PROBABILISTIC APPROACH TO MODELING UNDER CHANGING SCENARIOS

BY

ANDRES FELIPE PRADA SEPULVEDA

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Biological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Adviser:

Assistant Professor Maria L. Chu

ABSTRACT

The complexity of the hydrologic system challenges the development of models. One issue faced at model development stage is the uncertainty involved when calibrating and validating the model. Model inputs and parameters can introduce large amount of uncertainties that can be propagated non-linearly to the model outputs. Additionally, several sets of parameters may also exist that acceptably represent the system (i.e., equifinality). As a result, converting model outputs into important environmental decisions become challenging. The main objective of this study was to define a framework that facilitates model development while evaluating uncertainty to assess the impacts of land management practices at watershed scale. A two-step probabilistic approach to model calibration and parametrization was implemented using global uncertainty and sensitivity analysis. The Agricultural Policy/Environmental eXtender (APEX) model was developed for the Lake Creek Watershed in Oklahoma using probabilistic parameters to derive the spectrum of responses of the model for water yield and nitrogen loads. A variance-based sensitivity analysis was used to identify the most important parameters, and their ranges, that drive these spectrum of responses. The baseline APEX model, composed of twenty-seven different sets of parameters, was then applied to estimate the water yield and N loads in the watershed for 7 years (2007-2013) under different land management scenarios. The total monthly water yield was found to range from 0.17 to 41.5 mm with an uncertainty of 11%, while the total monthly Nitrogen loads can vary from 0 to 5.3 kg/ha with uncertainty of 50%. Four alternative land use scenarios (75% Pasture, 100% Pasture, 75% Winter Wheat, and 100% Winter Wheat) and two alternative land management scenarios (conventional tillage for grain, and conventional tillage for graze out) were proposed and simulated over the study area to observe their effects on the monthly N loads. Results suggested that changes in land use and land

management did not affect the total water yield at watershed scale. However, the N loads showed a high variability ranging from 0.1 to 1.2 kg/ha during summer and fall seasons, with uncertainty of up to 77%. The 100% Pasture scenario was the most effective alternative in reducing nitrogen, the N loads in this scenario did not exceed 0.9 kg/ha in any season. This methodology demonstrated that the modeling process, coupled with the evaluation of uncertainty and equifinality, facilitate the adjustment of input/parameters and quantify the uncertainties in the model outputs. By considering all possible parameter combinations that represent the response of the system, the most likely ranges of hydrologic outcomes can be established under changing scenarios while accounting for the associated uncertainty.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the following individuals, who helped and guided me through the process of writing this thesis:

Dr. Maria Chu, for all her patience and assistance during these two years. Without her guidance and leadership, this project would not have been successfully completed.

Dr. Rabin Bhattarai, Dr. Luis Rodriguez, and Dr. Richard Cooke, for the time they spent reading this thesis and the valuable feedback they provided.

Dr. Jorge Guzman, for his support on the computational tools used; e.g., scripts and executables to extract and analyze the model outputs.

Fellow ABE Graduate Students, for their company along these two years and their interest in this research.

Funding for this research was provided by the U.S. Department of Agriculture – Agricultural Research Services Research, education and Economics (REE) Cooperative Agreement (No. 59-3070-5-001). Findings, discussions, and conclusions presented in this document are those of the author(s) and do not necessarily reflect the views of the USDA-ARS.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The physical processes occurring at the watershed scale involve the interaction of many environmental and land management variables (e.g., rainfall, temperature, evapotranspiration, land use, crop) that continuously change in time and space. Understanding the temporal and spatial patterns of these variables is essential to assess and quantify the response of the system under changing conditions. This may be achieved through long-term monitoring and research (Starks, et al., 2014; Tomer & Locke, 2011) that can also support the evaluation of conservation practices and provide data for model development. However, long-term datasets with high spatio-temporal resolution are scarce and expensive to collect (Moriasi et al., 2014a). In addition, the complexity linked to agro-production systems does not allow social and cost-effective experimentation at large scales (Gassman, et al., 2007). A primary approach to assess these problems is through the use of numerical models and small scale experimentation. However, to better assess long-term systemic responses, the evaluation of expected model outcomes must include uncertainty analysis that reveals the full spectrum of system behavior driven by probabilities rather than monotonic evaluation.

The use and development of environmental models (e.g., hydrologic and water quality) implies dealing with uncertainties from different sources. Uncertainty is present in input data (e.g., climate data, elevation, land cover, soil), data processing (e.g., rainfall and runoff data aggregation and interpolation), model parameters, spatio-temporal discretization, model structure, etc. (Guzman, et al., 2015b). By using a model, uncertainties can be propagated non-linearly to the model outputs (e.g., runoff, nutrient concentration, sediment load). Evaluating the uncertainties in the model and how they are propagated to the model outputs will help ensure the reliability of these outputs, and thus build confidence in the model (Chu-Agor et al., 2011). On the other hand, ignoring uncertainty during model development may lead to bias conclusions and perhaps simulated responses outside the expected boundaries of the system being simulated (Ajami et al., 2007). Therefore, results should be presented with the full disclosure of the risks associated with the outputs uncertainty.

Model parameterization that involves calibration may introduce another layer of uncertainty to the results while improving model performance metrics. Multiple acceptable parameter combinations for a set of model inputs may exist that can represent the observed watershed systemic behavior (i.e., Equifinality). Equifinality makes it difficult to determine whether or not the selected set of parameters is the most appropriate to represent the system response. However, Beven (2006) argued that evaluation of equifinality should be given serious consideration not because of the difficulty of identifying parameter values but as an identification of multiple functional hypotheses about how the system is working.

One of the most critical impacts of uncertainty occurs when model results are used to support important environmental decisions and policies. It is crucial that the results are interpreted in light of the risk associated with model output uncertainty (Cariboni et al., 2007; Guzman et al., 2015b). Moreover, quantifying input-output uncertainties and equifinality during model parameterization can cultivate the consciousness of "model accountability" while helping define model outputs' range of validity. Identifying which components, and at what ranges, contribute the most to the uncertainty and equifinality of the model will facilitate parameter adjustments improving model robustness.

CHAPTER 2

OBJECTIVES

The objective of this study was to define a framework that facilitates model parameterization while evaluating uncertainty to assess the impacts of land management practices at watershed scale. A two-step probabilistic approach for model calibration and parametrization was implemented to the Agricultural Policy/Environmental eXtender (APEX) model using global uncertainty and sensitivity analysis. The model was then used to estimate the changes in Nitrogen loads under different land management practices scenarios in the Lake Creek watershed located in south-central Oklahoma, USA. The specific objectives of this study are to:

- Parameterize an APEX model for the Lake Creek watershed using a two-step probabilistic approach using global uncertainty and sensitivity analysis.
- Estimate the full spectrum of total monthly Water Yield and total monthly Nitrogen Loads at the Lake Creek watershed for the period of simulation.
- Simulate different Land Management Practices at the Lake Creek watershed and compare their effect on Nitrogen Loads.

CHAPTER 3

REVIEW OF LITERATURE

The study of water quality impairment due to the use of pesticides and fertilizers in agricultural practices in the United States has been one of the most important concerns of the U.S. Environmental Protection Agency (USEPA), the U.S. Department of Agriculture (USDA) and the U.S. Geological Survey (USGS) (Reibich & Demcheck, 2007). Since the 1930s, the USDA has promoted the implementation of agricultural conservation practices along the country to mitigate land degradation and reduce environmental impacts. Landowners receive financial incentives from USDA to implement these conservation practices. However, there were no controls over the actual environmental benefits of these practices (Tomer & Locke, 2011). Only after 1989 the Management Systems Evaluation Areas program (MSEA) (USDA, 1994) proposed the estimation of conservation practices effects on environmental quality at a field scale but the benefits at the watershed scales remained unknown (Richardson et al., 2008).

Agricultural impacts on water quality need to be addressed at watershed scales since the primary public benefit is observed at that scale (Richardson et al., 2008). However, field experiments to evaluate the effects of conservation practices at watershed scales are hampered by several factors such as private areas of agricultural lands, or the high cost of instrumentation and the long time that they may require. Experimentation at smaller scales (e.g., field scale) is more accessible but results should not be extrapolated to larger scales since large watershed responses are not replicable from field-scale experiments (Tomer & Locke, 2011). Then, computer-based watershed models are commonly used as a cost-effective alternative to simulate the watershed processes (Guzman, et al., 2015). Some of those models are based on empirical equations and may require few input data to represent the response of a watershed. This is the case of APEX model, which was designed to assess the influence of conservation practices on water quality applying equations such as the SCS curve number or the Hargreaves PET. These equations depend on measured precipitation or temperature data which can be retrieved from climate stations. These features make this model easy to apply, but at the same time accurate at watershed scale.

This literature review introduces the concept of hydrological and water quality models, and their application in assessing watershed responses under different conservation management practices. It describes the different types of models used to date, the process of parameterizing a model, and the interpretation of model outputs considering the uncertainty inherent in the modeling process.

3.1 Hydrological and water quality models

Hydrological and water quality models are mathematical and computed-based tools that support the assessment of watershed processes (Moriasi et al., 2007; Tomer & Locke, 2011). These processes include, for example, the estimation of the water yield at the outlet of the watershed and concentration of nutrients and contaminants. As the conditions, land uses, or agricultural practices in watersheds may vary in time and space, the processes in the watershed are also impacted. Then, hydrological and water quality models are also powerful at simulating the effect of land use and agricultural practices changes on hydrology and water quality over long-time periods (Tomer & Locke, 2011). The simulated results may guide environmental decisions and policies to mitigate soil and water degradation by determining suitable conservation programs (Moriasi et al., 2007; Guzman et al., 2015).

Different types of models have been developed according to the different needs in watershed analyses (Arnold et al., 1998). These needs may require a particular time step, different degree of accuracy, or certain spatial resolution. As computer hardware and software have advanced, models became more computationally efficient allowing more detailed assessments that can be continuous in time and space. The use of GIS/spatial interfaces have facilitated the model development process and the input of higher quality inputs, which enabled in simulation of larger areas. This section describes the different types of models available to date and the benefits and drawbacks of developing and using hydrological and water quality models.

3.1.1 Types of models

The classification adopted in this review divides the hydrological and water quality models mainly into two categories: conceptual, and physically-based models. Each category can be subdivided, according to the spatial discretization assumed in the model, into: lumped, semi-distributed, and distributed models. Then, a model can be defined as "conceptual and lumped" or "physically-based and distributed". The two categories and the subdivisions are intrinsically related (Wagener and Gupta, 2005). Conceptual models are those that represent hydrological processes by empirical algebraic equations or differential equations based on simplified hydraulic laws (Arnold et al., 1998). These equations may just simulate the processes as linear and additive, or capture the nonlinear and non-additive character of hydrologic system (Kirchner, 2006). If the processes are assumed linear and additive, as any curve can be considered linear over small segments, these approaches can provide good approximations over short periods of time and give reasonable answers to practical questions. But if the answers for the problems remains beyond the linear domain, these simplifications may provide unreliable predictions (Kirchner, 2006).

Physically-based models are those that represent hydrological processes by differential equations derived from conservation of mass, energy, and momentum laws (Arnold et al., 1998); for example, Darcy's law, Richards equation, or the advection-dispersion equation. The solution for these equations may require the use of grid networks with the spatial distribution of watershed parameters. This means that the data requirements are substantial compared to conceptual models (Jain et al., 1992). However, physically-based models may discretize the system into smaller units and each unit is an individual entity (distributed models), instead of treating the system as a whole and assuming homogeneity (lumped models). Then, the most complex models are both physically-based and distributed, while the simplest are both conceptual and lumped. Yet, using one type of model over other depends mainly on the quantity and quality of data available, and the temporal and spatial dimensions of the problem. Not always a physically-based model is preferred over a conceptual model. Sometimes a conceptual model may solve the problem easier than the physically-based and achieve similar accuracy. This means, both types of models have strengths as well as weaknesses.

3.1.2 Benefits and drawbacks of models

As stated previously, hydrological and water quality models can save time and money for watershed analyses compared to field experiments. They can also assess larger areas that field experiments may hardly cover. However, models cannot replace field monitoring or physical networks that measure the real response of watersheds. Models actually depend on monitoring networks in order to be calibrated and validated. Then, they complement each other. The advances in hydrologic sciences depend on new approaches to modeling hydrologic systems as well as on improvements in the measurement networks (Kirchner, 2006). There is no reason of developing more elaborated models if the monitoring systems are not improved. On the other hand, new methods for analyzing hydrologic data are also needed, especially for analyzing and quantifying the uncertainty of model outputs (Guzman, et al., 2015). Independently of how complex and elaborated the model is, it will always produce output with some level of uncertainty.

Uncertainty is inherent in the model development process and is always propagated to the model outputs. The uncertainty of model outputs can be estimated using statistical metrics that contrast the simulated data against the measured data. This gives an approximate value of how well the model fits the response of the real system. The process of getting an acceptable model fit through the statistical metric, i.e., the calibration process, is carried out by adjusting the different model parameters. This methodology is usually applied as an iterative and operational process that may result in mathematical success, but the internal assumptions and approximations of the model are rarely analyzed (Kirchner, 2006). This means the rationale of model outputs is not always questioned.

Model simplifications should be also considered, firstly because physical laws behind the equations solved in models have been derived for small scales. There is no certainty whether these laws are applicable to larger scales and heterogeneous systems. In fact, physically-based and distributed models can be calibrated to data from one time interval but if they are tested against data from another time interval with different patterns of rainfall, they often perform poorly (Seibert, 2003). This questions the model’s ability to predict the response of the same system under different conditions.

3.2 Model parameterization process

3.2.1 Input data

Hydrological and water quality models are mathematical tools that depend mainly on field data and its quality. Without sufficient and good quality data, a model will never represent properly a particular hydrologic system (Benedini & Tsakiris, 2013). Most of models require at least some climate data (e.g., rainfall and/or temperature), information of soil properties, and land use distribution. The quality of this data relies primarily on the accuracy of measurement instruments and the way the information is processed. The data is commonly measured and collected at some points within the study area and then extrapolated to the whole area. This extrapolation of data, at watershed scales, usually disregards the spatial heterogeneity of the measured variables (Guzman, et al., 2015). Since the hydrological variables are spatiotemporal dependent (Hiebeler & Michaud, 2012), the quality of data is being affected and more uncertainty is being added to this input information.

The input uncertainty is one of the sources of model output uncertainty. The input data uncertainty is propagated and reflected in the level of uncertainty and variability displayed in the model output. The impact that input uncertainty has on the model simulation accuracy is substantial (Moriasi & Starks, 2010); the model may result in simulated responses outside the expected boundaries of the system being simulated or undesired trends (Ajami et al., 2007). Therefore, the quality of observed data should be taken into serious consideration (Harmel et al., 2006). Aside from input data measured on field, models also require the parameterization of some constants defined in the equations internally solved. These values cannot be defined on field measurements, because either it is not possible or they are seldom estimated. Most of them were defined based on laboratory experiments for specific conditions and were included in empirical equations. Therefore, these unknowns are usually parameterized through the model calibration process (Guzman, et al., 2015) and may add another layer of uncertainty. This layer of uncertainty will depend on how accurate the model was calibrated.

3.2.2 Calibration and parameterization

The calibration and parameterization of a hydrological and water quality model is the process of adjustment and estimation of those model parameters that cannot be obtained from field observations and measurements. This process is carried out by comparing the model output with the measured output in the watershed (Moriasi, et al., 2007). The calibrated outputs can be streamflow, contaminant loads or concentrations, suspended sediments, etc. The calibration and parameterization process should follow a logical sequence, i.e. it is nonsense to calibrate the model, for example, for suspended sediment loads first and then for streamflow. It should be calibrated first for streamflow and then for suspended sediment loads, because the computations of suspended sediments depend on the streamflow, while the computations of streamflow are independent of the amount of sediments. Then, the parameters involved in the estimation of streamflow are commonly adjusted first based on approximate and feasible ranges of values often found in literature, or suggested by the user's manual of each particular model, or simply defined using some scientific intuition.

