



EPA

United States
Environmental Protection
Agency

Watershed Management Tool for Selection and Spatial Allocation of Non-Point Source Pollution Control Practices

Final Report

$$\hat{f}_i(\vec{\alpha} | \mathbf{B}) = \frac{m_i}{m} \times \left(\frac{1}{V_i}\right)$$

$$\sum_{\tau=1}^{td} \sum_i \left[(\Delta y_{i,\tau} \cdot b m_i) (1+s)^{(\tau-1)} \right]$$

$$z = \frac{\text{maximize}}{\max \left(c_0 (1+s)^{td} + c_0 \cdot rm \left[\sum_{\tau=1}^{td} (1+s)^{(\tau-1)} \right], 1 \right)}$$

Mazdak Arabi, Rao S. Govindaraju, Mohamed M. Hantush

EPA/600/R-08/036
January 2007

**WATERSHED MANAGEMENT TOOL FOR SELECTION
AND SPATIAL ALLOCATION OF NONPOINT SOURCE
POLLUTION CONTROL PRACTICES**

Prepared By

**Mazdak Arabi
Department of Agricultural and Biological Engineering
Purdue University
West Lafayette, Indiana 47907**

**Rao S. Govindaraju
School of Civil Engineering
Purdue University
West Lafayette, Indiana 47907**

Contract Number: 4C-R330-NAEX

Project Officer

**Mohamed M. Hantush
National Risk Management Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, Ohio 45268**

NOTICE

The U.S. Environmental Protection Agency through its Office of Research and Development funded the research described here. This report has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

All research projects making conclusions or recommendations based on environmentally related measurements and funded by the Environmental Protection Agency are required to comply with the Agency Quality Assurance Program. This report has been subjected to QA/QC review. The report presented a mathematical framework for modeling water quality in eutrophic water bodies and did not involve collection and analysis of environmental measurements.

FOREWORD

The U.S. Environmental Protection Agency (EPA) is charged by Congress with protecting the Nation's land, air, and water resources. Under a mandate of national environmental laws, the Agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. To meet this mandate, EPA's research program is providing data and technical support for solving environmental problems today and building a science knowledge base necessary to manage our ecological resources wisely, understand how pollutants affect our health, and prevent or reduce environmental risks in the future.

The National Risk Management Research Laboratory (NRMRL) is the Agency's center for investigation of technological and management approaches for preventing and reducing risks from pollution that threaten human health and the environment. The focus of the Laboratory's research program is on methods and their cost-effectiveness for prevention and control of pollution to air, land, water, and subsurface resources; protection of water quality in public water systems; remediation of contaminated sites, sediments and ground water; prevention and control of indoor air pollution; and restoration of ecosystems. NRMRL collaborates with both public and private sector partners to foster technologies that reduce the cost of compliance and to anticipate emerging problems. NRMRL's research provides solutions to environmental problems by: developing and promoting technologies that protect and improve the environment; advancing scientific and engineering information to support regulatory and policy decisions; and providing the technical support and information transfer to ensure implementation of environmental regulations and strategies at the national, state, and community levels.

This publication has been produced as part of the Laboratory's strategic long-term research plan. It is published and made available by EPA's Office of Research and Development to assist the user community and to link researchers with their clients.

Sally Gutierrez, Director
National Risk Management Research Laboratory

ABSTRACT

Distributed-parameter watershed models are often utilized for evaluating the effectiveness of sediment and nutrient abatement strategies through the traditional {calibrate→ validate→ predict} approach. The applicability of the method is limited due to modeling approximations. In this study, a computational method is presented to determine the significance of modeling uncertainties in assessing the effectiveness of best management practice (BMPs) in two small watersheds in Northeastern Indiana with the Soil and Water Assessment Tool (SWAT). The uncertainty analysis aims at (i) identifying the hydrologic and water quality processes that control the fate and transport of sediments and nutrients within watersheds, and (ii) establishing uncertainty bounds for model simulations as well as estimated effectiveness of BMPs. The SWAT model is integrated with a Monte-Carlo based methodology for addressing model uncertainties. The results suggested that fluvial processes within the channel network of the study watersheds control sediment yields at the outlets, and thus, BMPs that influence channel degradation or deposition are the more effective sediment control strategies. Conversely, implementation of BMPs that reduce nitrogen loadings from uplands areas such as parallel terraces and field borders appeared to be more crucial in reducing total N yield at the outlets. The uncertainty analysis also revealed that the BMPs implemented in the Dreisbach watershed reduced sediment, total P, and total N yields by nearly 57%, 33%, and 31%, respectively. Finally, a genetic algorithm (GA)-based optimization methodology is developed for selection and placement of BMPs within watersheds. The economic return of the selected BMPs through the optimization model was nearly three fold in comparison to random selection and placement of the BMPs.

TABLE OF CONTENTS

| | |
|---|------|
| Notice..... | ii |
| Foreword..... | iii |
| Abstract..... | iv |
| List of Figures..... | viii |
| List of Tables..... | ix |
| Acknowledgements..... | xi |
| | |
| SECTION 1. Introduction..... | 1 |
| 1.1. Rationale and Background..... | 1 |
| 1.2. Objectives..... | 5 |
| 1.3. Impacts of the Study..... | 5 |
| 1.4. Structure of the Report..... | 6 |
| | |
| SECTION 2. Case Study Watersheds and the Watershed Model..... | 7 |
| 2.1. Case Study Watersheds..... | 7 |
| 2.2. Choice of the Watershed Model: SWAT..... | 8 |
| 2.2.1. Model calibration..... | 11 |
| 2.2.2. Representation of BMPs..... | 12 |
| | |
| SECTION 3. Probabilistic Approach for Analysis of Uncertainty in the Evaluation of Watershed Management Practices..... | 14 |
| 3.1. Abstract..... | 14 |
| 3.2. Introduction..... | 15 |
| 3.3. Theoretical Considerations..... | 17 |
| 3.3.1. Morris's One-at A Time sensitivity analysis..... | 17 |
| 3.3.2. Generalized Likelihood Uncertainty Estimation (GLUE)..... | 17 |
| 3.4. Methodology..... | 18 |

| | |
|---|----|
| 3.4.1. OAT sensitivity analysis | 19 |
| 3.4.2. Range adjustment | 20 |
| 3.4.3. Uncertainty computations | 21 |
| 3.5. Results | 22 |
| 3.5.1. Range adjustment results..... | 22 |
| 3.5.2. Uncertainty analysis results..... | 26 |
| 3.6. Discussion..... | 31 |
| 3.7. Conclusions | 32 |
| SECTION 4. Sensitivity Analysis of Sediment and Nutrient Processes with a Watershed Model | 33 |
| 4.1. Abstract..... | 33 |
| 4.2. Introduction | 33 |
| 4.3. Theoretical Considerations | 36 |
| 4.3.1. Regionalized Sensitivity Analysis (RSA) | 36 |
| 4.3.2. Tree Structured Density Estimation (TSDE) | 36 |
| 4.4. Methodology..... | 38 |
| 4.4.1. Behavior Definition..... | 39 |
| 4.5. Results | 40 |
| 4.6. Discussion..... | 43 |
| 4.7. Conclusions | 44 |
| SECTION 5. Cost-Effective Allocation of Watershed Management Practices Using a Genetic Algorithm..... | 54 |
| 5.1. Abstract..... | 54 |
| 5.2. Introduction | 54 |
| 5.3. Theoretical Considerations | 58 |
| 5.3.1. Genetic Algorithm (GA) | 58 |
| 5.4. Methodology..... | 59 |
| 5.4.1. NPS Model | 60 |
| 5.4.2. BMP Representation | 61 |
| 5.4.3. Economic Component..... | 61 |
| 5.4.4. Optimization Component | 64 |
| 5.4.5. Optimization Cases | 70 |

| | |
|---------------------------------|--------|
| 5.5. Results | 71 |
| 5.6. Discussion..... | 77 |
| 5.7. Conclusions | 78 |
| SECTION 6. Conclusions..... | 80 |
| Bibliography | 83 |

LIST OF FIGURES

| Figure | Page |
|---|-------------|
| 2.1 Case Study Watersheds..... | 8 |
| 3.1 Spider plots for the most sensitive input factors..... | 24 |
| 3.2 Cumulative GLUE-likelihood for output variables simulated at the outlet of the Dreisbach watershed..... | 27 |
| 3.3 Cumulative GLUE-likelihood for output variables simulated at the outlet of the Smith Fry watershed..... | 28 |
| 4.1 Computational framework for sensitivity analysis of sediment and nutrient processes.... | 39 |
| 4.2 Posterior distributions of behavior and non-behavior sets of input factors along with their prior distribution (uniform distribution) for sediment at the outlet of Dreisbach and Smith Fry watersheds..... | 46 |
| 4.3 Posterior distributions of behavior and non-behavior sets of input factors along with their prior distribution (uniform distribution) for total N at the outlet of Dreisbach and Smith Fry watersheds..... | 47 |
| 4.4 TSDE diagram for sediment at the outlet of Dreisbach watershed..... | 48 |
| 4.5 TSDE diagram for sediment at the outlet of Smith Fry watershed..... | 48 |
| 4.6 TSDE diagram for total N at the outlet of Dreisbach watershed..... | 49 |
| 4.7 TSDE diagram for total N at the outlet of Smith Fry watershed..... | 49 |
| 5.1 Schematic of the optimization procedure..... | 60 |
| 5.2 Schematic of an optimization string..... | 65 |
| 5.3 Analysis of sensitivity of GA operating parameters..... | 71 |
| 5.4 (a) Case3 optimization outputs for the Dreisbach watershed over 2000–2009 period. (b) Case 4 optimization outputs for the Dreisbach watershed over 2000–2009 period. (c) Case 4 optimization outputs for the Smith Fry watershedover2000–2009period. | 74 |
| 5.5 Spatial allocation of BMPs from: (a) case 3 computations for Dreisbach watershed; (b) case 4 computations for Dreisbach; (c) case 4 computations for smith Fry watershed. | 76 |
| 5.6 Tradeoff between economic criterion (cost) and environmental criteria..... | 77 |

LIST OF TABLES

| Table | Page |
|---|-------------|
| 2.1 BMPs installed in the Dreisbach and Smith Fry watersheds..... | 8 |
| 2.2 Listing of SWAT parameters..... | 10 |
| 2.3 Representation of BMPs with SWAT..... | 13 |
| 3.1 Top five sensitive SWAT parameters for streamflow, sediment, total P, and total N computations..... | 23 |
| 3.2 Adjusted parameter ranges for the study watersheds..... | 25 |
| 3.3 Summary of the results from analysis of uncertainty..... | 29 |
| 4.1 Regionalized Sensitivity Analysis ranking of input factors for sediment based on observed-past behavior definition at the outlet of Dreisbach watershed..... | 50 |
| 4.2 Regionalized Sensitivity Analysis ranking of input factors for sediment based on observed-past behavior definition at the outlet of Smith Fry watershed..... | 50 |
| 4.3 Regionalized Sensitivity Analysis ranking of input factors for total N based on observed-past behavior definition at the outlet of Dreisbach watershed..... | 51 |
| 4.4 Regionalized Sensitivity Analysis ranking of input factors for total N based on observed-past behavior definition at the outlet of Smith Fry watershed..... | 51 |
| 4.5 Summary of TSDE terminal nodes for sediment related parameters based on observed-past behavior definition at the outlet of Dreisbach watershed..... | 52 |
| 4.6 Summary of TSDE terminal nodes for sediment related parameters based on observed-past behavior definition at the outlet of Smith Fry watershed..... | 52 |
| 4.7 Summary of TSDE terminal nodes for total N related parameters based on observed-past behavior definition at the outlet of Dreisbach watershed..... | 53 |
| 4.8 Summary of TSDE terminal nodes for total N related parameters based on observed-past behavior definition at the outlet of Smith Fry watershed..... | 53 |
| 5.1 Unit cost of establishment (c), maintenance rate (rm), and design life (td) of BMPs in the study along with number (nbmp) and quantity (abmp) of each type in the study watersheds..... | 62 |

| | |
|--|----|
| 5.2 Monetary benefits of unit reduction of sediment and nutrient loads in 2000 dollars | 64 |
| 5.3 Various combinations of GA operating parameters for sensitivity analysis..... | 69 |
| 5.4 Description of cases compared in this study..... | 70 |
| 5.5 Results for base-case and targeting cases. Mean and maximum value of monthly sediment, phosphorus, and nitrogen loads were estimated from SWAT simulations | 72 |
| 5.6 Results for optimization cases | 73 |

ACKNOWLEDGEMENTS

The U.S. Environmental Protection Agency through its Office of Research and Development funded the research described here. This support is acknowledged. The report benefited from the constructive comments of Dr. M. M. Hantush. The report was reviewed by Drs. Zhonglong Zhang and M. Dougherty, and their comments were utilized in preparing this final version. Their help is acknowledged.

SECTION 1. INTRODUCTION

1.1. Rationale and Background

Soil and water conservation practices are widely accepted as effective control measures for agricultural nonpoint sources of sediments and nutrients. The 2002 Farm Bill provided up to \$13 billion for conservation programs aimed at protecting water quality from agricultural nonpoint source (NPS) pollution (USDA, 2003). In addition, under the Clean Water Act Section 319 Nonpoint Source National Monitoring Program and wetland protection programs, the EPA supports programs to reduce the negative impacts of runoff from agricultural, urban, and industrialized areas. Similarly, the Natural Resources Conservation Service (NRCS) provides hundreds of millions of dollars in federal funds to support agricultural conservation practices in an effort to reduce the movement of pollutants into our waterways. Success of such programs, however, is contingent upon availability of efficient watershed-scale planning tools.

Watershed management plans are composed of individual structural and/or cultural soil and water conservation units that are often referred to as best management practices (BMPs). BMPs can substantially reduce transport of contaminants to surface water (Mostaghimi et al., 1997; Arabi et al., 2004), but, their implementation bears additional costs to stakeholders and watershed managers. Assessment of watershed management plans is challenged by complexities in incorporation of conflicting environmental, economic, and institutional criteria. Environmental assessments in watersheds hinge on resolving social benefits such as achieving the goal of swimable and fishable water bodies under the US EPA's total maximum daily load (TMDL) agenda. While BMPs facilitate achievement of such targets, their establishment bears additional cost for watershed management and/or agricultural producers. Since management practices are usually implemented under a limited budget, costs associated with unnecessary/redundant management actions may jeopardize attainability of designated water quality goals. Identifying optimal combinations of

watershed management practices requires systematic approaches that allow decision makers to quickly assess trade-offs among environmental and economic criteria.

Implementation of BMPs is rarely followed by long-term monitoring and data collection efforts to assess their effectiveness in meeting their intended goals. Long-term data on flow and water quality within watersheds, before and after placement of BMPs, is not generally available. Therefore, cost-effective evaluation of BMPs (especially new ones that have had little or no history of use) needs to be conducted with the help of simulation models. The assessment of these economic and environmental metrics requires an integrated approach often involving social objectives and competing interests from various stakeholder groups.

Arabi et al. (2004) developed a process based methodology, using the Soil and Water Assessment Tool (SWAT), to numerically represent the impacts of four structural BMPs on hydrologic/water quality processes in two small agricultural watersheds in Indiana. Various processes that were considered included: infiltration; surface runoff (peak and volume); upland erosion (sheet and rill erosion); gully and channel erosion; nutrient loadings from upland areas; and within-channel processes. Based on the function of a BMP, specific SWAT parameters that represent the impacted hydrologic/water quality processes were altered. While similar approaches can be followed for other BMPs, an important question for evaluation of BMPs at the watershed scale arises: *how can the tools of modeling and computational analysis be applied to identify the best configuration (type and location) of BMPs at the watershed scale?*

As stated earlier, this question calls for an integrated assessment involving process-based models, economics, and social considerations. Consequently, several models have to be utilized in conjunction to represent the various processes that will have a bearing on the modeling targets and the decision making process. Computational techniques can assist with quantification of modeling uncertainties, critical natural processes and key management actions, and optimal location of BMPs.

Uncertainty Analysis

Any modeling process will necessarily entail reducible and inherent uncertainty from data, model abstractions, and natural heterogeneity of watersheds. The National Research Council

report “Assessing the TMDL approach to water quality management; NRC, 2001) concluded that modeling uncertainty should be rigorously and explicitly addressed in development and application of models for environmental management, especially when stakeholders are affected by the decisions contingent upon model-supported analyses. The EPA’s guideline for TMDL development recommends accounting for the uncertainties embedded in model estimates by applying a margin of safety (MOS), i.e., $TMDL = \sum WLA + \sum LA + MOS$ where $TMDL$ =maximum pollutant load a water body can receive and still maintain water quality standards, $\sum WLA$ =point source waste load allocation, and $\sum LA$ = nonpoint source load allocation. Lack of systematic and efficient approaches for evaluation of uncertainties associated with cost-effectiveness of watershed plans has led to abandonment of explicit computation of the MOS (Dilks and Freedman, 2004).

The uncertainty estimates (i.e., MOS values) associated with absolute estimates of design quantities of agricultural NPS pollution tend to be very high because of data sparsity and model limitations (Osiede et al., 2003; Benaman and Shoemaker, 2004). Although the literature is replete with sensitivity analysis and uncertainty analysis methods, implications of uncertainty associated with model predictions have not been widely endorsed in the decision making process mainly as a result of large uncertainty estimates. The magnitude of uncertainty itself is a key factor in its acceptance as the cost of implementation of management actions such as the TMDL program may significantly increase with larger uncertainty estimates (Dilks and Freedman, 2004). The basic thesis of this report is that *if the goal of an integrated modeling study is to examine the impact of management scenarios on water quality of a study area, it may be neither practical nor necessary to incorporate large uncertainty of absolute predictions in the decision making process. It would be more feasible, and desirable, to communicate uncertainty of estimated effectiveness of management scenarios rather than uncertainty of absolute predictions.*

Identification of critical natural processes for selection of management actions

The current US EPA’s “draft handbook for developing watershed plans to restore and protect our waters” (US EPA, 2005a) recommends implementation of management practices in portions of the watershed that are believed to contribute intensively to NPS pollution. Locating critical areas, however, is complicated because contaminants are carried along with

the flow, and water movement over a watershed tends to be fairly dynamic, behaving in a nonlinear fashion. Potential interactions among hydrologic, fluvial, and nutrient processes and NPS control measures should be identified and included in designing watershed management scenarios. Nearly a hundred BMPs for NPS pollution control are recommended by the USDA Natural Resources Conservation Service (USDA NRCS, 2006). Evaluation of impacts of all these practices, even within small watersheds, is infeasible. However, it is likely that only a few management actions largely control fate and transport of contaminants within a given watershed system (Arabi et al., 2006a). We hypothesize that *the analysis of uncertainty of model simulations could highlight critical processes and key management actions that are key to control of NPS pollution in a given watershed system.*

Identification of optimal spatial configuration of management actions

Management of limited and fragile environmental resources in the face of conflicting interests is a challenging problem. An appropriate balance needs to be struck between protecting the environment and allowing users to access the natural resource base and the environment to assure their livelihoods. For example, farmers need to be able to pursue avenues that increase agricultural productivity without creating adverse environmental consequences. This includes safeguarding their plants from pests, augmenting soil fertility to raise productivity to levels required in today's modern agriculture, and accessing water. Systematic methods are required to assist policy makers in selecting strategies that lead to environmentally friendly solutions along with being cost-effective. To make informed decisions, policy makers should have the tools necessary for evaluation of alternatives and to assess trade-offs between environmental and socioeconomic criteria.

Research to date indicates the promise of heuristic optimization for cost-effective allocation of watershed management practices (Srivastava et al., 2002; Veith et al., 2004; Muleta and Nicklow, 2005). Unlike gradient-based approaches, heuristic techniques do not require linearity, continuity, or differentiability either for objective/constraint functions or for input parameters. Thus, they are well-suited for cost-effective allocation of watershed management plans. We hypothesize that *heuristic optimization techniques could facilitate spatial allocation of management actions in watersheds in a more cost-effective fashion than random and even more commonly used cost-sharing and targeting strategies.*

1.2. Objectives

The overall goal of this hypothesis-driven study is to develop a computational framework for selection and placement of BMPs in a cost-effective manner. Ultimately, the hope is that this tool will help watershed management programs, such as the TMDL program, attain the desired water quality goals for impaired water bodies. This goal is accomplished through the following specific objectives:

1. Development of an uncertainty analysis method for establishing uncertainty bounds for the estimated effectiveness of BMPs for sediment and nutrient control.
2. Development of a computational framework for identification of natural fluvial processes and BMPs that control fate and transport of sediments and nutrients at the watershed scale.
3. Development of an optimization tool for cost-effective allocation of BMPs within watersheds.

1.3. Impacts of the Study

The impacts of the present work in management of watershed systems are three fold. First, the computational analysis presented here will increase the capabilities of watershed managers to quantify and integrate modeling uncertainties in the decision making process. The tool is likely to promote stakeholder involvement through cost- a language that they understand and are able to identify with. The additional cost of implementation of watershed management plans due to incorporation of modeling uncertainties is key to adoption of scientific uncertainty estimates in the planning process. The present work presents an effective framework for analysis of uncertainties and its communication to stakeholders. Second, the analysis will allow watershed managers to asses what it takes to reduce pollutant loads/concentrations to meet water quality standards. For example, the computational analysis can be used to identify key management actions to reduce average annual nitrate concentration below 10 ppm, the EPA's drinking water standard. Third, the optimization framework presented herein could greatly increase the water quality benefits of dollars spent on conservation plans. Alternatively, the analysis could be used to attain the same level of

water quality benefits currently received from watershed management plans for substantially less costs.

1.4. Structure of the Report

The remainder of this report is organized into five sections. Section 2 provides a description of the study watersheds and the choice of the watershed model. The hypotheses and objectives of this study will be tested in two small agricultural watersheds in northeastern Indiana within the Maumee River basin located in the Midwestern portion of the U.S. A review of the choice of the watershed model, Soil and Water Assessment Tool (SWAT) will also be presented. In Section 3, a 3-step computational analysis based on Latin Hypercube Sampling, the Morris's One-at-A-Time sensitivity analysis and the Generalized Likelihood Uncertainty Estimation (GLUE) is developed and demonstrated for establishing uncertainty bounds associated with evaluation of BMPs. Application of multivariate multiobjective sensitivity analysis including Regionalized Sensitivity Analysis (RSA) and Tree-Structured Density Estimation (TSDE) for identification of critical natural processes and key management actions for sediment and nutrient control is presented in Section 4. This is followed in Section 5 by development and demonstration of a genetic algorithm (GA) search engine for identification of best location of management actions/BMPs at the watershed scale. Finally, major conclusions of the study and recommendations for future research are discussed in Section 6.

The report is primarily based on three journal papers: Section 3 is derived from Arabi et al. (2007a), Section 4 is derived from Arabi et al. (2007b), and Section 5 is derived from Arabi et al. (2006a).

SECTION 2. CASE STUDY WATERSHEDS AND THE WATERSHED MODEL

2.1. Case Study Watersheds

The utility of the optimization framework was examined for two subwatersheds in the Black Creek basin. The Black Creek watershed (Figure 2.1) located in northeast Indiana is a typical watershed in the upper Maumee River basin in the Midwestern portion of the United States. In mid 1970's and early 1980's, several BMPs were implemented in the watershed and detailed water quality monitoring was carried out at various locations within the watershed to examine short-term water quality impacts of soil and water conservation techniques (Lake and Morrison, 1977; Lake and Morison, 1978). Data collected from automated samplers at the outlets of Dreisbach (6.23 km²) and Smith Fry (7.3 km²) within the Black Creek watershed were the most complete and were used in this study. Black Creek watershed has been listed as an impaired water body in Indiana with nutrients, algal growth, and impaired biotic communities as major concerns (EPA, 2005 b).

The dominant hydrological soil group in the study watersheds is type C. Land use in Dreisbach is mostly pasture in the upper portion, while row crops are wide spread in the remainder of the watershed. Smith Fry is a predominantly agricultural watershed. Field borders, parallel terraces, grade stabilization structures, and grassed waterways were the structural BMPs installed in the watershed by targeting (see Figure 2.1). Table 2.1 summarizes the number and area under influence of each type of BMP installed in the study watersheds as shown in Figure 2.1. Available data, land use distribution, and other information for the watersheds can be obtained from Lake and Morrison (1977), Lake and Morison (1978), and Morrison and Lake (1983), and have been more recently described in Arabi et al. (2004).

