January 2016

INTEGRATED AQUIFER VULNERABILITY ASSESSMENT OF NITRATE CONTAMINATION IN CENTRAL INDIANA

Won Seok Jang
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Won Seok Jang

Entitled

INTEGRATED AQUIFER VULNERABILITY ASSESSMENT OF NITRATE CONTAMINATION IN CENTRAL INDIANA

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Dr. Bernard A. Engel
Chair

Dr. Jon Harbor

Dr. Indrajeet Chaubey

Dr. Venkatesh Merwade

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Dr. Bernard A. Engel

Approved by: Dr. Bernard A. Engel 10/13/2016

Head of the Departmental Graduate Program Date

INTEGRATED AQUIFER VULNERABILITY ASSESSMENT OF NITRATE CONTAMINATION IN CENTRAL INDIANA

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Won Seok Jang

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2016

Purdue University

West Lafayette, Indiana

“Therefore humble yourselves under the mighty hand of God,

that He may exalt you at the proper time”

- 1 Peter 5:6 -

ACKNOWLEDGEMENTS

Foremost, I would like to thank my Lord Jesus Christ who always gives me strength, many blessings, and unconditional love. Without Him, I couldn't do anything, and I couldn't meet my forever advisor and mentor Dr. Bernard Engel.

My deep gratitude goes to Dr. Bernard Engel, my advisor, who has expertly guided me and showed me his humane aspects during my graduate education. I am fully indebted to Dr. Bernard Engel for his understanding, wisdom, patience, and encouragement.

I would like to give thanks to Dr. Jon Harbor, Dr. Indrajeet Chaubey, and Dr. Venkatesh Merwarde for their valuable guidance, assistance, and suggestions throughout my research.

Without the encouragement by Dr. Kyoung Jae Lim (Professor at Kangwon National University), I couldn't even start my graduate study at Purdue University. I am deeply grateful for his great support and guidance.

I would like to give thanks to my family members and all my friends in the U.S. and Korea for always praying for me and helping me. I would like to thank the following

friends for their continued support and friendship: Jongyoon Park, Chulmin Yeum, Larry Theller, and all my friends.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Jang, Won Seok. Ph.D., Purdue University, December 2016. Integrated Aquifer Vulnerability Assessment of Nitrate Contamination on a Watershed Scale. Major Professor: Bernard A. Engel.

Groundwater is not easily contaminated, but it is difficult to restore once contaminated. Therefore, groundwater management is important to prevent pollutants from reaching groundwater. A common step in developing groundwater management plans is assessment of aquifer risk using computational models. Groundwater modeling with a geographic information system (GIS) for efficient groundwater management can provide maps of regions where groundwater is contaminated or may be vulnerable and also can help select the optimal number of groundwater monitoring locations.

For efficient groundwater resources management, integrated aquifer vulnerability assessment is required. Integrated aquifer vulnerability assessment is incorporated into a groundwater characterization and pollutant transport analysis with tiered approaches for intrinsic aquifer vulnerability assessment (intrinsic aquifer properties) and aquifer hazard assessment (pollutant transport properties). Intrinsic aquifer vulnerability was conducted by using high resolution data to create high resolution results with DRASTIC. Aquifer hazard assessment was performed using a watershed scale hydrological model (SWAT) and a machine learning technique (Geo-ANN) developed in this study. For accurate

estimation of aquifer hazard assessment, SWAT 2012 code was modified to simultaneously calibrate streamflow and baseflow using SUFI-2. With the DRASTIC, modified SWAT, and Geo-ANN, integrated aquifer assessment was performed in the Upper White River Watershed (UWRW) located in the East central IN.

The intrinsic aquifer vulnerability results from DRASTIC without calibration were validated with observed nitrate concentrations in wells. The results showed that approximately 35.3% of nitrate detections > 2 ppm are within "High" and "Very high" vulnerability areas (represent 3.2% of vulnerability area). The results from calibrated DRASTIC showed that approximately 42.2% of nitrate detections > 2 ppm are within DRASTIC "High" and "Very high" vulnerability areas which represent only 3.4% of the area. The calibrated DRASTIC better predicted vulnerability areas using based on observed well nitrate levels > 2 ppm.