The process of finding the appropriate value for each parameter, and the appropriate combination of those parameter values, in a way that the model output can mimic the real watershed output is usually carried out either by trial and error or by optimization algorithms. Since the 1970s, the Monte Carlo method has been applied in hydrology as an optimization algorithm for model calibration (e.g. Whitehead & Young, 1979; Hornberger & Spear, 1981). General procedures for model calibration have also been reported (e.g. Donigian, 2002; Moriasi et al., 2007). However, there are no universally accepted guidelines for model calibration and parameterization (Moriasi et al., 2012); each user may apply any method or strategy that seems convenient. Furthermore, each user may select a different statistical metric to compare the simulated and measured data, and different criteria for accepting the simulated values. All of this challenges the comparison of different studies.

The statistical metrics commonly used in hydrology for model calibration are: the Nash-Sutcliffe Efficiency Coefficient (NSEC; Nash & Sutcliffe, 1970), the Percent Bias (PBIAS), and the Root Mean Square Error (RMSE). Moriasi et al. (2007) proposed some general performance ratings to classify the statistical metric estimations during model calibration, in an attempt to standardize the criteria for acceptable simulations. They defined four categories of model performance: “Very good”, “Good”, “Satisfactory”, and “Unsatisfactory”. Thus, optimization algorithms simply evaluate one or more of the statistical metrics to find one set of parameters that satisfy the metric criteria. However, it has been accepted that not only one but many different parameter combinations may satisfy the criteria. This issue was called “equifinality” (Beven, 1993). Moreover, the parameter combinations that may acceptably represent the system for the calibrated period, may perform poorly if the watershed conditions are changed. Therefore, the calibration and parameterization of hydrological and water quality model is still a matter of research.

3.3 Model output interpretation

3.3.1 Uncertainty

Every output computed by a hydrological and water quality model is always associated with some degree of uncertainty (Shirmohammadi, et al., 2006). Uncertainty, in this context, can be defined as any deviation of the simulated data from the measured data (Guzman, et al., 2015). Moreover, the variability of the model output is another sign of uncertainty in the computations. Scattered outputs to represent the response of the watershed system may indicate lack of precision in the estimates. Uncertainty in modeling complex systems can arise at any stage of the model development (Kirchner, 2006), since it comes from different sources: input data, data processing, model parameters, spatio-temporal discretization, or model structure. The uncertainty can propagate non-linearly to the model output and be also amplified due to parameterization (Beven, 1993).

Model simulations should always be interpreted in light of the risk associated with the output uncertainty (Cariboni et al., 2007). Even though the model parameters can be adjusted and acceptable model metrics can be obtained, it does not mean that the model is now able to represent any response of the watershed system, and that its results can be extrapolated to predict future scenarios. Hydrological and water quality models usually produce very different estimates for the same hydrological system if the conditions of the system change (Kirchner, 2006). The model output and its accuracy depend always on the particular period of time used for calibration and the conditions establish during that period. Therefore, predictions should always be made after understanding the system behavior under changing conditions, considering that different set of parameters may be needed to represent new scenarios in the watershed, and estimating the output uncertainty in each case.

3.3.2 Equifinality

As it was stated previously, many model parameter combinations may acceptably represent the particular response of a hydrologic system; i.e., equifinality may exist in the modeling process. Then, several sets of parameters may give almost identical fits to the calibration data. Equifinality was firstly introduced into hydrological modeling by Beven (1993). It is explained as the result of the overparameterization of hydrological and water quality models, which means that these models depend on dozens of free parameters to estimate the outputs.

Equifinality hinders the selection of one set of parameters over other that may also represent the system response. However, Beven (2006) argued that evaluation of equifinality should be given serious consideration not because of the difficulty of identifying parameter values but as an identification of multiple functional hypotheses about how the system is working. The consideration of equifinality may also help to assess the uncertainty associated with model predictions. Beven and Binley (1992) proposed a methodology to address equifinality, which was called Generalized Likelihood Uncertainty Estimation (GLUE). This methodology was developed based on empirical studies that obtained good model fits to observed data. On the other hand, Kirchner (2006) claimed that the most appropriate way to solve the problem of equifinality is not by learning how to tackle it, but by reducing the overparameterization of the models. This can be achieved if reduced-form models (i.e., models in terms of very few parameters) are developed. This author cited the example of Jakeman and Hornberger (1993), who showed that a rainfall-runoff time series can be simulated with a model that considers only up to four free parameters.

3.3.3 Simulation of future scenarios

The Conservation Effects Assessment Project (CEAP) was the first project that tried to quantify the effects of land management practices at watershed scales in the United States (Richardson et al., 2008). The main reason behind this initiative was to carry out a national assessment of the benefits that the land management practices, promoted by the USDA, had over the environmental quality. Since at watershed scales field experiments are not feasible, the use of models was increasingly necessary. Then, many efforts were combined to tackle all issues that model development and implementation may have (e.g., output uncertainties, equifinality, loss of accuracy under changing scenarios), in order to make model outputs useful for water quality estimations, comparison of scenarios, and decision-making (e.g. Moriasi et al., 2007; Moriasi et al., 2012; Wagener and Gupta, 2005). If high-quality data is used, and a proper evaluation of the output uncertainty and equifinality is performed, models are then the best representation of the hydrological system available to date. It is also important to analyze the multiple functional hypotheses of each system's behavior and to interpret the output uncertainty in a way that it is clear how reliable the results are. This will definitely support each assessment and the predictions that can be made from the simulations.

CHAPTER 4

METHODOLOGY

4.1 Study area

The Lake Creek watershed is one of the three main sub-watersheds that composed the Fort-Cobb Reservoir Experimental Watershed (FCREW) located in southwestern Oklahoma. It drains an approximate area of 154 km^2 towards the Fort-Cobb reservoir located near the main FCREW outlet (Figure 1) (Guzman et al., 2015a). The FCREW region is mostly agricultural land composed of croplands and pastures. Soils are mostly fine silty loams of different erodibility (Steiner J. L., et al., 2008). The climate in southwestern Oklahoma is sub-humid with long and hot summers, and short and temperate winters. The mean daily temperature during summer is about 28°C while in winter is 3°C. The annual precipitation is approximately 800 mm with the largest monthly average at the end of spring (May-June) and beginning of fall (September-October) (Steiner J. L., et al., 2008).

The Fort-Cobb reservoir is an important source of public and domestic water supply. However, it has been added to the list of water bodies that do not meet the water quality standards as given in the Clean Water Act (Steiner J. L., et al., 2008). The agricultural practices in this watershed release nutrients, especially Nitrogen (N) and Phosphorus (P), to the surface streams that feed the Fort-Cobb reservoir resulting in eutrophication (Steiner J. L., et al., 2008). On the other hand, several agencies such the Oklahoma Water Resources Board, Oklahoma Department of Environmental Quality, and Oklahoma Conservation Commission recognized the FCEW as an experimental land to improve water quality through land conservation practices. In fact, several agronomic management practices have been adopted in the watershed such as no-tillage management, conversion of cropland to grassland, installation of fencing to exclude cattle from streams, and various structural and water management practices (Storm et al., 2006).

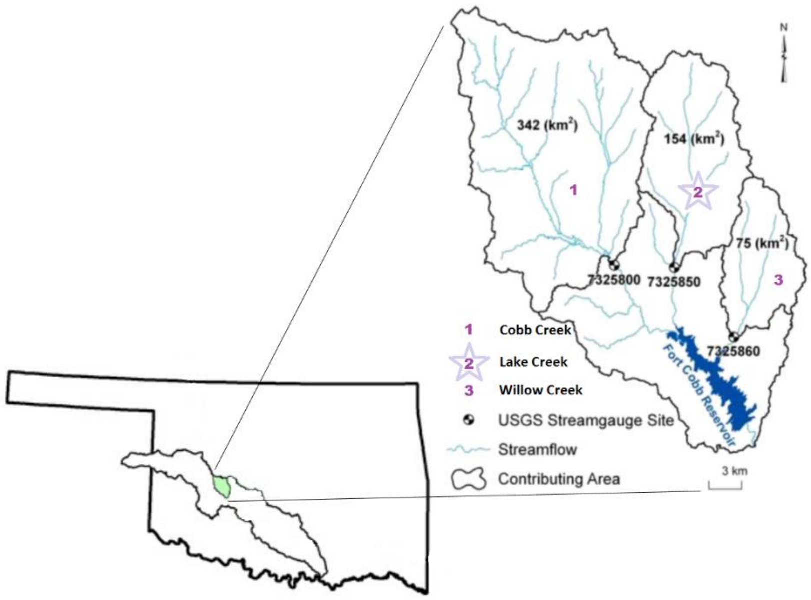

Figure 1. The Lake Creek watershed is one of the three sub-watersheds of the Fort-Cobb Reservoir Experimental Watershed (FCREW) located in southwestern Oklahoma.

4.2 Model set-up

The Agricultural Policy/Environmental eXtender (APEX) model (Williams et al., 1995) is a conceptual and distributed hydrologic and environmental model. It simulates the different hydrologic processes at a watershed scale while evaluating the impacts of conservation and best management practices on water quality (Wang, et al., 2012). The primary inputs to the model are elevation, soil, land use, and time series of climate variables. Outputs are time series of the computed hydrologic variables, nutrients, and crop yields at different time steps (annual, monthly, and daily) and different spatial scales (subareas or watershed).

The main outputs of APEX used to evaluate the impacts of the land management practices were the water yield (WYLD) and the Nitrogen load (N). The WYLD (in mm) was computed in

APEX model using the SCS curve number (CN) equation (USDA-Soil Conservation Service, 1972) given as follows:

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

where P is the daily rainfall (mm), and S is a retention parameter (mm). The parameter S implicitly depends on the curve number (CN) expressed as $S = 254(100/CN - 1)$. The N load (in kg/ha) was computed separately for surface runoff, lateral flow, quick return flow, and horizontal pipe flow (for drains) using the equation:

$$N = W\left(1 - e^{-\frac{Q_i}{kV}}\right) \tag{2}$$

where W is the nitrogen load contained in a layer at the beginning of the day (kg/ha), Q_i is the flow through the layer, V is water storage volume, and k is the fraction of V occupied by percolating water. The Nitrogen load in the stream is the sum of the four components.

In this study, the GIS interface for APEX (ArcAPEX) was used to build the model for the Lake Creek. This interface requires three different data layers: the Digital Elevation Model (DEM), soils, and land use. In addition, weather data (precipitation and temperatures) and information on land management operations were also needed. The DEM and land cover maps were obtained from USGS (USGS, 2016) while the soil map from the Web Soil Survey (USDA, 2016).

Measurement of hydrological variables and water quality in the FCREW began in late 2004, as part of the Conservation Effects Assessment Project (CEAP). This project was designed to estimate the environmental benefits of conservation practices implemented on agricultural lands (Mausbach & Dedrick, 2004). Fifteen climate stations, known as the Micronet stations, were installed in the FCREW to measure weather and soil observations (e.g., rainfall, solar radiation, and air temperature). Daily precipitation and temperature (minimum and maximum) time series for Lake Creek were obtained from these stations (Guzman et al., 2014). Land management data were collected from literature and reports of conservation and management schemes implemented in the watershed from 2005 onward (Storm et al., 2006). Nitrogen concentration collected at USGS streamflow sites was part of a water quality monitoring program in the FCREW for dissolved O_2, pH, total N, total P, and suspended sediment concentration (Starks et al., 2014b; Moriasi et al., 2014). Water yield, WYLD, and total Nitrogen loads, N, were simulated at daily time step for 9 years (2005-2013) and evaluated with observations. However, the first two years of simulations were used to warm-up the model and were not considered in the analysis.

4.3 Global uncertainty and sensitivity analysis

Sensitivity analyses are commonly used to optimize the process of model development (Saltelli, et al., 2008). They facilitate the adjustment of model parameters by establishing their impact on model output. The most common type of sensitivity analysis, the local one-at-a-time approach, is based on derivative and evaluates the impact of changing one parameter on the output while considering a constant value for other parameters. This approach is, however, problematic if the inputs are uncertain or the linearity of the model is unknown since this approach provides information only at the point where they were taken and do not provide for an exploration of the entire input space (Saltelli, et al., 2008). Most of the time, no a priori knowledge of the linearity of the model is available or the linearity can change with changes in model assumptions. As a result, the derivative equation used to quantify sensitivity may no longer apply. In contrast, the global approach explores the entire sampling space of the inputs and evaluates the impact of each parameter by varying simultaneously several parameters over a defined range (Zhou & Lin, 2008). This approach does not require a priori information of the linearity of the model and it also allows evaluation of the interactions among the different parameters.

In this study, a two-step probabilistic approach using global uncertainty and sensitivity analysis (GUSA) was implemented to develop the APEX model for Lake Creek watershed. The global uncertainty analysis quantified the output uncertainties propagated from uncertain APEX parameters. A variance-based sensitivity analysis was then performed to divide the output uncertainty according to the contribution of each of the model parameters (Figure 2). The general process of performing GUSA is described in Figure 2. The first step is to assign probability distribution functions (pdf) to the uncertain model input/parameters. Each pdf is then sampled using a particular method (e.g., Sobol method). These sample points were used to run the model from which a spectrum of model outputs (e.g., discharge, crop yield) can be obtained. Using the variance of the output of interest, sensitivity analysis determines the sensitivity indices of each contributing input/parameter.

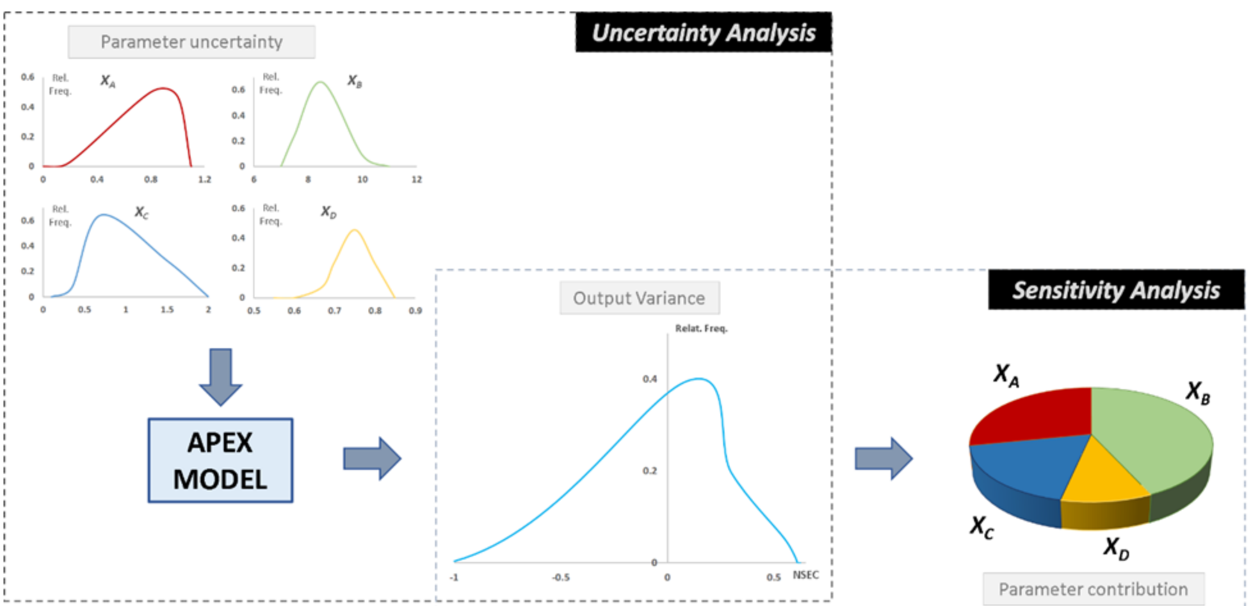

Figure 2. Schematic diagram of the Global Uncertainty and Sensitivity Analysis. Global uncertainty analysis determines the spectrum of model output due to uncertainty in model inputs/ parameters while global sensitivity analysis quantifies the contribution of each model parameter to output uncertainty.