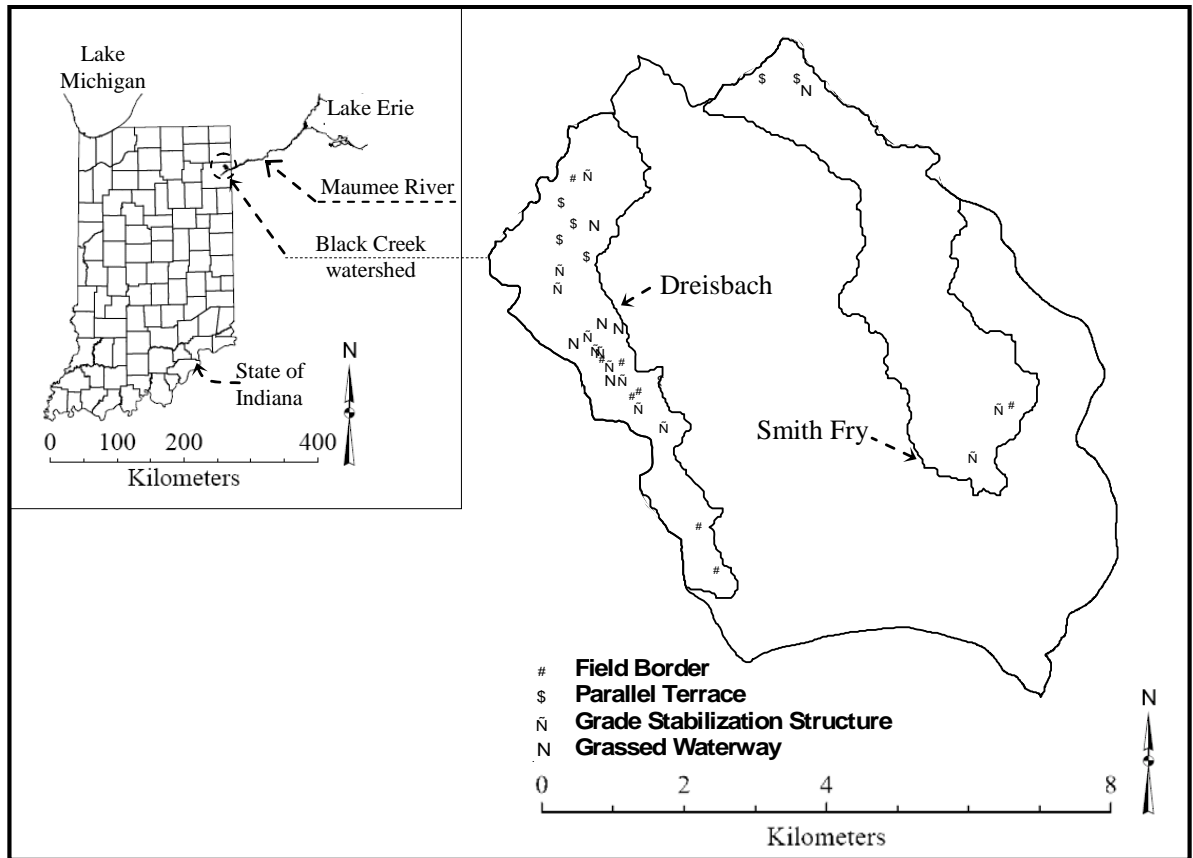


Figure 2.1 Case Study Watersheds.

| Table 2.1 BMPs installed in the Dreisbach and Smith Fry watersheds | | | | |
|---|-----------|--------------------|-----------|--------------------|
| BMP | Dreisbach | | Smith Fry | |
| | Number | Length/Area (unit) | Number | Length/Area (unit) |
| Field Border | 7 | 2600 (m) | 1 | 1800 (m) |
| Parallel Terrace | 4 | 2130 (m) | 2 | 480 (m) |
| Grassed Waterway | 5 | 3.50 (ha) | 1 | 0.95 (ha) |
| Grade Stabilization Structure | 10 | - | 2 | - |

2.2. Choice of the Watershed Model: SWAT

Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998; Arnold and Fohrer, 2005) is a process-based distributed-parameter simulation model, operating on a daily time step. The model was originally developed to quantify the impact of land management practices in

large, complex watersheds with varying soils, land use, and management conditions over a long period of time. SWAT uses readily available inputs and has the capability of routing runoff and chemicals through streams and reservoirs, and allows for addition of flows and inclusion of measured data from point sources. Moreover, SWAT has the capability to evaluate the relative effects of different management scenarios on water quality, sediment, and agricultural chemical yield in large, ungauged basins. Major components of the model include weather, surface runoff, return flow, percolation, evapotranspiration (ET), transmission losses, pond and reservoir storage, crop growth and irrigation, groundwater flow, reach routing, nutrient and pesticide loads, and water transfer. Table 2.2 provides a listing of important SWAT input parameters corresponding to the above-mentioned components. The lower and upper parameter bounds (LB and UP) were obtained from SWAT theoretical documentation (Neitsch et al., 2002) and similar studies in the literature (Lenhart et al., 2002; Benaman and Shoemaker, 2004; van Griensven et al., 2006).

For simulation purposes, SWAT partitions the watershed into subunits including subbasins, reach/main channel segments, impoundments on main channel network, and point sources. Subbasins are divided into hydrologic response units (HRUs) that are portions of subbasins with unique land use, management, and soil attributes. SWAT uses a modification of the SCS curve number method (USDA Soil Conservation Service, 1972) or Green and Ampt infiltration method (Green and Ampt, 1911) to compute surface runoff volume for each HRU. Peak runoff rate is estimated using a modification of the Rational Method. Daily or sub-daily rainfall data is used for calculations. Flow is routed through the channel using a variable storage coefficient method developed by Williams (1969) or the Muskingum routing method.

Erosion and sediment yield are estimated for each HRU with the Modified Universal Soil Loss Equation (MUSLE) (Williams, 1975). Sediment deposition and degradation are the two dominant channel processes that affect sediment yield at the outlet of the watershed. Whether channel deposition or channel degradation occurs depends on the sediment loads from upland areas and transport capacity of the channel network. If sediment load in a channel segment is larger than its sediment transport capacity, channel deposition will be the dominant process. Otherwise, channel degradation (i.e. channel erosion) occurs over the channel segment.

Table 2.2 Listing of SWAT parameters: LB and UB refer to lower and upper bounds of parameter vector, parameters identified by * were altered as a percentage of the default value, parameters identified by ** were considered in sensitivity analysis but not in uncertainty analysis.

| SWAT Symbol | Description | Units | Range | |
|-------------------|---|----------|-------|-------|
| | | | LB | UB |
| <i>CN2</i> * | SCS runoff curve number | % | -15 | 15 |
| <i>SOL_AWC</i> | available soil water capacity | mm/mm | 0.01 | 0.4 |
| <i>ESCO</i> | soil evaporation compensation factor | | 0 | 1 |
| <i>OV_N</i> | Manning's "n" value for overland flow | | 0.1 | 0.3 |
| <i>SLOPE</i> *,** | average slope steepness | m/m | 0 | 0.6 |
| <i>SLSUBSN</i> | average slope length | m | 0.15 | 1.2 |
| <i>GWQMN</i> | minimum threshold depth of water in the shallow aquifer for return | mm | 0 | 5000 |
| <i>REVAPMN</i> | minimum threshold depth of water in the shallow aquifer for "revap" | mm | 0 | 500 |
| <i>GW_REVAP</i> | groundwater "revap" coefficient | | 0.02 | 2 |
| <i>GW_DELAY</i> | groundwater delay | days | 0 | 500 |
| <i>ALPHA_BF</i> | baseflow alpha factor for recession constant | days | 0 | 1 |
| <i>SURLAG</i> | surface runoff lag time | | 1 | 12 |
| <i>SFTMP</i> | snowfall temperature | °C | -5 | 5 |
| <i>CH_S1</i> *,** | average slope for tributary channels | m/m | 0 | 10 |
| <i>CH_N1</i> | Manning's "n" value for tributary channels | | 0.008 | 0.065 |
| <i>CH_K1</i> | effective hydraulic conductivity in tributary channels | mm/hr | 0 | 150 |
| <i>CH_S2</i> *,** | average slope for the main channels | m/m | 0 | 10 |
| <i>CH_N2</i> | Manning's "n" value for the main channel | | 0.01 | 0.3 |
| <i>CH_K2</i> | effective hydraulic conductivity in the main channel | mm/hr | 0 | 150 |
| <i>CH_EROD</i> | channel erodibility factor | cm/hr/Pa | 0 | 0.6 |
| <i>CH_COV</i> | channel cover factor | | 0 | 0.6 |
| <i>PRF</i> | peak rate adjustment factor for in-stream channel routing | | 0 | 2 |
| <i>SPCON</i> | Linear coefficient for in-stream channel routing | | 0 | 0.001 |
| <i>SPEXP</i> | exponent coefficient for in-stream channel routing | | 1 | 1.5 |
| <i>BIOMIX</i> | biological mixing efficiency | | 0.01 | 1 |
| <i>BIOMIN</i> | minimum plant biomass for grazing | kg/ha | 0 | 1 |
| <i>USLE_P</i> | USLE equation support practice factor | | 0.1 | 1 |
| <i>USLE_C</i> * | minimum value of USLE equation cover factor | % | 0.001 | 0.5 |
| <i>USLE_K</i> | USLE equation soil erodibility factor | | 0.01 | 0.65 |
| <i>ORGP_AG</i> | initial organic P in soils for agriculture land use | mg/kg | 1 | 1000 |
| <i>ORGP_PAST</i> | initial organic P in soils for pasture land use | mg/kg | 1 | 500 |
| <i>LABP_AG</i> | initial soluble P in soils for agriculture land use | mg/kg | 1 | 100 |
| <i>LABP_PAST</i> | initial soluble P in soils for pasture land use | mg/kg | 1 | 50 |
| <i>ORGN_AG</i> | initial organic N in soils for agriculture land use | mg/kg | 1 | 10000 |
| <i>ORGN_PAST</i> | initial organic N in soils for pasture land use | mg/kg | 1 | 5000 |
| <i>SOLN_AG</i> | initial NO ₃ in soils for agriculture land use | mg/kg | 0.1 | 5 |
| <i>SOLN_PAST</i> | initial NO ₃ in soils for pasture land use | mg/kg | 0.1 | 3 |

Movement and transformation of several forms of nitrogen and phosphorus over the watershed are accounted for within the SWAT model. Nutrients are introduced into the main channel through surface runoff and lateral subsurface flow, and transported downstream with channel flow. Major phosphorous sources in mineral soil include organic phosphorus available in humus and mineral phosphorus that is not soluble. Phosphorus may be added to the soil in the form of fertilizer, manure, and residue application. Surface runoff is deemed as the major mechanism of phosphorus removal from a field. Unlike phosphorus that has low solubility, nitrogen is highly mobile. Major nitrogen sources in mineral soil include organic nitrogen available in humus, mineral nitrogen in soil colloids, and mineral nitrogen in solution. Nitrogen may be added to the soil in the form of fertilizer, manure, or residue application. Plant uptake, denitrification, volatilization, leaching, and soil erosion are the major mechanisms of nitrogen removal from a field.

It is worthwhile mentioning that the computational analyses proposed in this study do not necessarily require using the SWAT model as the NPS component. We have chosen SWAT because the case studies that will be performed will deal with sediments and nutrients. Soundness of mathematical representation of sediment, nutrient, and pesticide processes in SWAT has been validated in previous research (Santhi et al., 2001; Kirsch et al., 2002; Arabi et al., 2004; Vazquez-Amabile and Engel, 2006). Moreover, SWAT has been linked with optimization routines for automated calibration of the model (van Griensven and Bauwens, 2003) and for evaluation of efficiency of nonpoint source pollution regulatory programs (Whittaker et al., 2003). The optimization framework herein can be easily integrated with other hydrologic/water quality models to develop management plans for other types of pollutants.

2.2.1. Model calibration

Arabi et al. (2004) utilized a procedure adopted from Santhi et al. (2001) to calibrate and validate the SWAT model for the Dreisbach and Smith Fry watersheds. The results from this previous study indicated acceptable agreement between observed and simulated baseflow, streamflow, sediment yield, and total P and total N loads at the outlet of the study

watersheds. In this study, the adjusted value of SWAT parameters from the manual calibration in Arabi et al. (2004) were used as default values.

2.2.2. Representation of BMPs

For this study, a method presented by Arabi et al. (2004) was utilized to evaluate the water quality impacts of grassed waterways, grade stabilization structures, field borders and parallel terraces. The method was developed based on published literature pertaining to BMP simulation in hydrological models and considering the hydrologic and water quality processes simulated in SWAT. Based on the function of the BMPs and hydrologic and water quality processes that are directly modified by their implementation, corresponding SWAT parameters were selected and altered. The impacts of the BMPs on other hydrologic/water quality processes were indirectly accounted for through their functional representation within the SWAT model. For example, implementation of grassed waterways prevents gully erosion due to concentrated flow. Thus, channel cover and erodibility factors were set to zero for channel segments with grassed waterways. Also, grassed waterways increase deposition of sorbed pollutant loads in the channel network ((Fiener and Auerswald, 2003). This impact is captured by increasing the channel Manning's number to 0.3 (Fiener and Auerswald, 2006). For parallel terraces, Wischmeier and Smith (1978) recommended a USLE practice factor of 0.1 for terraces on slopes less than 15% with graded channels sod outlets. Table 2.3 summarizes SWAT parameters and their corresponding values for representation of the BMPs. More information on BMP representation procedure with SWAT model can be obtained from Arabi et al. (2004), and Arabi et al. (2006b).

It is worthwhile to mention that the numerical representation of BMP performances as presented in Table 2.3 bears some degree of uncertainty. For example, Fiener and Auerswald (2006) assumed CH_N2 ranges between 0.3-0.4 $s\ m^{-1/3}$ over the year for grassed waterways in case of dense grasses and herbs under non-submerged conditions. Assuming a fixed value for the representative parameters adds additional uncertainty in the BMP evaluation methodology. This uncertainty will perhaps be alleviated once precise numerical representation procedures are developed and validated with experimental data. We have

adopted the methodology provided by Arabi et al. (2004) without explicitly incorporating the uncertainty of BMP representation procedure.

| Table 2.3 Representation of BMPs with SWAT | | | | |
|---|----------------------------|-----------------------------|---------|----------------------------|
| BMP | Function | Representing SWAT Parameter | | |
| | | Parameter (SWAT input file) | Range | value when BMP implemented |
| Field Border | increase sediment trapping | <i>FILTERW</i> (.hru) | 0-5 (m) | 5 (m) |
| Parallel Terrace | reduce overland flow | <i>CN2</i> (.mgt) | 0-100 | ^a |
| | reduce sheet erosion | <i>USLEP</i> (.mgt) | 0-1 | 0.1 (terraced) |
| | reduce slope length | <i>SLSUBBSN</i> (.hru) | 10-150 | ^b |
| Grassed Waterway | increase channel cover | <i>CH_COV</i> (.rch) | 0-1 | 0.0 (fully protected) |
| | reduce channel erodibility | <i>CH_EROD</i> (.rch) | 0-1 | 0.0 (non-erosive) |
| | increase channel roughness | <i>CH_N2</i> (.rch) | 0-0.3 | 0.24 |
| Grade Stabilization Structure | reduce gully erosion | <i>CH_EROD</i> (.rch) | 0-1 | 0.0 (non-erosive) |
| | reduce slope steepness | <i>CH_S2</i> (.rch) | - | ^c |

^a Estimated based on land use and hydrologic soil group of the HRU where it is installed for terraced condition.

^b Estimated for each parallel terrace based on its features and SWAT assigned overland slope of the HRU where it is installed:

$$SLSUBBSN = (A \times S + B) \times 100 / S$$
where S is average slope of the HRU; $A=0.21$ and $B=0.9$ (ASAE, 2003).

^c Estimated for each grade stabilization structure based on its features and SWAT assigned slope, CH_S2_{old} , and length of the channel segment where it is installed:

$$CH_S2_{new} = CH_S2_{old} - D / CH_L2$$
where D is height of the structure (1.2 m) and CH_L2 is length of the channel segment.

SECTION 3. PROBABILISTIC APPROACH FOR ANALYSIS OF UNCERTAINTY IN THE EVALUATION OF WATERSHED MANAGEMENT PRACTICES

3.1. Abstract

A computational framework is presented for analyzing the uncertainty in model estimates of water quality benefits of best management practices (BMPs) in two small (<10 km²) watersheds in Indiana. The analysis specifically recognizes the significance of the difference between the magnitude of uncertainty associated with absolute hydrologic and water quality predictions, and uncertainty in estimated benefits of BMPs. The Soil and Water Assessment Tool (SWAT) is integrated with Monte Carlo-based simulations, aiming at (1) adjusting the suggested range of model parameters to more realistic site-specific ranges based on observed data, and (2) computing a scaled distribution function to assess the effectiveness of BMPs. A three-step procedure based on Latin Hypercube Sampling (LHS), the Morris's One-factor-At-a-Time (OAT) sensitivity analysis and the Generalized Likelihood Uncertainty Estimation (GLUE) was implemented for the two study watersheds. Results indicate that the suggested range of some SWAT parameters, especially the ones that are used to determine the transport capacity of channel network and initial concentration of nutrients in soils, required site-specific adjustment. It was evident that uncertainties associated with sediment and nutrient outputs of the model were too large, perhaps limiting its application for point estimates of design quantities. However, the estimated effectiveness of BMPs sampled at different points in the parameter space varied by less than 10% for all variables of interest. This suggested that BMP effectiveness could be ascertained with good confidence using models, thus making it suitable for use in watershed management plans such as the EPA's Total Maximum Daily Load (TMDL) program. The potential impact of our analysis on utility of models and model uncertainties in decision making process is discussed.

3.2. Introduction

The analysis of uncertainty associated with the utility of simulation models is an important consideration in the development of watershed management plans. Modeling uncertainty should be rigorously addressed in development and application of models, especially when stakeholders are affected by the decisions contingent upon model-supported analyses (NRC, 2001). Watershed models are commonly utilized to investigate rainfall-runoff generation, and fate and transport of contaminants resulting from nonpoint source activities. Nonpoint source activities are perceived to be the most important source of pollution in the United States (Ice, 2004). The evaluation of the success of best management practices (BMPs) in meeting their original goals has also been facilitated by watershed models (Griffin, 1995; Edwards et al., 1996; Mostaghimi et al., 1997; Saleh et al., 2000; Santhi et al., 2001; Kirsch et al., 2002; Santhi et al., 2003; Arabi et al., 2006b). Uncertainty associated with absolute estimates of design quantities tends to be very high because of data sparsity and model limitations (Osiedle et al., 2003; Benaman and Shoemaker, 2004). Thus, models are found to be more useful when making relative comparisons rather than making absolute predictions. It may be more meaningful to implement the uncertainty associated with effectiveness of BMPs in the planning process.

The common modeling approach entails the {calibrate → validate → predict} process. The thrust of the calibration procedure is to identify a set of model parameters by optimizing a goodness-of-fit statistic between observed and predicted values such as the Nash-Sutcliff coefficient of efficiency. The calibrated model is then used to examine the impact of various management scenarios on the future behavior of the system. Such an analysis is subject to identifiability, and non-uniqueness of the optimal (calibrated) parameter set (Beck, 1987), i.e. there may be several sets of model parameters that fit the observed data equally (Beven and Binley, 1992). Calibration of a simulation model for a given watershed will reduce, but not totally remove, modeling uncertainties associated with both structure of the model and parameter estimates. Even with the best model structure, parameter estimation contains residual uncertainty (Beck, 1987) that propagates forward into model predictions and evaluation of effectiveness of management practices.

Although the literature is replete with sensitivity analysis and uncertainty analysis methods (Spear and Hornberger, 1980; Beven and Binley, 1992; Spear et al., 1994; Saltelli et al., 2000), implications of uncertainty associated with model predictions have not been widely endorsed in the decision making process mainly as a result of large uncertainty estimates. The magnitude of uncertainty itself is a key factor in its acceptance as the cost of implementation of management actions such as the TMDL program may significantly increase with larger uncertainty estimates (Dilks and Freedman, 2004). In a case study in the Cannonsville Reservoir system watershed (1178 km²) located in upstate New York, Benaman and Shoemaker (2004) concluded that even in the presence of observed data it was not possible to reduce the uncertainty of absolute sediment predictions in their study to practical values for the TMDL program. The argument, however, is that if the goal of a modeling study is to examine the impact of management scenarios on water quality of a study area, it may be neither practical nor necessary to incorporate large uncertainty of absolute predictions in the decision making process. It would be perhaps more feasible (and more desirable) to communicate and implement uncertainty of estimated effectiveness of management scenarios rather than uncertainty of absolute predictions (Zhang and Yu, 2004). Moreover, the importance of such a formulation would be particularly appreciated when the reduction of a variable of concern (sediment, nutrients, etc) due to implementation of an abatement action is less than estimated uncertainty of absolute predictions. In such cases, evaluation of impact of management scenarios would not be inhibited by uncertainty of model outputs.

The impact of modeling uncertainties on evaluation of management practices has not been addressed sufficiently, as studies have generally focused on uncertainty of point predictions. Specifically, a computational procedure that can be used to establish uncertainty bounds for the estimated effectiveness of BMPs has not been developed to the best of our knowledge. In this study, a Monte Carlo-based probabilistic approach is utilized (i) to develop a computational procedure for analysis of uncertainty; (ii) to examine the effect of modeling uncertainties on evaluation of long-term water quality impacts of BMPs using a distributed watershed model, SWAT; and (iii) to provide a comparison between magnitudes of uncertainties associated with absolute predictions versus effectiveness of BMPs. The analysis

is demonstrated for two small watersheds in Indiana where water quality data were collected and several structural BMPs were implemented.

3.3. Theoretical Considerations

3.3.1. Morris's One-at A Time sensitivity analysis

The OAT is a sensitivity analysis technique that falls under the category of screening methods (Saltelli et al. 2000). In the OAT, each model run involves perturbation of only one parameter in turn. This way, the variation of model output can be unambiguously attributed to the perturbation of the corresponding factor. For each input parameter, local sensitivities are computed at different points of the parameter space, and then the global (main) effect is obtained by taking their average. The elementary effect of a small perturbation Δ of the i^{th} component of the p -dimensional parameter vector (α_i) at a given point in the parameter space $\alpha = (\alpha_1, \dots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \dots, \alpha_p)$ is (Morris, 1991):

$$d(\alpha_i | \alpha) = \frac{[y(\alpha_1, \dots, \alpha_{i-1}, \alpha_i + \Delta, \alpha_{i+1}, \dots, \alpha_p) - y(\alpha)]}{\Delta} \quad \text{Eq 3.1}$$

where $y(\alpha)$ corresponds to model output. The results are quantitative, elementary, and exclusive to the parameter α_i . However, the elementary effect computed from Eq 3.1, i.e. $d(\alpha_i | \alpha)$, is only a partial effect and depends on the values chosen for the other elements of the parameter vector (α_j). A finite distribution (F_i) of elementary effects of parameter α_i is obtained by sampling at different points of the space, i.e. different choices of parameter set α . The mean of the distributions is indicative of the overall influence of the parameter on the output, while the variance demonstrates interactions with other parameters and nonlinearity effects.

3.3.2. Generalized Likelihood Uncertainty Estimation (GLUE)

The GLUE methodology is based on recognition of the importance of the set of parameters to produce the behavior of the system, not individual parameters. The acceptable model realizations (i.e. behavior set) obtained from Monte Carlo simulations are given a likelihood

weight according to observed data and a likelihood function. Several choices for an appropriate likelihood function can be obtained from Beven and Freer (2001). A likelihood measure based on Nash-Sutcliff efficiency criterion with shaping factor N is defined as (Freer et al., 1996):

$$L(\alpha | y) = (1 - \frac{\sigma_{\varepsilon}^2}{\sigma_o^2})^N; \quad \sigma_{\varepsilon}^2 < \sigma_o^2 \quad \text{Eq 3.2}$$

where $L(\alpha|y)$ is the likelihood of parameter set (α), given the observed data (y). The quantities σ_{ε}^2 and σ_o^2 refer to the error variance between model simulations and observed data, and the variance of the observed data, respectively. For $N=1$, L in Eq 3.2 is the well-known Nash-Sutcliff efficiency coefficient that is often used for calibration of hydrologic and water quality models. A negative L value indicates that the corresponding model output is dissimilar to the behavior of the system under study, and the likelihood of such a simulation in mimicking the system behavior is zero. Therefore, the likelihood measure is rewritten as:

$$L(\alpha | y) = \max[(1 - \frac{\sigma_{\varepsilon}^2}{\sigma_o^2}), 0] \quad \text{Eq 3.3}$$

For each simulation from a random parameter set, a likelihood weight is obtained from Eq 3.3. Then, these weights are rescaled by dividing each of them by their total sum. The rescaled likelihood weights are used to construct a cumulative distribution for the output of interest, which can be used for estimation of uncertainty bounds associated with the output by computing its quantiles. The GLUE method is subjective not only to the choice of likelihood function, but also to the cutoff criterion used for classification of Monte-Carlo simulations to behavior and non-behavior sets.

3.4. Methodology

The computational framework utilized in this study is comprised of (1) a process-based watershed model (SWAT; Soil and Water Assessment Tool, Arnold et al., 1998) employed for simulating the fate and transport of sediment and nutrients in the study watersheds under two scenarios- before and after implementation of BMPs, (2) a BMP representation method

(adapted from Arabi et al., 2004) for incorporation of water quality impacts of BMPs, and (3) an uncertainty analysis methodology based on two sampling based methods: One-factor-At-a-Time (OAT) sensitivity analysis (Saltelli et al., 2000), and Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992). A computational method was developed to reduce the “suggested” range of input parameters to site-specific “adjusted” ranges to be used in the GLUE analysis. The methodology was tested on two different watersheds located in northeast Indiana, one with significantly larger number of installed BMPs. This confirmed the versatility of the computational analysis to quantify modeling uncertainties associated with estimated effectiveness of BMPs in study watersheds under influence of a large number as well as a small number of BMPs.

3.4.1. OAT sensitivity analysis

Identifying the most sensitive input parameters is not an essential component of the framework for analysis of uncertainty that is proposed in this study. The results of such an analysis, however, provide useful information pertaining to model parameters with large uncertainties. The uncertainty associated with parameters that are determined based on topographic, land use, soil, and management attributes could perhaps be reduced by using better spatial resolution for these attributes. Conversely, uncertainty of parameters that are not determined from landscape attributes and are usually estimated through calibration procedure may be reduced only through a systematic site-specific range adjustment process.

In this study, the following procedure was followed for the OAT sensitivity analysis. First, in the absence of any prior information with regard to the distribution of input parameters, a uniform distribution, with minima and maxima specified in Table 2.2, was considered for all parameters. The same approach was used by Beven and Freer (2001) and Benaman and Shoemaker (2004). Then, OAT sensitivity indices were computed. While a local sensitivity index (d) was computed by applying Eq 3.1 for each parameter, global sensitivities were determined by taking the average of these local sensitivities at 50 random points sampled across the entire parameter space. Finally, a rescaled sensitivity index (d_n) was determined by dividing the global OAT sensitivity indices by their total sum. The d_n values range from [0,1], with values closer to 1 indicating a more sensitive parameter.

