DISSERTATION

A MULTI CRITERIA DECISION SUPPORT SYSTEM FOR WATERSHED

MANAGEMENT UNDER UNCERTAIN CONDITIONS

Submitted by

Mahdi Ahmadi

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2012

Doctoral Committee:

    Advisor: Mazdak Arabi

    James C. Ascough II
    Darrell G. Fontane
    Dana L. Hoag

ABSTRACT

A MULTI CRITERIA DECISION SUPPORT SYSTEM FOR WATERSHED
MANAGEMENT UNDER UNCERTAIN CONDITIONS

Nonpoint source (NPS) pollution is the primary cause of impaired water bodies in the United States and around the world. Elevated nutrient, sediment, and pesticide loads to waterways may negatively impact human health and aquatic ecosystems, increasing costs of pollutant mitigation and water treatment. Control of nonpoint source pollution is achievable through implementation of conservation practices, also known as Best Management Practices (BMPs). Watershed–scale NPS pollution control plans aim at minimizing the potential for water pollution and environmental degradation at minimum cost. Simulation models of the environment play a central role in successful implementation of watershed management programs by providing the means to assess the relative contribution of different sources to the impairment and water quality impact of conservation practices.

While significant shifts in climatic patterns are evident worldwide, many natural processes, including precipitation and temperature, are affected. With projected changes in climatic conditions, significant changes in diffusive transport of nonpoint source pollutants, assimilative capacity of water bodies, and landscape positions of critical areas that should be targeted for implementation of conservation practices are also expected. The amount of investment on NPS pollution control programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans. Coupling of watershed models with regional climate projections can potentially provide answers to a variety of questions on the dynamic linkage between climate and ecologic health of water resources.

The overarching goal of this dissertation is to develop a new analysis framework for the development of optimal NPS pollution control strategy at the regional scale under projected future climate conditions. Proposed frameworks were applied to a 24,800 ha watershed in the Eagle Creek Watershed in central Indiana. First, a computational framework was developed for incorporation of disparate information from observed hydrologic responses at multiple locations into the calibration of watershed models. This study highlighted the use of multiobjective approaches for proper calibration of watershed models that are used for pollutant source identification and watershed management. Second, an integrated simulation–optimization approach for targeted implementation of agricultural conservation practices was presented. A multiobjective genetic algorithm (NSGA-II) with mixed discrete–continuous decision variables was used to identify optimal types and locations of conservation practices for nutrient and pesticide control. This study showed that mixed discrete–continuous optimization method identifies better solutions than commonly used binary optimization methods. Third, the conclusion from application of NSGA-II optimization followed by development of a multi criteria decision analysis framework to identify near–optimal NPS pollution control plan using *a priori* knowledge about the system. The results suggested that the multi criteria decision analysis framework can be an effective and efficient substitute for optimization frameworks. Fourth, the hydrologic and water quality simulations driven by an extensive ensemble of climate projections were analyzed for their respective changes in basin average temperature and precipitation. The results revealed that the water yield and pollutants transport are likely to change substantially under different climatic paradigms. And finally, impact of projected climate change on performance of conservation practice and shifts in their optimal types and locations were analyzed. The results showed that performance of NPS control plans under different climatic projections will alter substantially; however, the optimal types and locations of conservation practices remained relatively unchanged.

DEDICATION

This dissertation is dedicated to my parents and siblings for their unconditional love, support, and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Problem Statement and Research Goals

Managing water quality and quantity is essential for socioeconomic development and critical for a long-term commitment to environmental viability. Population and industrial growth intensifies competition for freshwater resources. Moreover, increasing demands from agricultural, energy, and environmental uses together with projected climate change and interaction between physical drivers exacerbate the conflicts and complexity of the problem. On the other hand, limited budget necessitates identifying a compromise solution between development and resource protection. Achieving such a solution requires the best efforts of scientists, engineers, economists, and policymakers together with the public to implement adaptive management strategies. When more than one decision criterion comes to play a role, decision making essentially becomes a *multi criteria decision making* (MCDM) problem. In particular, control of water resources pollution at the watershed scale is inherently a *multiobjective* problem.

In the United States, destructive land use is the major cause of environmental and water quality problems [*Horowitz et al.*, 2007]. Diffusive nature of the *nonpoint sources* (NPS) of pollutants from agricultural areas is the leading source of impairment in the nation's rivers, lakes, wetlands, and a major source of impairment in bays and estuaries [*US EPA*, 2009]. Agriculture's contribution has remained relatively unchanged over the past decade [*Ribaudo and Johansson*, 2006]. By growing

the national awareness of the increasingly dominant influence of nonpoint source pollutants on water quality, U.S. Congress amended the Clean Water Act in 1987 to focus greater national efforts on nonpoint sources. The goal of this amendment was to expedite control of nonpoint sources of pollution [*US Congress*, 2002]. Section 319 of the Clean Water Act encourages states to address water pollutions by assessing nonpoint source pollution causes and problems and support nonpoint source pollution control programs within the state [*US EPA*, 1993]. Control of agricultural nonpoint source pollution is achievable through implementation of *conservation practices*, commonly known as *Best Management Practices* (BMPs). The *Total Daily Maximum Load* (TMDL) program administered by the EPA attempts to balance the fluxes of pollutants from a contributing watershed to assimilative capacity of receiving water bodies without violating water quality standard.

While significant shifts in climatic patterns are evident worldwide, many natural processes, including precipitation and temperature and therefore frequency and magnitude of extreme hydrologic events will also alter. With projected changes in climatic conditions, significant changes in diffusive transport of nonpoint source pollutants, assimilative capacities of water bodies, and landscape position of critical areas that should be targeted for implementation of conservation practices are also expected [*Parker et al.*, 2008; *Jennings et al.*, 2009; *Kaini et al.*, 2010]. The amount of investment on nonpoint source pollution control programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans.

Watershed–scale conservation planning aims at minimizing the potential for water pollution and environmental degradation at minimum cost. An effective nonpoint source pollution control strategy should identify optimal type, location, and timing of the practices and provide information on effectiveness of the plan. However, it cannot be tested for all potential cases in a watershed scale. Thus, watershed

planners need to consult with engineers and professional resource managers to ensure that the actions being considered are realistic and capable of meeting water quality objectives [*US EPA*, 2008]. Engineers rely on computer models to provide an estimate of adaptation practices impacts on improving water quality. Watershed models are increasingly embedded in the decision making process to address a wide range of hydrologic and water quality issues. Simulation models of the environment play a central role in successful implementation of watershed management programs by providing the means to assess the relative contribution of different sources (i.e., stressors) to the impairment. Therefore, it is of keen interest to evaluate the performance validity of watershed models according to the past observations of fluxes of water and contaminants at multiple locations on the stream network. Coupling of watershed models with regional climate projections can potentially provide answers to a variety of questions on the dynamic linkage between climate and ecologic health of water resources.

The main goal of this dissertation is to present a new integrated simulation–optimization framework for identifying optimal conservation plans at the watershed scale under projected climatic condition. The results of this dissertation are significant in several aspects of the watershed management, including calibration and sensitivity analysis of watershed models, hybrid optimization of targeted conservation plans, impact analysis of climate change on hydrologic and contaminants fluxes, and performance of conservation plans under changing climate.

1.2 Background and Specific Objectives

Nonpoint source pollution is the primary cause of impaired water bodies in the United States and around the world *US EPA* [2008]. Elevated nutrient, sediment and pesticide loads to waterways may negatively impact human health and aquatic ecosystems, increasing costs of pollutant mitigation and water treatment

[*Novotny*, 1993]. Agricultural conservation practices are widely accepted control measures of nonpoint source pollutants [*Novotny*, 1993; *Ritter*, 2001; *Mostaghimi et al.*, 1997].The primary goal of watershed scale conservation plans is minimizing pollutants movement from landscapes to water bodies at minimum cost. Watershed plans for nonpoint source pollution control are developed using several approaches. Government agencies promote implementation of NPS pollution control practices by recommending cost-share program that is a field-by-field approach [*Veith et al.*, 2003]. While this approach might be effective in a field or farm level, maximum water quality benefits at the watershed scale is not guaranteed [*Arabi et al.*, 2006]. Thus, NPS pollution control can be enhanced through development of the watershed scale conservation plans [*Maringanti et al.*, 2011]. The critical source area targeting method is a watershed scale planning approach that suggests implementation of conservation practices in critical source areas within the watershed, which contribute larger amounts of NPS pollutants. A major drawback with critical source area approach is that it does not incorporate all important watershed processes and interactions, and therefore, does not guarantee the cost-effectiveness of the developed conservation plan. In addition, monitoring long-term impact of implemented conservation practices on water quality via field-studies is infeasible mainly because of the complexity of hydrologic and water quality processes and changes in annual weather pattern [*Arabi et al.*, 2007; *Veith*, 2002]. Moreover, impact of conservation practices can be delayed for several years [*Veith*, 2002]. Hence, estimation of the water quality benefits of NPS pollution control strategies necessitate the use of appropriate hydrologic and water quality models. Watershed models can simulate hydrologic and water quality responses within a watershed system of interest and can help to identify water quality impact of conservation practices.

Hydrologic and water quality processes are highly complex and comprise a network of nonlinear dynamics [*US EPA*, 2008; *Arnold et al.*, 1998]. Distributed watershed models are commonly used to simulate natural processes and the response

of the watershed to changes in land use, climate, and land and water management. The number of possible NPS pollution control scenarios within a watershed increases exponentially with the number of fields. Thus, it is not possible to evaluate performance of all possible scenarios at all fields within the watershed. Furthermore, incorporation of socioeconomic factors in the watershed planning process increases complexity of the procedure for identification of cost-effective nonpoint source pollution control plan. Recent studies have shown that optimization methods hold great promise for optimal allocation of NPS pollution control measures at the watershed scale [*Arabi et al.*, 2006; *Veith*, 2002; *Jha et al.*, 2009]. In this approach, a proper representation of conservation practices is required to predict water quality changes arising from adoption of conservation plans [*Easton et al.*, 2008].

1.2.1 Objectives

The objectives of the study were to:

O1. Present a framework for parameter estimation of the watershed models for multisite and many objective. The framework will include several single and multiobjective optimization methods, flexible and statistically correct formulation of likelihood function, and appropriate data transformations methods.

O2. Identify a set of optimal conservation plans for reducing pollutant loads at minimum cost for a range of feasible budget and predict the effectiveness of the plans with an improved formulation of the environmental and economic objectives and decision variables.

O3. Present a multi criteria decision analysis framework to identify near–optimal nonpoint source pollution control strategies for a range of feasible budget based on *a priori* knowledge about the system.

<u>O4.</u> Predict the changes in fluxes of water, sediment, nitrogen, phosphorus, and pesticide over the course of the next several decades in response to the projected climate conditions.

<u>O5.</u> Assess the effectiveness of watershed-scale agricultural conservation strategies for minimizing vulnerability under uncertain climatic conditions.

1.2.2 Proposition

The objectives will be achieved by means of a careful examination of the following propositions:

<u>P1.</u> Calibration of watershed models for many hydrologic and water quality responses at multiple locations requires a framework that can find solutions with minimum weighted errors, but more importantly, can adequately mimic the observed behavior of the system at all locations for all objectives. This needs a multiobjective optimization framework with a statistically correct likelihood function.

<u>P2.</u> Coupled simulation–optimization frameworks can effectively identify optimal type and placement of the conservation practices for a given amount of available budget ("pollutant load–cost" Pareto-optimal fronts).

   (a) Multiobjective optimization methods can explicitly identify the optimal siting of the conservation practices and identify the best solutions for a range of feasible budget.

   (b) Nonpoint source pollution planning does not necessarily include binary decision variables (0's and 1's), as commonly used in the literature. Definition of decision variables is substantially important in identifying more realistic, flexible, and more efficient conservation plans.

<u>P3.</u> Identifying near-optimal conservation plans is efficiently possible with aggregating the prior knowledge about the system and performance of the conservation practices in a multi criteria decision analysis framework.

<u>P4.</u> Changes in seasonal and long-term pattern of precipitation and temperature will be translated into substantial changes in hydrologic and water quality fluxes.

<u>P5.</u> Performance of conservation practices will be substantially different under the different projected climate scenarios. Thus, the optimal solutions of nonpoint source pollution control will be different, too.

1.2.3 Study Area: Eagle Creek Watershed, Indiana

The Eagle Creek Watershed (ECW) lies within the Upper White River Watershed of central Indiana, and drains into Eagle Creek Reservoir (Figure 1.1), which supplies drinking water to the City of Indianapolis. Much of ECW is poorly drained and the majority of the watershed likely contains tile-drain systems, which have been shown to be a significant pathway for pollutant loading to waterways in central Indiana [*Tedesco et al.*, 2005]. Elevated atrazine concentrations in ECW have resulted in atrazine levels in Eagle Creek Reservoir in excess of the EPA drinking water standard of 3 μg/L, and high sediment and nutrient loads have potential to degrade aquatic habitat [*Tedesco et al.*, 2005]. The importance of ECW for metropolitan water supplies and an extensive historic dataset make the basin well suited for examining watershed-scale effects of climate change on water quality fluxes.

Moreover, spatial scale of conservation planning depends upon numerous factors, including management objectives, available data resolution, dominant ecological processes, and potential sociopolitical constraints [*Walter et al.*, 2007; *Garen and Moore*, 2005]. To achieve the specific goals of water quality control, conservation practices targeting should be performed within a smaller geographic unit (12-digit
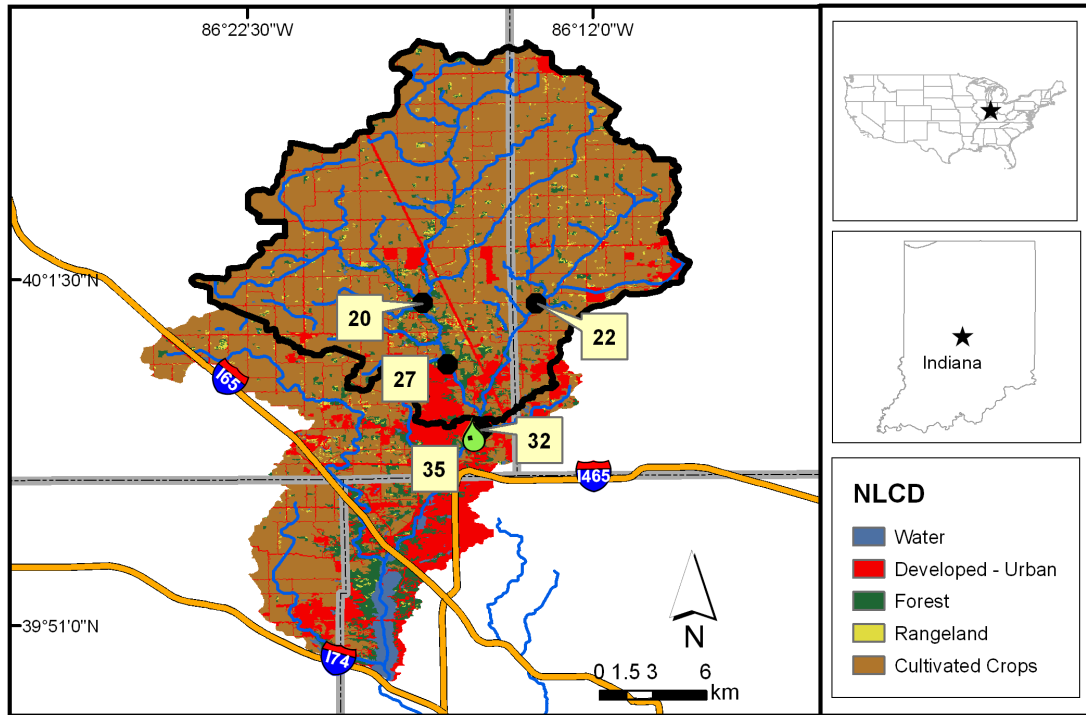
Figure 1.1: Location and landuse maps of the Eagle Creek Watershed

Hydrologic Unit Codes per se) which ultimately allows us to better evaluate targeted management plan.

1.3   Research Approach

Schematic of the proposed conservation plans analysis framework is presented in Figure 1.2. The framework consists of six main modules which are linked with a non-point source pollution modeling tool (Simulation): (i) a global sensitivity analysis tool; (ii) a parameter estimation tool (Calibration); (iii) an optimization framework for identifying optimal type and placement of conservation practices (Optimization); (iv) a conservation practices implementation tool (BMP Tool) including economic and environmental analysis component; and (v) a climate downscaling module; (vi) a multi criteria decision analysis framework (MCDA). The modules were developed and tested with MATLAB 2009b, 2010a, and 2011b and are expected to work with newer versions.
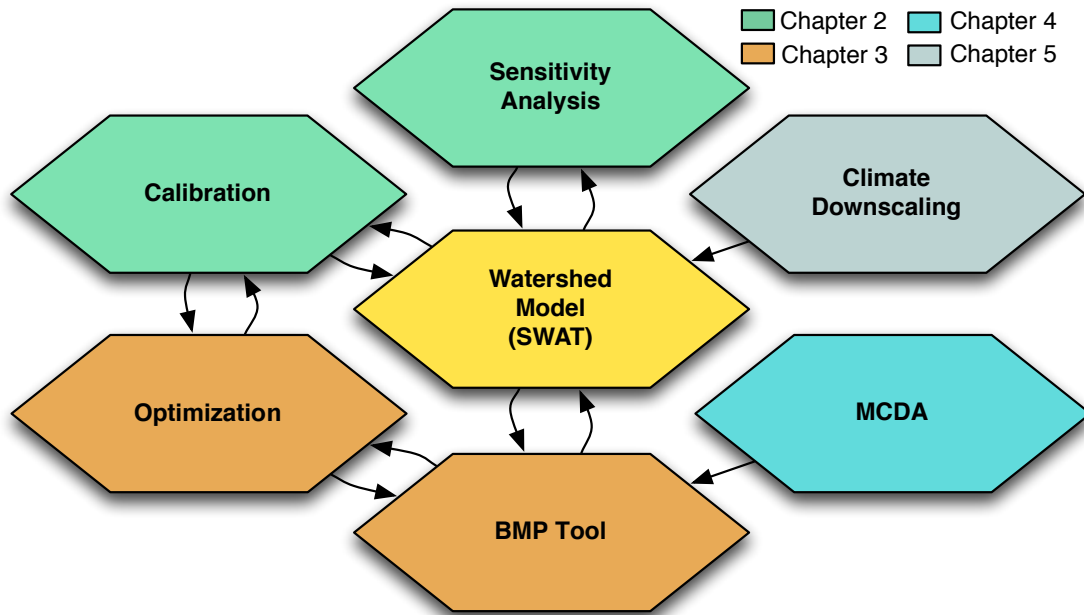
Figure 1.2: Schematic of the integrated simulation-optimization-MCDA framework

1.3.1 Hydrologic and Water Quality Modeling

Hydrologic and water quality modeling module consists of following three components:

1.3.1.1 Nonpoint Source Pollution Model

In order to fulfill the objectives of the study, the Soil and Water Assessment Tool (SWAT) [*Arnold et al.*, 1998] was used to represent hydrologic and water quality processes, analyze performance of nonpoint source pollution control plans, and assess hydrologic and water quality responses under climate projections. SWAT is a process based, distributed parameter, continuous time, and long-term watershed model that runs on a daily time step. It subdivides a watershed into subbasins connected by a stream network, and further delineates hydrologic response units (HRUs) consisting of unique combination of land cover and soils in each subbasin. SWAT can simulate

9

major nutrient processes and a selected pesticide, in addition to the hydrologic fluxes, within a watershed [*Neitsch et al.*, 2005].

1.3.1.2   Sensitivity Analysis

A global sensitivity analysis toolbox (GSATool) is developed for MATLAB with the support of commonly used global sensitivity analysismethods and factors probability distribution functions. The GSATool is an open source tool and can be modified for additional sensitivity analysis methods and factors pdf. The tool is tested and validated for several examples.

1.3.1.3   Parameter Estimation

Selected parameters from sensitivity analysis was used in a computational framework for incorporation of disparate information from observed hydrologic responses at multiple locations into the calibration of watershed models. The framework consists of four components: (i) an *a priori* characterization of system behavior; (ii) a formal and statistically correct formulation of objective function(s) of model errors; (iii) an optimization engine to determine the Pareto-optimal front for the selected objectives; and (iv) a multi criteria decision analysis tool to select optimal solutions from the Pareto-optimal front that are most consistent with the goals of the modeling study.

1.3.2   Conservation Practices Optimization Framework

To optimally locate conservation practices within the watershed a simulation–optimization framework was established. This framework consists of an (i) optimization method, (ii) a conservation practices implementation tool, and (iii) an economic/environmental component to calculated the objective function values.

1.3.2.1 Optimization Method

Global search methods are robust in finding optimal solutions by searching over the larger subset of the search space, and thereby escape being trapped in local optima. Nondominated Sorted Genetic Algorithm II (NSGA-II) was used as the primary search method in identifying optimal nonpoint source pollution control. A typical genetic algorithm starts with an initial population of solutions and then implements probabilistic and parallel exploration in the search space using the domain-independent genetic operators (i.e. chromosome reproduction) to find optimal solutions.

1.3.2.2 BMP Tool

A novelty of the proposed framework is, unlike the optimization-based conservation planning tools developed in other studies, its capability to operate on both binary-discrete and continuous decision variables (known as "mixed-chromosome" or "mixed-variable"). This will likely identify more realistic alternatives and higher flexibility in addition to exploring overall better fitness. It is also expected to achieve more diverse Pareto-optimal solutions through application of the mixed continuous-discrete optimization.

1.3.2.3 Decision Variables

A novelty of the proposed simulation–optimization framework is that, unlike the commonly used conservation practices targeting tools developed in other studies, it is capable of coping with both binary/discrete and continuous-decision variables.

1.3.3 Multi Criteria Decision Analysis

A multi criteria decision analysis (MCDA) framework is proposed, as a robust substitute for multiobjective optimization methods, in identifying optimal nonpoint source pollution control strategies at the watershed scale.

1.3.4 Climate Downscaling

The climate module is a statistical downscaling procedure to create projected regional climate scenarios (including minimum and maximum temperature and precipitation) at the meteorological stations and daily time intervals over the 2010-2099 period. The climate inputs were used as climate forcings to drive hydrologic response of the system.

1.4 Significance of this dissertation

This dissertation is significant in several aspects of the watershed management: (i) it presents a computational framework for incorporation of disparate information from observed hydrologic responses at multiple locations into the calibration of watershed models that are used for pollutant source identification and watershed management; (ii) it demonstrates importance of the availability of hydrologic and water quality data at multiple locations and highlights the use of multiple objective functions in identifying the most sensitive parameters of watershed models; (iii) it provides more realistic watershed scale conservation solutions with higher flexibility to decision makers by improving the commonly used simulation–optimization framework of conservation planning via reformulating the objective functions, upgrading the decision variables, and hybridizing optimization methods; (iv) it demonstrates a MCDA framework to identify cost-effective alternatives for nonpoint source pollution control in a watershed scale without any need for computationally intensive and iterative search algorithms; (v) it analyzes the hydrologic and water quality simulations under extensive ensemble of climate projections for their respective changes in basin average temperature and precipitation; (vi) it studies impact of climate change on performance of nonpoint source pollution control plans and shifts in their optimal type and placement in a watershed scale under different climate projections; and (vii) it presents a set of watershed modeling and management tools to facilitate decision making process for watershed scale conservation planning.

1.5   Organization of the Dissertation

This dissertation is organized into six chapters. The current chapter provides an introduction to the study problem. The second chapter presents a framework for multisite, many objective calibration of watershed models. The third chapter presents a modified simulation-optimization framework for optimal placement of conservation practices within a watershed. The fourth chapter demonstrates a MCDA framework for identifying near-optimal solutions for conservation planning. The fifth chapter assesses the impact of climate change on hydrologic and water quality fluxes within a watershed and analyzes the changes in performance of non-point source pollution control plans under projected climate. The sixth chapter provides a summary, conclusions, and recommendations for future work.

Bibliography

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

Arabi, M., R. S. Govindaraju, B. A. Engel, and M. M. Hantush (2007), Multiobjective sensitivity analysis of sediment and nitrogen processes with a watershed model, *Water Resources Research*, *43*(6), doi:10.1029/2006WR005463.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment, Part I: Model development, *Journal of the American Water Resources Association*, *34*(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x.

Easton, Z. M., M. T. Walter, and T. S. Steenhuis (2008), Combined monitoring and modeling indicate the most effective agricultural best management practices., *Journal of environmental quality*, *37*(5), 1798–809, doi:10.2134/jeq2007.0522.

Garen, D. C., and D. S. Moore (2005), Curve number hydrology in water quality modeling: uses, abuses, and future directions, *Journal of the American Water Resources Association*, *41*(2), 377–388, doi:10.1111/j.1752-1688.2005.tb03742.x.

Horowitz, J. K., L. Lynch, and A. Stocking (2007), Competition-Based Environmental Policy: An Analysis of Farmland Preservation in Maryland, *Tech. rep.*, University of Maryland, Collge Park, MD.

Jennings, E., N. Allott, D. C. Pierson, E. M. Schneiderman, D. Lenihan, P. Samuelsson, and D. Taylor (2009), Impacts of climate change on phosphorus loading from a grassland catchment: implications for future management., *Water research*, *43*(17), 4316–26, doi:10.1016/j.watres.2009.06.032.

Jha, M. K., S. Rabotyagov, and P. W. Gassman (2009), Optimal Placement of Conservation Practices Using Genetic Algorithm with SWAT, *International Agricultural Engineering Journal*, *18*(1-2), 41–50.

Kaini, P., J. W. Nicklow, and J. T. Schoof (2010), Impact of Climate Change Projections and Best Management Practices on River Flows and Sediment Load, in *World Environmental and Water Resources Congress 2010: Challenges of Change*, pp. 234–234, ASCE, doi:10.1061/41114(371)234.

Maringanti, C., I. Chaubey, M. Arabi, and B. A. Engel (2011), Application of a Multi-Objective Optimization Method to Provide Least Cost Alternatives for NPS Pollution Control, *Environmental management*, *48*(3), 448–61, doi:10.1007/s00267-011-9696-2.

Mostaghimi, S., S. Park, R. Cooke, and S. Wang (1997), Assessment of management alternatives on a small agricultural watershed, *Water Research*, *31*(8), 1867–1878, doi:10.1016/S0043-1354(97)00018-3.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams (2005), Soil and Water Assessment Tool (SWAT) Theoretical Documentation, Version 2005, (BRC Report 02-05).

Novotny, V. (1993), *Water Quality: Diffuse Pollution and Watershed Management*, 2 edition ed., 888 pp., Wiley.

Parker, G., R. Droste, and K. Kennedy (2008), Modeling the effect of agricultural best management practices on water quality under various climatic scenarios, *Journal of Environmental Engineering and Science*, *7*(1), 9–19, doi: 10.1139/S07-026.

Ribaudo, M. O., and R. Johansson (2006), AREI Chapter 2.2: Water QualityImpacts of Agriculture.

Ritter, W. F. (2001), Nonpoint Source Pollution and Livestock Manure Management, *Hydrology Nonpoint Source Pollution Watershed Management And*, pp. 135–168.

Tedesco, L., D. Pascual, L. Shrake, R. Hall, L. Casey, P. Vidon, F. Hernly, K. Salazar, and R. Barr (2005), Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality, *Tech. rep.*, IUPUI, Indianapolis.

US Congress (2002), Federal water pollution control act, As Amended Through P.L.

US EPA (1993), Guidance Specifying Management Measures for Sources of Nonpoint Pollution in Coastal Waters, EPA 840-B-92-002, *Tech. rep.*, U.S. Environmental Protection Agency, Office of Water.

US EPA (2008), Handbook for Developing Watershed Plans to Restore and Protect Our Waters, *Tech. rep.*, U.S. EPA, Office of Water, Washington, DC.

US EPA (2009), National Water Quality Inventory: 2009 Report to Congress, *Tech. rep.*, U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Veith, T. L. (2002), Agricultural Bmp Placement For Cost-Effective Pollution Control At The Watershed Level, Phd, Virginia Polytechnic Institute and State University.

Veith, T. L., M. L. Wolfe, and C. D. Heatwole (2003), Optimization procedure for cost effective BMP placement at a watershed scale, *Journal of the American Water Resources Association*, *39*(6), 1331–1343.

Walter, T., M. Dosskey, M. Khanna, J. Miller, M. Tomer, and J. Wiens (2007), The science of targeting within landscapes and watersheds to improve conservation effectiveness, in *Managing agricultural landscapes for environmental quality*, edited by S. M and C. C, Soil and Water Conservation Society, Ankeny, IA.

Chapter 2

TOWARD IMPROVED CALIBRATION OF WATERSHED MODELS: MULTISITE MANY OBJECTIVE MEASURES OF INFORMATION

Highlights

This paper presents a computational framework for incorporation of disparate information from observed hydrologic responses at multiple locations into the calibration of watershed models. The framework consists of four components: (i) an *a-priori* characterization of system behavior; (ii) a formal and statistically correct formulation of objective function(s) of model errors; (iii) an optimization engine to determine the Pareto-optimal front for the selected objectives; and (iv) a multi criteria decision analysis tool to select optimal solutions from the Pareto-optimal front that are most consistent with the goals of the modeling study. Application of the proposed framework for calibration of the Soil and Water Assessment Tool (SWAT) in the Eagle Creek watershed, Indiana, revealed that aggregation of streamflow and nitrate information measured at multiple locations within the watershed into a single measure of weighted errors would result in faster convergence to a solution with a lower overall objective function value than using multiple measures of information. However, none of the solutions from the single objective approach satisfied the conditions defined for characterizing the system behavior. In particular, aggregation of streamflow and nitrate responses undermined finding behavioral solutions for nitrate, primarily because of the substantially larger number of observations for streamflow. Aggregation of only nitrate responses into a single measure expedited

18

finding better solutions, although aggregation of data from nested sites appeared inappropriate because of correlated errors. This study demonstrates the importance of the availability of hydrologic and water quality data at multiple locations, and also highlights the use of multiobjective approaches for proper calibration of watershed models that are used for pollutant source identification and watershed management.

Keywords: Parameter estimation, hydrologic modeling, optimization, SWAT, water quality

2.1 Introduction

Watershed models are increasingly embedded in the decision making process to address a wide range of hydrologic and water quality issues. In the United States, federal law requires states to develop total maximum daily loads (TMDLs) for impaired water bodies to attain ambient water quality standards through the control of point and nonpoint sources [*National Research Council*, 2001]. Similarly, the European Water Framework Directive aims to enhance the water quality status of all water bodies within its jurisdiction [*Kaika*, 2003]. Simulation models of the environment play a central role in successful implementation of watershed management programs by providing the means to assess the relative contribution of different sources (i.e., stressors) to the impairment. Therefore, it is of keen interest to evaluate the performance validity of watershed models according to the past observations of fluxes of water and contaminants at multiple locations on the stream network.

Application of models that credibly represent important processes of the natural system presents a challenge [*Hantush et al.*, 2005; *Konikow and Bredehoeft*, 1992]. With the temptation to incorporate more parameters in models to represent a broader range of hydrologic and water quality processes has come an insidious effect: the ever-increasing complexity of model structures, which in turn gives rise to issues of lack of identifiability [*Beck*, 1987], lack of uniqueness [*Spear*, 1997], and equifinality [*Beven and Binley*, 1992]. Therefore, efficient and effective use of observed data is vital for calibration of complex spatially distributed process-based models. In the United States, daily or more frequent discharge measurements at the watershed outlets on many rivers and streams are available from the United States Geological Survey (USGS). On the other hand, nutrient concentrations are often measured by local watershed groups on less frequent (e.g., weekly or monthly) timesteps at the smaller subwatershed level. This is also the case for many other regions in the world. These hydrologic and water quality observations are characterized

by varying measurement errors/uncertainties, varying sample size, and are typically noncommensurable. These considerations must be taken into account when using the data in construction of the objective function(s) for calibration purposes.

The effectiveness of parameter estimation techniques depends greatly on the selection of a proper likelihood function [*Beven and Binley*, 1992; *Beven and Freer*, 2001; *Box and Tiao*, 1992; *Mantovan and Todini*, 2006; *Sorooshian and Dracup*, 1980; *van Griensven et al.*, 2008]. *Stedinger et al.* [2008] showed how the use of a formal Bayesian-based likelihood function can provide more acceptable and statistically valid prediction intervals for future observations. Similarly, *Vrugt et al.* [2009a] demonstrated that better coverage of observed data and more acceptable posterior distribution of parameters can be achieved when a formal likelihood function is used. *McMillan and Clark* [2009] also found that using formal likelihood measures provides a more complete exploration of the parameter space, a more accurate estimation of parametric uncertainty, and a better representation of the observed behavior of the system under study. However, in the context of watershed management, a statistically correct likelihood function may not exist [*Gupta et al.*, 1998], and subsequently, identification of a single set of model parameters that is optimal for all of the hydrologic and water quality variables of interest may be infeasible.

Model calibration at multiple sites and for many variables is inherently a multiobjective problem [*Gupta et al.*, 1998; *Madsen*, 2003; *Madsen et al.*, 2002]. Multiobjective optimization approaches enable the analyst to assess trade-offs associated with conflicting objectives. They determine a set of nondominated solutions that comprise the Pareto-optimal front. Any improvement in one objective among Pareto-optimal solutions will necessarily result in the degradation of at least one other objective [*Pareto*, 1971]. Strength Pareto Evolutionary Algorithm 2 (SPEA2) [*Zitzler et al.*, 2001] and Nondominated Sorted Genetic Algorithm II (NSGA-II)

[*Deb*, 2001] are among the most commonly used multiobjective approaches [*Bekele*, 2005]. Multi-algorithm methods, such as A Multi-ALgorithm Genetically Adaptive Multiobjective (AMALGAM), have also shown fast convergence to the Pareto front by combining attributes of the best available optimization algorithms [*Vrugt and Robinson*, 2007].

The complexity of multiobjective methods increases substantially with the increasing number of objectives in the optimization problem [*Brockhoff and Zitzler*, 2007]. Typically, these methods require more model simulations than single-objective techniques for convergence and are more difficult to implement [*Kollat and Reed*, 2007]. Therefore, an analyst may opt to use a single aggregated objective function of weighted errors [*Madsen*, 2000; *Seibert*, 2000; *van Griensven and Bauwens*, 2003], though this can lead to the loss of important information from observations [*Fenicia et al.*, 2008; *van Griensven and Bauwens*, 2003]. In doing so, the analyst may also relinquish the flexibility to identify Pareto-optimal solutions most consistent with his/her preferences. The Shuffled Complex Evolutionary (SCE) algorithm [*Duan et al.*, 1992] and its modifications such as the SCEM [*Vrugt*, 2003] and DREAM [*Vrugt et al.*, 2009a], Bayesian-based approaches [*Marshall*, 2004], Dynamically Dimensioned Search (DDS) algorithm [*Tolson and Shoemaker*, 2007], and Genetic Algorithms (GA) [*Holland*, 1975] are among the most popular single optimization methods for calibration of hydrological models. *Nicklow et al.* [2010] provided a comprehensive review of state-of-the-art evolutionary algorithms in water resources management.

An important and often neglected issue in calibration of complex models of the environment is that while optimization techniques facilitate the search for solutions with minimum errors, they do not necessarily ascertain model adequacy for mimicking the observed behavior of the system. When multiple hydrologic and water quality responses are involved, system behavior is defined according to the

model application goals [*Moriasi et al.*, 2007]. Literature is replete with studies that provide noncommensurable measures of performance for classification of model parameter sets to behavioral (i.e., good or acceptable) or non-behavioral (i.e., bad or unacceptable) solutions [*Beven and Binley*, 1992; *Blazkova and Beven*, 2009; *Gupta et al.*, 1998; *Klepper et al.*, 1991; *McMillan and Clark*, 2009; *Moriasi et al.*, 2007; *Nash and Sutcliffe*, 1970; *Spear and Hornberger*, 1980]. However, it is still unclear how these measures can be used in conjunction with Pareto-optimal solutions for selecting the optimal choice for the model parameters.

The primary goal of this study is to present a computational framework for multisite many objectives calibration of watershed models that can incorporate the priorities considered for model application. Two specific objectives are examined *en-route* to the overall goal of the study: (i) to examine the efficiency of single-objective and multiobjective approaches in finding optimal solutions while minimizing proper objective functions of hydrologic responses; and (ii) to evaluate the effectiveness of these methods in generating optimal solutions that are consistent with model application purposes. The paper also presents a discussion of benefits and limitations of procedures for aggregation of hydrologic and water quality objectives corresponding to different observation sites within the watersheds.

2.2 Methods and materials

A computational framework was utilized to reconcile the strengths of optimization algorithms with the flexibility provided by multi criteria decision analysis (MCDA) for multisite many-objectives calibration of watershed models. The framework encompasses four major components: (1) a quantitative definition of the system behavior that captures the priorities of the analyst according to the explicit modeling purpose, (2) a proper formulation of objective functions that reflects the model error structure for all objectives considered in the study, (3) an efficient multiobjective optimization algorithm to expose trade-offs amongst conflicting hydrologic

objectives, and (4) a multi criteria decision analysis (MCDA) to select alternative parameter sets consistent with the model application purposes reflected in the definition of system behavior. A case study is presented to demonstrate the application of the proposed computational framework for calibration of the Soil and Water Assessment Tool (SWAT) [*Neitsch et al.*, 2005] for hydrologic and water quality modeling in the Eagle Creek Watershed, Indiana, USA.

2.2.1 Behavioral solutions

Application of behavior/non-behavior classification in environmental modeling was first examined by *Spear and Hornberger* [1980]. Behavioral solutions of a model comprise a subset of conceptually plausible responses that are judged by the analyst to be satisfactory according to past observations of the system under study [*Schaefli et al.*, 2011]. This requires selection of performance criteria and behavioral thresholds such that the solution is rejected when model responses fall outside acceptable ranges or thresholds. Acceptance thresholds are subjectively chosen based on the experts' judgment according to model application goals and objectives. *Moriasi et al.* [2007] and *Rossi et al.* [2008] established guidelines for evaluating the performance of watershed models. Thresholds were defined based on dimensionless model evaluation statistics including Nash-Sutcliffe model efficiency coefficient (NSE) [*Nash and Sutcliffe*, 1970] and percent bias (PBIAS). The NSE measure determines the relative magnitude of the residual variance compared to the observed data variance. It indicates how well the plot of observed versus simulated values fits the 1:1 line. The coefficient can range from $-\infty$ to a perfect match of $+1$:

$$NSE = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{2.1}$$

where $\hat{y}_i$ and y_i denote, respectively, the simulated and observed responses for time-step i, n is the number of observations, and $\bar{y}$ denotes the mean of observed responses. NSE has been recommended for use by the American Society of Civil

Engineers (ASCE) [*ASCE*, 1993] and is a commonly used statistical measure for calibration of watershed models. *Servat and Dezetter* [1991] also found NSE to be the best objective function for reflecting the overall fit of a hydrograph.

PBIAS is a measure of the average tendency of simulated values to be larger or smaller than corresponding observed values and is obtained by summation of the ratios of absolute error to observed value at each time-step:

$$PBIAS = 100 \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{y_i} \tag{2.2}$$

PBIAS is also recommended by ASCE and has the ability to clearly indicate poor model performance [*Gupta et al.*, 1999]. PBIAS is a particularly useful measure for evaluation of model performance for nutrients, pesticides, and other contaminants that are addressed in watershed management programs such as TMDLs based on the average annual response [*Moriasi et al.*, 2007].

