varying one parameter at a time, as well as varying all the parameters at the same time for predicted streamflow, sediment yield, total nitrogen and total phosphorous for the Huanquisco River Watershed for 2001-2011. The results are shown in figure 46 where all the uncertainties are expressed through the 10^{th} – 90^{th} percentile range.

For streamflow, the predicted median is visually similar for all the individual soil parameters (0.70 m^3s^{-1}), the median was higher when all the parameters were varied at the same time (0.73 m^3s^{-1}). Slight visual differences are observed with the means of streamflow, means range between 0.68 m^3s^{-1} (saturated hydraulic conductivity) and 0.73 m^3s^{-1}(available water content and all the parameters at the same time). In addition, organic carbon content, albedo, and USLE soil erodibility did not show any impact on streamflow predictions. The Interquartile Range (IQR) for all the parameters at the same time (0.20 m^3s^{-1}) is wider than for bulk density (0.12 m^3s^{-1}), available water content (0.12 m^3s^{-1}), and saturated hydraulic conductivity (0.13 m^3s^{-1}); their 10^{th} – 90^{th} percentile ranges were similar. The 10^{th} – 90^{th} percentile range for bulk density was 0.17 m^3s^{-1}, for available water content was 0.24 m^3s^{-1}, and saturated hydraulic conductivity was 0.28 m^3s^{-1}. The 10^{th} – 90^{th} percentile range of all the parameters at the same time was 0.28 m^3s^{-1} was wider than the range of individual soil parameters.

There was no effect caused by organic carbon content and albedo on sediment yield predictions. The same tendency observed for streamflow was repeated for sediment yield when the soil parameters were varied individually, the median was approximately 2680 t per year and the mean ranged from ~2650 to ~2790 t per year. However, when all the parameters varied at the same time, the mean and median were slightly lower, 2340 and 2540 t per year respectively. The 10^{th} – 90^{th} percentile range was wider when all the parameters were varied at the same time (2697 t) compared to the individual parameters: bulk density was (563 t), for available water content was (1104 t), and saturated hydraulic conductivity was (1843 t).

Median for total nitrogen varied between 166.8 to 172.9 t, while means varied between 144.4 to 198.4 t per year. Albedo did not have any impact on total nitrogen. On the other hand, variation in organic carbon content had a big impact over the 10^{th} – 90^{th} percentile range of total nitrogen (425.8 t), while available water content had the lowest 10^{th} – 90^{th} percentile range (33.8 t). When all the parameters varied at the same time, the median and the mean were lower, 864.7 and 157.3 t respectively, than when soils parameters varied one by one. The predicted

10^{th} – 90^{th} percentile range for all the parameters at the same time was wider (355.0 t) than most of the individual soil parameters with the exception of soil carbon content.

Total phosphorous had almost the same behavior than total nitrogen. Medians of individual soil parameters varied between 22.5 and 25.6 t, while their means varied between 20.8 to 27.8 t per year. As in nitrogen, albedo did not have any impact on total phosphorous. Also, variation in organic carbon content had a big impact over the 10^{th} – 90^{th} percentile range of total phosphorous (51.2 t), while saturated hydraulic conductivity had the lowest 10^{th} – 90^{th} percentile range (5.2 t). When all the parameters varied at the same time, the median and the mean were lower, 12.67 and 20.0 t respectively, than when soils parameters varied one by one. The predicted 10^{th} – 90^{th} percentile range for all the parameters at the same time was wider (47.6 t) than most of the individual soil parameters but soil carbon content.

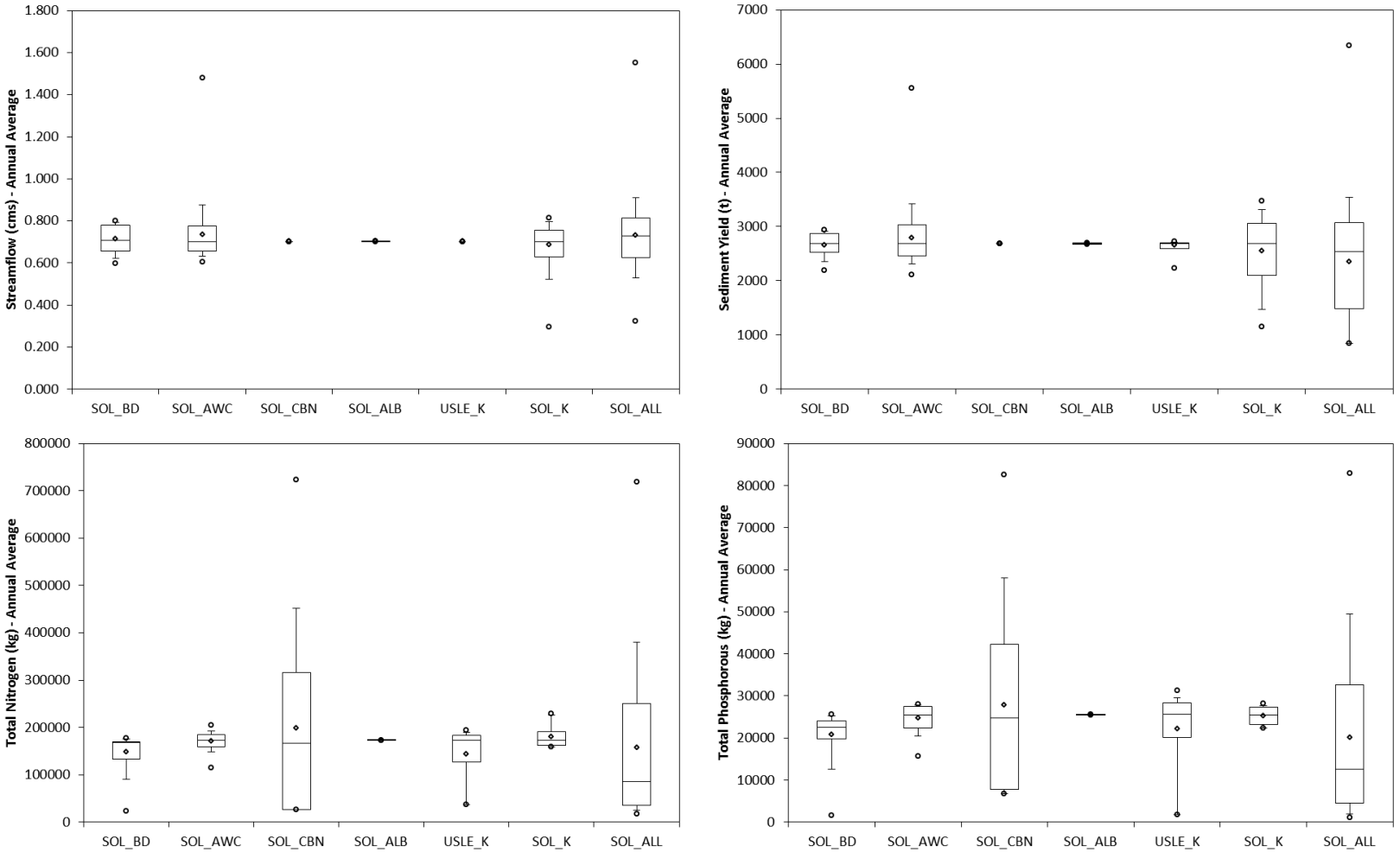


Figure 46. Modified Tukey box plots representing the effect of individual soils parameters (SOL_BD, SOL_AWC, SOL_K, SOL_CBN, SOL_ALB, USLE_K) and all of them at the same time over the uncertainty of SWAT outputs due to for Huanquisco River Watershed. Maximum and minimum values are shown as circles; average as a diamond; the upper and lower whiskers represent the 90th and 10th percentiles, respectively; the upper and lower limits of the boxes are the 3rd and 1st quartiles (75th and 25th percentiles); the line inside the box indicates the median.

4.3.4. Streamflow analysis

Streamflow measurements were available for only a six-month period for Huanquisco Watershed so a formal calibration and validation of the SWAT predictions was not possible. The model was run in a daily time step mode for the period December 31, 2008 to May 31, 2009 to compare measured and predicted streamflow values. Figure 47 shows the predicted SWAT values plotted with the information collected by Peñaranda et al. (2010).