3.4.2. Range adjustment

Adjustment of “suggested” ranges of sensitive input parameters was found to be an essential step prior to quantifying uncertainty of model outputs. Use of unrealistically large parameter ranges may result in a large number of simulations with negative likelihoods (L in Eq 3.2) that renders the GLUE analysis inefficient and prohibitively expensive. Particularly, ranges of calibration parameters that are not determined from landscape attributes should be investigated. For example, consider the model parameters that are used to determine the transport capacity of the channel network ($SPCON$, $SPEXP$, and PRF in Table 2.2). These parameters are not identified based on topographic and/or soil characteristics of the channel segments, but typically determined based on similar research in the study area or through a calibration procedure.

The range adjustment process for important model parameters entailed the following steps:

1. *Establish base-case values for input parameters:* the base-case values for the input parameters in Table 2.2 could be selected by one the following approaches: (i) insights gained from the OAT sensitivity analysis; (ii) model calibration procedure; (iii) a suggested value obtained from literature, previous studies in the study area, or prior experience of the analyst. In this analysis, base-case values were selected from a manual calibration (Arabi et al., 2004).
2. *Perform interval-spaced simulations:* the range of each parameter of concern was divided into 50 equal intervals. A SWAT model simulation was performed for a random realization of the parameter from each interval while other parameters were kept at their base-case values. The simulation timeline covered the 1974-1978 period when hydrologic and water quality data were collected at the outlets of the study watersheds.
3. *Adjust ranges of important parameters:* A method based on a goodness-of-fit measure (Nash-Sutcliff coefficient) was developed to reduce the suggested ranges of important parameters to narrower adjusted ranges for the study watersheds. The acceptable ranges of parameters were determined by plotting the Nash-Sutcliff efficiency coefficients (E_{N-S} ; Nash and Sutcliff, 1970) computed for interval-spaced simulations. The E_{N-S} is defined as:

$$E_{N-S} = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad \text{Eq 3.4}$$

where O_i and P_i are observed and predicted monthly output variables, respectively, while \bar{O} is the average of monthly observed values. E_{N-S} ranges from $-\infty$ to 1, with higher values indicating a better 1:1 fit between the observed and simulated values. Model simulations with negative E_{N-S} were considered “unacceptable” (Santhi et al., 2001), and subsequently, the range of the parameter was reduced to those that yielded non-negative E_{N-S} values.

3.4.3. Uncertainty computations

Adjusted parameter ranges from previous step were used to probabilistically determine the effectiveness of BMP implementation. In the present uncertainty analysis, 5000 model simulations were performed for the 1974-1978 period when water quality data were collected. All model parameters in Table 2.2 were included in the analysis. For those parameters that required range reduction, adjusted ranges were selected from the range adjustment process. For each realization of the parameter space, the uncertainty analysis entailed the following steps:

1. Select parameter values from their specified ranges in a random fashion;
2. Perform model simulation with the selected values to compute model outputs for the scenario without BMPs;
3. Compute likelihood of the simulation for monthly output variables by applying Eq 3.3;
4. Represent BMPs by changing appropriate model parameters from Table 2.3;
5. Compute model outputs for the scenario with BMPs in place;
6. Using GLUE-likelihood weights computed at step (3), generate a cumulative density function for outputs of interest for both model simulations with and without BMPs.

The main assumption in the numerical procedure described above is that the same cumulative likelihood can be used for the scenarios with BMPs and without BMPs. Water quality data utilized in this study were representative of the state of the watersheds before implementation

of BMPs. Thus, the likelihood of a set of model parameters to mimic the behavior of the system was computed for the scenario without BMPs. When adequate hydrologic and water quality data are available for both scenarios with BMPs and without BMPs, the GLUE-likelihood functions should be computed separately.

It should be noted that likelihood functions associated with each set of model parameters in the GLUE analysis are likely to be approximately the same before and after implementation of BMPs, if other watershed characteristics remain the same. Representation of BMPs entails only modification of some model parameters for the fields and/or streams where BMPs are implemented. BMPs are implemented in a relatively small area within the watershed. For example, BMPs in Figure 2.1 cover less than 5% of total upland fields and the channel network in the Dreisbach and Smith Fry watersheds. Thus, in the absence of water quality data for computation of GLUE-likelihoods for the scenario with or without BMPs, it is suggested that the same likelihood be used for both scenarios.

3.5. Results

3.5.1. Range adjustment results

The results of the range adjustment analysis indicated that none of hydrology-related parameters of the SWAT model needed range reduction. On the other hand, three sediment-related parameters, one total P-related parameter, and one total N parameter required site specific adjustments. Table 3.1 provides a listing of top five sensitive SWAT parameters in a descending order for monthly streamflow, sediment, total P, and total N output variables. The term “min (E_{N-S})” in the table refers to the minimum of Nash-Sutcliffe coefficients computed for 50 interval-spaced model simulations corresponding to each parameter. Shaded values indicate parameters with negative “min (E_{N-S})” value for which the range adjustment procedure was applied.

Table 3.1 Top five sensitive SWAT parameters for streamflow, sediment, total P, and total N computations: d_n is the OAT sensitivity index, “min (E_{N-S})” refers to minimum of Nash-Sutcliff coefficient computed for each 50 simulations in the OAT analysis. Description of parameters can be found from Table 2.2.

| Variable | SWAT Symbol | d_n | min (E_{N-S}) | |
|------------|-----------------|-------|-------------------|-----------|
| | | | Dreisbach | Smith Fry |
| Streamflow | <i>SOL_AWC</i> | 0.30 | 0.19 | 0.13 |
| | <i>CN</i> | 0.20 | 0.38 | 0.13 |
| | <i>GW_REVAP</i> | 0.17 | 0.27 | 0.36 |
| | <i>ESCO</i> | 0.09 | 0.61 | 0.56 |
| | <i>CH_KI</i> | 0.09 | 0.40 | 0.60 |
| Sediment | <i>SPCON</i> | 0.30 | -30.59 | -4.9 |
| | <i>PRF</i> | 0.12 | -0.31 | -0.85 |
| | <i>SLOPE</i> | 0.09 | -2.07 | 0.11 |
| | <i>CH_N2</i> | 0.06 | -0.27 | -0.2 |
| | <i>CN</i> | 0.06 | 0.60 | 0.01 |
| Total P | <i>SLOPE</i> | 0.43 | -110.02 | -10.66 |
| | <i>ORGP_AG</i> | 0.25 | -1.69 | -0.19 |
| | <i>LABP_AG</i> | 0.07 | 0.03 | 0.4 |
| | <i>CN</i> | 0.06 | 0.09 | 0.12 |
| | <i>USLE_K</i> | 0.05 | 0.12 | 0.05 |
| Total N | <i>SLOPE</i> | 0.51 | -132.94 | -11.77 |
| | <i>ORGN_AG</i> | 0.11 | 0.03 | -0.5 |
| | <i>USLE_K</i> | 0.06 | 0.12 | 0.05 |
| | <i>CN</i> | 0.07 | 0.18 | 0.05 |
| | <i>SOL_AWC</i> | 0.06 | 0.14 | 0.34 |

Figure 3.1 illustrates one dimensional response surface of model outputs at the outlets of the study watersheds to parameters with reduced ranges. The spider-plots (dashed lines) have been generated by varying one parameter at a time, while other parameters were kept constant at their base-case value. Each panel represents 50 model simulations, where the model output is shown on the left y-axis and the right y-axis displays the Nash-Sutcliff coefficient in Eq 3.4. The x-axis is the normalized value of each parameter $\bar{\alpha}_i$, determined from its absolute value α_i , and its respective upper (U_i) and lower (L_i) bounds summarized in Table 2.2:

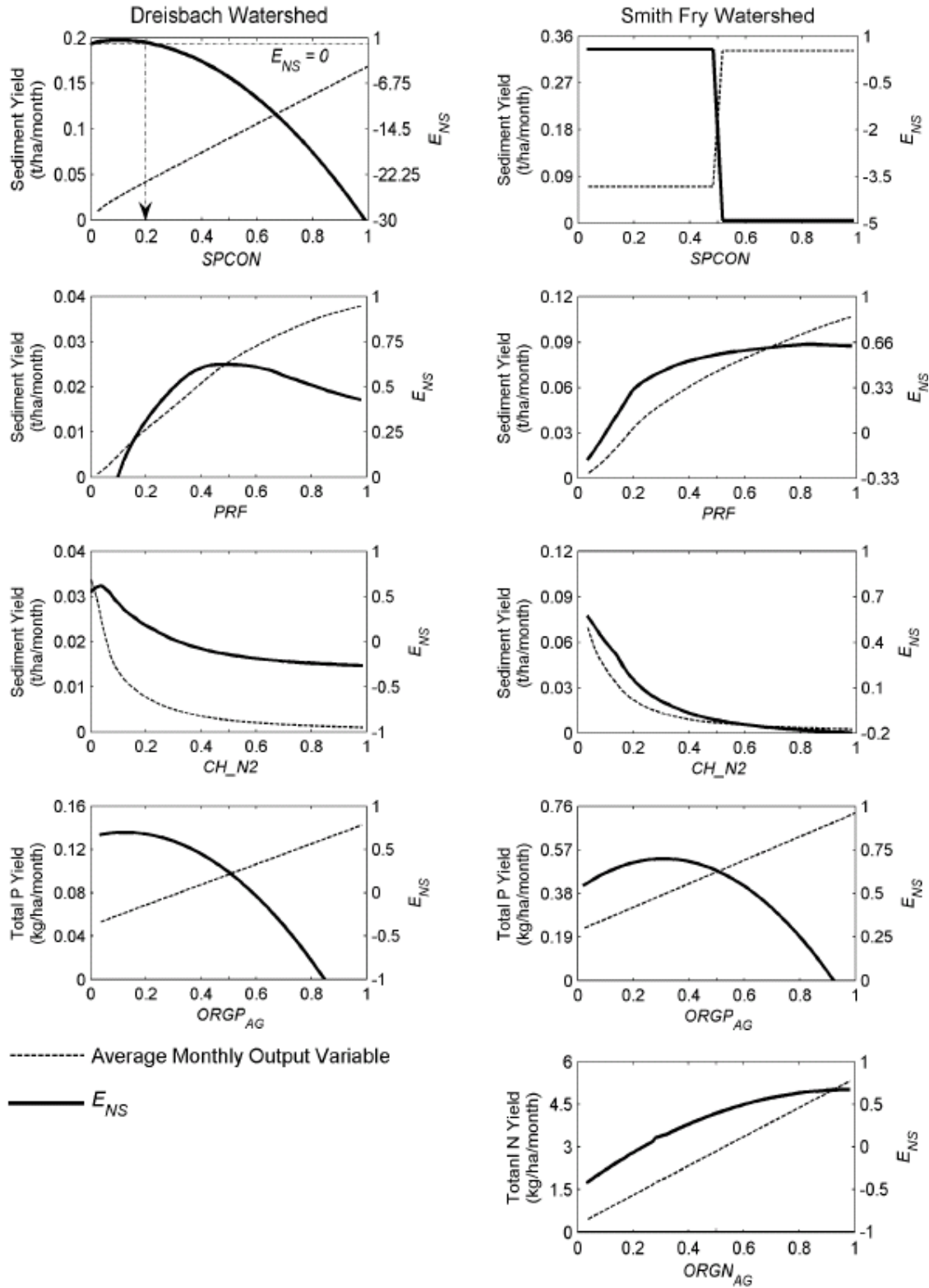


Figure 3.1 Spider plots for the most sensitive input factors related to water quality computations in SWAT with adjusted ranges in Table 3.2. Each panel represents 50 model realizations. Subscript “AG” refers to agricultural land use.

$$\bar{\alpha}_i = \frac{\alpha_i - L_i}{U_i - L_i} \quad \text{Eq 3.5}$$

Equation 3.5 can be used to back calculate the absolute parameter values corresponding to the normalized values shown in Figure 3.1. A negative “min (E_{N-S})” value indicated the necessity for range reduction. For each panel, a horizontal line was drawn at E_{N-S} (right y-axis) equal to zero. The range of the corresponding parameter vector was reduced to the portion that lies above this line. The top panel to the left in Figure 3.1 demonstrates how parameter ranges were adjusted. The new parameter ranges for the study watersheds are presented in Table 3.2.

Table 3.2 Adjusted parameter ranges for the study watersheds: both suggested ranges and adjusted ranges are based on absolute parameter values. Description of the parameters can be found in Table 2.2.

| Parameter | Units | Limiting Variable | Suggested Range | Adjusted Range | |
|----------------|-------|-------------------|-----------------|----------------|------------|
| | | | | Dreisbach | Smith Fry |
| <i>SPCON</i> | - | sediment | 0-0.001 | 0-0.0002 | 0-0.0005 |
| <i>PRF</i> | - | sediment | 0-2 | 0.2-2 | 0.2-2 |
| <i>CH_N2</i> | - | sediment | 0.008-0.3 | 0.008-0.1 | 0.008-0.1 |
| <i>ORGP_AG</i> | mg/kg | total P | 0-1000 | 0-500 | 0-950 |
| <i>ORGN_AG</i> | mg/kg | total N | 0-10000 | 0-10000 | 2000-10000 |

Sediment outputs were most sensitive to model parameters that are used to estimate the transport capacity of the channel network, usually determined through calibration procedure and not based on land and/soil/management characteristics. The slope of upland areas represented by parameter *SLOPE* was the most sensitive parameter for total P and total N simulations. However, the range of this parameter was not reduced, because it is directly estimated from Digital Elevation Model (DEM). In the absence of field measurements, the initial concentration of phosphorus and nitrogen in soils were reduced to the values reported in Table 3.2.

Several issues should be considered in interpretation of the results from range adjustment analysis. First, the threshold value of Nash-Sutcliff coefficient that is used to reduce

parameter ranges significantly impacts the method. Obviously, a lower E_{N-S} value would result in acceptance of wider parameter ranges. In this study, this threshold was set at zero because the likelihood of model simulations in GLUE uncertainty analysis with negative E_{N-S} values is assumed to be zero (see Eq 3.3). Second, the number of intervals to be used in the range adjustment process should be selected such that the plot of the goodness-of-fit measure (E_{N-S}) over the entire parameter range (e.g. Figure 3.1) forms a relatively smooth curve. Our experience from this study indicated that 20-50 intervals should be adequate. The results of the OAT sensitivity analysis and range adjustment process are site specific, and may be different if the watershed conditions are significantly different than those discussed here. We note that the results of the sensitivity analysis are in general agreement with previous studies such as Lenhart et al. (2002), Heuvelmans et al. (2004), Muleta and Nicklow (2005), and White and Chaubey (2005).

3.5.2. Uncertainty analysis results

The computational procedure described previously was performed to quantify the uncertainty associated with both model simulations and estimated effectiveness of BMPs. Figures 3.2 and 3.3 depict the cumulative likelihood distribution of average monthly values for various output constituents at the outlet of the Dreisbach and Smith Fry watersheds, respectively. A likelihood weight from Eq 3.3 was assigned to each of 5000 model realizations resulting from Monte Carlo simulations. Each panel in the figures contains two cumulative probability distributions for the two scenarios. The dashed lines correspond to results from the GLUE method for the scenario without BMPs (scenario A), while the solid lines demonstrate the results for the scenario with BMPs represented in the model (scenario B). Two major trends are evident. First, the difference between GLUE-likelihoods for scenarios A and B was marginal for streamflow, but substantial for sediment, total P, and total N computations. Second, the trends were quite different in the two study watersheds.

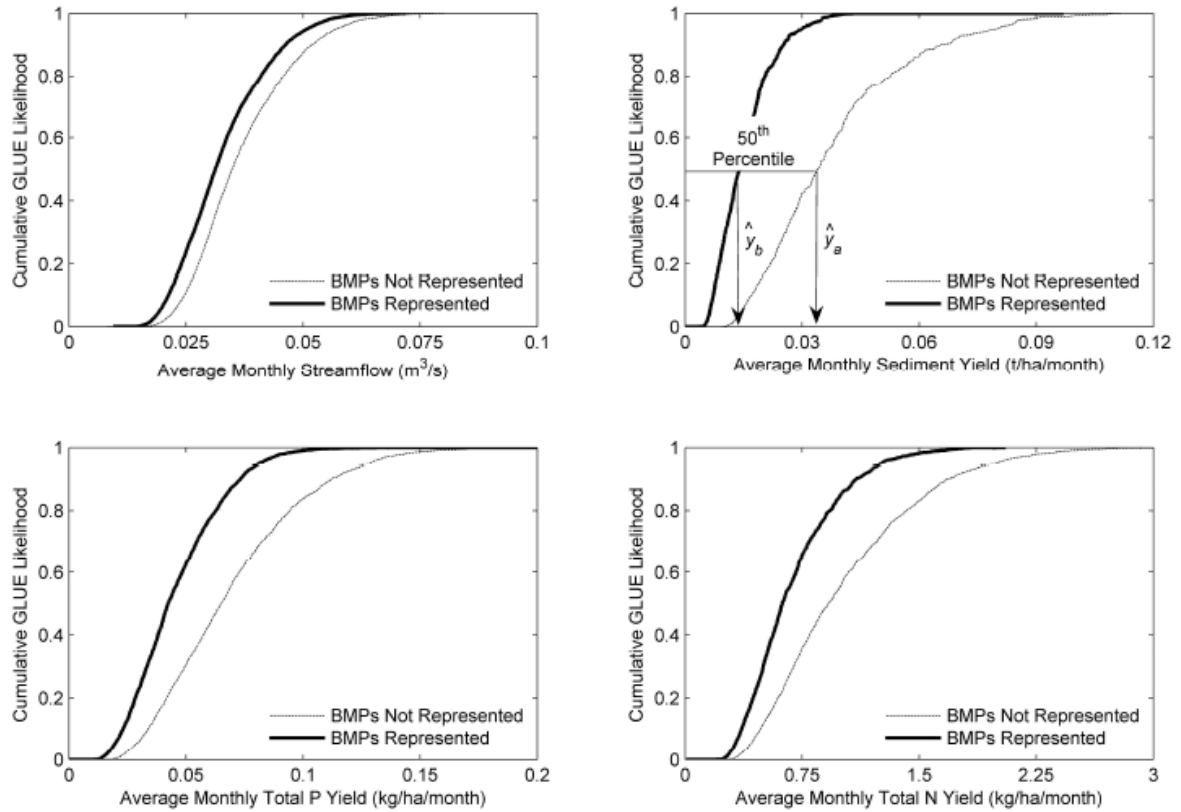


Figure 3.2 Cumulative GLUE-likelihood for variables simulated at the outlet of the Dreisbach watershed, Indiana. Each panel represents 5000 model realizations. The top panel in the right demonstrates how various percentiles were determined. Variable \hat{y}_a is the expected value for scenario A (BMPs not represented), and \hat{y}_b is the expected value for scenario B (BMPs represented).

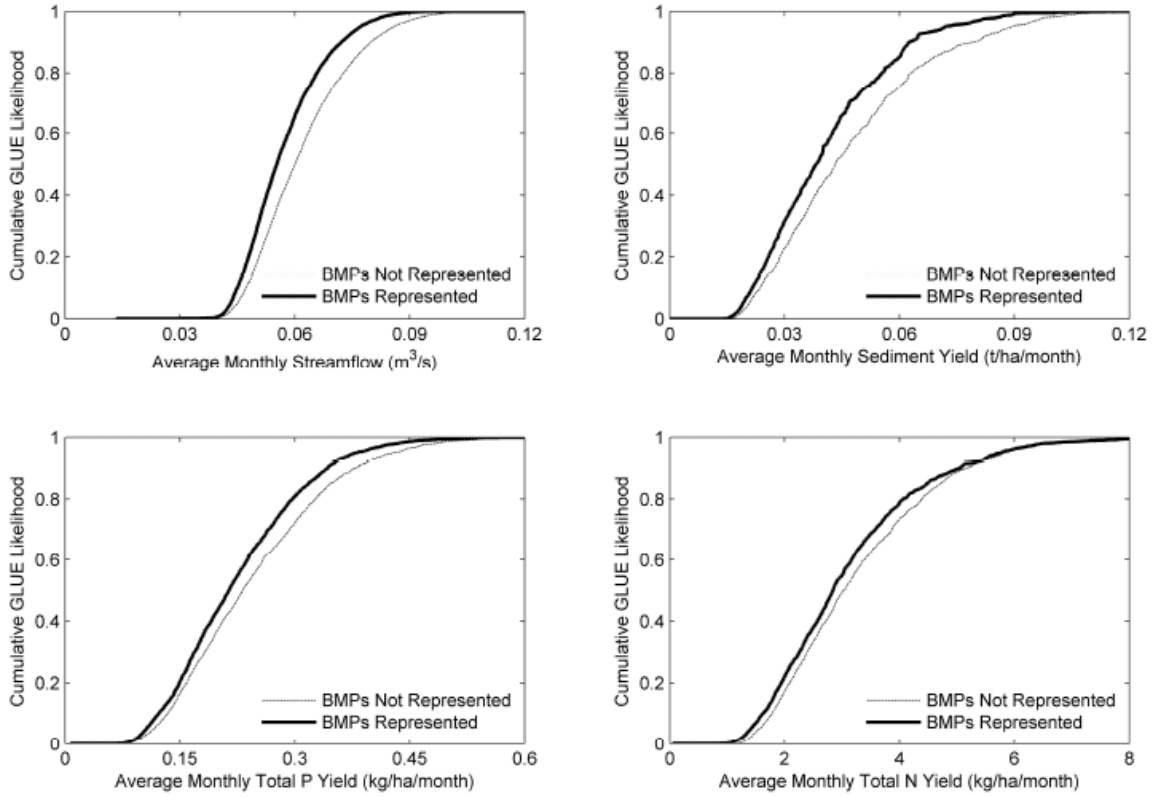


Figure 3.3 Cumulative GLUE-likelihood for variables simulated at the outlet of the Smith Fry watershed, Indiana. Each panel represents 5000 model realizations.

Following Figures 3.2 and 3.3, the results are summarized in Table 3.3 for further discussion. Lower and upper bounds in Table 3.3 refer to 5th and 95th percentiles of the cumulative likelihoods, respectively, while the median represents the 50th percentile. The top panel on the right in Figure 3.2 demonstrates how various percentiles were determined. For scenario A, comparison of median values to the average of monthly observed data, also presented in Table, indicate that averages of monthly observed values for streamflow were within $\pm 15\%$ of the median in both study watersheds, and fell well within the uncertainty ranges from GLUE. Sediment yield measurements at the outlet of the Dreisbach and Smith Fry watersheds were within $\pm 20\%$ of the median values, also well covered by the uncertainty bounds. The median values for total P appeared to adequately match the average of monthly observations as they underpredicted by only 27% and 20% in Dreisbach and Smith Fry, respectively. Moreover, the uncertainty bounds covered the data. The median for total N determined for Dreisbach watershed differed from the average of monthly observed values by only 7%, indicating satisfactory results from GLUE. Total N at the outlet of Smith Fry was the only case where the uncertainty bounds from GLUE did not contain the observed value. The corresponding median underpredicted total N yield by nearly 50%, perhaps because of structural uncertainties involved in simulations. In particular, a lack of proper linkage between fluvial channel processes and in-stream nutrient processes in SWAT could be responsible for these results.

Table 3.3 Summary of the results from analysis of uncertainty: variable \hat{y} refers to 50th percentile (median) of output variables for 5000 Monte Carlo simulations of the SWAT model, \hat{r} is the 50th percentile (median) of estimated reduction of output variables due to implementation of BMPs computed by 3.3 for 5000 Monte Carlo simulations of the SWAT model, and “range” refers to the 5th and 95th percentile of the Monte Carlo simulations.

| Watershed | Variable | Units | Observed Value | Scenario A Without BMPs | | Scenario B With BMPs | | Reduction (%) | |
|-----------|------------|-------------------|----------------|----------------------------|---------------|-------------------------|---------------|---------------|-------------|
| | | | | \hat{y}_a | range | \hat{y}_b | range | \hat{r}_i | range |
| Dreisbach | Streamflow | m ³ /s | 0.039 | 0.041 | [0.026,0.062] | 0.037 | [0.023,0.056] | 9.8 | [9.70,11.5] |
| | Sediment | t/ha/month | 0.031 | 0.035 | [0.015,0.083] | 0.015 | [0.007,0.033] | 57.1 | [53.3,60.2] |
| | Total P | kg/ha/month | 0.095 | 0.069 | [0.030,0.125] | 0.046 | [0.015,0.086] | 33.3 | [31.2,4,50] |
| | Total N | kg/ha/month | 1.228 | 1.145 | [0.487,2.385] | 0.788 | [0.360,1.502] | 31.2 | [26.1,37.0] |
| Smith Fry | Streamflow | m ³ /s | 0.054 | 0.06 | [0.045,0.086] | 0.055 | [0.042,0.077] | 8.3 | [6.70,10.5] |
| | Sediment | t/ha/month | 0.093 | 0.043 | [0.021,0.107] | 0.037 | [0.018,0.090] | 14 | [12.8,15.9] |
| | Total P | kg/ha/month | 0.336 | 0.267 | [0.115,0.551] | 0.243 | [0.098,0.502] | 9 | [8.90,14.8] |
| | Total N | kg/ha/month | 6.112 | 3.026 | [1.552,5.838] | 2.847 | [1.450,5.560] | 5.9 | [4.70,6.60] |

These values may now be compared to corresponding quantities for scenario B listed in Table 3.3. Comparison of the expected values for scenarios A and B revealed the estimated effectiveness of BMPs, shown as a percent reduction. The percent reduction of each constituent as a result of implementation of BMPs was determined as:

$$r_i = \frac{\hat{y}_{i,a} - \hat{y}_{i,b}}{\hat{y}_{i,a}} \times 100 \quad \text{Eq 3.6}$$

where i is the constituent of interest, i.e., streamflow, sediment yield, total P load, or total N load, r_i is percent reduction of constituent i , $\hat{y}_{i,a}$ is the median of constituent i for scenario A, and $\hat{y}_{i,b}$ is the median of constituent i for scenario B. The results indicated that implementation of BMPs as shown in Figure 2.1 would not reduce streamflows significantly. This did not come as a surprise because the BMPs implemented in the watersheds were mostly sediment control BMPs. Implementation of the 26 BMPs in the Dreisbach watershed would lower median sediment yield by nearly 57%, from 0.035 t/ha/month to 0.015 t/ha/month. The estimated reduction rates for total P and total N at the Dreisbach outlet were 33% and 31%, respectively. The estimated reduction rates at the outlet of Smith Fry were nearly 14%, 9%, and 6% for sediment yield, and total P and total N loads (Table 3.3). These rates suggested that implementation of the BMPs in Smith Fry was almost four times less effective than the ones in Dreisbach, which was anticipated because of smaller number of BMPs in the former.