2.2.2 Likelihood function derivation

Let's consider a watershed model M with a vector of p parameters ($\boldsymbol{\theta}$) within the feasible parameter space ($\boldsymbol{\Theta}$) that simulates the response vector of the watershed ($\hat{y}$) as follows:

$$\hat{y} = M(\boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^n \tag{2.3}$$

The discrete stochastic time-series vector of model residuals is:

$$\mathcal{E}(\boldsymbol{\theta}) = \hat{y} - y_i = M(\boldsymbol{\theta}) - y \tag{2.4}$$

where y is the vector of observed (measured) output response. The goal of calibration procedures is to estimate θ such that the residuals are as close to zero as possible.

Bayesian methods are commonly used for parameter estimation and uncertainty analysis of hydrological and water quality models [*Stedinger et al.*, 2008; *Vrugt et al.*,

2006, 2009b]. Bayesian statistics consider model parameters as stochastic components, and the joint posterior probability distribution of the parameters conditioned on the observed response is expressed as [*Box and Tiao*, 1992]:

$$P(\boldsymbol{\theta}|y) = c \cdot P(\boldsymbol{\theta}) \cdot \ell(\mathcal{E}|\boldsymbol{\theta}) \tag{2.5}$$

where c is a normalization constant, ℓ is called the likelihood function and represents the likelihood of producing model residuals ($\mathcal{E}$) for a given set of model parameters (θ), while θ denotes the prior probability density function of θ that is assigned before assimilation of any observed data. Assuming that residuals are normally and independently distributed (NID) with mean equal to zero and unknown but constant standard deviation σ_e, the likelihood function l will then take the following form [*Box and Tiao*, 1992]:

$$\ell(\mathcal{E}|\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_e^2}} exp\left[-\frac{(\hat{y}_i(\boldsymbol{\theta}) - y_i)^2}{2\sigma_e^2}\right] \tag{2.6}$$

where n is the number of observations. Since monotone transformations do not affect maxima, the same parameter values that maximize the likelihood function also maximize the logarithm of likelihood function, referred to as log-likelihood function (ℓ^*):

$$\ell^*(\mathcal{E}|\boldsymbol{\theta}) = -\frac{n}{2}ln(2\pi) - \frac{1}{2}ln\sigma_e^{2n} - \frac{1}{2}\sigma_e^{-2} \cdot \sum_{i=1}^{n} [\hat{y}_i(\boldsymbol{\theta}) - y_i]^2 \tag{2.7}$$

Since in most cases the errors in hydrological and water quality modeling are not normally distributed, independent, and homoscedastic, suitable transformations must be applied to account for error characteristics that are not consistent with assumptions made for deriving Eq. (2.7). The first-order autoregressive (AR-1) transformation of the residuals can be used to account for correlated errors:

$$\mathcal{E}_i = \rho\mathcal{E}_{i-1} + \nu_i, \quad i = 1, ..., n \tag{2.8}$$

where ρ is the lag-1 serial correlation coefficient for the residuals $\mathcal{E}$, and $\nu \sim N(0, \sigma_\nu^2)$ is the innovation term with zero mean and constant variance ν^2. Consequently, the corrected time series of residuals after the AR-1 transformation will take the following form:

$$\delta(\boldsymbol{\theta}, \rho|y) = \mathcal{E}_i(\boldsymbol{\theta}) - \rho\mathcal{E}_{i-1}(\boldsymbol{\theta}) \tag{2.9}$$

Sorooshian and Dracup [1980] showed that substituting the AR-1 transformation into the log-likelihood function of Eq. (2.7) yields:

$$\begin{aligned}
\ell^*(\mathcal{E}|\boldsymbol{\theta}) = &-\frac{n}{2}ln(2\pi) - \frac{1}{2}ln\frac{\sigma_\nu^{2n}}{1-\rho^2} - \frac{1}{2}(1-\rho^2) \cdot \sigma_\nu^{-2}[\hat{y}_1(\boldsymbol{\theta}) - y_1]^2 \\
&- \frac{1}{2}\sigma_\nu^{-2} \cdot \sum_{i=2}^{n} \{(y_i - \rho y_{i-1}) - (\hat{y}_i(\boldsymbol{\theta}) - \hat{y}_{i-1}(\boldsymbol{\theta}))\}^2
\end{aligned} \tag{2.10}$$

The terms σ_ν^2 and ρ can be estimated using the Bayesian approach [*Vrugt et al.*, 2009a] or can be assigned based on prior knowledge.

In the case that residuals do not have a stable variance, one can use a suitable transformation of the residuals [*Box et al.*, 2008; *Kuczera and Mroczkowski*, 1998; *Stedinger et al.*, 2008] or different objective functions for different portions of the hydrologic record. The transformation proposed by *Box and Cox* [1964] and alternate power transformations have been commonly used in hydrologic modeling. The extended form of the Box-Cox transformation of the simulated and observed hydrologic fluxes takes the following form [*Yeo and Johnson*, 2000]:

$$\tau(y) = \begin{cases} \frac{(y+\lambda_2)^{\lambda_1}-1}{\lambda_1}, & if \quad \lambda_1 \neq 0 \\ log(y + \lambda_2), & if \quad \lambda_1 = 0 \end{cases} \tag{2.11}$$

where λ_1 and λ_2 are transformation parameters. λ_2 should be chosen such that $y + \lambda_2 > 0$ (or $\lambda_2 > -y$) and λ_1 can be estimated using maximum likelihood function. The likelihood function is constructed based on the assumption that the transformed data, $\tau(y)$, are normally distributed and then this function is maximized with respect to the unknown value (λ_1).

2.2.3 Multisite many objective optimization algorithms

The multisite many objectives calibration problem can be stated as an optimization problem as follows [*Gupta et al.*, 1998]:

$$\text{minimize } F(\boldsymbol{\theta}) = \{F_1(\boldsymbol{\theta}), ..., F_m(\boldsymbol{\theta})\}, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^n \qquad (2.12)$$

where m is total number of objective functions considering all hydrologic and water quality variables of interest at all sampling sites. Each of the objectives in Eq. (2.12) can be formulated using the negative of the log-likelihood function of Eq. (2.10) after applying proper transformations of responses. Two general approaches can be taken to search for the choice of optimal parameter values: a single objective approach or a truly multiobjective search method. Single-objective parameter estimation techniques require fewer model evaluations to find the minimum value of the aggregated objective function of the weighted errors. The efficiency of these approaches is an important consideration, particularly for complex watershed models with long computational time for each model evaluation. However, they require aggregation of the objectives using a statistically coherent scheme. Several methods are available in the literature for aggregating objective functions, including minimization of the total sum of squared residuals [*Madsen*, 2003; *van Griensven and Bauwens*, 2003]. Other methods use statistically-coherent aggregation techniques such as Bayesian statistics [*Ajami et al.*, 2007; *Stedinger et al.*, 2008; *van Griensven and Meixner*, 2007].

Assuming the NID residuals for all variables at all observation sites, *van Griensven and Meixner* [2007] showed that the likelihood function 2.6 can be stated as:

$$\ell(\mathcal{E}|\boldsymbol{\theta}) = \prod_{j=1}^{m}\prod_{i=1}^{n_j} \frac{1}{\sqrt{2\pi\sigma_{e,j}^2}} exp\left[-\frac{(\hat{y}_{i,j}(\boldsymbol{\theta}) - y_{i,j})^2}{2\sigma_{e,j}^2}\right] \qquad (2.13)$$

where n_j denotes the length of observed data for output variable j. The variance of residuals for the objectives may be obtained from previous experience or a close

investigation of the residuals. Taking the logarithms of Eq. (2.13) and applying the AR-1 transformation scheme of Eq. (2.9), a proper likelihood function for multiple outputs similar to the single output case of Eq. (2.10) can be derived as follows:

$$
\ell^*(\mathcal{E}|\boldsymbol{\theta}) = \sum_{j=1}^{m} \left\{ -\frac{n_j}{2} ln(2\pi) - \frac{1}{2} ln \frac{\sigma_{\nu,j}^{2n_j}}{1-\rho_j^2} - \frac{1}{2}(1-\rho_j^2) \cdot \sigma_{\nu,j}^{-2}[\hat{y}_{1,j}(\boldsymbol{\theta}) - y_{1,j}]^2 \right.
$$
$$
\left. -\frac{1}{2}\sigma_{\nu,j}^{-2} \cdot \sum_{i=2}^{n_j} \left\{ (y_{i,j} - \rho y_{i-1,j}) - (\hat{y}_{i,j}(\boldsymbol{\theta}) - \hat{y}_{i-1,j}(\boldsymbol{\theta}))^2 \right\} \tag{2.14}
$$

On the other hand, a truly multiobjective search algorithm does not require aggregation of objectives [*Deb et al.*, 2002; *Zitzler and Thiele*, 1999]. They can be used to simultaneously minimize two or more conflicting objectives, resulting in determination of a set of nondominated solutions that comprise the Pareto-optimal front. Consequently, the analyst is faced with choosing the best alternative from a set of nondominated solutions [*Khu and Madsen*, 2005]. MCDA methods provide a computational framework to rank or group alternatives that are most consistent with model application purposes.

2.2.4 Multi Criteria Decision Analysis

Multi criteria decision analysis (MCDA) is a numerical procedure to compare or score alternatives on a comparable scale [*Figueira et al.*, 2004; *Jacquet-Lagrèze*, 2001]. Typically, a MCDA method aims at one of the following goals: (i) choice: to find the best alternative, (ii) sorting: to group the alternatives into well-defined categories, (iii) ranking: to rank the alternatives in order of total preference, and (iv) description: to describe alternatives in terms of their performance on meeting the predefined criteria [*Jacquet-Lagrèze*, 2001; *Triantaphyllou and Baig*, 2005]. Combination of a MCDA method and behavioral thresholds can be employed to sort the optimization solutions in groups based on the well-defined criteria. For example, suggested error statistics and thresholds from *Moriasi et al.* [2007] can be used to classify model simulations into "satisfactory", "good", and "very good" for each objective included in the Pareto optimal surface.

2.3 Case Study

The proposed computational framework was used for calibration of the Soil and Water Assessment Tool (SWAT) in the Eagle Creek Watershed, Indiana (Figure 2.1). Two single objective methods (SCE and DDS) and one multiobjective optimization method (NSGA-II) were employed to determine optimal choices of SWAT parameter values for modeling fluxes of water and nitrate at five locations within the watershed. The SCE algorithm was selected because of its numerous successful applications for parameter estimation of complex watershed models such as SWAT (e.g. *Gassman et al.* [2007]; *Sharma et al.* [2006]; *Zhang et al.* [2009]). The performance of the DDS algorithm was also evaluated for its demonstrated efficiency and fast convergence for calibration purposes [*Tolson and Shoemaker*, 2007]. Moreover, the NSGA-II was used as a truly multiobjective approach.

2.3.1 Study area and data availability

The Eagle Creek Watershed (ECW), located in central Indiana, has a drainage area of 248.1 km^2 and lies within the Upper White River Watershed, extending from 40°01'24" to 40°04'16" north latitudes and 86°15'43" to 86°16'45" west longitudes. The watershed consists of 52% croplands, 27% pasture, 12% low and high density urban areas, and 9% forest. The predominant crops are corn and soybeans. ECW drains into Eagle Creek Reservoir, which supplies drinking water for the city of Indianapolis. Figure 2.1 presents the location and land cover for the watershed. The soils are generally poorly draining and developed from glacial materials with thin loess over loamy glacial till and alluvial materials depositions. The dominant soils are the Crosby-Treaty-Miami in the headwaters and Miami-Crosby-Treaty along the downstream areas. The mean annual precipitation for the Eagle Creek Watershed area is 1052 mm. Monthly mean temperatures for this area from 1971-2000 show
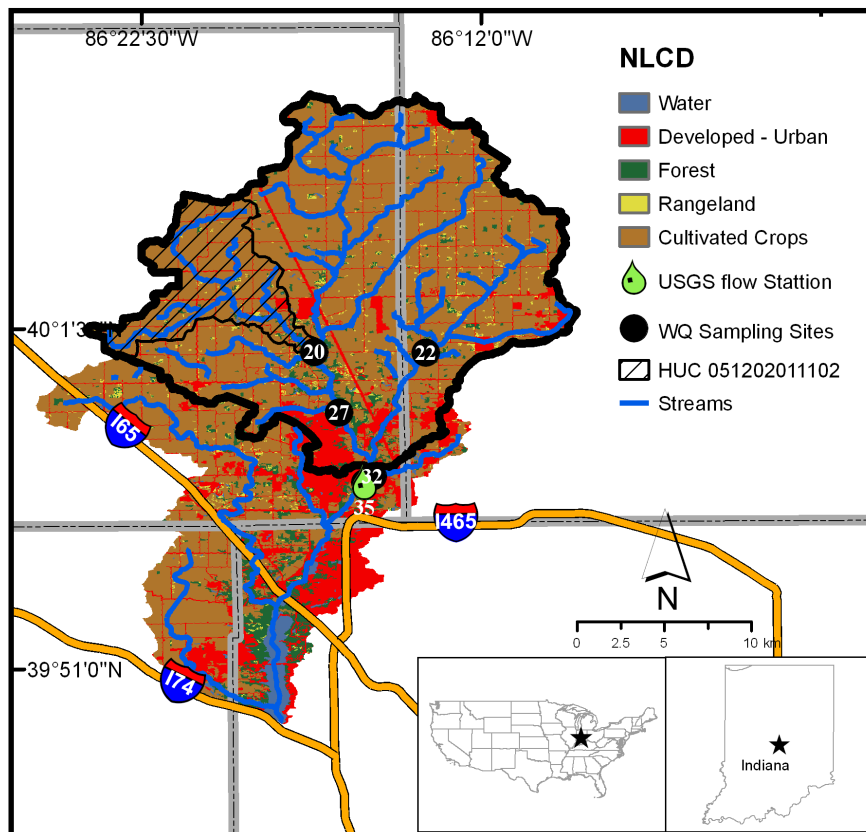
Figure 2.1: Eagle Creek Watershed Map

Table 2.1: Performance measure of model evaluation (adopted from *Moriasi et al.* [2007]

Performance Rating	Percent Bias (PBIAS-%)		Nash-Sutcliffe Efficiency (NSE)	
	Daily Streamflow	Monthly NOx	Daily Streamflow	Monthly NOx
Satisfactory	< 20	< 25	> 0.60	> 0.45
Good	< 15	< 20	> 0.65	> 0.50
Very Good	< 10	< 15	> 0.70	> 0.55

January as having the lowest average temperature of -3.3°C and July as the being the warmest month with an average temperature of 23.7°C [*Tedesco et al.*, 2005].

Daily streamflow data was available at the watershed outlet (outlet 35 in Figure 2.1, USGS Gauge # 03353200) from a USGS gauging station. Instantaneous samples of nitrate (NO_3) and nitrite (NO_2) were available at multiple sites on bi-weekly and monthly basis from Indiana Department of Environmental Management (IDEM). NO_2 and NO_3 (NO_2+NO_3, referred to as NOx hereafter) data from four sites (outlets 20, 22, 27, and 32 as shown in Figure 2.1) were used in this study. Monthly NOx loads were estimated from concentration data using the LOADEST program [*Runkel et al.*, 2004]. Behavioral classifications of "satisfactory", "good", and "very good" are identified by a range of values for statistical measures of PBIAS and NSE in Table 2.1, having stricter values for "very good" ratings.

2.3.2 Watershed model description: SWAT

The Soil and Water Assessment Tool (SWAT) [*Arnold et al.*, 1998] was used to represent hydrologic and water quality processes in the Eagle Creek watershed. SWAT is a process based, distributed parameter, continuous time, and long-term watershed model that runs on a daily time step. It subdivides a watershed into subbasins connected by a stream network, and further delineates hydrologic response units (HRUs) consisting of unique combinations of land cover and soils in

each subbasin. Watershed processes simulated by SWAT include snow accumulation and melt, evapotranspiration, infiltration, percolation losses, surface run-off, and groundwater flows [*Neitsch et al.*, 2005]. SWAT can simulate major nutrient processes within a watershed. The nitrogen (N) cycle is simulated in five pools: inorganic (including ammonium and nitrate) and organic (including fresh, stable, and active). The main N processes are mineralization, decomposition, and immobilization. Nutrients are introduced into the main channel through surface runoff and lateral flow and transported downstream with channel flow. Plant uptake, denitrification, volatilization, leaching, and soil erosion are the major mechanisms of N removal from a field. The transport rate of organic N with sediment is calculated with a loading function developed by *McElroy et al.* [1976] and modified by *Williams and Hann* [1978] for application to individual runoff events. The loading function estimates daily organic N runoff loss based on the concentrations of constituents in the topsoil layer, sediment yield, and an enrichment ratio. Nutrient transformations in the stream are controlled by the in-stream water quality component of the model that is adapted from the QUAL2E in-stream water quality model [*Brown and Barnwell*, 1987]. More detailed description of the nutrient components of SWAT can be found in *Neitsch et al.* [2005].

A 30-m resolution DEM from USGS National Elevation Dataset [*USGS NED*, 2010], National Land Cover Dataset (NLCD) 1992 and 2001 [*USGS NLCD*, 2001] for urban areas, National Agriculture Statistics Service (NASS) Cropland Data Layer 2000-2003 [*USDA NASS*, 2003] for croplands, and SSURGO data from national resources conservation service (NRCS) [*USDA NRCS*, 2010] were used for watershed subdivision and delineating HRUs in the SWAT model. The ECW was subdivided into 35 subwatersheds and a total of 446 hydrologic HRUs.

SWAT has many input parameters used for predicting water quality and quantity output responses. While automatic optimization techniques can construct pre-

Table 2.2: Selected SWAT parameters for streamflow and nitrogen processes

Symbol	Description	Min	Max
ALPHA_BF	Base flow recession coefficient, days	0	1
CANMX	Maximum canopy storage, mm	0	10
CDN	Denitrification exponential rate coefficient	0	0.3
CH_K(1)	Tributary effective hydraulic conductivity, mm/hr	0	300
CH_K(2)	Channel Effective hydraulic conductivity, mm/hr	-0.01	500
CH_N(1)	Tributary channels' Manning's "n"	0.008	0.3
CH_N(2)	Main channel's Manning's "n"	0.01	0.3
CMN	Mineralization rate of organic nutrients	0.001	0.003
CN_F	Curve number, %	-10	10
ESCO	Soil Evaporation Coefficient	0.01	1
ORGN	Initial organic N concentration, kg-N/ha	1	10000
SFTMP	Snowfall temperature, °C	-5	5
SMFMN	Snow melt factor, mm/°C-day	0	10
SMTMP	Snow melt base temperature, °C	-5	5
SNO50COV	Snow volume fraction for 50% snow cover, mm	0	0.9
SNOCOVMX	Snow water content for 100% snow cover, mm H_2O	0	650
SOL_AWC	Available water capacity, %	-10	20
SURLAG	Surface lag, day	1	12

dictive models with high accuracy by searching within the high-dimensional parameters space, it is still of interest in many applications to reduce the dimension of the parameters prior to optimization [*Fodor*, 2002]. A two-step sensitivity analysis (SA) was used to select the most important input parameters. First, a local screening SA method was employed on 118 SWAT input parameters affecting hydrology and nitrogen processes to eliminate parameters with no influence on the selected responses (stream discharge and nitrate). As a result, 34 parameters were selected. Then, the more computationally expensive global SA method of Sobol' [*Saltelli and Sobol'*, 1993] was used to select influential model parameters listed in Table 2.2. Selected important parameters are in accordance with results of other studies (e.g. *Ullrich and Volk* [2010]; *Zhang et al.* [2009]).

2.3.3 Available data and computational setup

Daily streamflow data at the watershed outlet (site 35) and monthly Nitrate and Nitrite loads at four sites (sites 20, 22, 27 and 32) within the watershed were used in this study. The sample size was 2922 for daily stream discharge in cubic meter per second at the USGS site (site 35) and 45 for monthly NOx loads in tons/ha at each of the other sites. The SCE and DDS techniques were implemented using the single aggregated objective function of Eq. (2.14) according to streamflow response at site 35 and NOx responses at sites 20, 22, 27 and 32. The NSGA-II method was implemented using two strategies: (1) five independent objectives, including one streamflow objective at site 35 and four NOx objectives at sites 20, 22, 27, and 32; and (2) two objectives, including one streamflow objective at site 35 and one aggregated NOx responses for sites 20, 22, 27, and 32 using the likelihood function of Eq. (2.14). Simulations were performed for a period of 11 years from 1993 through 2003, including three years as the warm-up period. For consistency, the maximum number of function evaluations for all optimization methods were set to 5,000 as one of the termination conditions. Population size for NSGA-II was set to 50.

2.3.4 Results and discussion

Overall, the results of this study suggest that single objective approaches (SCE and DDS) are more efficient in finding a lower value for the aggregated single objective function of weighted errors than the multiobjective approach (NSGA-II). However, multiobjective approaches are more likely to find solutions that are consistent with the observed behavior of the system for all objectives of interest at all sites within the watershed. In particular, the results indicated that the use of multiobjective optimization methods in conjunctions with a proper MCDA procedure is favorable when stricter definition of system behavior is considered.
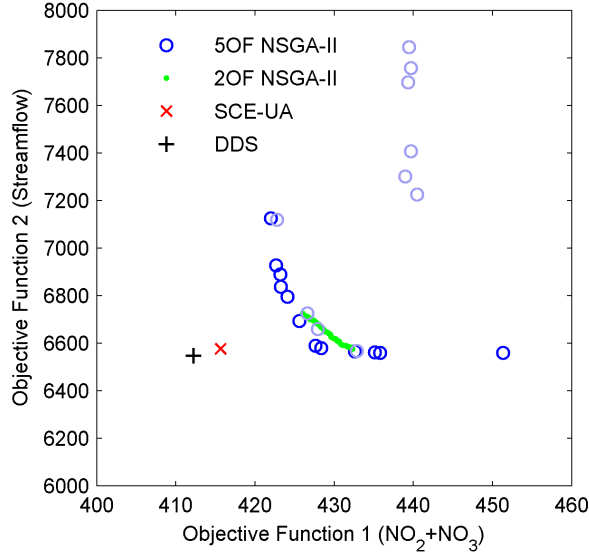
Figure 2.2: Optimal solution from different optimization techniques in a two-dimensional response space

Figure 2.2 depicts the final optimal solutions from all optimization techniques mapped into a two-dimensional space. The x-axis shows the global objective function for NOx (Eq. 2.14) for four outlets and the y-axis shows the streamflow objective function. The results demonstrated that the single-objective optimization techniques outperform the multiobjective techniques in finding solutions that have lower overall objective functions values. Furthermore, DDS exhibits greater efficiency and faster convergence to the optimal solution than the multiobjective NSGA-II method. Figure 2.3 illustrates the efficiency and convergence of optimization techniques to the aggregated objective function. It shows that the fewer the number of objective functions, the faster the convergence to the optimal solution for the aggregated objective function, this is more clear from the first 2,000 model simulations. DDS has the fastest convergence rate with five-objective NSGAII (5OF NSGA-II) converging with the slowest rate. This is in agreement with other studies (e.g. [*Chen et al.*, 2002]) that found the "curse of dimensionality" may decrease the convergence rate of the optimization algorithms.

36
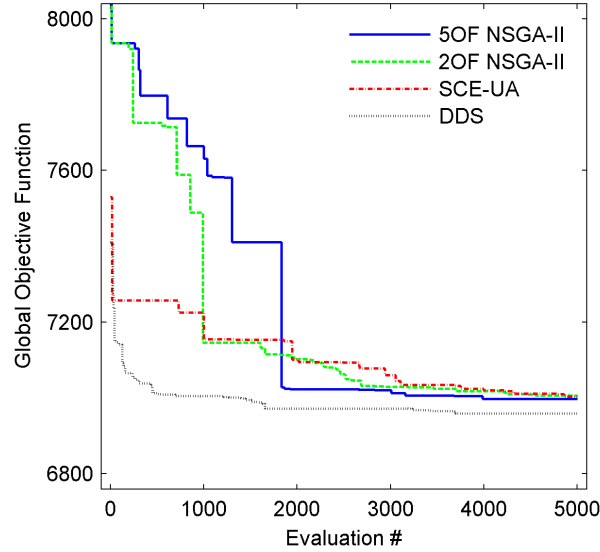
Figure 2.3: Efficiency and Convergence rate of different optimization techniques. The y-axis is the aggregated objective function of weighted errors of Eq. (2.14)

Efficiency is an important consideration for parameter estimation of complex environmental models. However, effectiveness in finding solutions that provide an acceptable representation of natural systems for all objectives and sites is of keen interest. Here, we used the performance measures presented in Table 2.1 to classify optimal solutions from all methods (including Pareto-optimal parameter sets) to satisfactory, good, and very good solutions (Figure 2.4). None of the single objective optimal solutions were behavior-giving (i.e. behavioral), while all of the Pareto solutions from two-objective NSGA-II (2OF NSGA-II) and 33% of the Pareto solutions from 5OF NSGA-II were within the "satisfactory" behavioral region (Figure 2.4-a).

The 5OF NSGA-II was the only method that included behavioral solutions when stricter behavior definitions were used. For the "good" behavior definitions for all output variables at all sites, no solutions were obtained from the single-objective methods (Figure 2.4-b). With "very good" behavioral thresholds for streamflow

37
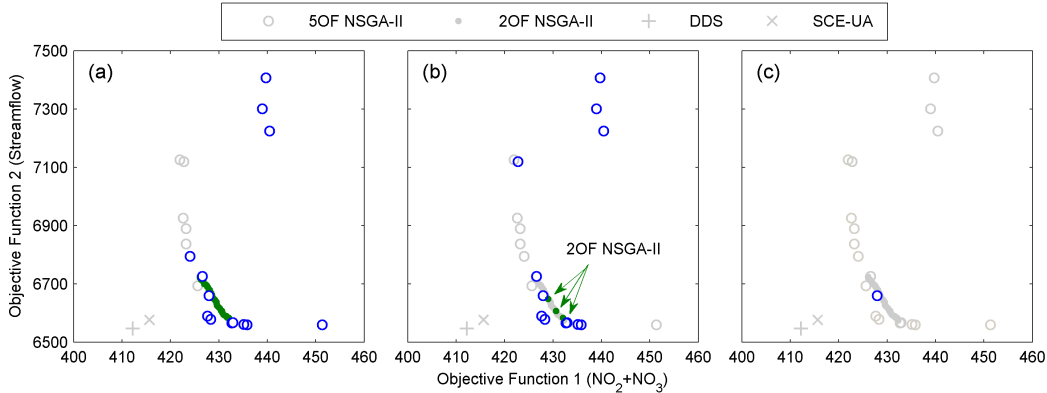
Figure 2.4: Behavioral Pareto solutions for different performance levels: (a) both streamflow and nitrate satisfactory; (b) both streamflow and nitrate good; and (c) very good streamflow and satisfactory nitrate. Non-behavior Pareto solutions are depicted in light gray

and "satisfactory" behavioral thresholds for NOx, the 5OF NSGA-II was the only method that included a behavioral Pareto-optimal solution (Figure 2.4-c).

The results of this study indicate that a low fraction of the behavioral solutions from multiobjective methods belong to the Pareto-optimal front. Therefore, it may be necessary to examine dominated (i.e., sub-optimal) solutions to identify solutions that are consistent with the observed responses at multiple sites. Table 2.3 presents the percentage of solutions that were classified as behavioral according to the performance measures in Table 2.1. Figure 2.5 depicts the regions encompassing streamflow and NOx objective functions corresponding to the "satisfactory" performance level. In this case, the size of behavioral response region decreased as the number of objective functions decreased, with SCE having the smallest region. 2OF NSGA-II encompassed the highest number of satisfactory simulations (i.e. 80%) while other algorithms produced less than 15% satisfactory simulations. This can be explained by the fact that a trade-off exists between number of objectives in the optimization problem and convergence rate of the algorithms. 5OF NSGA-II had fewer "satisfactory" solutions than the 2OF NSGA-II because of the

Table 2.3: Percentage (%) of behavioral solutions obtained for each algorithm

Method	Performance Level		
	Satisfactory	Good	Very Good
DDS	12.6	0.02	0
SCE	12	0	0
2OF NSGA-II	80	6.3	0
5OF NSGA-II	10.5	1.3	0.1

slower convergence rate towards the final optimal solutions. 2OF NSGA-II found the Parteo-optimal front in a fewer number of model evaluations, and thus more satisfactory solutions were identified. However, 5OF NSGA-II was the only method that identified "very good" behavioral solutions (Table 2.3). On the other hand, aggregation of the objective function values in single objective optimization methods resulted in domination of the streamflow objective due to its substantially higher number of observations and failure to meet the satisfactory thresholds of NOx responses. This can be also noticed from clustering of behavioral simulations around the minimum of objective function 2 (corresponding to the streamflow objective) and extension of the behavioral response region box along the objective function 1 in Figure 2.5.

The Pareto-optimal front from NSGA-II provided several nondominated parameter sets that can be selected as the choice of optimal solution. *Gupta et al.* [1998] suggested that Pareto-optimal solutions can be used to investigate the uncertainty of model parameters. Figure 2.6 shows the uncertainty in parameter values from Pareto front solutions. The y-axis reflects parameter values normalized to the initial parameter ranges listed in Table 2.2, i.e., $\boldsymbol{\theta}_{NORM} = (\boldsymbol{\theta} - \boldsymbol{\theta}_L)/(\boldsymbol{\theta}_U - \boldsymbol{\theta}_L)$. In the case of 2OF NSGA-II, the parameter values exhibit negligible uncertainties except for two parameters: main channels Mannings roughness coefficient (CH-N2) and mineralization rate of organic nutrients (CMN). This indicates that the variations in the two objective functions examined in this case could be primarily apportioned to
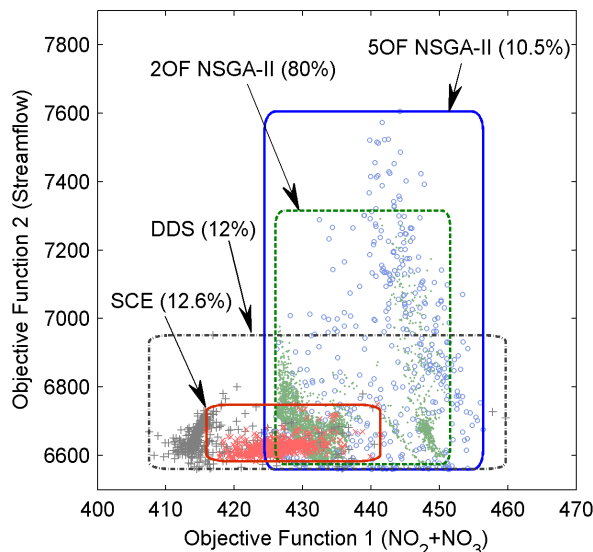
Figure 2.5: The range of objective function values corresponding to behavioral solutions based on the "satisfactory" performance level for both streamflow and nitrate. Percentages in the parenthesis are the percentage of behavioral solutions obtained from each algorithm

the uncertainty in the values of these two parameters. On the other hand, a larger number of parameters influenced the uncertainties in the five objectives examined in the 5OF NSGA-II case. Clearly, aggregating information from multiple stations into one objective underestimates the uncertainty in the parameters.

It should be also noted that given current developments of parallel and distributed computing technologies, overall runtime can be reduced substantially for optimization methods that can utilize individual parallel processing units. For example, NSGA-II is inherently well suited for parallelization within each generation to accelerate the convergence to the Pareto-optimal front. In addition, simulations using the SCE optimization method can be parallelized for each individual complex at a time. Unfortunately, similar to Markov Chain Monte Carlo methods, a new DDS run in a chain cannot be initiated before completion of the previous simulation. For this reason NSGA-II and SCE optimization runs were 6-7 times faster than DDS for an 8-core CPU.

40
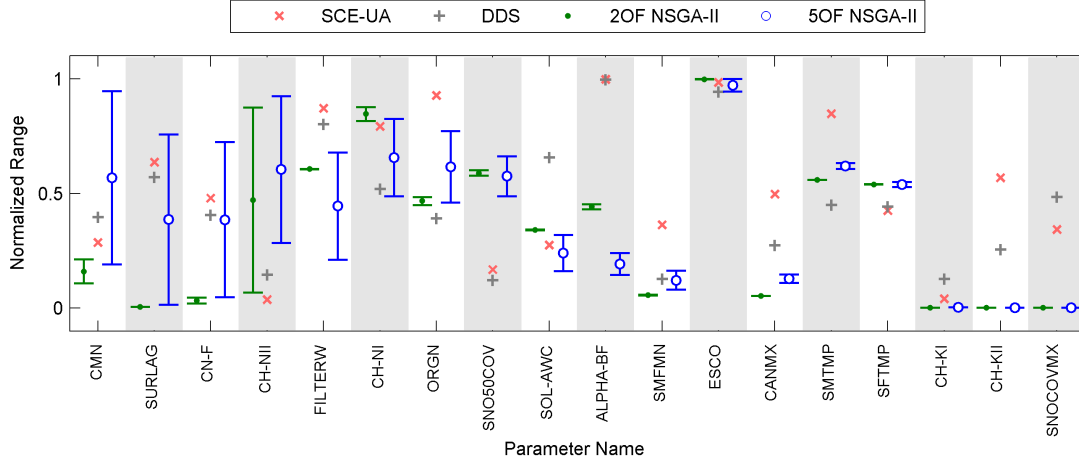
Figure 2.6: Uncertainty of model parameters obtained from solutions on the Pareto-optimal front

Table 2.4: Transformation parameter values

Variable	Outlet	Correlation Coefficient (ρ)	σ_ν	Box-Cox transformation parameter (λ_1)
Streamflow	35	0.89	0.48 (m^3/s)	0.06
Total Nitrate (NO$_2$+NO$_3$)	20	0.35	0.98 (kg/ha-month)	
	22	0.14	0.87 (kg/ha-month)	
	27	0.26	0.87 (kg/ha-month)	0.12
	32	0.26	0.82 (kg/ha-month)	

A careful analysis of the transformed residuals showed that the consecutive employment of the Box-Cox transformation and the AR-1 model was useful for satisfying the assumptions made for the derivation of Eq. (2.14). The Box-Cox transformation stabilized the variance of the residuals and reduced the heteroscedasticity, while the AR-1 transformation removed the autocorrelation in the model errors. Figure 2.7 shows the residual analysis results for all output variables of interest and Table 2.4 summarizes parameter values of the transformations used for proper application of Eq. (2.14) in the ECW. Interestingly, the standard deviation of innovations (σ_ν) and the Box-Cox transformation parameter for NOx at all sites were within the same range, while the correlation coefficients (ρ_ν) were slightly different.

41
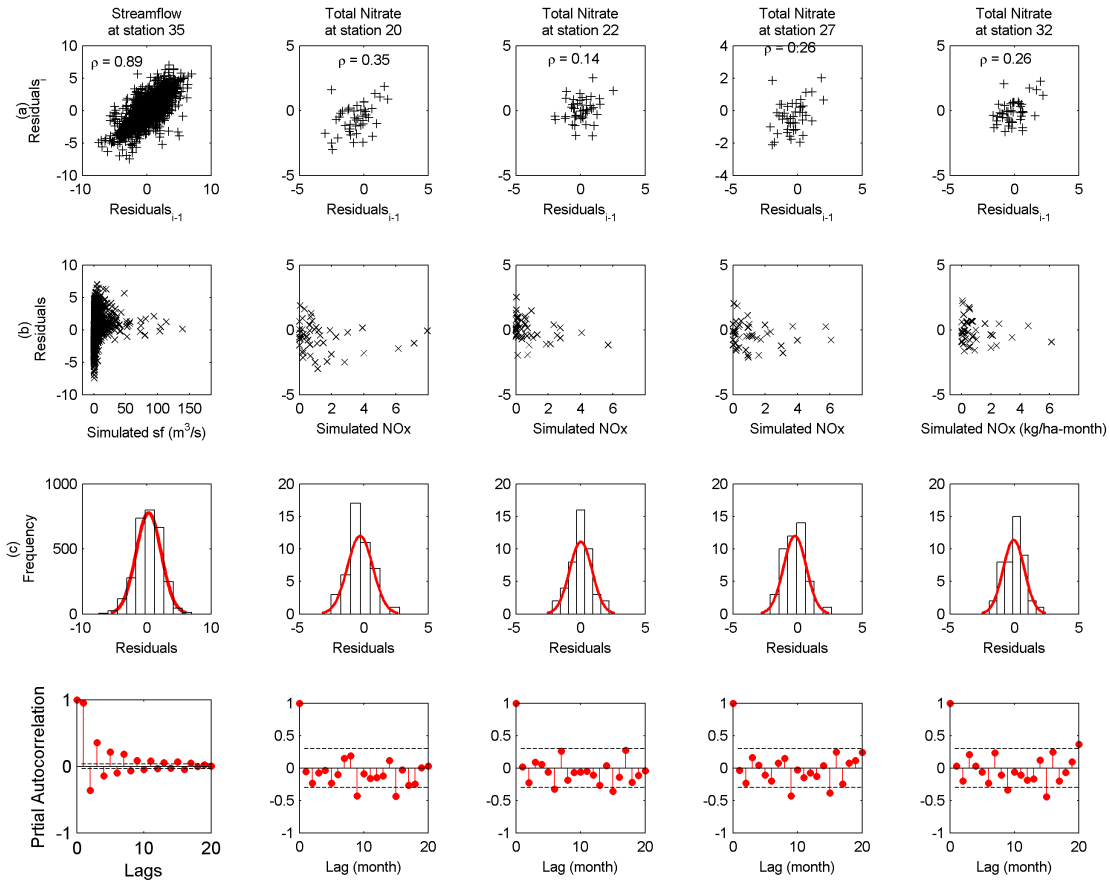
Figure 2.7: Analysis of residual for individual outputs after Box-Cox and AR-1 transformations: (a) lag-1 correlation of residuals, (b) residuals versus simulated values fordepiction of heteroscedasticity, (c) actual (bars) and fitted (solid line) frequency of residuals, and (d) partial auto-correlation coefficients (stems) of residuals with 95% significance levels (solid horizontal lines)

42
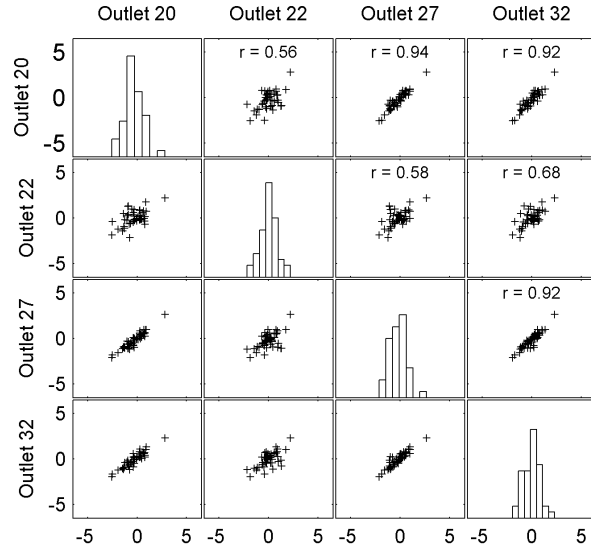
Figure 2.8: Scatter plot matrix of transformed residual errors (kg/ha-month) at four NO_2+NO_3 observation sites within the ECW

Although the above-mentioned transformations effectively removed the auto-correlation among residuals for an output variable at a given site, they did not address potential cross-correlation between model errors at different sites. This issue is particularly important in the context of watershed management where water quality data are typically collected at nested observation sites. Indeed, our analysis revealed the presence of strong correlation between model errors at various sites within the ECW, with coefficients of correlation ranging from 0.56 to 0.94 (Figure 2.8). The correlation was stronger for nested sites (e.g., sites 20 and 27). This highlights the necessity of using multiobjective approaches for calibration of watershed models at multiple sites within a given watershed system. Figure 2.8 summarizes the cross-correlation analysis of residuals for NOx data. To use Eq. (2.14) for aggregation of likelihood function from multiple sites, additional statistical methods should be employed to remove the cross-correlation between residuals.