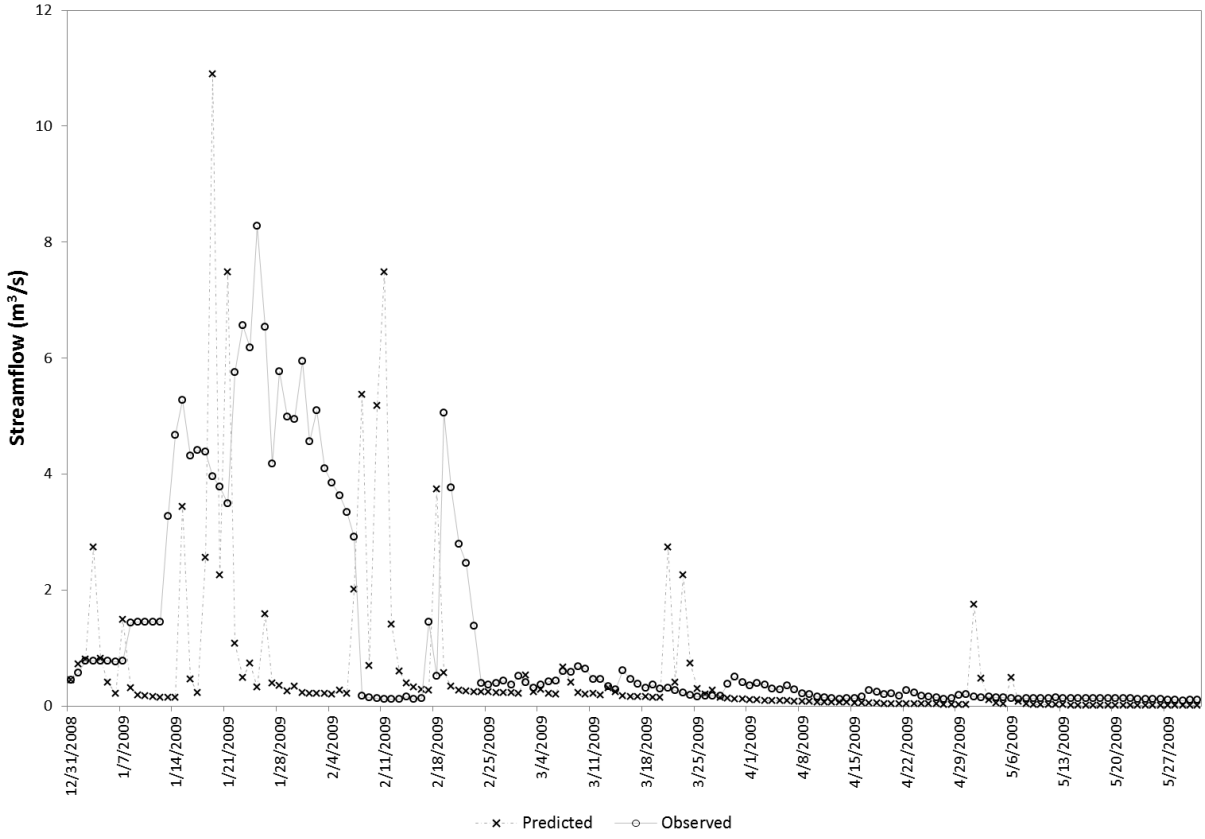


Figure 47. Observed and simulated daily streamflow (m^3s^{-1}) for Huanquisco River Watershed (December, 2008 to May, 2009).

The graph shows a mismatch between observed and predicted values for most of the time period. Differences can be observed in the time of occurrence of some peaks. Additionally, for the time period between January 21 and February 5, differences in the predicted and observed streamflow were ~6 m^3s^{-1} (fig. 47). The trend of the data was the same between predicted and observed streamflow values; when the observed values increased the predicted values did as well and vice versa. In addition, the model performed better during the low flow period than during high flow period. Major differences are observable in January, whereas, a

closer relationship is observed during March, April and May. Despite these differences, the predicted and observed values were the same order of magnitude.

One possible reason for the observed lack of agreement between the measured and predicted streamflow is the location of the weather station used for the climatic input values; the Ancoraimes station falls outside the watershed. In mountainous regions variations in intensity and distribution of rainfall from one place to another are quite common; therefore, weather properties obtained from the Ancoraimes station may not be exactly similar to the properties for the Huanquisco watershed. Secondly, even though SWAT developers stated that the model does not require calibration, several studies demonstrated that SWAT provides more accurate prediction when the model is calibrated against measured data (Gassman et al., 2010); however, as mentioned in the first part of this study, due to the lack of measured data for the study area, calibration was not an option for most places in data-poor environments.

A further analysis includes the 10^{th} – 90^{th} percentile range for each of the daily predictions (figure 48). Those intervals were obtained from 1000 model runs for each day. The predicted-observed plot showed that SWAT underestimated streamflow for most of the wet periods, but there were also some overpredicted values.

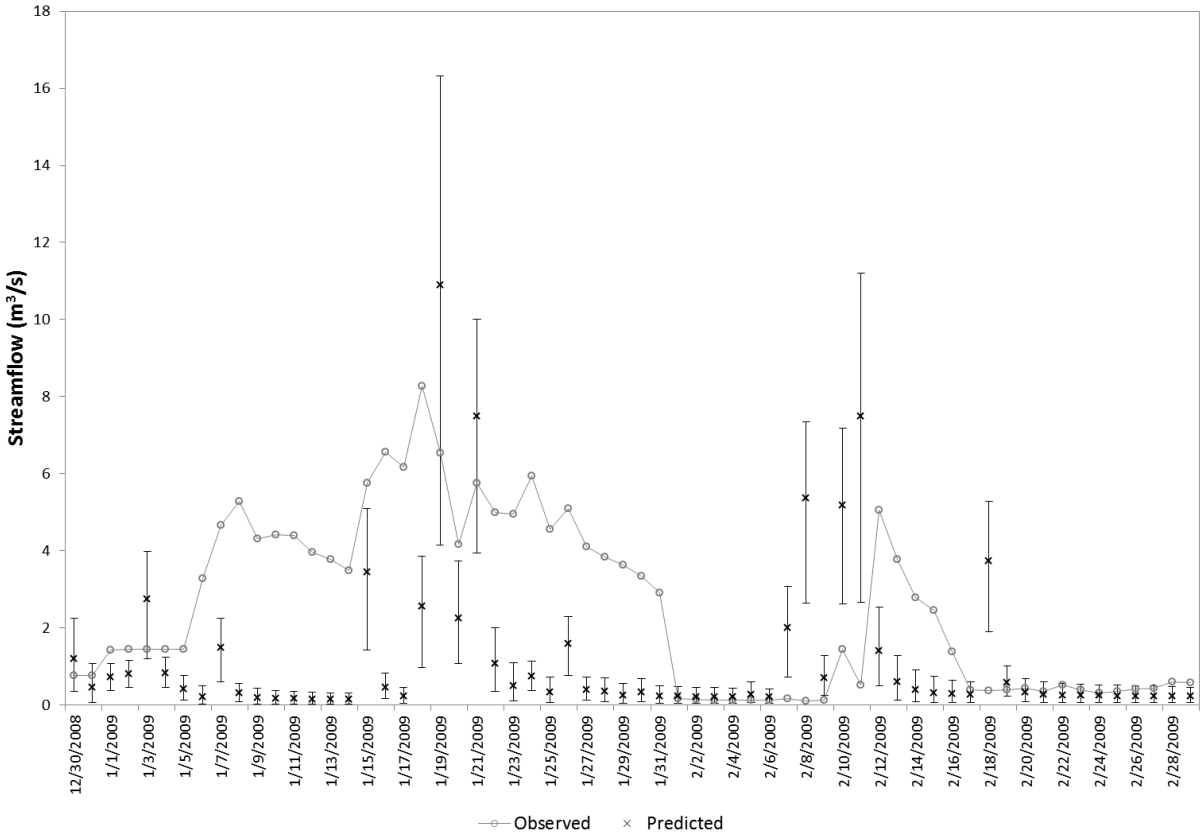


Figure 48. Observed and simulated daily streamflow (m^3s^{-1}) for Huanquisco river Watershed (December, 2008 to April, 2009). Predicted daily streamflow shows the 10^{th} – 90^{th} percentile interval form the uncertainty analysis.

The SWAT model was applied to a mountainous watershed in the Andean region of Bolivia to predict streamflow, sediment, nitrogen and phosphorous. Streamflow predicted values were much better than might have been expected, especially considering that the soil input data were created based on scarce available data and simple field measurements. As there is limited data available from the region of study, the framework developed herein proved helpful to assess different management options for data-poor environments.

The proposed methodologies presented here showed how available knowledge can be employed to generate data for modeling purposes in order to enable SWAT predictions and give the opportunity to incorporate uncertainty in the decision making process in data-poor environments. Making a decision implies that there are options to be considered. Very few decisions are made with absolute certainty because of incomplete knowledge about all of the alternatives; consequently, decisions are often based on intuition. Quantification of uncertainty provides information about the reliability of the predictions, so that decisions based on model predictions are made with knowledge of the model's limitations. Therefore, quantification of

uncertainty helps the process of which decision to adopt. Knowing the range of uncertainty represented by the options allows the decision-maker to determine the alternatives and to make a well-founded and objective decision.

Chapter VI. Conclusions

For this research, a "data-poor environment" was defined as a place where there are not enough data to accurately characterize and evaluate the watershed under study. The lack of adequate data limits the capacity to represent and assess agricultural dominated systems in order to identify the problems to be addressed by a watershed management plan. Therefore, the work presented here addressed the problem of how to enable data-intensive models to support planning and decision-making under situations where data are scarce.