The uncertainty bounds associated with absolute predictions of the SWAT model were much larger than the ones corresponding to the estimated effectiveness of BMPs. For example, the uncertainty bound for monthly sediment yield at the outlet of Dreisbach for the simulation period under scenario A (scenario with no BMP represented) was determined to be [0.015,0.083] with a median value of 0.035 t/ha/month (Table 3.3). Incorporation of this large uncertainty in decision making and management may be extremely costly and not feasible. However, as is evident in Table 3.3, the estimated 5th and 95th uncertainty bounds for estimated effectiveness of BMPs in the Dreisbach watershed in reducing sediment yield at its outlet were [53%,60%] with a median (50th percentile) of 57%. Likewise, the estimated

effectiveness of BMPs in Smith Fry was estimated to range between [13%, 16%] with a median of 14%. Similar results were observed for streamflow, total P, and total N computations for the study watersheds.

3.6. Discussion

The conceptual simplicity is an attractive feature of the computational analysis developed in this study. The utilized likelihood measure, Nash-Sutcliff coefficient, is widely used by modelers for calibration and validation of watershed models. Also, the OAT-GLUE methodology will, more than likely, produce various sets of model parameters that adequately satisfy calibration criteria that are usually set based on Nash-Sutcliff coefficient. This helps modelers avoid the cumbersome practice of manual calibration. That the uncertainty bounds of various constituents encompassed the corresponding observed values (except for only one case for total N in the Smith Fry watershed) strongly supports this suggestion. It appears that for the Dreisbach and Smith Fry watersheds, sediment and nutrient outputs of the SWAT model bear more uncertainty than streamflow simulations.

Consolidation of results in Tables 3.1 and 3.3 from the three-step procedure indicate that reducing the uncertainty associated with absolute sediment and nutrient outputs of the SWAT model to practical ranges may not be feasible. Sediment outputs of the model were found to be most sensitive to transport capacity of channel network. Some of the parameters that are used to determine the transport capacity of channel segments (*SPCON*, *SPEXP*, and *PRF* in Table 2.2) cannot be measured in the field and are usually calibrated from a broad “suggested” range. With availability of more data, the three-step procedure used in this study may result in more narrow uncertainty bounds for absolute predictions. Nevertheless, it may still not be adequate for reducing the uncertainty associated with these absolute model predictions to small enough ranges to be useful for practical management decisions. However, *comparison* of sediment and nutrient outputs from model simulations with and without representation of BMPs would not suffer from such limitations as demonstrated here. It is evident that incorporation of modeling uncertainty in informed decision making is more practical through communicating the uncertainty associated with evaluation of effectiveness of management actions.

The foregoing results indicate that the computational framework developed in this study could be endorsed by watershed management programs such as the TMDL as advocated in NRC (2001). Once the level of protection to be provided by the Margin of Safety (MOS) is specified by decision makers, the probabilistic framework presented in this study could be applied to verify the success of a particular management action in achieving its designated goals. Implementation of the estimated MOS is more feasible in this context, as opposed to the MOS computed from uncertainty analysis of absolute model predictions where the additional cost of implementation may render its consideration impractical. Finally, the methodology can be easily incorporated in watershed models such as SWAT that are commonly used for TMDL development. Land use in the study watershed has changed since data were collected. Thus, application of the results presented herein is limited to demonstration of the methodology.

3.7. Conclusions

A computational framework was developed in which investigation of uncertainty provides complementary quantitative and qualitative information to support management and decision making. The analysis focused specifically on two issues. First, ranges of some model parameters may require site-specific adjustments with available data. This was particularly evident for parameters that are not determined from landscape characteristics. Second, the uncertainty associated with estimated effectiveness of BMPs is substantially smaller than the uncertainty associated with absolute predictions. The computational procedure for analysis of uncertainty was performed for two Indiana watersheds with relatively similar spatial scale, and landscape characteristics, but different number of installed BMPs. It was demonstrated that the computational framework is capable of quantifying uncertainty of effectiveness of implemented BMPs in both watersheds. Future work will focus on investigation of the utility of the developed methodology in the TMDL process. Especially, the means for coping with the Margin of Safety (MOS) in the TMDL formula should be explored.

SECTION 4. SENSITIVITY ANALYSIS OF SEDIMENT AND NUTRIENT PROCESSES WITH A WATERSHED MODEL

4.1. Abstract

This section presents a computational analysis for evaluating critical nonpoint source sediment and nutrient processes and management actions at the watershed scale. In the analysis, model parameters that bear key uncertainties are presumed to reflect the importance of natural processes and/or management actions that they represent. The Regionalized Sensitivity Analysis (RSA) and the Tree-Structured Density Estimation (TSDE) procedures were integrated with the Soil and Water Assessment Tool (SWAT) to investigate correlation structure in the parameter space while accounting for multiple objectives. The computational analysis was applied to the Dreisbach and Smith Fry watersheds in Indiana in the Midwestern portion of the United States. Results showed that incorporation of parameter interactions is essential to obtaining conclusive information about critical system processes and management actions. Interactions between surface runoff volume and within-channel processes were critical to describe transport of sediments in the study watershed. Key management actions for nutrient control were found to be fertilizer application and upland farming practices such as parallel terraces. The sensitivity analysis reported herein could be used to derive a list of key nonpoint source best management practices for development of watershed management plans.

4.2. Introduction

Identification of natural processes and management actions that control nonpoint source (NPS) pollution is essential for development of watershed management plans. The current handbook of the Environmental Protection Agency for developing watershed plans (EPA, 2005a) recommends implementation of best management practices (BMPs) in portions of the watershed that are believed to contribute intensively to NPS pollution. Locating critical areas, however, is complicated because contaminants are carried along with the flow, and water

movement over a watershed tends to be fairly dynamic, behaving in a nonlinear fashion. Potential interactions among hydrologic, fluvial, and nutrient processes and NPS control measures should be identified and included in designing watershed management scenarios. Such information could be obtained by integration of computational techniques with hydrologic/water quality models.

Previous studies have utilized optimization methods to design near optimal combination of BMPs (Srivastava et al., 2002; Veith et al., 2004). These methods search for the optimal location of selected management practices for NPS pollution control in a watershed system. As the number of selected BMPs increases, the number of model evaluations for identifying the optimal solution increases exponentially. Thus, these optimization studies have focused on spatial allocation of only a few BMPs. Nearly a hundred BMPs for NPS pollution control are recommended by the USDA Natural Resources Conservation Service (USDA NRCS, 2006). Evaluation of impacts of all these practices, even within small watersheds, is infeasible. In this section, we present a novel approach to systematically abridge the list of key BMPs for sediment and nutrient control at the watershed scale. The methodology is based on the hypothesis that analysis of uncertainty of model simulations could underscore critical processes and key management actions for NPS control.

Inverse modeling approaches based on sensitivity analysis (SA) have been used in the past to obtain information about important system processes in a variety of disciplines. Osidele et al. (2003) utilized a sensitivity analysis to investigate the importance of sediment and nutrient processes in a section of the Chattahoochee River (nearly 115 river miles) in Atlanta. The univariate regionalized sensitivity analysis (RSA; Spear et al., 1980) was used in conjunction with the multivariate tree structured density estimation (TSDE; Spear et al., 1994) to account for the correlation structure in the parameter space. The RSA was also utilized by Zheng and Keller (2006) for sensitivity analysis of the Watershed Analysis Risk Management Framework (WARMF) model and its management implications.

The one-at-a time sensitivity analysis (Pitman, 1994), factorial design, and the Fourier amplitude sensitivity test (FAST) (Saltelli et al., 2000) are some other SA methods. These SA methods determine parameter sensitivities for a single model response (e.g. average

annual sediment yield). Bastidas (1998) developed a multiobjective generalized sensitivity analysis (MOGSA) based on RSA to consider the influence of multiple criteria on parameter sensitivities. Previous research studies based on MOGSA (Bastidas et al., 1999; Meixner et al., 1999; Liu et al., 2004, Demarty et al., 2005), while incorporating multiple criteria, were limited in revealing the multivariate correlation structure in the parameter space. Moreover, these studies have been applied only to models up to medium level of complexity (Zheng and Keller, 2006).

The work of Osidele et al. (2003) and van Griensven et al. (2006) are examples of studies that applied multivariate approaches for sensitivity analysis of sediment and nutrient processes. However, they evaluated model sensitivities based on a single output or its trajectory, and did not explicitly address the interactions between flow, sediment, and nutrient criteria. To our knowledge, both multivariate interactions and multiple criteria have not been incorporated simultaneously for performing a sensitivity analysis with a complex watershed model such as SWAT (Arnold et al., 1998). Also, implications of sensitivity of NPS sediment and nutrient processes/BMPs in holistic watershed management have not been discussed. These shortcomings are addressed in the present work.

In this section, we investigate the utility of computational approaches for identifying critical NPS processes and key BMPs that are likely to control fate and transport of sediments and nutrients in watershed systems. To this end, the RSA and TSDE procedures are linked with a comprehensive watershed model (SWAT) to (i) develop a new sensitivity analysis framework that can handle interactions in the parameter space as well as multiple trajectories of model outputs, and (ii) demonstrate the application of the computational analysis to highlight critical sediment and nutrient processes in an agricultural watershed in Indiana. With simultaneous incorporation of multicriteria and multivariate approaches, the developed framework provides a novel capability for dealing with important issues for holistic watershed management. Critical NPS sediment and nutrient processes, key agricultural BMPs, and their potential interactions can be evaluated. Watershed management programs such as the Total Maximum Daily Load (TMDL) and the USDA's Environmental Quality Incentive Program (EQIP) would significantly benefit from such an analysis. Another feature of the present work is that real system response data are utilized in the analysis.

4.3. Theoretical Considerations

4.3.1. Regionalized Sensitivity Analysis (RSA)

The Regionalized Sensitivity Analysis (RSA; Spear and Hornberger, 1980) aims at identification of critical uncertainties in order to evaluate the relative importance of individual parameters that exert the most influence on system behavior. The procedure utilizes a uniform sampling of the parameter space and involves (i) a qualitative definition of the behavior of the system under study (criterion function C), and (ii) a binary classification of the parameter space (S) into good (behavior B) or bad (nonbehavior \bar{B}) regions. The strength of the method lies in the classification scheme, which facilitates the application of multivariate statistical methods to explore the level of significance that the posterior probability distribution function of each element of the parameter vector $\vec{\alpha}$ in the behavior region $f_m(\vec{\alpha} | B)$ deviates from the one in the nonbehavior region $f_n(\vec{\alpha} | \bar{B})$. A Kolmogorov-Smirnov two sample test is performed to test the hypothesis that for a given parameter $\alpha_k \in \vec{\alpha}$, $f_m(\alpha_k | B)$ separates from $f_n(\alpha_k | \bar{B})$:

$$\begin{aligned} H_0: & \quad f_m(\alpha_k | B) = f_n(\alpha_k | \bar{B}) \\ H_1: & \quad f_m(\alpha_k | B) \neq f_n(\alpha_k | \bar{B}) \end{aligned} \tag{Eq 4.1}$$
$$\text{Kolmogorov-Smirnov statistic: } d_{m,n} = \sup_x |F_m(\alpha_k | B) - F_n(\alpha_k | \bar{B})|$$

where $F_m(\alpha_k | B)$ and $F_n(\alpha_k | \bar{B})$ are, respectively, the cumulative density functions corresponding to $f_m(\alpha_k | B)$ and $f_n(\alpha_k | \bar{B})$ for m behavior and n nonbehavior model simulations. The \sup_x notation refers to the largest vertical separation between $F_m(\alpha_k | B)$ and $F_n(\alpha_k | \bar{B})$. The $d_{m,n}$ statistic can be used to determine the relative importance of the uncertainty associated with each element of the parameter vector, with higher values indicating higher influence on model outputs.

4.3.2. Tree Structured Density Estimation (TSDE)

Spear et al. (1994) recognized that despite the conceptual simplicity of the RSA procedure, its applicability may be limited as a result of different correlation structures among

parameters. The tree structured density estimation (TSDE) methodology was developed to obtain information relevant to the interactions between model parameters in the complex non-uniform behavior region. The TSDE procedure is as follows. The m behavior parameter vectors are treated as independent samples from an unknown probability distribution function $f(\vec{\alpha} | B)$. To construct an adequate approximation of this unknown distribution, the behavior parameter space (S_B) is partitioned into q sub-spaces ($S_B = \bigcup_{i=1}^q S_{B,i}$). A local estimate of f for the i^{th} sub-space ($S_{B,i}$) with volume (V_i) can be defined as (Spear et al., 1994):

$$\hat{f}_i(\vec{\alpha} | B) = \frac{m_i}{m} \times \left(\frac{1}{V_i}\right) \quad \text{Eq 4.2}$$

where m_i is the number of behavior parameter vectors in $S_{B,i}$. With p reflecting the number of individual model parameters, the $[m \times p]$ parameter hypercube can be split into q sub-spaces in $q \times p$ ways. The search algorithm for finding the optimal split involves minimizing a loss function (L) that mimics deviation of $\hat{f}(\alpha_k | B)$ from $f(\alpha_k | B)$ and can be defined as (Spear et al., 1994):

$$L = -\sum_{i=1}^q (\hat{f}_i^2 \times V_i) \quad \text{Eq 4.3}$$

If the parameters were uniformly distributed in the parameter space S_B (i.e. no split was required), the first estimated density function (\hat{f}_0) and the first loss function (L_0) would be equal to $1/V$ and $-1/V$, respectively, where V reflects the volume of the behavior space (B). The splitting process is performed successively, and begins with splitting the parameter space into two sub-spaces. While the loss function corresponding to the first split L_1 is always less than L_0 , $L_0 - L_1$ is a measure of accuracy of approximation of the density function. Therefore, maximizing $L_0 - L_1$ is synonymous with minimizing the errors associated with density estimation. The parameter space is split on the axis of the parameter that produces the largest increase in the accuracy criterion ($L_0 - L_1$). Likewise, in the second recursion, each of the two sub-spaces constructed in the first step is split into two new sub-spaces. This procedure is

repeated until either the accuracy of density estimate does not increase significantly or the density of each sub-space is less than some critical value.

The TSDE procedure finally yields a tree structure that reveals the multivariate correlations among model parameters. The top (origin) node in the tree represents the original sample space with normalized relative density, defined as the volume of each sub-space to the volume of the original space, equal to 1. The density of end (terminal) nodes indicates the relative importance of the corresponding intermediate parameters and their interactions for matching the behavior of the system under study. Beginning from the origin node and ending with a terminal node, each branch graphically depicts the interactions between model parameters.

4.4. Methodology

The computational analysis presented in this section was aimed at identification of hydrologic and water quality processes that are likely to control transport of sediments and nutrients. Model parameters served as surrogates for processes that they represent in the model's mathematical structure. The resulting ranking and classification of the parameters based on the importance of their uncertainties indicated the relative importance of the system processes they represent. Inferences relevant to key management actions for sediment and nutrient control were drawn.

The computational procedure shown in Figure 4.1 was implemented as follows. First, the natural process and/or management action represented by each input parameter of the SWAT model was identified. Model parameters and their suggested ranges are presented in Table 2.2. For example, *CH_N2* represents the impact of roughness of the channel network on sediment and nutrient transport. Likewise, *USLE_P* indicates the importance of upland farming practices such as parallel terraces (Renard et al., 1997). Next, 5000 parameter vectors were randomly generated with a Latin Hypercube Sampling (LHS; McKay et al., 1979) strategy. The SWAT model was used to simulate monthly flow, sediment and nutrient outputs corresponding to each parameter vector. Model inputs and outputs were integrated with the RSA and the TSDE methods to investigate significance of input factors.

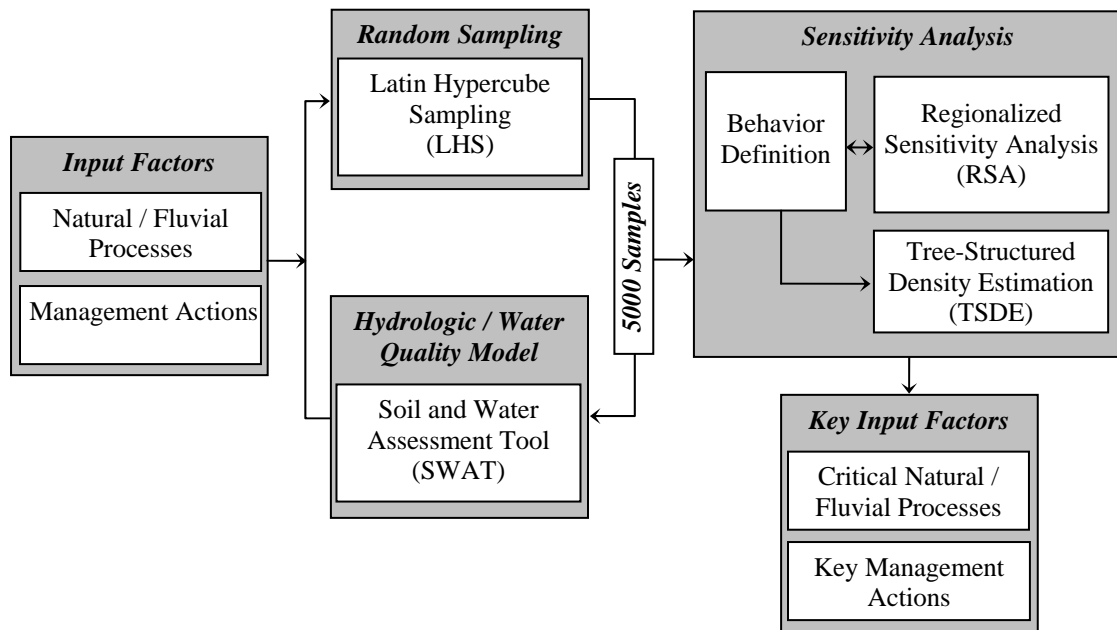


Figure 4.1 Computational framework for sensitivity analysis of sediment and nutrient processes.

4.4.1. Behavior Definition

The definition of the behavior of a given system refers to a combination of thresholds, extremes, and time scales derived from available or proposed information about the system conditions. This combination defines a *corridor* through which the model output trajectory must pass in order to qualify as a behavior simulation (Beck, 1987). Depending on the goals of the study, any of the following can define the system behavior:

- analysis of observed data,
- speculations with regard to future state of the system,
- desired state of the system based on regulatory standards, and
- target values for TMDL implementation.

The definition of the behavior of a given watershed system depends on the goal of the problem at hand (Bastidas et al., 1999). In the present work, model simulations were compared to the observed data and classified as behavior or non-behavior based on the Nash-Sutcliffe efficiency coefficient E_{N-S} . Available data for the watersheds can be obtained from Arabi et al. (2004). For streamflow, E_{N-S} associated with streamflow output should be greater

than or equal to zero in order to classify the simulation as behavior. For sediment yield, E_{N-S} associated with both streamflow and sediment output should be greater than or equal to zero for a behavior simulation. A simulation was deemed behavior for the total N constituent only if the E_{N-S} values associated with all three streamflow, sediment, and total N outputs were greater than or equal to zero. The Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970) that is commonly used for calibration of hydrologic models can be defined as:

$$E_{N-S} = 1.0 - \left\{ \left[\sum_{i=1}^N (y_i - \hat{y}_i)^2 \right] / \left[\sum_{i=1}^N (y_i - \bar{y})^2 \right] \right\} \quad \text{Eq 4.4}$$

where y and \hat{y} refer to measured and simulated output variables, respectively, and \bar{y} is the average of the N measured values ($\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$).

It should be noted that in the current version of the SWAT model, the in-stream nutrient processes have not been linked with the model components that pertain to sediment routing in the channel network. This is problematic, particularly for total P computations, because phosphorus is typically bound to sediment and is carried out of most catchments with sediments. In this study, the results obtained for sediment yield are used for identification of control processes and critical management actions for control of total P yield.

4.5. Results

Figures 4.2-4.3 depict the cumulative marginal distributions for the sediment and total N related parameters with highest rank based on the Kolmogorov-Smirnov ($d_{m,n}$) statistic for the system behavior. The $d_{m,n}$ test statistic was determined from largest vertical separation between the cumulative marginal distribution of behavior and non-behavior parameter distributions. Kolmogorov-Smirnov test statistic ($d_{m,n}$) has been graphically illustrated on the first panel in Figure 4.2. Tables 4.1-4.4 summarize the results of the RSA procedure for sensitivity of input factors listed in Table 2.2. Figures 4.4-4.7 depict the TDSE diagram for sediment and total N at the outlet of the study watersheds, while Tables 4.5-4.8 provide a summary of the diagrams.

The top three ranks for sediment yield at the outlet of Dreisbach and Smith Fry watersheds were channel Manning's number $CH-N2$, peak rate adjustment factor PRF , and average main channel slope $CH-S2$. The ranking indicated that fluvial processes within the main channel of the study watersheds were the control processes for meeting the specified behaviors. Main channel sediment processes within the SWAT model include channel deposition and channel erosion (degradation). Whether a channel segment is undergoing erosion or deposition is determined by comparing its estimated transport capacity with the sediment concentration in the streamflow. The transport capacity of channel segments is estimated as a function of peak flow velocity. SWAT uses Manning's equation for estimation of flow velocity in the channel network. Then, a peak rate adjustment factor is applied to determine the peak velocity in the segment. Channel Manning's number $CH-N2$, peak rate adjustment factor PRF , and average slope of the channel segment $CH-S2$ interact directly in the mathematical expression of peak flow velocity of channel segments. Such formulation allows for a negative correlation between channel Manning's number and peak rate adjustment factor. This negative correlation can be graphically identified by interpretation of their spider-plots in Figure 4.2. Sediment yield at the outlet of the study watersheds decreased with increasing $CH-N2$ value, while an inverse trend was observed for PRF . In the case of directly correlated model parameters, without desire for their *true* values, it is sufficient to adjust only one of the interacting parameters, and leave the others fixed. Nevertheless, the relative importance of these parameters leads to the same inferences about the control processes and management actions. Indirect correlations occur among state variables of the model such as surface runoff and sediment yield at the outlet. The RSA procedure is essentially a univariate analysis of system behavior and only suggests the *main* effects of input factors. Therefore, it is incapable of revealing any form of correlation structure among directly or indirectly correlated parameters. A multivariate analysis such as the TSDE procedure was utilized for such purposes.

Figures 4.4-4.5 show the TSDE trees for sediment yield at the outlet of Dreisbach and Smith Fry for system behavior specified by the observed past. The summary of main attributes of the terminal nodes is provided in Tables 4.5-4.6, sorted in order of descending relative density. High Density Terminal Nodes (HDTN) were selected from the top of the list such that they contained at least 75% of the input parameter sets (points); i.e. the sum of their

corresponding percentage volumes was more than 75 (%). All three HDTNs for the Dreisbach watershed were defined by the sediment transport process parameters {*CH-N2*; *PRF*; *CH-S2*}. The factors related to percolation {*SOL-AWC*} and surface runoff {*CN2*} processes defined two HDTNs, but at a lower level in the tree. Likewise the sediment transport process parameters {*CH-N2*; *PRF*; *CH-S2*} defined both HDTNs identified for the Smith Fry watershed, while the percolation process parameter {*SOL-AWC*} defined only one of them at a lower level. All of these input factors appeared in the RSA classification tables (Tables 4.1-4.4), except for the available water capacity of the soil layer (*SOL-AWC*) for the observed-past behavior defined for Smith Fry. Thus, the multivariate TSDE analysis confirmed the rankings obtained from the univariate RSA method. Based on the RSA-TSDE procedure, it can be concluded that sediment transport processes were the control processes for sediment yield at the outlet of the study watersheds. Therefore, implementation of Best Management Practices (BMPs) that influence fluvial processes within the channel network such as grassed waterways and grade stabilization structures would be appropriate choices to reduce sediment yield at the outlet of the Dreisbach and Smith Fry watersheds. The same conclusions could possibly be drawn for total P yield because phosphorus is mainly transported as a sediment bound constituent. A similar analysis can be performed for total P, once the computation of phosphorus transport through the channel network in SWAT is linked with sediment transport components of the model.