43

2.4 Summary and Conclusions

The application of watershed models for pollutant source identification, non-point source pollution control, and watershed management requires assimilation of observed hydrologic and water quality responses with varying characteristics (i.e., quality and frequency) collected at different locations along the stream network. Often, observed data are reconciled using an aggregated objective function of weighted errors, and a single objective automatic calibration algorithm is used to identify the set of parameter values that minimizes the objective function. Recent studies have demonstrated that the selected objective function must be statistically correct to effectively determine optimal parameter values [*Mantovan and Todini*, 2006; *van Griensven et al.*, 2008]. However, calibration of a watershed model for many output variables at multiple sites is a multi criteria decision process. The analyst must choose a solution that not only minimizes errors, but more importantly, results in model simulations that adequately simulate the system behavior.

To this end, a computational framework was suggested to facilitate multisite many objectives calibration of watershed models, which includes: (i) an a-priori characterization of system behavior; (ii) a formal and statistically correct formulation of objective function(s) of model errors; (iii) an optimization engine to determine the Pareto-optimal front for the selected objectives; and (iv) a multi criteria decision analysis tool to select optimal solutions from the Pareto-optimal front that are most consistent with the goals of the modeling study. The framework was demonstrated for calibration of the SWAT model for streamflow response at one site and nitrate response at four locations within the Eagle Creek Watershed in Indiana. Both single objective (SCE and DDS) and multiobjective (NSGA-II) algorithms were examined to estimate 18 important model parameters important for estimating streamflow and nutrients.

This study revealed that for multisite many objective automatic calibration of a watershed model, both a formal likelihood function considering the structure of residuals and a multiobjective optimization approach are essential, particularly when a strict definition of system behavior is considered. A proper likelihood function was derived using Bayesian statistics that can be used to reconcile observed hydrologic time series for disparate objectives at multiple stream locations. The Box-Cox transformation and first-order autoregressive model were employed in sequence to reduce heteroscedasticity and eliminate correlation between residuals. It became evident that single objective calibration methods (SCE and DDS) find a lower (better) value for the aggregated objective function of weighted errors while requiring fewer model evaluations. However, the use of the solutions from single objective techniques was limited because the simulations did not mimic the observed behavior of the system for all objectives at all sites. Based on a satisfactory, good, or very good classification of model simulations, multiobjective methods were the only methods that yielded behavioral solutions. Satisfying a stricter definition of the system behavior required incorporation of a separate objective function for each response at each location within the multiobjective optimization framework.

It was evident that as the number of objectives in the calibration procedure increases, the convergence to the Pareto-optimal front becomes slower. The aggregation of information for the same response variable (nitrate in this study) at different observational sites using the proposed likelihood function appeared as a pragmatic approach for enhancing the speed of convergence to the Pareto-optimal front. However, residuals for nested sites tended to be highly correlated. Therefore, aggregation of information even for the same response should be conducted with a careful examine of residuals.

Bibliography

Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resources Research*, *43*(1), doi:10.1029/2005WR004745.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment, Part I: Model development, *Journal of the American Water Resources Association*, *34*(1), 73–89, doi:10.1111/j.1752-1688. 1998.tb05961.x.

ASCE (1993), Criteria for Evaluation of Watershed Models, ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management, Irrigation, and Drainage Division, *Journal of Irrigation and Drainage Engineering*, *119*(3), 429, doi:10.1061/(ASCE)0733-9437(1993)119: 3(429).

Beck, M. B. (1987), Water quality modeling: A review of the analysis of uncertainty, *Water Resources Research*, *23*(8), 1393, doi:10.1029/WR023i008p01393.

Bekele, E. G. (2005), Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms, *Water Resources Research*, *41*(10), doi:10.1029/2005WR004090.

Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, *6*(3), 279–298, doi: 10.1002/hyp.3360060305.

Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, *249*(1-4), 11–29, doi:10.1016/S0022-1694(01)00421-8.

Blazkova, S., and K. Beven (2009), A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, *Water Resources Research*, *45*, doi:10.1029/2007WR006726.

Box, G. E. P., and D. Cox (1964), An analysis of transformations, *Journal of the Royal Statistical Society Series B*, *26*, 1–78.

Box, G. E. P., and G. C. Tiao (1992), *Bayesian Inference in Statistical Analysis*, Wiley-Interscience.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel (2008), *Time Series Analysis: Forecasting and Control (Wiley Series in Probability and Statistics)*, 784 pp., Wiley.

Brockhoff, D., and E. Zitzler (2007), Operations Research Proceedings 2006, waldmann, ed., pp. 423–429, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-540-69995-8.

Brown, L. C., and T. O. Barnwell (1987), The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual, *Tech. rep.*, EPA document EPA/600/3-87/007. USEPA, Athens, GA.

Chen, J. H., D. E. Goldberg, S.-Y. Ho, and K. Sastry (2002), Fitness Inheritance in Multi-Objective Optimization, in *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 319–326.

Deb, K. (2001), *Multi-Objective Optimization Using Evolutionary Algorithms*, 518 pp., Wiley.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multi-objective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197, doi:10.1109/4235.996017.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, *28*(4), 1015, doi:10.1029/91WR02985.

Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008), Learning from model improvement: On the contribution of complementary data to process understanding, *Water Resources Research*, *44*(6), doi:10.1029/2007WR006386.

Figueira, J., S. Greco, and M. Ehrogott (Eds.) (2004), *Multiple Criteria Decision Analysis: State of the Art Surveys (International Series in Operations Research & Management Science)*, 1081 pp., Springer.

Fodor, I. . K. (2002), A survey of dimension reduction techniques (LLNL technical report UCRL-ID-148494), *Tech. rep.*, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory.

Gassman, P. W., M. R. Reyes, C. H. Green, and J. G. Arnold (2007), The Soil and Water Assessment Tool: Historical development applications, and future research directions, *Transactions of the American Society of Agricultural and Biological Engineers*, *50*, 1211–1240, doi:10.1.1.88.6554.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, *34*(4), 751, doi:10.1029/97WR03495.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1999), Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration, *Journal of Hydrologic Engineering*, *4*(2), 135, doi:10.1061/(ASCE)1084-0699(1999)4: 2(135).

Hantush, M. M., T. Dai, and J. Koenig (2005), TMDL Model Evaluation and Research Needs, EPA/600/R-05/149 (NTIS PB2006 101508), *Tech. rep.*, U.S. Environmental Protection Agency, Washington, Washington, DC.

Holland, J. H. (1975), *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, Michigan.

Jacquet-Lagrèze, E. (2001), Preference disaggregation: 20 years of MCDA experience, *European Journal of Operational Research*, *130*(2), 233–245, doi:10.1016/ S0377-2217(00)00035-7.

Kaika, M. (2003), The water framework directive: A new directive for a changing social, political and economic European framework, *Geography*, *11*(3), 299–316, doi:10.1080/0965431032000070802.

Khu, S. T., and H. Madsen (2005), Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration, *Water Resources Research*, *41*(3), doi:10.1029/2004WR003041.

Klepper, O., H. Scholten, and J. P. G. D. Van Kamer (1991), Prediction uncertainty in an ecological model of the oosterschelde estuary, *Journal of Forecasting*, *10*(1-2), 191–209, doi:10.1002/for.3980100111.

Kollat, J., and P. M. Reed (2007), A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications, *Advances in Water Resources*, *30*(3), 408–419, doi:10.1016/j.advwatres.2006.05.009.

Konikow, L. F., and J. D. Bredehoeft (1992), Ground-water models cannot be validated, *Advances in Water Resources*, *15*(1), 75–83, doi:10.1016/0309-1708(92) 90033-X.

Kuczera, G., and M. Mroczkowski (1998), Assessment of Hydrologic Parameter Uncertainty and the Worth of Multiresponse Data, *Water Resources Research*, *34*(6), 1481–1489, doi:10.1029/98WR00496.

Madsen, H. (2000), Automatic calibration of a conceptual rainfallrunoff model using multiple objectives, *Journal of Hydrology*, *235*(3-4), 276–288, doi:10.1016/ S0022-1694(00)00279-1.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, *26*(2), 205–216, doi:10.1016/S0309-1708(02)00092-1.

Madsen, H., G. Wilson, and H. C. Ammentorp (2002), Comparison of different automated strategies for calibration of rainfall-runoff models, *Journal of Hydrology*, *261*(1-4), 48–59, doi:10.1016/S0022-1694(01)00619-9.

Mantovan, P., and E. Todini (2006), Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology, *Journal of Hydrology*, *330*(1-2), 368–381, doi:10.1016/j.jhydrol.2006.04.046.

Marshall, L. (2004), A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling, *Water Resources Research*, *40*(2), doi: 10.1029/2003WR002378.

McElroy, A., S. Chiu, J. Nebgen, A. Aleti, and F. Bennett (1976), Loading functions for assessment of water pollution from nonpoint sources, *Tech. rep.*, EPA 600/2-76-151. USEPA, Athens, GA.

McMillan, H., and M. P. Clark (2009), Rainfall-runoff model calibration using informal likelihood measures within a Markov chain Monte Carlo sampling scheme, *Water Resources Research*, *45*(4), W04,418, doi:10.1029/2008WR007288.

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

Nash, J. E., and J. Sutcliffe (1970), River flow forecasting through conceptual models part I A discussion of principles, *Journal of Hydrology*, *10*(3), 282–290, doi: 10.1016/0022-1694(70)90255-6.

National Research Council (2001), *Assessing the TMDL Approach to Water Quality Management*, The National Academies Press, Washington, DC.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams (2005), Soil and Water Assessment Tool (SWAT) Theoretical Documentation, Version 2005, (BRC Report 02-05).

Nicklow, J. W., P. M. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh, and E. Zechman (2010), State of the art for genetic algorithms and beyond in water resources planning and management, *Journal of Water Resources Planning and Management*, *136*(4), 412, doi:10.1061/(ASCE)WR.1943-5452.0000053.

Pareto, V. (1971), *Manual of Political Economy*, 504 pp., Augustus M. Kelley, Publishers.

Rossi, C. G., T. J. Dybala, D. N. Moriasi, J. G. Arnold, C. Amonett, and T. Marek (2008), Hydrologic calibration and validation of the Soil and Water Assessment Tool for the Leon River watershed, *Journal of Soil and Water Conservation*, *63*(6), 533–541, doi:10.2489/jswc.63.6.533.

Runkel, R., C. Crawford, and T. Cohn (2004), Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers, in *U.S. Geological Survey Techniques and Methods Book 4, Chapter A5*, p. 69.

Saltelli, A., and I. M. Sobol' (1993), Sensitivity analysis for non-linear mathematical models, *Mathematical Modeling & Computational Experiment (Engl. Transl.)*, *1*(4), 407–414.

Schaefli, B., C. J. Harman, M. Sivapalan, and S. J. Schymanski (2011), HESS Opinions: Hydrologic predictions in a changing environment: behavioral modeling, *Hydrology and Earth System Sciences*, *15*(2), 635–646, doi:10.5194/hess-15-635-2011.

Seibert, J. (2000), Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, *Hydrology and Earth System Sciences*, *4*(2), 215–224, doi:10.5194/hess-4-215-2000.

Servat, E., and A. Dezetter (1991), Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese savannah area, *Hydrological Sciences Journal*, *36*(4), 307–330.

Sharma, V., D. Swayne, D. Lam, and W. Schertzer (2006), *Parallel Shuffled Complex Evolution Algorithm for calibration of hydrological models*, 30–30 pp., IEEE, doi:10.1109/HPCS.2006.34.

Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resources Research*, *16*(2), 430, doi:10.1029/WR016i002p00430.

Spear, R. C. (1997), Large simulation models: calibration, uniqueness and goodness of fit, *Environmental Modelling & Software*, *12*(2-3), 219–228, doi:10.1016/S1364-8152(97)00014-5.

Spear, R. C., and G. M. Hornberger (1980), Eutrophication in peel inlet- II. Identification of crictical uncertainties via generalized analysis, *Water Research*, *14*(1), 43–49.

Stedinger, J. R., R. M. Vogel, S. U. Lee, and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resources Research*, *44*, doi:10.1029/2008WR006822.

Tedesco, L., D. Pascual, L. Shrake, R. Hall, L. Casey, P. Vidon, F. Hernly, K. Salazar, and R. Barr (2005), Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality, *Tech. rep.*, IUPUI, Indianapolis.

Tolson, B. A., and C. A. Shoemaker (2007), Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, *Water Resources Research*, *43*(1), doi:10.1029/2005WR004723.

Triantaphyllou, E., and K. Baig (2005), The impact of aggregating benefit and cost criteria in four MCDA methods, *IEEE Transactions on Engineering Management*, *52*(2), 213–226, doi:10.1109/TEM.2005.845221.

Ullrich, A., and M. Volk (2010), Influence of different nitrate-N monitoring strategies on load estimation as a base for model calibration and evaluation., *Environmental Monitoring and Assessment*, doi:10.1007/s10661-009-1296-8.

USDA NASS (2003), USDA-National Agricultural Statistics Service, Cropland Data Layer.

USDA NRCS (2010), Soil Data Mart.

USGS NED (2010), 1 arc Second Digital Elevation Model.

USGS NLCD (2001), 2001 Land Cover Data.

van Griensven, A., and W. Bauwens (2003), Multiobjective autocalibration for semidistributed water quality models, *Water Resources Research*, *39*(12), doi: 10.1029/2003WR002284.

van Griensven, A., and T. Meixner (2007), A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models, *Journal of Hydroinformatics*, *9*(4), 277, doi:10.2166/hydro.2007.104.

van Griensven, A., T. Meixner, R. Srinivasan, and S. Grunwald (2008), Fit-for-purpose analysis of uncertainty using split-sampling evaluations, *Hydrological Sciences*, *53*(5), 1090–1103.

Vrugt, J. A. (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research*, *39*(8), doi:10.1029/2002WR001642.

Vrugt, J. A., and B. A. Robinson (2007), Improved evolutionary optimization from genetically adaptive multimethod search, *Proceedings of the National Academy of Sciences of the United States of America*, *104*(3), 708–11, doi:10.1073/pnas.0610471104.

Vrugt, J. A., M. P. Clark, C. G. H. Diks, Q. Duan, and B. A. Robinson (2006), Multi-objective calibration of forecast ensembles using Bayesian model averaging, *Geophysical Research Letters*, *33*(19), doi:10.1029/2006GL027126.

Vrugt, J. A., C. J. F. Braak, C. Diks, B. Robinson, J. Hyman, and D. Higdon (2009a), Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling, *International Journal of Nonlinear Sciences and Numerical Simulation*, *10*(3), 273 – 290.

Vrugt, J. A., C. J. F. Braak, V. K. Gupta, and B. A. Robinson (2009b), Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stochastic Environmental Research and Risk Assessment*, *23*(7), 1011–1026, doi:10.1007/s00477-008-0274-y.

Williams, J. R., and R. Hann (1978), Optimal operation of large agricultural watersheds with water quality constraints, *Tech. rep.*, Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Yeo, I.-K., and R. A. Johnson (2000), A new family of power transformations to improve normality or symmetry, *Biometrika*, *87*(4), 954–959, doi:10.1093/biomet/87.4.954.

Zhang, X., R. Srinivasan, and D. Bosch (2009), Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging, *Journal of Hydrology*, *374*(3-4), 307–317, doi:10.1016/j.jhydrol.2009.06.023.

Zitzler, E., and L. Thiele (1999), Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation*, *3*(4), 257–271, doi:10.1109/4235.797969.

Zitzler, E., M. Laumanns, and L. Thiele (2001), SPEA2: Improving the Strength Pareto Evolutionary Algorithm, *Tech. rep.*

Chapter 3

A MIXED DISCRETE-CONTINUOUS VARIABLE MULTIOBJECTIVE GENETIC ALGORITHM FOR TARGETED IMPLEMENTATION OF NONPOINT SOURCE POLLUTION CONTROL PRACTICES

Highlights

Planning for the implementation of nonpoint source pollution control strategies at the watershed scale hinges on abating pollutant movement from the landscape to water bodies at minimum cost. This paper presents an integrated simulation-optimization approach for targeted implementation of agricultural conservation practices at the watershed scale. A multiobjective genetic algorithm (NSGA-II) with mixed discrete-continuous decision variables was coupled with a distributed watershed model, Soil and Water Assessment Tool (SWAT), to identify optimal types and locations of conservation practices for nutrient and pesticide control at the watershed scale. While management options for NPS pollution control could be characterized as discrete or continuous decision variables, pervious optimization studies have only used binary representation of these choices. In this study, a novel discrete-continuous decision variable representation was used to find more realistic solutions for nonpoint source pollution control planning. Application of the proposed framework in the Eagle Creek Watershed, Indiana, indicated that while the types and locations of conservation practices from the mixed-variable NSGA-II were more effective in meeting water quality targets at lower costs than binary-variable

56

optimization, the procedure was considerably slower in finding the Pareto-optimal solutions and the tradeoffs between environmental and economic factors. A method for hybridization of binary- and mixed-variable NSGA-II methods in the context of NPS pollution control practices was developed to enhance the computational efficiency of the optimization procedure. As a result, the number of model simulation required for convergence to the Pareto-optimal solutions was reduced by 96 percent. The conceptual complexity and computational requirements of optimization-based approaches are impediments to their wider application for targeted implementation of NPS pollution control strategies. The methods and finding of this study address these issues and could result in a more effective implementation of management strategies at the watershed scale.

3.1 Introduction

Nonpoint source (NPS) pollution control practices, also known as Best Management Practices (BMPs), are widely accepted control measures of nonpoint sources of pollutants from agricultural landscapes [*Novotny*, 1993; *Ritter and Shirmohammadi.*, 2001; *Mostaghimi et al.*, 1997]. The primary goal of watershed scale conservation plans is minimizing pollutants movement from fields to water bodies at minimum cost. Thus, watershed planning is inherently a multiobjective problem. Watershed plans for nonpoint source pollution control are developed using several approaches. Government agencies promote implementation of NPS pollution control practices by recommending cost-share program that is a field-by-field approach [*Veith et al.*, 2003]. While this approach might be effective in a field or farm level, maximum water quality benefits at the watershed scale is not guaranteed [*Arabi et al.*, 2006]. Thus, NPS pollution control can be enhanced through development of the watershed scale conservation plans [*Maringanti et al.*, 2011]. The critical source area targeting method is a watershed scale planning approach that suggests implementation of conservation practices in critical source areas within the watershed, which contribute larger amounts of NPS pollutants. Targeted implementation of conservation practices in this approach is often subjective and based on the experts recommendations [*Veith et al.*, 2004].

Another drawback with critical source area approach is that it does not incorporate all important watershed processes and interactions, and therefore, does not guarantee the cost-effectiveness of the developed conservation plan. In addition, monitoring long-term impact of implemented conservation practices on water quality via field-studies is infeasible mainly because of the complexity of hydrologic and water quality processes and changes in annual weather pattern [*Arabi et al.*, 2007; *Veith*, 2002]. Moreover, impact of conservation practices can be delayed for several years [*Veith*, 2002]. Hence, estimation of the water quality benefits of NPS

pollution control strategies necessitate the use of appropriate hydrologic and water quality models. Watershed models can simulate hydrologic and water quality responses within a watershed system of interest and can help to identify water quality impact of conservation practices.

Hydrologic and water quality processes are highly complex and comprise a network of nonlinear dynamics. Distributed watershed models are commonly used to simulate natural processes and the response of the watershed to changes in land use, climate, and land and water management. The number of possible NPS pollution control scenarios within a watershed increases exponentially with the number of fields. Thus, it is not possible to evaluate performance of all possible scenarios at all fields within the watershed. Furthermore, incorporation of socioeconomic factors in the watershed planning process increases complexity of the procedure for identification of cost-effective nonpoint source pollution control plan. Recent studies have shown that optimization methods hold great promise for optimal allocation of NPS pollution control measures at the watershed scale [*Arabi et al.*, 2006; *Veith*, 2002; *Maringanti et al.*, 2011; *Jha et al.*, 2009; *Rabotyagov et al.*, 2010].

The watershed planning process involves many objectives. Optimal allocation of conservation practices requires incorporation of water quality criteria as well as socioeconomic factors in the planning process. However, aggregation of disparate and often conflicting environmental and socioeconomic criteria into a single objective is a pragmatic approach for the evaluation of proposed management actions and comparisons amongst them. When a single objective function is used, the objective function must reflect the compromise between often conflicting environmental, economic, and social objectives [*Ngatchou et al.*, 2005]. However, finding a satisfactory compromise solution that satisfy all stakeholders' preferences is difficult if not impossible [*Srinivas and Deb*, 1994]. Multiobjective optimization methods present an alternative approach that can effectively explore the tradeoffs between socioeconomic and environmental factors. While earlier studies utilized single-objective

59

optimization approach for optimal allocation of conservation practices [*Veith et al.*, 2003; *Arabi et al.*, 2006], more recent efforts have focused on the development and application of multiobjective approaches [*Jha et al.*, 2009; *Maringanti et al.*, 2008].

The type, size, and capacity of conservation practices are usually defined explicitly as binary options and optimization algorithms are operated to find the optimal combination of decision variables. However, many real world decisions problems consist of continuous decision variables. While, most of the conservation practice optimization studies considered binary decision variables, recent studies have dealt with discrete (integer) decision variables [*Maringanti et al.*, 2011]. However, inclusion of continuous decision variables may lead to the selection of more realistic and better solutions in terms of the final "optimal" set of type and placement conservation practices, also known as "Pareto-optimal front" solutions. It must be noted that representation of decision variables in continuous form is not always a practical choice. For example, length of stabilized river bank, size of ponds and wetlands, and percentage decrease in application of chemicals can be considered as the continuous decision variables, although discrete decision variable might be more preferable for width of grassed waterways and height of grade stabilization structures. Discrete-continuous optimization, also referred to as "mixed-chromosome" or "mixed-variable" optimization in evolutionary algorithms, is an active research topic with applications in a variety of engineering and scientific disciplines [*Gantovnik et al.*, 2003]. Several studies have shown that implementation of evolutionary algorithms with binary decision variables is more efficient than implementation with continuous variables by virtue of its reduced objective space to finite sets of points [*Gantovnik et al.*, 2003; *Brockhoff and Zitzler*, 2007; *Andriyenko et al.*, 2012]. However, the choices of types and location of nonpoint source pollution control strategies include both discrete and continuous options. Hence, implementation of evolutionary algorithms using continuous variables will likely result in identifying solutions with overall better fitness.

Genetic algorithm (GA) and its extensions have been used as efficient and effective tools for optimal placement of conservation practices because, among other pragmatic considerations, they do not require differentiability and continuity for the objective and constraint functions. GA, originally proposed by *Holland* [1975], mimics the process of natural evolution such as selection, crossover, and mutation, referred to as genetic operators. The GA optimization and its extensions have shown to be particularly suited for optimization problems consisting several, often conflicting objectives [*Goldberg*, 1989; *Zitzler and Thiele*, 1999]. GAs perform well in finding optimal solutions within large search spaces and, if properly implemented, can efficiently explore (broad search) and exploit (local search) the search space [*El-Mihoub et al.*, 2006]. The GA-based optimization methods guarantee "convergence" but not "optimality" [*Ingber and Rosen*, 1992]. In other words, although it is not possible to guarantee that the global optimal solution is found, identifying regions that encompass some good solutions can be ascertained [*Muleta and Nicklow*, 2005; *Vecchietti et al.*, 2003].

The computational requirements and technical complexity of optimization approaches are barriers to adoption of GA for cost-effective implementation of NPS pollutions control practices [*Vecchietti et al.*, 2003]. Several methods have been proposed to facilitate faster convergence of evolutionary algorithm optimization methods by hybridizing search algorithms [*Renders and Flasse*, 1996; *Joines et al.*, 1997; *Durand and Alliot*, 1999; *Sotiropoulos et al.*, 1997] and/or decision variables [*Lempitsky et al.*, 2008]. Hybridization of GAs with a local search method have shown to enhance the computational efficiency and the overall performance of evolutionary algorithms [*Sinha and Goldberg*, 2003; *van Hentenryck and Milano*, 2011]. Local search methods are employed to search for a better set of solutions within a marginal proximity of the results obtained from the global method. Hybridization of decision variables has also shown successful application in optimizing complex problems by reducing the computational time [*Lempitsky et al.*, 2008].

61

The primary goal of this study is to present an integrated simulation-optimization framework to identify the set of optimal types and locations of NPS pollution control practices at the watershed scale. Two specific objectives are defined: (1) to develop a novel genetic algorithm-based multiobjective optimization method, which will focus on improved selection of decision variables and versatile formulation of environmental and economic objectives; (2) to determine if mixed-variable optimization method can identify more cost-effective NPS pollution control plans than binary optimization method; and (3) to examine hybridization of GA for faster convergence to optimal solutions. The proposed framework is demonstrated for optimal allocation of agricultural conservation practices that minimize nutrient and pesticide loads at minimum cost in the Eagle Creek Watershed, Indiana, USA.

3.2   Methods and Materials

The proposed simulation-optimization framework for optimal allocation of NPS pollution control practices include three main component: (1) watershed model for simulation of hydrologic and water quality processes under scenarios with and without NPS pollution control measures; (2) an optimization engine for reconciliation of environmental and socioeconomic factors; and (3) an economic module for cost-benefit analysis of the implemented conservation plans. The framework was tested for optimization of conservation plans using binary- and mixed discrete-continuous variables for minimizing atrazine and nitrated loads at minimum cost in the Eagle Creek Watershed, Indiana, USA. Hybridization of search algorithms and decision variables were also examined to enhance the computational efficiency of the optimization procedure.

3.2.1   Multiobjective Optimization Framework

The framework developed in this study incorporates water quality benefits expressed in terms of the reduction of sediment, nutrient, and pesticide fluxes, on-

site and offsite costs and benefits, and water quality vulnerability. The proposed optimization approach for targeting NPS pollution control practices includes the following objective functions and constraint functions:

$$\begin{cases} \text{minimize } y = f(\boldsymbol{x}|\boldsymbol{\theta}, I, c_s, t_d, T) & \text{pollutant load(s)} \\ \text{minimize } C = g(\boldsymbol{x}|\boldsymbol{\theta}, I, p, c_s, t_d, T) & \text{cost(s)} \\ \text{maximize } \pi = h(\boldsymbol{x}|\boldsymbol{\theta}, I, r, c_s, t_d, T) & \text{profit(s)} \end{cases} \quad (3.1)$$

subject to:

$$\begin{cases} y \leq y_{allowable} & \text{water quality targets} \\ C \leq C_{max} & \text{economic constraints} \\ N(\boldsymbol{x}|m, s) \geq 0 & \text{management and social constraints} \end{cases}$$

where $\boldsymbol{x}$ denotes discrete-continuous decision variables reflecting the type and location of practices, y represents pollutant load(s) over the assessment period T estimated by the watershed model f, $\boldsymbol{\theta}$ is the vector of calibrated watershed model parameters, I represents input forcing, t_d is the design lifetime of conservation practices, C represents the total on-site and off-site cost of implementation of nonpoint source pollution control practices over the assessment period T, p is the unit cost of conservation practices, π is the total on-site and off-site profit computed for the conservation strategy over the assessment period T, r is the unit price of beneficial products of conservation strategy. The term $y_{allowable}$ indicates water quality standard, C_{max} is the maximum available budget for implementation of the nonpoint source pollution control strategy $\boldsymbol{x}$ The term N represents management preferences (m) and social values (s), and.

The objective functions in Eq. (3.1) are often conflicting and incommensurable. For example, implementation of a larger number of conservation practices would likely result in lower pollutant loads, but the cost for implementation and maintenance of practices would increase. Hence, the optimal solution(s) for each objective could substantially differ from the optimal solution(s) for the other objectives. Multiobjective optimization approaches can determine a set of nondominated

solutions that comprise the "Pareto-optimal front". Nondominated solutions are set of solutions in the search space that are better than any other solution in space in one or more objective [*Srinivas and Deb*, 1994]. Any improvement in one objective among Pareto-optimal solutions will essentially result in the degradation of at least one other objective [*Pareto*, 1971]. Nondominated Sorted Genetic Algorithm II (NSGA-II) [*Deb*, 2001] is among the most commonly used multiobjective global optimization methods with numerous successful application in watershed management [*Bekele*, 2005; *Nicklow et al.*, 2010]. Hence, the NSGA-II was used as the search method in identifying optimal solutions for allocation of conservation practices.

3.2.1.1 Environmental Criteria

Water quality and environmental impacts of NPS pollution control practices are evaluated by means of an appropriate distributed watershed model. Minimizing pollutant loads, or alternatively maximizing reduction of pollutant loads, is a key objective in nonpoint source pollution control planning. The reduction of pollutant load z can be calculated as:

$$\Delta L_z = \frac{L_{z,base} - L_{z,BMP}}{L_{z,base}} \times \%100 \tag{3.2}$$

where ΔL_z is the estimated percent reduction of pollutant load z, while $L_{z,base}$ and $L_{z,BMP}$ represent the pollutant loads before and after implementation of NPS pollutions control practices, respectively. Water quality standards ($y_{allowable}$ in Eq. (3.1) are expressed differently depending on whether the transport of the pollutant of concern would cause chronic or acute contamination. Nutrient and pesticide load (or concentration) reduction targets are typically stated on an average annual basis [*US EPA*, 2008]. U. S. Environmental Protection Agency (EPA) has set maximum contaminant level of 3 *ppb* and 10 *mgN/L* for atrazine and nitrate in drinking water, respectively [*US EPA*, 1988].

3.2.1.2 Economic objective

Nonpoint source pollution control plans yield benefits in water quality and wildlife habitats, but their implementation and maintenance come at a cost to stakeholders that should be considered in decision making process. Several graphical and mathematical methods have been used for economic evaluation of agricultural technology and conservation practices [*World Bank*, 2001; *FAO*, 2001; *Whittaker*, 2003; *Arabi et al.*, 2006]. In particular, Data Envelopment Analysis (DEA) [*Charnes et al.*, 1978; *Xu and Prato*, 1995] provides a methodology for economic analysis of the agricultural technology and conservation practices. In this approach, each of the K decision making units uses a set of M inputs $x = (x_1, \cdots, x_M) \in \mathbb{R}_+^M$ (e.g., conservation practices and chemicals application) to produce a set of N outputs $u = (u_1, \cdots, u_N) \in \mathbb{R}_+^N$ (e.g., environmental benefits and crop yield). Representing input unit prices by $p \in \mathbb{R}_+^M$ and output unit prices by $r \in \mathbb{R}_+^N$, the profit (net return) throughout the study area is computed as [*Whittaker*, 2003]:

$$\pi(x, u) = \sum_{k=1}^{K} \left\{ \sum_{n=1}^{N} [(r^{n,k} + \alpha_r^{n,k})(u^{n,k} + \beta_u^{n,k})] - \sum_{m=1}^{M} [(p^{m,k} + \alpha_p^{m,k})(x^{m,k} + \beta_x^{m,k})] \right\}$$
(3.3)

where α_r and α_p represent changes in the unit prices of outputs and inputs, respectively. The terms β_u and β_x denote, respectively, changes in magnitude of outputs and inputs due to the application of conservation plan.

Equation (3.3) can be simplified using a partial budgeting approach, which is suited for evaluation of the economic effects of adjustments in some portion of the business [*Roth and Hyde*, 2002]. The method only considers the costs and returns that will be changed by implementation of technology or management practices. For example, to study economic impacts of the nutrient reduction in an agricultural field, only the "changes" in crop production and chemical application are considered. Thus, assuming constant input and output prices, Eq. (3.3) simplifies to[*Whittaker*,

2003]:

$$\pi(x,u) = \overbrace{\sum_{k=1}^{K}\sum_{n=1}^{N} r^{n,k}\beta_u^{n,k}}^{\text{benefit}} - \overbrace{\sum_{k=1}^{K}\sum_{m=1}^{M} p^{m,k}\beta_x^{m,k}}^{\text{cost}} \tag{3.4}$$

Cost of conservation practices consists of (i) implementation (or establishment), (ii) maintenance, (iii) damage, and (iv) opportunity costs. Implementation cost includes cost of installation of conservation practice and technical and field assistance. Maintenance cost is usually evaluated as a percentage of the establishment cost, r_M [*Arabi et al.*, 2006]. Damage cost may include off-site damages, such as pollutants delivered to streams, and on-site damage. Opportunity costs account for any gain or loss from crop production foregone where conservation practices are established. For individual conservation practices, the cost (C) is evaluated by:

$$C = C_0 + r_M \times C_0 \left(\frac{1 - (1-i)^{-t_d}}{i}\right) + C_D + C_O \tag{3.5}$$

where C_0 is the implementation cost and r_M is the maintenance cost as a percentage of C_0, which is capitalized to be consistent with present value of implementation. The term i denotes interest rate and t_d is the design lifetime of the conservation practice in years. Damage cost, C_D, and opportunity cost, C_O, are included for completeness, however set to zero hereafter. Total watershed cost of a nonpoint source pollution control plan is computed by summing the costs of individual conservation practices implemented throughout the watershed.

A objective of NPS pollution control policy instruments is to maximize the expected net economic benefits to the society from pollution control [*Ribaudo et al.*, 1999]. The damage cost of nonpoint source pollutions is difficult to measure, largely because of the unknown relationship between the transport of NPS pollutant loads and the corresponding economic damage, and considerable time lags between causes and effects of NPS pollution [*Ribaudo et al.*, 1999; *USDA NRCS*, 2011]. Moreover, site-specific estimation of offsite damage from nonpoint source pollution is data intensive and cost prohibitive [*Piper and Martin*, 2001]. The benefit transfer method

Table 3.1: Summary of offsite damage cost of environmental pollutants

Pollutant	Cost (2012$/metric ton)	Reference
Erosion	10.0	*Hansen and Ribaudo* [2008]
Total Nitrogen	6,000	*Ribaudo et al.* [1999]
Total Phosphorus	3,650	*Fang and Easter* [2003]
Pesticide	16,000	*Pretty et al.* [2001]

provides a fairly inexpensive approach for estimating the damage cost of NPS pollution [*Brookshire and Neill*, 1992], where economic values of damage/benefit from a "study area" are extended to a "policy area" in another time and location [*Walsh et al.*, 1992; *Desvousges et al.*, 1992]. Previous studies have used the benefit transfer approach to estimate offsite cost of soil erosion and the cost associated with the transport of nitrogen, phosphorus, and pesticides at different geographical scales [*Hansen and Ribaudo*, 2008; *Ribaudo et al.*, 1999]. Table 3.1 summarizes the external costs of major environmental pollutants from other studies [*Hansen and Ribaudo*, 2008; *Ribaudo et al.*, 1999; *Fang and Easter*, 2003; *Pretty et al.*, 2001]. All values are adjusted using 4 percent average annual inflation rate to year 2012 dollars.

While offsite benefit of conservation practices are determined based on cost of offsite damage, onsite benefits are estimated based on the direct benefits of implementation of conservation practices in the field. Onsite benefits could be negative (e.g. due to the reduced crop yield) or positive (e.g. through increased unit price of crops). Thus, the economic benefits of implementing pollution control strategies include: (1) offsite benefits, equivalent to the monetary values of the water quality improvement benefits; and (2) onsite benefits to agricultural production as a results of implementing conservation practices. Monetary values of water quality improvement are estimated based on the predicted sediment, nitrogen, phosphorus, and pesticide loads reduction (a direct output of watershed model) and associated onsite and offsite benefits, and can be expressed as:
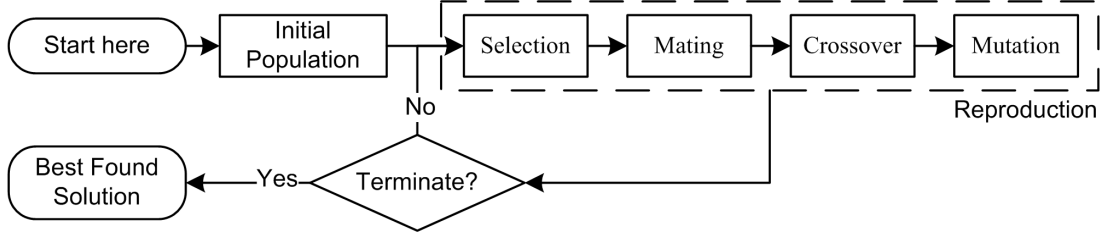
Figure 3.1: Flowchart of a typical genetic algorithm optimization method

$$B = MV_z \times \Delta L_z + \sum_{k=1}^{K} \sum_{n=1}^{N} r^{n,k} \beta_u^{n,k} \tag{3.6}$$

where MV_z is monetary value of water quality improvement for pollutant z (refer to Table 3.1 for NPS pollutants considered in this study).

3.2.1.3 Nondominated Sorted Genetic Algorithm II (NSGA-II)

The Nondominated Sorted Genetic Algorithm II (NSGA-II) is a heuristic multiobjective genetic algorithm developed by *Deb et al.* [2002]. The NSGA-II starts with an initial population of solutions and then implements probabilistic and parallel exploration in the search space using the domain-independent genetic operators (i.e. chromosome reproduction) to find optimal solutions [*Arabi et al.*, 2006]. Figure 3.1 illustrates the flowchart for implementation of a typical genetic algorithm. The procedure starts with an initial population of solutions that are typically generated randomly. The fitness of individual solutions in successive generations increases through selection, crossover, and mutation [*Goldberg*, 1989]. The procedure stops when a set of pre-defined termination conditions is met. Similar to other heuristic optimization algorithms, satisfying the optimality conditions cannot be guaranteed [*Lakshmanan*, 2000]. Therefore, defining termination criteria for NSGA-II is a hard and subjective task. The commonly used termination criteria are based on the maximum number of function evaluations and the number of successive generations without a significant improvement in the objective function values.

Planning of conservation practices at the watershed scale always consists of large number of decision variables. The computational time required to find the optimal solution(s) increases as the number and complexity of the decision variables and objectives increase [*Brockhoff and Zitzler*, 2007]. In particular, a larger number of decision variables would require a larger population size, which in turn would necessitate a larger number of function evaluations.

3.2.1.4 Selection of Decision Variables for Representation of NPS pollution Control options

While many options in NPS pollution control problem are discrete or continuous decision variables, most of the conservation practice optimization studies considered binary decision variables (1's and 0's respectively indicating that the corresponding conservation practice "be" or "not be" implemented). Incorporating continuous decision variables in NPS pollution control plan optimization may lead to the selection of more realistic and overall better fitness. A novelty of the proposed framework in this study is its capability to operate on both binary-discrete and continuous decision variables (known as "mixed-chromosome" or "mixed-variable"). It is also expected to achieve more diverse Pareto-optimal solutions through application of mixed-variable optimization. Figure 3.2 demonstrates the mixed-variable structure for placement of conservation practices in the optimization method. The length of each decision variable set (chromosome, n) corresponds to the total number of genes, i.e., combination of conservation practice types and locations that are considered in the optimization problem.