An efficient flow calibration regime (EFCR) created by incorporating modified SWAT 2012 code and SUFI-2 was developed for accurate streamflow and baseflow estimation by calibrating streamflow and baseflow simultaneously. The results of the calibration and validation in the UWRW showed that the simulated streamflow and baseflow agreed well with the observed data. With the EFCR, for the calibration period (1990 - 2001), NSE / R^2 / PBIAS for streamflow were 0.85 / 0.87 / 3.90 and NSE / R^2 / PBIAS for baseflow were 0.63 / 0.73 / 16.7. For the validation period (2002 - 2010), NSE / R^2 / PBIAS for streamflow and baseflow showed 0.88 / 0.92 / 1.50 (streamflow) and 0.65 / 0.70 / 13.8 (baseflow). These values indicate that the model is more than "Satisfactory"

for all periods. For baseflow-related studies, such as analysis of nitrate leaching for aquifer hazard assessment, simultaneous streamflow and baseflow calibration would be a reasonable approach.

For integrated aquifer vulnerability assessment in the UWRW, an integrated aquifer vulnerability map was produced by combining the intrinsic aquifer vulnerability map from DRASTIC and the aquifer hazard map from SWAT and Geo-ANN. The results of integrated aquifer vulnerability assessment were validated with observed nitrate concentrations in wells. Approximately 81.0% of well nitrate detections > 2 ppm were within "High" and "Very high" vulnerability areas that represented only 5.8% of the area. Approximately 12% of the nitrate detections were within the "Moderate" vulnerability class (30.7% of area), and 6.9% of the nitrate detections were within the "Low" vulnerability class (50.7% of area). Well nitrate levels > 2 ppm were not detected within the "Very low" vulnerability class (12.8% of area). The results indicate that integrated aquifer vulnerability assessment performed well. The integrated aquifer vulnerability assessment considers both intrinsic aquifer properties and pollutant transport properties. Thus, the overall assessment of aquifer vulnerability can be performed using the integrated aquifer vulnerability assessment technique provided in this study. Moreover, this approach is expected to be an efficient guide for managing groundwater resources for policy makers and groundwater-related researchers.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Groundwater is the primary source of water for over 1.5 billion people worldwide (Alley et al., 2002). Moreover, approximately one-half of the U.S. population depends on groundwater for its supply of potable water; approximately 36% of all municipal public drinking water supply systems and 95% of all rural populations draw potable water from groundwater resources (Conservation Foundation, 1985). Groundwater is a vital resource in Indiana as well with approximately 60% of the state's drinking water coming from groundwater. Despite its widespread use as drinking water in Indiana and globally, groundwater is a poorly understood resource by most people (Solly et al. 1998). Groundwater is also a critical component of the global environment. It offers human populations a variety of services, including water for drinking and irrigation. However, groundwater systems have been increasingly threatened, directly and indirectly, by human activities (Hamblin and Christiansen 2004).

Groundwater is typically not easily contaminated yet once this occurs, it is difficult to restore. Furthermore, groundwater pollution is not visible and is detected only when a well or spring becomes noticeably polluted or the pollutant is discharged into surface waters (Novotny 2003). Groundwater management is necessary to maintain clean

groundwater. Groundwater management has usually been facilitated by either modeling aquifer vulnerability with computational models or monitoring aquifers using groundwater sampling.

Compared with groundwater monitoring and sampling, groundwater modeling is less complex and costly, and allows evaluation of broad areas. Groundwater modeling can help select the optimal number of monitoring locations and their spatial distribution for detecting pollution in groundwater aquifers and can be useful to assess groundwater quality and provide a guide to manage groundwater efficiently (Wu 2004; Chadalavada and Datta 2008; Baalousha 2010). Therefore, if monitoring is conducted after identifying the most vulnerable areas by modeling techniques as an initial screening tool, potential monitoring sites and areas where Best Management Practices (BMPs) for groundwater quality protection can be determined in an effective and economic manner (Fienen et al. 2011).

There are two methods for effective estimation of groundwater contamination for drinking water with regard to aquifer intrinsic and transport properties (Brouyére et al., 2001). Aquifer intrinsic and transport properties are analyzed by intrinsic aquifer vulnerability assessment, and aquifer hazard assessment, respectively. Intrinsic aquifer vulnerability (intrinsic properties) is defined as natural susceptibility to contamination based on the properties of the land and subsurface, and aquifer hazard (transport properties) is regarded as pollution potential with respect to pollutant transport from the land surface to aquifers (Brouyére et al., 2001). Intrinsic aquifer vulnerability assessment

is usually conducted by overlay and index GIS models, and aquifer hazard assessment is implemented by numerical models (Pacheco and Sanches Fernandes, 2013; Chen et al., 2013; Akhavan et al., 2011; Akhavan et al., 2010).