Average slope steepness of upland fields, parameter *SLOPE*, had the top RSA rank for total N load for observed-past behavior in the Dreisbach watershed (see Table 4.5). Variable *SLOPE* is a representative topographic attribute of upland areas in the watershed, and is used by SWAT to estimate the USLE topographic factor used for computation of sheet erosion from upland areas by Modified Universal Soil Estimation Equation (MUSLE). Other factors in the MUSLE equation including USLE support practice factor and USLE cover factor also appear in the list of top RSA ranks. Therefore, sheet erosion from upland areas could be deemed as the critical process that controls total N yield at the outlet of the Dreisbach watershed for the specified behavior. Another high ranking parameter was the initial concentration of organic nitrogen in the soil layer in agricultural areas, denoted by *ORGN-AG*. This indicated that the nitrogen loads generated at the agricultural fields were important to match system behavior. Although *ORGN-AG* refers to the initial concentration of N in the

soil layer, its importance reveals the prominence of anthropogenic agricultural activities such as fertilizer and manure application, which increase the amount of nitrogen susceptible to being transported by surface runoff in upland areas. Figure 4.6 depicts the TSDE tree for total N yield at the outlet of Dreisbach for the observed-past behavior, while Table 4.7 summarizes the main attributes of the terminal nodes and three identifiable HDTNs. The topographic attribute of upland areas represented by parameter {*SLOPE*} defined all HDTNs, suggesting that implementation of BMPs such as parallel terraces that reduce upland slope steepness would be critical management actions for control of total nitrogen yield at the outlet of Dreisbach watershed. The other sheet erosion process parameters {*USLE-K*}, and also parameters representing agricultural activities that increase concentration of nitrogen in the soils such as fertilizer application {*ORGN-AG*} defined two HDTNs in a lower tree level. Consolidation of the RSA and TSDE results suggest that management actions that focus on reduction of nitrogen loadings from upland areas are critical to meet the desired target values in the Dreisbach watershed.

The results of the TSDE procedure for total N yield at the outlet of the Smith Fry watershed are shown in Figures 4.7 and Table 4.8. Suggestions similar to those made for the Dreisbach watershed can be inferred for the Smith Fry watershed.

4.6. Discussion

Consolidation of the RSA-TSDE results shows great promise for identification of critical processes and key management actions through analysis of uncertainty of model simulations. Particularly, evaluation of effectiveness of non-point source pollution control scenarios such as implementation of Best Management Practices (BMPs) could be achieved by applying this method. The common modeling approach is to rely on outputs of a calibrated process-based model to investigate the impact of management scenarios on pollutant transport within watersheds. However, this approach can be criticized as a result of its inability to identify a unique set of process parameters that describe the behavior of the system, often referred to as problems of identifiability and multiple optima. The success of the combined RSA-TSDE method is in its ability to identify the critical processes and management actions without being handicapped by limitations of common parameter estimation procedures. Model

realizations obtained from the Monte Carlo routine in the RSA-TSDE procedure can be further utilized for calculating the margin of safety (MOS) in development of total daily maximum loads (TMDLs) as suggested in NRC (2001) and Benaman and Shoemaker (2004).

4.7. Conclusions

The conventional utility of process-based watershed models {calibrate → validate → predict} for evaluating the impact of management scenarios on fate and transport of sediments and nutrients was demonstrated in previous sections. In this section, a computational framework was developed in which investigation of uncertainty provides complementary quantitative and qualitative information in support of management and decision making. The analysis focused specifically on two issues. First, the univariate Regionalized Sensitivity Analysis (RSA) ranking method was applied in conjunction with the multivariate Tree Structured Density Estimation (TSDE) analysis to identify critical hydrologic and water quality processes that control the ability to attain specified water quality conditions. The implications of such analysis provide benefits for watershed management programs such as TMDL. Under TMDL agenda, Best Management Practices (BMPs) should be implemented to reduce pollutant loads into the water bodies. Because BMPs are typically implemented under a restricted budget, identification of critical processes and management actions becomes highly desirable. Application of the RSA-TSDE procedure for two small watersheds in Indiana, Dreisbach (6.25 km²) and Smith Fry (7.3 km²), showed that fluvial channel processes appeared to control sediment yield at the outlet of the watersheds. Therefore, implementation of within-channel BMPs such as grassed waterways and grade stabilization structures are likely more effective for reducing sediment loads at the outlet of the study watersheds. However, total nitrogen yield at the outlets was controlled by upland nitrogen loading, indicating that implementation of parallel terraces or fertilizer application strategies would be more effective. Also, EPA has been advised to promote research on translating effectiveness needs into Use Attainability Analysis (UAA) implications (NRC, 2001). The adoption of UAA, which is applied for setting new water quality standards or revising the existing ones, has been restricted due to inadequate technical guidance from EPA. The results of the TSDE method can provide a numerical probability

value that can be utilized to answer pragmatic questions such as how feasible a target value is for a desired constituent.

The developed methodology can be utilized to inform management about attainability of desired goals for a given watershed, and key management action(s) for sediment and nutrient non-point source control. Optimization techniques should be applied to determine the [near] optimal selection of these management actions based on cost-benefit analysis. For example, the results of the RSA-TSDE procedure indicated that within-channel BMPs such as grassed waterways and grade stabilization structures would be most effective for sediment control in both Dreisbach and Smith Fry watersheds. On the other hand, implementation of upland nitrogen control plans such as parallel terraces and field borders was deemed more effective for total nitrogen reduction.

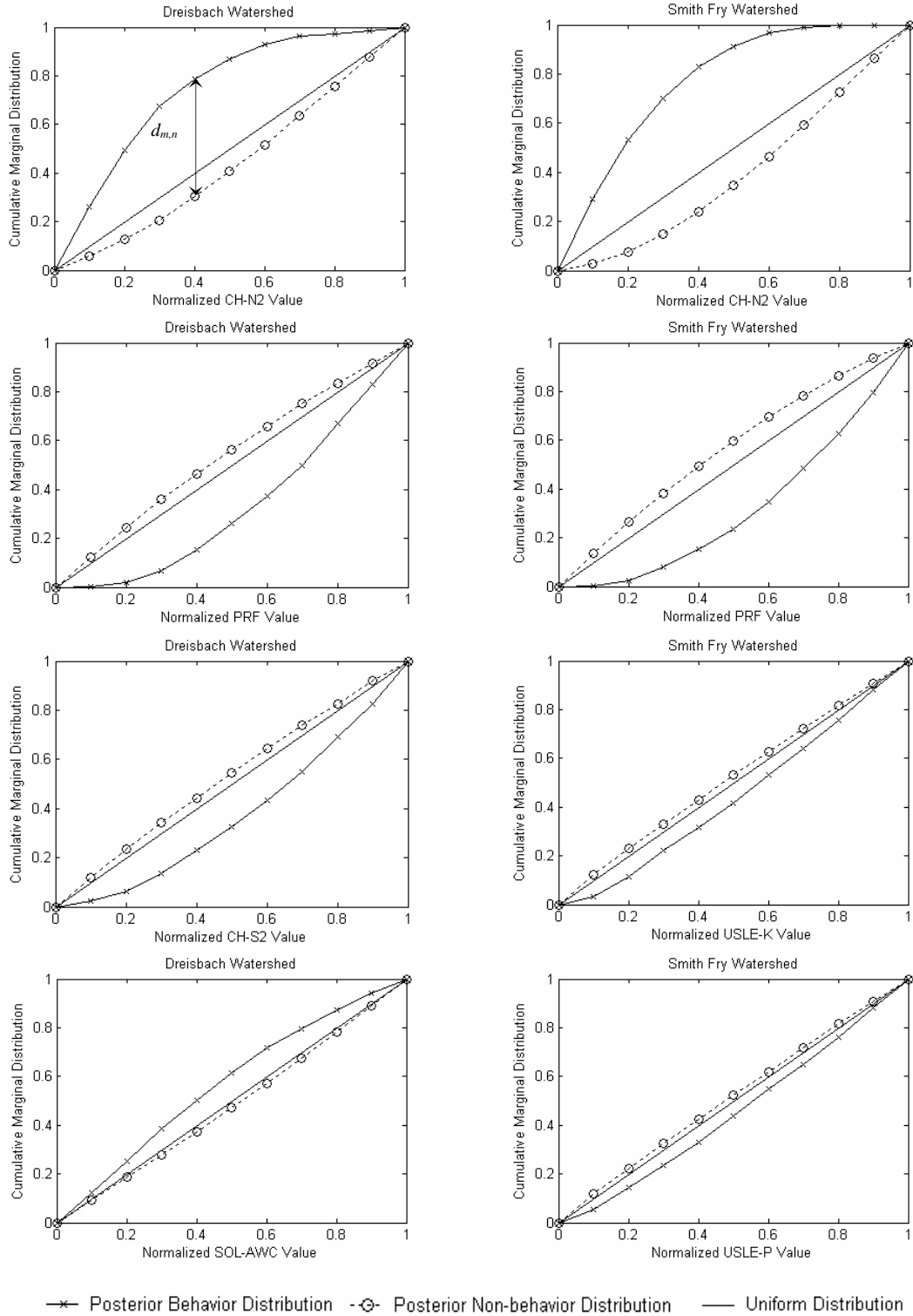


Figure 4.2 Posterior distributions of behavior and non-behavior sets of input factors along with their prior distribution (uniform distribution) for sediment at the outlet of Dreisbach and Smith Fry watersheds.

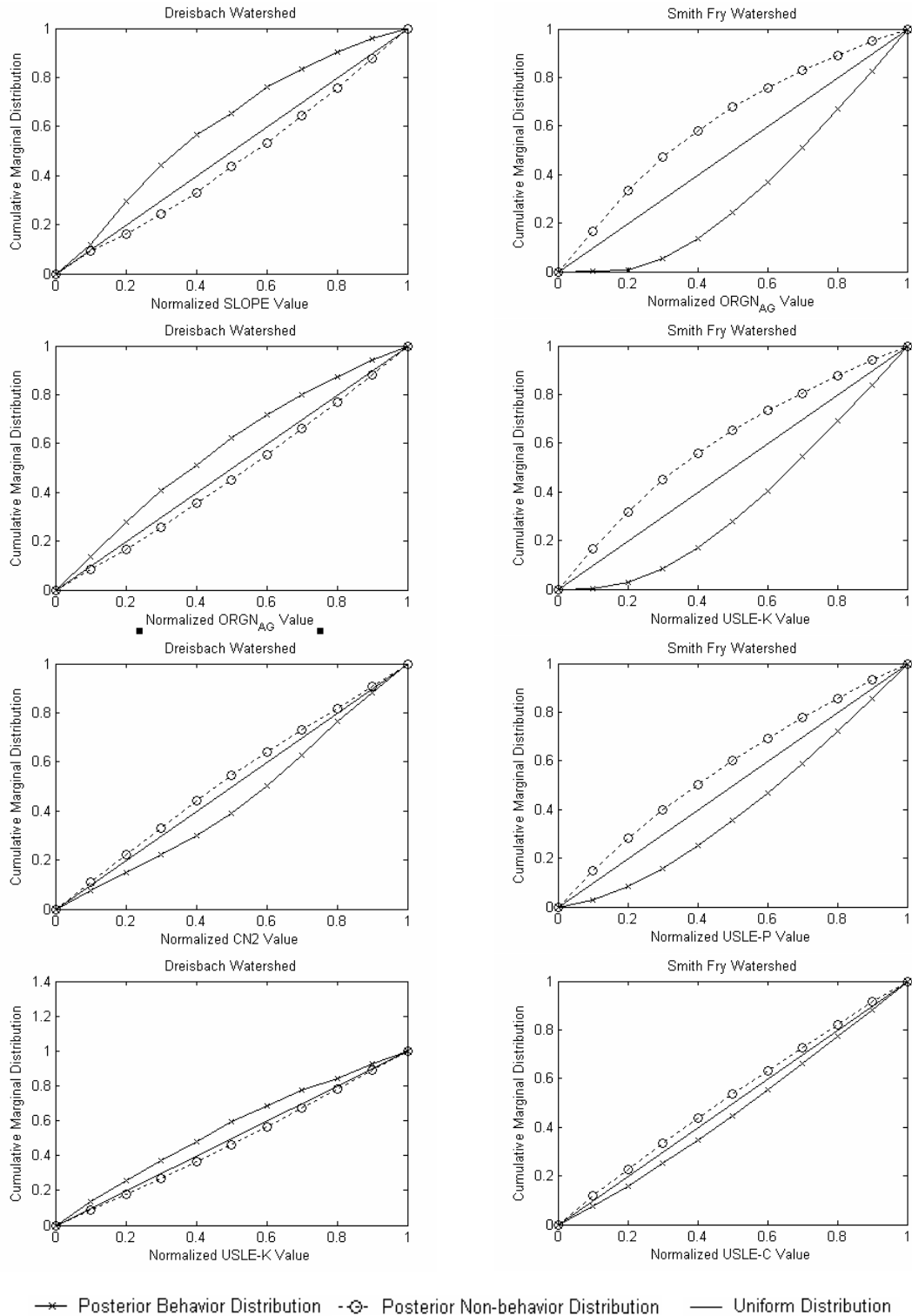


Figure 4.3 Posterior distributions of behavior and non-behavior sets of input factors along with their prior distribution (uniform distribution) for total N at the outlet of Dreisbach and Smith Fry watersheds.

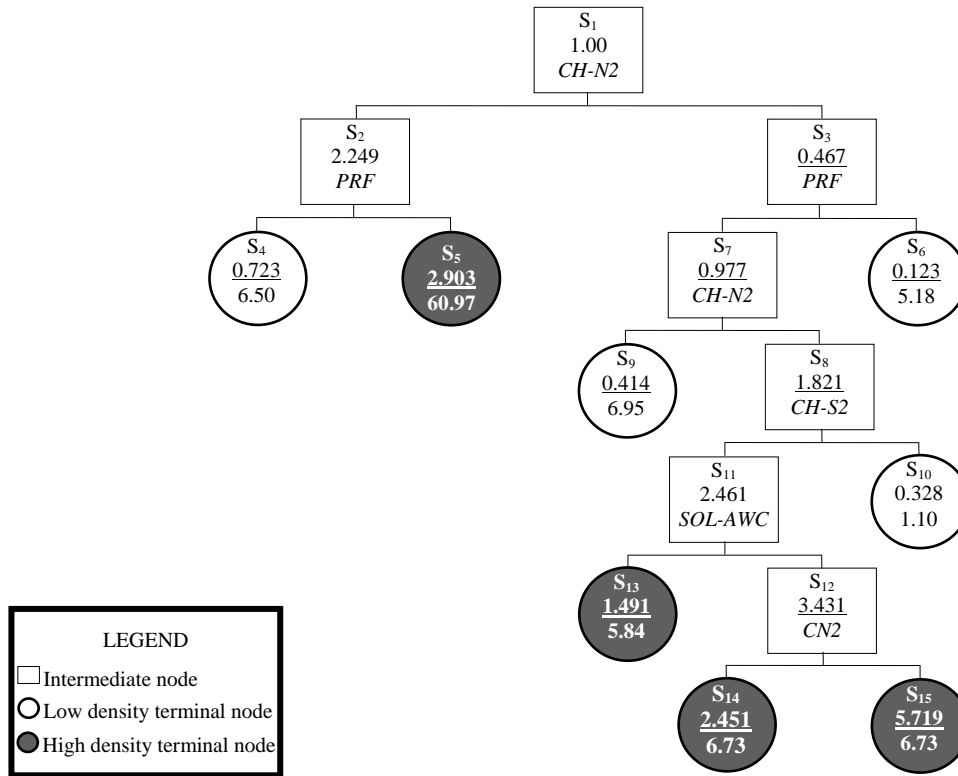


Figure 4.4 TSDE diagram for sediment at the outlet of Dreisbach watershed: 1st line-terminal number, 2nd line-relative density of the node, 3rd line-input factor that splits the intermediate nodes, or percentage volume of terminal nodes.

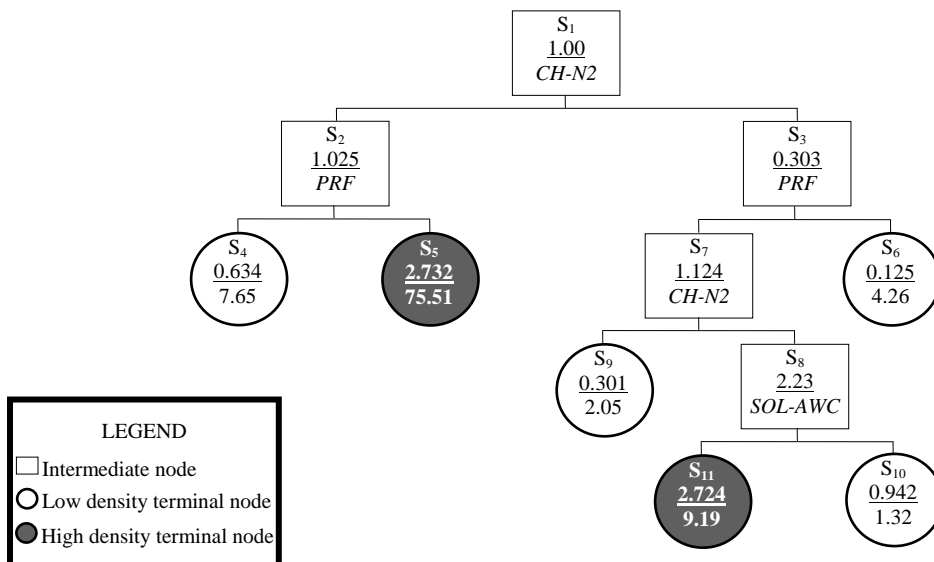


Figure 4.5 TSDE diagram for sediment at the outlet of Smith Fry watershed, description of values in each line as in Figure 4.4.

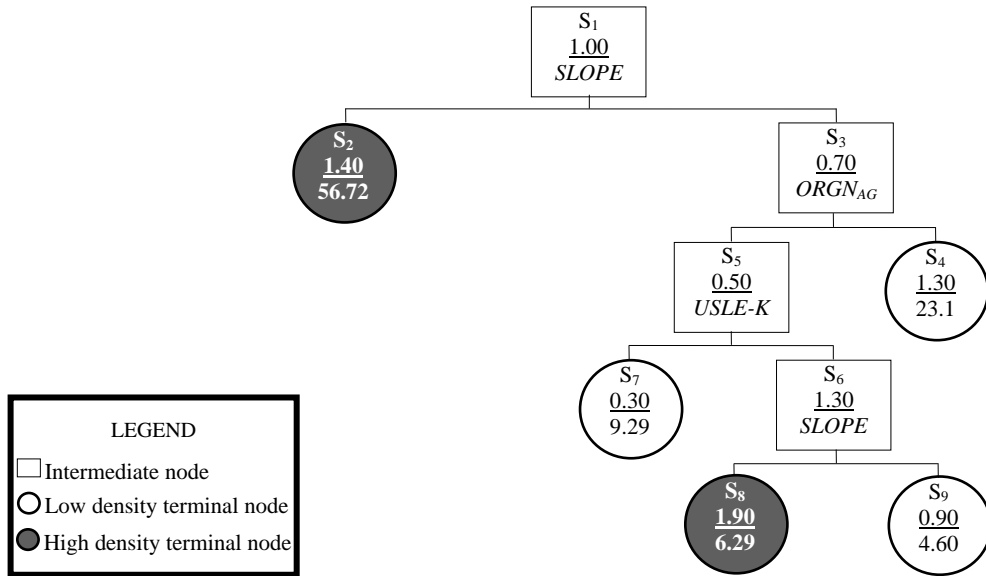


Figure 4.6 TSDE diagram for total N at the outlet of Dreisbach watershed, description of values in each line as in Figure 4.4.

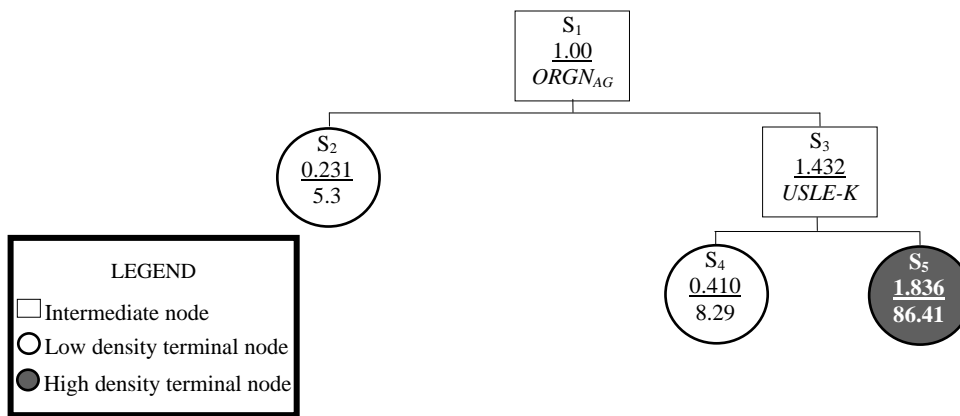


Figure 4.7 TSDE diagram for total N at the outlet of Smith Fry based on observed-past behavior definition, description of values in each line as in Figure 4.4.

Table 4.1 Regionalized Sensitivity Analysis ranking of input factors for sediment based on observed-past behavior definition at the outlet of Dreisbach watershed

| Symbol | Short Description | Rank |
|---------------------|---|---|
| | | $d_{m,n}$ |
| <i>CH-N2</i> | Channel Manning's number | 0.4822 |
| <i>PRF</i> | Peak rate adjustment factor | 0.3090 |
| <i>CH-S2</i> | Average main channel slope | 0.2189 |
| <i>SOL-AWC</i> | Available water capacity of soil layer | 0.1457 |
| <i>CN2</i> | SCS curve number | 0.1346 |
| <i>ESCO</i> | Soil evaporation compensation factor | 0.1056 |
| <i>SPEXP</i> | Exponent coefficient for sediment routing | 0.0730 |
| <i>ALPHA-BF</i> | Baseflow alpha factor | 0.0532 |
| Summary Statistics: | | |
| | | Number of Behavior Simulations (m) : 1010 |
| | | Total number of Simulation ($m+n$) : 5000 |
| | | Success rate : 20.2% |

Table 4.2 Regionalized Sensitivity Analysis ranking of input factors for sediment based on observed-past behavior definition at the outlet of Smith Fry watershed

| Symbol | Short Description | Rank |
|---------------------|------------------------------|---|
| | | $d_{m,n}$ |
| <i>CH-N2</i> | Channel Manning's number | 0.5929 |
| <i>PRF</i> | Peak rate adjustment factor | 0.3616 |
| <i>CH-S2</i> | Average main channel slope | 0.2010 |
| <i>USLE-K</i> | USLE soil erodibility factor | 0.1172 |
| <i>USLE-P</i> | USLE support practice factor | 0.0970 |
| Summary Statistics: | | |
| | | Number of Behavior Simulations (m) : 1360 |
| | | Total number of Simulation ($m+n$) : 5000 |
| | | Success rate : 27.2% |

Table 4.3 Regionalized Sensitivity Analysis ranking of input factors for total N based on observed-past behavior definition at the outlet of Dreisbach watershed

| Symbol | Short Description | Rank |
|---------------------|--|---|
| | | $d_{m,n}$ |
| <i>SLOPE</i> | Average slope steepness | 0.2353 |
| <i>ORGNAG</i> | Initial organic N in soil, agricultural land use | 0.1755 |
| <i>CN2</i> | SCS curve number | 0.1539 |
| <i>USLE-K</i> | USLE soil erodibility factor | 0.1345 |
| <i>SOL-AWC</i> | Available water capacity of soil layer | 0.1266 |
| <i>USLE-P</i> | USLE support practice factor | 0.1096 |
| <i>ORGNPAST</i> | Initial organic N in soil, pasture land use | 0.0877 |
| <i>SLSUBBSN</i> | Average slope length | 0.0593 |
| <i>USLE-C</i> | Maximum value of USLE cover factor | 0.0550 |
| Summary Statistics: | | |
| | | Number of Behavior Simulations (m) : 1448 |
| | | Total number of Simulation ($m+n$) : 5000 |
| | | Success rate : 29.0% |

Table 4.4 Regionalized Sensitivity Analysis ranking of input factors for total N based on observed-past behavior definition at the outlet of Smith Fry watershed

| Symbol | Short Description | Rank |
|---------------------|--|---|
| | | $d_{m,n}$ |
| <i>ORGN_AG</i> | Initial organic N in soil, agricultural land use | 0.4436 |
| <i>USLE-K</i> | USLE soil erodibility factor | 0.3848 |
| <i>USLE-P</i> | USLE support practice factor | 0.2514 |
| <i>USLE-C</i> | Maximum value of USLE cover factor | 0.0908 |
| <i>CH-K2</i> | Channel effective hydraulic conductivity | 0.0518 |
| Summary Statistics: | | |
| | | Number of Behavior Simulations (m) : 2039 |
| | | Total number of Simulation ($m+n$) : 5000 |
| | | Success rate : 40.8% |

Table 4.5 Summary of TSDE terminal nodes for sediment related parameters based on observed-past behavior definition at the outlet of Dreisbach watershed

| Terminal Node | Relative Density | Tree level (root node = level 1) | | | | | |
|---------------|------------------|----------------------------------|------------|--------------|--------------|----------------|------------|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| S_{15} | 5.719 | CH-N2 | <i>PRF</i> | <i>CH-N2</i> | <i>CH-S2</i> | <i>SOL-AWC</i> | <i>CN2</i> |
| S_5 | 2.903 | CH-N2 | <i>PRF</i> | | | | |
| S_{14} | 2.451 | CH-N2 | <i>PRF</i> | <i>CH-N2</i> | <i>CH-S2</i> | <i>SOL-AWC</i> | <i>CN2</i> |
| S_{13} | 1.491 | CH-N2 | <i>PRF</i> | <i>CH-N2</i> | <i>CH-S2</i> | <i>SOL-AWC</i> | |
| S_4 | 0.723 | CH-N2 | <i>PRF</i> | | | | |
| S_9 | 0.414 | CH-N2 | <i>PRF</i> | <i>CH-N2</i> | | | |
| S_{10} | 0.328 | CH-N2 | <i>PRF</i> | <i>CH-N2</i> | <i>CH-S2</i> | | |
| S_6 | 5.18 | CH-N2 | <i>PRF</i> | | | | |

High Density Terminal Nodes (HDTNs): { S_{15} ; S_5 ; S_{14} ; S_{13} }

Table 4.6 Summary of TSDE terminal nodes for sediment related parameters based on observed-past behavior definition at the outlet of Smith Fry watershed

| Terminal Node | Relative Density | Tree Level (root node = level 1) | | | |
|---------------|------------------|----------------------------------|------------|--------------|----------------|
| | | 1 | 2 | 3 | 4 |
| S_5 | 1.90 | <i>CH-N2</i> | <i>PRF</i> | | |
| S_{11} | 1.40 | <i>CH-N2</i> | <i>PRF</i> | <i>CH-S2</i> | <i>SOL-AWC</i> |
| S_{10} | 0.942 | <i>CH-N2</i> | <i>PRF</i> | <i>CH-S2</i> | <i>SOL-AWC</i> |
| S_4 | 0.64 | <i>CH-N2</i> | <i>PRF</i> | | |
| S_9 | 0.301 | <i>CH-N2</i> | <i>PRF</i> | <i>CH-S2</i> | |
| S_6 | 0.125 | <i>CH-N2</i> | <i>PRF</i> | | |

High Density Terminal Nodes (HDTNs): { S_5 ; S_{11} }

Table 4.7 Summary of TSDE terminal nodes for total N related parameters based on observed-past behavior definition at the outlet of Dreisbach watershed

| Terminal Node | Relative Density | Tree level (root node = level 1) | | | |
|---|------------------|----------------------------------|--------------------------|---------------|--------------|
| | | 1 | 2 | 3 | 4 |
| S_8 | 1.90 | <i>SLOPE</i> | <i>ORGN_{AG}</i> | <i>USLE-K</i> | <i>SLOPE</i> |
| S_2 | 1.40 | <i>SLOPE</i> | | | |
| S_4 | 1.30 | <i>SLOPE</i> | <i>ORGN_{AG}</i> | | |
| S_9 | 0.90 | <i>SLOPE</i> | <i>ORGN_{AG}</i> | <i>USLE-K</i> | <i>SLOPE</i> |
| S_7 | 0.3 | <i>SLOPE</i> | <i>ORGN_{AG}</i> | <i>USLE-K</i> | |
| High Density Terminal Nodes (HDTNs): { S_8, S_2 } | | | | | |

Table 4.8 Summary of TSDE terminal nodes for total N related parameters based on observed-past behavior definition at the outlet of Smith Fry watershed

| Terminal Node | Relative Density | Tree level (root node = level 1) | |
|--|------------------|----------------------------------|---------------|
| | | 1 | 2 |
| S_5 | 1.836 | <i>ORGN_{AG}</i> | <i>USLE-K</i> |
| S_4 | 0.410 | <i>ORGN_{AG}</i> | <i>USLE-K</i> |
| S_2 | 0.231 | <i>ORGN_{AG}</i> | |
| High Density Terminal Nodes (HDTNs): { S_5 } | | | |

SECTION 5. COST-EFFECTIVE ALLOCATION OF WATERSHED MANAGEMENT PRACTICES USING A GENETIC ALGORITHM

5.1. Abstract

Implementation of conservation programs are perceived as being crucial for restoring and protecting waters and watersheds from nonpoint source pollution. Success of these programs depends to a great extent on planning tools that can assist the watershed management process. Herein, a novel optimization methodology is presented for deriving watershed-scale sediment and nutrient control plans that incorporate multiple, and often conflicting, objectives. The method combines the use of a watershed model (SWAT), representation of best management practices, an economic component, and a genetic algorithm-based spatial search procedure. For two small watersheds in Indiana located in the Midwestern portion of the United States, selection and placement of best management practices by optimization was found to be nearly three times more cost-effective than targeting strategies for the same level of protection specified in terms of maximum monthly sediment, phosphorus, and nitrogen loads. Conversely, for the same cost, the optimization plan reduced the maximum monthly loads by a factor of two when compared to the targeting plan. The optimization methodology developed in this study can facilitate attaining water quality goals at significantly lower costs than commonly used cost-share and targeting strategies.