3.2.2 Hybridization of NSGA-II

The slow convergence to the optimal-front solution(s) is the biggest drawback with genetic algorithms in complex problems [*Vecchietti et al.*, 2003], especially when continuous decision variables are included. The computational requirements
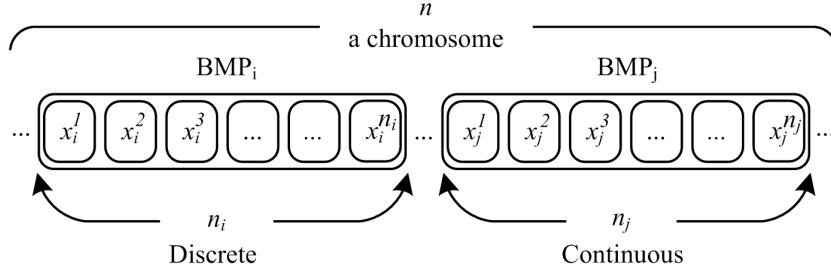
Figure 3.2: Schematic of a mixed discrete-continuous decision variable (also referred to as "mixed-chromosome"). n is the length of chromosome

and technical complexity of optimization approaches are barriers to adoption of NSGA-II for cost-effective implementation of NPS pollutions control practices [*Vecchietti et al.*, 2003]. However, the NSGA-II algorithm can be implemented using parallel computing procedures to address the increasing computational demand for complex environmental problems. In addition, several methods have been proposed to facilitate faster convergence of evolutionary algorithm optimization methods by hybridizing decision variables [*Brockhoff and Zitzler*, 2007] and/or search algorithms [*Renders and Flasse*, 1996; *Joines et al.*, 1997; *Durand and Alliot*, 1999; *Sotiropoulos et al.*, 1997]. In hybridization of decision variables, following termination of binary optimization, binary decision variables will evolve to a mixed decision variables to search for a better set of solutions. Therefore, a proper mechanism for translating binary decision variable to mixed discrete-continuous decision variables was used. Hybridization of search algorithm, on the other hand, is based on the well-known characteristics of the genetic algorithm guarantees convergence but not optimality. Given GA solutions as an initial guess, a local search method can be used to search a smaller region for further improvement in optimal solutions. NSGA-II method was hybridized with two multiobjective local optimization methods of goal programming and minimax to fine-tune the solutions obtained by the evolutionary algorithm in search for better local solutions.
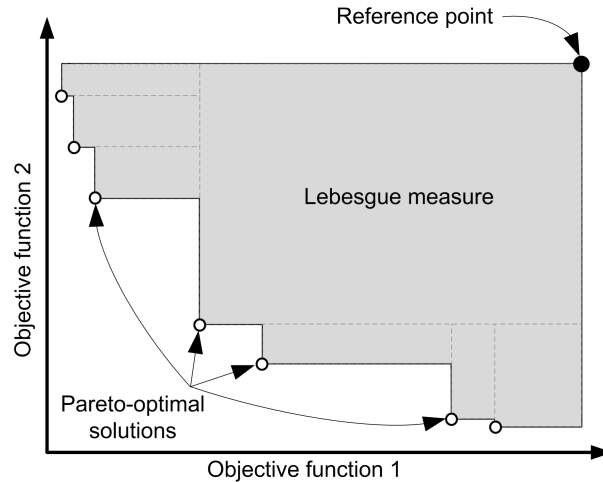
Figure 3.3: Schematic of a Lebesgue measure for a two-objective minimization problem. Area of the shaded space is the Lebesgue measure of the Pareto-optimal solutions with respect to the reference point

3.2.3 Comparison of Search Algorithms

Pareto-front solutions of mixed-variable and binary-variable optimization algorithms will be compared by means of hypervolume values estimated based on the Lebesgue measure of the nondominated solutions. Lebesgue measure, also called n-dimensional volume (Euclidean space), is the standard way of mapping a set of Pareto optimal points to a scalar [*Solovay*, 1970; *Fleischer*, 2003]. For n= 1, 2, or 3, it coincides with the standard measure of length, area, or volume. Lebesgue can be measured with respect to a lower or upper bound reference point for maximization or minimization problems. For a minimization problem with upper bound reference point, a better optimal front will have higher Lebesgue measure. Lebesgue measure has also the advantage of taking diversity of the solutions into account [*Fleischer*, 2003]. Figure 3.3 shows schematic of a Lebesgue measure calculated for a two-objective minimization problem.

71

3.3 Application of the Optimization Approach for Targeted implementation of Agricultural Conservation Practices

The proposed optimization framework was applied for the targeted implementation of agricultural conservation practices for nutrient and pesticide control in the Eagle Creek Watershed, Indiana, USA. The Soil and Water Assessment Tool (SWAT) was used to simulate hydrologic and water quality processes. The model was first calibrated and tested for a nine year period from 1995 to 2003 for daily streamflow, monthly nitrate loads and monthly atrazine loads at multiple locations within the watershed. Then, the calibrated model was used to evaluate water quality benefits of conservation scenarios examined during the optimization procedures. Since water quality standards are typically expressed in terms of average annual loads (or concentrations) for nutrients and pesticides, water quality benefits of management actions were calculates as the average annual reduction of loads during the simulation period. Specific objectives of the optimization approach were to simultaneously (i) minimize nitrate loads, (ii) minimize atrazine loads, and (ii) minimize cost of implementation. The optimization procedures were implemented using a population size of 108, a crossover probability of 0.5, and a mutation rate of 0.005.

3.3.1 Study Area

The Eagle Creek Watershed (ECW), located in central Indiana, has a drainage area of 248.1 km2 and lies within the Upper White River Watershed, extending from 40°01'24" to 40°04'16" north latitudes and 86°15'43" to 86°16'45" west longitudes.. According to the 2001 national Land Cover Dataset [*USGS NLCD*, 2001], the watershed consists of 52% croplands, 27% pasture, 12% low and high density urban areas, and 9% forest. The predominant crops are corn and soybeans. ECW drains into Eagle Creek Reservoir, which supplies drinking water for the city of Indianapolis. Figure 3.4 presents the location and land cover for the watershed. The
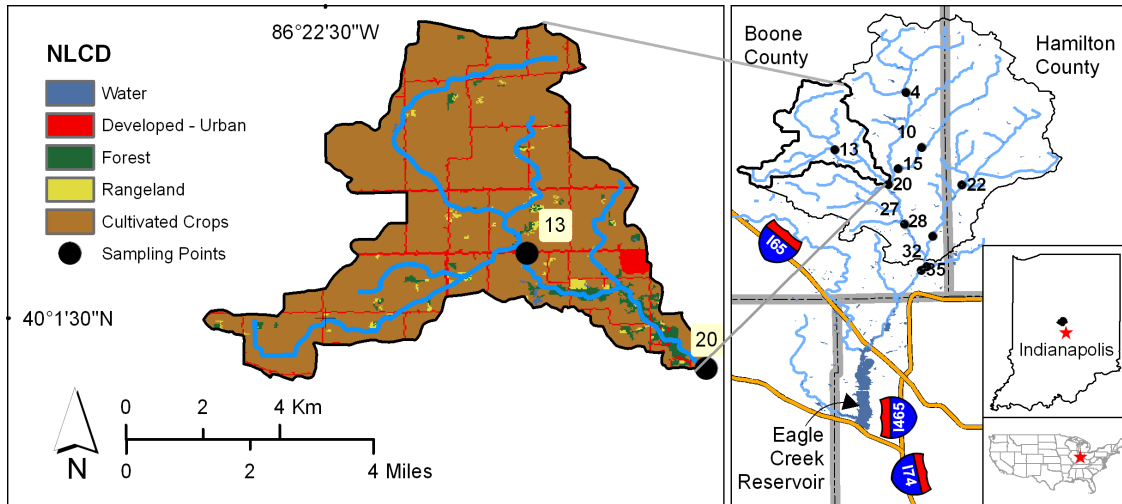
Figure 3.4: Location and landuse maps of the HUC 051202011102

soils are generally poorly draining and developed from glacial materials with thin loess over loamy glacial till and alluvial materials depositions. The dominant soils are the Crosby-Treaty-Miami in the headwaters and Miami-Crosby-Treaty along the downstream areas. The mean annual precipitation for the Eagle Creek Watershed area is 1052 mm. Monthly mean temperatures for this area from 1971-2000 shows January as having the lowest average temperature of -3.3°C and July as the being the warmest month with an average temperature of 23.7°C [*Tedesco et al.*, 2005].

Spatial scale of conservation planning depends upon numerous factors, including management objectives, available data resolution, dominant ecological processes, and potential sociopolitical constraints [*Walter et al.*, 2007; *Garen and Moore*, 2005]. To achieve the specific goals of water quality control, conservation practices targeting should be performed within a smaller geographic unit which ultimately allows us to better evaluate targeted management plan. Using 12-digit Hydrologic Unit Codes (HUCs) is recommended for the watershed-scale evaluation and planning of conservation practices to encourage participation of stakeholders in implementing conservation practices and enable monitoring water quality changes within subwatershed [*Haggard et al.*, 2010]. The ECW encompasses five 12-digit HUCs. The

HUC 051202011102 subwatershed with a drainage area of 41.2 km^2 was selected for the application of the integrated simulation-optimization framework in this study. The subwatershed consists of predominantly cropland (88%), and also contributes the largest amounts of non-point source nitrate and atrazine loads from agricultural lands to the Eagle Creek Reservoir (approximately 23% of nitrate loads and 28% of the atrazine loads).

Similar to many other agricultural watersheds in the Midwestern U.S., an extensive network of subsurface drainage systems (also known as tile drains) has been installed in the Eagle Creek watershed in areas with poor drainage capacity. Nearly 80 percent of agricultural fields in the watershed have tile drains, which substantially alter movement of water and chemicals in the watershed [*Tedesco et al.*, 2005]. Although tile drainage systems have many benefits, they increase the potential for loss of chemicals by short circuiting the natural flow of water. Thus, tile drains contributed to elevated nitrate and atrazine concentration and loss of natural wetlands in Indiana watersheds. *Tedesco et al.* [2005] have developed pollutants load reduction targets in Eagle Creek watershed that consists of 40% and 36% reduction in atrazine and nitrate loads, respectively.

3.3.2 Watershed Model Description

The Soil and Water Assessment Tool (SWAT; *Arnold et al.* [1998]) was used to represent hydrologic and water quality processes in the ECW. Hydrologic processes simulated by SWAT include snow accumulation and melt, evapotranspiration, infiltration, percolation losses, surface runoff, and groundwater flows [*Neitsch et al.*, 2005]. SWAT is a physically-based watershed-scale, distributed-parameter, continuous time, and long-term, model that runs on a daily time step. It subdivides a watershed into subbasins connected by a stream network, and further delineates hydrologic response units (HRUs) consisting of unique combinations of land cover and soils in each subbasin.

74

SWAT can simulate major nutrient processes within a watershed. The nitrogen (N) cycle is simulated in five pools: inorganic (including ammonium and nitrate) and organic (including fresh, stable, and active). The main N processes are mineralization, decomposition, and immobilization. Nutrients are introduced into the main channel through surface runoff and lateral flow and transported downstream with channel flow. Plant uptake, denitrification, volatilization, leaching, and soil erosion are the major mechanisms of N removal from a field. The transport rate of organic N with sediment is calculated with a loading function developed by *McElroy et al.* [1976] and modified by Williams and Hann [*Williams and Hann*, 1978] for application to individual runoff events. The loading function estimates daily organic N runoff loss based on the concentrations of constituents in the top soil layer, sediment yield, and an enrichment ratio. Nutrient transformations in the stream are controlled by the in-stream water quality component of the model that is adapted from QUAL2E in-stream water quality model [*Brown and Barnwell*, 1987]. More detailed description of the nutrient components of SWAT can be found in *Neitsch et al.* [2005]. SWAT uses algorithms from GLEAMS (Ground Water Loading Effects on Agricultural Management Systems) [*Leonard et al.*, 1987] and EPIC (Erosion Productivity Impact Calculator) [*Williams.*, 1990] to model pesticide's overland fate and transport and movement from land to streams. It also incorporates a simple mass-balance method developed by *Chapra* [2008] to model the transformation and transport of pesticides in streams.

A 30-m resolution DEM from USGS National Elevation Dataset [*USGS NED*, 2010], National Land Cover Dataset (NLCD) 1992 and 2001 [*USGS NLCD*, 2001] for urban areas, National Agriculture Statistics Service (NASS) Cropland Data Layer 2000-2003 [*USDA NASS*, 2003] for croplands, and SSURGO data from national resources conservation service (NRCS) [*USDA NRCS*, 2010] were used for watershed subdivision and delineating HRUs in the SWAT model. The ECW was subdivided

Table 3.2: SWAT performance indices for daily streamflow, monthly nitrate, and monthly total pesticide simulation during 1993-2004 including 2 years of warmup period. PBIAS and NSE indicate percent bias and Nash-Sutcliff Efficiency coefficient, respectively.

Gauging Station	Constituent	Calibration Period Criteria (1995-1999)			Validation Period Criteria (2000-2003)		
		PBIAS (%)	R^2	NSE	PBIAS (%)	R^2	NSE
20	Nitrate	7.9	0.94	0.83	16.9	0.85	0.67
	Atrazine	-6	0.81	0.34	-14	0.7	0.41
22	Nitrate	-22.3	0.89	0.78	1.24	0.74	0.36
	Atrazine	42	0.69	0.44	-0.1	0.5	0.28
27	Nitrate	0.59	0.93	0.85	18.3	0.78	0.59
	Atrazine	13	0.66	0.35	-30	0.51	0.19
32	Nitrate	-7.9	0.92	0.84	8.4	0.76	0.55
	Atrazine	42.3	0.75	0.52	33.1	0.51	0.14
35	Streamflow	-12.2	0.78	0.61	4.3	0.78	0.56

into 35 subwatersheds and a total of 446 hydrologic HRUs. HUC 051202011102 includes 4 subwatersheds and 40 HRUs (Figure 5.1. Watershed model is calibrated and validated for entire ECW model for predicting streamflow at watershed outlet, and nitrate and Atrazine at gauging stations 20, 22, 27, and 32. Table 5.2 presents performance indices of the SWAT model in predicting daily streamflow and monthly nitrate, and Atrazine.

3.3.3 Representation of conservation practices

In this study, water quality impacts of fertilizer management, grassed waterways, grade stabilization structures, and tillage/residue management were evaluated. Only row crops, including corn and soybean, were considered for implementation of the nonpoint source pollution control plan. Conservation practices were represented using numerical procedures from the published studies [*Arabi et al.*, 2004; *Arnold et al.*, 2011; *White et al.*, 2010]. SWAT includes explicit functions for representation of fertilizer management and tillage/residue management, grassed water ways [*Arnold et al.*, 2011]. Table 3.3 summarizes the relevant SWAT management operations and parameters and their corresponding values for representation

Table 3.3: SWAT parameter values for representation of conservation practices

Practice	Representation Parameters	Unit	Value
Fertilizer management	Fertilizer application rate reduction, FRT_KG (*.mgt)	%	20
Grassed Waterways	Manning's n, GWATN (*.ops)		0.35
	Width, GWATW (*.ops)	m	15
	Length, GWATL (*.ops)	km	_[A]
Grade Stabilization	Reduce slope steepness, CH_S1 (*.sub)	m	_[B]
Tillage/Residue Mgt.	Reduce curve number, CN2 (*.mgt)		2
	Harvest Efficiency, HARVEFF (*.mgt)		1 [C]
	Overland Manning's number, OV_N (*.hru)		0.1 [D]

[A] Estimated using equation $0.5 \times (area)^{1/2}$

[B] Estimated based on the height of structure (default is GSSH = 1.2 meters) with the minimum of 0

[C] Estimated based on the percent cover for different types of tillage. HARVEFF for conventional, conservation, and no-till is assumed to be 1, 01, and 0.5, respectively

[D] For conventional, conservation, and no-till is assumed to be assumed to be 0.1, 0.2, and 0.3 respectively

of BMPs. Implementation of grade stabilization structures are only considered in combination with grassed waterways. Conservation practices in the binary optimization approach were represented by default values presented in Table 3.3. Detailed description of the conservation practices representation methods can be found in *Arabi et al.* [2004], *Arnold et al.* [2011], and *White et al.* [2010]. In the mixed-variable genetic algorithm, continuous decision variables for fertilizer management and discrete decision variables for grassed waterways, grade stabilization structure, and tillage/residue management were considered and represented as shown in Table 3.4. Implementation cost, operation & maintenance rate, and design life for conservation practices are also summarized in Table 3.4. Interest rate of 4 percent was also used in calculation of total cost of nonpoint source pollution control plan.

3.3.4 Results and Discussion

Overall, results of optimizing conservation practices type and placement using binary- and mixed-variable NSGA-II suggest that the mixed-variable NSGA-II method was able to find Pareto-optimal solutions with better fitness than the commonly used binary optimization method, but it was computationally more demanding.

Table 3.4: SWAT parameter values for representation of conservation practices in mixed discrete-continuous framework

Practice	Representation Parameters	Unit	Value	Cost	r_{OM} [A]	t (years) [B]
Fertilizer management	Fertilizer application rate reduction, FRT_KG (*.mgt)	%	20			
Grassed Waterways	Manning's n, GWATN (*.ops)		0.35	11,400 ($/ha)	3	10
	Width, GWATW (*.ops)	m	15			
	Length, GWATL (*.ops)	km	-[C]			
Grade Stabilization	Reduce slope steepness, CH_S1 (*.sub)	m	-[D]	4,800 ($/structure)	2	15
Tillage/Residue Mgt.	Reduce curve number, CN2 (*.mgt)		2		0	1
	Harvest Efficiency, HARVEFF (*.mgt)		1	75 ($/ha)		
	Overland Manning's number, OV_N (*.hru)		0.1			

[A] Operation and maintenance rate, percentage of capital cost
[B] design lifetime of conservation practices
[C] Estimated using equation $0.5 \times (area)^{1/2}$
[D] Estimated based on the height of structure (default is GSSH = 1.2 meters) with the minimum of 0

Figure 3.5: Initial population (top) and Pareto-front solutions (bottom) from binary-
and mixed- variable optimization methods for Atrazine and nitrate load reduction

First the SWAT model was calibrated and tested for daily stream discharge, monthly nitrate loads, and monthly atrazine loads over a nine year period from 1995 to 2003. The calibrated SWAT model for the ECW was linked with both the binary-variable and mixed discrete-continuous NSGA-II optimization methods to expose the tradeoffs between water quality and economic criteria to meet the targets. Figure 3.5 shows the initial population and optimal front solutions for both binary and mixed-variable multiobjective optimization methods. The mixed-variable NSGA-II method finally converged to a better set of alternatives. At the same cost, mixed-variable NSGA-II results were up to 5% and 3% more effective than binary optimization in reducing atrazine and nitrate loads, respectively. Interestingly, the Pareto-optimal solutions were similar at lower costs. At higher costs (nearly $50,000), the optimal solutions diverged.

Lebesgue measures of Pareto-front solutions from binary- and mixed-variable NSGA-II mehods, measured with a reference to cost=$600,000, Atrazine load reduction= 0%, and nitrate load reduction= -30%, were respectively 4.54×10^8 and 5.71×10^8. In a minimization problem, Pareto-optimal front with larger Lebesgue measures indicates better fitness function values and/or diverse optimal solutions (refer to Figure 3.3). So, larger Lebesgue measure of Pareto-optimal front obtained from mixed-variable NSGA-II implies that for the same cost higher pollutant reductions are achievable.

In addition to tradeoff between implementation cost and pollutant load reduction, there was a tradeoff between atrazine and nitrate loads, too. The tradeoff was seen when tillage/residue management were implemented at the fields. This can be explained by the well-known impact of tillage practices on increased drainage and subsequently elevated nitrate leaching into the shallow groundwater that eventually flows into surface waters [*Kanwar*, 2006; *Malone et al.*, 2007]. In particular, this was highly expected in Eagle Creek Watershed that extensive network of tile drains has been installed. However, both optimization methods demonstrated successful performance in identifying a set of solutions without negative impacts on nitrate load.

In addition to the efficiency in converging to the Pareto-optimal front, the diversity of solutions on the optimal front is an important consideration in characterizing the tradeoffs between objectives [*Deb*, 2001]. The mixed-variable NSGA-II outperformed the binary NSGA-II in finding diversity of the solutions on Pareto-optimal front. This was evident with wider extension of the Pareto-optimal front (shown in Figure 3.5), larger Lebesgue measure (5.71×10^8 versus 4.54×10^8), and lager number of Pareto-optimal solutions (46 versus 38). Higher diversity indicates that solutions are more likely on global optima rather than local optima.

While mixed-variable NSGA-II showed more effectiveness in finding optimal solutions, it was less efficient in converging to Pareto-optimal solutions. Mixed-variable NSGA-II methods are efficient in exploring a wide search space and detecting promising areas that optimal solution can be found (also called "valley" or "near optimal region"), but it needs more computational time to search for the bottom of this valley [*Chelouah and Siarry*, 2000; *Vecchietti et al.*, 2003]. Figure 3.6 shows the progress of the Pareto-optimal fronts during the optimization for both binary- and mixed-variable NSGA-II methods. In both cases as optimization proceeded, a larger number of non-dominated solutions on the Pareto-optimal front were found. However, termination conditions for binary- and mixed-variable methods were satisfied in 116 and 6,380 generations, respectively. The application of the mixed-variable NSGA-II for optimal allocation of NPS pollution control practices in the ECW is highly complex, resulting in a slow to the Pareto-optimal front. The full enumeration of the three practice types in all fields comprising the twenty three hydrologic response units in the study watershed using the binary-variable approach would require $2^{3 \times 23}$ scenario evaluations. Mixed-variable NSGA-II, on the other hand, involves continuous decision variables and therefore had infinite number of scenarios.

To improve convergence of optimization methods toward optimal solutions, two hybridization methods were examined: hybridizing decision variables and search algorithms. In hybridization of decision variables, following termination of binary-variable optimization, binary decision variables was evolved to a mixed decision variables. Therefore, a proper mechanism for translating binary-variable to mixed-variable was used. This type of hybridization demonstrated substantial improvement in finding diverse of solutions, extended Pareto-front, and convergence to the Pareto-optimal front. Hybrid NSGA-II method terminated in 66 generations, that was more than 60 time faster than the mixed-variable NSGA-II method by itself.
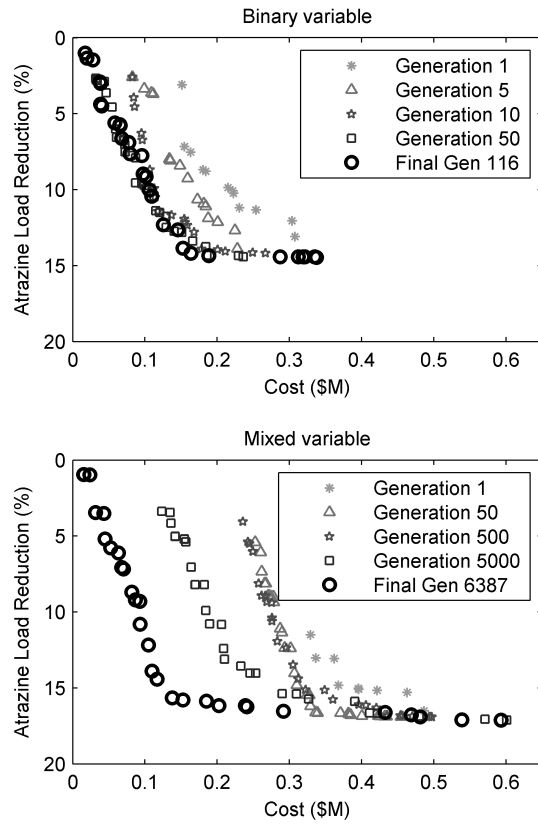
Figure 3.6: Progress of the optimal fronts during the optimization for binary- and mixed- variable optimization
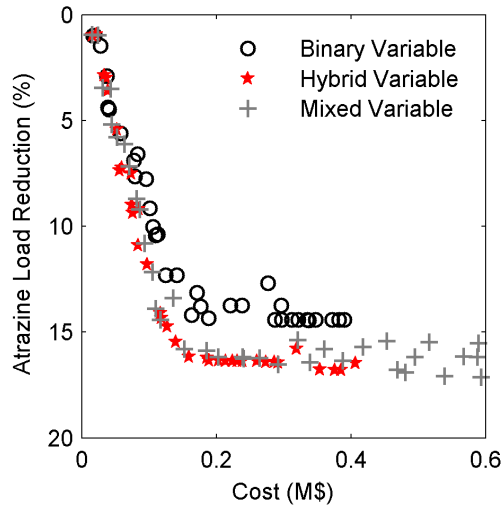
Figure 3.7: Comparison of Pareto-fronts from three different optimization settings

Results from mixed-variable and hybrid NSGA-II methods closely match regarding the objective functions values, and type and placement of conservation practices. Lebesgue measure was $\$5.78 \times 10^8$ that suggests slight improvement in the results in considerably lower model runs. Figure 3.7 compares Pareto-front from three optimization settings for atrazine removal.

Hybridization of search algorithm was based on the well-known characteristics of the genetic algorithm guarantees convergence but not optimality. Given GA solutions as an initial guess, a local search method can be used to search a smaller region for further improvement in optimal solutions. Multiobjective local optimization methods are gradient-based search methods that require an increase or decrease in the slopes of sigmoids, but gradient procedures are not stable during abrupt transitions in objective function shape [*Duch et al.*, 1999]. In our case study, inclusion of the discrete decision variables created step-like behavior in the objective function shape and therefore local optimization methods could not be improve the solutions. This was confirmed by hybridized of genetic algorithm with two multiobjective local optimization methods of goal programming and minimax, in which results were not improved any of the solutions from genetic algorithm method.

83

The shape of the tradeoff curve shifted from relatively smooth curves corresponding to the initial population and early generations to a Pareto-optimal curve with sharp edges in the final generation (shown in Figure 3.6). An obvious breakpoint in the Pareto-optimal curve is often considered a good compromise between objectives [*Madsen*, 2003]. In this study, sharp front reflects the fact that greatest degree of pollutant load reduction is achievable through the application of limited number of conservation practices with lower costs. For example on mixed-variable optimal front, the first 14 percent reduction in atrazine loads was possible with approximately $100,000, while to increase atrazine load reduction from 14 percent to 17 percent additional $100,000 was required. The breakpoint solution on mixed-variable Pareto-optimal front, that is considered the compromise solution, resulted in 14 and 20 percent reduction in atrazine and nitrate loads, respectively, with approximately $100,000 investment.

Each solution on the Pareto-optimal front explicitly corresponds to a specific optimal allocation of NPS pollution control practices in the study watershed. Investigation of the spatial distribution of the optimal conservation practices revealed that successive solutions on Pareto-optimal front have many solutions in common. Figure 3.8 demonstrates the type and location of the conservation practices associated with several scenarios on Pareto-optimal solutions from binary optimization. Grassed waterways received the highest priorities for both pesticide and nitrate load control. As expected, residue/tillage practices did not appeared in solutions at lower costs, mainly because of their adverse impact on nitrate loads. With the incremental increase in the conservation expenditure, most of the conservation practices and their locations from solutions with lower costs on the Pareto-optimal front remained unchanged. This observation confirms the need for targeting conservation practices to portions of the watershed that are more vulnerable to NPS pollution and therefore more suitable for mitigating the inverse impacts of human activities with less
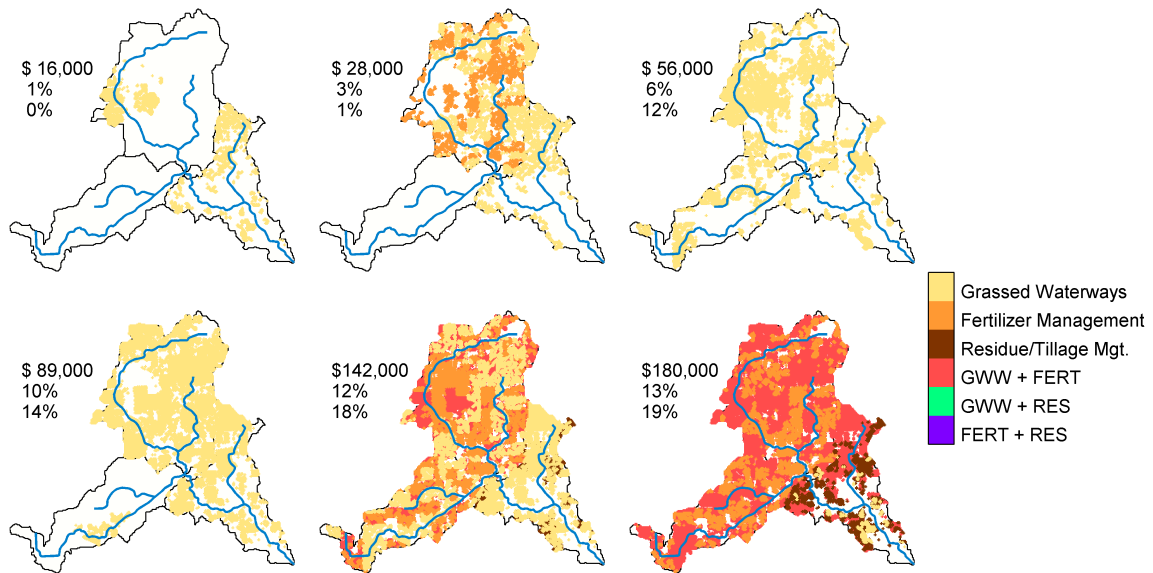
Figure 3.8: Spatial distribution of conservation practices for selected solutions on Pareto-optimal front. Numbers in boxes are present value of implementing cost ($), atrazine load reduction (%), and nitrate load reduction (%)
, respectively

expenditure [*Walter et al.*, 2007]. This also provides an opportunity for modular planning of watershed scale conservation practices, while giving higher priorities to more effective and less expensive alternativ'1es.

3.4 Summary and Conclusions

An integrated simulation-optimization framework for optimal placement of agricultural conservation practices is presented. A novel mixed-variable multiobjective genetic algorithm based on the commonly-used NSGA-II method was coupled with a spatially distributed watershed model, Soil and Water Assessment Tool (SWAT), and was used to realize the Pareto-optimal sets of conservation practices at the watershed scale. A case study is presented to demonstrate the application of the proposed framework in realizing optimal suite of conservation practices in the Eagle Creek Watershed, Indiana, USA.

Nondominated Sorted Genetic Algorithm II (NSGA-II) is used to find optimal suite of conservation practices type, size, and location with both binary and mixed decision variables. We also analyzed efficiency and effectiveness of the optimizations in terms of convergence rate, diversity, and optimality of the solutions. Two approaches were also taken to improve efficiency of GA algorithm by (i) updating binary to mixed-variable during the optimization, and (ii) hybridizing GA with a local search algorithm. Spatial distribution of the conservation practices type and location were also studied. Soil and Water Assessment Tool (SWAT) was used to simulate runoff and water quality and assess performance of the nonpoint source pollution control strategies. Objectives were to simultaneously minimize nitrate load, atrazine load, and cost of implementation of nonpoint source pollution plan. Results from implementing different optimization setting showed that

1. For an optimal placement of conservation practices in a watershed-scale, discrete-continuous decision variable, referred to as "mixed-variable", optimization method identified a set of solutions which is more effective than solutions obtained from commonly used binary optimization method for the same amount of cost

2. Mixed-variable optimization provided more realistic alternatives and higher flexibility to the decision makers

3. Using mixed-variable optimization increased complexity of the optimization problem that increases computational time by several orders of magnitude. Using hybrid optimization algorithms substantially improved the efficiency of mixed-variable optimization methods

Investigating spatial distribution of the optimal conservation practices showed that there is an obvious overlay of the conservation practices type and location within consecutive solutions on Pareto-front. This overlay provides an opportunity

86

for modular planning of watershed scale conservation strategies, while giving higher priorities to more effective and less expensive alternatives.

Bibliography

Andriyenko, A., K. Schindler, and S. Roth (2012), Discrete-Continuous Optimization for Multi-Target Tracking, in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, Rhode Island.

Arabi, M., R. S. Govindaraju, and M. M. Hantush (2004), Impact of Best Management Practices on Water Quality of Two Small Watersheds in Indiana: Role of Spatial Scale, *Tech. rep.*, US EPA, National Risk Management Research Laboratory, Cincinnati, OH.

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

Arabi, M., R. S. Govindaraju, B. A. Engel, and M. M. Hantush (2007), Multiobjective sensitivity analysis of sediment and nitrogen processes with a watershed model, *Water Resources Research*, *43*(6), doi:10.1029/2006WR005463.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment, Part I: Model development, *Journal of the American Water Resources Association*, *34*(1), 73–89, doi:10.1111/j.1752-1688. 1998.tb05961.x.

Arnold, J. G., J. R. Kiniry, R. Srinivasan, J. R. Williams., E. B. Haney, and S. L. Neitsch (2011), Soil and Water Assessment Tool Input/Output File Documenta-

tion, Version 2009, Report No. 365, *Tech. rep.*, Texas Water Resources Institute, College Station, Tx.

Bekele, E. G. (2005), Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms, *Water Resources Research*, *41*(10), doi:10.1029/2005WR004090.

Brockhoff, D., and E. Zitzler (2007), Operations Research Proceedings 2006, waldmann, ed., pp. 423–429, Springer Berlin Heidelberg, Berlin, Heidelberg, doi: 10.1007/978-3-540-69995-8.

Brookshire, D. S., and H. R. Neill (1992), Benefit transfers: Conceptual and empirical issues, *Water Resources Research*, *28*(3), 651, doi:10.1029/91WR02590.

Brown, L. C., and T. O. Barnwell (1987), The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual, *Tech. rep.*, EPA document EPA/600/3-87/007. USEPA, Athens, GA.

Chapra, S. C. (2008), *Surface Water-Quality Modeling*, 844 pp., Waveland Pr Inc.

Charnes, A., W. Cooper, and E. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research*, *2*(6), 429–444, doi:10.1016/0377-2217(78)90138-8.

Chelouah, R., and P. Siarry (2000), A continuous genetic algorithm designed for the global optimization of multimodal functions, *Journal of Heuristics*, *6*, 191–213.

Deb, K. (2001), *Multi-Objective Optimization Using Evolutionary Algorithms*, 518 pp., Wiley.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multi-objective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197, doi:10.1109/4235.996017.

Desvousges, W. H., M. C. Naughton, and G. R. Parsons (1992), Benefit transfer: Conceptual problems in estimating water quality benefits using existing studies, *Water Resources Research*, *28*(3), 675, doi:10.1029/91WR02592.

Duch, W., K. Grabczewski, and K. G. abczewski (1999), Searching for optimal MLP, in *IN PROC. 4TH CONF. NEURAL NETWORKS AND THEIR APPLICATIONS*, pp. 65 – 70.

Durand, N., and J.-M. Alliot (1999), A Combined Nelder Mead Simplex and Genetic Algorithm, in *Genetic and Evolutionary Computation Conference GECCO99*, pp. 1–7.

El-Mihoub, T., A. A. Hopgood, L. Nolle, and A. Battersby (2006), Hybrid genetic algorithms - a review, *Engineering Letters*, *13*(2).

Fang, F., and K. W. Easter (2003), Pollution Trading to Offset New Pollutant Loadings A Case Study in the Minnesota River Basin, in *Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting*, Montreal, Canada.

FAO (2001), The economics of conservation agriculture.

Fleischer, M. (2003), The measure of Pareto optima applications to multi-objective metaheuristics, pp. 519–533.

Gantovnik, V. B., C. M. Anderson-Cook, Z. Gürdal, and L. T. Watson (2003), A genetic algorithm with memory for mixed discretecontinuous design optimization, *Computers & Structures*, *81*(20), 2003–2009, doi:10.1016/S0045-7949(03)00253-0.

Garen, D. C., and D. S. Moore (2005), Curve number hydrology in water quality modeling: uses, abuses, and future directions, *Journal of the American Water Resources Association*, *41*(2), 377–388, doi:10.1111/j.1752-1688.2005.tb03742.x.

Goldberg, D. E. (1989), *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, Reading, Massachusett.

Haggard, B. E., A. N. Sharpley, L. B. Massey, N. Hardiman, and K. Teague (2010), HUC 12 Watershed Prioritization Along a Water Quality and Land Use gradient: Upper Illinois River Watershed Management Plan, in *2010 Land Grant and Sea Grant National Water Conference.*

Hansen, L., and M. O. Ribaudo (2008), Economic measures of soil conservation benefits: regional values for policy assessment, *Tech. rep.*, U.S. Dept. of Agriculture,Econ. Res. Serv.

Holland, J. H. (1975), *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, Michigan.

Ingber, L., and B. Rosen (1992), Genetic Algorithms And Very Fast Simulated Reannealing: A Comparison, *Mathematical and Computer Modelling*, *16*(11), 87–100.

Jha, M. K., S. Rabotyagov, and P. W. Gassman (2009), Optimal Placement of Conservation Practices Using Genetic Algorithm with SWAT, *International Agricultural Engineering Journal*, *18*(1-2), 41–50.

Joines, J. A., J. A. Joines, M. G. Kay, and R. E. King (1997), A Hybrid-Genetic Algorithm for Manufacturing Cell Design.

Kanwar, R. S. (2006), Effects of cropping systems on NO3-N losses to tile drain, *Journal of the American Water Resources Association*, *42*(6), 1493–1501, doi: 10.1111/j.1752-1688.2006.tb06016.x.

Lakshmanan, V. (2000), Using a Genetic Algorithm to Tune a Bounded Weak Echo Region Detection Algorithm, *Journal of Applied Meteorology*, *39*(2), 222–230, doi:10.1175/1520-0450(2000)039⟨0222:UAGATT⟩2.0.CO;2.

Lempitsky, V., S. Roth, T. Darmstadt, and C. Rother (2008), FusionFlow: Discrete-Continuous Optimization for Optical Flow Estimation, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society.

Leonard, R., W. Knisel, and D. Still. (1987), GLEAMS: Groundwater loading effects on agricultural management systems, *Transactions of the ASAE*, *30*(5), Transactions of the ASAE.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, *26*(2), 205–216, doi:10.1016/S0309-1708(02)00092-1.

Malone, R., L. Ma, D. Karlen, T. Meade, D. Meek, P. Heilman, R. Kanwar, and J. Hatfield (2007), Empirical analysis and prediction of nitrate loading and crop yield for cornsoybean rotations, *Geoderma*, *140*(3), 223–234, doi: 10.1016/j.geoderma.2007.04.007.

Maringanti, C., I. Chaubey, M. Arabi, and B. Engel (2008), A multi-objective optimization tool for the selection and placement of BMPs for pesticide control, *Hydrology and Earth System Sciences Discussions*, *5*(4), 1821–1862, doi: 10.5194/hessd-5-1821-2008.

Maringanti, C., I. Chaubey, M. Arabi, and B. A. Engel (2011), Application of a Multi-Objective Optimization Method to Provide Least Cost Alternatives for NPS Pollution Control, *Environmental management*, *48*(3), 448–61, doi:10.1007/s00267-011-9696-2.

McElroy, A., S. Chiu, J. Nebgen, A. Aleti, and F. Bennett (1976), Loading functions for assessment of water pollution from nonpoint sources, *Tech. rep.*, EPA 600/2-76-151. USEPA, Athens, GA.

Mostaghimi, S., S. Park, R. Cooke, and S. Wang (1997), Assessment of management alternatives on a small agricultural watershed, *Water Research*, *31*(8), 1867–1878, doi:10.1016/S0043-1354(97)00018-3.

Muleta, M. K., and J. W. Nicklow (2005), Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model, *Journal of Hydrology*, *306*(1-4), 127–145, doi:10.1016/j.jhydrol.2004.09.005.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams (2005), Soil and Water Assessment Tool (SWAT) Theoretical Documentation, Version 2005, (BRC Report 02-05).