The Sobol method (Sobol, 1993) is a variance-based sensitivity analysis method that evaluates the sensitivity of each parameter based on the principle of variance decomposition. The sensitivity analysis is computed via the first sensitivity index (S_i) that represents the direct contribution of each parameter to the variance of the output. It is expressed as:

$$S_i = v_i / V \tag{3}$$

where v_i is the part of the variance due to the input parameter x_i, and V is the total variance of the model output. This variance-based analysis is also capable of quantifying the influence of the full range of variation of each parameter and their interaction effects (Saltelli, et al., 2008). On the other hand, the difference $100-\sum S_i$ can be used as an indicator of the linearity of the model. If this difference is equal to 0, this means that the model is linear, otherwise it is nonlinear. The total effect index (S_{Ti}) accounts for the total contribution to the outputs variation of each input parameter x_i, i.e., its first-order effect (S_i) plus all the higher-order effects due to interactions (Saltelli, et al., 2008). For example, for three input parameters x_1, x_2, and x_3, the S_{Ti} of parameter x_i can be expressed as:

$$S_{Ti} = S_1 + S_{12} + S_{13} + S_{123} \tag{4}$$

where S_{T1} is the total sensitivity index of x_i, S_1 is the first-order effect of x_1, S_{12} is the interaction effect between x_1 and x_2, and S_{123} is the interaction effect between x_1, x_2, and x_3. Using equation (2), S_{T1}–S_1 provides a measure of how much x_1 is involved in interactions with any other input factors (Saltelli, et al., 2008). The number of sample points (N) required for Sobol method is given as:

$$N = M(2k + 2) \tag{5}$$

where M is the sample size of each index, typically taken between 500 and 1000 (Chu-Agor et al., 2011; Chu-Agor et al., 2012), and k is the number of parameters.

Upon evaluation of the different parameters in APEX, 44 parameters were found to be uncertain and independent and were used to estimate water yield, WYLD, (i.e., the runoff) (Table 1) and 15 to estimate Nitrogen loads (Table 2). The Nash-Sutcliffe Efficiency Coefficient (NSEC) was used as the main output of interest and as a metric in filtering the simulations that are considered acceptable. The global uncertainty and sensitivity analysis were performed using SimLab (version 2.2), a software designed for Monte-Carlo based uncertainty analysis (Saltelli et al., 2004).

Table 1. APEX parameters for water yield (WYLD) simulation

No.	Parameter	Description	Distribution[a]
1	acsf	Adjusts climatic stress factor	U(40-100)
2	BWD	Channel bottom width/depth (m/m)	U(0-40)
3	CHSO	Average upland slope in watershed (m/m)	U(0.0001-1)
4	cmeq	Coefficient in MUST equation	U(1-4)
5	cnrp	Expands CN Retention Parameter	U(0-3)
6	dlhc	Estimates drainage system lateral hydraulic conductivity	U(0.00001-20)
7	DTHY	Time interval for flood routing (hours)	U(0.1-24)
8	ecri	Exponential coefficient used to account for rainfall intensity on CN	U(0-3)
9	FCW	Floodplain width/channel width (m/m)	U(1-100)
10	fevl	Flood evaporation limit	U(0.001-2)
11	FPSC	Floodplain saturated hydraulic conductivity (mm/hr)	U(0.0001-20)
12	GWSO	Maximum groundwater storage (mm)	U(0-200)
13	gwst	Groundwater storage threshold	U(0.001-1)
14	hdvp	Hydrograph development parameter	U(0.01-2)
15	hpec	Hargreaves PET equation coefficient	U(0.0023-0.0032)
16	hpee	Hargreaves PET equation exponent	U(0.4-0.8)
17	lwtm	Limits daily water table movement	U(0.001-1)
18	QCF	Exponent in watershed area flow rate equation	U(0.1-1)
19	QG	Channel capacity flow rate (mm/hr)	U(0.1-200)
20	QTH	Routing threshold (mm)	U(0-500000)
21	rcfc	RUSL C-factor coefficient	U(0.1-2)
22	rcfx	RUSL C-factor coefficient 2	U(0.1-2)
23	rcia	Runoff CN Initial Abstraction	U(0.01-1)

[a] Probability distribution function and its parameters: U = Uniform distribution (left boundary, right boundary).

Table 1 (cont.). APEX parameters for water yield (WYLD) simulation

No.	Parameter	Description	Distribution[a]
24	rfic	Rainfall interception coefficient	U(0.01-0.5)
25	RFPO	Return flow/(return flow + deep percolation)	U(0-1)
26	RFTO	Groundwater residence time (days)	U(0-365)
27	rgss	Root Growth Soil Strength	U(0-3)
28	ripc	Maximum rainfall interception by plant canopy (mm)	U(0-30)
29	rrap	Runoff CN Residue Adjustment Parameter	U(0-0.5)
30	rrfs	Reduces NRCS runoff CN retention parameter for frozen soil	U(0.01-1)
31	rvad	Runoff volume adjustment for direct link	U(0.1-2)
32	scsc	SCS CN index coefficient	U(0.1-3)
33	secf	Soil Evaporation plant Cover Factor	U(0-1.5)
34	sevc	Soil Evaporation Coefficient	U(0.1-5)
35	ssff	Subsurface flow factor	U(1-10000)
36	STND	VSC routing used when storage > standing	U(0-500000)
37	swll	Soil Water Lower Limits	U(0-1)
38	swul	Soil water upward flow limit	U(0.01-2)
39	ulrp	Upper limit of CN retention parameters	U(0.1-4)
40	wdrm	Winter dormancy	U(0-1)
41	wswc	Water stress weighting coefficient	U(0-2)
42	wtrc	Water table recession coefficient	U(0.001-10)
43	wtre	Water table recession exponent	U(0.1-2)
44	YWI	Number of years maximum monthly 0.5 hour rainfall available (years)	U(0-50)

[a] Probability distribution function and its parameters: U = Uniform distribution (left boundary, right boundary).

Table 2. APEX parameters for Nitrogen load (N) simulation

No.	Parameter	Description	Distribution[a]
1	BME	Biological mixing efficiency	U(0-1)
2	DNSW	denitrification soil-water threshold	U(0.9-1.1)
3	MDBM	Max. depth for biological mixing (m)	U(0-1)
4	NERC	N enrichment ratio coefficient for routing	U(0-1)
5	NERE	N enrichment ratio exponent for routing	U(0-1)
6	NFIX	Nitrogen fixation	U(0-1)
7	NLR	Nitrate leaching ratio	U(0-1)
8	PLR	Pesticide Leaching Ratio	U(0-1)

[a] Probability distribution function and its parameters: U = Uniform distribution (left boundary, right boundary).

Table 2 (cont.). APEX parameters for Nitrogen load (N) simulation

No.	Parameter	Description	Distribution[a]
9	PNFG	Partitions Nitrogen flow from groundwater	U(0-20)
10	SRTC	Sediment routing travel time coefficient	U(0.5-10)
11	ULDN	Upper Limit of Daily Denitrification rate	U(0.0001-0.5)
12	ULNV	Upper limit of Nitrification/Volatilization	U(0-1)
13	UNFL	Upper Nitrogen Fixation Limit (kg/ha/day)	U(0.1-20)
14	VNPC	Volatilization/nitrification partitioning coefficient	U(0-1)
15	WSNL	Water Storage N Leaching	U(0-1)

[a] Probability distribution function and its parameters: U = Uniform distribution (left boundary, right boundary).

4.4 Modeling framework

A modeling framework was developed to facilitate model parameterization and validation. The framework was divided into three main parts: (I) parameterization for hydrology, (II) parameterization for contaminants (or nutrients), and (III) definition of the baseline model (Figure 3). Each part transitions to the next part through global uncertainty and sensitivity analysis (GUSA). The GUSA is used to define the variability of the outputs, identify the most important parameters that drive this variability, and identify the simulations that satisfied the requirements for acceptable model performance. The starting ranges for hydrology parameters were determined based on acceptable default values, field estimates, or as reported from the literature. Uniform probability distribution functions were assigned to these parameters when only the base value was known, the range was considered finite, and no explicit knowledge of the distribution was available (McKay, 1995). This conservative assumption allows an equal probability of occurrence of the parameters along the probability range (Muñoz-Carpena et al., 2010).

The first GUSA analysis (GUSA 1 in Figure 3) was performed to identify the most important hydrology parameters that controlled the variability of the hydrology output. The ranges of these important parameters were further narrowed down by filtering the simulations that resulted in acceptable model outcomes based on metrics used to evaluate the model (e.g., NSEC). The ranges of parameters from these simulations were used as basis in defining the new narrowed ranges. The important parameters identified in GUSA 1 plus the starting nutrient parameters were used in the second GUSA run (GUSA 2; Figure 3). From GUSA 2, new list of important parameters and their new probability distribution functions were identified and used to define the baseline model. The baseline model consisted of several APEX models, each of which, has different parameter combination that resulted in acceptable model performance. Using the baseline models, WYLD and N were simulated resulting in a family of time series each one representing the possible systemic response of the watershed.

The development of the APEX model for the Lake Creek watershed was performed using the following steps:

• Model run 1: Global uncertainty analysis of APEX was performed using the 44 WYLD parameters (Table 1) to determine the spectrum of NSEC and identify the most important (sensitive) parameters for WYLD (i.e., hydrology).

• Model run 2: The ranges of the important parameters identified in run 1 were adjusted based on the acceptability of the model (i.e. NSEC > 0.5). Global uncertainty analysis was then performed using the adjusted important parameters plus the 15 uncertain Nitrogen parameters (Table 2) to determine the spectrum of NSEC for both WYLD and N. The new set of important parameters were also determined.

• Model run 3: The baseline model was defined using the important parameters for both WYLD and N and the updated probability distribution functions (ranges and distribution) of these important parameters.

4.5 Land management practices scenarios

The land management scenarios adopted in the study were divided into two general categories: change in management operation and change in land use (Table 3). The impacts of land management practices were quantified using six additional scenarios, two with changes in management operations (scenarios 2 and 3) and four with changes in land use (scenarios 4-7). The baseline model (scenario 1) was defined according to the land use characteristics presented in Moriasi et al. (2014a).

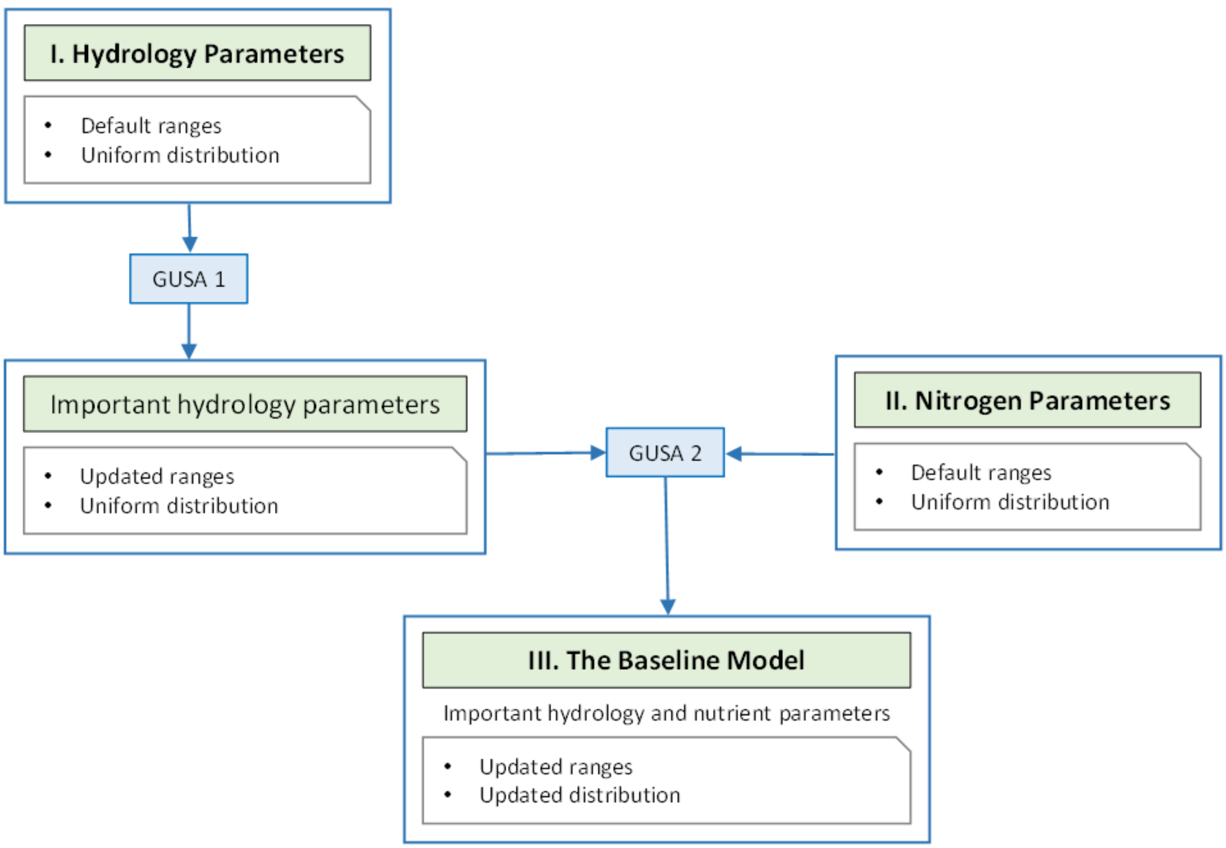

Figure 3. Modeling framework used to parameterize the APEX model. The important hydrology parameters were first identified through GUSA 1. These important parameters were then combined with the starting nutrient parameters for GUSA 2. The final list of important parameters and their updated pdf were used to define the baseline model.

The main land uses in the Lake Creek watershed were winter wheat crops and pasture. About one half of the watershed area is covered with winter wheat crops and the other half with pasture. For this reason, 50% of the subareas in the APEX model were assigned with winter wheat and the other 50% with pasture. The land management practices defined for the baseline model were based on a survey of agricultural land types and practices in the FCREW carried out by Storm at al. (2006). For the winter wheat, the crops from previous harvests were killed in June and then the land was prepared for new planting using tandem disk plows. In august, approximately 67 kg/ha of N (46-0-0) and 4.5 Kg/ha of P205 were applied to the crops (winter wheat). Planting takes place in the middle September. For pasture, grazing activities take place all year around.

In APEX, the agricultural management practices were simulated through the Operational Schedule file (*.OPS). APEX generates as many OPS files as there are crops defined in the model set-up process. Each file contains the dates and equipment used in each stage of crop growth and management (e.g., planting, fertilizer application, and harvesting). The two alternative management scenarios (scenarios 2 and 3) started also in June with the killing of the previous winter wheat crops followed by the land preparation for new planting. In scenario 2 (conventional), the land was tilled in early July using springtooth harrows and the same fertilizer and amount used in scenario 1 (approximately 67 Kg/ha of N (46-0-0) and 4.5 Kg/ha of P205) was applied in August. Planting occurred also in the middle September. In scenario 3 (graze out), the land was tilled in early July using springtooth harrows. Fertilizer application and planting followed the same schedule and amounts used for scenario 2, however, there was an extra fertilizer application of approximately 45 kg/ha, N (46-0-0), on the second week of February. In the case of subareas with pasture, grazing activities was assumed over the entire year for both scenarios (2 and 3).

The other four alternative scenarios (scenarios 4-7) were defined considering the same management practices for winter wheat and pasture used in the baseline model, but different land use distribution. The land use percentages were redistributed into the subareas (Table 3). The OPS of the baseline models was replaced with that of the individual scenarios to simulate the effects of different land management practices on WYLD and N.

Table 3. Summary of the land management scenarios used.