Among various groundwater models, DRASTIC, an overlay and index GIS model, has been widely used to evaluate environmental impact associated with groundwater pollution with the use of different ratings criteria, and the strength of the vulnerability concept is that it is performed by classifying a geographical area regarding its susceptibility to groundwater contamination (Babiker et al. 2005; Akhavan et al. 2011). Advantages of the DRASTIC model include the method's low cost of application (Aller et al. 1987; Akhavan et al. 2011) and relative accuracy of model results for extensive regions (from regional scale to global scale) with a complex geological structure (Kalinski et al. 1994; McLay et al. 2001). Moreover, DRASTIC requires limited input data and has small computational needs, because there is no complex numerical analysis which requires many parameters and there is no complicated simulation process (Barbash and Resek 1996).

SWAT has been widely used for water resources management in part because SWAT has a user friendly Graphical User Interface (GUI) and a well-organized database. SWAT simulates surface flow and shallow groundwater dynamics based on hydrological response units (HRUs), which are the smallest computational units in SWAT. Surface runoff, shallow groundwater dynamics, soil water content, nutrient cycles, and sediment erosion are simulated for each HRU, and then HRUs are summed together in subbasins

(Neitsch et al., 2011). For accurate estimation of aquifer hazard using the watershed scale hydrological model (SWAT), surface and groundwater hydrology should be calibrated first, and then simulation of pollutant leaching to aquifers should be conducted (Santhi et al., 2001; Arnold et al., 2012) because nitrate leaching to aquifers is interdependent with baseflow or vice versa. Thus, streamflow and baseflow calibrations should be conducted simultaneously for accurate estimation of aquifer hazard.

Sometimes when a model cannot address a research or project problem, two or more models can be used, or a combined model can be utilized to enhance the physical representation of hydrologic processes for better estimation. Many studies related to hydrologic and water quality have been conducted to determine efficient water management using a coupled model (Noori and Kalin, 2016; Chen and Wu, 2012; Maxwell et al., 2015). The machine learning technique, ANN, is a data-driven model based on data experienced in the real world phenomena of a specific system. In contrast to analytical or numerical models, data-driven models can be used to solve problems in the field of hydrology and water resources engineering where knowledge of the physical behavior of the system and data are limited (Solomatine and Ostfeld, 2008). Thus, for the simulation of complex systems, data-driven models are useful to define the patterns within the behavior of the system (Araghinejad, 2014). ANN does not require detailed knowledge of the internal functions of a system to identify the complex, dynamic and non-linear relationships from given patterns by input and output (Ha and Stenstrom, 2003).

For efficient groundwater resources management, integrated aquifer vulnerability assessments are required. Integrated aquifer vulnerability assessments are incorporated into a groundwater characterization and risk analysis with tiered approaches for intrinsic aquifer vulnerability (aquifer intrinsic properties) assessment by DRASTIC and aquifer hazard (pollutant transport properties) assessment by SWAT and Geo-ANN (Brouyére et al., 2001). This research focuses on developing and evaluating novel techniques for integrated aquifer vulnerability assessment at the watershed scale.

1.2 Objectives

The overall goals of this study are to: 1) develop and evaluate integrated aquifer vulnerability assessment for efficient groundwater management in the Midwestern United States, and 2) modify and develop additional models (GA, SUFI-2, and Geo-ANN) for better prediction of groundwater contamination. Groundwater management has usually been conducted by either modeling aquifer vulnerability and aquifer hazard with computational models or monitoring aquifer and subsurface vulnerability by direct groundwater sampling. In this study, different modeling approaches were used to evaluate integrated aquifer vulnerability without groundwater sampling and monitoring, but groundwater data were used from different organizations such as USGS, EPA, IDEM, and Heidelberg University. The specific research objectives in this study are as follows:

1. Assessment of intrinsic aquifer vulnerability for sustainable groundwater management using DRASTIC: 1) to conduct intrinsic aquifer vulnerability assessment with DRASTIC using high resolution data, 2) to calibrate DRASTIC

weights using a binary classifier calibration method with a genetic algorithm (Bi-GA), and 3) to identify areas of potential high aquifer vulnerability and select potential aquifer monitoring and management sites for effective monitoring planning and areas where BMPs to prevent groundwater contamination might be considered.

2. Development of efficient flow calibration regime (EFCR) for accurate estimation of hydrologic and water quality components using a watershed scale hydrological model: 1) to provide the EFCR for accurate baseflow estimation with SUFI-2 and modified SWAT 2012 code, and 2) to evaluate the performance of the EFCR by streamflow and baseflow estimation.