Ngatchou, P., A. Zarei, and A. El-Sharkawi (2005), Pareto Multi Objective Optimization, in *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems*, pp. 84–91, IEEE, doi:10.1109/ISAP.2005.1599245.

Nicklow, J. W., P. M. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh, and E. Zechman (2010), State of the art for genetic algorithms and beyond in water resources planning and management, *Journal of Water Resources Planning and Management*, *136*(4), 412, doi:10.1061/(ASCE)WR.1943-5452.0000053.

Novotny, V. (1993), *Water Quality: Diffuse Pollution and Watershed Management*, 2 edition ed., 888 pp., Wiley.

Pareto, V. (1971), *Manual of Political Economy*, 504 pp., Augustus M. Kelley, Publishers.

Piper, S., and W. E. Martin (2001), Evaluating the accuracy of the benefit transfer method: a rural water supply application in the USA., *Journal of environmental management*, *63*(3), 223–35, doi:10.1006/jema.2001.0464.

Pretty, J., C. Brett, D. Gee, R. Hine, C. Mason, J. Morison, M. Rayment, G. Van Der Bijl, and T. Dobbs (2001), Policy Challenges and Priorities for Internalizing the Externalities of Modern Agriculture, *Journal of Environmental Planning and Management*, *44*(2), 263–283, doi:10.1080/09640560123782.

Rabotyagov, S. S., M. Jha, and T. D. Campbell (2010), Nonpoint-Source Pollution Reduction for an Iowa Watershed: An Application of Evolutionary Algorithms, *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, *58*(4), 411–431, doi:10.1111/j.1744-7976.2010.01198.x.

Renders, J. M., and S. P. Flasse (1996), Hybrid methods using genetic algorithms for global optimization., *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, *26*(2), 243–58, doi:10.1109/3477.485836.

Ribaudo, M. O., R. D. Horan, and M. E. Smith (1999), Economics of Water Quality Protection from Nonpoint Sources: Theory and Practice Agricultural Economic Report No. (AER782), *Tech. rep.*, U.S. Dept. of Agriculture, Econ. Res. Serv.

Ritter, W., and A. Shirmohammadi. (Eds.) (2001), *Agricultural Nonpoint Source Pollution- Watershed Management and Hydrology*, 342 pp., Lewis Publishers, Washington, D.C.

Roth, G., and J. Hyde (2002), Partial Budgeting for Agricultural Businesses, *Tech. rep.*, Penn State University. CAT UA366., University Park, PA.

Sinha, A., and D. E. Goldberg (2003), A Survey of hybrid genetic and evolutionary algorithms, *Tech. rep.*, Illinois Genetic Algorithms Laboratory (ILLIGAL) Technical Report 2003004.

Solovay, R. M. (1970), A Model of Set-Theory in Which Every Set of Reals is Lebesgue Measurable, *The Annals of Mathematics*, *92*(1), 1, doi:10.2307/1970696.

Sotiropoulos, D. G., E. C. Stavropoulos, and M. N. Vrahatis (1997), A new hybrid genetic algorithm for global optimization, *Machine Learning and Cybernetics 2005 Proceedings of 2005 International Conference on*, *30*(7), 4529–4538, doi:10.1109/ ICMLC.2005.1527462.

Srinivas, N., and K. Deb (1994), Muiltiobjective optimization using nondominated sorting in genetic algorithms, *Evolutionary Computation*, *2*(3), 221–248.

Tedesco, L., D. Pascual, L. Shrake, R. Hall, L. Casey, P. Vidon, F. Hernly, K. Salazar, and R. Barr (2005), Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality, *Tech. rep.*, IUPUI, Indianapolis.

US EPA (1988), quality criteria for water, *Tech. rep.*, U.S. Environmental Protection Agency, Washington D.C.

US EPA (2008), Handbook for Developing Watershed Plans to Restore and Protect Our Waters, *Tech. rep.*, U.S. EPA, Office of Water, Washington, DC.

USDA NASS (2003), USDA-National Agricultural Statistics Service, Cropland Data Layer.

USDA NRCS (2010), Soil Data Mart.

USDA NRCS (2011), Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Great Lakes Region, *Assessment*, (August).

USGS NED (2010), 1 arc Second Digital Elevation Model.

USGS NLCD (2001), 2001 Land Cover Data.

van Hentenryck, P., and M. Milano (2011), *Hybrid Optimization, Springer Optimization and Its Applications*, vol. 45, Springer New York, New York, NY, doi: 10.1007/978-1-4419-1644-0.

Vecchietti, A., S. Lee, and I. E. Grossmann (2003), Modeling of discrete/continuous optimization problems: characterization and formulation of disjunctions and their relaxations, *Computers & Chemical Engineering*, *27*(3), 433–448, doi:10.1016/S0098-1354(02)00220-X.

Veith, T. L. (2002), Agricultural Bmp Placement For Cost-Effective Pollution Control At The Watershed Level, Phd, Virginia Polytechnic Institute and State University.

Veith, T. L., M. L. Wolfe, and C. D. Heatwole (2003), Optimization procedure for cost effective BMP placement at a watershed scale, *Journal of the American Water Resources Association*, *39*(6), 1331–1343.

Veith, T. L., M. L. Wolfe, and C. D. Heatwole (2004), Cost-effective bmp placement: optimization versus targeting, *Transactions of the ASAE*, *47*(5), 1585–1594.

Walsh, R. G., D. M. Johnson, and J. R. McKean (1992), Benefit transfer of outdoor recreation demand studies, 19681988, *Water Resources Research*, *28*(3), 707, doi:10.1029/91WR02597.

Walter, T., M. Dosskey, M. Khanna, J. Miller, M. Tomer, and J. Wiens (2007), The science of targeting within landscapes and watersheds to improve conservation effectiveness, in *Managing agricultural landscapes for environmental quality*, edited by S. M and C. C, Soil and Water Conservation Society, Ankeny, IA.

White, M. J., D. E. Storm, P. Busteed, S. Stoodley, and S. J. Phillips (2010), Evaluating conservation program success with Landsat and SWAT., *Environmental management*, *45*(5), 1164–74, doi:10.1007/s00267-010-9458-6.

Whittaker, G. W. (2003), Spatial evaluation of alternative nonpoint nutrient regulatory instruments, *Water Resources Research*, *39*(4), 1079, doi:10.1029/2001WR001119.

Williams., J. R. (1990), The Erosion-Productivity Impact Calculator (EPIC) Model: A Case History, *Philosophical Transactions: Biological Sciences*, *329*(1255), 421–428.

Williams, J. R., and R. Hann (1978), Optimal operation of large agricultural watersheds with water quality constraints, *Tech. rep.*, Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

World Bank (2001), Economic Analysis of Incentives for Soil Conservation, in *ncentives in Soil Conservation: From Theory to Practice*, edited by D. Sanders, P. Huszar, S. Sombatpanit, and T. Enters, pp. 41–56, Science Publishers, Enfield, NH.

Xu, F., and T. Prato (1995), Measuring Technical Efficiency Using Data Envelopment Analysis: A Simulation Study of K Films and N Inputs, *Journal of Economics*, *xx*(2), 81–88.

Zitzler, E., and L. Thiele (1999), Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE Transactions on Evolutionary Computation*, *3*(4), 257–271, doi:10.1109/4235.797969.

Chapter 4

HOW *A PRIORI* KNOWLEDGE CAN HELP US IN IDENTIFYING OPTIMAL TYPE AND PLACEMENT OF CONSERVATION PRACTICES? APPLICATION OF MULTI CRITERIA DECISION MAKING

4.1   Highlights

Control of agricultural non-point sources of pollution is achievable through implementation of conservation practices at the farm or field level. There are several approaches to achieve a targeted implementation of conservation practices at the watershed scale. Recent studies have shown that optimization methods hold great promise for optimal allocation of non-point source pollution control measures at the watershed scale. However, the use optimization is a computationally intensive task and ultimately depends upon availability of automated optimization tools and expertise to analyze the results. In this study, a multi criteria decision analysis (MCDA) framework is proposed to identify near-optimal type and location of conservation practices at the watershed scale using *a priori* knowledge about the system. The proposed framework requires: (i) selecting a set of criteria that should be considered in ranking the alternative(s), which depends upon the objectives of the study; (ii) constructing an evaluation matrix; and (iii) a computational MCDA to aggregate the criteria and rank the alternatives. The framework was used to identify optimal placement of four types of conservation practices for nutrient and pesticide load control at minimum cost in the Eagle Creek Watershed, Indiana, USA. Results were

compared with optimal solutions obtained from an optimization framework coupled with the Soil and Water Assessment Tool (SWAT). The results of this study showed that the proposed framework can be an effective and efficient substitute for the optimization frameworks in identifying near-optimal solutions for nonpoint source pollution control. The MCDA framework outperformed the optimization method by identifying the similar solutions with more diversity without any need for iterative search algorithms. For complex problems or poorly established evaluation matrix, the MCDA framework may fail to identify near-optimal solutions; however, it can effectively serve as an ideal initial population in a hybrid MCDA and optimization framework. A hybrid framework substantially improved efficiency of the search algorithm, optimality of the front, and diversity of the solutions. This study also highlighted the importance of the defining proper decision variables and accurate scoring of the conservation practices for a successful watershed planning.

4.2 Introduction

Nonpoint source pollution is the primary cause of impaired water bodies in the United States today [*Horowitz et al.*, 2007]. Elevated nutrient, sediment, and pesticide loads to waterways may negatively impact human health and aquatic ecosystems [*Novotny*, 2003]. Control of agricultural non-point sources (NPS) of pollution is achievable through implementation of conservation practices, commonly known as Best Management Practices (BMPs), at farm or field level. In practice, several approaches have been taken to develop proper watershed scale nonpoint source pollution control strategies at minimum cost. Government agencies promote implementation of NPS pollution control practices by recommending cost-share program that is a field-by-field approach [*Veith et al.*, 2003]. Most of the early applications used cost-share or critical source area (CSA) targeting approaches [*Sullivan and Batten*, 2007; *Ripa et al.*, 2006]; however, more recent studies have utilized optimization methods to identify optimal type and placement of the conservation practices [*Arabi et al.*, 2006; *Maringanti et al.*, 2011]. Cost sharing with landowners is promoted by government agencies for implementation of conservation practices through site investigation, monitoring, and field-scale modeling [*Arabi et al.*, 2006]. The critical source area targeting method, on the other hand, suggests implementation of the conservation practices in critical source areas, which contribute higher amounts of nonpoint source pollutants. The major disadvantages of cost-share and critical source area approaches is their unknown impact on water quality at the watershed scale, primarily because of the site-specific performance of conservation practices. Moreover, they don't consider duplicative impacts of conservation practices on water quality that reduces the potential benefit for nonpoint source pollution control at the watershed scale [*Arabi et al.*, 2006]. Research to date shows that an optimization approach has better performance than other approaches by identifying more cost-effective solutions [*Arabi et al.*, 2006; *Veith*, 2002; *Jha et al.*, 2009; *Rabotyagov et al.*, 2010].

Optimization algorithms should be coupled with a spatially distributed watershed model and use a proper representation of conservation practices to predict water quality changes due to the adoption of conservation plans [*Easton et al.*, 2008].

The primary goal of watershed scale conservation plans is minimizing pollutants movement from landscapes to water bodies at minimum cost. Implementation of conservation practices is technically possible and economically profitable; however, noneconomic factors, such as psychological motivations and social norms, should also be taken into account [*Cary and Wilkinson*, 1997]. Thus, conservation planning is inherently a multiobjective problem. While most of the earlier studies focused on development and application of single-objective optimization methods [*Veith et al.*, 2003; *Arabi et al.*, 2006], more recent efforts have utilized multiobjective approaches [*Jha et al.*, 2009; *Maringanti et al.*, 2008]. Multiobjective optimization algorithms attempt to explore the tradeoffs between incommensurable and often conflicting socioeconomic and environmental factors by finding the optimal type, size, and location of conservation practices within the watershed, resulting in a set of non-dominated (also referred to as "Pareto-optimal") solutions. Although, optimization is an effective tool in identifying optimal set of solutions, it ultimately depends upon availability of the automated optimization tools and expertise on analyzing the results. Moreover, optimization is a computationally intensive task and, depending on the complexity of the problem, may need thousands to millions of model runs.

While optimization algorithms demonstrated successful application in optimal conservation planning, they have mostly ignored the fact that certain portions of the watersheds are more vulnerable to human activities [*Walter et al.*, 2007]. Unlike the optimization approaches, critical source area targeting method received maximum benefit from this fact. Chapter 3 showed that spatial distribution of optimized conservation practices explicitly supports this concept. Furthermore, I showed that in a set of Pareto-optimal solutions obtained from a multiobjective optimization algorithm, moving in the direction of increasing cost, each of the solutions contained

most of the conservation practices from previous solutions. Considering the curvature of the Pareto-solutions, it can be concluded that higher priorities are given to more effective and less expensive conservation practices. Analyzing the type and distribution of several optimized nonpoint source pollution control strategies implied that identifying a set of effective and near-optimal conservation plans is achievable by considering concept of critical source area along with the other information on effectiveness of the conservation practices in the context of multi criteria decision making (MCDM) process. Using a priori knowledge of the system and experts recommendation may provide worthwhile information for near-optimal placement of conservation practices [*Srivastava*, 2002].

Decision making in environmental management typically receives information from several sources in different forms. Therefore, a computational scheme is required to aggregate the information to identify the most preferred alternative. Multicriteria decision analysis (MCDA) provides such a framework by reflecting the complexity and dynamics of the conflicting criteria and sorting the alternatives from the most to the least preferred [*Brown et al.*, 2010]. The following steps should be taken to rank alternatives in the MCDA approach: (i) selecting appropriate evaluation criteria; (ii) assigning values to criteria based on the relative importance in the alternatives; and (iii) ranking the options after aggregating criteria using a mathematical MCDA. Ranks of the alternatives describe how well each alternative meets all the criteria. MCDA is extensively utilized in different aspects of the environmental decision making process [*Hajkowicz and Collins*, 2007; *Mirchi et al.*, 2010]. *Hajkowicz and Collins* [2007] have reviewed the most commonly used MCDA methods water resources management.

In this study, we propose a MCDA framework to identify near-optimal solutions for nonpoint source pollution control at the watershed scale, based on the prior knowledge of the system and effectiveness of conservation practices. Specific

objectives of this paper are (1) to present a computational framework to rank conservation practices type and placement alternatives at the watershed scale based on the socioeconomic and environmental objectives; and (2) to determine how results from MCDA framework compares to the results from application of multiobjective optimization methods for the given problem.

4.3 Methods and Materials

The proposed MCDA framework for near-optimal allocation of NPS pollution control practices include three main component: (1) watershed model for simulation of hydrologic and water quality processes under scenarios with and without NPS pollution control measures; (2) a computational MCDA method to aggregate environmental and socioeconomic factors and rank alternatives from the most to the least favorable; and (3) an economic module for cost-benefit analysis of the implemented conservation plans. The framework was tested for identifying near-optimal conservation plans for minimizing atrazine and nitrated loads at minimum cost in the Eagle Creek Watershed, Indiana, USA. Optimality of the solutions will be tested using optimal conservation practices obtained from application of a simulation-optimization approach proposed in Chapter 3.

4.3.1 Multi Criteria Decision Analysis (MCDA) Framework

Decision making in watershed and water resources management typically receives technical inputs from four types of sources: modeling and monitoring studies, risk assessment, economic analysis, and stakeholders preferences [*Kiker et al.*, 2005]. This information comes in different forms (quantitative or qualitative) and various units (monetary, load/concentration, volume, or dimensionless). MCDA provides a framework to aggregate the information from different sources and rank alternatives from the most to the least favorable [*Brown et al.*, 2010]. In addition to defining

a set of available alternatives, ranking in MCDA method requires: (i) a set of criteria that should be considered in choosing the best alternative(s), which depends on objectives of the study (e.g. implementation cost, environmental benefits, effectiveness, adoptability, and etc); (ii) constructing evaluation matrix, that assigns "scores" for the criteria based on the relative importance of each criterion for each alternative; and (iii) ranking the alternatives after aggregating the criteria using a mathematical MCDA. Pareto-front will then be constructed based on the ranks of alternatives.

4.3.1.1 Conservation Practices Scoring Criteria

Selection of type and placement of conservation practices depends upon several biophysical, agro-environmental, socioeconomic, political, educational, and ethical factors. Different goals and objectives, ownership status, landscape characteristics, and dynamics of the weather conditions can also form a site-specific feature for the cases under study [*Tomer*, 2010; *Kiker et al.*, 2005; *Hansen et al.*, 1987]. Table 4.1 summarizes the important criteria that should be considered for developing conservation plans. A successful nonpoint source pollution control strategy should address these criteria in selection of the proper type and location of conservation practices. Depending on the specific objectives of the study and data availability, analysts may neglect some of them or add new criteria.

4.3.1.2 Constructing An Evaluation Matrix

A MCDA problem with m alternatives and n criteria can be defined by a generic evaluation matrix (EM), P, as [*Hipel*, 1992; *Madani and Lund*, 2011]

$$P_{m \times n} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix} \tag{4.1}$$

Table 4.1: Influential factors on selection of type and placement of NPS pollution control practices, adopted from *Tomer* [2010]; *Lacroix et al.* [2005]; *Kiker et al.* [2005]

Resource-specific features (ecosystem services)	Landscape Spatial features (terrain modeling)	Temporal features (climate forecast)	Behavioral/social features (watershed and farm)
Soil type/quality	Farm/field size	Seasonal climate pattern	Economic incentives
Land use/land cover	Proximity to rivers	Climate change	Monetary benefits of practicess
Crop diversity	Slope/pollutant movement		Willingness to take risks
practicess Effectiveness	Spatial precipitation pattern		Farmers age
			Education level
			Income
			Ownership
			Environmental awareness

where p_{ij} is the value or "score" of alternative i under criterion j. The evaluation matrix may contain a mix of ordinal and cardinal data [*Hajkowicz and Higgins*, 2008]. In the nonpoint source pollution control planning, alternatives are combination of "type" and "location" of conservation practices. For example, in a problem with a fields and b conservation practices, there would be $m = a \times b$ alternatives available for decision makers. The importance of each criterion should also be given in a one-dimensional weights vector, W:

$$W_{n \times 1} = [w_1, \cdots, w_n]^T \tag{4.2}$$

where $\sum_{j=1}^{n} w_j = 1, 0 \leq w_j \leq 1$, and w_j denotes the weight assigned to the jth criterion. Ranking of the alternatives is highly sensitive to the weights assigned to the criteria and should be assigned precisely according to the specific objectives of the study. For example if criterion i is 3 times more important than criterion j, then $w_i = 3w_j$.

4.3.1.3 Mathematical MCDA Techniques

The ultimate goal of MCDA methods is to identify the rank order of the alternatives, R,

$$R = f(P, W) = [r_1, \cdots r_n] \tag{4.3}$$

where r_j is the rank of the jth alternative [*Hajkowicz and Higgins*, 2008]. Several MCDA techniques are available to perform mathematical MCDA computations. Weighted Summation (WS) is the most simple and widely applied technique of MCDA [*Howard*, 1991]. The Weighted Summation ranks alternatives on the basis of utility score u_k determined by

$$u_k = \sum_{j=1}^{n} p_{kj} w_j \tag{4.4}$$

106

where $k = 1, \cdots, m$ and $P_{kj} = [p_{k1}, p_{k2}, \cdots, p_{kn}]$. Alternatives with higher utility scores rank higher, and vice versa. Simplifying assumptions in the Weighted Summation method can potentially lead to inaccurate results [*Rowe and Pierce*, 1982]; however, Hajkowicz and Higgins [*Hajkowicz and Higgins*, 2008] reported strong agreement between results of Weighted Summation and other MCDA methods. Study of different MCDA methods is beyond the scope of this study and may require further investigation. Furthermore, when more than one decision maker is involved in decision making, a group MCDA method could be utilized that is a general form of generic MCDA methods.

After ranking the n alternatives (from r_1 to r_n), the near-optimal front for multiple objectives can then be constructed by a successive addition of alternatives starting from the higher ranks. Figure 4.1 depicts schematic of the near-optimal front constructed after ranking the alternatives using a MCDA framework. In this figure, objective functions are subsets or functions of the selected criteria. If scores are assigned properly, MCDA may guarantee near-optimality of the solutions by assigning higher ranks to more favorable conservation plans. Several environmental and economic objectives should be considered in targeting of nonpoint source pollution control practices. Water quality impacts of the conservation practices could be evaluated using a distributed watershed models, such as Soil and Water Assessment Tool (SWAT).

4.3.2 Optimization of Conservation Plans

Identifying the best compromise solution between multiple incommensurable and often conflicting objectives in nonpoint source pollution control strategy is naturally a multiobjective problem. To verify optimality of the results from the MCDA framework, a Nondominated Sorted Genetic Algorithm II (NSGA-II) was used to identify a set of optimal type and placement of conservation practices for the given

107
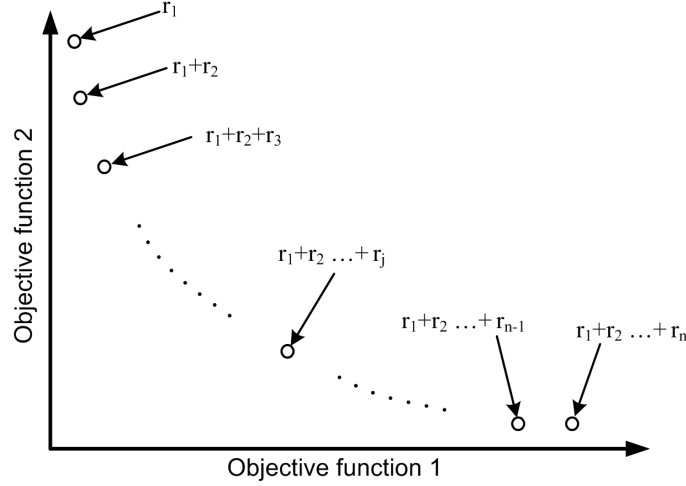
Figure 4.1: Schematic of Pareto-front constructed using MCDA framework. r_j is the conservation practice ranked jth

problem. Detailed description of the optimization framework, objective functions, and selection of the decision variables are presented in Chapter 3. The framework developed in Chapter 3 incorporates water quality benefits expressed in terms of the reduction of sediment, nutrient, and pesticide fluxes, onsite and offsite costs and benefits, and water quality vulnerability. The proposed optimization approach for targeting NPS pollution control practices includes the following objective functions and constraint functions:

$$
\begin{cases}
\text{minimize } y = f(\boldsymbol{x}|\boldsymbol{\theta}, I, c_s, t_d, T) & \text{pollutant load(s)} \\
\text{minimize } C = g(\boldsymbol{x}|\boldsymbol{\theta}, I, p, c_s, t_d, T) & \text{cost(s)} \\
\text{maximize } \pi = h(\boldsymbol{x}|\boldsymbol{\theta}, I, r, c_s, t_d, T) & \text{profit(s)}
\end{cases}
\tag{4.5}
$$

subject to:

$$
\begin{cases}
y \leq y_{allowable} & \text{water quality targets} \\
C \leq C_{max} & \text{economic constraints} \\
N(\boldsymbol{x}|m, s) \geq 0 & \text{management and social constraints}
\end{cases}
$$

where $\boldsymbol{x}$ denotes discrete-continuous decision variables reflecting the type and location of practices, y represents pollutant load(s) over the assessment period T estimated by the watershed model f, $\boldsymbol{\theta}$ is the vector of calibrated watershed model

parameters, I represents input forcing, t_d is the design lifetime of conservation practices, C represents the total on-site and off-site cost of implementation of nonpoint source pollution control practices over the assessment period T, p is the unit cost of conservation practices, π is the total on-site and off-site profit computed for the conservation strategy over the assessment period T, r is the unit price of beneficial products of conservation strategy. The term $y_{allowable}$ indicates water quality standard, C_{max} is the maximum available budget for implementation of the nonpoint source pollution control strategy $\boldsymbol{x}$ The term N represents management preferences (m) and social values (s), and.

Pareto-front solutions of mixed-variable and binary-variable optimization algorithms will be compared by means of hypervolume values estimated based on the Lebesgue measure of the nondominated solutions. The Lebesgue measure, also called n-dimensional volume (Euclidean space), is the standard way of mapping a set of Pareto optimal points to a scalar [*Solovay*, 1970; *Fleischer*, 2003]. For n= $1, 2,$ or 3, it coincides with the standard measure of length, area, or volume. Lebesgue can be measured with respect to a lower or upper bound reference point for maximization or minimization problems. For a minimization problem with upper bound reference point, a better optimal front have higher Lebesgue measure. The Lebesgue measure has also the advantage of taking diversity of the solutions into account [*Fleischer*, 2003]. Figure 3.3 shows schematic of a Lebesgue measure calculated for a two-objective minimization problem.

4.4 Application of the MCDA Approach for Targeted Implementation of Agricultural Conservation Practices

The proposed MCDA framework was applied for the targeted implementation of agricultural conservation practices for nutrient and pesticide control in the Eagle Creek Watershed, Indiana, USA. The Soil and Water Assessment Tool (SWAT) was used to simulate hydrologic and water quality processes. The model was first

calibrated and tested for a 10-year period from 1994 to 2003 for daily streamflow, monthly nitrate loads and monthly atrazine loads at multiple locations within the watershed. Then, the calibrated model was used to evaluate water quality benefits of conservation scenarios obtained from MCDA procedures. Since water quality standards are typically expressed in terms of average annual loads (or concentrations) for nutrients and pesticides, water quality benefits of management actions were calculated as the average annual reduction of loads. Results from the coupled simulation-optimization approach were also used to verify optimality of the results from the MCDA approach. Both binary and discrete-continuous decision variables (as described in Chapter 3) were examined. The optimization procedures were implemented using a population size of 108, a crossover probability of 0.5, and a mutation rate of 0.005. Objectives of the MCDA and optimization approaches were to simultaneously (i) minimize nitrate loads, (ii) minimize atrazine loads, and (ii) minimize cost of implementation. The results from the MCDA approach were compared with results from the optimization approach.

4.4.1 Study Area

The Eagle Creek Watershed (ECW), located in central Indiana, has a drainage area of 248.1 km^2 and lies within the Upper White River Watershed, extending from 40°01'24" to 40°04'16" north latitudes and 86°15'43" to 86°16'45" west longitudes. According to the 2001 national Land Cover Dataset [*USGS NLCD*, 2001], the watershed consists of 52% croplands, 27% pasture, 12% low and high density urban areas, and 9% forest. The predominant crops are corn and soybeans. ECW drains into Eagle Creek Reservoir, which supplies drinking water for the city of Indianapolis. Figure 4.2 presents the location and land cover for the watershed. The soils are generally poorly draining and developed from glacial materials with thin loess over loamy glacial till and alluvial materials depositions. The dominant soils

are the Crosby-Treaty-Miami in the headwaters and Miami-Crosby-Treaty along the downstream areas. The mean annual precipitation for the Eagle Creek Watershed area is 1052 mm. Monthly mean temperatures for this area from 1971-2000 shows January as having the lowest average temperature of -3.3°C and July as the being the warmest month with an average temperature of 23.7°C [*Tedesco et al.*, 2005].

Spatial scale of conservation planning depends upon numerous factors, including management objectives, available data resolution, dominant ecological processes, and potential sociopolitical constraints [*Walter et al.*, 2007; *Garen and Moore*, 2005]. To achieve the specific goals of water quality control, conservation practices targeting should be performed within a smaller geographic unit which ultimately allows better evaluation of the targeted management plan. Using 12-digit Hydrologic Unit Codes (HUCs) is recommended for the watershed-scale evaluation and planning of conservation practices to encourage participation of stakeholders in implementing conservation practices and enable monitoring water quality changes within subwatershed [*Haggard et al.*, 2010]. The ECW (shown in Figure 4.2) encompasses five 12-digit HUCs. The HUC 051202011102 subwatershed with a drainage area of 41.2 km^2 was selected for the application of the integrated simulation-optimization framework in this study. The subwatershed consists of predominantly cropland (88%), and also contributes the largest amounts of non-point source nitrate and atrazine loads from agricultural lands to the Eagle Creek Reservoir (more than 23% and 28% of the total watershed nitrate and atrazine loads, respectively).

4.4.2 Watershed Simulation Model Description

The Soil and Water Assessment Tool (SWAT) [*Arnold et al.*, 1998] was used to represent H/WQ processes in the Eagle Creek watershed. Hydrologic processes simulated by SWAT include snow accumulation and melt, evapotranspiration, infiltration, percolation losses, surface runoff, and groundwater flows [*Neitsch et al.*,
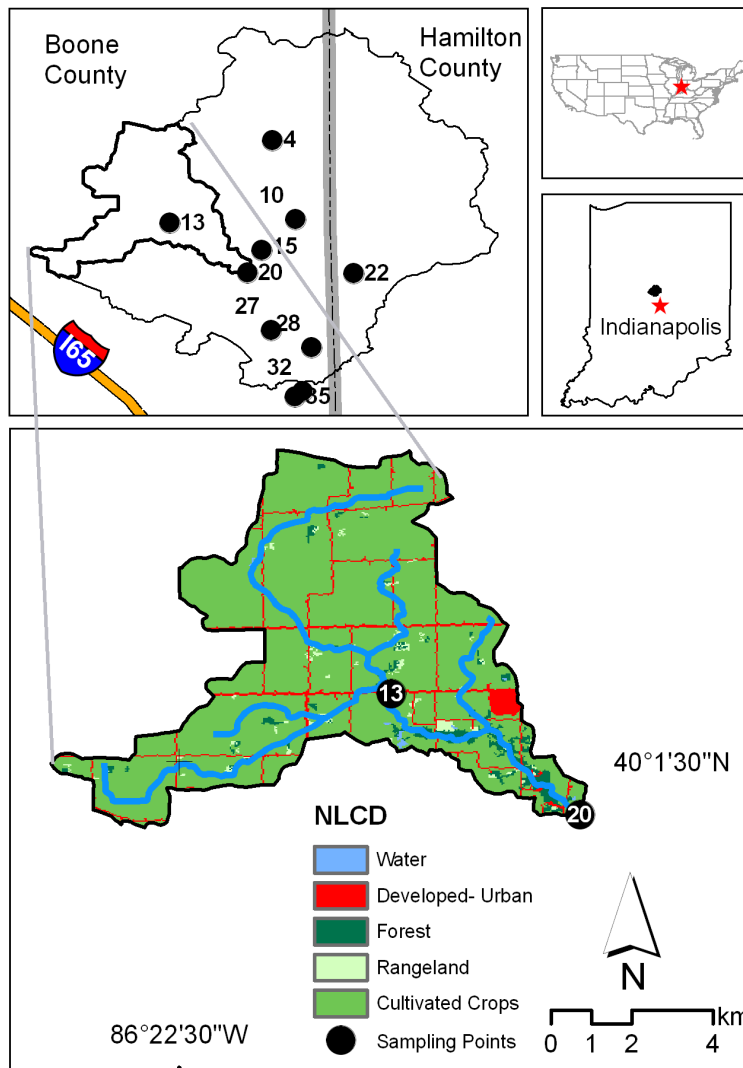
Figure 4.2: Location and landuse maps of the Eagle Creek Watershed, HUC 051202011102

2005]. SWAT is a physically-based watershed-scale, distributed-parameter, continuous time, and long-term, model that runs on a daily time step. It subdivides a watershed into subbasins connected by a stream network, and further delineates hydrologic response units (HRUs) consisting of unique combinations of land cover and soils in each subbasin.

SWAT can simulate major nutrient processes within a watershed. The nitrogen (N) cycle is simulated in five pools: inorganic (including ammonium and nitrate) and organic (including fresh, stable, and active). The main N processes are mineralization, decomposition, and immobilization. Nutrients are introduced into the main channel through surface runoff and lateral flow and transported downstream with channel flow. Plant uptake, denitrification, volatilization, leaching, and soil erosion are the major mechanisms of N removal from a field. The transport rate of organic N with sediment is calculated with a loading function developed by *McElroy et al.* [1976] and modified by *Williams and Hann* [1978] for application to individual runoff events. The loading function estimates daily organic N runoff loss based on the concentrations of constituents in the top soil layer, sediment yield, and an enrichment ratio. Nutrient transformations in the stream are controlled by in-stream water quality component of the model that is adapted from the QUAL2E in-stream water quality model [*Brown and Barnwell*, 1987]. More detailed description of the nutrient components of SWAT can be found in [*Neitsch et al.*, 2005]. SWAT uses algorithms from GLEAMS (Ground Water Loading Effects on Agricultural Management Systems) [*Leonard et al.*, 1987] and EPIC (Erosion Productivity Impact Calculator) [*Williams.*, 1990], to model pesticide's overland fate and transport and movement from land to streams. It also incorporates a simple mass-balance method developed by *Chapra* [2008] to model the transformation and transport of pesticides in streams.

A 30-m resolution DEM from USGS National Elevation Dataset [*USGS NED*, 2010], National Land Cover Dataset (NLCD) 1992 and 2001 [*USGS NLCD*, 2001] for

Table 4.2: SWAT performance criteria for daily streamflow, monthly nitrate, and monthly total pesticide simulation during 1993-2004 including 2 years of warmup period. PBIAS and NSE indicate percent bias and Nash-Sutcliff Efficiency coefficient, respectively.

Gauging Station	Constituent	Calibration Period Criteria (1995-1999)			Validation Period Criteria (2000-2003)		
		PBIAS (%)	R^2	NSE	PBIAS (%)	R^2	NSE
20	Nitrate	7.9	0.94	0.83	16.9	0.85	0.67
	Atrazine	-6	0.81	0.34	-14	0.7	0.41
22	Nitrate	-22.3	0.89	0.78	1.24	0.74	0.36
	Atrazine	42	0.69	0.44	-0.1	0.5	0.28
27	Nitrate	0.59	0.93	0.85	18.3	0.78	0.59
	Atrazine	13	0.66	0.35	-30	0.51	0.19
32	Nitrate	-7.9	0.92	0.84	8.4	0.76	0.55
	Atrazine	42.3	0.75	0.52	33.1	0.51	0.14
35	Streamflow	-12.2	0.78	0.61	4.3	0.78	0.56

urban areas, National Agriculture Statistics Service (NASS) Cropland Data Layer 2000-2003 [*USDA NASS*, 2003] for croplands, and SSURGO data from national resources conservation service (NRCS) [*USDA NRCS*, 2010] were used for watershed subdivision and delineating HRUs in the SWAT model. The ECW was subdivided into 35 subwatersheds and a total of 446 hydrologic HRUs. HUC 051202011102 has 4 subwatersheds and 40 HRUs that encompasses 23 raw crop (corn and soybean in this case) HRUs. Watershed model was calibrated and validated for streamflow at watershed outlet (outlet 35), and nitrate and Atrazine at gauging stations 20, 22, 27, and 32 to create a plausible model for the study (Chapter 2). Table 4.2 presents performance indices of the SWAT model in predicting daily streamflow and monthly nitrate, and Atrazine.

4.4.3 Representation of Conservation Practices

In this study, water quality impacts of fertilizer management, grassed waterways, grade stabilization structures, and tillage/residue management were evaluated. Only row crops, including corn and soybean, were considered for implementation of the nonpoint source pollution control plan. Conservation practices were

represented using numerical procedures from the published studies [*Arabi et al.*, 2004; *Arnold et al.*, 2011; *White et al.*, 2010]. SWAT includes explicit functions for representation of fertilizer management and tillage/residue management, grassed water ways [*Arnold et al.*, 2011]. Table 4.3 summarizes the relevant SWAT management operations and parameters and their corresponding values for representation of BMPs. Implementation of grade stabilization structures are only considered in combination with grassed waterways. Conservation practices in the binary optimization approach were represented by default values presented in Table 4.3. Detailed description of the conservation practices representation methods can be found in *Arabi et al.* [2004], *Arnold et al.* [2011], and *White et al.* [2010]. In the mixed-variable genetic algorithm, continuous decision variables for fertilizer management and discrete decision variables for grassed waterways, grade stabilization structure, and tillage/residue management were considered and represented as shown in Table 4.4. Implementation cost, operation & maintenance rate, and design life for conservation practices are also summarized in Table 4.4. The interest rate of 4 percent was also used in calculation of total cost of the nonpoint source pollution control plan. It was assumed that more than one conservation practice can be applied in the same field; and therefore, considering 23 raw crop HRUs were obtained $23 \times 3 = 69$ alternatives in this study.

4.4.4 Decision making Criteria and Evaluation Matrix

To reflect specific objectives of the study in reducing atrazine and nitrate loads at minimum cost, three environmental and economic criteria were selected accordingly: (1) effectiveness of conservation practices on atrazine loads, (2) effectiveness of conservation practices on nitrate loads, and (3) implementation cost of conservation practices. We assigned performance scores for conservation practices in reduction of atrazine and nitrate loads in the range of 1-5, based on our prior knowledge, as shown

Table 4.3: SWAT parameter values for representation of conservation practices

Practice	Representation Parameters	Unit	Value
Fertilizer management	Fertilizer application rate reduction, FRT_KG (*.mgt)	%	20
Grassed Waterways	Manning's n, GWATN (*.ops)		0.35
	Width, GWATW (*.ops)	m	15
	Length, GWATL (*.ops)	km	_[A]
Grade Stabilization	Reduce slope steepness, CH_S1 (*.sub)	m	_[B]
Tillage/Residue Mgt.	Reduce curve number, CN2 (*.mgt)		2
	Harvest Efficiency, HARVEFF (*.mgt)		1 [C]
	Overland Manning's number, OV_N (*.hru)		0.1 [D]

[A] Estimated using equation $0.5 \times (area)^{1/2}$

[B] Estimated based on the height of structure (default is GSSH = 1.2 meters) with the minimum of 0

[C] Estimated based on the percent cover for different types of tillage. HARVEFF for conventional, conservation, and no-till is assumed to be 1, 01, and 0.5, respectively

[D] For conventional, conservation, and no-till is assumed to be assumed to be 0.1, 0.2, and 0.3 respectively

in Table 4.5. Conservation practices were assumed to be insensitive to landuse, soil types, and climatic conditions and therefore the same type of conservation practices received the similar score regardless of the biophysical conditions. However, if enough information was available, different scores would be preferred. For economic criteria, implementation cost of conservation practices for each alternative was estimated and then normalized in a way that the least and the most expensive ones were respectively scored 5 and 1 and other alternatives received decimal scores in the range of 1-5. All three criteria were given equal weights, i.e. $W = [1/3, 1/3, 1/3]$, to emphasis equal importance of the criteria in decision making process.