No.	Scenarios	% of pasture	% of winter wheat	Tillage practices
1	Baseline	50	50	conservation
2	Conventional	50	50	conventional
3	Graze out	50	50	conventional
4	75% Pasture	75	25	conservation
5	75% Winter Wheat	25	75	conservation
6	100% Pasture	100	0	conservation
7	100% Winter Wheat	0	100	conservation

The OPS files of the baseline models were modified to incorporate the land management practices implemented for scenarios 2 and 3 while retaining their land use distribution. For scenarios 4 to 7, the OPS file of the baseline models were left unchanged while the land use distribution was modified accordingly. Each scenario, like the baseline models, is composed of several parameter combination models.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Model calibration and parameterization

The probabilistic approach for model calibration and parameterization, described in Figure 3, was divided into three model runs. For the first model run (model run 1) the model was run varying the 44 parameters considered for WYLD (Table 1) resulting in 46,080 simulations. The output of interest (i.e. NSEC) was computed for each simulation to generate the full spectrum of model performance (Figure 4). This spectrum ranged from negative NSEC values up to a maximum of 0.51. The contribution of each parameter to the variability of the output was also computed using the Sobol's 1st order sensitivity index (S_i) (Table 4). From the 44 parameters, only 20 of them directly influenced the estimation of the water yield, WYLD. The Expands CN Retention Parameter (cnrp) was found to be the most important parameter contributing approximately 50% to the model output variability over a total of 79%. The sum of all S_i was less than 100% which indicates that the models are non-linear and hence, interactions between parameters exist.

The water yield, WYLD, in APEX model is computed using the SCS curve number (CN) equation (USDA-Soil Conservation Service, 1972). This empirical equation is in terms of the daily rainfall and the CN retention parameter. The role of the "cnrp" parameter (i.e., the most sensitive in this case) in the SCS CN equation is to expand the CN retention parameter. Since the computed WYLD is inversely proportional to the retention parameter, values of cnrp greater than 1.0 will reduce the water yield. The initial range of cnrp considered in run 1 was 0-3 (Table 1) however, simulations with NSEC greater than 0.4 had cnrp values between 1.5 and 1.8.

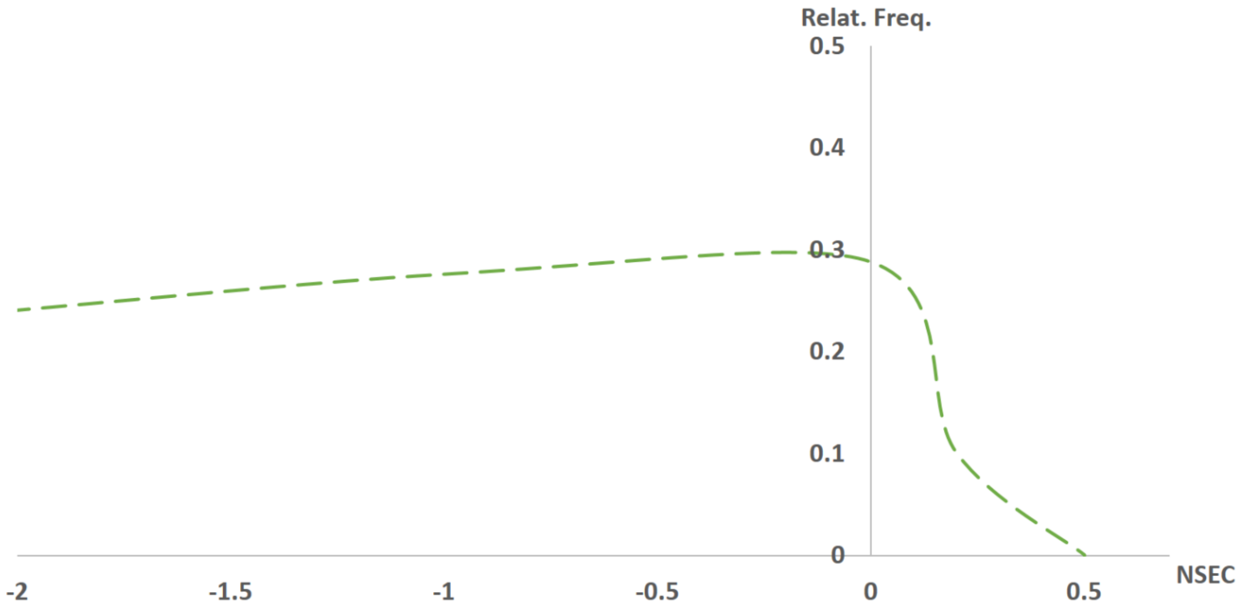

Figure 4. Model spectrum performance for daily Water Yield in model run 1.

The variance-based analysis showed the presence of interactions between parameters and the non-linear property of the APEX model; i.e., the empirical equations (e.g. SCS CN equation) used in the model are capturing the non-additive behavior of the hydrologic system at the Lake Creek watershed. This behavior is expected in most watershed response since the hydrologic system is not linear and additive at all.

Table 4. Direct contribution of the WYLD parameters.

No.	Parameter	S_i	No.	Parameter	S_i
1	cnrp	48.6%	23	rcfx	0%
2	ulrp	9.5%	24	CHSO	0%
3	rgss	9.3%	25	acsf	0%
4	scsc	2%	26	cmeq	0%
5	wswc	1.8%	27	dlhc	0%
6	rrap	1.6%	28	DTHY	0%
7	hpee	1.2%	29	hdvp	0%
8	rfic	0.97%	30	lwtm	0%

Table 4 (cont.). Direct contribution of the WYLD parameters.

No.	Parameter	S_i	No.	Parameter	S_i
9	YWI	0.76%	31	QTH	0%
10	rcia	0.71%	32	rrfs	0%
11	ripc	0.66%	33	rvad	0%
12	hpec	0.6%	34	STND	0%
13	ecri	0.56%	35	swul	0%
14	FCW	0.52%	36	wtrc	0%
15	sevc	0.35%	37	wtre	0%
16	FPSC	0.12%	38	gwst	0%
17	secf	0.09%	39	RFPO	0%
18	ssff	0.04%	40	wdrm	0%
19	RFTO	0.02%	41	GWSO	0%
20	swll	0.01%	42	BWD	0%
21	rcfc	0%	43	fevl	0%
22	QCF	0%	44	QG	0%
TOTAL		**79.2%**			

For the second run (model run 2), the 20 WYLD sensitive parameters were considered along with the 15 Nitrogen parameters (Table 2) to simulate the N loads. In total, 35 parameters were sampled for this second run requiring a minimum of 36,864 simulations. The ranges of the 20 WYLD parameters were redefined according to the model performance obtained in run 1. From the full spectrum, almost the 10% of the simulations (4,701/46,080) had a NSEC≥0, however, only 2 of them reached NSEC>0.5. These two sets of parameters were then used to narrow down the initial wide ranges of the 20 WYLD parameters (Table 5). The ranges were redefined towards those values that produced higher NSEC. The original uniform distribution of these parameters was kept.

Table 5. Redefinition of sensitive WYLD parameters ranges.

No.	Parameter	Initial Range	New Range
1	cnrp	0 - 3	1.3 - 2
2	ulrp	0.1 - 4	2 - 3.2
3	rgss	0 - 3	0.4 - 0.75
4	scsc	0.1 - 3	0.3 - 0.6
5	wswc	0 - 2	1.5 - 1.7
6	rrap	0 - 0.5	0.3 - 0.5
7	hpee	0.4 - 0.8	0.55 - 0.8
8	rfic	0.01 - 0.5	0.01 - 0.2
9	YWI (years)	0 - 50	15 - 40
10	rcia	0.01 - 1	0.65 - 0.85
11	ripc (mm)	0 - 30	2.5 - 26
12	hpec	0.0023 - 0.0032	0.0023 - 0.003
13	ecri	0 - 3	0.1 - 1.5
14	FCW (m/m)	1 - 100	15 - 100
15	sevc	0.1 - 5	1.8 - 3.8
16	FPSC (mm/hr)	0.0001 - 20	7 - 10
17	secf	0 - 1.5	1.2 - 1.5
18	ssff	1 - 10000	1500 - 10000
19	RFTO (days)	0 - 365	150 - 200
20	swll	0 - 1	0.1 - 1

In model run 2, the model performance (NSEC) was evaluated separately for WYLD and N. After narrowing down the ranges of the 20 WYLD parameters, the highest NSEC value computed for daily WYLD estimation improved from 0.51 to 0.63. More than 5,800 simulations (near 15%) obtained NSEC greater than 0.4 (Figure 5). In addition, 134 simulations (0.36%) had NSEC>0.6. The strategy of narrowing the parameter ranges down produced an increase of the model equifinality but a substantial improvement of the model performance for water yield estimation. As a result, there were more than 100 simulations that acceptably represent the hydrologic outcome for the period of simulation. However, the variability of the spectrum is high (i.e., the NSEC values were distributed from -8.0 to 0.63) with more than 10,000 simulations (approximately 28%) resulted in NSEC lower than -2.0.

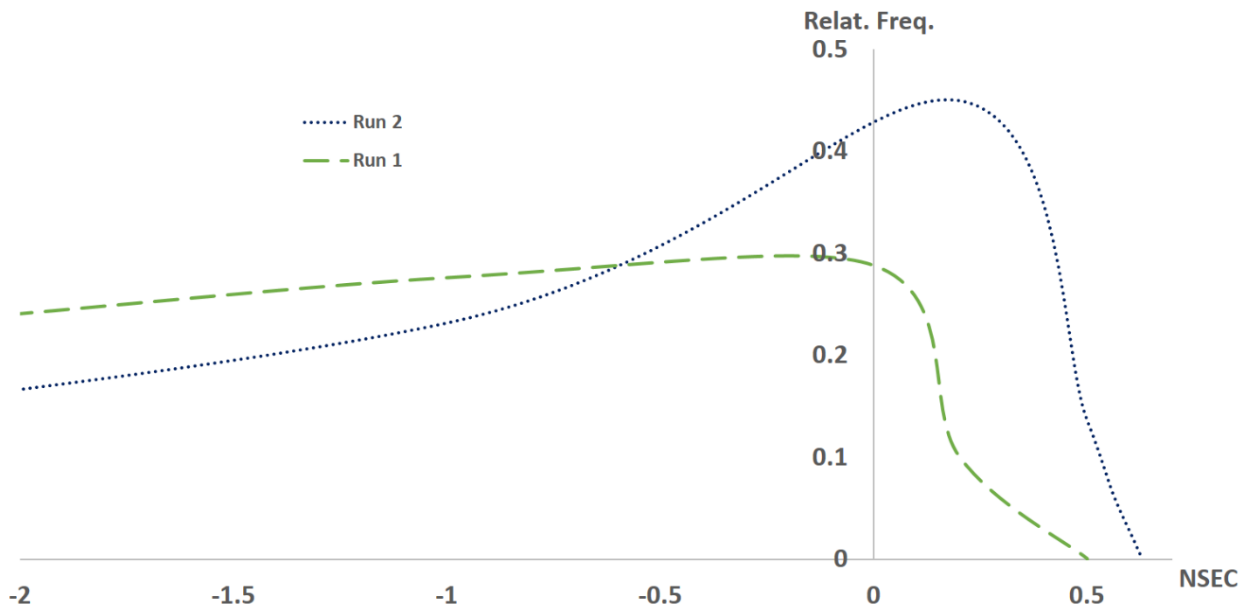

Figure 5. Comparison of model spectrum performance for daily Water Yield between run 1 and 2.

In the case of Nitrogen, the variability of the spectrum (Figure 6) was even higher than the WYLD's spectrum in model run 2. More than 27,000 simulations (75%) obtained NSEC lower than -2.0. However, 4,600 (12.5%) had positive NSEC and, among them, 163 simulations (0.44%) were acceptable (i.e., reached NSEC>0.6) with the highest metric of 0.74. This high variability was expected since the 15 N parameters were sampled with the wide initial ranges and were varied along with the 20 WYLD parameters. The large variability in the input was propagated to the outputs through the model.

A new sensitivity analysis was performed using the Nitrogen spectrum (Table 6). Only 10 parameters directly influenced the variability of the N load estimation. Out of the 10 parameters, just three were exclusively N parameters. These three parameters were the Nitrate Leaching

Ratio (NLR), which is the ratio of nitrate concentration in surface runoff to nitrate concentration in percolate; the Volatilization/nitrification partitioning coefficient (VNPC), which is the fraction of process allocated to volatilization; and the Upper Limit of Daily Denitrification rate (ULDN), which is the maximum fraction of NO3 in a soil layer subject to denitrification. From these three, the NLR had the greatest contribution to the N load output (approximately 6%). This means that the partition between surface runoff and percolation is partly driving the estimation of N loads. However, the two parameters that mostly influenced the N model output were the Hargreaves PET equation parameters (hpee and hpec). They controlled more than the 60% of output variability over a total of 73%.

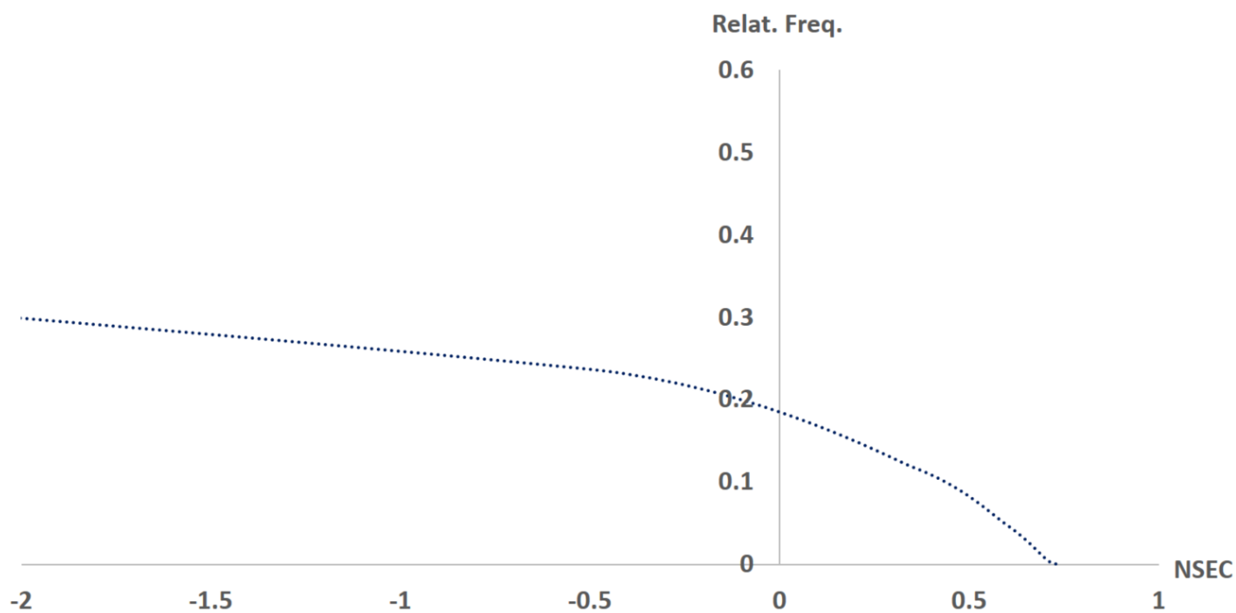

Figure 6. Model spectrum performance for monthly Nitrogen in model run 2.

In APEX model, the Hargreaves method (Hargreaves & Samani, 1985), which estimates the potential evapotranspiration as a function of the solar radiation and air temperature, was modified for the local US conditions by replacing the fixed temperature difference exponent (usually used as 0.5) with the “hpee” and the incoming solar radiation with the “hpec” parameters. In APEX high values of hpee and hpec result in high evapotranspiration. Steglich & Williams (2013) suggested a range of 0.5-0.6 for hpee and 0.0023-0.0032 for hpec. The suggested range for hpec was kept but the range of hpee was widen to 0.4-0.8 in model run 1 (Table 1) and narrowed down to 0.55-0.8 for model run 2 (Table 5). The acceptable simulations in model run 2 showed that the hpee values were mostly ranged from 0.7 to 0.8. This means that the evapotranspiration should be higher to fit the observed N loads. The reason for this is, since plants take up nitrogen during transpiration, the nitrogen consumed should be proportional to the transpiration; i.e., the higher the transpiration is, the more nitrogen is taken up.