3. Evaluation of integrated aquifer vulnerability using DRASTIC, a watershed scale hydrological model (SWAT), and Geo-ANN: 1) to generate meaningful data related to groundwater pollution by SWAT as significant input variables for aquifer hazard assessment, 2) to develop and provide a modeling guideline for aquifer hazard assessment by combining SWAT and Geo-ANN for efficient groundwater management, 3) to develop Geo-ANN which is compatible with GIS/RS data formats for a flexible hydrology and water quality modeling, and 4) to conducted integrated aquifer vulnerability assessment by incorporating intrinsic aquifer vulnerability and aquifer hazard assessment.

1.3 Dissertation Organization

This dissertation contains five chapters. Chapter 1: Introduction, focuses on providing research needs, and research objectives in this study. Chapters 2 to 4 discuss the methods and results related to the proposed objectives in the Chapter 1. These chapters are reformatted from the journal articles which are ready to submit in various journals. Chapter 2 covers objective 1 to conduct intrinsic aquifer vulnerability assessment for sustainable groundwater management using DRASTIC. Chapter 3 covers objective 2 to perform efficient flow calibration regime for accurate estimation of hydrologic and water quality components using the watershed scale hydrologic model (SWAT). Chapter 4 covers objective 3 to conduct integrated aquifer assessment using DRASTIC, SWAT, and Geo-ANN. Chapter 5, Summary and conclusions, provide expected significant findings of the study and presents recommendations for future studies.

1.4 References

Akhavan, S., Abedi-Koupai, J., Mousavi, S.F., Afyuni, M., Eslamian, S.S., Abbaspour, K.C., 2010. Application of SWAT model to investigate nitrate leaching in Hamadan–Bahar Watershed, Iran. Agriculture, Ecosystems & Environment 139, 675–688.

Akhavan, S., Mousavi, S.F., Abedi-Koupai, J., Abbaspour, K.C., 2011. Conditioning DRASTIC model to simulate nitrate pollution case study: Hamadan–Bahar plain. Environmental Earth Sciences 63, 1155–1167.

Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., Hackett, G., 1987. DRASTIC: A standardized system for evaluating groundwater potential using hydrogeologic settings. USEPA Rep.600/287/035. Ada, Oklahoma.

Alley W.M., Healy R.W., LaBaugh J.W., Reilly T.E., 2002. Flow and storage in groundwater systems. Science 296, 1985-1991.

Araghinejad, S., 2014. Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering, Springer 67, 1-5.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M. K., 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE 55, 1491-1508.

Baalousha, H., 2010. Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. Agricultural Water Management 97, 240-246.

Babiker, I.S., Mohamed, M.A.A., Hiyama, T., Kato, K., 2005. A GIS based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. Science of the Total Environment 345: 127–140.

Barbash, J.E., Resek, E.A., 1996. Pesticides in ground water: distribution, trends, and governing factors. Chelsea, M17 Ann Arbor Press.

Brouyére, S., Jeannin, P.Y., Dassargues, A., Goldscheider, N., Popescu, C., Sauter, M., Vadillo, I., Zwahlen, F., 2001. Evaluation and validation of vulnerability concepts using a physically based approach. 7th Conference on Limestone Hydrology and Fissured Media, Besnacon 20–22 Sep. 2001, Sci. Tech. Envir., Mém. H.S. No. 13, 67-72.

Chadalavada, S., Datta, B., 2008. Dynamic Optimal Monitoring Network Design for Transient Transport of Pollutants in Groundwater Aquifers. Water Resour. Manage. 22, 651-670.

Chen, J., Wu, Y., 2012. Advancing representation of hydrologic processes in the Soil and Water Assessment Tool (SWAT) through integration of the TOPographic MODEL (TOPMODEL) features, Journal of Hydrology 420-421, 319-328.

Chen, S.K., Jang, C.S., Peng, Y.H., 2013. Developing a probability-based model of aquifer vulnerability in an agricultural region. Journal of Hydrology 486, 494–504.

Conservation Foundation, 1985. Groundwater - Saving the unseen resource: Proposed conclusions and recommendations, CF, Washington, DC.

Ha, H., Stenstrom, M.K., 2003. Identification of land use with water quality data in stormwater using a neural network. Water Research 37, 4222–4230.

Hamblin W.K., Christiansen E.H., 2004. Earth's dynamic systems (10th ed.), Prentice Hall.