4.4.5 Results and Discussion

Overall, the results of this study suggest that a MCDA framework in nonpoint source pollution control planning, if implemented properly, can effectively and efficiently identify Pareto-optimal front (or near-optimal front) for multiple objectives. Success of the MCDA framework was demonstrated using both binary and mixed decision variables. Both MCDA and optimization approaches were used to identify the optimal nonpoint source pollution control plans within the watershed. Figure 4.3 compares the solutions obtained from proposed MCDA approach and the

Table 4.4: SWAT parameter values for representation of conservation practices in mixed discrete-continuous framework

Practice	Representation Parameters	Unit	Value	Cost	r_{OM}[A]	t (years)[B]
Fertilizer management	Fertilizer application rate reduction, FRT_KG (*.mgt)	%	20			
Grassed Waterways	Manning's n, GWATN (*.ops)		0.35	11,400 ($/ha)	3	10
	Width, GWATW (*.ops)	m	15			
	Length, GWATL (*.ops)	km	_[C]			
Grade Stabilization	Reduce slope steepness, CH_S1 (*.sub)	m	_[D]	4,800 ($/structure)	2	15
Tillage/Residue Mgt.	Reduce curve number, CN2 (*.mgt)		2		0	1
	Harvest Efficiency, HARVEFF (*.mgt)		1	75 ($/ha)		
	Overland Manning's number, OV_N (*.hru)		0.1			

[A] Operation and maintenance rate, percentage of capital cost

[B] design lifetime of conservation practices

[C] Estimated using equation $0.5 \times (area)^{1/2}$

[D] Estimated based on the height of structure (default is GSSH = 1.2 meters) with the minimum of 0

Table 4.5: Performance matrix of BMPs in pollutants load reduction

BMP	Atrazine load reduction	Nitrate load reduction
Fertilizer management	1	5
Grassed waterways	4	4
Residue/tillage management	5	2

multiobjective genetic algorithm framework with binary decision variables. While the MCDA approach identified near-optimal front in 69 SWAT model runs, binary NSGA-II terminated at 116 generations that is equivalent to 12,528 SWAT model runs. Thus, MCDA framework used 99.5 percent less SWAT runs to identify the near-optimal front for the given problem. The MCDA and optimization approaches identified very similar front in reduction of atrazine loads and comparable solutions in removal of nitrate loads. In particular, MCDA resulted in more diverse solutions in the "Cost-Atrazine Load reduction" plane with a higher number of solutions on its breakpoints (Figure 4.3-b). For nitrate load reduction, however, both methods identified several solutions with negative impact on nitrate load. This can be described by the well-known impact of tillage practices on increased drainage and subsequently elevated nitrate leaching into the shallow aquifers and tile drain systems that eventually flows into the surface waters [*Kanwar*, 2006; *Malone et al.*, 2007]. While the optimization framework automatically searched to eliminate some of the solutions with negative impact on nitrate loads, there was not such a mechanism with the MCDA framework.

Lebesgue measures of the Pareto-fronts from the MCDA and optimization approaches, measured with a reference to $600,000 cost, 0% atrazine load reduction, and 30% nitrate load reduction, was respectively 3.86×10^8 and 4.54×10^8 meaning that for the same amount of money spent on conservation plan, solutions from optimization framework were more effective as compared to the solutions from the MCDA approach. The lower Lebesgue measure of the MCDA results is mainly
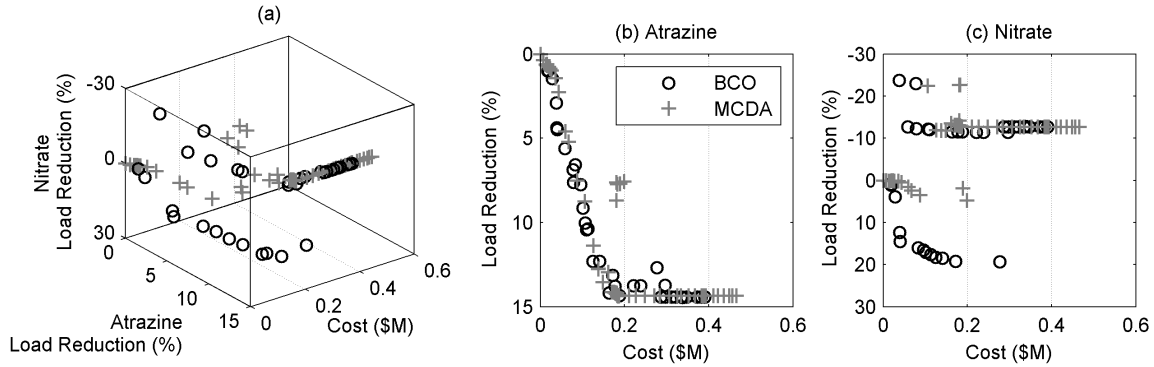
Figure 4.3: Comparison of solutions of MCDA approach and Pareto-optimal front of a binary NSGA-II (BCO) algorithm

because of the poor performance of the MCDA approach in identifying diverse solutions for nitrate load reduction. On the other hand, the MCDA framework is a more appropriate method, since it was based on our prior knowledge of the system and did not use iterative search algorithm to identify the solutions.

Hybridized MCDA and the genetic algorithm, on the other hand, resulted in more promising solutions. In this approach, the genetic algorithm initialized using results from the MCDA approach and terminated at the 37th generation, which means 3 times faster convergence than the standalone genetic algorithm. The Pareto-optimal fronts from the standalone and hybrid genetic algorithm are shown in Figure 4.4. Fast convergence of the hybrid approach was mainly because of diversity in its initial population. In spite of having less coverage of the objective space, the solution from hybrid approach had a larger Lebesgue measure, 4.67×10^8, which demonstrates the improved performance of the optimization framework. In addition, the standalone optimization method had no solution that dominates any of the solutions from the hybrid framework.

Investigating spatial distribution of the conservation practices associated with results from the hybrid framework disclosed another advantage of using this approach. Figure 4.5 demonstrates the type and location of the conservation practices

119
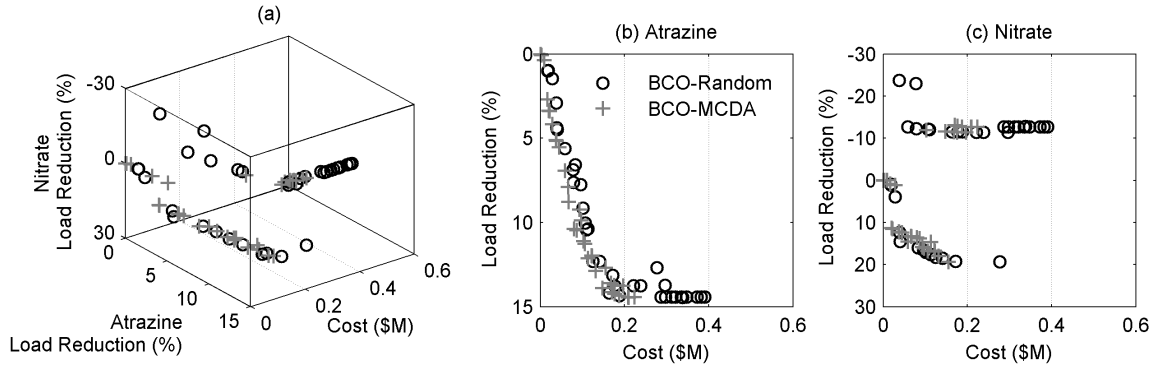
Figure 4.4: Comparison of Pareto-optimal solutions of a binary-variable NSGA-II (BCO) with random and MCDA-based initialization

associated with several scenarios on Pareto-front. Comparing spatial distribution of the results from the hybrid framework and standalone optimization framework (shown in Figure 3.8), demonstrated more uniformly distributed conservation practices regarding the type and geographical location of the conservation. This provides an opportunity for modular implementation of the watershed scale conservation plans, as discussed in chapter 3. Results also show that optimal plans only contained a few residue management practices, which is mostly because of their inverse impact on nitrate loads emerging to the streams.

While binary NSGA-II demonstrated successful performance in optimizing conservation plans, many real world decision making problems consist of continuous and discrete decision variables. For example, length of stabilized river bank, size of ponds and wetlands, and percentage decrease in application of chemicals can be considered as continuous decision variables, although discrete decision variable might be more preferable for width of grassed waterways and height of grade stabilization structures. Further improvement in the initial and Pareto-optimal solutions was achieved by employing discrete-continuous variable (also referred to as "mixed-chromosome" decision variables). Initial population and Pareto-optimal solutions from the genetic algorithm and hybrid frameworks with mixed-variables are presented in Figure 4.6.

120
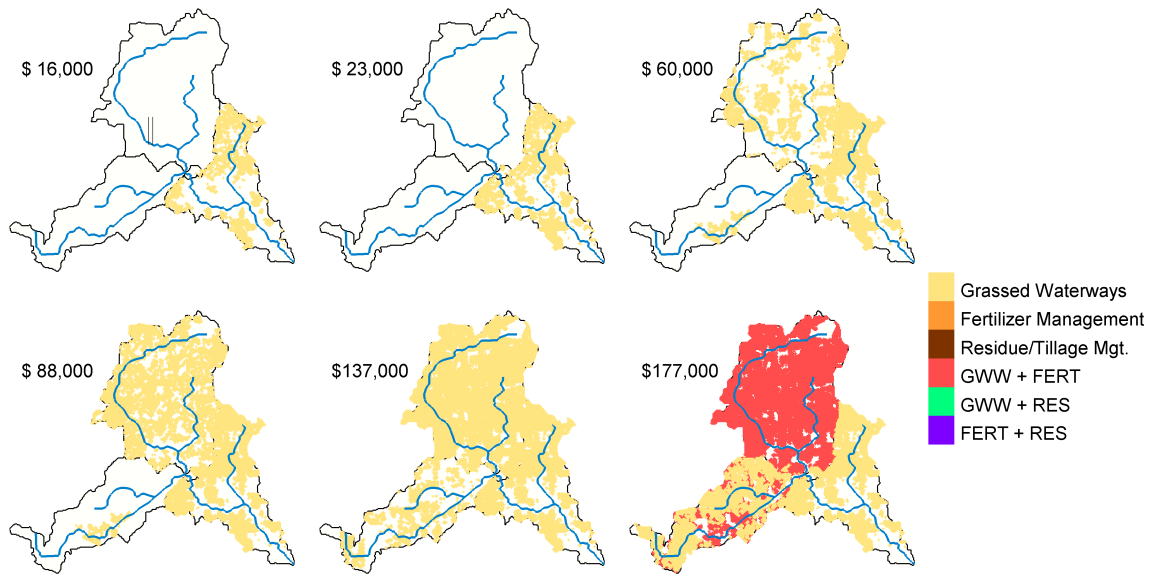
Figure 4.5: Spatial distribution of conservation practices selected for different costs for MCDA-based initialization scheme

Initial population of the hybrid framework, in addition to better fitness values and higher diversity, contains considerably higher number of nondominated solutions than the standalone genetic algorithm (67% versus 9%). Moreover, the standalone genetic algorithm and hybrid frameworks terminated in 6,380 and 44 generations, respectively, that means 145 times faster convergence of the hybrid framework. For an 8-core CPU, the difference was overall runtime of 28 days versus 4.5 hours. Performance of the hybrid framework can be actually seen as a fine tuning of the MCDA results.

Another interesting observation in optimization convergence, that is theoretically expected, was the changes in Pareto-front shape from a relatively smooth curve to a sharp curve as the optimization proceeds. In sharp Pareto fronts, breakpoint(s) are often considered as a good compromise between objectives [*Madsen*, 2003]. The hybrid framework outperformed the standalone genetic algorithm in identifying breakpoint solutions. In this study, a sharp Pareto-optimal front reflects
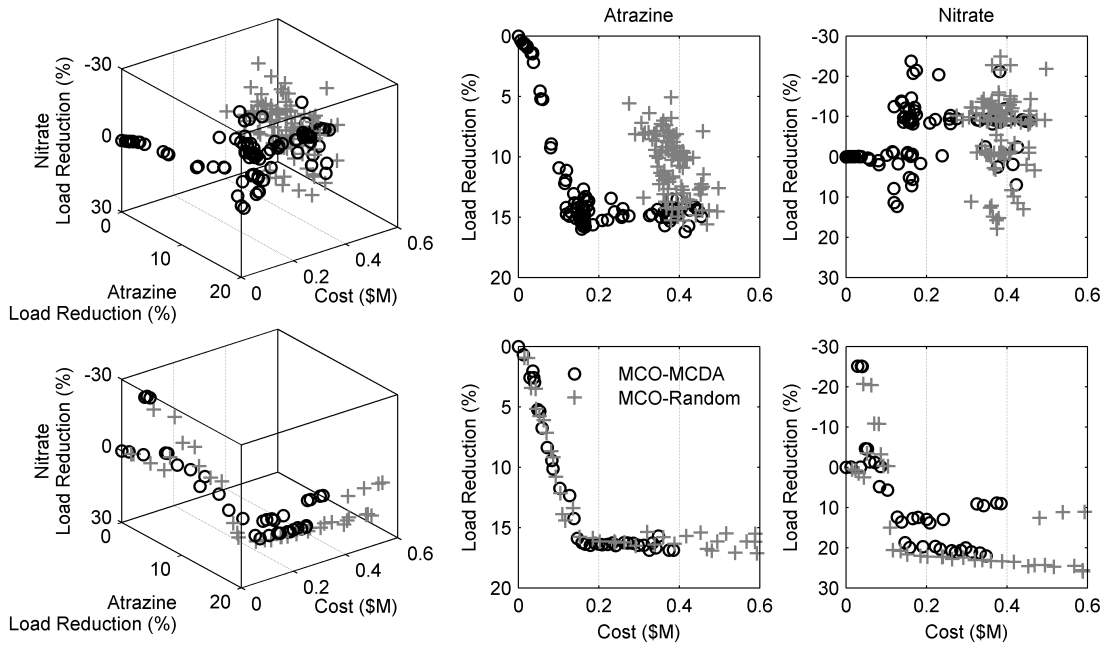
Figure 4.6: Comparison of initial population (top) and Pareto-front solutions (bottom) from a mixed-variable optimization frameworks with random initial population (MCO-Random) and MCDA-based initial population (MCO-MCDA)

the fact that the greatest degree of pollutant load reduction is achievable through the application of limited number of conservation practices with lower costs.

It should also be noted that success of the MCDA framework, however, depends on the relative importance scores assigned to the alternatives in the evaluation matrix. Figure 4.7 compares initial population generated using three different scoring methods: (i) "good scoring" that was based on the proper knowledge of the system as presented in Table 5; (ii) "equal scoring" in which all conservation practices received the same score; and (iii) "inverse scores" that was the opposite of the "good scoring" (i.e. the conservation practices scored 5 and 4 in a 1-5 scale in "good scoring" scored 1 and 2 in "inverse scoring", respectively). Inverse scoring resulted in a set of initial population far away from the Pareto-front and did not provide a good estimate of the nondominated front. The initial population from equal weighting
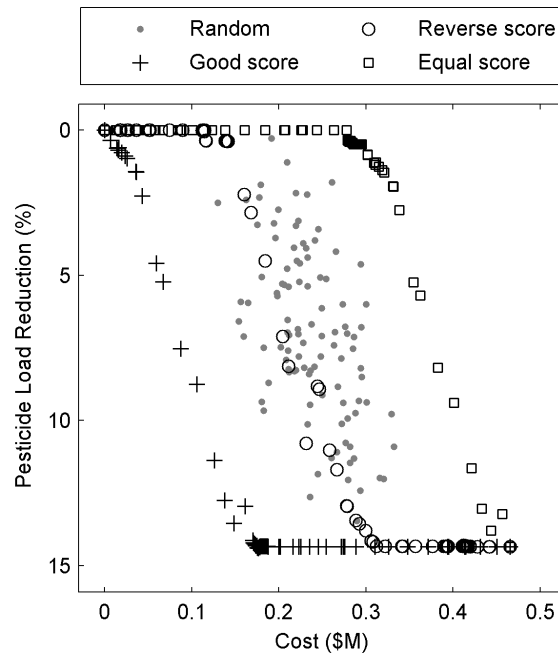
Figure 4.7: Comparison of initial population using three different weighting schemes

was also far from optimality; although, it provided better initial solutions than the random population in terms of the diversity of solutions.

4.5 Summary and Conclusions

Control of agricultural non-point sources of pollution is achievable through implementation of conservation practices, commonly known as Best Management Practices (BMPs), at farm or field level. In practice, watershed plans for non-point source pollution control can be achieved using several approaches. Research to date showed that optimization approach outperforms other strategies regarding the both environmental and socioeconomic factors. Optimization of the nonpoint source pollution control strategy at the watershed scale aims at prevention of water quality degradation at the minimum cost. Although, optimization is an effective computational tool, it ultimately depends upon the availability of the automated

optimization tools and expertise on analyzing the results and is a computationally intensive task.

In this paper, a multi criteria decision analysis (MCDA) framework was proposed to identify a set of near-optimal solutions for nonpoint source pollution control planning at the watershed scale. The framework was established based on a priori knowledge about the system and effectiveness of the conservation practices on reducing pollutants load. Selection of the type and location of conservation practices depends upon several socioeconomic and environmental criteria. MCDA provided a framework to aggregate these incommensurable and often conflicting criteria and rank the alternatives from the most to the least preferred. Ranking the alternatives in MCDA framework requires: (i) selecting a set of criteria that should be considered in ranking the alternative(s), which depends on objectives of the study (e.g. cost, effectiveness, adoptability, and etc); (ii) constructing an evaluation matrix, that assigns "scores" for the criteria based on the relative importance of each criterion for each alternative; and (iii) ranking the alternatives after aggregating the criteria using a mathematical MCDA.

The framework was demonstrated for identifying optimal types and locations of conservation practices in the Eagle Creek Watershed, Indiana. The goal was to find a set of optimal type and location of four types of conservation practices, including grassed waterways, grade stabilization, fertilizer management, and tillage/residue management, to minimize nitrate and atrazine loads at minimum cost. Thus, three criteria were selected according to the objectives of the study that include: (1) effectiveness of conservation practices on nitrate load reduction, (2) effectiveness of conservation practices on atrazine load reduction, and (3) implementation cost of conservation practices. Appropriate scores were assigned to the alternatives based on the performance of conservation practices in reducing pollutants load. The results from the proposed framework were compared with results of an optimization

124

framework for the given problem. The hybridization of the methods was also examined to further improve the results. Both MCDA and optimization frameworks were coupled with the Soil and Water Assessment Tool (SWAT) to simulate the impact of conservation plans on water quality.

The study revealed that the proposed MCDA framework can be an effective and efficient alternative for optimization frameworks in identifying near-optimal solutions for nonpoint source pollution control plans. The MCDA framework outperformed the optimization framework by identifying more diverse solutions within a marginal proximity of the Pareto-optimal front without any need for iterative and computationally intensive search algorithms. For complex problems or in the case of a poorly established evaluation matrix, the MCDA framework may fail to identify near-optimal solutions; however, it can effectively serve as ideal initial population in a hybrid framework. The hybrid framework outperformed the standalone optimization framework in terms of convergence (i.e. efficiency), optimality (i.e. effectiveness), and diversity of the solutions. In addition, results from different scorings of conservation practices highlighted the importance of using credible information in establishing the evaluation matrix.

Bibliography

Arabi, M., R. S. Govindaraju, and M. M. Hantush (2004), Impact of Best Management Practices on Water Quality of Two Small Watersheds in Indiana: Role of Spatial Scale, *Tech. rep.*, US EPA, National Risk Management Research Laboratory, Cincinnati, OH.

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

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment, Part I: Model development, *Journal of the American Water Resources Association*, *34*(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x.

Arnold, J. G., J. R. Kiniry, R. Srinivasan, J. R. Williams., E. B. Haney, and S. L. Neitsch (2011), Soil and Water Assessment Tool Input/Output File Documentation, Version 2009, Report No. 365, *Tech. rep.*, Texas Water Resources Institute, College Station, Tx.

Brown, K., W. L. Hall, M. H. Snook, and K. Garvin (2010), *Sustainable Land Development and Restoration: Decision Consequence Analysis*, 528 pp., Butterworth-Heinemann.

Brown, L. C., and T. O. Barnwell (1987), The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual, *Tech. rep.*, EPA document EPA/600/3-87/007. USEPA, Athens, GA.

Cary, J. W., and R. L. Wilkinson (1997), Perceived profitability and farmers conservation behaviour, *Journal of Agricultural Economics*, *48*(1-3), 13–21, doi: 10.1111/j.1477-9552.1997.tb01127.x.

Chapra, S. C. (2008), *Surface Water-Quality Modeling*, 844 pp., Waveland Pr Inc.

Easton, Z. M., M. T. Walter, and T. S. Steenhuis (2008), Combined monitoring and modeling indicate the most effective agricultural best management practices., *Journal of environmental quality*, *37*(5), 1798–809, doi:10.2134/jeq2007.0522.

Fleischer, M. (2003), The measure of Pareto optima applications to multi-objective metaheuristics, pp. 519–533.

Garen, D. C., and D. S. Moore (2005), Curve number hydrology in water quality modeling: uses, abuses, and future directions, *Journal of the American Water Resources Association*, *41*(2), 377–388, doi:10.1111/j.1752-1688.2005.tb03742.x.

Haggard, B. E., A. N. Sharpley, L. B. Massey, N. Hardiman, and K. Teague (2010), HUC 12 Watershed Prioritization Along a Water Quality and Land Use gradient: Upper Illinois River Watershed Management Plan, in *2010 Land Grant and Sea Grant National Water Conference.*

Hajkowicz, S., and K. Collins (2007), A Review of Multiple Criteria Analysis for Water Resource Planning and Management, *Water Resources Management*, *21*(9), 1553–1566, doi:10.1007/s11269-006-9112-5.

Hajkowicz, S., and A. Higgins (2008), A comparison of multiple criteria analysis techniques for water resource management, *European Journal of Operational Research*, *184*(1), 255–265, doi:10.1016/j.ejor.2006.10.045.

Hansen, D. O., J. M. Erbaugh, and T. L. Napier (1987), Factors related to adoption of soil conservation practices in the Dominican Republic, *Journal of Soil and Water Conservation*, *42*(5), 367.

Hipel, K. W. (1992), Multiple objective decision making in water resources, *Journal of the American Water Resources Association*, *28*(1), 3–12, doi:10.1111/j.1752-1688.1992.tb03150.x.

Horowitz, J. K., L. Lynch, and A. Stocking (2007), Competition-Based Environmental Policy: An Analysis of Farmland Preservation in Maryland, *Tech. rep.*, University of Maryland, Collge Park, MD.

Howard, A. (1991), A critical look at multiple criteria decision making techniques with reference to forestry applications, *Canadian journal of forest research*, *21*(11).

Jha, M. K., S. Rabotyagov, and P. W. Gassman (2009), Optimal Placement of Conservation Practices Using Genetic Algorithm with SWAT, *International Agricultural Engineering Journal*, *18*(1-2), 41–50.

Kanwar, R. S. (2006), Effects of cropping systems on NO3-N losses to tile drain, *Journal of the American Water Resources Association*, *42*(6), 1493–1501, doi:10.1111/j.1752-1688.2006.tb06016.x.

Kiker, G. a., T. S. Bridges, A. Varghese, P. T. P. Seager, and I. Linkov (2005), Application of multicriteria decision analysis in environmental decision making., *Integrated environmental assessment and management*, *1*(2), 95–108.

Lacroix, A., N. Beaudoin, and D. Makowski (2005), Agricultural water nonpoint pollution control under uncertainty and climate variability, *Ecological Economics*, *53*(1), 115 – 127.

Leonard, R., W. Knisel, and D. Still. (1987), GLEAMS: Groundwater loading effects on agricultural management systems, *Transactions of the ASAE*, *30*(5), Transactions of the ASAE.

Madani, K., and J. R. Lund (2011), A Monte-Carlo game theoretic approach for Multi-Criteria Decision Making under uncertainty, *Advances in Water Resources*, *34*(5), 607–616, doi:10.1016/j.advwatres.2011.02.009.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, *26*(2), 205–216, doi:10.1016/S0309-1708(02)00092-1.

Malone, R., L. Ma, D. Karlen, T. Meade, D. Meek, P. Heilman, R. Kanwar, and J. Hatfield (2007), Empirical analysis and prediction of nitrate loading and crop yield for cornsoybean rotations, *Geoderma*, *140*(3), 223–234, doi: 10.1016/j.geoderma.2007.04.007.

Maringanti, C., I. Chaubey, M. Arabi, and B. Engel (2008), A multi-objective optimization tool for the selection and placement of BMPs for pesticide control, *Hydrology and Earth System Sciences Discussions*, *5*(4), 1821–1862, doi: 10.5194/hessd-5-1821-2008.

Maringanti, C., I. Chaubey, M. Arabi, and B. A. Engel (2011), Application of a Multi-Objective Optimization Method to Provide Least Cost Alternatives for NPS Pollution Control, *Environmental management*, *48*(3), 448–61, doi:10.1007/s00267-011-9696-2.

McElroy, A., S. Chiu, J. Nebgen, A. Aleti, and F. Bennett (1976), Loading functions for assessment of water pollution from nonpoint sources, *Tech. rep.*, EPA 600/2-76-151. USEPA, Athens, GA.

Mirchi, A., D. W. J. Watkins, and K. Madani (2010), Modeling for Watershed Planning, Management, and Decision Making, in *Watersheds: Management, Restoration and Environmental Impact*, edited by J. C. Vaughn, p. 420.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams (2005), Soil and Water Assessment Tool (SWAT) Theoretical Documentation, Version 2005, (BRC Report 02-05).

Novotny, V. (2003), *Water quality: diffuse pollution and watershed management.*, second ed., 864 pp., Hoboken, NJ: J. Wiley.

Rabotyagov, S., M. Jha, and T. Campbell (2010), Searching for Efficiency: Least Cost Nonpoint Source Pollution Control with Multiple Pollutants, Practices, and Targets.

Ripa, M. N., A. Leone, M. Garnier, and A. Lo Porto (2006), Agricultural land use and best management practices to control nonpoint water pollution., *Environmental management*, *38*(2), 253–66, doi:10.1007/s00267-004-0344-y.

Rowe, M. D., and B. L. Pierce (1982), Sensitivity of the weighting summation decision method to incorrect application, *Socio-Economic Planning Sciences*, *16*(4), 173–177, doi:10.1016/0038-0121(82)90036-2.

Solovay, R. M. (1970), A Model of Set-Theory in Which Every Set of Reals is Lebesgue Measurable, *The Annals of Mathematics*, *92*(1), 1, doi:10.2307/1970696.

Srivastava, P. (2002), Watershed optimization of best management practices using AnnAGNPS and a genetic algorithm, *Water Resources Research*, *38*(3), doi:10.1029/2001WR000365.

Sullivan, D. G., and H. L. Batten (2007), Little River Experimental Watershed, Tifton, Georgia, United States: A historical geographic database of conservation practice implementation, *Water Resources Research*, *43*(9), doi:10.1029/2007WR006143.

Tedesco, L., D. Pascual, L. Shrake, R. Hall, L. Casey, P. Vidon, F. Hernly, K. Salazar, and R. Barr (2005), Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality, *Tech. rep.*, IUPUI, Indianapolis.

Tomer, M. D. (2010), How do we identify opportunities to apply new knowledge and improve conservation effectiveness?, *Journal of Soil and Water Conservation*, *65*(4), 261–265, doi:10.2489/jswc.65.4.261.

USDA NASS (2003), USDA-National Agricultural Statistics Service, Cropland Data Layer.

USDA NRCS (2010), Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Great Lakes Region, Phd, Purdue University.

USGS NED (2010), 1 arc Second Digital Elevation Model.

USGS NLCD (2001), 2001 Land Cover Data.

Veith, T. L. (2002), Agricultural Bmp Placement For Cost-Effective Pollution Control At The Watershed Level, Phd, Virginia Polytechnic Institute and State University.

Veith, T. L., M. L. Wolfe, and C. D. Heatwole (2003), Optimization procedure for cost effective BMP placement at a watershed scale, *Journal of the American Water Resources Association*, *39*(6), 1331–1343.

Walter, T., M. Dosskey, M. Khanna, J. Miller, M. Tomer, and J. Wiens (2007), The science of targeting within landscapes and watersheds to improve conservation effectiveness, in *Managing agricultural landscapes for environmental quality*, edited by S. M and C. C, Soil and Water Conservation Society, Ankeny, IA.

White, M. J., D. E. Storm, P. Busteed, S. Stoodley, and S. J. Phillips (2010), Evaluating conservation program success with Landsat and SWAT., *Environmental management*, *45*(5), 1164–74, doi:10.1007/s00267-010-9458-6.

Williams., J. R. (1990), The Erosion-Productivity Impact Calculator (EPIC) Model: A Case History, *Philosophical Transactions: Biological Sciences*, *329*(1255), 421–428.

Williams, J. R., and R. Hann (1978), Optimal operation of large agricultural watersheds with water quality constraints, *Tech. rep.*, Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Chapter 5

ASSESSMENT OF OPTIMAL AGRICULTURAL CONSERVATION PLANS UNDER UNCERTAIN CONDITIONS: IMPACT OF CLIMATE CHANGE

Highlights

The amount of investment on nonpoint source pollution control programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans. In this paper, hydrologic and water quality simulations driven by an extensive ensemble of climate projections were analyzed for their respective changes in basin average temperature and precipitation in the Eagle Creek Watershed, Indiana. Impact of climate change on performance of conservation practice and shifts in their optimal type and placement were also assessed. Nondominated Sorted Genetic Algorithm II (NSGA-II) was used to identify a set of optimal conservation practices in minimizing atrazine and nitrate loads at minimum cost. The results of this study revealed that substantial changes in water yield and pollutants transport are expected under different climate projections. Streamflow, sediment, total nitrogen, total phosphorus, and Atrazine loads respectively showed 15%, 40%, 20%, 32%, and 50% net increase at the end of 21st century with considerably high peaks (up to 250% higher than historical observations) and low fluxes (up to 40% lower than historical observation). In addition, performance of nonpoint source pollution control plans under different climatic projections were altered substantially from what they have designed for.

Optimization of conservation practices type and placement under several climatic projections also indicated considerable shift in Pareto-optimal front shape and position. However, the Pareto-front obtained for historic climate data represented the average performance of the optimal conservation plans for future climate projections. Analysis of the spatial distribution of optimal conservation plans showed that despite altered Pareto-optimal front position, spatial pattern and type of the selected conservation practices remained relatively the same. Performance of the nonpoint source pollution control plans in the course of their lifetime, considering lower changes in the climate conditions, was expected to be sustained.

5.1 Introduction

The global atmospheric concentration of carbon dioxide and other greenhouse gases (GHGs) has changed throughout the history; although based on the growing evidence, scientists are becoming confident that the current global warming trend is very likely induced by anthropogenic activities [*IPCC*, 2007]. Thus, many natural processes, including precipitation and temperature and therefore frequency and magnitude of extreme hydrologic events, are affected. Climate change can also significantly change diffusive nonpoint source pollution (NPS) transport and nutrients cycles through changes in physical, chemical, and biological responses [*Jennings et al.*, 2009]. Altered precipitation and temperature patterns, which in turn have implications on water quality and quantity, have been particularly compelling to trigger several interdisciplinary climate change research [*Brekke et al.*, 2009; *Chang*, 2004].

Climate change studies typically aim at (i) developing tools and methods to improve understanding of the human-environment interactions, (ii) supporting effective adaptation strategies to cope with the uncertain conditions, and (iii) developing decision support systems to mitigate policies and facilitate long-range decision making process [*National Research Council*, 2010a; *Antle*, 2009]. Existing studies concerning the impacts of climate change on the hydrology and water resources have typically been undertaken at coarse spatial scales. General circulation models (GCMs) are the basic tools to provide information regarding the response of the global climate system to increasing GHG emission path scenarios [*Rummukainen et al.*, 2001]. Emissions scenarios are described by the Intergovernmental Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) [*Nakicenovic et al.*, 2000]. However, application of GCM outputs in regional hydrologic modeling bear several key challenges because of inadequate accuracy and coarse spatial and temporal resolutions [*Fowler et al.*, 2007]. Thus, considerable effort has

focused on development of the "downscaling" techniques. Downscaled outputs of GCMs have widely used to study hydrologic and water quality impacts of climate change around the world [*Lettenmaier et al.*, 1999; *Bouraoui et al.*, 2002; *Abbaspour et al.*, 2009; *Brekke*, 2011]. Vulnerability to climate change impacts is shown to exist across the world [*National Research Council*, 2010b]. While direction and magnitude of the future impacts can not be determined precisely, adaptation strategies can be taken to reduce vulnerability of the certain socioeconomic and environmental systems [*National Research Council*, 2010c]. Watershed management decisions play a key role in climate change adaptation [*Lal et al.*, 2011] in which conservation practices, commonly known as Best Management Practices (BMPs), are widely accepted control measures of nonpoint sources of pollutants [*Novotny*, 1993; *Ritter and Shirmohammadi.*, 2001].

With projected changes in climatic conditions, significant changes in assimilative capacities of water bodies and landscape position of critical areas that should be targeted for implementation of conservation practices are also expected [*Parker et al.*, 2008; *Kaini et al.*, 2010]. The magnitude of money spent on nonpoint source pollution control programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans. In practice, several approaches have been taken to accomplish proper watershed scale conservation plans. Research to date shows that optimization approach outperforms other strategies regarding the both environmental and socioeconomic factors [*Arabi et al.*, 2006; *Jha et al.*, 2009; *Rabotyagov et al.*, 2010; *Maringanti et al.*, 2011].

Optimization algorithms should be coupled with a spatially distributed watershed model and use a proper representation of conservation practices to predict water quality changes arising from adoption of conservation practices [*Easton et al.*, 2008]. Watershed models have increasingly been recognized as an important tool for

136

improved understanding of the environmental processes [*Singh and Frevert*, 2005]. The primary goal of conservation practices optimization is to identify a set of optimal type and placement of conservation practices by minimizing pollutant movement from landscapes to water bodies at minimum cost. Thus, conservation practices targeting is inherently a multiobjective problem. Multiobjective optimization methods attempt to explore the tradeoffs between incommensurable and often conflicting socioeconomic and environmental factors by identifying the optimal type, size, and location of conservation practices within the watershed, resulting in a set of non-dominated (also referred to as Pareto-optimal) solutions. Although, optimization is an effective tool in finding optimal set of solutions, it ultimately depends upon availability of the automated optimization tools and expertise on developing and utilizing them and analyzing the results. Moreover, optimization is a computationally intensive task and, depending on the complexity of the problem, may need thousands to millions of model runs. In Chapter 4 a computational framework based on the multicriteria decision analysis (MCDA) was presented to effectively identify a set of optimal type and placement of conservation practices within a watershed. Solutions from MCDA approach can further be improved by using optimization methods in substantially less runs than standalone optimization methods.

The primary goal of this study is to evaluate the impact of climate change and variability on performance of agricultural nonpoint source pollution control practices in reducing NPS pollutants. Three specific objectives are (i) to identify the direction and degree of potential impact of climate change on hydrologic and water quality responses of the watersheds, (ii) to analyze vulnerability of optimal conservation plans to changes in climatic conditions in a watershed scale, and (iii) to address optimal placement of nonpoint source pollution control plans for uncertain future climate projections with low, moderate, and high emission scenarios. A broad Bias-Correction Spatial Disaggregation (BCSD) climate projections composed of 16

GCMs covering 3 emissions path scenarios [*Maurer et al.*, 2007] from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset [*WCRP*, 2012] were used in the study. Each scenario was temporally downscaled to a daily time step using a resampling and scaling/incrementing technique [*Wood et al.*, 2002, 2004]. An extensive discussion of impact of climate change on sediment, nutrients, and Atrazine loads and alteration of conservation practices performance is also provided.

5.2 Methods and materials

Hydrologic and water quality simulations driven by several climate projections were analyzed for their respective changes in basin average temperature and precipitation in the Eagle Creek Watershed, Indiana, USA. An extensive ensemble of future climate datasets from 1950-2099 were obtained and temporally and spatially downscaled. Performance of conservation practices that were optimized for historical data in Chapter 4 were also analyzed under different climate projections. Impact of several selected projections on identifying a set of optimal type and location of conservation practices were also assessed. Nondominated Sorted Genetic Algorithm II (NSGA-II) in combination with binary decision variables were used.

5.2.1 Study area

The Eagle Creek Watershed (ECW), located in central Indiana, has a drainage area of 248.1 km2 and lies within the Upper White River Watershed, extending from 40°01'24" to 40°04'16" north latitudes and 86°15'43" to 86°16'45" west longitudes.. According to the 2001 national Land Cover Dataset [*USGS NLCD*, 2001], the watershed consists of 52% croplands, 27% pasture, 12% low and high density urban areas, and 9% forest. The predominant crops are corn and soybeans. ECW drains into Eagle Creek Reservoir, which supplies drinking water for the city of Indianapolis. Figure 5.1 presents the location and land cover for the watershed. The
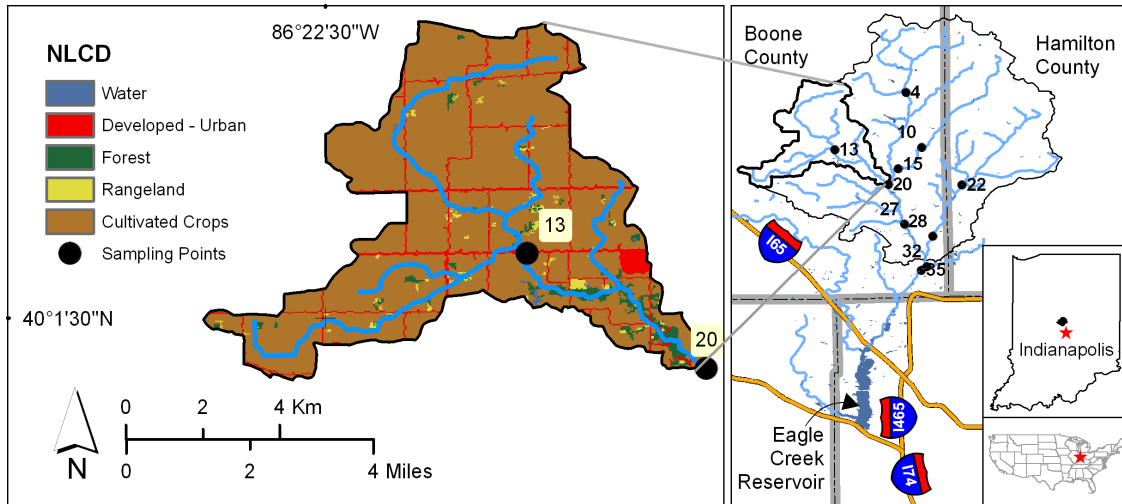
Figure 5.1: Eagle Creek Watershed Map- HUC 051202011102

soils are generally poorly draining and developed from glacial materials with thin loess over loamy glacial till and alluvial materials depositions. The dominant soils are the Crosby-Treaty-Miami in the headwaters and Miami-Crosby-Treaty along the downstream areas. The mean annual precipitation for the Eagle Creek Watershed area is 1052 mm. Monthly mean temperatures for this area from 1971-2000 shows January as having the lowest average temperature of -3.3°C and July as the being the warmest month with an average temperature of 23.7°C [*Tedesco et al.*, 2005].

Spatial scale of conservation planning depends upon numerous factors, including management objectives, available data resolution, dominant ecological processes, and potential sociopolitical constraints [*Walter et al.*, 2007; *Garen and Moore*, 2005]. To achieve the specific goals of water quality control, conservation practices targeting should be performed within a smaller geographic unit which ultimately allows us to better evaluate targeted management plan. Using 12-digit Hydrologic Unit Codes (HUCs) is recommended for the watershed-scale evaluation and planning of conservation practices to encourage participation of stakeholders in implementing conservation practices and enable monitoring water quality changes within subwatershed [*Haggard et al.*, 2010]. The ECW encompasses five 12-digit HUCs. The

HUC 051202011102 subwatershed with a drainage area of 41.2 km^2 was selected for the application of the integrated simulation-optimization framework in this study. The subwatershed consists of predominantly cropland (88%), and also contributes the largest amounts of non-point source nitrate and atrazine loads from agricultural lands to the Eagle Creek Reservoir (approximately 23% of nitrate loads and 28% of the atrazine loads).