The first objective of the study was to characterize the uncertainty in SWAT's outputs due to input parameters likely to be lacking in data-poor environments. This objective was accomplished as follows. Information from Little River Experimental Watershed (Georgia, USA) was gathered to characterize and evaluate the uncertainty of the SWAT model's predictions of streamflow, sediment yield, and total nitrogen and phosphorous due to inputs likely to be absent in data-poor environments, such as soil parameters. Calibration and validation of streamflow was successfully achieved for annual and monthly values for the time periods 1987-1992 and 1993-1999, respectively. Streamflow, sediment yield, total nitrogen and total phosphorous predictions were found to be sensitive ($Sr \geq \pm 1$) to bulk density, available water capacity, organic carbon content, soil albedo, hydraulic conductivity, and USLE soil erodibility. Quantification of uncertainty due to these soil parameters was performed by varying one soil parameter at a time and all six parameters at the same time. The individual effect of the two soils datasets used in this part of the study, STATSGO and SSURGO, caused significant differences ($p < 0.001$) in the predictions of streamflow, sediment yield, total N and P for all soil parameters. The different DEM resolutions, 10, 30, and 90 m, led to similar results for all resolutions for streamflow and sediment yield, while differences in the predictions of both nutrients, N and P, were seen between 90-m resolution and the other two resolutions (10-m and 30-m). Going from 10 m DEM to 30 m did not increase uncertainty. The effect of the interaction between soils database and DEM resolution was different for each predicted output. However, when all of the soil parameters were varied at the same time, there were no significant differences in predictions as a function of soil database or DEM resolution or their interaction.

Conclusions:

- SWAT predictions were shown to be sensitive to the type of soils data used to characterize the watershed.

- Bulk density, available water capacity, organic carbon content, saturated hydraulic conductivity, and USLE soil erodibility are the key input data on which research should be focused to reduce predictive uncertainty in data-poor environments.
- The uncertainty of model predictions increases as the input data gets coarser, except for 10-m to 30-m DEM, for all individual soil parameters. The combination of SSURGO and 30 m DEM proved to adequately balance the level of uncertainty and the quality of the input data sets.
- To facilitate application of SWAT in data-poor environments, alternatives are needed to downscale 90-m DEM to 30 m and to generate an equivalent soils map to SSURGO.

The SSURGO soils database is available and free of charge for the continental US, Hawaii, Alaska, Puerto Rico and the Virgin Islands. Equivalent soils information can be found for Canada (National Soil Data Base – NSDB: http://sis.agr.gc.ca/cansis/nsdb/), most of Eastern Europe (European Soil Data Centre – ESDB: http://eusoils.jrc.ec.europa.eu/) and Australia (Australian Soil Resource Information System – ASRIS: http://www.asris.csiro.au/); however, most places in Central and South America, Africa and Asia do not have the same detailed soils information or it is not freely available. Digital elevation models at 90 m resolution are freely available for most of the world. However, 90 m resolution is considered coarse and not typically recommended for use in hydrologic analysis. On the other hand, it is reported that there is a threshold beyond which higher resolution of data does not produce better predictions and that the extra cost and labor to obtain DEMs with 10 m resolution is not justified to obtain more accurate prediction. Because of that, a high resolution source, such as 10 m, may not be the best alternative for data-poor environments. In addition to that, this study shows that there is no increase in uncertainty going from 10 m DEM to 30 m. Therefore, a 30 m grid gives a better representation of the surface and can be resampled through resampling techniques, and, thus, is recommended. Consequently, the next part of the research was focused on developing alternatives to downscale 90 m DEM to 30 m and to generate an equivalent soils map to SSURGO.

The second objective of this study was to develop methods to gather qualitative and quantitative data that will allow modeling agricultural watershed systems under data-poor environments. This objective was achieved as follows. First, the 90 m DEM for Little River watershed was obtained from CGIAR-CSI. It was resampled by direct projection using a cubic convolution method. A comparison between the USGS 30 m DEM and the resampled 30 m DEM from CGIAR-CSI showed differences in the minimum, maximum and average elevations

and the number of sub watersheds for Little River. Overall, the resampling method applied here, produced an acceptable elevation map, however, the CGIAR-CSI DEM had higher elevations in the range of 0.1 to 9 m and also derived eight additional sub watersheds. A method for soils dataset generation (SATSLP) was developed based on overlapping a lithologic map derived from LANDSAT imagery and slope derived from a DEM. Differences were seen in the number of soil types for each map: 5, 18 and 9 soil units for STATSGO, SSURGO, and the SATSLP maps, respectively.

Measures of uncertainty of SWAT outputs due to changes in DEM datasets (USGS DEM and CGIAR-CSI DEM) and to the source of the soils database (SSURGO, STATSGO and SATSLP) and their combined effect were compared using an F-test for selected soil parameters. The individual effect of DEM and soils database source caused significant differences ($p < 0.001$) in the predictions of streamflow, sediment yield, total nitrogen and total phosphorous for all soil parameters. The combination of the derived soils map (SATSLP) and the resampled DEM produced a narrower 10^{th} - 90^{th} percentile interval regardless of the soil parameter than the other combinations of DEM and soils database. When all the soil parameters were varied at the same time, streamflow, sediment yield and both nutrient predictions based on the derived soils map were comparable to the predictions based on the SSURGO map. The use of the derived soils map (SATSLP) and the resampled DEM tended to lower the values of the mean and median for streamflow, sediment yield, nitrogen and phosphorous when compared to results of the combinations of USGS 30 m DEM and US soils databases (STATSGO and SSURGO). The use of the resampled DEM lowered the values of all statistics regardless of the soils dataset.

Conclusions:

- The soils dataset generation method for data-poor environments (SATSLP) had comparable results for predicted mean, median and 10^{th} - 90^{th} percentile range as the SSURGO database for all predicted outputs.
- Downscaling of coarser DEM, from 90 m to 30 m, can be applied for application of data-intensive watershed models.
- A soils map that produces similar SWAT predictions to the SSURGO dataset was created using the developed method.
- Soils map development through the use of Landsat and SRTM data is a viable alternative approach for application of data-intensive watershed models.

Results from this part of the study suggest that soil mapping and downscaling of coarser DEM can be done through the use of Landsat and SRTM data as an alternative approach that proved to be simple and cost-effective for watershed modeling purposes in data-poor conditions. The methodologies to overcome insufficient data to characterize the watershed represent a viable option to enable data-intensive models for conditions in which data is scarce, incomplete or nonexistent. The methods presented here are based on freely available data and procedures that are easy to reproduce. However, the methodologies require the use of specialized software, to process remotely sensed data and GIS information, which can be expensive. The proposed methodologies were tested in a data-rich environment, the results showed that resampling of 90 m digital elevation models to a finer resolution (30 m) and the use of satellite imagery to create soils datasets were viable alternatives to overcome with data deficiencies. The method developed for the digital soils mapping can be further improved by comparing the resulting soils map with ground truth data. Also, selection of the soil attributes based on field textural classes could be improved by including pedotransfer functions to predict additional soil parameters in order to narrow down the matching process.

The third objective was to determine to what extent the proposed methodology developed in objective 2 to use SWAT with limited data will be able to represent water quality impacts of agricultural watershed systems in data-poor environments. This objective was accomplished as follows.

Field work was conducted in Ancoraimes, Bolivia for soil sampling and applying a questionnaire. The field soil sampling results were used to create a geospatial database of sampling points and textural proportions of the top soil. The textural analysis performed on each of these samples was proven to be simple, easy to be reproduce and inexpensive; however, time consuming. The SWAT model was applied to a mountainous watershed in the Andean region of Bolivia to predict runoff, sediment, nitrogen and phosphorous to quantify uncertainty due to created input data sets (DEM resampled and soils map modeling). For streamflow, the predicted median and mean values were very similar when varying each of the soil parameters. Organic carbon content, soil albedo and USLE soil erodibility did not show any impact on streamflow predictions. Bulk density, available water capacity, and hydraulic conductivity are the soils parameters that contribute the most to uncertainty of the model predictions. For streamflow and sediment yield, the 10^{th} - 90^{th} percentile range was wider when all the soils parameters were varied at the same time than when they were varied individually. For nutrients, organic

carbon content generated wider 10^{th} - 90^{th} percentile intervals than the other individual parameters and when all parameters were varied at the same time.

A comparison of observed and simulated daily streamflow for a 6-month period (December, 2008 to May, 2009) showed a mismatch between observed and uncalibrated predicted values for most of the time period. However, the tendency of the data maintains a relationship between predicted and observed streamflow values, when the observed values rise the predicted values raise as well and vice versa. Also, the model performed better during the low flow period than during high flow period. Generally, the predicted and observed values have values with the same order of magnitude.