Table 6. Direct contribution of the 20 WYLD sensitive + 15 Nitrogen parameters at computing N loads.

No.	***Parameter***	***S_i***	***No.***	***Parameter***	***S_i***
1	hpee	49.8%	19	NFIX	0%
2	hpec	14%	20	PLR	0%
3	NLR	5.8%	21	UNFL	0%
4	scsc	1.6%	22	DNSW	0%
5	cnrp	1.2%	23	SRTC	0%
6	secf	0.1%	24	NERC	0%
7	rfic	0.075%	25	NERE	0%
8	rcia	0.067%	26	FPSC	0%
9	VNPC	0.02%	27	WSNL	0%
10	ULDN	0.01%	28	RFTO	0%
11	MDBM	0%	29	FCW	0%
12	rgss	0%	30	ssff	0%
13	BME	0%	31	ecri	0%
14	YWI	0%	32	swll	0%
15	wswc	0%	33	ripc	0%
16	PNFG	0%	34	rrap	0%

Table 6 (cont.). Direct contribution of the 20 WYLD sensitive + 15 Nitrogen parameters at computing N loads.

No.	***Parameter***	***S_i***	***No.***	***Parameter***	***S_i***
17	ULNV	0%	35	sevc	0%
18	ulrp	0%			
TOTAL		**72.8%**			

The two sensitivity analyses performed at this point were used to determine the ranges of input/parameters to be used in deriving the baseline model for the Lake Creek watershed. 20 WYLD parameters were identified in model run 1, and 10 Nitrogen parameters were added in model run 2. Of the 10 N parameters, 7 also belong to WYLD (Figure 7) and only three were exclusive to Nitrogen. The variability of the system response, as measured by NSEC, is controlled by 23 parameters (Figure 7). This indicated that the rest of the WYLD and N parameters (36 parameters) can take the default values without affecting the outputs (i.e., NSEC).

The original 59 parameters (44 WYLD + 15 N) were sampled using a uniform distribution. This distribution was used since only the parameter ranges were known. After two runs, some parameter combinations have resulted in NSEC$\geq$0.6 for both WYLD and N load estimations. Considering only these acceptable simulations (i.e. 134 simulations for WYLD and 163 for N; some of them in common), the values of the 23 final model parameters were then plotted separately and their *pdfs* were then constructed (Figure 8). Most of the identified distributions were either normal or triangular distributions (Table 7).

For the third run (model run 3) the 23 parameters (24,576 simulations) were sampled considering the updated distributions listed in Table 7. In this run, the NSEC was estimated separately for WYLD and N, as in run 2. The model performance spectrums can be found in Figures 9 and 10 respectively.

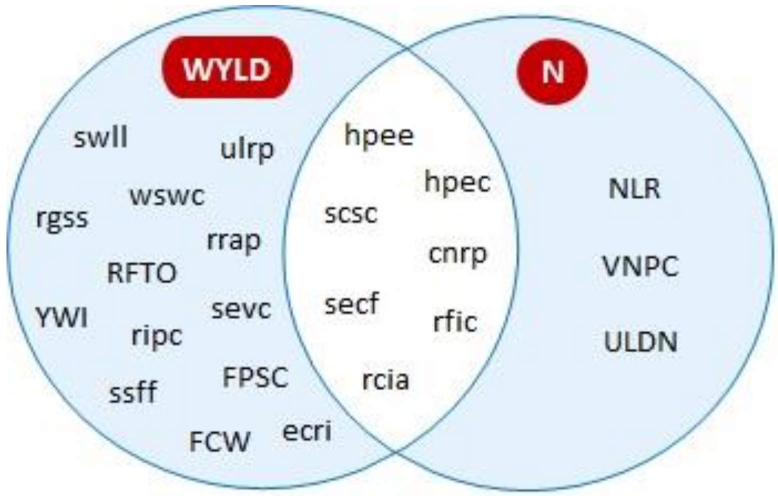

Figure 7. Parameters shared by WYLD and N computations and considered for the baseline model.

In the case of WYLD, there was a reduction in the output variability and negative NSEC values. In this run, only 0.8% of the simulations obtained a NSEC≤ -1.0 as opposed to run 2 where 33% of its simulations had NSEC≤ -1.0. Also the equifinality (i.e., number of acceptable models) was reduced. In run 2, 0.36% of the simulations obtained a NSEC≥0.6 while in run 3 only 0.2% reached that value. In the case of Nitrogen, the reduction in the output variability and negative NSEC values was even higher than in WYLD (run 3). In model run 3, 4% of the simulations obtained a NSEC≤ -1.0 while in run 2 more than 78% of the simulations had NSEC≤ -1.0. Also, the model performance was improved from NSEC=0.74 in run 2 to a highest NSEC=0.78 in run 3.

Model run 3 produced 49 acceptable simulations for WYLD and 853 for N. Since the baseline model will be used to estimate both water yield and nitrogen loads at different land management scenarios, it was necessary to identify the simulations that can predict both WYLD

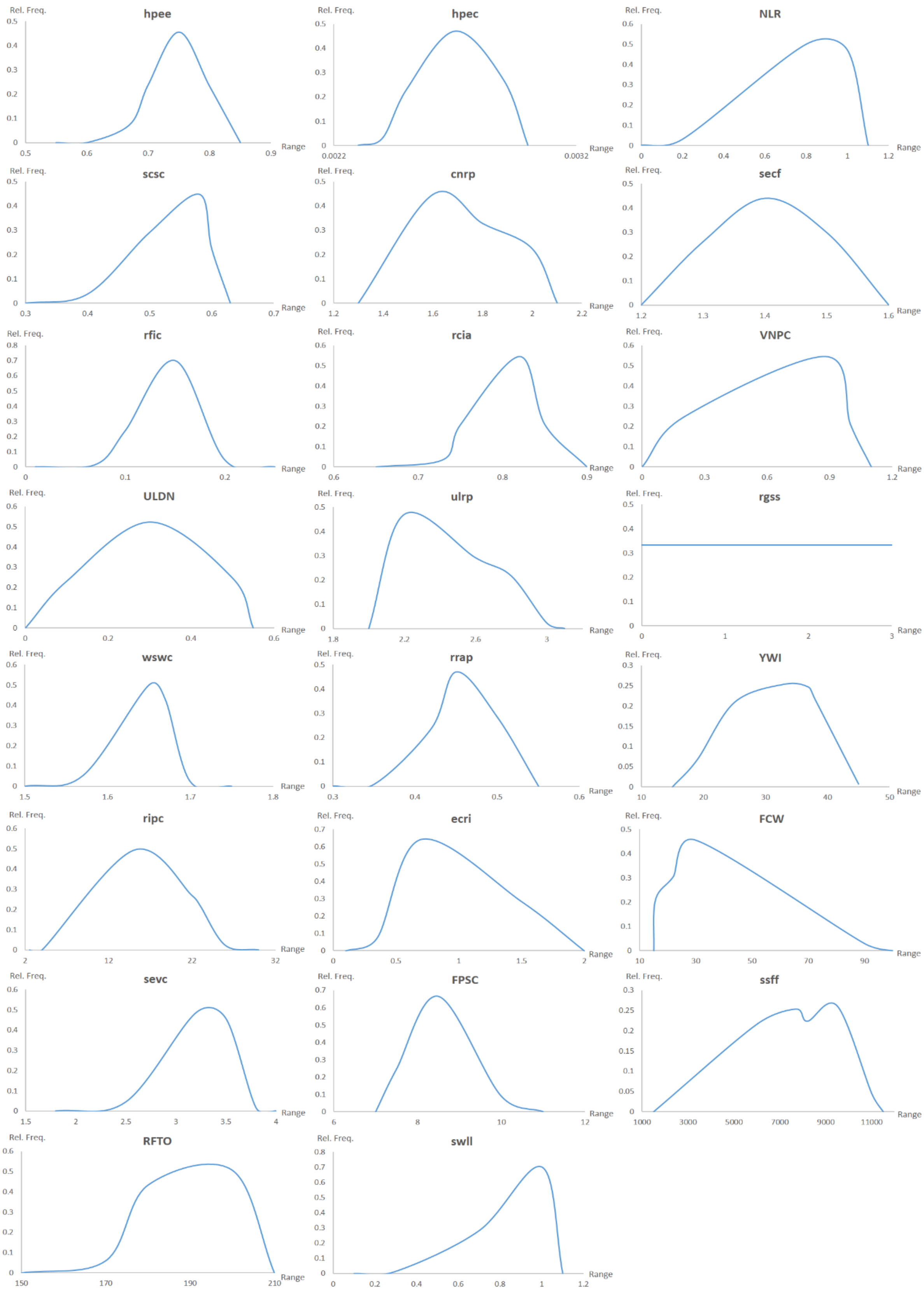

Figure 8. Parameter distribution for Baseline model.

Table 7. Final parameter distribution for Baseline model.

No.	Parameter	Description	Distribution[a]
1	hpee	Hargreaves PET equation exponent	N(0.76, 0.04)
2	hpec	Hargreaves PET equation coefficient	T(0.0024, 0.0027, 0.003)
3	NLR	Nitrate leaching ratio	T(0.015, 0.9, 1.1)
4	scsc	SCS CN index coefficient	T(0.38, 0.58, 0.63)
5	cnrp	Expands CN Retention Parameter	T(1.3, 1.65, 2.1)
6	secf	Soil Evaporation plant Cover Factor	T(1.2, 1.4, 1.6)
7	rfic	Rainfall interception coefficient	N(0.15, 0.02)
8	rcia	Runoff CN Initial Abstraction	T(0.7, 0.83, 0.89)
9	VNPC	Volatilization/nitrification partitioning coefficient	T(0, 0.9, 1.1)
10	ULDN	Upper Limit of Daily Denitrification rate	T(0, 0.3, 0.55)
11	ulrp	Upper limit of CN retention parameters	T(2, 2.3, 3.1)
12	rgss	Root Growth Soil Strength	U(0, 3)
13	wswc	Water stress weighting coefficient	N(1.65, 0.03)
14	rrap	Runoff CN Residue Adjustment Parameter	T(0.35, 0.45, 0.55)
15	YWI	Number of years maximum monthly 0.5 hour rainfall available (years)	N(33, 5)
16	ripc	Maximum rainfall interception by plant canopy (mm)	T(4, 15, 26)
17	ecri	Exponential coefficient used to account for rainfall intensity on CN	T(0.4, 0.7, 2)
18	FCW	Floodplain width/channel width (m/m)	T(15, 30, 95)
19	sevc	Soil Evaporation Coefficient	N(3.4, 0.2)
20	FPSC	Flood saturated hydraulic conductivity (mm/hr)	N(8.5, 0.8)
21	ssff	Subsurface flow factor	T(1500, 9000, 11500)
22	RFTO	Groundwater residence time (days)	N(190, 8)
23	swll	Soil Water Lower Limits	T(0.3, 1, 1.1)

[a] Approximate distributions and their parameters: N = Normal distribution (mean, standard deviation); T = Triangular distribution (minimum, peak, maximum); U = Uniform distribution (left boundary, right boundary).

and N with the same parameter combinations; i.e., the simulations in common for water yield and nitrogen. Due to the differences in sensitive parameters between the two estimations (e.g. cnrp controlled WYLD while hpee controlled N) the acceptable WYLD simulations did not match the acceptable N ones at all. Most of the acceptable N simulations (i.e., those with NSEC>0.6) were related to WYLD simulations with NSEC between 0.4 and 0.5. On the other hand, most of the acceptable WYLD simulations produced N outcomes with NSEC between 0.5 and 0.6. Therefore, the criteria for acceptable N simulations was redefined to NSEC>0.5. Thus, those simulations that simultaneously estimated the WYLD with NSEC>0.6 and the N loads with NSEC>0.5 were accepted as the baseline model. In summary, 27 parameter combinations satisfied this condition.

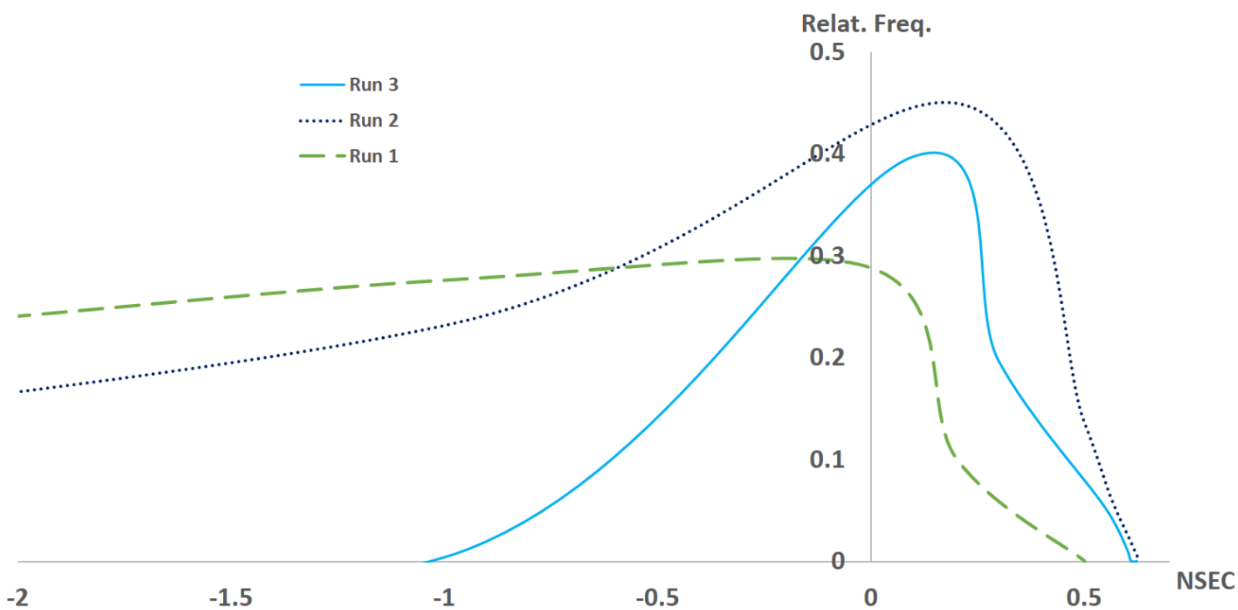

Figure 9. Model spectrum performance for daily Water Yield in model run 3 in comparison with runs 1 and 2.

5.2. Output uncertainty and model estimates

Once the Lake Creek watershed model was calibrated using the two-step probabilistic approach, the estimation of water yield (WYLD in mm) and Nitrogen loads (N in kg/ha) was carried out. The baseline model of Lake Creek was composed of 27 different acceptable parameter combinations, i.e., the total model output consisted of 27 time series of WYLD and total Nitrogen loads.

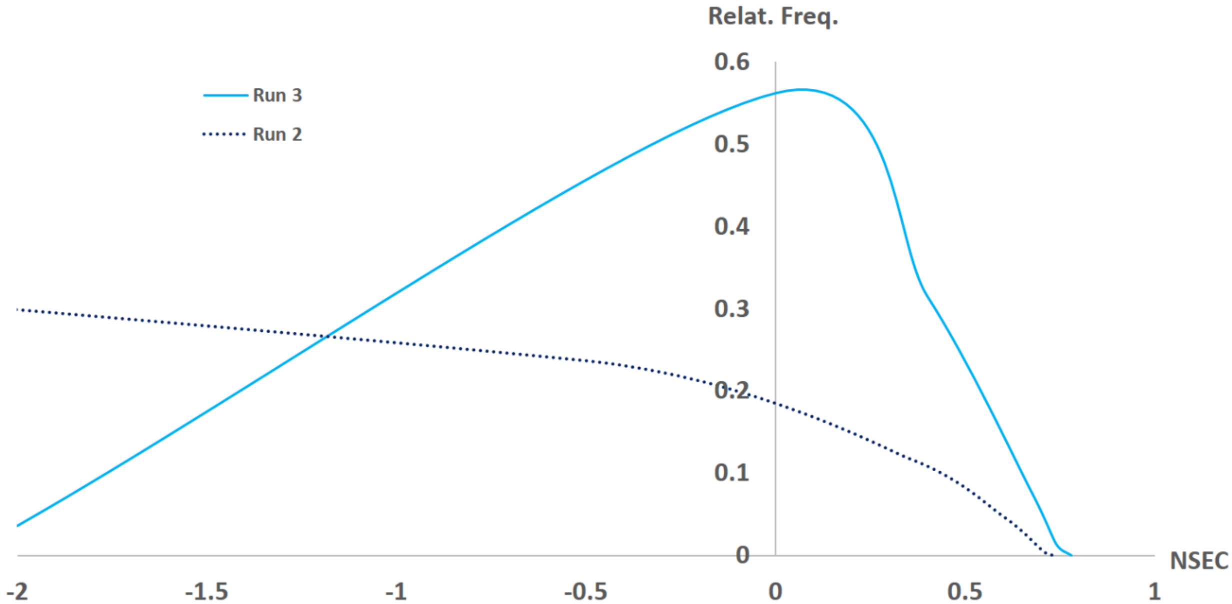

Figure 10. Model spectrum performance for monthly Nitrogen in model runs 2 and 3.