5.2. Introduction

Best management practices (BMPs) are widely accepted as effective control measures for agricultural nonpoint sources of sediments and nutrients. The 2002 Farm Bill provided up to \$13 billion for conservation programs aimed at protecting water quality from agricultural nonpoint source (NPS) pollution (USDA, 2003). In addition, under the Clean Water Act Section 319 Nonpoint Source National Monitoring Program and wetland protection programs, the EPA supports programs to reduce the negative impacts of runoff from agricultural, urban, and industrialized areas. Similarly, the Natural Resources Conservation

Service (NRCS) provides hundreds of millions of dollars in federal funds to support agricultural best management practices (BMPs) in an effort to reduce the movement of pollutants into our waterways. Success of such programs, however, is contingent upon availability of efficient watershed-scale planning tools.

Implementation of BMPs is challenged by complexities in incorporation of conflicting environmental, economic, and institutional criteria. Environmental assessments in watersheds hinge on resolving social benefits such as achieving the goal of swimable and fishable water bodies under the EPA's Total Maximum Daily Load (TMDL) agenda. While BMPs facilitate achievement of such targets, their establishment bears additional cost for watershed management and/or agricultural producers. Since management practices are usually implemented under a limited budget, costs associated with unnecessary/redundant management actions may jeopardize attainability of designated water quality goals. Identifying optimal combinations of watershed management practices requires systematic approaches that allow decision makers to quickly assess trade-offs among environmental and economic criteria.

Cost-sharing with landowners is promoted by government agencies for BMP implementation in agricultural fields (EPA, 2003). BMPs are implemented through site investigation, monitoring, and field-scale modeling. While cost-share programs may improve water quality standards at a field scale, their impact at a watershed scale typically remains unknown for two main reasons. First, impact of BMPs may duplicate or overlap each other, thereby reducing the potential benefit for the watershed. Interactions between BMPs may also significantly affect their individual performances at a watershed scale. Second, water quality impacts of BMPs are site-specific, greatly influenced by landscape characteristics such as land use, soils, and management actions. Thus, the same BMP is likely to have varying efficacies at different locations within a watershed. The direction of many regulatory programs such as the EPA's Clean Water Act (CWA) has now shifted from a source-by-source, pollutant-by-pollutant approach to a more holistic watershed-scale strategy (EPA, 2003).

An example of watershed scale approaches is targeting where pollution control measures are allocated to critical source areas. Critical sources are portions of the watershed that are believed to contribute intensively to nonpoint source pollution. Previous studies in the same watershed, watershed data (i.e. land use, soils, and management characteristics), physical characteristics such as topography and proximity to a stream, and instream monitoring data are typically analyzed to identify critical sources. This approach has been endorsed in the recent draft of EPA's "Handbook for developing watershed plans to restore and protect our waters" (EPA, 2005a). However, identification of an optimal pollution control strategy through targeting becomes infeasible in large complex watersheds, because the number of possible scenarios increases exponentially with the number of fields. The search for an optimal watershed pollution control plan needs to be conducted in a more efficient manner.

Development of increasingly powerful computers has facilitated application of mathematical programming heuristics for solving complex and computationally cumbersome problems. A few previous studies have used evolutionary algorithms (EA) to optimize placement of watershed management actions. Srivastava et al. (2002) linked a genetic algorithm (GA) with a continuous NPS model (AnnAGNPS) to optimize selection of crop rotation practices in a 7.25 km² USDA experimental watershed in Pennsylvania. The optimized crop management scheme obtained after 150 GA generations decreased pollution load by nearly 56% and increased net annual return by nearly 110% from a random selection of crop rotation scenarios. The objective function of the optimization procedure was formulated such that either pollution reduction or net return, but not both at the same time, was maximized. Veith et al. (2004) developed a GA-based optimization model for cost-effective allocation of land use and tillage practices to upland fields that minimized cost (economic criteria) while sediment load at the outlet of watershed was constrained to a predefined value. The model was tested in a 10.14 km² case study watershed in Virginia. Results of the study showed that the optimization plan achieved the same sediment reduction as a targeting plan at a lower cost. The major setback with optimizing only one of the decision objectives (economic, environmental, and/or institutional) is that the procedure can not demonstrate the tradeoff between multiple objectives.

Incorporation of multiple objectives in the search for alternative agricultural landscapes was studied by Bekele and Nicklow (2005) and Muleta and Nicklow (2005). A multiobjective evolutionary algorithm (SPEA; Strength Pareto Evolutionary Algorithm, Zitzler and Thiele., 1999) was linked with a watershed model for selection of crop rotation and tillage operations in a 133 km² watershed in southern Illinois. Multiple objectives included minimizing pollutant loads, and maximizing net profit. Though both studies showed the effectiveness of an integrated approach in providing tradeoff between multiple objectives, explicit incorporation of water quality and budget constraints was not examined. These two studies do not offer any clear guidance for selection of operating parameters and termination criteria as they rely on the previously developed SPEA. Moreover, neither optimization results were compared with the more commonly used targeting strategies, nor was convergence of pollutant loads/net profit tested. Hence, appraisal of efficiency of the optimization method in deriving solutions better than targeting strategies was curtailed.

Research to date indicates the promise of heuristic optimization for cost-effective allocation of watershed management practices. Unlike gradient-based approaches, heuristic techniques do not require linearity, continuity, or differentiability either for objective/constraint functions or for input parameters. Thus, they are well-suited for cost-effective allocation of watershed management plans. However, several questions still defy answers. A decision making tool that can clearly accommodate economic, environmental and institutional criteria is still lacking. The means for imposing target values for pollutant loads, and total watershed cost of implementation of management plans needs to be explored. As discussed by Srivastava et al. (2002), a more versatile formulation is needed to maximize pollution reduction and minimize cost at the same time. Moreover, previous studies have only focused on alternative agricultural landscapes. Research on economic evaluation of structural BMPs that are installed both in upland areas and within the channel network is lacking. In such cases, the interaction between the BMPs could be critical to reach water quality goals. Finally, selection of GA's operating parameters, and termination criteria need further study.

The main goal of this section is to develop an optimization framework that enhances decision makers' capacity to evaluate a range of agricultural and environmental management alternatives. The tool will be designed to identify near optimal watershed plans that reduce

pollutant loads at a watershed outlet to below regulatory or target values with minimum cost. We hypothesize that reductions of pollutants at watershed outlets can be attained at significantly lower cost by optimized implementation of conservation practices than by current cost-share and targeting approaches. This overall goal is achieved by the following specific steps:

1. Development of a novel genetic algorithm-based spatial search model. This step will focus on formulating versatile objective and constraint functions for the optimization model that can handle multi-criteria and landscape characteristics.
2. Integration of an NPS model (SWAT; Soil and Water Assessment Tool), a new BMP representation method, and a cost-benefit economic relationship with the GA-based spatial optimization model to identify optimal spatial allocation of best management practices;
3. Provision of guidance for selection of GA operating parameters and termination criteria;
4. Conducting case studies to determine how the optimization plan compares to plans from cost-share and targeting strategies.

5.3. Theoretical Considerations

5.3.1. Genetic Algorithm (GA)

Genetic algorithms belong to the evolutionary class of artificial intelligence (AI) techniques. GAs are based on natural selection of chromosomes from a population for mating, reproduction of offspring by crossover, and mutation to ascertain diversity- ideas borrowed from biology. Each chromosome string in the population corresponds to a solution for the problem at hand, with each variable being represented by a gene (a specific position in the string). The values of the genes, known as allele, can be binary, real-valued, or character-valued.

The GA begins with an initial population containing randomly generated chromosomes or strings, each representing a possible solution to the problem. The next population is generated by mating the fittest solutions (i.e. chromosomes) in the previous population.

Based on the principle of survival of the fittest, the higher the solution's fitness, the more likely it will be chosen for reproduction. The search for the optimal solution continues until a predefined termination criterion is reached.

The GA is essentially a search algorithm for solving difficult nonlinear optimization problems and is not intended to examine all possible solutions. Regardless of how long the algorithm searches, its convergence to an optimum can not be guaranteed. However, it has been shown that the GA converges to near optimal solutions for a variety of problems (Winston and Venkataramanan, 2003). Efficiency of the algorithm depends on the optimization's operating parameters and the convergence criterion. The higher the number of individual evaluations for converging to the optimum, the less efficient is the procedure. The values of the operating parameters are problem dependent, and can be determined by performing a sensitivity analysis.

The GA is a well-suited optimization technique for spatial allocation of BMPs, because unlike gradient-based methods it does not require linearity, continuity, or differentiability of either the objective function or the constraint function.

5.4. Methodology

The optimization model developed in this study is comprised of the SWAT model for simulating pollutant loads, a BMP representation tool, an economic component, and a GA-based spatial optimization technique. A MATLAB (The Mathworks, Natick, MA) computer program was developed to provide the linkage among various components of the model as shown in Figure 5.1. The model was tested for optimization of the location of field borders, parallel terraces, grassed waterways, and grade stabilization structures in the Dreisbach and Smith Fry watersheds.

SWAT simulations were performed for a 10-year period from January 1st, 2000 through December 31st, 2009. In the analysis, 1991-2000 precipitation data, 2000 USDA-National Agriculture Statistics Service (NASS) land use, and 2002 Soil Survey Geographical Database (SSURGO) were utilized to establish a base-case SWAT run. Parameter values in the base-case run were selected from a manual calibration (Arabi et al., 2005). Portions of the

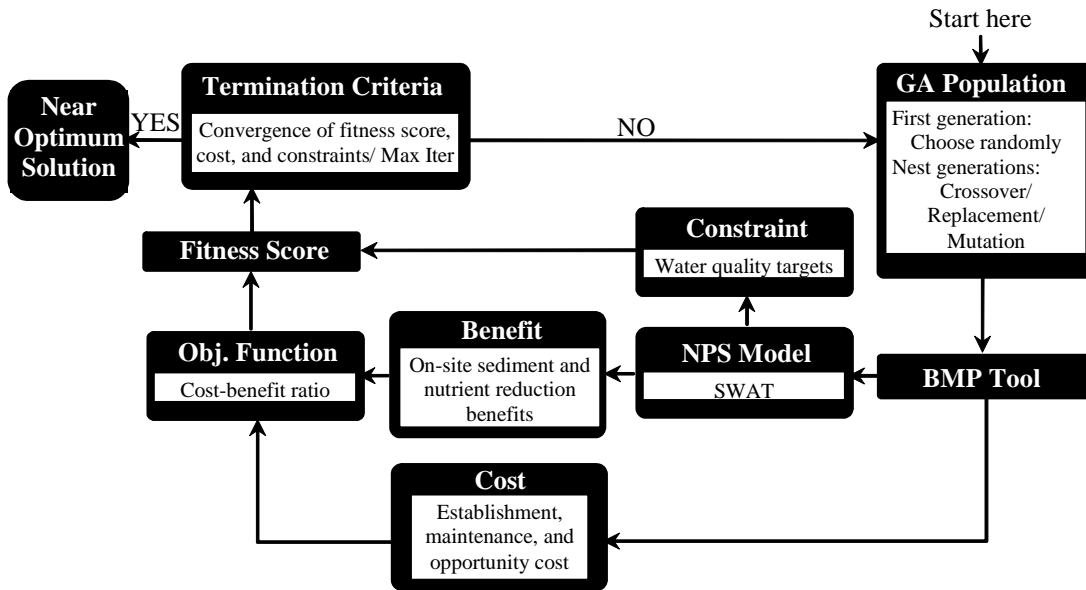


Figure 5.1 Schematic of the optimization procedure.

watershed classified as urban and forested areas were not considered for implementation of BMPs.

5.4.1. NPS Model

The optimization tool developed here does not necessarily require using the SWAT model as the NPS component. The SWAT model was chosen because the case studies will deal with sediments, and nutrients. Soundness of mathematical representation of sediment, nutrient, and pesticide processes in SWAT has been validated in previous research (Santhi et al., 2001; Kirsch et al., 2002; Arabi, 2005; Vazquez-Amabile, 2005). Moreover, SWAT has been linked with optimization routines for automated calibration of the model (van Griensven and Bauwens, 2003) and for evaluation of efficiency of nonpoint source pollution regulatory programs (Whittaker et al., 2003). The optimization framework herein can be easily integrated with other hydrologic/water quality models to develop management plans for other types of pollutants, and for other watersheds as well.

5.4.2. BMP Representation

In this study, a method presented by Arabi et al. (2004) and Bracmort et al. (2006) was utilized to evaluate the water quality impacts of grassed waterways, grade stabilization structures, field borders and parallel terraces. The method was developed based on published literature pertaining to BMP simulation in hydrological models and considering the hydrologic and water quality processes simulated in SWAT. Based on the function of the BMPs and hydrologic and water quality processes that are modified by their implementation, corresponding SWAT parameters were selected and altered. Table 2.3 summarizes the SWAT parameters and their corresponding values for representation of the BMPs. Arabi et al. (2004) provides a detailed description of the method used for representation of field borders, parallel terraces, grassed waterways, and grade stabilization structures.

5.4.3. Economic Component

An economic component was developed for the optimization model that is comprised of a cost function in addition to an economic return (benefit) function. Both cost and benefit are a function of watershed characteristics and time. The benefit function reflects the impact of BMPs on sediment and nutrient reductions.

Cost Function

The total cost of implementation of BMPs was evaluated by establishment, maintenance, and opportunity costs. Establishment costs included the cost of BMP installation, and technical and field assistance. Maintenance cost is usually evaluated as a percentage of establishment cost. The opportunity cost is a dollar value that would be produced over the BMP design life as a result of investing the establishment and maintenance costs by purchasing saving bonds. For each individual BMP, the total cost (ct) was evaluated by the following equation:

$$ct(x, t, td) = c_0(1 + s)^{td} + c_0 \cdot rm \left[\sum_{\tau=1}^{td} (1 + s)^{(\tau-1)} \right] \quad \text{Eq 5.1}$$

$$c_0 = c_0(x, t) \quad \text{Eq 5.2}$$

where c_0 is the establishment cost that is a function of state of the watershed (i.e. x , landscape and physical characteristics) at time t ; s is the interest rate; td is BMP design life; and rm is the ratio of maintenance cost to establishment cost. Interest rate (s) of 6.5% was used in computations. The design life of a BMP is defined by Natural Resources Conservation Service (NRCS) as “the intended period of time that the practice will function successfully with only routine maintenance determined during design phase” (USDA-NRCS, 2005). Design life, establishment cost, and maintenance rate for BMPs in this study are summarized in Table 5.1. The table also provides information with regard to the number and area (i.e. quantity of each type of BMP) allocated within the study watersheds through targeting (see Figure 2.1). For each management plan, establishment cost (c_0) of each type of BMP was obtained by multiplying its unit establishment cost (c) by its quantity ($abmp$). The total watershed cost of BMPs was computed by summing up individual costs from Eq 5.1.

Table 5.1 Unit cost of establishment (c), maintenance rate (rm), and design life (td) of BMPs in the study along with number ($nbmp$) and quantity ($abmp$) of each type in the study watersheds as shown in Figure 2.1 (i.e. targeting strategy).

| BMP | c^a (\$) | rm^a (%) | td (yr) | Dreisbach | | Smith Fry | |
|---------------------|-------------------|---------------|--------------|-----------|------------------|-----------|------------------|
| | | | | $nbmp$ | $abmp$ (unit) | $nbmp$ | $abmp$ (unit) |
| Field Border | 4.6 /m | 1 | 10 | 7 | 2600 (m) | 1 | 1800 (m) |
| Parallel Terrace | 26 /m | 3 | 10 | 4 | 2130 (m) | 2 | 480 (m) |
| Grassed Waterway | 6200 /ha | 3 | 10 | 5 | 3.50 (ha) | 1 | 0.95 (ha) |
| Grade Stabilization | 10,000 /structure | 2 | 15 | 10 | - | 2 | - |

^a Adapted from Indiana Environmental Quality Incentives Program (2004).

Benefit Function

The economic return of implementation of BMPs was determined by assigning monetary values to onsite and offsite benefits of sediment and nutrient reductions. Reduction of sediments and nutrients as a result of implementation of a management practice is a function of landscape characteristics. Likewise, benefits gained by BMP implementation vary with landscape properties of the site where the BMP is installed. While offsite benefit of BMPs were determined based on cost of offsite damage that would be caused by sediment and nutrient loads, onsite benefits were estimated based on the quantity of reduced nutrient loads from fields under the influence of BMPs.

Ribaudo et al. (1989) equated the benefit of reducing soil loss to offsite damage in \$ per ton of eroded soil based on freshwater recreation, marine recreation, water storage, navigation flooding, roadside ditches, irrigation ditches, freshwater commercial fishing, marine commercial fishing, municipal water treatment, and municipal and industrial use. A \$1.15/t damage cost estimate was suggested for the Corn-Belt region in the United States. Additionally, Cangelosi et al. (2001) evaluated the benefit of reducing soil erosion for dredging in the Maumee River basin to be \$0.87 for a ton of eroded soil. In a study by Fang and Easter (2003), a social cost of \$2.65/kg of removed phosphorous was used to quantify the cost-effectiveness of a water quality trade in southeastern Minnesota. Onsite water quality benefit of BMP implementation was evaluated by the monetary value of nutrient reduction from upland areas. Buckner (2001) estimated the benefit of reducing phosphorus from agricultural lands by the bulk rate cost of triple super phosphate (00-46-00) that, in 2003, was \$0.26/kg (Heartland Co-op, IN, personal communication). Table 5.2 summarizes the monetary benefits (*bm*) of unit reduction of sediment, phosphorus, and nitrogen deliveries expressed in 2000 dollars used for the Dreisbach and Smith Fry watersheds. The values used in this study are site specific and may vary for other regions.

After on-site and off-site benefits of BMPs were quantified by monetary values from related literature, the total benefit (*bt*) was calculated for individual BMPs over their design life as:

$$bt(x, t, td, \alpha) = \sum_{\tau=1}^{td} \sum_i [(\Delta y_{i,\tau} bm_i) (1+s)^{(\tau-1)}] \quad \text{Eq 5.3}$$

$$\Delta y = y(x, t, t_d, \alpha) \quad \text{Eq 5.4}$$

where $\Delta y_{i,\tau}$ is annual reduction of constituent *i* (i.e. sediment in t/yr, phosphorus in kg/yr, and nitrogen in kg/yr) for year τ at the outlet of the study watershed, *td* (yr) is BMP design life, and *bm* is the monetary benefit of BMP as a result of reducing constituent *i* from Table 5.2. In 5.3, *x* is the state of the watershed at any given time *t*, α denotes the watershed management plan, and function *y* reflects mathematical relationships in the SWAT model that are used for representation of hydrologic and water quality processes.

Table 5.2 Monetary benefits of unit reduction of sediment and nutrient loads in 2000 dollars

| Variable | Benefit (<i>bm</i>) (\$) |
|------------|-------------------------------|
| Sediment | 2.5 /t |
| Phosphorus | 2.86 /kg |
| Nitrogen | 3.5 /kg |

It is worthwhile noting that bm_i serves as a weighting factor that incorporates the relative importance of reduction of pollutant constituent i in development of watershed management plans. A higher value indicates that larger benefits can be gained from unit reduction of the corresponding constituent. In the absence of regional and site-specific data for the sort of analysis that was used in this study, the term bm can be assigned by judgment of the analyst. If all constituents bear the same importance, bm should be 1 for all constituents of concern.

5.4.4. Optimization Component

A genetic algorithm (GA) was employed to optimize the spatial allocation of BMPs. In this GA component, each optimization string (vector α_j^i) corresponds to a specific watershed management plan. Figure 5.2 is a demonstration of an optimization string for placement of BMPs. The length of each string (m) corresponds to the total number of genes, i.e., individual management actions ($\alpha_{j,l}^i$) that are considered in the optimization. For example, in a watershed with 50 fields (i.e., $nhru=50$) considered for implementation of field borders and/or parallel terraces, and 20 reach segments (i.e., $nch=20$) considered for implementation of grassed waterways and/or grade stabilization structures, the total number of genes on each management string is equal to $m= [2 \times 50 + 2 \times 20 =]$ 140. The alleles are binary values, with “1” or “zero” indicating that the corresponding BMP “be” or “not be” implemented. The i^{th} optimization generation (P^i) with $nstr$ solution vectors is expressed as:

$$P^i = \begin{bmatrix} \alpha_1^i \\ \alpha_2^i \\ \vdots \\ \alpha_{nstr}^i \end{bmatrix} = \begin{bmatrix} \alpha_{1,1}^i & \cdots & \alpha_{1,m}^i \\ \alpha_{2,1}^i & \cdots & \alpha_{2,m}^i \\ \vdots & \ddots & \vdots \\ \alpha_{nstr,1}^i & \cdots & \alpha_{nstr,m}^i \end{bmatrix} \quad \text{Eq 5.5}$$

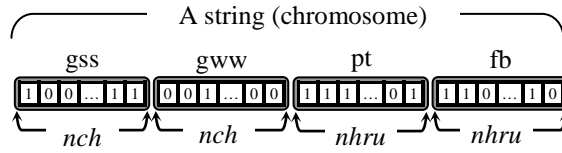


Figure 5.2 Schematic of an optimization string (chromosome) representing grade stabilization structures (gss), grassed waterways (gww), parallel terraces (pt), and field borders (fb). nch and nhru refer to the number of channel segments and fields (HRUs) in the watershed, respectively.

where

$$\forall j \in \{1, 2, \dots, nstr\}, \forall l \in \{1, 2, \dots, m\}: \alpha_{j,l}^i = 0/1 \quad \text{Eq 5.6}$$

where α_j^i denotes the j^{th} management vector in P^i , and $\alpha_{j,l}^i$ reflects the l^{th} gene (an individual management action) in α_j^i . For a total of $ngen$ optimization generations, the entire solution space (A) is obtained from the union of solutions in optimization generations:

$$A = \bigcup_{i=1}^{ngen} P^i \quad \text{Eq 5.7}$$

Mathematical Formulation

A general multiobjective optimization problem can be formally expressed as (Zitzler and Thiele, 1999):

$$\text{Minimize/Maximize } z(\alpha) = f(\alpha) = (f_1(\alpha), f_2(\alpha), \dots, f_n(\alpha)) \quad \text{Eq 5.8}$$

subject to

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m) \in \Omega \quad \text{Eq 5.9}$$

$$z = (z_1, z_2, \dots, z_n) \in Z \quad \text{Eq 5.10}$$

where z denotes a vector function f that maps a set of m decision variables α to a set of n objectives; Ω and Z are, respectively, the decision parameter space and objective space (refer to Ringuest, 1992 for more information).