Conclusions:

- Results of the field survey were useful to characterize the watershed with SWAT.
- The textural analysis performed was proven to be simple, easy to reproduce and inexpensive; however, time consuming.
- The approach to generate the soils map and attribute table proved to be simple and cost effective.
- Based on the information used to characterize Huanquisco River watershed, SWAT predictions for streamflow were shown to be in the same order of magnitude as measured data.

The proposed methodologies presented here showed how available knowledge can be employed to generate data for modeling purposes in order to enable SWAT predictions and give the opportunity to incorporate uncertainty in the decision making process in data-poor environments.

Responses to research questions include the following:

What is the capability of SWAT to generate reliable predictions in data-poor environments?

- The lack of enough measured data has limited the capability to evaluate SWAT predictions for the Huanquisco watershed with conventional statistical procedures that compare predicted values against observed data. However, the same methodology presented here was applied to a data-rich condition in which data likely to be absent in data-poor conditions was replaced to mimic a data-poor condition. The results of this exercise compared with the original data set provided satisfactory outcomes. Therefore, the predictions obtained for the Huanquisco Watershed are somehow proved to be reliable. SWAT model was able to represent the watershed in data-poor environments.

This is notable because two of the most important input data sets for the model were created from either coarse data or using alternative sources.

- For streamflow, based on the results of this study, SWAT showed an acceptable performance taking in account that there was no attempt to calibrate the model.

How can available knowledge be employed to generate data for modeling purposes?

- The current study shows different ways to cope with the lack of available data for modeling purposes. The proposed methodologies presented here will enable SWAT predictions in data-poor environments, and give the opportunity to incorporate uncertainty in the decision making process.

What are the key input data requirements to reduce predictive uncertainty in data-poor environments?

- Soils map and attribute table are the most important input data set for the model.
- Bulk density, available water capacity, organic carbon content, and USLE soil erodibility are the key soil parameters.

Does characterization of uncertainty help to overcome the problem of not being able to validate a model in data-poor environments?

- Characterized uncertainty may not help to overcome the problem of being unable to validate a model in data-poor environments; however, it introduces the concept of uncertainty into the decision-making process. Making a decision implies that there are options to be considered. Very few decisions are made with absolute certainty because of incomplete knowledge about all the alternatives; consequently, decisions are often based on intuition. Quantification of uncertainty provides information about the reliability of the predictions, so that decisions based on model predictions are made with knowledge of the model's limitations. The 10^{th} - 90^{th} percentile range was used in this study to quantify uncertainty. Ideally, this range should be as narrow as possible; the smaller the value is (tending to zero), the less uncertain the predicted output is; therefore, quantification of uncertainty helps the process of which decision to adopt. Knowing the range of uncertainty represented by the options allows the decision-maker to determine the alternatives and to make a well-founded and objective decision.

The overall objective of this research was to enable development of watershed management plans for agricultural dominated systems in situations where data are scarce. The framework developed herein showed how available knowledge can be employed to generate

data for modeling for practical management purposes. Therefore, SWAT can be applied to enable development of watershed management plans for agricultural dominated systems under situations where data are scarce.

References

Abbaspour, K. C., J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, and R. Srinivasan. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *J. Hydrology* 333(2): 413–430.

Acron, D. F. 1965. The relationship of pattern and gradient of slopes to soil type. *Can. J. Soil Sci.* 45(1): 96-101.

Aksoy, E., g. Özsoy, and M. Sabri Dirim. 2009. Soil mapping approach in GIS using Landsat satellite imagery and DEM data. *African J. Ag.* 4(11):1295-1302.

Alberti, A., V. Alessandro, U. Pieruccini, and E. Pranzini. 1993. Landsat TM data processing for lithological discrimination in the Caraculo area (Namibe Province, SW Angola). *J. African Earth Sci.* 17(3): 261-274.

Arabi, M., R.S. Govindaraju, M.M. Hantush, and B.A. Engel. 2006. Role of watershed subdivision on modeling the effectiveness of best management practices with SWAT. *Journal of the American Water Resources Association* 42(2): 513-528.

Arnold, J. G., and J. R. Williams. 1987. Validation of SWRRB: Simulator for water resources in rural basins. *J. Water Resour. Plan. Manage. ASCE* 113(2): 243-256.

Arnold, J. G., J. R. Williams, and D. R. Maidment. 1995. Continuous- time water and sediment routing model for large basins. *J. Hydrol. Eng. ASCE* 121(2): 171-183.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J. R. Williams. 1998. Large area hydrologic modeling and assessment: Part I. Model development. *JAWRA* 34(1): 73-89.

Arnold, J. G. and N. Fohrer. 2005. SWAT2000: Current capabilities and research opportunities in applied watershed modeling. *Hydrol. Process.* 19(1): 563–572.

ASCE. 1993. Criteria for evaluation of watershed models. *J. Irrigation Drainage Eng.* 119(3): 429-442.

Barlund, I., T. Kirkkala, O. Malve, and J. Kmri. 2007. Assessing the SWAT model performance in the evaluation of management actions for the implementation of the Water Framework Directive in a Finnish catchment. *Environ. Model. Soft.* 22(5): 719- 724.

Barrera-Bassols, N. and J. A. Zinck. 2003. Ethnopedology: a world view on the soil knowledge of local people. *GEODERMA* 111(1): 171-195.

Barrios E. and Trejo, M.T., 2003: Implications of local soil knowledge for integrated soil management in Latin America. *GEODERMA* 111(1): 217–231.

Batchily, A. K., D. F. Post, R. B. Bryant, and D. J. Breckenfeld. 2003. Spectral reflectance and soil morphology characteristics of Santa Rita experimental range soils. USDA Forest Service Proceedings. 175-182.

Behera, S., and R. K. Panda. 2006. Evaluation of management alternatives for an agricultural watershed in a sub- humid subtropical region using a physical process model. *Agric. Ecosys. Environ.* 113(1): 62-72.

Bennett, S. A., W. W. Atkinson, and F. A. Kruse. 1993. Use of thematic mapper imagery to identify mineralization in the Stanta Teresa district, Sonora, Mexico. *International Geology Review* 35(3): 1009-1029.

Birhanu, B.Z., Ndomba, P.M and F.W. Mtalo. 2007. Application of SWAT model for mountainous catchment. *FWU Water Resources Pub*. 6(1): 181-187.

Boettinger, J. L., R.D. Ramsey, J. M. Bodily, N. J. Cole, S. Kienast-Brown, S. J. Nield, A. M. Saunders and A. K. Stum. 2007. Chapter 16. Landsat Spectral Data for Digital Soil Mapping. In: Lagacherie, P., A. B. McBratney, and M. Voltz (Eds.), Digital Soil Mapping, an introductory perspective. Developments in soil science, vol. 31. Elsevier, Amsterdam, pp. 193–202.

Borah, D. K., and M. Bera. 2003. Watershed- scale hydrologic and nonpoint- source pollution models: Review of mathematical bases. *Trans. ASAE* 46(6): 1553-1566.

Bosch, D. D., J. M. Sheridan, R. R. Lowrance, R. K. Hubbard, T. C. Strickland, G. W. Feyereisen, and D. G. Sullivan. 2007. Little River Experimental Watershed database. *Water Resour. Res.* 43(9): 1-4.

Bosch, D. D. and J. M. Sheridan. 2007. Stream discharge database, Little River Experimental Watershed, Georgia, United States. *Water Resour. Res.* 43(9): 1-4.

Bouraoui, F., S. Benabdallah, A. Jrad, and G. Bidoglio. 2005. Application of the SWAT model on the Medjerda River basin (Tunisia). *Phys. Chem. Earth* 30(10): 497-507.

Bouraoui, F., L. Galbiati, and G. Bidoglio. 2002. Climate change impacts on nutrient loads in the Yorkshire Ouse catchment (UK). *Hydrol. Earth System Sci.* 6(2): 197-209.

Boyle, D. P., H. V. Gupta, and S. Sorooshian. 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Res.* 36(12): 3663-3674.

Brady, N. and R. Weil. 2008. Soils spatial variability in the field. In The nature and properties of soils, 840-845. Columbus, USA. Pearson Prentice Hall.

Cao, W., W. B. Bowden, T. Davie, and A. Fenemor. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Proc.* 20(5): 1057-1073.

Carre, F., and M. C. Girard. 2002. Quantitative mapping of soil types based on regression kriging of taxonomic distances with landform and land cover attributes. *GEODERMA* 110(1): 241– 263.

Chang, C. L. 2009. The impact of watershed delineation on hydrology and water quality simulation. *Environ Monit Assess* 148(1):159–165.

Chanasyk, D.S., Mapfumo, E., and Williams, W. 2003. Quantification and simulation of surface runoff from fescue grassland watersheds. *Agricultural Water Managment.* 59(2): 137-153.

Chapra S. C., and K. Reckhow. 1979. Expressing the phosphorous loading concept in probabilistic terms. *J. Sish. Res. Bd. Can.* 36(1): 225-229.