5.2.2 Climate data

Ensemble of 112 Bias-Correction Spatial Disaggregation (BCSD) climate projections composed of 16 GCMs covering 3 emissions path scenarios [*Maurer et al.*, 2007], were used in this study. Monthly climate projections from 1950-2099 were obtained from *WCRP* [2012] in 1/8° spatial resolution. Three different emissions scenarios are named A2, A1B, and B1, which respectively correspond to the high, balanced (or moderate), and low GHG emissions. Table 5.1 summarises the information of the GCMs utilized in this study. Each of the GCMs have one or more runs for each emission scenario depending upon the 20th century "control" simulations [*WCRP*, 2012].

Coarse temporal resolutions of GCM outputs is not adequate for regional watershed modeling [*Fowler et al.*, 2007]. Moreover, SWAT setup requires use of point data at meteorological stations, whereas GCMs data are served as monthly data on 1/8° grids. Therefore, each scenario in this study was corrected from areal average to point bias (spatial correction) and then downscaled to a daily time step using a resampling and scaling/incrementing technique (temporal downscaling; [*Wood et al.*, 2002, 2004]).

To correct bias between grid average and point location, monthly precipitation and temperature data from NOAA Whitestown, Indiana station and GCM outputs (for a single grid cell encompassing the station) for 1950-2010 period were used. As

Table 5.1: List of GCMs utilized in this study

Modeling Group	Country	Name	Runs A2	A1B	B1	Reference
Bjerknes Centre for Climate Research	Norway	BCCR-BCM2.0	1	1	1	[Furevik et al., 2003]
Canadian Centre for Climate Modeling Analysis	Canada	CGCM3.1 (T47)	5	5	5	[Flato and Boer, 2001]
Meteo-France/ Centre National de Recherches Meteorologiques	France	CNRM-CM3	1	1	1	[Salas-Mélia et al., 2005]
CSIRO Atmospheric Research	Australia	CSIRO-Mk3.0	1	1	1	[Gordon et al., 2000]
US Dept. of Commerce/ NOAA/ GFDL	USA	GFDL-CM2.0	1	1	1	[Delworth et al., 2006]
US Dept. of Commerce/ NOAA /GFDL	USA	GFDL-CM2.1	1	1	1	[Delworth et al., 2006]
NASA/ Goddard Institute for Space Studies	USA	GISS-ER	1	2	1	[Russell et al., 2000]
Institute for Numerical Mathematics	Russia	INM-CM3.0	1	1	1	[Diansky and Volodin, 2002]
Institut Pierre Simon Laplace	France	IPSL-CM4	1	1	1	[Marti et al., 2006]
Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC)	Japan	MIROC3.2 (medres)	3	3	3	[Developers, 2004]
Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA	Germany/ Korea	ECHO-G	3	3	3	[Legutke and Voss, 1999]
Max Planck Institute for Meteorology	Germany	ECHAM5/ MPI-OM	3	3	3	[Jungclaus et al., 2006]
Meteorological Research Institute	Japan	MRI-CGCM2.3.2	5	5	5	[Yukimoto et al., 2001]
National Center for Atmospheric Research	USA	CCSM3	4	6	7	[Collins et al., 2006]
National Center for Atmospheric Research	USA	PCM	4	4	2	[Washington et al., 2000]
Hadley Centre for Climate Prediction and Research/ Met Office	UK	UKMO-HadCM3	1	1	1	[Gordon, 2002]

described by *Wood et al.* [2002] a quantile map was created from the probability thresholds between cumulative distribution functions of GCM and observed grids (on a monthly basis) for both precipitation and temperature during the overlapping period (1950-2010). Bias-correction was then completed by adjusting each climate projection (both past and future simulations) using the previously defined quantile maps (a single scale for precipitation and a single increment for temperature were used). This step ensured that adjusted GCM datasets were statistically consistent with the observed datasets during the overlapping period.

Likewise, a resampling method proposed by *Wood et al.* [2002] was utilized to downscale monthly total precipitation and average temperature data into daily values of precipitation, maximum temperature, and minimum temperature. In this method precipitation values are generated randomly for each future month within each year in each ensemble and then scaled so that the sum of daily precipitation data for each month is equal to the future monthly forecast. The same way, T_{max} and T_{min} for each month were generated and shifted such that their average, $(T_{min} + T_{max})/2$, resembles the future monthly forecast projection. This method created daily values of climatic variables that both preserve observed spatial and temporal correlations and aggregated to future monthly projections. In order to avoid resampling a "wet-cool" historical month to generate a daily series of climatic variables in a "dry-warm" projected month, or vice-versa, a wetness and warmth classification scheme was used. Detailed description of the temporal downscaling method can be found in *Wood et al.* [2002] and *Bureau of Reclamation* [2009].

5.2.3   Watershed model description: SWAT

The Soil and Water Assessment Tool (SWAT; *Arnold et al.* [1998]) was used to represent hydrologic and water quality processes in the ECW. Hydrologic processes simulated by SWAT include snow accumulation and melt, evapotranspiration, infiltration, percolation losses, surface runoff, and groundwater flows [*Neitsch et al.*,

2005]. SWAT is a physically-based watershed-scale, distributed-parameter, continuous time, and long-term, model that runs on a daily time step. It subdivides a watershed into subbasins connected by a stream network, and further delineates hydrologic response units (HRUs) consisting of unique combinations of land cover and soils in each subbasin.

SWAT can simulate major nutrient processes within a watershed. The nitrogen (N) cycle is simulated in five pools: inorganic (including ammonium and nitrate) and organic (including fresh, stable, and active). The main N processes are mineralization, decomposition, and immobilization. Nutrients are introduced into the main channel through surface runoff and lateral flow and transported downstream with channel flow. Plant uptake, denitrification, volatilization, leaching, and soil erosion are the major mechanisms of N removal from a field. The transport rate of organic N with sediment is calculated with a loading function developed by *McElroy et al.* [1976] and modified by Williams and Hann [*Williams and Hann*, 1978] for application to individual runoff events. The loading function estimates daily organic N runoff loss based on the concentrations of constituents in the top soil layer, sediment yield, and an enrichment ratio. Nutrient transformations in the stream are controlled by the in-stream water quality component of the model that is adapted from QUAL2E in-stream water quality model [*Brown and Barnwell*, 1987]. More detailed description of the nutrient components of SWAT can be found in *Neitsch et al.* [2005]. SWAT uses algorithms from GLEAMS (Ground Water Loading Effects on Agricultural Management Systems) [*Leonard et al.*, 1987] and EPIC (Erosion Productivity Impact Calculator) [*Williams.*, 1990] to model pesticide's overland fate and transport and movement from land to streams. It also incorporates a simple mass-balance method developed by *Chapra* [2008] to model the transformation and transport of pesticides in streams.

A 30-m resolution DEM from USGS National Elevation Dataset [*USGS NED*, 2010], National Land Cover Dataset (NLCD) 1992 and 2001 [*USGS NLCD*, 2001] for

Table 5.2: SWAT performance indices for daily streamflow, monthly nitrate, and monthly total pesticide simulation during 1993-2004 including 2 years of warmup period. PBIAS and NSE indicate percent bias and Nash-Sutcliff Efficiency coefficient, respectively.

Gauging Station	Constituent	Calibration Period Criteria (1995-1999)			Validation Period Criteria (2000-2003)		
		PBIAS (%)	R^2	NSE	PBIAS (%)	R^2	NSE
20	Nitrate	7.9	0.94	0.83	16.9	0.85	0.67
	Atrazine	-6	0.81	0.34	-14	0.7	0.41
22	Nitrate	-22.3	0.89	0.78	1.24	0.74	0.36
	Atrazine	42	0.69	0.44	-0.1	0.5	0.28
27	Nitrate	0.59	0.93	0.85	18.3	0.78	0.59
	Atrazine	13	0.66	0.35	-30	0.51	0.19
32	Nitrate	-7.9	0.92	0.84	8.4	0.76	0.55
	Atrazine	42.3	0.75	0.52	33.1	0.51	0.14
35	Streamflow	-12.2	0.78	0.61	4.3	0.78	0.56

urban areas, National Agriculture Statistics Service (NASS) Cropland Data Layer 2000-2003 [*USDA NASS*, 2003] for croplands, and SSURGO data from national resources conservation service (NRCS) [*USDA NRCS*, 2010] were used for watershed subdivision and delineating HRUs in the SWAT model. The ECW was subdivided into 35 subwatersheds and a total of 446 hydrologic HRUs. HUC 051202011102 includes 4 subwatersheds and 40 HRUs (Figure 5.1. Watershed model is calibrated and validated for entire ECW model for predicting streamflow at watershed outlet, and nitrate and Atrazine at gauging stations 20, 22, 27, and 32. Table 5.2 presents performance indices of the SWAT model in predicting daily streamflow and monthly nitrate, and Atrazine.

5.2.4 Representation of Conservation Practices

IIn this study, water quality impacts of fertilizer management, grassed waterways, grade stabilization structures, and tillage/residue management were evaluated. Only row crops, including corn and soybean, were considered for implementation of the nonpoint source pollution control plan. Conservation practices were

144

represented using numerical procedures from the published studies [*Arabi et al.*, 2004; *Arnold et al.*, 2011; *White et al.*, 2010]. SWAT includes explicit functions for representation of fertilizer management and tillage/residue management, grassed water ways [*Arnold et al.*, 2011]. Table 5.3 summarizes the relevant SWAT management operations and parameters and their corresponding values for representation of conservation practices. Implementation of grade stabilization structures are only considered in combination with grassed waterways. Conservation practices in the binary optimization approach were represented by default values presented in Table 5.3. Detailed description of the conservation practices representation methods can be found in *Arabi et al.* [2004]; *Arnold et al.* [2011], and *White et al.* [2010]. We assumed selected conservation practices can coexist in the same field; and therefore, there exist $23 \times 3 = 69$ alternatives for this study that makes 2^{69} possible combination of conservation practices type and location.

5.2.5   Multiobjective Optimization

Watershed scale nonpoint source pollution control planning aims at minimizing the potential for water pollution and environmental degradation at minimum cost. Hence, selection of type and placement of the conservation practices is inherently a multiobjective problem. A multiobjective optimization problem can be stated in the following mathematical form:

$$\text{minimize } F(\boldsymbol{\theta}) = \{F_1(\boldsymbol{\theta}), ..., F_m(\boldsymbol{\theta})\}, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^n \qquad (5.1)$$

where m is total number of objective functions and $\boldsymbol{\theta}$ denotes a decision variable vector within the feasible decision space of $\boldsymbol{\Theta}$. Multiobjective search algorithms can simultaneously optimize two or more conflicting objectives, resulting in a set of nondominated (also referred to as "Pareto-optimal front") solutions. Global search methods are robust in finding optimal solutions by searching over the larger subset of the search space, and thereby escape being trapped in local optima [*Gitau et al.*,

Table 5.3: SWAT parameter values for representation of conservation practices

Practice	Representation Parameters	Unit	Value	Cost	r_{OM}	t (years)
Fertilizer management	Fertilizer application reduction, FRT_KG (*.mgt)	%	20			
Grassed Waterways	Manning's n, GWATN (*.ops)		0.35	11,400 ($/ha)	3	10
	Width, GWATW (*.ops)	m	15			
	Length, GWATL (*.ops)	km	(A)			
Grade Stabilization	Reduce slope steepness, CH_S1 (*.sub)	m	(B)	4,800 ($/structure)	2	15
	Reduce curve number, CN2 (*.mgt)		2		0	1
Tillage/Residue Mgt.	Harvest Efficiency, HARVEFF (*.mgt)		1	75 ($/ha)		
	Overland Manning's number, OV_N (*.hru)		0.1			

(A) Estimated using equation $0.5 \times sqrt(HRUarea)$

(B) Estimated based on the height of structure (default is GSSH = 1.2 meters) with the minimum of 0

2004]. Nondominated Sorted Genetic Algorithm II (NSGA-II) was used as the primary search method in identifying optimal solutions. A typical GA starts with an initial population of solutions and then implements probabilistic and parallel exploration in the search space using the domain-independent genetic operators (i.e. chromosome reproduction) to find optimal solutions [*Arabi et al.*, 2006; *Perez-Pedini et al.*, 2005].

The initial population of individuals is typically generated randomly. A GA does its tuning in stages called generations. GA's chromosome reproduction stops if any of the pre-defined termination conditions are met. Unlike most of the optimization algorithms, GA promises convergence but not optimality [*Lakshmanan*, 2000]. Therefore, defining termination criteria for GA is a hard and subjective task. The commonly used GA termination criteria are maximum number of function evaluations and number of successive generations without considerable improvement. The first criterion needs a prior knowledge about the problem to allow the estimation of a reasonable maximum search length. On the other hand, the second criterion has an adaptive nature and does not require such knowledge [*Safe et al.*, 2004], however needs definition of substantial improvement. In multiobjective optimization problems the computational time required to find the optimal solution(s) increases as the number and complexity of the decision variables and objectives increase. In particular, larger number of decision variables requires larger population size that implies larger number of function evaluation. Parallel computing was adopted in NSGA-II algorithm to reduce the overall computational time.

In this study, decision variables are defined as a binary chromosome consisting of 1's and 0's respectively indicating that the corresponding conservation practice "be" or "not be" implemented. Similar to any GA algorithm, the algorithm considers different candidate solution sets as the individuals within a population. Each

individual consists of genes coded within a chromosome that defines the corresponding fitness. Optimization initializes using a MCDA-based scheme (as proposed in Chapter 4) and then reproduces the chromosomes to find optimal solutions.

Minimizing pollutants load, or alternatively maximizing pollutants load reduction, is the key objective in nonpoint source pollution control planning. The load reduction of pollutant z can be calculated as:

$$\Delta L_z = \frac{L_{z,base} - L_{z,BMP}}{L_{z,base}} \times \%100 \tag{5.2}$$

where ΔL_z is the estimated percent reduction of pollutant load z, while $L_{z,base}$ and $L_{z,BMP}$ represent the pollutant loads before and after implementation of NPS pollutions control practices, respectively. Both $L_{z,base}$ and $L_{z,BMP}$ would be estimated from hydrologic and/or water quality responses of the watershed model. Depending on the chronic or acute nature of the pollutant z, different approached can be taken to calculate pollutant load L_z. For pollutants with chronic impacts (such as sediments and nutrients) long-term loading (e.g. monthly, annual) are appropriate in nonpoint source pollution control strategies [*US EPA*, 2008].

Nonpoint source pollution control plans yield benefits in water quality and wildlife habitats, but impose costs on stakeholders as well, that should be considered in decision making process. Data Envelopment Analysis (DEA) [*Charnes et al.*, 1978] provides a methodology for economic analysis of the agricultural technology and conservation practices. In a basic DEA, there are $k = 1, , K$ decision making unit (DMUs). Each DMU uses $x = (x_1, , x_M) \in \mathbb{R}_+^M$ inputs to produce $u = (u_1, , u_N) \in \mathbb{R}_+^N$ outputs. Hence profit (net return) within the entire study area can be computed as:

$$\pi(x,u) = \overbrace{\sum_{k=1}^{K}\sum_{n=1}^{N} r^{n,k}\beta_u^{n,k}}^{\text{benefit}} - \overbrace{\sum_{k=1}^{K}\sum_{m=1}^{M} r^{m,k}\beta_x^{m,k}}^{\text{cost}} \tag{5.3}$$

α_r and α_p are changes in the unit prices of outputs and inputs, respectively. Cost of conservation plans consists of implementation, maintenance, damage, and opportunity costs. Economic benefit also includes monetary values of the water quality improvement benefits (off-site benefits), and benefits from changes in agricultural production as a results of the implementing new conservation practices. In this study three objective functions were selected as follows: (1) minimize conservation practices implementation cost including implementation, management, and damage costs, (2) maximize Atrazine load reduction, and (3) maximize nitrate load reduction. Benefit was con considered in this study, because it is a direct function of the pollutants load reduction.

Performance of the optimal conservation plans were analyzed for an extensive 112 projections. With projected changes in climatic conditions and subsequently altered hydrologic regimes, assimilative capacities of water bodies and landscape position of critical areas were also expected to alter. Thus, impact of projected climatic conditions on optimal placement of conservation practices were also analyzed in this study. Baseline scenario solutions were obtained from optimizing type and placement of conservation practices for historical climate conditions (from 1994-2003). In addition to the shift in Pareto-optimal front shape, degree of similarity in type and spatial distribution of the conservation practices were also analyzed. Two pairwise distance measures of Jaccard and Spearman were used for test of similarity. The "Jaccard index", also known as the "Jaccard similarity coefficient" [*Jaccard*, 1908; *McCormick et al.*, 1992] expresses the similarity, or dissimilarity, of nonzero datasets X and Y as

$$J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|} \tag{5.4}$$

where $J(X,Y)$ is the Jaccard similarity coefficient between datasets X and Y, $|X \cap Y|$ is size of the intersection between dataset X and dataset Y, $|X \cup Y|$ is size of the

union between the two datasets. For binary datasets, Jaccard index simplifies to

$$J(X,Y) = \frac{p}{p+q+d} \tag{5.5}$$

where, p is number of attributes that are 1 for both, q is number of attributes that are 1 in X but not in Y, and d represents number of attributes that are 1 in Y but not in X. Jaccard index of 1 indicates the perfect matching similarity.

Likewise, Spearman's rank correlation coefficient is a measure of statistical dependence between two variables *Lehmann* [1975]

$$\rho = 1 - \frac{6\sum_{i=1}^{n}(x_i - y_i)^2}{n(n^2 - 1)} \tag{5.6}$$

where x_i and y_i are rank of ith attribute 1 in X and Y, respectively and n represents number of attributes in X and Y. A perfect Spearman correlation coefficient of $+1$ or -1 occurs when each of the variables in X is a perfect monotone function of Y.

5.3 Results and discussion

Overall, the results of this study showed that water yield and pollutant transport substantially changed under different climatic paradigms. The direction, magnitude, and variability of the climate and its impacts were analyzed by combining SRES emissions scenarios and computing ensemble means, medians, and quartiles. In addition, conservation practices showed variable performance under different climatic projection. Figure 5.2 shows the annual basin averaged temperature and precipitation from 2010-2099. Each individual projected ensemble shows considerably variable minimum and maximum temperature over the nine decade, although follow a general pattern when grouped by SRES emissions scenario. The ensemble averaged minimum and maximum temperature for each of the emission scenarios appears to increase at approximately the same rate until 2040. From 2040-2070 the increase in temperature in B1 is substantially lower than that of A1B and A2 that
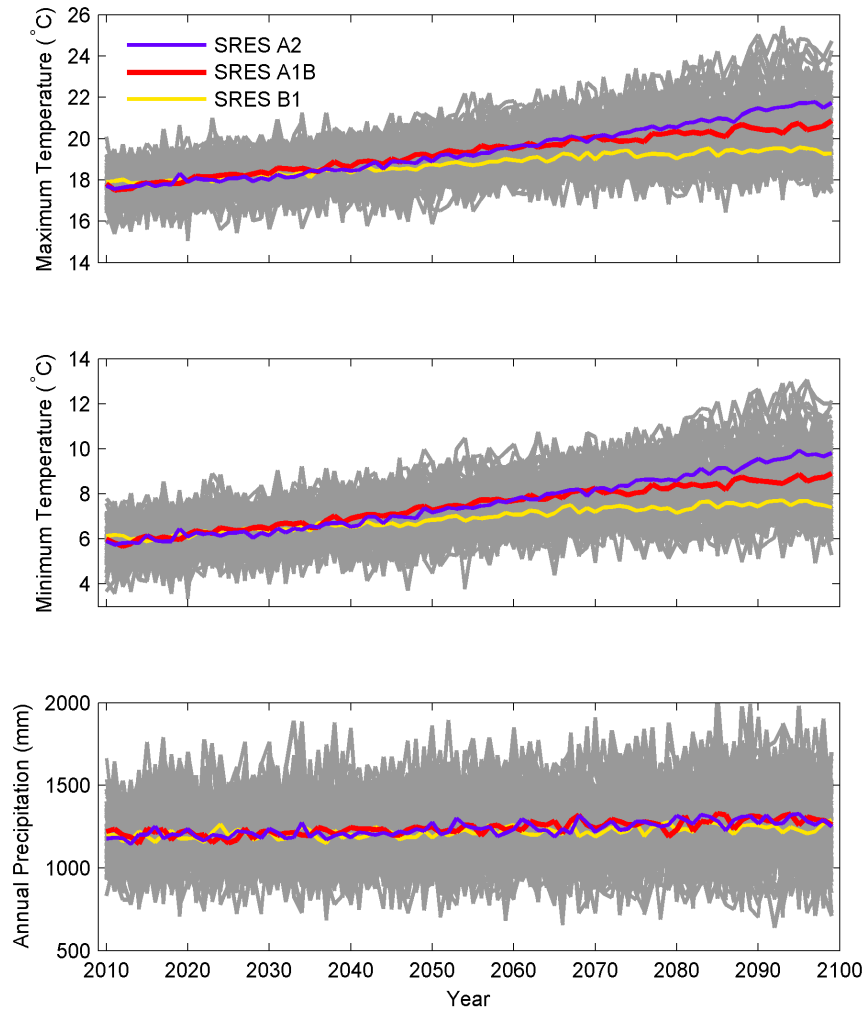
Figure 5.2: Basin averaged annual precipitation, minimum temperature and maximum temperature. The grey lines represent annual values from each of the 112 projections and the colored lines represent the ensemble averaged annual values from SRES emissions scenarios A2, A1B, and B1.

continue increasing with higher rates. After 2070, temperature in A2 increased with higher rate than A1B and continued by the end of the century.

Precipitation projections, on the other hand, showed more steady and gradually increasing throughout the century. Averaged ensemble for each of the scenarios showed approximately 15 percent increase over the century. However, precipitation fluctuation shows slight increase as it gets closer to the end of century which means intense flood and drought conditions should be expected. Figure 5.3 depicts the variability in projected temperature and precipitation within different emission scenarios at three time periods, including near-future (2010–2019), mid-century (2050–2059), and late-century (2090–2099). Box plot with whiskers extends from minimum to maximum of values and bottom and top of the box shows lower and upper quartiles, respectively. Increased median and wider ranges indicate higher variability over the time. Higher variability is always associated with higher uncertainty. Thus, higher variability in hydrologic regimes and pollutants fluxes and consequently higher prediction uncertainty is also expected as the simulations progresses into the future.

Analysis of monthly averaged temperature and precipitation also showed seasonal variability in the time-series (Figure 5.4) which increases the uncertainty. Precipitation projections showed minimum increase from June-August and higher in January and March-May of each year with higher variations for A2 emission scenario. Near future (i.e. 2010s) precipitation showed slight decrease than historic data in June for SRES A2. Likewise, seasonal temperature showed systematic increase with the highest increase of 14 percent during spring (i.e. March, April, and May) followed by a 10 percent increase in winter (December, January, and February) and negligible net increase during warmer seasons of summer and fall .

As expected, variability in climatic conditions were translated into a range of watershed responses via SWAT simulations. Figure 5.5 shows the variability in resulting streamflow, sediment load, major nitrogen and phosphorus pools, and

Figure 5.3: Variability in annual average temperature and precipitation in the Eagle Creek watershed from SRES B1, A1B, and A2 emissions scenarios over the three near-, mid-, and late-century. The dashed red line represents the average annual value over the 1950–2009 historic period.
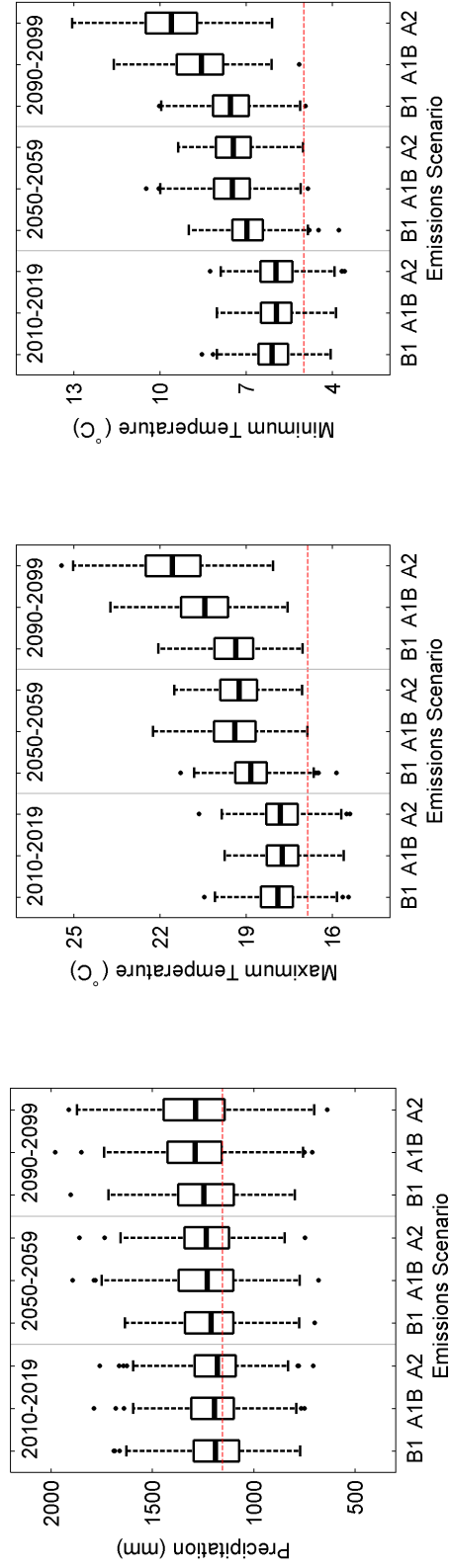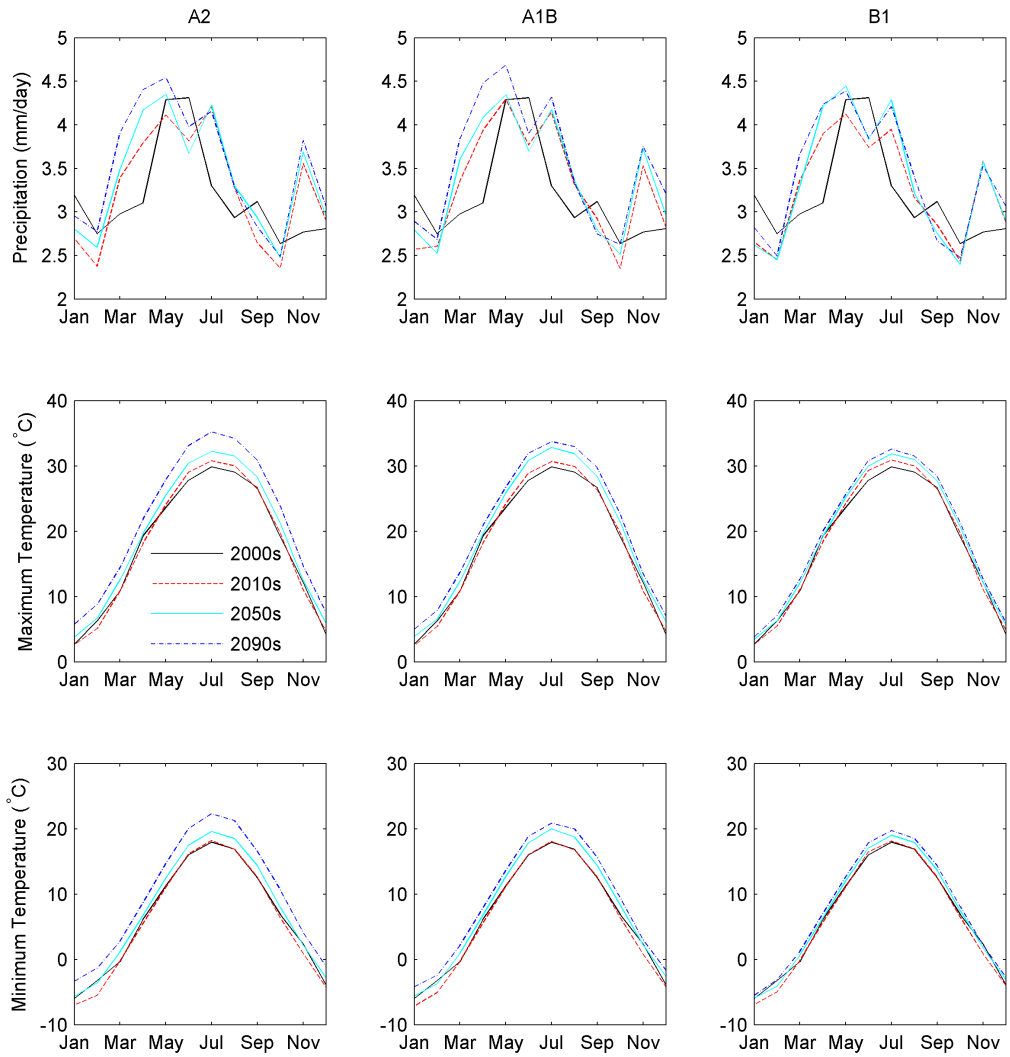
Figure 5.4: Seasonal variability in monthly average temperature and precipitation in the Eagle Creek watershed from SRES B1, A1B, and A2 emissions scenarios over the three near-, mid-, and late century.

Atrazine at the watershed outlet. Streamflow showed similar pattern and ranges as the precipitation, that is consistent with results of other studies (e.g. *Stonefelt et al.* [2000]). Each of the three emission scenarios showed 25 percent increase in streamflow generation with the higher increase rate in the near-century (2010s). Increased precipitation in the winter in the form of snow is also expected to increase early snowmelt with higher rates in spring as the temperature starts to increase. Hence, naturally, higher and more frequent storms are also expected. Each of the projected scenarios have demonstrated several higher peaks (up to 2 times higher than historical peaks) by the end of the century. On the other hand, extreme drought are also expected. Each of the emission scenarios indicated several extremely low flow conditions which are up to 45 percent lower than that of historical observations with greatest variations for A2 emission scenario.

A detailed analysis of water quality constituents variation under climate projections were also performed. Water quality responses were highly variable with a consistent net increase throughout the century. SRES averaged data for sediment, total phosphorus, total nitrogen, and total pesticide follow an increasing pattern with respectively up to 40%, 20%, 32%, and 50% net increase at the end of century. Peak values are expected to be up to 180% higher for sediment and total nitrogen and less than 250% higher for phosphorus and Atrazine constituents than historical high observations. While total nitrogen and total phosphorus loads demonstrated gradually increasing trend, organic phosphorus and organic nitrogen are substantially decreasing (40% decrease by the end of century). Thus, net increase in total nitrogen loads is mainly due to the increase in nitrate load. This can be explained by increased rates of organic-N decomposition, ammonification, and nitrification due to the increased projected temperature and more moisture available. Increased sediment and total Atrazine loads was also a direct consequence of the increased runoff that led to the elevated wash-out potential.
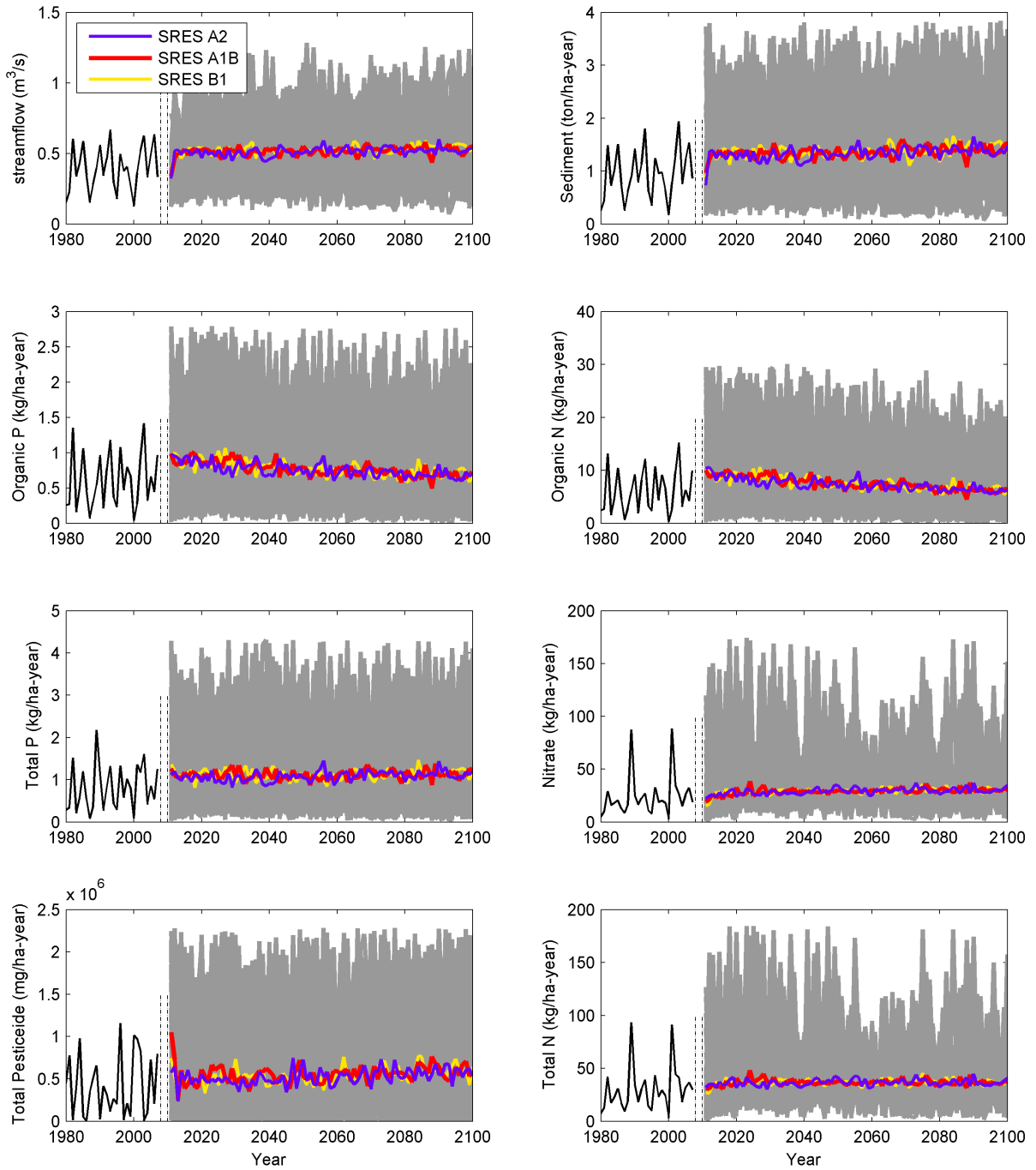
Figure 5.5: Variability in SWAT major outputs from 112 projections in the Eagle Creek watershed
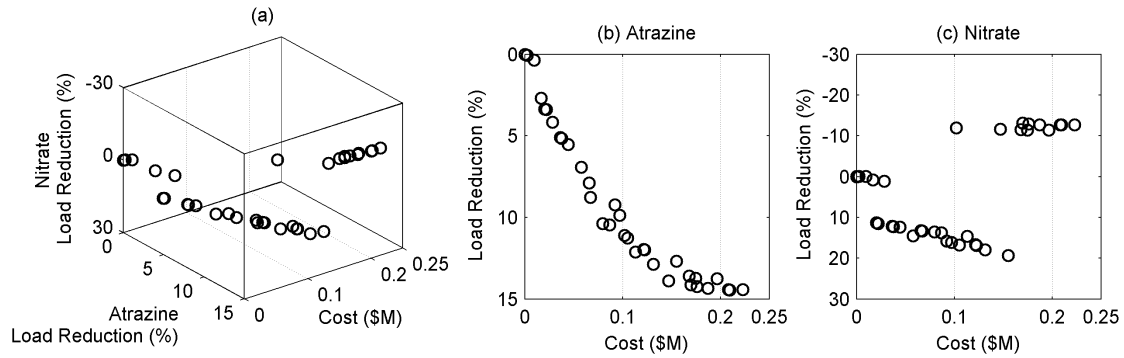
Figure 5.6: Pareto-optimal front solutions for maximizing pollutants load reduction and minimizing implementation cost of conservation practices.

Analysis of the results made me confident that climate change and variability can substantially change the nutrients, sediment, and Atrazine fluxes within the watershed. Accordingly, efficiency of nonpoint source pollution control plans in reducing pollutant loads was expected to alter. To study impact of climate projections on performance of conservation plans, Pareto-optimal front solutions were identified for historical climatic data (1994-2003) for maximizing nitrate and Atrazine loads reduction and minimizing implementation cost of conservation plans in Chapter 4. Figure 5.6 depicts the Pareto-optimal front solutions for the problem that consists of 38 nondominated solutions. Each one of the solutions were then tested under 112 climate projections and the results are presented in Figure 5.7 in which x-axis presents implementation cost of conservation plans associated with the optimal solutions. Conservation plans showed highly variable performance in reducing the pollutants load under different projections. While the average performance was degraded for Atrazine load reduction, several expensive solutions were improved in reducing nitrate load. Uncertainty in performance of the conservation practices, which is represented by the difference between top and bottom whiskers, increased as more money was spent.

Knowing the impact of climate change on efficiency of conservation plans, it is natural to take the next step by considering climate change projections in identi-
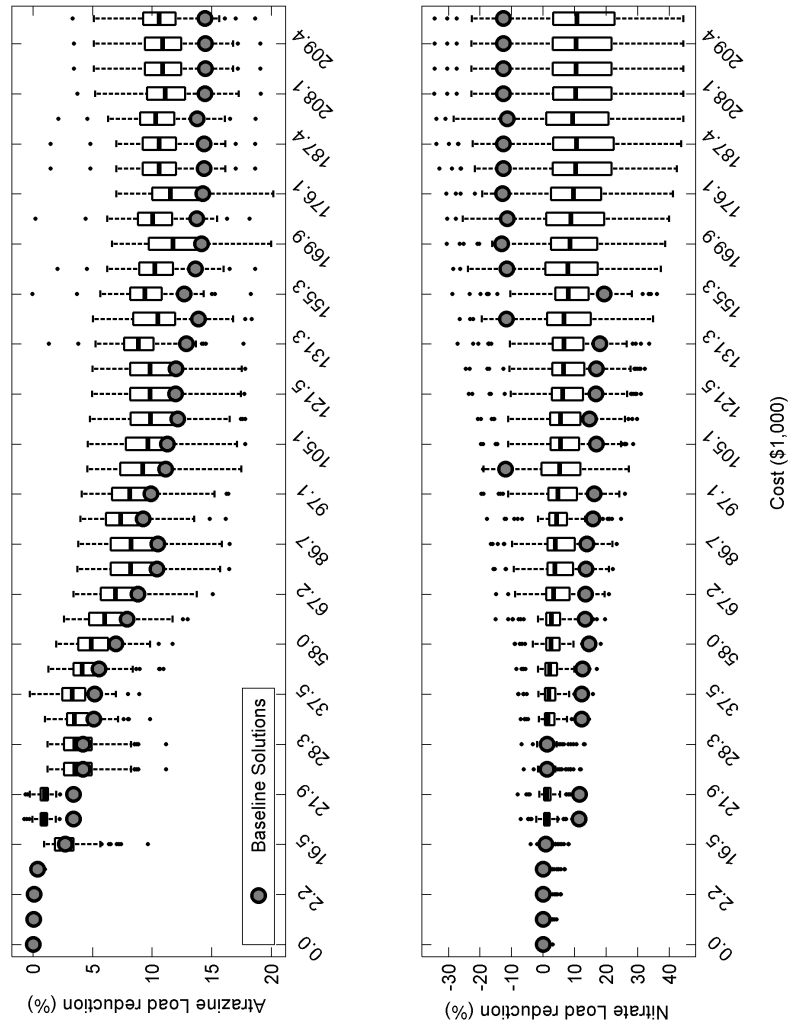
Figure 5.7: Impact of climate change on performance of optimal conservation plans in reducing Atrazine and nitrate loads in Eagle Creek Watershed, Indiana. Solutions from optimization are sorted in the order of increasing cost in x-axis. Box plots show median, range and variability in pollutants loads under 112 climate projections. Negative load reduction means increased load.

fying optimal type and placement of conservation practices within the watershed. For this purpose, 2 climate projections with minimum and maximum precipitation were selected for each of the SRES emission scenarios. Optimization was ran for 90 years for minimizing Atrazine and nitrate load reductions at minimum cost. Figure 5.8 depicts how optimal solutions may change under different climate projections. Depending upon the selected climate projections, optimal solutions might be more effective (solutions below baseline Pareto-front in Figure 5.8) or less effective (above the baseline Pareto-front) in reducing pollutant loads. Interestingly, for all selected emission scenarios, projections with low precipitation identified less efficient conservation plans in reducing both Atrazine and nitrate loads than projections with high precipitation. In the case of Atrazine load reduction, Pareto-front solutions associated with low precipitation scenario settled above baseline solutions (i.e. was less effective) and vice versa. On the other hand, for nitrate load reduction, both optimal sets in SRES B2 (low emission scenario) were less effective and both optimal sets in SRES A1 (high emission scenario) were more effective than baseline solutions. Solutions in the selected SRES A1B scenarios showed lower efficiency for low precipitation and higher efficiency for higher precipitation.