Several common methods have been previously used for handling multiobjective optimization problems (Deb, 2001). The most commonly used approach is the weighted sum approach. In this method, a set of objectives are aggregated into a single objective by multiplying each objective by a user-defined weight (w). These weights reflect the importance of each objective in the context of the optimization problem:

$$z = f(\alpha) = \sum_{i=1}^n w_i f_i(\alpha) \quad \text{Eq 5.11}$$

The BMP spatial allocation problem was formulated with an aggregate single-objective based on multiple criteria. The multiple criteria included minimizing implementation cost of BMPs and maximizing pollutant load reductions. Monetary benefits of unit reduction of sediment, phosphorus, and nitrogen loads (Table 5.2) were used to determine the weighted sum of onsite and offsite benefits gained by reducing pollutant loads (bt). Then, the objective function was evaluated with a benefit (bt) to cost (ct) ratio at the watershed level. Pollutant loads were constrained to ascertain that sediment and nutrient yields would not exceed allowable values and/or that total available budget is not exceeded by entailed costs. Allowable pollutant loads are typically specified by regulatory agencies. Since the available budget for implementation of BMPs is usually limited, an additional constraint was imposed to search for solutions that can be implemented with a cost less than the specified budget. The mathematical representation of the objective function used in this study was:

$$\text{Maximize } z = \frac{bt}{\max(ct, 1)} = \frac{\sum_{\tau=1}^{td} \sum_i [(\Delta y_{i,\tau} b m_i) (1+s)^{(\tau-1)}]}{\max\left(c_0 (1+s)^{td} + c_0 \cdot rm \left[\sum_{\tau=1}^{td} (1+s)^{(\tau-1)} \right], 1\right)} \quad \text{Eq 5.12}$$

subject to water quality constraints:

$$\Gamma_1(x,t,td,\alpha) = y_{max_i}(x,t,td,y,\alpha) - yt_i \leq 0 \quad \text{Eq 5.13}$$

and budget constraints:

$$\Gamma_2(x,t,td,\alpha) = ct(x,t,td,\alpha) - wcost \leq 0 \quad \text{Eq 5.14}$$

where y_{max_i} is maximum delivery of pollutant constituent i after implementation of BMP combination (α) estimated with SWAT simulations over period td ; and yt_i is the allowable load of constituent i . Variable ct is the total cost of implementation of α estimated by applying Eq 5.1, and $wcost$ is the total available budget for implementation of management plans. The denominator in 5.12 was designed such that it will never be zero.

The optimization constraints in 5.13 and 5.14 are typically defined by regulatory and implementation agencies. For example, allowable sediment and nutrient loads (yt_i) may be obtained from a Total Maximum Daily Load (TMDL) for a given watershed. While yt_i and y_{max_i} can be expressed on a daily, monthly, or annual basis, as loads or concentrations, their units should be consistent. The cost constraint represents available budget for implementation of watershed management scenarios and may be specified by implementation agencies.

The fitness score (fs) for each string was evaluated by the objective function (z) associated with the string from Eq 5.12. Infeasible solutions, i.e., solutions that do not satisfy the constraint functions Γ_1 and Γ_2 in 5.13 and 5.14, were penalized by applying a penalizing factor k as:

$$fs = k \times z; \quad k = \begin{cases} 10^{-5} & ; \Gamma_1 \vee \Gamma_2 > 0 \\ 1 & ; \Gamma_1 \wedge \Gamma_2 \leq 0 \end{cases} \quad \text{Eq 5.15}$$

Operating Parameters

Selection of the size ($nstr$) of the optimization population (P) and their initial values ($\alpha_{j,i}^i$) is complex and is likely to vary for different problems (Deb, 2001). A large population provides the GA with an adequate sampling of the search space. However, there are some trade-offs

between the population size and the number of generations needed for convergence. A small population size may cause the GA to become trapped in a local optimum, whereas a large size may be computationally inefficient and take too long to converge. Winston and Venkataramanan (2003) suggest that a population of 50 or 100 is typical for GAs. In this study, population size ($nstr$) was determined through a sensitivity analysis.

The initial population (P_0) was generated by assigning random binary values (i.e. 0 or 1) for management actions in each watershed plan as shown in Figure 5.2. Subsequent generations were produced through the following steps:

1. Replacement: The strings of the previous population characterized by the best fitness scores were propagated to the next generation unchanged. The rest of the population was replaced by new chromosomes generated from mating. The replacement rate (rr) was determined through a sensitivity analysis.
2. Selection: The parents from the previous generation were selected to reproduce the offspring through the principle of survival of the fittest. A probability score (F) was assigned to each string based on its fitness score (fs in Eq 5.15) as:

$$F_j^i = \frac{fs_j^i}{\sum_{j=1}^{nstr} fs_j^i} \quad \text{Eq 5.16}$$

where $nstr$ is the population size. Then, the strings were sorted in descending order of the assigned probabilities, and their cumulative probabilities were determined. Two random numbers were generated for selection of two parent chromosomes for mating.

3. Mating: The two selected parents were used for production of two offspring. A crossover point was selected randomly so that the offspring consists of a portion of each parent. In this study, a random crossover point was chosen for each type of BMP. One point was chosen to crossover the genes corresponding to parallel terraces, one point for field borders, one for grassed waterways, and another one for grade stabilization structures.

4. Mutation: The genetic makeup of the offspring was altered randomly to avoid local optima. Typically 1% to 5% of the genes mutate per iteration. A 2% mutation rate (mr) was used here. The alleles for the randomly chosen genes changed from a “1” to “0”, or vice versa.
5. Repeat: Previous steps were repeated until the algorithm converged.

The optimization procedure was terminated once the convergence criteria, defined below, were met: (1) the fitness score of the best solution of ten consecutive generations did not vary; (2) the median fitness score of solutions in 10 consecutive generations did not vary more than 5%; and (3) values of constraint functions did not vary for ten consecutive generations. A minimum of 2000 individual model evaluations were considered.

Efficiency of the optimization algorithm for various combinations of operation parameters was evaluated through a sensitivity analysis. Table 5.3 summarizes different combinations of operating parameters (labeled as setup 1 to setup 6) that were examined in the analysis. A total of 2000 individual model evaluations were performed for each combination of operating parameters. As a result, all setups had almost identical computational runtime. The number of generations ($ngen$) in Table 5.3 was determined as:

$$ngen = \frac{nmax}{nstr \cdot rr} \quad \text{Eq 5.17}$$

where $nmax$, $nstr$, and rr refer to total number of model evaluations, population size, and replacement rate, respectively. The population size, and replacement rate with the highest efficiency was utilized to identify the optimal spatial allocation of BMPs in the study watersheds. The setup with the highest objective function value over all model simulations was appraised to be the most efficient.

| Table 5.3 Various combinations of GA operating parameters for sensitivity analysis | | | | | | | |
|---|--------|-------|-----|----|----|-----|-----|
| Parameter | Symbol | Setup | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| Population Size | $nstr$ | 100 | 100 | 50 | 50 | 20 | 20 |
| Replacement Rate (%) | rr | 70 | 90 | 70 | 90 | 70 | 90 |
| Number of Generations | $ngen$ | 28 | 22 | 58 | 45 | 142 | 110 |

5.4.5. Optimization Cases

BMPs depicted in Figure 2.1 were implemented in the Dreisbach and Smith Fry watersheds through targeting strategies. This provided the opportunity to compare the cost-effectiveness of BMP combination obtained from the optimization procedure with the combination allocated through targeting. Four cases (summarized in Table 5.4) were examined in this study. *Case 1* represented the base-case simulation with no BMPs in place. Cost-effectiveness of BMPs allocated by targeting was determined in *case 2*. The results of the optimization procedure were compared with targeting in two ways. In *case 3*, the optimization procedure was used to allocate BMPs such that maximum sediment, phosphorus, and nitrogen loads over the simulation period (2000-2009) (y_{max_i} in Eq 5.13) did not exceed the ones corresponding to the targeting plan. The purpose of comparing cases 2 and 3 was to compare the total watershed cost (c_t in Eq 5.14) of the two cases while providing the same level of water quality protection. For simulations in *case 4*, the optimization constraints consisted of only the cost constraint. This case aimed at investigating environmental benefits of BMPs from optimization and targeting plans limited to the same budget for implementation. Therefore, the optimization procedure was designed to terminate once the total watershed cost of the near optimal solution reached the cost of the targeting plan.

| Case | Strategy | Description |
|------|----------------|---|
| 1 | Base-case | No BMP implemented in the watershed. |
| 2 | Targeting | BMPs spatially allocated as in Figure 2.1. |
| 3 | Optimization 1 | BMPs allocated such that maximum monthly sediment and nutrient loads did not exceed maximum monthly values from targeting strategy. No cost constraint was imposed. |
| 4 | Optimization 2 | BMPs allocated such that total watershed cost did not exceed cost of targeting strategy. |

5.5. Results

Results of sensitivity analysis for the GA operating parameters indicated that the algorithm is more efficient with smaller population size and replacement rate. Figure 5.3 shows sensitivity of the optimization procedure to combinations of operating parameters listed in Table 5.3. Model evaluations were performed for the Dreisbach watershed over the 2000-2009 period. The trend for each setup was determined based on a total number of 2000 individual model evaluations. The runtime corresponding to 2000 model evaluations for the study watersheds was nearly 15 hours on a 2.0 GHz PC. The corresponding number of generations for each setup was determined from Eq 5.17. A mutation rate of 2% was used for all setups. The results indicate that with the same number of model evaluations and computational cost, the setup with the highest number of generations (population size of 20 with 70% replacement rate, i.e. setup 5 in Table 5.3) outperformed the other combinations. Similar results were obtained for the Smith Fry Watershed. Therefore, these values were used for optimizing the location of grassed waterways, grade stabilization structures, parallel terraces, and field borders in the study watersheds.

BMPs allocated with the optimization procedure were more cost-effective than randomly selected BMPs and BMPs allocated through targeting strategies. Table 5.5 summarizes results for base-case and targeting cases, *case 1* and *case 2*. Estimated average and maximum

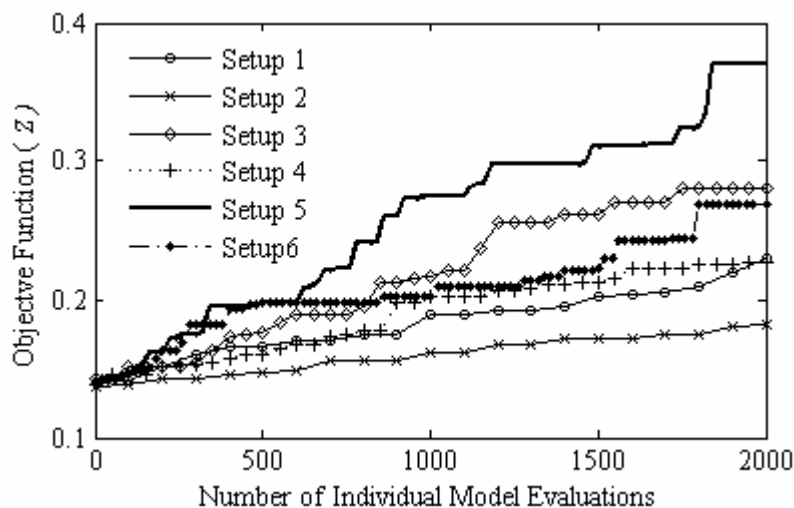


Figure 5.3 Analysis of sensitivity of GA operating parameters. Table 5.3 shows values of operating parameters corresponding to each setup.

Table 5.5 Results for base-case and targeting cases. Mean and maximum value of monthly sediment, phosphorus, and nitrogen loads were estimated from SWAT simulations.

| Watershed | Variable | Symbol | Units | Base-case (Case 1) | Targeting (Case 2) | |
|------------------|--------------------|-----------|-------------|-----------------------|----------------------------|-------------------------|
| <u>Dreisbach</u> | Sediment | Mean | y_s | t/ha/m | 0.074 | |
| | | Max | y_{max_s} | t/ha/m | 0.55 | 0.17^a |
| | | Reduction | r_s | % | - | 70 |
| | Phosphorus | Mean | y_p | kg/ha/m | 0.044 | 0.033 |
| | | Max | y_{max_p} | kg/ha/m | 0.20 | 0.15^a |
| | | Reduction | r_p | % | - | 25 |
| | Nitrogen | Mean | y_n | kg/ha/m | 0.395 | 0.298 |
| | | Max | y_{max_n} | kg/ha/m | 2.69 | 2.1^a |
| | | Reduction | r_n | % | - | 25 |
| | objective function | | z | \$/ | - | 0.12 |
| watershed cost | | ct | \$ | - | 414,690^b | |
| <u>Smith Fry</u> | Sediment | Mean | y_s | t/ha/m | 0.116 | 0.105 |
| | | Max | y_{max_s} | t/ha/m | 1.62 | 1.5 |
| | | Reduction | r_s | % | - | 10 |
| | Phosphorus | Mean | y_p | kg/ha/m | 0.32 | 0.306 |
| | | Max | y_{max_p} | kg/ha/m | 2.129 | 2.06 |
| | | Reduction | r_p | % | - | 5 |
| | Nitrogen | Mean | y_n | kg/ha/m | 4.545 | 4.368 |
| | | Max | y_{max_n} | kg/ha/m | 34.52 | 33.31 |
| | | Reduction | r_n | % | - | 4 |
| | objective function | | z | \$/ | - | 1.4 |
| watershed cost | | ct | \$ | - | 60,610^b | |

^a Used as water quality constraint for case 3 summarized in Table 5.6.

^b Used as cost constraint for case 4 summarized in Table 5.6.

monthly sediment and nutrient yields simulated by SWAT, and total watershed cost (c_i) are provided for the study watersheds. Variable z in the table refers to the benefit-cost ratio as defined in Eq 5.12. Percent reduction of constituent i (r_i) as a result of implementation of BMPs was determined as:

$$r_i = \frac{y_{i,1} - y_{i,2}}{y_{i,1}} \times 100 \quad \text{Eq 5.18}$$

where i is the constituent of interest (i.e., sediment, phosphorus, or nitrogen load), $y_{i,1}$ is the average monthly load of constituent i for the base-case (case 1), and $y_{i,2}$ is the average

monthly load of constituent i for the targeting case (*case 2*). In the Dreisbach watershed, implementation of the targeting plan (depicted in Figure 2.1) would cost \$414,690 over the 2000-2009 period, with a benefit-cost ratio (z in Eq 5.12) of 0.12. As a result, average monthly sediment, phosphorus, and nitrogen yields would be reduced from the base-case by nearly 70%, 25%, and 25%, respectively. Likewise, estimated maximum monthly sediment, phosphorus, and nitrogen yields at the outlet of the Dreisbach watershed would respectively reduce to 0.17 t/ha, 0.15 kg/ha, and 2.1 kg/ha. These values were used as the allowable sediment and nutrient loads (yt_i in Eq 5.13) for the first optimization case (*case 3*).

A summary of results for optimization cases in the study watersheds is provided in Table 5.6. Highlighted values reflect optimization constraints in each case that were obtained from targeting results in Table 5.5. Comparison of the results reveals that in the Dreisbach watershed, BMPs selected and placed by optimization (*case 3*) would yield nearly three times better benefit-cost ratio and would cost 2.5 times less than the targeting combination, while providing the same level of protection against phosphorus and providing even higher protection against sediment and nitrogen pollution.

| Watershed | Variable | | Symbol | Units | Optimization | |
|--------------------|--------------------|----------------------|-------------|---------------------------|----------------------------|---------------|
| | | | | | <i>case 3</i> | <i>case 4</i> |
| Dreisbach | Sediment | allowable | yt_s | t/ha/m | 0.17^a | - |
| | | maximum ^b | y_{max_s} | t/ha/m | 0.06 | 0.085 |
| | Phosphorus | allowable | yt_p | kg/ha/m | 0.15^a | - |
| | | maximum ^b | y_{max_p} | kg/ha/m | 0.15 | 0.082 |
| | Nitrogen | allowable | yt_n | kg/ha/m | 2.10^a | - |
| | | maximum ^b | y_{max_n} | kg/ha/m | 1.55 | 1.00 |
| Objective function | | z | \$\$ | 0.36 | 0.21 | |
| Watershed cost | | ct | \$ | 165,370 | 414,690^a | |
| Smith Fry | Sediment | allowable | yt_s | t/ha/m | was not investigated | - |
| | | maximum ^b | y_{max_s} | t/ha/m | | 0.48 |
| | Phosphorus | allowable | yt_p | kg/ha/m | | - |
| | | maximum ^b | y_{max_p} | kg/ha/m | | 1.69 |
| | Nitrogen | allowable | yt_n | kg/ha/m | | - |
| | | maximum ^b | y_{max_n} | kg/ha/m | | 26.3 |
| | Objective function | | z | \$\$ | | 3.12 |
| Watershed cost | | ct | \$ | 60,610^a | | |

^a From targeting strategy summarized in Table 5.5.
^b Estimated from SWAT simulations.

Results for *case 3* simulations in the Dreisbach watershed are shown in Figure 5.4(a). In this case, only maximum monthly sediment and nutrient loads were constrained to their allowable values obtained from the targeting strategy. A total number of 150 optimization generations each with a population of 20 strings, were computed. In the top panel in Figure 5.4(a), the left y-axis reflects benefit-cost ratio for all model evaluations (dots), and right y-axis is total cost of implementation of the best solution in each optimization generation (dashed line). The first generation represents a random selection and placement of BMPs, while the last generation shows the near optimum solution. It is evident that maximum and median fitness of generations improved as optimization progressed to future generations. This points to the efficiency of the developed algorithm. Accordingly, the total watershed cost associated with the best solutions of GA generations generally reduced in successive generations. The fitness of the best solution of GA generations did not improve once one of the pollutant constituents reached its allowable value.

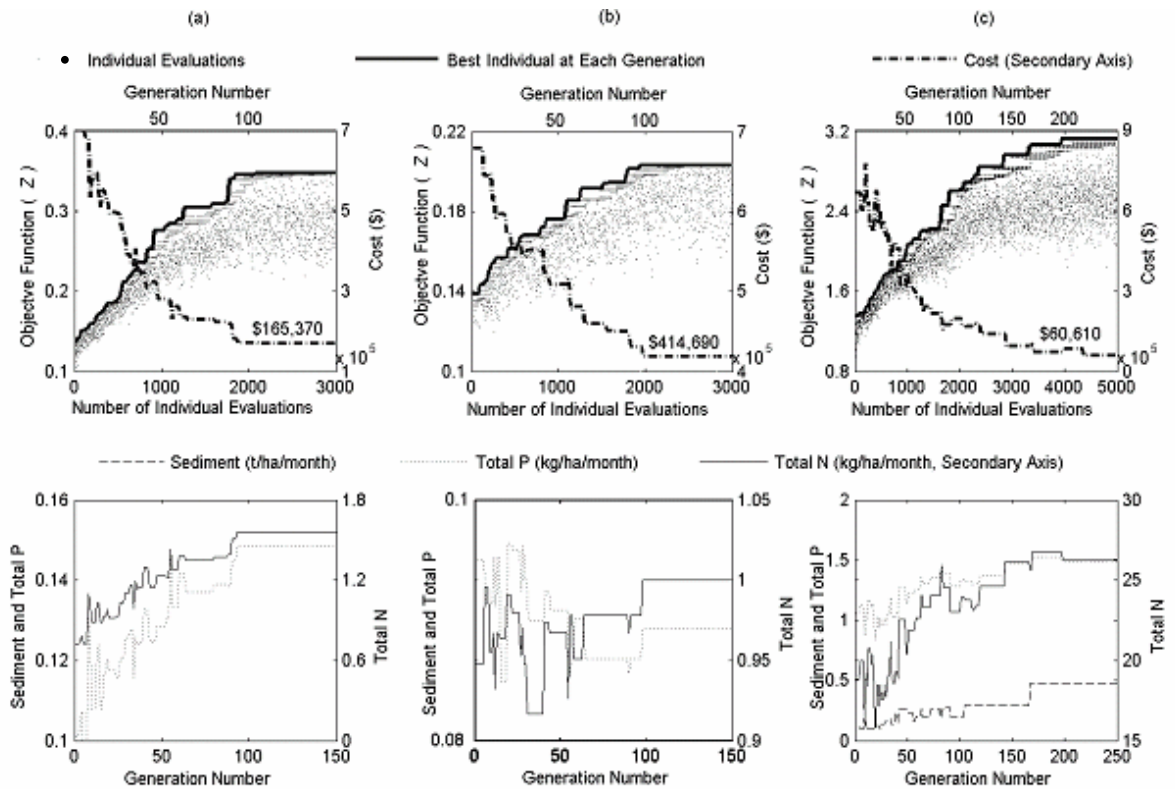


Figure 5.4 (a) Case3 optimization outputs for the Dreisbach watershed over 2000–2009 period. (b) Case 4 optimization outputs for the Dreisbach watershed over 2000–2009 period. (c) Case 4 optimization outputs for the Smith Fry watershed over 2000–2009 period.

The bottom panel in Figure 5.4(a) demonstrates phosphorus and nitrogen constraints for the best solution of optimization generations. The values in the figure reflect maximum monthly pollutant loads simulated over the 2000-2009 period. At the near optimal solution (generation 150), maximum monthly phosphorus load (y_{max_p}) approached its allowable value (0.15 kg/ha). Estimated maximum monthly nitrogen load (1.55 kg/ha) and maximum monthly sediment load (0.12 t/ha) were well below allowable values. This indicated that the GA converged to a near optimal solution with phosphorus being the restricting pollutant. Note that exceedence of only one of the pollutant loads over its allowable value was sufficient to penalize a solution. Hence, once monthly phosphorus load approached its allowable value, regardless of sediment and nitrogen loads being below their allowable values, the fitness of subsequent optimization generations did not improve, indicating that a near optimal solution was identified.

Figure 5.4(b) depicts the results of GA computations for *case 4*, where only the cost constraint of Eq 5.13 was imposed. The aim of this case was to derive an optimization plan that would cost the same as the targeting plan. Therefore, the algorithm was designed to terminate once the total cost of the optimization plan reached the total cost of the targeting plan (\$414,690). A summary of the results for *case 4* computations is provided in Table 5.6. It appeared that for the same watershed cost, the solution from optimization would yield nearly half the maximum monthly sediment and nutrient loads. Also, the benefit-cost ratio for near optimal solution obtained from *case 4* computations was nearly two times higher than the one corresponding to the targeting solution.

In the Smith Fry watershed, implementation of BMPs allocated by targeting as shown in Figure 2.1 would cost \$60,610, while reducing average monthly sediment, phosphorus, and nitrogen loads by nearly 10%, 4.5%, and 3.9%, respectively (see Table 5.5). Since reduction of pollutant loads was not significant in comparison to the Dreisbach watershed, *case 3* was not investigated for the Smith Fry watershed. The optimization tool was applied to investigate only *case 4*, for which only cost constraint was imposed. The analysis aimed at identification of near optimal combination of BMPs that would cost no more than \$60,610, equal to the cost of implementation of the targeting plan.

Figure 5.4(c) shows the results of optimization evaluations for *case 4* in the Smith Fry watershed. Results, also summarized in Table 5.6, indicate that the estimated reduction rate (Eq 5.18) of maximum monthly sediment, phosphorus, and nitrogen loads from the optimization procedure would be nearly 5 times higher than the ones estimated for the targeting plan. Similarly, the benefit-cost ratio would be more than two fold.

A sample of spatial allocation of the optimally placed BMPs in the study watersheds is depicted in Figure 5.5. In Dreisbach, *case 4* was more expensive than *case 3* with stricter sediment and nutrient constraints that were met by allocating additional field borders and also a few parallel terraces. Field borders appeared to be more widely allocated in the optimization plans. This is perhaps the case because the cost of implementation of field borders is significantly less than other types of BMPs studied here (see Table 5.1). Conversely, grade stabilization structures that are relatively more expensive than the others were not prescribed in any of the plans. Interestingly, a grassed waterway was delineated at the very downstream segment of the channel network in all three optimization plans performed for the study watersheds.

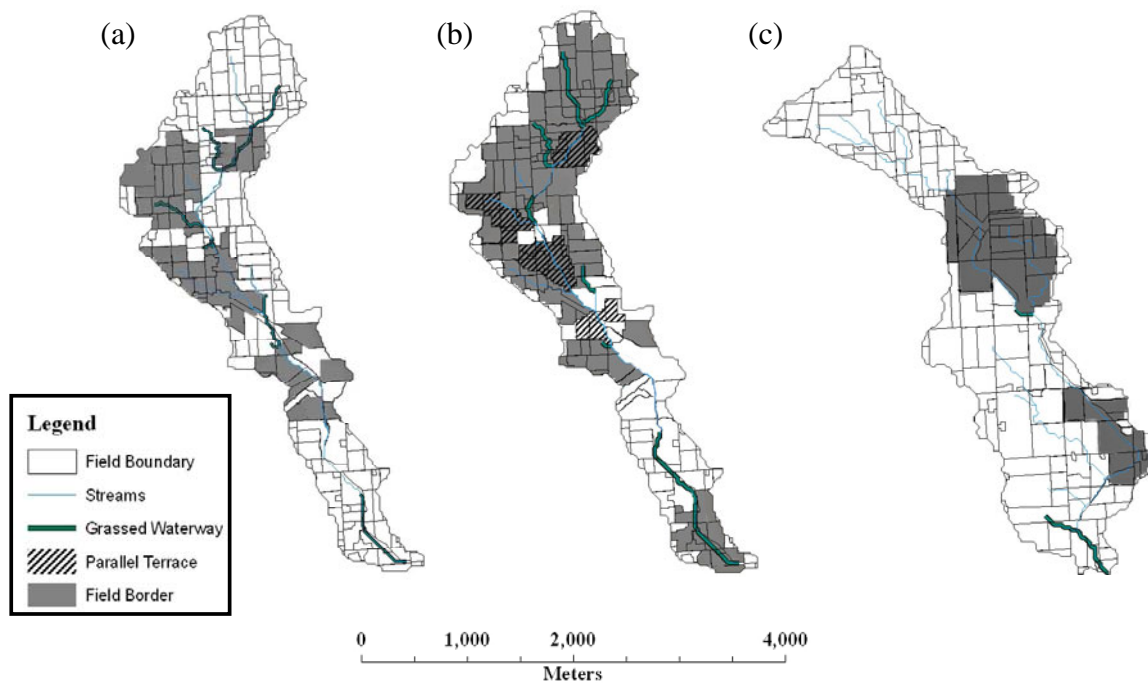


Figure 5.5 Spatial allocation of BMPs from: (a) *case 3* computations for Dreisbach watershed; (b) *case 4* computations for Dreisbach; (c) *case 4* computations for Smith Fry watershed.