Chaubey, I. and W.L. White. 2005. Influence of hydrologic response units (HRU) distribution on SWAT model flow and sediment predictions. On proceedings of the 3rd conference of watershed management to meet water quality standards and emerging TMDL. Atlanta, USA. Ed. Gassman, March 5, 2005. ASAE Pub #701P0105.

Chaubey, I., A. S. Cotter, T. A. Costello, and T. S. Soerens. 2005. Effect of DEM data resolution on SWAT output uncertainty. *Hydrol. Process*. 19(2): 621–628.

Cheng, H., W. Ouyang, F. Hao, X. Ren, and S. Yang. 2006. The nonpoint- source pollution in livestock- breeding areas of the Heihe River basin in Yellow River. *Stoch. Environ. Res. Risk Assess.* 21(2): 213-221.

Cho, J., R.R. Lowrance, D.D. Bosch, T.C. Strickland, Y. Her, and G. Vellidis. 2010. Effect of watershed subdivision and filter width on SWAT simulation of a coastal plain watershed. *JAWRA* 46(1):311-319.

Cotter, A. S., I. Chaubey, T. A. Costello, T. S. Soerens, and M. A. Nelson. 2003. Water quality model output uncertainty as affected by spatial resolution of input data. *JAWRA* 39(4): 977-986.

Crippen, R.E. 1989. Selection of Landsat TM band and band ratio combinations to maximize lithologic information in colour composite displays Proceedings of the 7e Thematic Conference on Remote Sensing for Exploration in Geology, Calgary, Alta., 2–6 Octobre 1989.

Elshorbagy, A., and L. Ormsbee. 2006. Object-oriented modeling approach to surface water quality management. *Envi. Modelling & Software* 21(3): 689-698.

Feyereisen, G. W. , T. C. Strickland, D. D. Bosch, and D. G. Sullivan. 2007. Evaluation of SWAT manual calibration and input parameter sensitivity in the little river watershed. *Trans. ASABE* 50(3): 843−855.

Frazier, B. E. 1989. Use of LANDSAT thematic mapper band ratios for soil investigations. *Adv. Space Res.* 9(1): 155-158.

FitzHugh, T.W. and D.S. Mackay. 2001. Impact of sub watershed partitioning on modeled source and transport – limited sediment yields in an agricultural nonpoint source pollution model. *J. Soil And Water Conservation* 56(2): 137-143.

Fry, J.A., M. J. Coan, C. G. Homer, D. K. Meyer, and J. D. Wickham. 2009. Completion of the National Land Cover Database (NLCD) 1992–2001 Land cover change retrofit product: U.S. Geological Survey Open-File Report 2008–1379, 18 p.

Gad, S. and T. Kusky. 2006. Lithological mapping in the Eastern Desert of Egypt, the Barramiya area, using Landsat thematic mapper (TM). *J. African Earth Sci.* 44(): 196–202.

Ghaffari, G. 2011. The Impact of DEM Resolution on Runoff and Sediment Modelling Results. Research *J. Env. Sci.* 5(1): 691-702.

Gassman, P.W. , J. G. Arnold, R. Srinivasan and M. Reyes. 2010. The worldwide use of the SWAT model: technological drivers, networking impacts, and simulation trends. ASABE 21st Century Watershed Technology: Improving Water Quality and Environment CD-Rom Proceedings, 21-24 February 2010, Costa Rica. 9p.

Gassman, P.W., M. Reyes, C.H. Green, and J.G. Arnold. 2007. The Soil and Water Assessment Tool: Historical development, applications, and future directions. *Trans. ASABE*. 50(4): 1211-1250.

Ge, Y., J. A. Thomasson, R. Sui. Front. 2011. Remote sensing of soil properties in precision agriculture: A review. *Earth Sci.* 5(3): 229–238.

Grayson, R. B., G. Bloschl, A. W. Western, T. A. McMahon. 2002. Advances in the use of observed spatial patterns of catchment hydrological response. *Advances in Water Resources* 25(4): 1313–1334.

Gupta, H. V., T. Wagener, and Y. Liu. 2008. Reconciling theory with observations: Elements of a diagnostic approach to model evaluation. *Hydrol. Processes* 22(3): 3802–3813.

Haan, C. T. 1989. Parametric uncertainty in hydrologic modeling. *Trans. ASAE* 32(1): 137-146.

Haimes, Y. Y. (2004). Risk Modeling, Assessment, and Management (2nd ed.). New York: John Wiley & Sons.

Hillel, D. 1998. Spatial variation of properties and processes. In *Environmental soil physics*, 611-613. San Diego, USA. Academic Press.

Hattis, D. and D. Burmaster. 1994. Assessment of variability and uncertainty distributions for practical risk analysis. *Risk Analysis* 14(5): 713-729.

Helton, J. C., J. D. Johnson, C. J. Sallaberry, and C. B. Storlie. 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety* 91(10-11): 1175–1209.

Helton, J.C. 1994. Treatment of uncertainty in performance assessments for complex systems. *Risk Analysis* 14(4): 483–511.

Her, Y. and C. Heatwole. 2008. Assessment of interpolation methods for refining SRTM and DEM. Proceedings of the ASABE Annual International Meeting, Providence, RI.

Hession, W.C., D.E. Storm, and C.T. Haan. 1996. Two-phase uncertainty analysis: an example using the Universal Soil Loss Equation, *Trans. ASAE* 39(4): 1309-1319.

Hession, W.C., and D.E. Storm. 2000. Watershed-level uncertainties: Implications for phosphorous management and eutrophication*. J. Environment Quality* 29(4): 1172-1179.

Jacobs, J.H., and R. Srinivasan. 2005. Application of SWAT in developing countries using readily available data. Proceedings of the 3rd International SWAT conference, Zurich, 2005.

Jarvis, A., H.I. Reuter, A. Nelson, and E. Guevara. 2008. Hole-filled SRTM for the globe Version 4. CGIAR-CSI SRTM 90 m Database. Available at: http://srtm.csi.cgiar.org. Accessed 14 October 2011.

Jayakrishnan, R., R. Srinivasan, C. Santhi and J. G. Arnold. 2005. Advances in the application of the SWAT model for water resources management. *Hydrol. Process.* 19: 749–762.

Jenny, H., 1941. Factors of Soil Formation. McGraw-Hill, New York, Reprinted 1994, Dover Publications, New York.

Jha, M., P.W. Gassman, S. Secchi, R. Gu, and J. Arnold,.2004. Effect of watershed subdivision on SWAT flow, sediment, and nutrient predictions. *JAWRA* 40(3): 811-825.

Juston, J., J. Seibert, and P. O. Johansson. 2009. Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrological Processes* 23(4): 3093-3109.

Keys, R. 1981. Cubic convolution interpolation for digital image processing. *Trans. Acoustics, Speech and Signal Proc.* 29(6): 1153-1160.

Knisel, W. G. 1980. CREAMS, a field-scale model for chemicals, runoff, and erosion from agricultural management systems. USDA Conservation Research Report No. 26. Washington, D.C.: USDA.

Krysanova. V., F. Hattermann, J. Post, A. Habeck, and F. Wechsung. 2005. Prerequisites for application of ecohydrological river basin models in ungauged basins and large regions. Proceedings of the 3rd International SWAT conference, Zurich, 2005.

Lagacherie, P., and A.B. McBratney. 2007. Chapter 1. Spatial soil information systems and spatial soil inference systems: perspectives for Digital Soil Mapping. In: Lagacherie, P., A. B. McBratney, and M. Voltz (Eds.), Digital Soil Mapping, an introductory perspective. Developments in soil science, vol. 31. Elsevier, Amsterdam, pp. 3–24.

Lagacherie, P. 2008. Chapter 1. Digital Soil Mapping: A State of the Art. In: Hartemink, A. E., A. B. McBratney, M. Mendonça-Santos (Eds.), Digital Soil Mapping with Limited Data. Springer, New York, pp. 3-14.

Legates, D. R., and G. J. McCabe. 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Res.* 35(1): 233-241.

Lenhart, T., K. Eckhardt, N. Fohrer, and H.-G. Frede. 202. Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth* 27(9): 645 – 654.

Leon, C. T., D. R. Shaw, M. S. Cox, M. J. Abshire, B. Ward, M. C. Wardlaw and C. Watson. 2003. Utility of Remote Sensing in Predicting Crop and Soil Characteristics. *Precision Ag.* 4(4): 359-384.

Leonard, R. A., W. G. Knisel, and D. A. Still. 1987. GLEAMS: Groundwater loading effects of agricultural management systems. *Trans. ASAE* 30(5): 1403-1418.

Letcher, R.A., W.S. Merrit, A.J. Jakeman, and B. Baginska. 1999. Modelling water quality in data poor catchments: A combined modelling approach. In: Oxley, L. and Scrimgeour, F., Editors, 1999. Proceedings of the International Congress on Modelling and Simulation (MODSIM99), Hamilton, New Zealand vol. 1, pp. 203–208.