The WYLD time series were computed at a daily time step while the N loads at monthly time step. The purpose of considering all acceptable sets of parameters is to account for the multiple functional hypotheses of response that the system may have, and as a result, establish the most probable values of WYLD and N loads at a given time under alternative land use and land management scenarios.

For the purpose of visually presenting the results of WYLD for the baseline model, monthly time step was used to plot the total WYLD for the whole period of simulation (7 years, 2007-2013) (Figure 11). This figure contains only the maximum, the minimum, and the mean of the computed values from all 27 time series at each time step. At first glance, the differences between the maximum and minimum values through the 7 years seem small. The highest difference observed was 2.7 mm while the average difference was 0.84 mm. However, the average of all computed values was 7.4 mm which means that the WYLD was estimated with an uncertainty of roughly 11% (0.84*100/7.4) at each monthly time step. The highest WYLD magnitude estimated (17 mm, Figure 11) occurred the first year (2007) after the warm-up period (2005-2006). This peak value was considerably high compared to the rest of the peak values; all of the rest were lower than 14 mm. This is due to the fact that 2007 was an extremely wet year for the FCREW region.

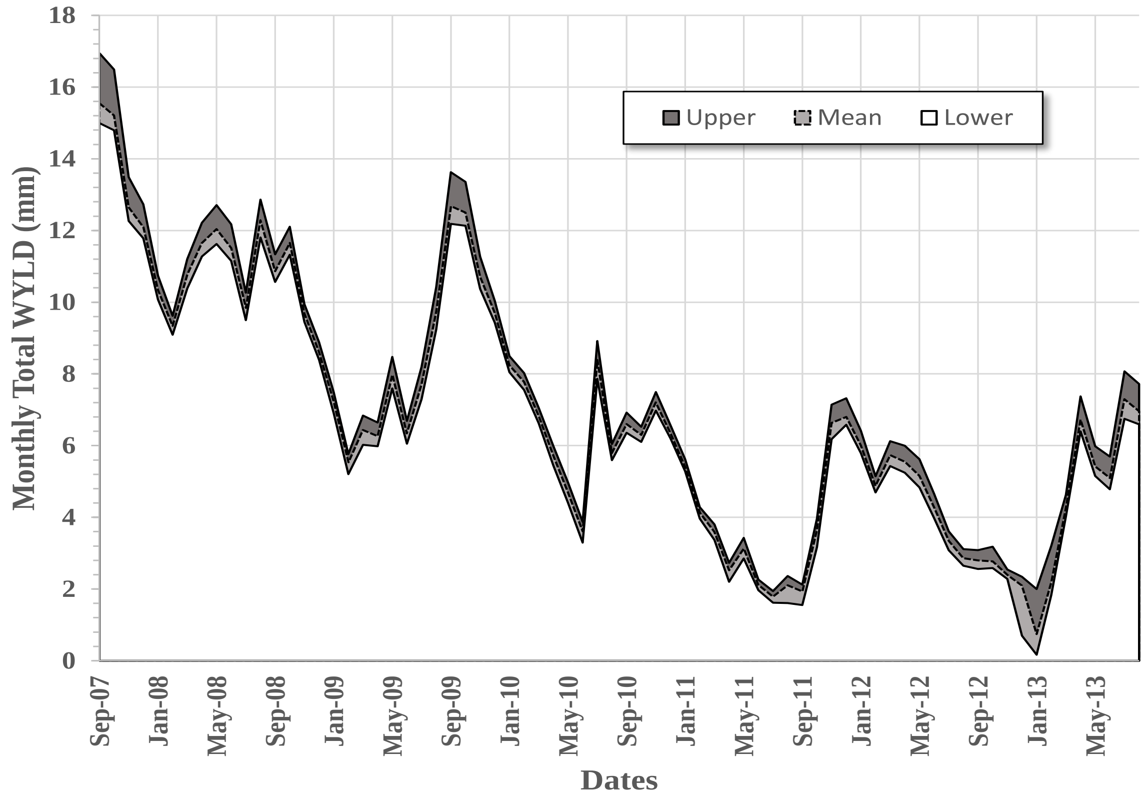

Figure 11. Monthly time series of simulated WYLD for baseline model. Only maximum, mean, minimum values were plotted.

The 2,133 computed WYLD values (79 months *x* 27 time series) were then plotted as a frequency distribution function (Figure 12). The values are spread from 0.17 mm to 41.5 mm, however, the most likely WYLD values in the Lake Creek watershed range between 5 and 25 mm per month.

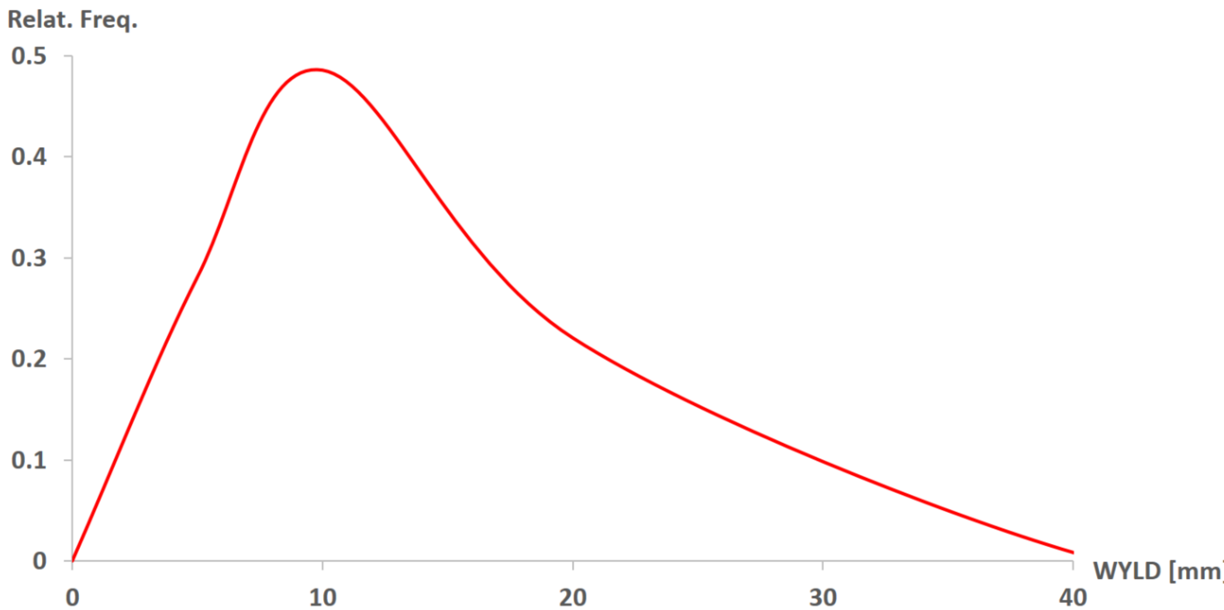

Figure 12. Monthly WYLD distribution of Baseline model.

A similar analysis was performed to the estimated Nitrogen loads. The maximum, the minimum, and the mean N loads at monthly time step were also plotted for the whole period of simulation (Figure 13). Like the WYLD time series, the highest N load value was observed in 2007 (5.3 kg/ha). This showed a direct relationship between the water yield and the nitrogen in the computations. The more the water yield produced in the watershed, the higher the nitrogen loads observed. Some other peak values in both time series may confirm this relationship. For example, the peak WYLD values in 2008 and 2009 coincided with peak values in the N time series (compare Figure 11 and 13).

The uncertainty in the nitrogen estimations was also analyzed by evaluating the differences between the maximum and minimum values. The highest difference observed was 0.42 kg/ha while the average difference was 0.07 kg/ha. Considering that the average of all computed N loads was 0.14 kg/ha, they were estimated with an approximate uncertainty of 50% (0.07*100/0.14) at monthly time step. This uncertainty was much higher than the uncertainty of

WYLD estimation (just 11%); but this was expected since the nitrogen estimations were accepted at a lower NSEC value. The magnitudes of monthly nitrogen loads ranged from 0 to 5.3 kg/ha. However, a frequency distribution plot showed that the most likely N load values in the Lake Creek watershed ranged approximately between 0.1 and 0.9 kg/ha (Figure 14).

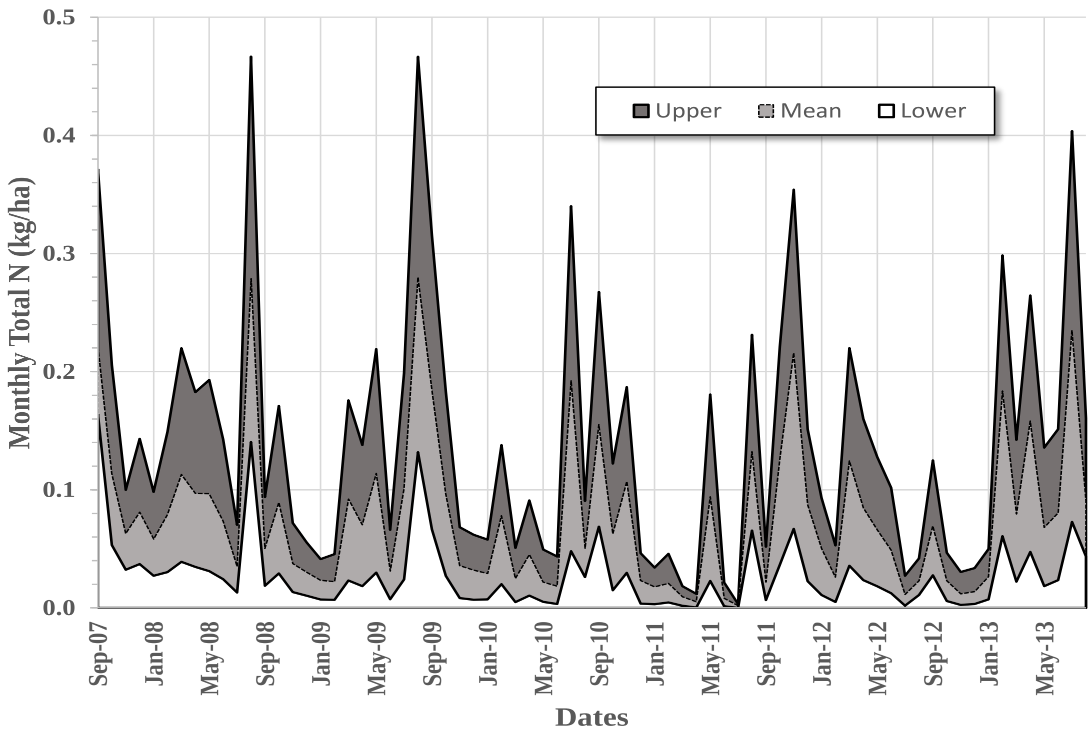

Figure 13. Monthly time series of simulated N loads for baseline model. Only maximum, mean, minimum values were plotted.

A more detailed analysis of the N loads was carried out by seasons. The 27 times series were rearranged by seasons and then plotted as frequency distribution functions (Figure 15). It can be observed that the distributions came out very similar for spring and winter, where the maximum N loads were found to be lower than in fall or summer. The probability to have N loads greater than 0.8 kg/ha in summer and fall is higher than in winter or spring. This can be due to the fact that land preparation for winter wheat planting and pesticide application at FCREW region starts in summer.

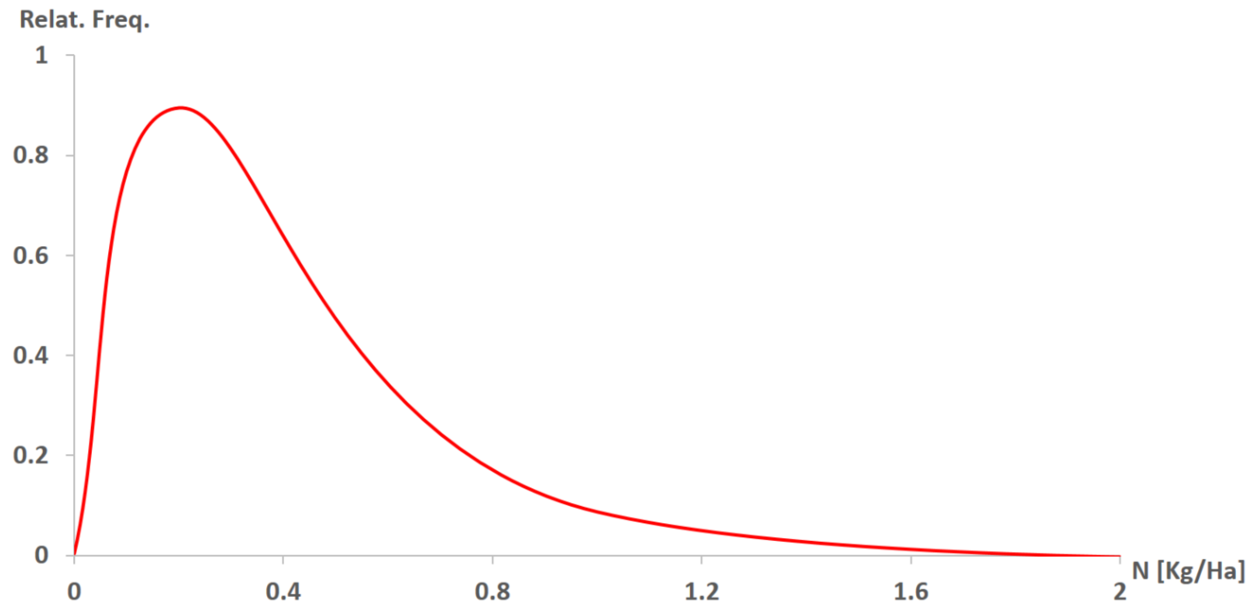

Figure 14. Monthly Nitrogen loads distribution of baseline model.