Analysis of the optimal solutions also showed that despite substantially different efficiency of the Pareto-optimal solutions in pollutant load reduction, optimized conservation plans and their spatial locations are highly correlated regardless of the selected climate projection. For this purpose optimal solutions were grouped in 5 groups based on their associated costs from $0 to $250,000. Then their "type and placements" (binary datasets) were tested for similarity using pairwise Jaccard coefficient. Similarity of the selected conservation practices' "type" was also tested by means of pairwise Spearman's rank coefficient. In total, 594 pairwise comparison between future scenarios and baseline solutions were performed. Figure 5.9 depicts
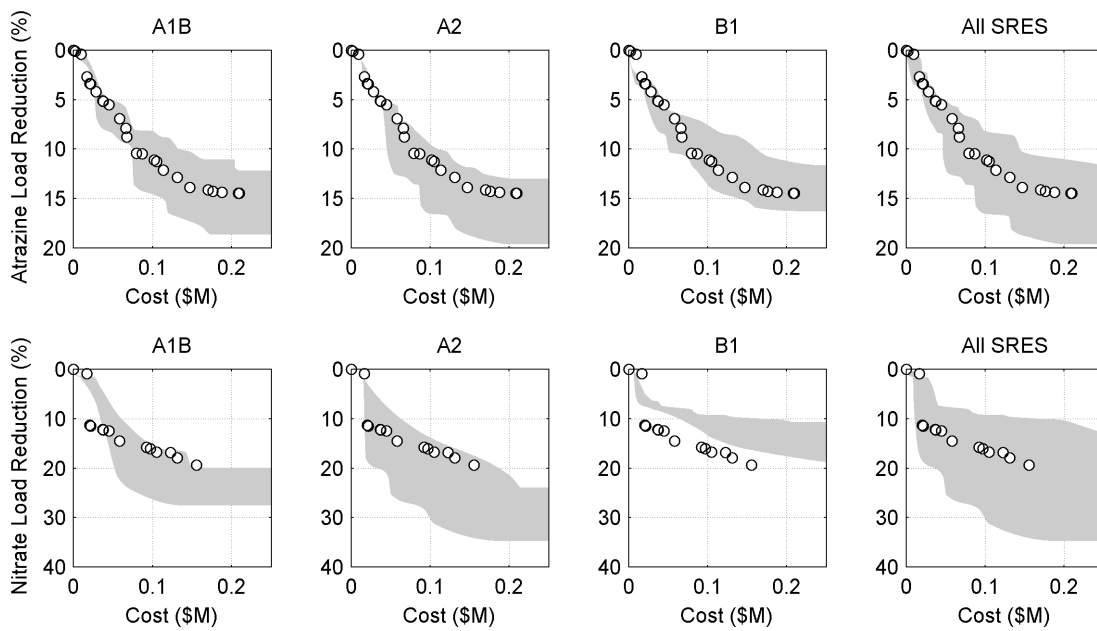
Figure 5.8: Impact of climate change on identifying optimal conservation plans in reducing Atrazine and nitrate loads. Circles (**O**) represent Pareto-optimal front solutions for baseline (1994-2003). Gray area determines objective space in which optimal solutionsor projections from SRES A1B, A2, and B1 are distributed.

summary of the results obtained from similarity tests. Line plots show SRES averaged similarity index and boxplots depict the median, range and quintile of Jaccard indices.

Jaccard test suggested that similarity between solutions for future and baseline scenarios increases as nonpoint source pollution control plan gets more effective. Higher dissimilarity in lower costs was mostly due to the diversity of the solutions obtained for all scenarios, whereas we saw several high Jaccard's index within 0-0.05 and 0.05-0.1 scales. In addition, dissimilarity between scenarios reveals the fact that critical areas that were targeted for implementation of conservation practices were slightly shifted under different climate projections. Results from Spearman's rank correlation also strongly supports the results from Jaccard test. higher Spearman's coefficient suggested that, regardless of the selected climate projection, selected conservation practices types within optimal solutions are highly correlated.

5.4 Summary and conclusion

With projected changes in climatic conditions, considerable changes in assimilative capacities of water bodies and landscape position of critical areas that should be targeted for implementation of conservation practices are also expected. The magnitude of money spent on conservation programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans. Hydrologic and water quality simulations driven by an extensive ensemble of climate projections were analyzed for their respective changes in basin average temperature and precipitation in the Eagle Creek Watershed, Indiana, USA. Impact of climate change on performance of conservation practices and their optimal placement were also assessed. Nondominated Sorted Genetic Algorithm II (NSGA-II) coupled with the Soil and Water Assessment Tool (SWAT) was used for optimizing type and placement of conservation plans with the goal of reducing Atrazine and nitrate load with minimum cost.
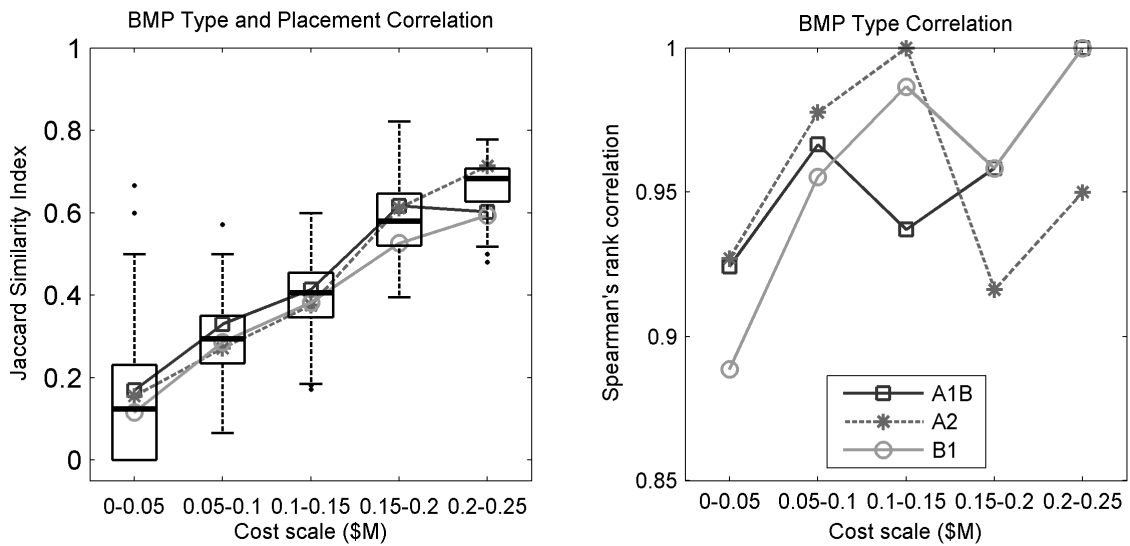
Figure 5.9: Similarity Index of optimal placement (left) and type (right) of conservation practices for different climate projection and five cost scale. The placement similarity was tested using Jaccard coefficient on pairwise comparison of binary sets. Spearman's rank coefficient was used to test correlation between number of conservation practices with the same type for each cost scale. Line plots show SRES averaged similarity index and boxplots show the median, range and quintile of Jaccard indices.

The results of this study showed that water yield and pollutants transport changes substantially under different climatic paradigms. The direction, magnitude, and variability of the climate and its impacts were analyzed by combining SRES emissions scenarios and computing ensemble means, medians, and quartiles. Streamflow, sediment, total nitrogen, total phosphorus, and Atrazine loads respectively showed 15%, 40%, 20%, 32%, and 50% net increase at the end of 21st century with peaks up to 250% higher than historical observations. In addition, performance of conservation plans under different climatic projections altered greatly from what they have designed for, based on the historic climate data. Optimization of type and placement of conservation practices for several climatic projections also indicated considerable shift in Pareto-optimal front from solutions optimized for historical data. Although, investigating the type and spatial distribution of the selected conservation practices showed that despite altered Pareto-front for climate projections, spatial pattern and type of the selected conservation practices remained relatively the same. Jaccard's and Spearman's similarity tests showed that as conservation plans became more effective (i.e. more money was spent), similarity of the type and location of conservation plans under different climate projections increases. Existing slight difference can be explained by shifted critical areas as a result of climate change.

Overall, the Pareto-front solutions obtained based on the historic data represented the average performance of the optimal nonpoint source pollution control plans from different climate projections. Thus, watershed scale conservation planning based on the historic data can be asserted for the future by confidence. Performance of the conservation practices in the course of their lifetime (that are between 10-30 years), considering lower changes in the climate conditions in near century, was expected to be sustained. However, revision of the adaptation measures should be considered on a regular basis to accordingly respond to the impacts of climate change and the risks associated with these impacts.

Bibliography

Abbaspour, K. C., M. Faramarzi, S. S. Ghasemi, and H. Yang (2009), Assessing the impact of climate change on water resources in Iran, *Water Resources Research*, *45*(10), doi:10.1029/2008WR007615.

Antle, J. M. (2009), Agriculture and the Food System: Adaptation to Climate Change, Resources For the Future, *Tech. rep.*, Washington D.C.

Arabi, M., R. S. Govindaraju, and M. M. Hantush (2004), Impact of Best Management Practices on Water Quality of Two Small Watersheds in Indiana: Role of Spatial Scale, *Tech. rep.*, US EPA, National Risk Management Research Laboratory, Cincinnati, OH.

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

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment, Part I: Model development, *Journal of the American Water Resources Association*, *34*(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x.

Arnold, J. G., J. R. Kiniry, R. Srinivasan, J. R. Williams., E. B. Haney, and S. L. Neitsch (2011), Soil and Water Assessment Tool Input/Output File Documentation, Version 2009, Report No. 365, *Tech. rep.*, Texas Water Resources Institute, College Station, Tx.

Bouraoui, F., L. Galbiati, and G. Bidoglio (2002), Climate change impacts on nutrient loads in the Yorkshire Ouse catchment (UK), *Hydrology and Earth System Sciences*, *6*(2), 197–209, doi:10.5194/hess-6-197-2002.

Brekke, L. D. (2011), Addressing Climate Change in Long-Term Water Resources Planning and Management: User Needs for Improving Tools and Information, *Tech. rep.*, U.S. Department of the Interior and U.S. Army Corps of Engineers.

Brekke, L. D., J. Kiang, J. Olsen, R. Pulwarty, D. Raff, D. Turnipseed, R. Webb, and K. White (2009), Climate Change and Water Resources Management: A Federal Perspective, *Tech. rep.*, U.S. Department of the Interior and U.S. Geological Survey, Reston, Virginia.

Brown, L. C., and T. O. Barnwell (1987), The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual, *Tech. rep.*, EPA document EPA/600/3-87/007. USEPA, Athens, GA.

Bureau of Reclamation (2009), long-term planning hydrology based on various blends of instrumental records, paleoclimate, and projected climate information, *Tech. rep.*

Chang, H. (2004), Water Quality Impacts of Climate and Land Use Changes in Southeastern Pennsylvania*, *The Professional Geographer*, *56*(2), 240–240–257–257, doi:10.1111/j.0033-0124.2004.05602008.x.

Chapra, S. C. (2008), *Surface Water-Quality Modeling*, 844 pp., Waveland Pr Inc.

Charnes, A., W. Cooper, and E. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research*, *2*(6), 429–444, doi:10.1016/0377-2217(78)90138-8.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, B. P. Briegleb, C. M. Bitz, S.-J. Lin, and M. Zhang (2006), The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3), *Journal of Climate*, *19*(11), 2144–2161, doi:10.1175/JCLI3760.1.

Delworth, T. L., A. J. Broccoli, A. Rosati, R. J. Stouffer, V. Balaji, J. A. Beesley, W. F. Cooke, K. W. Dixon, J. Dunne, K. A. Dunne, J. W. Durachta, K. L. Findell, P. Ginoux, A. Gnanadesikan, C. T. Gordon, S. M. Griffies, R. Gudgel, M. J. Harrison, I. M. Held, R. S. Hemler, L. W. Horowitz, S. A. Klein, T. R. Knutson, P. J. Kushner, A. R. Langenhorst, H.-C. Lee, S.-J. Lin, J. Lu, S. L. Malyshev, P. C. D. Milly, V. Ramaswamy, J. Russell, M. D. Schwarzkopf, E. Shevliakova, J. J. Sirutis, M. J. Spelman, W. F. Stern, M. Winton, A. T. Wittenberg, B. Wyman, F. Zeng, and R. Zhang (2006), GFDL's CM2 Global Coupled Climate Models. Part I: Formulation and Simulation Characteristics, *Journal of Climate*, *19*(5), 643–674, doi:10.1175/JCLI3629.1.

Developers, K.-. (2004), K-1 coupled model (MIROC) description. K-1 Technical Report 1. H. Hasumi and S. Emori (eds.), *Tech. rep.*, Center for Climate System Research, University of Tokyo, Tokyo, Japan.

Diansky, N., and E. Volodin (2002), Simulation of present-day climate with a coupled atmosphere-ocean general circulation model, *J Clim*, *38(6)*(6), 1057–1072.

Easton, Z. M., M. T. Walter, and T. S. Steenhuis (2008), Combined monitoring and modeling indicate the most effective agricultural best management practices., *Journal of environmental quality*, *37*(5), 1798–809, doi:10.2134/jeq2007.0522.

Flato, G. M., and G. J. Boer (2001), Warming asymmetry in climate change simulations, *Geophysical Research Letters*, *28*(1), 195–198, doi:10.1029/2000GL012121.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, *International Journal of Climatology*, *27*(12), 1547–1578, doi:10.1002/joc.1556.

Furevik, T., M. Bentsen, H. Drange, I. K. T. Kindem, N. G. Kvamst, and A. Sorteberg (2003), Description and evaluation of the bergen climate model: ARPEGE coupled with MICOM, *Climate Dynamics*, *21*(1), 27–51, doi:10.1007/s00382-003-0317-5.

Garen, D. C., and D. S. Moore (2005), Curve number hydrology in water quality modeling: uses, abuses, and future directions, *Journal of the American Water Resources Association*, *41*(2), 377–388, doi:10.1111/j.1752-1688.2005.tb03742.x.

Gitau, M. W., T. L. Veith, and W. J. Gburek (2004), Farmlevel optimization of bmp placement for costeffective pollution reduction, *ransactions of the ASABE*, *47*(6), 1923–1931.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Climate Dynamics*, *16*(2-3), 147–168.

Gordon, H. (2002), The CSIRO Mk3 Climate System Model.

Haggard, B. E., A. N. Sharpley, L. B. Massey, N. Hardiman, and K. Teague (2010), HUC 12 Watershed Prioritization Along a Water Quality and Land Use gradient: Upper Illinois River Watershed Management Plan, in *2010 Land Grant and Sea Grant National Water Conference*.

IPCC (2007), Climate Change 2007: Synthesis Report, *Tech. rep.*, Intergovernmental Panel on Climate Change, Geneva, Switzerland.

Jaccard, P. (1908), recherches sur la distribution florale., *Bulletin de la Société Vaudoise des Sciences Naturelles*, *44*, 223–270.

Jennings, E., N. Allott, D. C. Pierson, E. M. Schneiderman, D. Lenihan, P. Samuelsson, and D. Taylor (2009), Impacts of climate change on phosphorus loading from a grassland catchment: implications for future management., *Water research*, *43*(17), 4316–26, doi:10.1016/j.watres.2009.06.032.

Jha, M. K., S. Rabotyagov, and P. W. Gassman (2009), Optimal Placement of Conservation Practices Using Genetic Algorithm with SWAT, *International Agricultural Engineering Journal*, *18*(1-2), 41–50.

Jungclaus, J. H., N. Keenlyside, M. Botzet, H. Haak, J.-J. Luo, M. Latif, J. Marotzke, U. Mikolajewicz, and E. Roeckner (2006), Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM, *Journal of Climate*, *19*(16), 3952–3972, doi:10.1175/JCLI3827.1.

Kaini, P., J. W. Nicklow, and J. T. Schoof (2010), Impact of Climate Change Projections and Best Management Practices on River Flows and Sediment Load, in *World Environmental and Water Resources Congress 2010: Challenges of Change*, pp. 234–234, ASCE, doi:10.1061/41114(371)234.

Lakshmanan, V. (2000), Using a Genetic Algorithm to Tune a Bounded Weak Echo Region Detection Algorithm, *Journal of Applied Meteorology*, *39*(2), 222–230, doi:10.1175/1520-0450(2000)039⟨0222:UAGATT⟩2.0.CO;2.

Lal, R., J. a. Delgado, P. M. Groffman, N. Millar, C. Dell, and A. Rotz (2011), Management to mitigate and adapt to climate change, *Journal of Soil and Water Conservation*, *66*(4), 276–285, doi:10.2489/jswc.66.4.276.

Legutke, S., and R. Voss (1999), The Hamburg atmosphere-ocean coupled circulation model - ECHO-G, *Tech. rep.*

Lehmann, E. L. (1975), *Nonparametrics: Statistical Methods Based on Ranks*, xvi+457 pp., Holden-Day.

Leonard, R., W. Knisel, and D. Still. (1987), GLEAMS: Groundwater loading effects on agricultural management systems, *Transactions of the ASAE*, *30*(5), Transactions of the ASAE.

Lettenmaier, D. P., A. W. Wood, R. N. Palmer, E. F. Wood, and E. Z. Stakhiv (1999), Water resources implications of global warming: A US regional perspective, *Climatic Change*, *43*(3), 537–579, doi:10.1023/A:1005448007910.

Maringanti, C., I. Chaubey, M. Arabi, and B. A. Engel (2011), Application of a Multi-Objective Optimization Method to Provide Least Cost Alternatives for NPS Pollution Control, *Environmental management*, *48*(3), 448–61, doi:10.1007/s00267-011-9696-2.

Marti, O., P. Braconnot, J. Bellier, R. Benshila, S. Bony, P. Brockmann, P. Cadule, A. Caubel, S. Denvil, J.-L. Dufresne, L. Fairhead, M.-A. Filiberti, M.-A. Foujols, T. T. Fichefet, P. Friedlingstein, H. Gosse, J.-Y. Grandpeix, F. F. Hourdin, G. Krinner, C. Lévy, G. Madec, I. Musat, N. de Noblet, J. Polcher, and C. Talandier (2006), The new IPSL climate system model: IPSL-CM4.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy (2007), Fine-Resolution Climate Projections Enhance Regional Climate Change Impact Studies, *Eos, Transactions American Geophysical Union*, *88*(47), 504, doi:10.1029/2007EO470006.

McCormick, W. P., N. I. Lyons, and K. Hutcheson (1992), Distributional properties of jaccard's index of similarity, *Communications in Statistics Theory and Methods*, *21*(1), 51–68, doi:10.1080/03610929208830764.

McElroy, A., S. Chiu, J. Nebgen, A. Aleti, and F. Bennett (1976), Loading functions for assessment of water pollution from nonpoint sources, *Tech. rep.*, EPA 600/2-76-151. USEPA, Athens, GA.

Nakicenovic, N., J. Alcamo, and G. Davis (2000), *IPCC Special Report on Emissions Scenarios (SRES)*, 595 pp., Cambridge University Press.

National Research Council (2010a), *Advancing the Science of Climate Change*, The National Academies Press, Washington, D.C.

National Research Council (2010b), *Informing an Effective Response to Climate Change*, The National Academies Press, Washington, D.C.

National Research Council (2010c), *Adapting to the Impacts of Climate Change*, The National Academies Press, Washington, D.C.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams (2005), Soil and Water Assessment Tool (SWAT) Theoretical Documentation, Version 2005, (BRC Report 02-05).

Novotny, V. (1993), *Water Quality: Diffuse Pollution and Watershed Management*, 2 edition ed., 888 pp., Wiley.

Parker, G., R. Droste, and K. Kennedy (2008), Modeling the effect of agricultural best management practices on water quality under various climatic scenarios, *Journal of Environmental Engineering and Science*, *7*(1), 9–19, doi: 10.1139/S07-026.

Perez-Pedini, C., J. F. Limbrunner, and R. M. Vogel (2005), Optimal Location of Infiltration-Based Best Management Practices for Storm Water Management, *Journal of Water Resources Planning and Management*, *131*(6), 441, doi:10.1061/ (ASCE)0733-9496(2005)131:6(441).

Rabotyagov, S. S., M. Jha, and T. D. Campbell (2010), Nonpoint-Source Pollution Reduction for an Iowa Watershed: An Application of Evolutionary Algorithms, *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, *58*(4), 411–431, doi:10.1111/j.1744-7976.2010.01198.x.

Ritter, W., and A. Shirmohammadi. (Eds.) (2001), *Agricultural Nonpoint Source Pollution- Watershed Management and Hydrology*, 342 pp., Lewis Publishers, Washington, D.C.

Rummukainen, M., J. Räisänen, B. Bringfelt, A. Ullerstig, A. Omstedt, U. Willén, U. Hansson, and C. Jones (2001), A regional climate model for northern Europe: model description and results from the downscaling of two GCM control simulations, *Climate Dynamics*, *17*(5-6), 339–359, doi:10.1007/s003820000109.

Russell, G. L., J. R. Miller, D. Rind, R. A. Ruedy, G. A. Schmidt, and S. Sheth (2000), Comparison of model and observed regional temperature changes during the past 40 years, *Journal of Geophysical Research*, *105*(D11), 14,891–14,898, doi:10.1029/2000JD900156.

Safe, M., J. Carballido, I. Ponzoni, and N. Brignole (2004), Advances in Artificial Intelligence SBIA 2004, chap. On Stoppin, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/b100195.

Salas-Mélia, D., F. Chauvin, M. Déqué, H. Douville, J. Gueremy, P. Marquet, S. Planton, J. Royer, and S. Tyteca (2005), Description and validation of the CNRM-CM3 global coupled model, *Climate Dynamics*, *103*, 1–36.

Singh, V. P., and D. K. Frevert (Eds.) (2005), *Watershed Models*, 680 pp., CRC Press.

Stonefelt, M. D., T. A. Fontaine, and R. H. Hotchkiss (2000), Impacts of climate change on water yield in the upper Wind River basin, *Journal of the American Water Resources Association*, *36*(2), 321–336, doi:10.1111/j.1752-1688.2000.tb04271.x.

Tedesco, L., D. Pascual, L. Shrake, R. Hall, L. Casey, P. Vidon, F. Hernly, K. Salazar, and R. Barr (2005), Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality, *Tech. rep.*, IUPUI, Indianapolis.

US EPA (2008), Handbook for Developing Watershed TMDLs - Draft.

USDA NASS (2003), USDA-National Agricultural Statistics Service, Cropland Data Layer.

USDA NRCS (2010), Soil Data Mart.

USGS NED (2010), 1 arc Second Digital Elevation Model.

USGS NLCD (2001), 2001 Land Cover Data.

Walter, T., M. Dosskey, M. Khanna, J. Miller, M. Tomer, and J. Wiens (2007), The science of targeting within landscapes and watersheds to improve conservation effectiveness, in *Managing agricultural landscapes for environmental quality*, edited by S. M and C. C, Soil and Water Conservation Society, Ankeny, IA.

Washington, W. M., J. W. Weatherly, G. A. Meehl, A. J. Semtner Jr., T. W. Bettge, A. P. Craig, W. G. Strand Jr., J. Arblaster, V. B. Wayland, R. James, and Y. Zhang (2000), Parallel climate model (PCM) control and transient simulations, *Climate Dynamics*, *16*(10-11), 755–774, doi:10.1007/s003820000079.

WCRP (2012), Bias Corrected and Downscaled WCRP CMIP3 Climate Projections.

White, M. J., D. E. Storm, P. Busteed, S. Stoodley, and S. J. Phillips (2010), Evaluating conservation program success with Landsat and SWAT., *Environmental management*, *45*(5), 1164–74, doi:10.1007/s00267-010-9458-6.

Williams., J. R. (1990), The Erosion-Productivity Impact Calculator (EPIC) Model: A Case History, *Philosophical Transactions: Biological Sciences*, *329*(1255), 421–428.

Williams, J. R., and R. Hann (1978), Optimal operation of large agricultural watersheds with water quality constraints, *Tech. rep.*, Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier (2002), Long-range experimental hydrologic forecasting for the eastern United States, *Journal of Geophysical Research*, *107*(D20), 4429, doi:10.1029/2001JD000659.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, *Climatic Change*, *62*(1-3), 189–216, doi:10.1023/B: CLIM.0000013685.99609.9e.

Yukimoto, S., A. Noda, A. Kitoh, M. Sugi, Y. Kitamura, M. Hosaka, K. Shibata, S. Maeda, and T. Uchiyama (2001), The New Meteorological Research Institute Coupled GCM(MRI-CGCM2). Model Climate and Variability., *Papers in Meteorology and Geophysics*, *51*(2), 47–88, doi:10.2467/mripapers.51.47.

Chapter 6

SUMMARY AND CONCLUSION

Nonpoint source (NPS) pollution is the primary cause of impaired water bodies in the United States today. Control of nonpoint source pollutions is achievable through implementation of conservation practices, commonly known as Best Management Practices (BMPs). Implementation of nonpoint source pollution control practices at the watershed scale hinges on abating pollutant movement from the landscape to water bodies while minimizing cost. Thus, watershed management is inherently a multiobjective problem. An effective conservation plan should identify optimal type, location, and timing of the practices and provide information on effectiveness of the plan. However, it cannot be tested for all potential cases in a watershed scale. Thus, engineers and decision makers rely on computer models to provide an estimate of management practices impact on improving water quality. Models need to be calibrated and tested to evaluate the performance validity of watershed models according to the past observations of fluxes of water and contaminants at multiple locations on the stream network.

While significant shifts in climatic patterns are evident worldwide, many natural processes, including precipitation and temperature and therefore frequency and magnitude of extreme hydrologic events will also alter. Thus, substantial changes in diffusive transport of nonpoint source pollutants, assimilative capacities of water bodies, and landscape position of critical areas that should be targeted for implementation of NPS pollution control practices are also expected. The amount of

174

money spent on conservation programs makes it all but vital to assure the conservation benefits of practices will be sustained under the shifting climatic paradigms and challenges for adoption of the plans. Coupling of watershed models with regional climate projections can potentially provide answers to a variety of questions on the dynamic linkage between climate and ecologic health of water resources.

A decision support system was developed to present a computational framework for multiobjective decision making in nonpoint source pollution control planning at the watershed scale. The framework consists of six major modules that were linked with a nonpoint source pollution modeling model, Soil and Water Assessment Tool (SWAT): (i) a global sensitivity analysis tool to identify the most important model parameters for simulation of streamflow and water quality; (ii) a parameter estimation tool to calibrate and test performance of SWAT model in predicting the past observations of hydrologic and water quality fluxes at multiple locations; (iii) an optimization framework for identifying optimal type and placement of conservation practices; (iv) a conservation practice implementation tool that was linked with optimization engine and performed environmental and economic analysis; (v) a climate downscaling module to spatially and temporally downscale the projected climate data to study impact of climate projections on hydrologic and water quality fluxes at the watershed scale; and (vi) a multi criteria decision analysis framework, as an efficient alternative for time intensive optimization approach.

The primary objectives of the study, presented at the beginning of the Chapter 1, were met by developing frameworks and applying to a 25,000 ha watershed in the Eagle Creek Watershed (ECW) in central Indiana. Eagle Creek Watershed drains into Eagle Creek Reservoir, which supplies drinking water to the City of Indianapolis. Intense agricultural activities resulted in elevated nitrate and atrazine in excess of the EPA drinking water standards. Having extensive historic streamflow and water quality dataset makes the basin well suited for examining multisite many

objective sensitivity analysis and calibration and assessing the watershed-scale non-point source pollution control strategies under changing climate, including strategies for decreasing annual loading of nitrate and atrazine at minimum cost. The objective of developing an improved framework for multisite, many objective calibration of watershed models was met in Chapter 2. The SWAT model was calibrated and tested for daily streamflow, and monthly nitrate and pesticide observations. The calibrated SWAT model was then used for optimal placement of conservation practices. For this purpose, a multiobjective optimization framework for optimization of nonpoint source pollution control strategies at the watershed scale was presented in Chapter 3. Analyzing the results of optimization framework motivated for developing an efficient targeting framework for near-optimal placement of conservation practices. Hence, a targeting framework was proposed in Chapter 4 using *a priori* knowledge about the system. With projected changes in future climatic conditions, impact assessment of climate change on hydrologic and water quality fluxes of watershed and performance of conservation practices were studied in Chapter 5. Key findings are summarized, as follows.

6.1 Toward improved calibration of watershed models: Multisite many objective measures of information

A computational framework was presented for incorporation of disparate information from observed hydrologic responses at multiple locations into the calibration of watershed models. The framework consists of four components: (i) an *a-priori* characterization of system behavior; (ii) a formal and statistically correct formulation of objective function(s) of model errors; (iii) an optimization engine to determine the Pareto-optimal front for the selected objectives; and (iv) a multi criteria decision analysis tool to select optimal solutions from the Pareto-optimal front that are most consistent with the goals of the modeling study. Two single objective methods (Shuffled Complex Evolutionary and Dynamically Dimensioned Search)

and one multiobjective optimization method (Nondominated Sorted Genetic Algorithm II) were employed to determine optimal choices of SWAT parameter values for modeling fluxes of water and nitrate at five locations within the watershed. A proper likelihood function was derived using Bayesian statistics that can be used to reconcile observed hydrologic time series for disparate objectives at multiple stream locations. The Box-Cox transformation and first-order autoregressive model were employed in sequence to reduce heteroscedasticity and eliminate correlation between residuals. Application of the proposed framework for calibration of the Soil and Water Assessment Tool (SWAT) in the Eagle Creek watershed, Indiana, revealed that

- For a multisite many objective automatic calibration of a watershed model, both a formal likelihood function considering the structure of residuals and a multiobjective optimization approach are essential, particularly when a strict definition of system behavior is considered.

- Single objective calibration methods find a lower (better) value for the aggregated objective function of weighted errors while requiring fewer model evaluations. However, the use of the solutions from single objective techniques was limited because the simulations did not mimic the observed behavior of the system for all objectives at all sites.

- Based on a satisfactory, good, or very good classification of model simulations, multiobjective methods were the only methods that yielded behavioral solutions. Satisfying a stricter definition of the system behavior required incorporation of a separate objective function for each response at each location within the multiobjective optimization framework.

- Based on a satisfactory, good, or very good classification of model simulations, multiobjective methods were the only methods that yielded behavioral

177

solutions. Satisfying a stricter definition of the system behavior required incorporation of a separate objective function for each response at each location within the multiobjective optimization framework.

- The aggregation of information for the same response variable (nitrate in this study) at different observational sites using the proposed likelihood function appeared as a pragmatic approach for enhancing the speed of convergence to the Pareto-optimal front. However, residuals for nested sites tended to be highly correlated. Therefore, aggregation of information even for the same response should be conducted with a careful examine of residuals.

6.2 A mixed–chromosome genetic algorithm for optimal placement of conservation practices

An integrated simulation-optimization framework for optimal placement of agricultural conservation practices is presented. A novel mixed-variable multiobjective genetic algorithm based on the commonly-used NSGA-II method was coupled with a spatially distributed watershed model, Soil and Water Assessment Tool (SWAT), and was used to realize the Pareto-optimal sets of conservation practices at the watershed scale. Nondominated Sorted Genetic Algorithm II (NSGA-II) is used to find optimal suite of conservation practices type, size, and location with both binary and mixed decision variables. We also analyzed efficiency and effectiveness of the optimizations in terms of convergence rate, diversity, and optimality of the solutions. Two approaches were also taken to improve efficiency of GA algorithm by (i) updating binary to mixed-variable during the optimization, and (ii) hybridizing GA with a local search algorithm. Spatial distribution of the conservation practices type and location were also studied. Soil and Water Assessment Tool (SWAT) was used to simulate runoff and water quality and assess performance of the nonpoint

source pollution control strategies. Objectives were to simultaneously minimize nitrate load, atrazine load, and cost of implementation of nonpoint source pollution plan. Results from implementing different optimization setting showed that

1. For an optimal placement of conservation practices in a watershed-scale, discrete-continuous decision variable, referred to as "mixed-variable", optimization method identified a set of solutions which is more effective than solutions obtained from commonly used binary optimization method for the same amount of cost

2. Mixed-variable optimization provided more realistic alternatives and higher flexibility to the decision makers

3. Using mixed-variable optimization increased complexity of the optimization problem that increases computational time by several orders of magnitude. Using hybrid optimization algorithms substantially improved the efficiency of mixed-variable optimization methods

Investigating spatial distribution of the optimal conservation practices showed that there is an obvious overlay of the conservation practices type and location within consecutive solutions on Pareto-front. This overlay provides an opportunity for modular planning of watershed scale conservation strategies, while giving higher priorities to more effective and less expensive alternatives.

6.3 How *a priori* knowledge can help us in identifying optimal type and placement of conservation practices? Application of multi criteria decision making

A multi criteria decision analysis (MCDA) framework was proposed to identify a set of near-optimal solutions for nonpoint source pollution control planning at the watershed scale. The framework was established based on *a priori* knowledge about

179

the system and effectiveness of the conservation practices on reducing pollutant loads. The framework was demonstrated for identifying optimal type and location of conservation practices in the Eagle Creek Watershed, Indiana. The goal was to find a set of optimal type and location of four types of conservation practices, including grassed waterways, grade stabilization, fertilizer management, and tillage/residue management, to minimize nitrate and atrazine loads at minimum cost. Thus, three criteria were selected according to the objectives of the study that include: (1) effectiveness of conservation practices on nitrate load reduction, (2) effectiveness of conservation practices on atrazine load reduction, and (3) implementation cost of conservation practices. The study revealed that

- The proposed MCDA framework can be an effective and efficient substitute for optimization frameworks in identifying near-optimal solutions for nonpoint source pollution control plans.

- The MCDA framework outperformed optimization framework by identifying more diverse solutions within a marginal proximity of the Pareto-optimal front without any need for iterative and computationally intensive search algorithms.

- For complex problems or poorly established evaluation matrix, MCDA framework may fail to identify near-optimal solutions; however, it can effectively serve as ideal initialize population in a hybrid framework. Hybridized MCDA and NSGA-II framework outperformed the standalone optimization framework in terms of convergence (i.e. efficiency), optimality (i.e. effectiveness), and diversity of the solutions.

6.4 Assessment of optimal agricultural conservation plans under uncertain conditions: impact of climate change

Hydrologic and water quality simulations driven by an extensive ensemble of climate projections were analyzed for their respective changes in basin average temperature and precipitation in the Eagle Creek Watershed, Indiana, USA. Impact of climate change on performance of conservation practices and their optimal placement were also assessed. Nondominated Sorted Genetic Algorithm II (NSGA-II) coupled with the Soil and Water Assessment Tool (SWAT) was used for optimizing type and placement of conservation plans with the goal of reducing atrazine and nitrate load at minimum cost. The results of this study showed that

- Water yield and pollutants transport changes substantially under different climatic paradigms. Streamflow, sediment, total nitrogen, total phosphorus, and atrazine loads respectively showed 15%, 40%, 20%, 32%, and 50% net increase at the end of 21st century with peaks up to 250% higher than historical observations.

- Performance of conservation plans under different climatic projections altered considerably from what they have designed for.

- Optimization of the conservation plans for projected climatic conditions showed that despite altered Pareto-optimal solutions, spatial pattern and type of the selected conservation practices remained relatively the same.

- Similarity tests showed that as conservation plans became more effective (i.e. more money was spent), the type and location of conservation practices optimized under different climate projections became more an more similar. Existing slight difference can be explained by shifted critical areas as a result of climate change. Overall, the Pareto-optimal solutions obtained based on the

181

historic data, represented the average performance of the optimal nonpoint source pollution control plans from different climate projections.

- Watershed scale conservation planning based on the historic data can be asserted for the future by confidence. Performance of the conservation practices in the course of their lifetime (that are between 10-30 years) was expected to be sustained

6.5 Future Work

The current study was based on a number of simplifying assumptions and lacks several key analysis. Here are a few recommendations for future studies:

- This study addressed impact of climate change on performance of conservation plans. However, uncertainty in nonpoint source pollution control planning for current and future climatic and biophysical conditions could arise from several other sources, including land use/ land cover change and uncertainty in modeling. Modeling uncertainty itself includes uncertainty in model parameters, input data, and model structure. These uncertainty sources could be assessed individually and in combination with each other to provide enough information on likely effectiveness of conservation plans.

- Climate data downscaling is also associated with several types of uncertainty that could be addressed in nonpoint source pollution control planning. Different statistical and dynamic downscaling methods could be used in assessment of conservation plan's performance under different climate projections. Coupling SWAT model with regional climate models could also provide more reliable prediction on a regional scale.

- In the current study, it was assumed that representation of conservation practices in SWAT is sufficient . However, SWAT has several limitations in modeling impact of conservation practices. Some of them are due to the limitation in simulating the routing between hydrologic response units. Improvement in SWAT routing can considerably improve simulation of conservation practices which in turn can enhance the confidence on planning of nonpoint source pollution control strategies.

- In application of multi criteria decision analysis framework I listed generic factors may impact placement of conservation practices at the watershed scale. In the example presented for the Eagle Creek Watershed, I used my own judgments on scoring the conservation practices based on the previous experiences. An extensive importance scores of conservation practices on reducing pollutant loads can be obtained by analyzing the impact of individual conservation practices on reducing individual pollutants. This can also provide information on impact of biophysical and agro-environmental factors on performance of conservation plans.

- In this study, I assumed that all farmers will cooperate in implementing the conservation practices. However, chance of adoption of different conservation practices by farmers depends upon several factors, including age, income, education, awareness of environmental degradation, previous experience, adoption by neighbor farmers, and etc. Planning a nonpoint source pollution control strategy should reflect adoption behavior of the different stakeholders. Impact of financial incentives and providing educational programs should also be assessed.

- In the current study, conservation planning was performed at the 12-digit HUC that is highly recommended. To assess the cumulative impact of optimal conservation plans in a watershed scale (for example an 8-digit HUC that

encompasses several 12-digit sub-watersheds), optimization results for each sub-watershed should be aggregated in a simulation framework that reflects overall impact at the watershed scale. This may help analysts and decision makers to prioritize sub-watersheds in addition to the placement of the conservation practices.