Li, P. and Z. Li. 2001. Lithological discrimination of Altun area in northwest China using Landsat TM data and geostatistical textural information. *Geosciences J.* 5(4): 293-300.

Loague, K., and D. L. Corwin. 1996. Chapter 7: Uncertainty in regional-scale assessment of non-point source pollutants. In Applications of GIS to the Modeling of Non-point Source Pollutants in the Vadose Zone, 131-154. SSSA Special Publication Number 48. Madison, W.I.: Soil Science Society of America, Inc.

Lopes, V. L. and H. E. Canfield. 2004. Effects of watershed representation on runoff and sediment yield modeling. *JAWRA* 40(2): 586-602.

Macintosh, D. L., G. W. Suter, and F. O. Hoffman. 1994. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. *Risk Analysis* 14(4): 405-419.

Mamillapalli, S., R. Srinivasan, J. G. Arnold, B. A. Engel. 1996. Effect of spatial variability on basin scale modeling. Proceedings of the 3rd International Conference on Intagreting GIS and Environmental Modeling. January 21-26, Santa Fe, NM., USA.

Mapfumo, E., D. S. Chanasyk, and W. D. Willms. 2004. Simulating daily soil water under foothills fescue grazing with the Soil and Water Assessment Tool model (Alberta, Canada). *Hydrol. Process.* 18(3): 2787- 2800.

Marse K. J. and S. D. Roberts. 1983. Implementing a portable FORTRAN Uniform (0, 1) generator. *Simulation* 41(4): 135-139.

Mather, P.M., Tso, B. and Koch, M., 1998, An evaluation of Landsat TM spectral data and SAR-derived textural information for lithological discrimination in the Red Sea Hills, Sudan. International. *J. Remote Sensing* 19(1): 587−604.

McBratneya, A. B., M. L. Mendonca Santos, B. Minasny. 2003. On digital soil mapping. *GEODERMA* 117(1): 3– 52.

Mshiu, E. E. 2011. Landsat remote sensing data as an alternative approach for geological mapping in Tanzania: a case study in the Rungwe volcanic province, south-western Tanzania. *Tanz. J. Sci.* 37(1): 26-36.

Moore, I. D., P. E. Gessler, G. A. Nielsen, and G. A. Peterson. 1993. Soil attribute prediction using terrain analysis. *Soil Science Society of America J.* 57(3): 443-452.

Morgan, M. G., and M. Henrion (1990). Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis, Cambridge University Press, Cambridge, 47-72.

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

Mowrer, H. T. 2000. Uncertainty in natural resource decision support systems: sources, interpretation, and importance. *Computers and Electronics in Agriculture* 27(1): 139-154.

Muleta, M.K. and J.W. Nicklow. 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology* 306(1): 127-145.

Neff, E.L. 1965. Principles of precipitation design for intensive hydrologic investigations. Symposium on Design of Hydrological Networks, IASH Publication 67(1): 49-55.

Neitsch S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2005. Soil and Water Assessment Tool User's manual. Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, TX 76502. 455 p.

Newham, L.T.H., R.A. Letcher, A.J. Jakeman, and T. Kobayashi. 2004. A framework for integrated hydrologic, sediment and nutrient export modelling for catchment-scale management. *Environmental Modelling & Software* 19(5): 1029–1038.

O'Hagan, A. 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety* 91 (10-11): 1290–1300.

Olivera, F., M. Valenzuela, R. Srinivasan, J. Choi, H. Cho, S. Koka, and A. Agrawal. 2006. ArcGIS- SWAT: A geodata model and GIS interface for SWAT. *JAWRA* 42(2): 295- 309.

Palacios-Orueta, A., and S. L. Ustin. 1998. Remote Sensing of Soil Properties in the Santa Monica Mountains I. Spectral Analysis. *Remote Sens. Environ.* 65(1): 170–183.

Palisade. 2011. Monte Carlo simulation. Available at: http://www.palisade.com/risk/monte_carlo_simulation.asp. Accessed 10 September 2011.

Pawluk, R. R., J. A. Sandor, and J. A. Tabor. 1992. The role of indigenous soil knowledge in agricultural development. *J. of Soil and Water Cons.* 47(4): 298-302.

Peñaranda, M., J. Osorio, and C. Heatwole. 2010. Assessing climate impacts on water resources through monitoring and modeling water balance in an Andean catchment in Andean catchment in Ancoraimes, Bolivia. Quality and the Environment Conference at Earth University, Costa Rica (February 2010).

Rawlsa, W. J., Y. A. Pachepsky, J. C. Ritchie, T. M. Sobecki, and H. Bloodworth. 2003. Effect of soil organic carbon on soil water retention. *GEODERMA* 116(1): 61-76.

Ricchetti, E, 2000, Multispectral satellite image data and ancillary data integration for geological classification. *Photogrammetric Engineering and Remote Sensing* 66(1): 429−435.

Riley, J. 2001. Multidisciplinary indicators of impact and change. Key issues for identification and summary. *Agriculture, Ecosystems and Environment* 87(1): 245–259.

Rouhani, H., J. Feyen and P. Willems. 2006. Impact of Watershed Delineations on the SWAT Runoff Prediction: A Case of Study in the Grote Nete Catchment, Flanders, Belgium. Proceedings of the 2006 IASME/WSEAS Int. Conf. on Water Resources & Hydrology, Chalkida, grecce, May 11-13, 2006: 36-41.

Schafer, D., and G. Hanlon. 2001. Data Collection for Watershed Management. ASCE Conf. Proc. Pp. 111-323.

Santhi, G., J. G. Arnold, J. R. Williams, W. A. Dugas, R. Srinivasan, and L. M Hauck. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. *JAWRA* 37(5): 1169 – 1188.

Schetselaar, E.M., Chung, C.F. and Kim, K.E., 2000, Integration of Landsat TM, Gamma-ray, magnetic, and filed data to discriminate lithological units in vegetated Granite-Gneiss Terrain. *Remote Sensing of Environment* 71(1): 89−105.

Scull, P., J. Franklin, O. A. Chadwick, and D. McArthur. 2003. Predictive soil mapping: a review. *Progress in Physical Geography* 27(2): 171–197.

Seyfried, M., R. Harris, D. Marks, and B. Jacob. 2001. Geographic database, Reynolds Creek experimental watershed, Idaho, United States. *Water Resources Research* (37)11: 2825–2829.

Seibert, J., and K. Beven. 2009. Gauging the ungauged basin: how many discharge measurements are needed?. *Hydrol. Earth Syst. Sci. Discuss.* 6(2): 2275–2299.

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

Singh, J., H. V. Knapp, J. G. Arnold, and M. Demissie. 2005. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. *JAWRA* 41(2):361-375.

Sivapalan, M. 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrol. Process.* 17(3): 3163–3170.

Sivapalan, M., K. Takeuchi, S. W. Franks, V. K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. J. McDonnell, E. M. Mendiondo, P. E. O'Connell, T. Oki, J. W. Pomeroy, D. Schertzer, S. Uhlenbrook, E. Zehe. 2003. IAHS decade on predictions in ungauged basins (PUB), 2003-2012: shaping an exciting future for the hydrological sciences. Hydrological Sciences Journal 48(6): 857-880.

Slaughter, C. W., D. Marks, G. N. Flerchinger, S. S. Van Vactor, and M. Burgess. 2001. Thirty-five years of research data collection at the Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research* 37(11): 2819-2823.

Sohrabi, T.M., A. Shirmohammadi, T.W. Chu, H. Montas, and A.P. Nejadhashemi. 2003. Uncertainty Analysis of Hydrologic and Water Quality Predictions for a Small Watershed Using SWAT2000. *Environmental Forensics* 4(1): 229–238.

Soils Conservation Service. 1986. Urban hydrology for small watersheds. Technical Release No. 55. Soils Conservation Service, U.S. Department of Agriculture, Washington, D.C. 164p.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Idaho and Georgia. Available online at http://soildatamart.nrcs.usda.gov Accessed January 15, 2011.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Idaho and Georgia. Available online at http://soildatamart.nrcs.usda.gov Accessed January 15, 2011.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soils data for the Conterminous United States Derived from the NRCS State Soil Geographic (STATSGO) Data Base. Available online at http:// water.usgs.gov/GIS/metadata/usgswrd/XML/ussoils.xml#stdorder. Accessed August 15, 2011.

Soil Survey Division Staff. 1993. Soil survey manual. Soil Conservation Service. U.S. Department of Agriculture Handbook 18.

Srivastav, R. K., K. P. Sudheer, and I. Chaubey. 2007. A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. *Water Resour. Res.* 43(1): 1- 12.

Sullivan, D. G., H. L. Batten, D. Bosch, J. Sheridan, and T. Strickland. 2007. Little River Experimental Watershed, Tifton, Georgia, United States: A geographic database. *Water Resour. Res.* 43(9): 1-4.

Sultan, M., R. E. Arvidson, N. C. Sturchio, and E. A. Guiness. 1987. Lithologic mapping in arid regions with Landsat thematic mapper data: Meatiq dome, *Egypt. Geo. Soc. America Bull.* 9(2): 748-762.