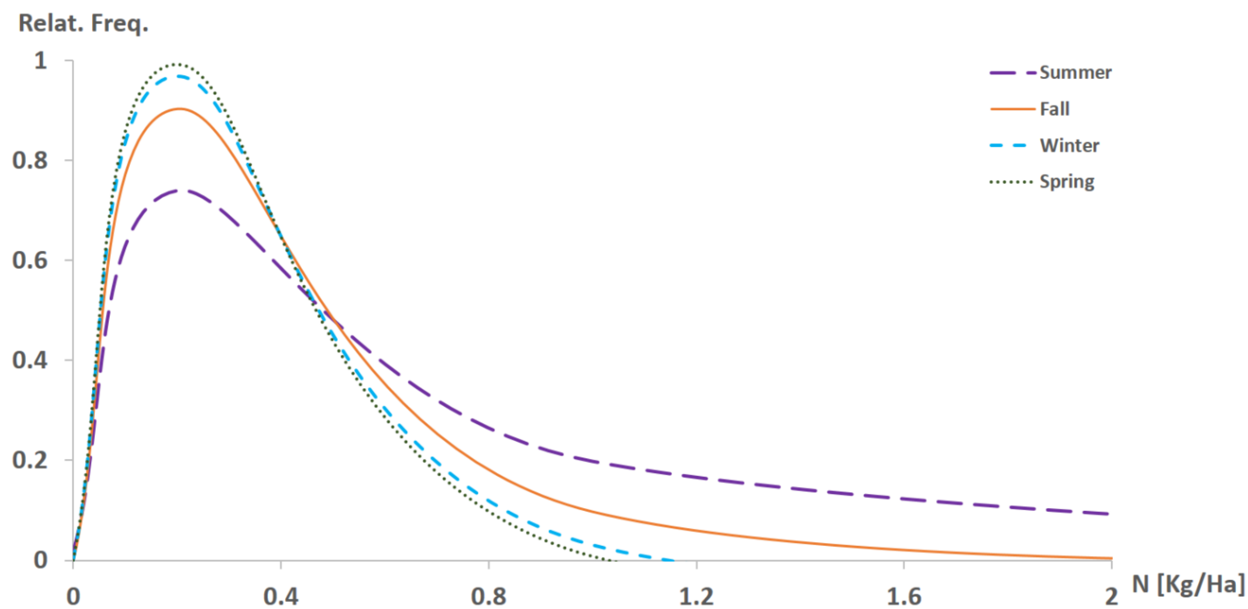

Figure 15. Seasonal Nitrogen loads distribution of baseline model.

Frequency distribution functions at a monthly step were also plotted to compare the alternative scenarios (defined in Table 3) with the baseline model for WYLD and N load outputs (Figures 16 and 17). The overall period of simulation (2,133 values) was considered when plotting the *pdf* for each scenario for both outputs. For WYLD (Figure 16), the baseline model and the different land management scenarios resulted in approximately the same distribution with a most likely value between 5 to 25 mm per month. The differences in land preparation methods and land uses did not particularly affect the water yield at the Lake Creek watershed. This was expected since land management practices implemented in FCREW were meant for the purpose of reducing nutrients and not flow.

In the case of N loads, the differences between the baseline model and the scenarios, and between the scenarios themselves, were more evident. It can be observed that the *pdfs* were increasingly spreading out (i.e., their ranges were increasing) from the 100% Pasture scenario to the 100% Winter Wheat scenario (Figure 17). The 100% Pasture *pdf* was narrower than the 100% Winter Wheat one (the most scattered). However, their most likely values were in fact the same. It seems that management practices only affected the higher extreme values not the mean. Even though the 100% Pasture scenario did not include the application of pesticides or fertilizers, the computed did not show any major changes since grazing activities also generate certain amount of nitrogen that is added to the nutrient cycle.

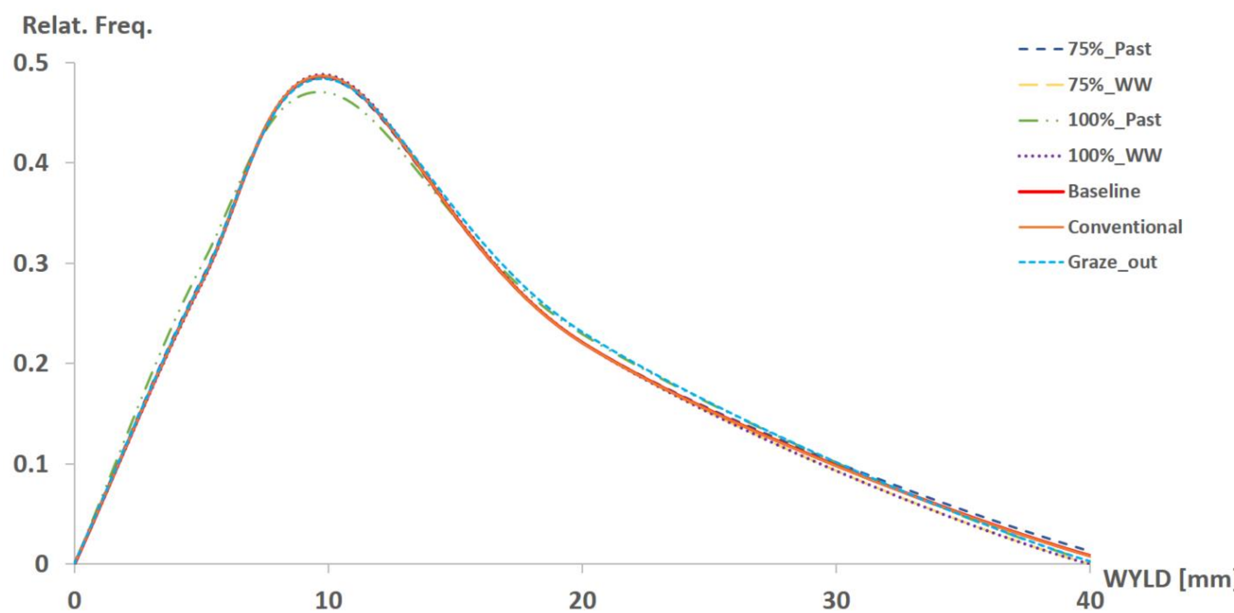

Figure 16. Monthly WYLD distribution of Scenarios.

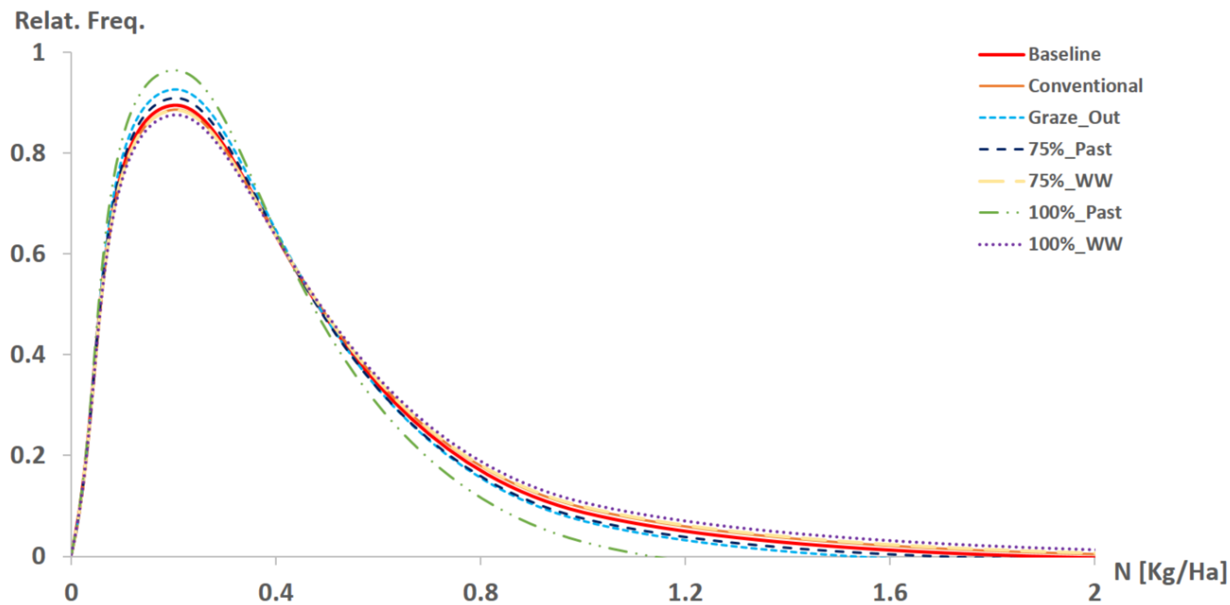

Figure 17. Monthly Nitrogen loads distribution of scenarios.

This general analysis of the nitrogen output, performed for the whole period of simulation, at the Lake Creek watershed showed minor differences with regards to the overall N load magnitudes between the baseline scenario and the alternative scenarios. The highest difference was observed between two of the land use change scenarios; the 100% Pasture and the 100% Winter Wheat scenarios which are the two opposing ends of the management spectrum. This difference may represent a reduction of 0.2 kg/ha in nitrogen loads in the 100% Pasture scenario compared to the 100% Winter Wheat. In the case of the land management scenarios (i.e., conventional and graze out), their *pdfs* were not exactly the same but the load magnitudes did not differ from each other; which may imply that the land management practices were not primarily driving the nitrogen amounts estimated for the whole period of 7 years. However, a seasonal analysis revealed greater differences between the scenarios, especially for the fall season.

The 27 times series of each scenario were rearranged by seasons (as it was done before for the baseline scenario) and then plotted as frequency distribution functions (Figure 18). For winter and spring, there were no notable differences between the scenarios but it can be noticed that the N loads in these seasons were smaller than the ones estimated for the overall period. It was stated before that N loads between 0.1 and 0.9 kg/ha can be expected at any scenario except for the 100% Pasture scenario, where N loads ranged between 0.1 and 0.7 kg/ha. In case of winter and spring, the N loads ranged approximately between 0.1 and 0.7 kg/ha (including the 100% Pasture scenario). For summer and fall the N loads were eventually greater. In summer, where the land preparation for winter wheat planting and pesticide application at the FCREW region begin, there was an evident increase in N loads. In any scenario (except 100% Pasture) the N loads may be even greater than 1.2 kg/ha. This magnitude was very unlikely in the analysis of the overall period of simulation.

For fall, the N load magnitudes diminished compared to summer in most of the scenarios but the differences between the *pdfs* were higher, especially between the Graze Out and the conventional and baseline scenarios. The Graze Out scenario displayed a broader range of possible N loads at this season compared to the others. The N magnitudes may exceed 1.2 kg/ha similarly to summer time. This is due to the fact that the tillage operations in the Graze Out occur only one week before the application of fertilizer while in the conventional and baseline scenarios occur more than one month in advance. In the case of Graze Out, the soil may be more exposed to nitrogen, which can mix with rainfall and reach the streams easier if the soil has been recently eroded. Even though the land preparation and fertilizer application activities were carried out in August (still considered summer), nutrients can reside longer in the stream until the fall season.

Another comparison of the total monthly N load with the baseline model was carried out by taking the mean, minimum, and maximum values of the 27 simulations comprising each scenario and contrasting them with the banded baseline N load (Figure 18). The deviation of the maximum and minimum N of each scenario (dots) from the baseline scenario (band) represents the impacts of the land management practices taking into account the uncertainties brought about by parameter selection. The uncertainties in N loads due to parameter estimation for the different land management practices can reach up to 0.36 Kg/ha for the conventional scenario. The lowest uncertainty is for 100% pasture at 0.24 Kg/ha while the rest of the scenarios remained the same as the baseline uncertainty.

The rise and fall in N loads for both the conventional and graze out scenarios follow the same pattern as the baseline model where increases happened during the early fall following fertilizer application for the winter wheat (Figure 19a and b). Increasing the amount of winter wheat in the area also increased N loads (Figure 19d and f) higher than the baseline model while

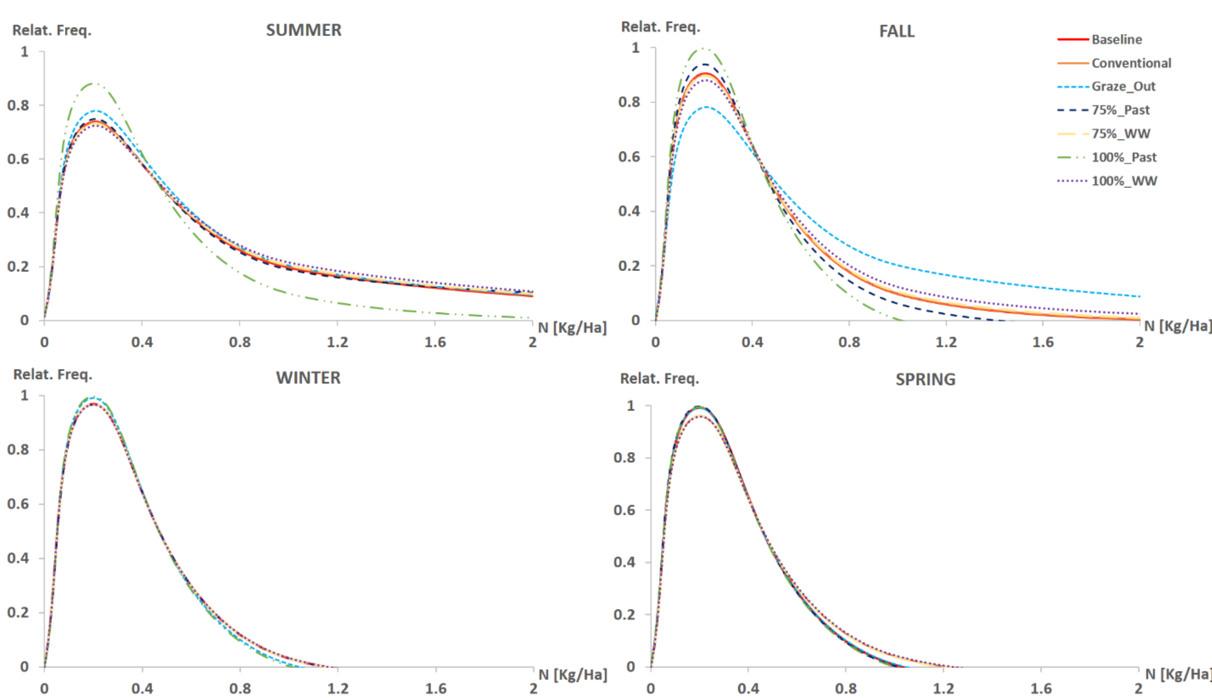

Figure 18. Seasonal Nitrogen loads distribution of scenarios.

the opposite happened when pasture cover was increased (Figure 19c and e). In general, changes in land use (Figure 19c to f) had more notable impacts on the N loads than the conservation practices did (Figure 19a and b).

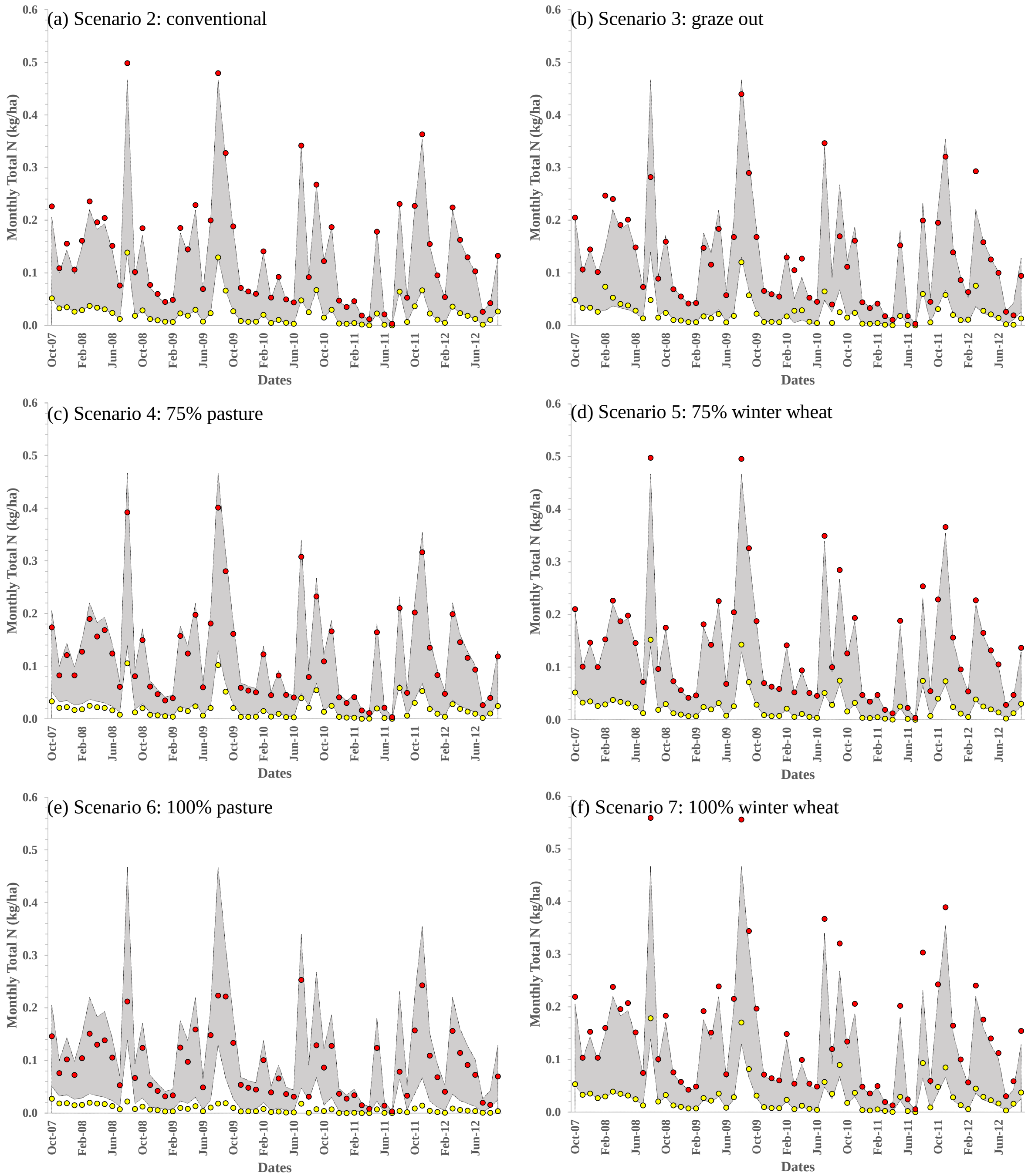

Figure 19. Comparison of the six land management scenarios with the baseline model. The band represents the 27 parameter combination of the baseline model while the dots the maximum and minimum values of the 27 models comprising each scenario.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The most challenging part in modeling is understanding the capabilities and limitations of the model given the uncertainties involved in the process. The challenge of finding the most appropriate combinations of input parameters to represent the response of the system is an issue faced by environmental scientists and managers. In particular, the simulation of a complex environmental system like the hydrologic system requires the consideration of the interaction of many variables with varying ranges and distributions. This study developed a methodology to address this issue. A two-step probabilistic approach, using global uncertainty and sensitivity analysis, was implemented to develop a hydrologic model that was used to simulate the systemic response of the watershed. Probabilistic inputs were used to derive the full spectrum of system response and, therefore, the multiple functional hypothesis of the system behavior.