5.6. Discussion

The developed optimization model shows promise for developing watershed restoration and management plans. This watershed-scale model is well suited to establish relationships between agricultural practices (particularly structural BMPs), watershed health, and satisfying aggregate policy objectives. Figure 5.6 shows a sample of reduction of pollutant loads estimated for the best solution of optimization generations versus their associated watershed cost for the Dreisbach watershed. Reductions of sediment, phosphorus, and nitrogen loads at the outlet were weighted by applying weighting factors in Table 5.2 and were rescaled to [0,1] range such that the maximum value was 1.0. These results clearly illustrate the capacity of the optimization model to handle the tradeoff amongst environmental and economic criteria in the objective function. Generating several maps similar to those of Figure 5.5 enhances the capacity of local managers and any organization undertaking a watershed planning effort to evaluate attainability of designated water quality targets at various costs. The foregoing results indicate that the optimization model is likely to converge to near optimal watershed plans that would achieve the same level of protection against NPS as targeting strategies at significantly lower costs.

The optimization model benefits from the ability to incorporate spatial heterogeneity in evaluation of both cost and benefit of BMPs. Critical source areas where BMPs yield utmost

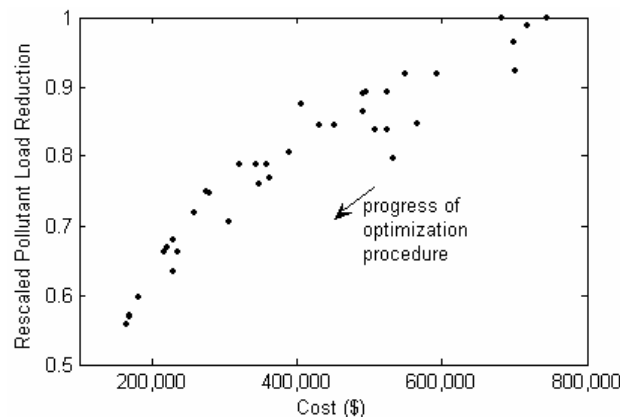


Figure 5.6 Tradeoff between economic criterion (cost) and environmental criteria (weighted pollutant load reduction) corresponding to *case 3* optimization computations for Dreisbach watershed. Weighted pollutant loads were rescaled to [0,1] range such that the maximum value was 1.

benefits are implicitly identified as the GA progresses to the next generations. Therefore, benefits of BMPs implemented in critical areas are determined as a function of landscape characteristics. Although unit cost of implementation of BMPs is the same throughout the watershed, the total cost is also a function of topographic, soil and land use properties. Moreover, the optimization procedure handles the interaction between BMPs. The results of investigated cases in Figure 5.5 indicated that grassed waterways that are implemented within the channel network were prescribed in near optimal schemes. Conversely, grade stabilization structures, while very effective in reducing sediment loads (Arabi et al., 2004), were not allocated in the watersheds, perhaps because of relatively high unit cost of establishment (see Table 5.1).

The optimization procedure can be performed on a daily, monthly, or annual basis for as many pollutant constituents as desired. This spatial optimization tool can be linked with other NPS models through a similar methodology to simulate for constituents that currently cannot be simulated by SWAT.

The major concern with regard to practicality of the proposed optimization framework is the CPU time required for computations. The search for near optimal solutions is computationally demanding mainly because of recurring NPS model simulations. Although the experienced gained from this research provides guidance as to how select the GA operating parameters, future studies shall focus on exploring the means to expedite the procedure. Parallel computing techniques hold great promise in this regard. Also, improvements in the structure of NPS model(s) that simulates pollutant loads and BMP representation method would enhance the accuracy of near optimal management plans.

5.7. Conclusions

A GA-based optimization procedure was developed for selection and placement of BMPs. The sensitivity of the model to different combinations of GA operating parameters, including population size and replacement rate, was tested in order to identify the most efficient combination that converges rapidly for a given runtime. For two small watersheds in Indiana, a setup with a higher number of generations and lower population size was more efficient.

However, these results may be site-specific and vary for watersheds with different spatial scale and characteristics.

The cost effectiveness of the optimized BMPs was compared to that of BMPs prescribed through targeting strategies in the study watersheds. Targeting strategies are developed based on field and modeling studies and intend to allocate BMPs in portions of the watershed that are deemed to extensively contribute to pollutant loads. In the absence of systematic approaches, however, identifying such locations and also designing most effective BMP combinations become infeasible at the watershed scale. It was demonstrated that the BMPs from optimization would achieve the same level of sediment and nutrient reductions with nearly one third of the cost required for implementation of the targeting scenario. Conversely, it was shown that an optimized management scheme would likely provide nearly twice the protection against sediment and nutrient loads for the same amount of money that would be spent for implementation of the targeting plans in these watersheds.

In all optimization plans, a grassed waterway was allocated to the very downstream of the channel network and immediately upstream of the watershed outlet where water quality standards were imposed. Field borders were the most commonly allocated BMPs, perhaps because of their relatively lower unit cost. Conversely, grade stabilization structures that cost significantly more than the other types of BMPs in this study were not included in any of the optimized plans.

SECTION 6. CONCLUSIONS

A computational framework was developed and tested on two small watersheds in Indiana to help watershed management agencies identify effective management practices and their spatial allocations for sediment and nutrient control. The method utilized simulations of a distributed-parameter watershed model, Soil and Water Assessment Tool (SWAT), in conjunction with statistical and optimization techniques.

In this study, the results of a previous study by Arabi et al. (2004) in the same study watersheds were adopted. First, the SWAT model was calibrated and validated for the study watersheds. The results indicated that SWAT streamflow, sediment, total P and total N simulations adequately represent observations. Two error statistics, coefficient of determination and Nash-Sutcliff efficiency coefficient, were used for calibration purposes. Model calibration was performed until a coefficient of determination greater than or equal to 0.6 and a Nash-Sutcliff coefficient greater than or equal to 0.5 was obtained for each constituent. Second, SWAT model parameters were used to represent BMPs based on their functionality and hydrologic and water quality processes that are modified by their implementation. Field borders, parallel terraces, grassed waterways, and grade stabilization structures were the specific BMPs modeled in this study. It would appear that SWAT model is an appropriate model for representation of agricultural management scenarios. The effectiveness of BMPs was evaluated by comparing model simulations for two scenarios representing conditions with and without the presence of BMPs. It was observed that implementation of the BMPs in the Dreisbach watershed would result in a significant reduction of sediment and nutrient yields. These reductions would be mainly due to implementation of grade stabilization structures and grassed waterways that appreciably reduce the transport capacity of channel network.

A systematic method based on the Generalized Likelihood Uncertainty Estimation (GLUE) methodology was developed for the estimation of uncertainty bounds for model predictions.

The procedure can be utilized for quantification of the margin of safety (MOS) in the TMDL allocation formula as an explicit uncertainty analysis method. In this study, the Generalized Likelihood Uncertainty Estimation method was utilized on the two Indiana watersheds. The forgoing results indicated that the GLUE methodology is suitable not only for determination of a numeric value for the MOS, but also for addressing the issues that render utility of the conventional approach for evaluation of BMPs subjective to the questions of *uniqueness* of calibrated parameter set, *identifiability*, and *equifinality* as described in Beven and Binley (1992).

Additionally, a computational framework was developed in which investigation of uncertainty provides complementary quantitative and qualitative information in support of management and decision making. The analysis focused specifically on two issues. First, the univariate Regionalized Sensitivity Analysis (RSA) ranking method was applied in conjunction with the multivariate Tree Structured Density Estimation (TSDE) analysis to identify critical hydrologic and water quality processes that control the ability to attain specified water quality conditions. The implications of such analysis provide benefits for watershed management programs such as TMDL. Application of the RSA-TSDE procedure for two small watersheds in Indiana, Dreisbach (6.25 km²) and Smith Fry (7.3 km²), showed that fluvial channel processes appeared to control sediment yield at the outlet of the watersheds. Therefore, implementation of within-channel BMPs such as grassed waterways and grade stabilization structures are likely more effective for reducing sediment loads at the outlet of the study watersheds. However, total nitrogen yield at the outlets was controlled by upland nitrogen loading, indicating that implementation of parallel terraces or fertilizer application strategies would be more effective. Also, EPA has been advised to promote research on translating effectiveness needs into Use Attainability Analysis (UAA) implications (NRC, 2001). The use of UAA, which is applied for setting new water quality standards or revising the existing ones, has been restricted due to inadequate technical guidance from EPA. The results of the TSDE method can provide a numerical probability value that can be utilized to answer pragmatic questions such as how feasible a target value is for a desired constituent.

Complementary to the RSA-TSDE approach that facilitates identification of effective management practices, a GA-based optimization model was developed for BMP placement within watersheds. The cost effectiveness of the optimized BMPs for the study watersheds was compared to a combination of BMPs that was implemented nearly 30 years ago through targeting. Results indicated that the benefit-cost ratio of the combination from optimization was more than two times higher than the one from targeting, providing the same sediment and nutrient reduction at the outlet. The optimization procedure can be utilized to produce a number of near optimal solutions, providing flexibility for management and decision making to meet water quality target values in an economic manner. Also, the sensitivity of the model to different combinations of operating parameters including population size and replacement rate was tested in order to identify the most efficient combination which converges faster for a given runtime. For two small watersheds in Indiana, it would appear that a setup with a higher number of generations and lower population size was more efficient.

While the foregoing results may be site-specific and limited to watersheds with spatial scale and characteristics similar to the watersheds used in this study, the developed sensitivity analysis, uncertainty estimation, and optimization procedures can be applied to in other watershed studies.

BIBLIOGRAPHY

- Arabi, M., Govindaraju, R.S., Hantush, M.M. (2007a). A probabilistic approach for analysis of uncertainty in evaluation of watershed management practices.” *Journal of Hydrology*, 333: 459-471.
- Arabi, M., R.S. Govindaraju, B. Engel, Hantush, M.M. (2007b). “Multiobjective sensitivity analysis of sediment and nutrient processes with a watershed model.” *Water Resources Research*, 43, W06409, doi:10.1029/2006WR005463.
- Arabi, M., Govindaraju, R.S., Hantush, M.M. (2006a). “Cost-effective allocation of watershed management practices using a genetic algorithm.” *Water Resources Research*, 42, W10429, doi:10.1029/2006WR004931.
- Arabi, M., Govindaraju, R.S., Hantush, M.M. (2006b). “Role of watershed subdivision on evaluation of long-term impact of best management practices on water quality.” *Journal of American Water Resources Association*, 42(2): 513-528.
- Arabi, M., Govindaraju, R.S., Hantush, M.M. (2004). “Source identification of sediments and nutrients in watersheds-Final report.” *EPA/600/R-05/080*, National Risk Management Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH 45268, 120 pp. Available online at: www.epa.gov/ORD/NRMRL/pubs/600r05080/600r05080.pdf.
- Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R. (1998). “Large area hydrologic modeling and assessment part I: model development.” *Journal of the American Water Resources Association*, 34(1): 73-89.
- Arnold, J.G. and Fohrer, N. (2005). “SWAT2000: Current capabilities and research opportunities in applied watershed modeling.” *Hydrological Processes*, 19(3):563-572.

-
- ASAE, 2003. *Design, Layout, Construction and Maintenance of Terrace Systems*. ASAE Standards S268.4 FEB03, St. Joseph, MI.
- Bastidas, L.A. (1998). "Parameter estimation for hydrometeorological models using multicriteria methods." Ph.D. dissertation, Department of Hydrology and Water Resources, University of Arizona, Tuscan, AZ.
- Bastidas, L.A., Gupta, H.V., Sorooshian, S., Shuttleworth, W.J., Yang, Z.L. (1999). "Sensitivity analysis of a land surface scheme using multicriteria methods." *Journal of Geophysical Research*, 104(D16):19,481-19,490.
- Beck, M.B. (1987). "Water quality modeling: a review of the analysis of uncertainty." *Water Resources Research*, 23(8): 1393-1442.
- Bekele, E.G. and Nicklow, J.W. (2005). "Multiobjective management of ecosystem services by interactive watershed modeling and evolutionary algorithms." *Water Resources Research*, 41, W10406, doi:10.1029/2005WR004090.
- Benaman, J. and Shoemaker, C.A. (2004). "Methodology for analyzing range of uncertain model parameters and their impact on total maximum daily load process." *Journal of Environmental Engineering*, ASCE, 130(6): 648-656.
- Beven, K.J. and Binely, A.M. (1992). "The future of distributed models: model calibration and uncertainty prediction." *Hydrological Processes*, 6: 279-298.
- Beven, K.J. and Freer, J. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology." *Journal of Hydrology*, 249(2001): 11-29.
- Bracmort, K.S., Arabi, M., Frankenberger, J.R., Engel, B.A., and Arnold, J.G. (2006). "Modeling long-term water quality impact of structural BMPs." *Transactions of the ASABE*, 49(2): 367-374.

-
- Buckner, E.R. (2001). "An evaluation of alternative vegetative filter strip models for use on agricultural lands of the Upper Wabash river watershed." Ph.D. Dissertation, Purdue University, West Lafayette, IN 47907.
- Cangelosi, A., Wether, R., Taverna, J., and Cicero, P. (2001). *Revealing the economic value of protecting the Great Lakes*, Northeast-Midwest Institute and the National Oceanic and Atmospheric Administration, Washington, DC.
- Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*, Wiley, Chichester, NY.
- Demarty, J., Otle, C., Braud, I., Oliosio, A., Frangi, J.P., Gupta, H.V., Bastidas, L.A. (2005). "Constraining a physically based soil-vegetation-atmosphere transfer model with surface water content and thermal infrared brightness temperature measurements using a multiobjective approach." *Water Resources Research*, 41, W01011, 15 pp.
- Dilks, D.W. and Freedman, P.L. (2004). "Improved consideration of the Margin of Safety in Total Maximum Daily Load development." *Journal of Environmental Engineering*, ASCE, 130(6):690-694.
- Edwards, D.R., Daniel, T.C., Scott, H.D., Murdoch, J.F., Habiger, M.J., Burkes, H.M. (1996). "Stream quality impacts of Best Management Practices in a Northwestern Arkansas basin." *Journal of the American Water Resources Association*, 32(3): 499-509.
- EPA (2003). *Introduction to the Clean Water Act*, Available online at:
<http://www.epa.gov/watertrain/cwa/>.
- EPA (2005a). "Chapter 7. Analyze data to characterize the watershed and pollutant sources." in *Draft Handbook for Developing Watershed Plans to Restore and Protect Our Waters*, EPA 841-B-05-005, October 2005, Available online at:
http://www.epa.gov/owow/nps/watershed_handbook/.
- EPA (2005b). *2002 Section 303(d) List Fact Sheet for INDIANA*, Available online at:
http://oaspub.epa.gov/waters/state_rept.control?p_state=IN, accessed on August 2005.

-
- Fang, F. and Easter, W.E. (2003). "Pollution trading to offset new pollutant loadings – A case study in the Minnesota River basin." *American Agricultural Economics Association Annual Meeting*, Montreal, Canada.
- Fiener, P. and Auerswald, K. (2006). "Seasonal variation of grassed waterway effectiveness in reducing runoff and sediment delivery from agricultural watersheds in temperate Europe." *Soil and Tillage Research*, 87: 48-58.
- Fiener, P. and Auerswald, K. (2003). "Concept and effects of a multi-purpose grassed waterway." *Journal of Environmental Quality*, 32:927–936.
- Freer, J. and Beven, K. (1996). "Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach." *Water Resources Research*, 32(7): 2161-2173.
- Green, W.H. and Ampt, G.A. (1911). "Studies on soil physics: 1. the flow of air and water through soils." *Journal of Agricultural Sciences*, 4: 11-24.
- Griffin, C.B. (1995). "Uncertainty analysis of BMP effectiveness for controlling nitrogen from urban nonpoint sources." *Journal of the American Water Resources Association*, 31(6): 1041-1050.
- Heuvelmans, G., Muys, B., Feyen, J. (2004). "Evaluation of hydrological model parameter transferability for simulating the impact of land use on catchment hydrology." *Physics and Chemistry of the Earth*, 29(2002): 739-747.
- Ice, G. (2004). "History of innovative best management practice development and its role in addressing water quality limited waterbodies." *Journal of Environmental Engineering*, ASCE, 130(6): 684-689.
- Indiana Environmental Quality Incentives Program (2004). *2004 EQIP Cost List Information* Available online at: http://www.in.nrcs.usda.gov/programs/2004eqip/state_costlist.html.
- Kirsch, K., Kirsch, A., Arnold, J.G. (2002). "Predicting sediment and phosphorous loads in the Rock River basin using SWAT." *the Transactions of ASAE*, 45(6): 1757-1769.

-
- Lake, J. and Morrison, J. (1977). "Environmental impacts of land use on water quality, Final report on the Black Creek project-project data." *EPA-905/9-77-007-C*, U.S. Environmental Protection Agency, Chicago, IL.
- Lake, J., Morrison, J. (1978). "Environmental impact of land use on water quality- Final report on the Black Creek project- Technical report." *EPA-905/9-77-007-B*, U.S. Environmental Protection Agency, Chicago, IL.
- Lenhart, T., Eckhardt, K., Fohrer, N., Frede, H.-G. (2002). "Comparison of two approaches of sensitivity analysis." *Physics and Chemistry of the Earth*, 27 (2002): 645-654.
- Liu, Y., Gupta, H.V., Sorooshian, S., Bastidas, L.A., Shuttleworth, W.J. (2004). "Exploring parameter sensitivities of the land surface using a locally coupled land-atmosphere model." *Journal of Geophysical Research*, 109, D21101, 13 pp.
- McKay, M.D., Beckman, R.J., Conover, W.J. (1979). "A Comparison of Three Methods of Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics*, 21: 239-245.
- Meixner, T., Gupta, H.V., Bastidas, L.A., Bales, R.C. (1999). "Sensitivity analysis using mass flux and concentration." *Hydrological Processes*, 13:2233-2244.
- Morris, M.D. (1991). "Factorial sampling plans for preliminary computational experiments." *Technometrics*, 33: 161-174.
- Morrison, J. and Lake, J. (1983). "Environmental impacts of land use on water quality, Black Creek Project- Final report." Allen County Soil and Water Conservation District, Fort Wayne, IN.
- Mostaghimi, S., Park, S.W., Cook, R.A., Wang, S.Y. (1997). "Assessment of management alternatives on a small agricultural watershed." *Water Research*, 31(8): 1867-1878.
- Muleta, M.K. and Niklow, J.W. (2005). "Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model." *Journal of Hydrology*, 306 (2005): 127-145.

-
- Nash, J.E. and Sutcliffe, J.V. (1970). "River flow forecasting through conceptual models part 1- A discussion of principles." *Journal of Hydrology*, 10(1970): 282-290.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., King, K.W. (2002). *Soil and Water Assessment Tool Theoretical Documentation, Version 2000*. Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, TX.
- NRC (2001). *Assessing the TMDL Approach to Water Quality Management*, Committee to Assess the Scientific Basis of the Total Maximum Daily Load Approach to Water Pollution Reduction, Water Science and Technology Board, Division on Earth and Life Studies, National Research Council, Washington, D.C.
- Osidele, O.O., Zeng, W., Beck, M.B. (2003). "Coping with uncertainty: a case study in sediment transport and nutrient load analysis." *Journal of Water Resources Planning and Management*, 129(4): 345-355.
- Pitman, A.J. (1994). "Assessing the sensitivity of land-surface scheme to the parameter values using a single column model." *Journal of Climate*, 7(12)1856-1869.
- Ribardo, M.O., Colacicco, D., Barbarika, A., and Young, C.E. (1989). "The economic efficiency of voluntary soil conservation programs." *Journal of Soil and Water Conservation*, 44: 40-43.
- Ringuest, J.L. (1992). *Multiobjective Optimization: Behavioral and Computational Considerations*, Boston, MA: Kluwer.
- Saleh, A., Arnold, J.G., Gassman, P.W., Hauch, L.M., Rosenthal, W.D., Williams, J.R., McFarland, A.M.S. (2000). "Application of SWAT for the Upper North Bosque River watershed." *the Transactions of ASAE*, 43(5): 1077-1087.
- Saltelli, A., Chan, K., Scott, E.M. (2000). *Sensitivity Analysis*, John Wiley & Sons Ltd., West Sussex PO19 8SQ, England.

-
- Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M. (2001). "Validation of the SWAT model on a large river basin with point and nonpoint sources." *Journal of the American Water Resources Association*, 37(5): 1169-1188.
- Santhi, C., Srinivasan, R., Arnold, J.G., Williams, J.R. (2003). "A modeling approach to evaluate the impacts of water quality management plans implemented in the Big Cypress Creek Watershed." second conference on *Watershed Management to Meet Emerging TMDL Environmental Regulations*, Albuquerque, NM, pp. 384-394.
- Spear, R.C. and Hornberger, G.M. (1980). "Eutrophication in Peel Inlet II: Identification of critical uncertainties via generalized sensitivity analysis." *Water Research*, 14: 43-49.
- Spear, R.C., Grieb, T.M., Shang, N. (1994). "Parameter uncertainty and interaction in complex environmental models." *Water Resources Research*, 30(11): 3159-3169.
- Srivastava, P., Robillard, P.D., Hamlet, J.M., and Day, R.L. (2002). "Watershed optimization of best management practices using AnnAGNPS and a genetic algorithm." *Water Resources Research*, 38(3): 1021.
- USDA Soil Conservation Service (1972). "Section 4- Hydrology, Chapters 4-10." *National Engineering Handbook*, USDA, Washington, DC.
- USDA (2003). *The 2002 Farm Bill: provisions and economic implications*, Available online at: <http://www.ers.usda.gov/Features/FarmBill/>.
- USDA-NRCS (2005). "National Conservation Practice Standards- NHCP." *Electronic Field Office technical Guide (eFOTG)*, Section IV-Practice Standards and Specifications, Available online at: <http://www.nrcs.usda.gov/Technical/Standards/nhcp.html>, Accessed on August 2006.
- van Griensven, A. and Bauwens W. (2003). "Multiobjective autocalibration for semidistributed water quality models." *Water Resource Research*, 39(12):1348, doi:10.1029/2003WR002284, 2003.

-
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R. (2006). "A global sensitivity analysis tool for the parameters of multi-variable catchment models." *Journal of Hydrology*, 324(2006):10-23.
- Vazquez-Amabile, G., Engel, B.A., Flanagan, D.C. (2006). "Modeling and risk analysis of nonpoint-source pollution caused by atrazine using SWAT." *Transactions of the ASABE*, 49(3): 667–678.
- Veith, T.L., Wolfe, M.L., and Heatwole, C.D. (2004). "Cost-effective BMP placement: optimization versus targeting." *Transactions of the ASAE*, 47(5): 1585-1594.
- White, K.L. and Chaubey, I. (2005). "Sensitivity analysis, calibrations, and validation for a multisite and multivariate SWAT model." *Journal of the American Water Resources Association*, 41 (5): 1077-1089.
- Whittaker, G., Fare, R., Srinivasan, R. and Scott, D.W. (2003). "Spatial evaluation of alternative nonpoint nutrient regulatory instruments." *Water Resources Research*, 39(4) 1079, doi:10.1029/2001 WR001119, 2003.
- Williams, J.R. (1969). "Flood routing with variable travel time or variable storage coefficients." *the Transactions of the ASAE*, 12(1): 100-103.
- Williams, J.R. (1975). "Sediment-yield prediction with universal equation using runoff energy factor, in Present and prospective technology for predicting sediment yield and sources." *Proceedings of the Sediment Yield Workshop*, Oxford, MS, pp. 244-252.
- Winston, W.L. and Venkataramanan, M. (2003). *Introduction to Mathematical Programming* Brooks/Cole, California 93950, USA.
- Wischmeier, W.H. and Smith, D.D. (1978). "Predicting Rainfall Erosion Losses, a Guide to Conservation Planning." *Agriculture Handbook NO. 537*, US Department of Agriculture, Washington D.C.

Zhang, H.X. and Yu, S.L. (2004). "Applying the first-order error analysis in determining the margin of safety for total maximum daily load computations." *Journal of Environmental Engineering*, 130(6): 664-673.

Zheng, Y. and Keller, A.A. (2006). "Understanding parameter sensitivity and its management implications in watershed-scale water quality modeling." *Water Resources Research*, 42, W05402, 14 pp.

Zitzler, E. and Thiele, L. (1999). "Multiobjective evolutionary algorithms: a comparative case study and the Strength Pareto approach." *IEEE Transactions on Evolutionary Algorithms*, 3(4): 257-271.



***United States
Environmental Protection
Agency***

***National Risk Management
Research Laboratory
Cincinnati, OH 45268***

***Official Business
Penalty for Private Use
\$300***

***Contract Number: 4C-R330-NAEX
January 2007***