Tripathi, M.P., N.S. Raghuwanshi and G.P. Rao. 2006. Effect of watershed subdivision on simulation of water balance components*. Hydrol. Process*. 20(3): 1137-1156.

U.S. EPA. 2002. Guidance for quality assurance project plans for modeling. EPA QA/G-5M. Report EPA/240/R-02/007. Washington, D.C.: U.S. EPA, Office of Environmental Information.

Vale, J. F., C. E. Schafer, J. A. Viera. 2007. Ethnopedology and knowledge transfer: dialogue between Indians and soil scientist in the Malacacheta Indian Territory, Roraima, Amazon. *Rev. Bras Cienc. Solo.* 31(2): 403-412.

van Griensven, A., and T. Meixner. 2006. Methods to quantify and identify the sources of uncertainty for river basin water quality models. Water Science Technology 53(1): 51–59.

van Griensven, A., L. Breuer, M. Di Luzio, V. Vandenberghe, P. Goethals, T. Meixner, J. Arnold, and R. Srinivasan. 2006. Environmental and ecological hydroinformatics to support the implementation of the European Water Framework Directive for river basin management. *J. Hydroinformatics* 8(4): 239-252.

Walker, W.E., P. Harremoes, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen, M.P. Krayer von Krauss. 2003. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment* 4(1): 5-17.

Watson, B.M., S. Selvalingam and M. Ghafouri. 2003. Evaluation of SWAT for modelling the water balance of the Woady Yaloak River catchment, Victoria, in Post, David (eds), MODSIM 2003 : International Congress on Modelling and Simulation, Jupiters Hotel and Casino, 14-17 July 2003 : integrative modelling of biophysical, social and economic systems for resource management solutions : proceedings, pp. 1-6, Modelling and Simulation Society of Australia and New Zealand Inc., Canberra, A.C.T.

Wheater, H.S., A.J. Jakeman, and K.J. Beven. 1993. Progress and directions in rainfallrunoff modelling. In: Jakeman, A.J., M.B. Beck and M.J. McAleer (Eds.), Modelling Change in Environmental Systems. Wiley, NY, pp. 101-132.

Williams, J. R., A. D. Nicks, and J. G. Arnold. 1985. Simulator for Water Resources in Rural Basins. *J. Hydraulic Engineering* 111(6): 970-986.

Williams, J. R. 1990. The erosion productivity impact calculator (EPIC) model: A case history. Phil. *Trans. R. Soc. London* 329(1255): 421- 428.

Yang, J., P. Reichert, K.C. Abbaspour, J. Xia, and H. Yang. 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J. Hydrology* 358(1): 1–23.

Yilmaz K. K., 2007: Effectively transforming data into information; A Student's View, PUB Newsletter, Issue 2.4 (December 2007), downloadable at: (http://pub.iwmi.org)

Yilmaz, K. K., 2007: Towards improved modeling for hydrologic predictions in poorly gauged basins, Ph.D. Dissertation, University of Arizona, Tucson, Arizona, 263 pages.

Zambrano, M. 2010. Package HydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. Version: 0.2-1. Available at: http://cran.r-project.org/web/ packages/hydroGOF/index.html. Accessed 19 January 2011.

Zhang, X., R. Srinivasan, and F. Hao. 2007. Predicting hydrologic response to climate change in the Luohe River basin using the SWAT model. *Trans. ASABE* 50(3): 901- 910.

Zhang, X., R. Srinivasan, and D. Bosch. 2009. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *J. Hydrology* 374(1): 307–317.

Zilli Bacic, I.L., D. G. Rossiter and C. M. Mannaerts. 2008. Applicability of a distributed watershed pollution model in a data-poor environment in Santa Catarina State, *Brazil. R. Bras. Ci. Solo* 32(4):1699-1712.

Appendices

Appendix A. In-situ textural classification

The percentages of sand, silt, and clay particles in a soil were estimated according to the following procedure:

1. Select a straight-sided bottle and fill approximately 3/5 full of soil (about 500 ml or 2 cups)
2. Add water up to 1 inch from the top. Put a strip of masking tape up one side of the bottle.
3. Add 5-6 drops of ammonium hydroxide (household ammonia) to separate the sand, silt and clay.
4. Mix for 20 minutes by rotating through 180 degrees.
5. Set the bottle down, let it stand for 45 seconds, and mark the sediment level. It represents the sand.
6. At the end of 30 minutes mark the sediment level in the bottle; this is the silt line. Note how slowly the soil particles are settling. The clay is still in suspension.
7. After two or three days, study the soil in the bottle carefully. The sand will be easily identifiable; if you can see the particles then it is sand. Sand will be very near the 45-second mark. The silt should have formed after 30 minutes. The clay will be a thin third layer lying on top of the silt.
8. Measure the total depth of soil in the bottle, in millimeters (mm). Then measure the depth of each layer. If the clay fraction is still in part suspended (as indicated by cloudy water), add 1-2 mm to the clay reading depending on the density of the suspension. Record these readings.
9. Using the following bulk density factors calculate the % by weight of each of the primary soil separates.

   Sand = 1.5 bulk density

   Silt = 1.3 bulk density

   Clay = 1.2 bulk density

   To calculate % by weight:

   A. mm of depth of separate layer x bulk density = weight relationship (WR).

   B. Add the 3 weight relationships to obtain a total weight relationship (TWR).

   C. (WR / TWR) x 100 = % by weight of each separate.

10. Using these percentages by weight for each separate, determine the soil class by use of the textural triangle.

Appendix B. Texture classes and classification

ID	CLAY	SILT	SAND	TEXTURE	ID	CLAY	SILT	SAND	TEXTURE
1	2.7	49.9	47.4	SANDY LOAM	71	18.0	21.6	60.3	SANDY LOAM
2	2.9	34.3	62.9	SANDY LOAM	72	27.1	7.3	65.6	SANDY CLAY LOAM
3	4.5	37.3	58.2	SANDY LOAM	73	22.6	24.0	53.4	SANDY CLAY LOAM
4	4.5	37.3	58.2	SANDY LOAM	74	31.7	33.3	35.0	CLAY LOAM
5	3.1	35.1	61.8	SANDY LOAM	75	27.1	23.6	49.2	SANDY CLAY LOAM
6	3.8	32.9	63.3	SANDY LOAM	76	15.8	13.3	70.9	SANDY LOAM
7	3.8	32.9	63.3	SANDY LOAM	77	11.1	41.3	47.6	LOAM
8	33.1	28.0	38.9	CLAY LOAM	78	22.6	30.0	47.4	LOAM
9	3.1	35.1	61.8	SANDY LOAM	79	17.1	29.3	53.6	SANDY LOAM
10	4.3	37.7	58.0	SANDY LOAM	80	9.5	48.9	41.6	LOAM
11	6.1	51.6	42.3	SILT LOAM	81	22.6	28.0	49.4	LOAM
12	21.8	55.3	22.9	SILT LOAM	82	13.7	11.6	74.7	SANDY LOAM
13	2.1	35.8	62.1	SANDY LOAM	83	19.8	44.0	36.2	LOAM
14	2.1	35.8	62.1	SANDY LOAM	84	44.6	28.0	27.4	CLAY
15	18.8	34.3	46.9	LOAM	85	19.7	46.7	33.6	LOAM
16	2.1	35.8	62.1	SANDY LOAM	86	26.2	27.5	46.3	SANDY CLAY LOAM
17	6.0	54.2	39.8	SILT LOAM	87	17.7	28.0	54.3	SANDY LOAM
18	6.0	54.2	39.8	SILT LOAM	88	17.4	21.0	61.6	SANDY LOAM
19	21.7	33.3	45.0	LOAM	89	25.1	35.6	39.2	LOAM
20	42.8	24.0	33.2	CLAY	90	17.7	28.0	54.3	SANDY LOAM
21	26.2	27.5	46.3	SANDY CLAY LOAM	91	19.1	37.3	43.6	LOAM
22	42.8	24.0	33.2	CLAY	92	19.1	37.3	43.6	LOAM
23	26.2	27.5	46.3	SANDY CLAY LOAM	93	54.7	12.7	32.6	CLAY
24	7.5	33.2	59.3	SANDY LOAM	94	54.7	12.7	32.6	CLAY
25	6.6	47.2	46.2	SANDY LOAM	95	7.0	12.2	80.8	LOAMY SAND
26	6.6	47.2	46.2	SANDY LOAM	96	6.9	12.1	81.0	LOAMY SAND
27	6.6	47.2	46.2	SANDY LOAM	97	7.0	12.3	80.7	LOAMY SAND
28	7.9	51.2	40.9	SILT LOAM	98	6.7	19.5	73.8	SANDY LOAM
29	6.0	54.2	39.8	SILT LOAM	99	6.5	19.8	73.7	SANDY LOAM
30	6.0	54.2	39.8	SILT LOAM	100	6.6	19.6	73.9	SANDY LOAM
31	6.6	47.2	46.2	SANDY LOAM	101	23.4	22.0	54.6	SANDY CLAY LOAM
32	6.6	47.2	46.2	SANDY LOAM	102	4.2	28.3	67.5	SANDY LOAM
33	6.6	47.2	46.2	SANDY LOAM	103	4.1	28.5	67.4	SANDY LOAM
34	2.1	35.8	62.1	SANDY LOAM	104	7.0	12.2	80.8	LOAMY SAND
35	2.1	35.8	62.1	SANDY LOAM	105	4.9	27.3	67.8	SANDY LOAM
36	7.8	59.1	33.1	SILT LOAM	106	4.8	27.6	67.6	SANDY LOAM
37	3.8	32.9	63.3	SANDY LOAM	107	2.8	42.7	54.5	SANDY LOAM
38	19.1	37.3	43.6	LOAM	108	2.7	42.5	54.8	SANDY LOAM
39	2.9	34.3	62.9	SANDY LOAM	109	2.8	42.7	54.5	SANDY LOAM
40	33.5	46.0	20.5	CLAY LOAM	110	3.1	25.2	71.7	SANDY LOAM
41	17.1	36.4	46.5	LOAM	111	3.5	40.6	55.9	SANDY LOAM
42	21.8	28.6	49.6	LOAM	112	3.4	40.9	55.7	SANDY LOAM
43	33.1	28.0	38.9	CLAY LOAM	113	3.4	40.9	55.7	SANDY LOAM
44	25.1	32.4	42.5	LOAM	114	4.3	37.7	58.0	SANDY LOAM
45	25.5	24.9	49.6	SANDY CLAY LOAM	115	4.3	37.7	58.0	SANDY LOAM
46	33.8	40.4	25.8	CLAY LOAM	116	4.5	37.3	58.2	SANDY LOAM
47	23.1	26.0	50.9	SANDY CLAY LOAM	117	6.1	51.6	42.3	SILT LOAM
48	21.1	30.0	48.9	LOAM	118	3.1	35.1	61.8	SANDY LOAM
49	27.1	47.3	25.6	CLAY LOAM	119	7.3	22.0	70.7	SANDY LOAM
50	33.8	49.3	16.9	SILTY CLAY LOAM	120	2.9	42.5	54.5	SANDY LOAM
51	28.6	36.0	35.4	CLAY LOAM	121	2.9	34.3	62.9	SANDY LOAM
52	16.8	20.7	62.5	SANDY LOAM	122	2.7	49.9	47.4	SANDY LOAM
53	15.7	39.6	44.7	LOAM	123	2.7	49.9	47.4	SANDY LOAM
54	36.6	38.0	25.4	CLAY LOAM	124	3.8	32.9	63.3	SANDY LOAM
55	24.8	35.1	40.2	LOAM	125	3.8	32.9	63.3	SANDY LOAM
56	29.5	21.6	48.9	SANDY CLAY LOAM	126	18.8	34.3	46.9	LOAM
57	48.5	21.9	29.6	CLAY	127	18.8	34.3	46.9	LOAM
58	4.4	22.7	72.9	SANDY LOAM	128	7.1	33.5	59.4	SANDY LOAM
59	19.1	22.0	58.9	SANDY LOAM	129	7.1	33.5	59.4	SANDY LOAM
60	37.5	26.9	35.6	CLAY LOAM	130	7.5	33.2	59.3	SANDY LOAM
61	21.8	55.3	22.9	SILT LOAM	131	7.1	33.5	59.4	SANDY LOAM
62	24.2	26.9	48.9	SANDY CLAY LOAM	132	6.9	32.7	60.4	SANDY LOAM
63	19.5	25.6	54.9	SANDY LOAM	133	6.6	47.2	46.2	SANDY LOAM
64	23.8	53.3	22.9	SILT LOAM	134	6.6	47.2	46.2	SANDY LOAM
65	42.8	24.0	33.2	CLAY	135	2.9	43.4	53.7	SANDY LOAM
66	23.5	8.0	68.5	SANDY CLAY LOAM	136	2.9	43.4	53.7	SANDY LOAM
67	19.1	32.0	48.9	LOAM	137	3.0	31.0	66.0	SANDY LOAM
68	21.7	33.3	45.0	LOAM	138	3.0	31.0	66.0	SANDY LOAM
69	25.1	32.4	42.5	LOAM	139	2.8	32.0	65.2	SANDY LOAM
70	19.1	22.0	58.9	SANDY LOAM	140	6.2	55.2	38.6	SILT LOAM

Appendix C. Huanquisco RW – Soils attribute table

TEXTURE CLASS	MUID	SEQN	SNAM	S5ID	CMPPCT	NLAYERS	HYDGRP	SOL_ZMX	ANION_EXCL	SOL_CRK	TEXTURE	SOL_Z	SOL_BD	SOL_AWC	SOL_K	SOL_CBN	CLAY	SILT	SAND	ROCK	SOL_ALB	USLE_K	SOL_EC
SILT LOAM	WI068	2	BASCO	WI0264	13	4	C	1524	0.5	0.5	SIL-SICL-SC-WB	152.4	1.38	0.19	5.5	1.74	16.0	54.3	29.7	4.0	0.01	0.32	0
												406.4	1.60	0.18	1.1	0.58	30.0	52.2	17.9	4.7	0.08	0.43	0
												838.2	1.75	0.10	1.0	0.19	47.5	2.8	49.7	5.1	0.16	0.32	0
												1524.0	1.70	0.05	550.0	0.06	20.0	40.0	40.0	90.0	0.20	0.00	0
CLAY	AR024	17	DESHA	AR0043	7	3	D	1828.8	0.5	0.5	C-C-SIC	177.8	1.30	0.15	0.4	1.45	55.0	27.8	17.2	0.0	0.01	0.37	0
												1397.0	1.30	0.09	0.1	0.29	70.0	21.1	8.9	0.0	0.13	0.28	0
												1828.8	1.38	0.13	0.2	0.29	47.5	47.0	5.5	0.0	0.13	0.32	0
SANDY LOAM	ME005	14	WAUMBEK	NH0016	3	3	B	1651	0.5	0.5	STV-FSL-GRV-LFS-CB--COS	254.0	0.95	0.09	300.0	0.00	3.0	35.2	61.8	40.0	0.23	0.17	0
												660.4	1.10	0.07	400.0	0.00	3.0	16.7	80.3	23.1	0.23	0.17	0
												1651.0	1.67	0.02	450.0	0.00	1.5	6.5	92.0	57.9	0.23	0.17	0
LOAM	ID207	6	CHAYSON	ID6026	1	4	C	1625.6	0.5	0.5	L-CL-IND-WB	203.2	1.20	0.16	28.0	1.45	18.0	38.8	43.2	3.5	0.01	0.28	0
												685.8	1.38	0.15	8.5	1.16	29.0	37.0	34.0	11.5	0.02	0.24	0
												711.2	2.20	0.01	210.0	0.44	5.0	25.0	70.0	0.0	0.10	0.00	0
												1625.6	1.70	0.05	84.0	0.15	20.0	40.0	40.0	90.0	0.17	0.00	0
SANDY LOAM	AZ059	1	WINKEL	UT0034	45	4	D	533.4	0.5	0.5	GR--FSL-CBV-FSL-IND-UWB	152.4	1.42	0.11	100.0	0.29	6.0	34.1	59.9	33.0	0.13	0.10	1
												406.4	1.48	0.08	81.0	0.10	6.0	34.1	59.9	55.0	0.19	0.02	1
												508.0	2.20	0.01	100.0	0.03	5.0	25.0	70.0	0.0	0.22	0.00	0
												533.4	2.50	0.01	650.0	0.01	5.0	25.0	70.0	98.0	0.23	0.00	0
SANDY LOAM	OR152	17	ADKINS	WA0416	8	3	B	1524	0.5	0.5	FSL-VFSL-VFSL	177.8	1.23	0.14	260.0	0.49	6.0	34.1	59.9	0.0	0.09	0.37	0
												609.6	1.35	0.15	250.0	0.29	6.0	35.0	59.0	0.0	0.13	0.55	0
												1524.0	1.55	0.15	190.0	0.29	6.0	35.0	59.0	0.0	0.13	0.55	1
SANDY LOAM	ID319	6	ROSEBERRY	ID0305	43	4	D	1524	0.5	0.5	COSL-LCOS-COS-VFSL	330.2	1.52	0.17	140.0	2.33	8.5	24.1	67.5	6.0	0.01	0.28	0
												889.0	1.52	0.11	200.0	0.78	6.0	10.9	83.1	2.9	0.05	0.10	0
												1397.0	1.80	0.05	270.0	0.26	5.0	3.9	91.1	10.7	0.14	0.10	0
												1524.0	1.65	0.14	31.0	0.09	11.5	25.2	63.3	3.2	0.20	0.10	0