The output of interest in the calibration process was the Nash-Sutcliffe Efficiency Coefficient (NSEC). This coefficient was used to measure the model performance in estimating the water yield (i.e., the hydrologic component, WYLD) and the Nitrogen loads in the watershed. The different uncertain and independent parameters of the model were sampled using uniform distributions within wide, but physically feasible, ranges. These inputs generated model output spectrums with high variability that was gradually reduced as new parameter distributions and narrower ranges were defined through the evaluation of uncertainty and sensitivity indexes. The model performance was also improved with the use of new parameter ranges and distributions. It allowed the definition of criteria for acceptable simulations (i.e., those parameter combinations that can acceptably represent the system response) which composed the baseline model of the study area.

This study demonstrated that quantification of input-output uncertainties and equifinality can be incorporated in the model development process. This coupled process facilitated the adjustment of input/parameters as well as the accountability of model uncertainties. The use of probabilistic inputs resulted in a more efficient way to deal with the many input/parameters considered in hydrological models. In the case of the APEX model developed for the Lake Creek watershed, 20 parameters were found to control the estimation of WYLD and 10 the estimation of N loads (7 of them also WYLD parameters). However, the computation of the Sobol's 1st order sensitivity index (S_i) also revealed that only one parameter for WYLD and only two parameters for N loads controlled over 50% (out of 80%) of the full spectrum of response of each output. The most important parameter that controlled WYLD was the Expands CN Retention Parameter (cnrp) while for Nitrogen were the two Hargreaves PET equation parameters (the Hargreaves PET equation exponent (hpee), and the Hargreaves PET equation coefficient (hpec)).

Results showed that 27 sets of parameters can acceptably represent the hydrologic responses of Lake Creek watershed. The total monthly water yield can range from 0.17 to 41.5 mm estimated with an uncertainty of 11%. The months of highest uncertainty were the first three months of the first year (2007) after the warm-up period (2005-2006). The total Nitrogen loads can vary from 0 to 5.3 kg/ha at monthly time step with uncertainty of 50%. The months of highest N loads and highest uncertainty were also the first three months of the first year (2007).

Results from the simulation of land management practices scenarios suggested that changes in land use and land preparation did not affect the total water yield at watershed scale. As the alternative scenarios were proposed to reduce the nitrogen loads, the WYLD was not impacted. In fact, the same range of probable total monthly WYLD for the baseline and all alternative scenarios was identified. The model simulated a water yield between 5 and 25 mm per month even if the physical and operational conditions were partially or completely modified, according to the scenarios defined.

The seasonal analysis of N loads at the Lake Creek watershed revealed that summer and fall seasons generated the highest variability between the scenarios. This seasons coincided with the land preparation season, summer in this case, and its subsequent season (i.e., fall). In summer and fall, N loads greater than 1.2 kg/ha can be expected in any scenario, except the 100% Pasture; while in winter and spring, the N loads did not exceed 0.7 kg/ha in any scenario (including the 100% Pasture). The Graze out scenario displayed the broadest range of possible N loads and the highest probability to register N loads greater than 1.2 kg/ha in fall season. This was due to the shorter time lag between plowing and application of fertilizers.

The methodology developed in this study was able to quantify the full spectrum of system responses, the uncertainty associated with them, and the most important inputs that drive their variability. The full spectrum of model outputs can provide robust information on the achievable responses of the watershed given the different land management practices in place and planned. Similarly, by knowing the parameters that drive the variability of the outputs, future research can be prioritized to collect more information about these parameters resulting in a more efficient use of resources. Results from this study can be used to develop strategic decisions on the risks and tradeoffs associated with different management alternatives that aim to increase productivity while also minimizing their environmental impacts.

REFERENCES

Ajami, N. K., Duan, Q., & Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour. Res., 43,* W01403.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment, part 1: model development. *J. of the Amer. Water Res. Assoc. 34(1)*, 73-89.

Benedini, M., & Tsakiris, G. (2013). *Water quality modelling for rivers and streams.* Springer: Water science and technology library.

Beven, K. J. (1993). Prophecy, reality, and uncertainty in distributed hydrological modeling. *Adv. Water Resources 16(1)*, 41-51.

Beven, K. J. (2006). A manifesto for the equifinality thesis. *J. Hydrol. 320(1-2)*, 18-36.

Beven, K. J., & Binley, A. M. (1992). The future of distributed models: Model calibration and predictive uncertainty. *Hydrol. Processes, 6*, 279-298.

Board, O. W. (2003). *Oklahoma's benefitial use monitoring program: Lakes, 2003 report.* Oklahoma City: Okla. Water Resour. Board. Retrieved from http://www.owrb.ok.gov/quality/monitoring/bump/pdf_bump/archives/2003BUMPLakes.pdf

Cariboni, J., Gatelli, D., Liska, R., & Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. *Ecol. Model.*, 167-182.

Chu-Agor, M. L., Munoz-Carpena, R., Kiker, G. A., Aiello-Lammens, M. E., Akcakaya, H. R., Convertino, M., & Linkov, I. (2012). Simulating the fate of Florida Snowy Plovers with sea-level rise: Exploring research and management priorities with a global uncertainty and sensitivity analysis perspective. *Ecological Modelling 224*, 33-47.

Chu-Agor, M. L., Munoz-Carpena, R., Kiker, G. A., Emanuelsson, A., & Linkov, I. (2011). Exploring sea level rise vulnerability of coastal habitats through sensitivity and uncertainty analysis. *Environ. Modell. Soft. 26*, 593-604.

Diaz, R. J., & Rosenberg, R. (2008). Spreading Dead Zones and Consequences for Marine Ecosystems. *Science*, *321*, 926-929.

Donigian, A. S. (2002). Watershed model calibration and validation: The HSPF experience. *WEF. Natl. TMDL science and policy* (pp. 44-73). Alexandria, VA: Water Environment Federation.

Gassman, P. W., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). The soil and water assessment tool - historical development applications, and future research directions. *Trans. of the ASABE 50(4)*, 1211-1250.

Guzman, J. A., Moriasi, D. N., Gowda, P. H., Steiner, J. L., Starks, P. J., Arnold, J. G., & Srinivasan, R. (2015). A model integration framework for linking SWAT and MODFLOW. *Envir. Model. Soft. 73*, 103 - 116.

Guzman, J. A., Shirmohammadi, A., Sadeghi, A., Wang, X., Chu, M. L., Jha, M. K., . . . Hernandez, J. (2015). Uncertainty Considerations in Calibration and Validation of Hydrologic and Water Quality Models. *Trans. of ASABE*, 1745-1762.

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engr. Agric., 1*, 96-99.

Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., & Arnold, J. G. (2006). Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE 49(3)*, 689-701.

Hiebeler, D. E., & Michaud, I. J. (2012). Quantifying spatial and temporal variability of spatial correlated disturbances. *Ecol. Model. 240*, 64-73.

Hornberger, G. M., & Spear, R. C. (1981). An approach to the preliminary analysis of environmental systems. *J. Environmental Management 12*, 7-18.

Jain, S. K., Storm, B., Bathurst, J. C., Refsgaard, J. C., & Singh, R. D. (1992). Application of the SHE to Catchments in India. Part 2. Field Experiments and Simulation studies with the SHE on the Kolar Subcatchment of the Narmada river. *Journal of Hydrology 140*, 25-47.

Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resour. Res. 29*, 2637-2649.

King, K. W., Fausey, N. R., & Williams, M. R. (2014). Effect of subsurface drainage on streamflow in an agricultural headwater watershed. *Journal of Hydrology 519* , 438–445.

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res. 42*, W03S04.

Loomis, R. S., Rabbinge, R., & Ng, E. (1979). Explanatory models in crop physiology. *Annu. Rev. Plant Physiol. 30*, 339-367.

Mausbach, M.J., and Dedrick, A.R. (2004). The length we go: Measuring environmental benefits of conservation practices. *J. Soil Water Conserv. 59(5)*, 96A-103A.

McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Elliot, R. L., Kilby, J. R., Grimsley, D. L., . . . Demko, D. B. (2007). Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmosp. and Ocean. Tech 24(3)*, 301-321.

Moriasi, D. N., & Starks, P. J. (2010). Effects of the resolution of soil dataset and precipitation dataset on SWAT2005 streamflow calibration parameters and simulation accuracy. *J. of Soil and Water Conserv. 65 (2)*, 63-78.

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

Moriasi, D. N., Guzman, J. A., Steiner, J. L., Starks, P. J., & Garbrecht, J. D. (2014a). Seasonal sediment and nutrient transport patterns. *J. of Environ. Qual*, 1334-1344.

Moriasi, D. N., Starks, P. J., Guzman, J. A., Garbrecht, J. D., Steiner, J. L., Stoner, J. C., . . . Naney, J. W. (2014b). Upper washita river experimental watersheds: reservoir, groundwater, and streamflow data. *J. Environ. Qual. 43*, 1262-1272.

Moriasi, D. N., Wilson, B. N., Douglas-Mankin, K. R., Arnold, J. G., & Gowda, P. H. (2012). Hydrologic and water quality models: use, calibration, and validation. *Trans. of the ASABE 55(4)*, 1241-1247.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: Part 1. A discussion of principles. *J. Hydrology 10(3)*, 282-290.

NCDC. (2015, 05 07). *Gauge data for Westerville, Ohio, GHCND:USC00338951*. Retrieved from National climatic data center: http://www.ncdc.noaa.gov/

Reibich, R. A., & Demcheck, D. K. (2007). *Trends in nutrient and sediment concentrations and loads in major river basins of the south-central Unites States, 1993-2004.* Reston, VA: Sci. Invest. Report, 2007-5090, USGS.

Richardson, C. W., Bucks, D. A., & Sadler, E. J. (2008). The Conservation Effects Assessment Project benchmark watersheds: Synthesis of preliminary findings. *Journal of Soil and Water Conservation 63 (6)*, 590-604.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., . . . Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer.* The Atrium, Southern Gate, Chichester, West Sussex, England.: John Wiley & Sons Ltd.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models.* Chichester, UK.: John Wiley and Sons.

Seibert, J. (2003). Reliability of model predictions outside calibration conditions. *Nord. Hydrol. 34*, 477-492.

Shirmohammadi, A., Chaubey, I., Harmel, R. D., Bosch, D. D., Muñoz-Carpena, R., Dharmasri, C., . . . Sohrabi, T. M. (2006). Uncertainty in TMDL models. *Trans. ASABE 49(4)*, 1033-1049.

Sobol, I. M. (1993). Sensitivity estimates for non-linear mathematical models. *Math modell. Comput. Exp. 4(I)*, 407-414.

Starks, P. J., Steiner, J. L., Moriasi, D. N., Guzman, J. A., Garbrecht, J. D., Allen, P. B., & Naney, J. W. (2014). Upper Washita River Experimental Watersheds: Nutrient water quality data. *J. of Environ. Quality*, 1280-1297.

Steglich, E. M., & Williams, J. W. (2013). *Agricultural Policy/Environmetal eXtender Model User's manual.* Temple, Texas: Blackland Research and Extension Center.

Steiner, J. L., Starks, P. J., Daniel, J. A., Garbrecht, J. D., Moraisi, D., McIntyre, S., & Chen, J. S. (2008). Environmental effects of agricultural conservation: A framework for research in two watersheds in Oklahoma's Upper Washita River Basin. *Journal of Soil and Water Conservation 63(6)*, 443-452.

Steiner, J. L., Starks, P. J., Garbrecht, J. D., Moriasi, D. N., Zhang, X., Schneider, J. M., . . . Osei, E. (2014). Long-Term Environmental Research: The Upper Washita River Experimental Watersheds, Oklahoma, USA. *J of Environ Qual*, 1227-1238.

Storm, D. E., Busteed, P. R., & White, M. J. (2006). *Fort-Cobb Basin. Modeling and land cover classification.* Retrieved from Draft reported to the Oklahoma Department of Environmental Quality: http://www.deq.state.ok.us/WQDnew/tmdl/fort_cobb/osu_fort_cobb_modeling_jan_2006.pdf

Tomer, M. D., & Locke, M. A. (2011). The challenge of documenting water quality benefits of conservation practices: a review of USDA-ARS's conservation effects assessment project watershed studies. *Water Science & Technology 64(1)*, 300-310.

USDA. (1994). Water Quality Research Plan for Management Systems Evaluation Areas. Washington, DC: USDA Agricultural Research Service: Agricultural Research Service Bulletin ARS-123.

USDA. (1995). *Data Base: Data Use information.* Soil Survey Geographic (SSURGO), National Coil Survey Center. Lincoln, NE: Natural Resources Conservation Services (NRCS).

USDA. (2016). The Web Soil Survey (WSS). http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm

USDA-Soil Conservation Service. (1972). National Engineering Handbook. In *Hydrology Section 4* (pp. Chapters 4-10).

USGS. (2016). The USGS Store. http://www.usgs.gov/pubprod/

Wagener, T. (2003). Evaluation of catchment models. *Hydrol. Processes, 17,* 3375-3378.

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environ. Res. Risk Assess., 19(6)*, 378-387.

Wang, X., Williams, J. R., Gassman, P. W., Baffaut, C., Izaurralde, R. C., Jeong, J., & Kiniry, J. R. (2012). EPIC and APEX: model use, calibration, and validation. *Trans. of the ASABE 55(4)*, 1447-1462.

Whitehead, P. G., & Young, P. C. (1979). Water quality in river systems: Monte-Carlo analysis. *Water Resources Research 15*, 451-459.

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

Williams, J. R., Jones, C. A., Gassman, P. W., & Hauck, L. M. (1995). Simulation of animal waste management with APEX. In *Innovations and New Horizons in Livestock and Poultry* (pp. 22-26). College Station: J. McFarland, ed. Austin, Tex.: Texas A&M University, Texas Agricultural Extension Service.