Kalinski, R.J., Kelly, W.E., Bogardi, I., Ehrman, R.L., Yamamoto, P.O., 1994. Correlation between DRASTIC vulnerabilities and incidents of VOC contamination of municipal wells in Nebraska. Ground Water 32(1), 31–34.

Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci. Model Dev. 8, 1-15.

McLay, C.D.A., Dragden, R., Sparling, G., Selvarajah, N., 2001. Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches. Environ. Pollut. 115, 191-204.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool, Theoretical Documentation, Version 2009. Blackland Research Center, Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, TX.

Noori, N., Kalin, L., 2016. Coupling SWAT and ANN models for enhanced daily streamflow prediction. Journal of Hydrology 533, 141–151.

Novotny, V., 2003. Water quality (2nd Ed.) - Diffuse pollution and watershed management. Wiley.

Pacheco, F.A.L., Sanches Fernandes, L.F., 2013. The multivariate statistical structure of DRASTIC model. Journal of Hydrology 476, 442–459.

Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M., 2001. Validation of the Swat Model on a Large River Basin with Point and Nonpoint Sources1. Journal of the American Water Resources Association 37, 1169-1188.

Solly W.B, Pierce R.R., Perlman H.A., 1998. Estimated use of water in the United States in 1995. U.S. Geological Survey Circular 1200.

Solomatine, D.P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. J of Hydroinformatics 10(1), 3-22.

Wu, Y., 2004. Optimal design of a groundwater monitoring network in Daqing, China. Environmental Geology 45, 527-535.

CHAPTER 2. INTRINSIC AQUIFER VULNERABILITY ASSESSMENT FOR SUSTAINABLE GROUNDWATER MANAGEMENT USING DRASTIC

2.1 Abstract

Groundwater is a vital resource in Indiana, with approximately 60% of the state's drinking water coming from groundwater. Groundwater systems are increasingly threatened, directly and indirectly, by human activities. Groundwater management is necessary to maintain clean groundwater. Groundwater management has typically been facilitated by either modeling aquifer vulnerability with computational models or monitoring aquifers through groundwater sampling. The DRASTIC model, which uses overlay and index methods, has been used by many researchers for groundwater quality assessment because it uses simple, straightforward methods. DRASTIC is useful as an initial screening tool to evaluate aquifer vulnerability in broad areas. Intrinsic aquifer vulnerability mapping identifies areas with high pollution potential, and in turn, areas for priority management and monitoring. The objectives of this study are to conduct intrinsic aquifer vulnerability assessment with DRASTIC using high resolution data, calibrate DRASTIC weights using a binary classifier calibration method with a genetic algorithm (Bi-GA), identify areas of high potential aquifer vulnerability, and select potential aquifer monitoring sites using spatial statistics. The intrinsic aquifer vulnerability results from DRASTIC using Bi-GA were validated with a well database of observed nitrate concentrations. The DRASTIC results using Bi-GA showed that approximately 42.2% of

nitrate detections > 2 ppm are within “High” and “Very high” vulnerability areas (represent 3.4% of area) as simulated by DRASTIC. Moreover, 53.4% of the nitrate detections were within the “Moderate” vulnerability class (26.9% of area), and 4.3% of the nitrate detections were within the “Low” vulnerability class (60.1% of area). In intrinsic aquifer vulnerability assessment, nitrates > 2 ppm were not detected within the “Very low” vulnerability class (9.6% of area). Intrinsic aquifer vulnerability assessment using calibration with Bi-GA better predicted nitrate detections than DRASTIC without calibration. Therefore, “High” and “Very high” vulnerability areas should be regarded as priority areas to conduct groundwater monitoring and apply practices to prevent groundwater contamination. The results of this study are expected to provide information for use as an efficient guide for managing groundwater resources by policy makers and groundwater-related researchers.

2.2 Introduction

Although groundwater accounts for a small percentage of the Earth's total water, groundwater comprises approximately thirty percent of the Earth's freshwater. Groundwater is the primary source of water for over 1.5 billion people worldwide (Alley et al. 2002). Moreover, approximately one-half of the U.S. population depends on groundwater for its supply of potable water; approximately 36% of all municipal public drinking water supply systems and 95% of all rural populations draws potable water from groundwater resources in the U.S. (Conservation Foundation 1985). Groundwater is a vital resource in Indiana with approximately 60% of the state’s drinking water coming

from groundwater. Despite its widespread use as drinking water in Indiana and globally, groundwater is a poorly understood resource by most people (Solly et al. 1998).

Groundwater is also a critical component of the global environment. It offers human populations a variety of services, including water for drinking and irrigation. However, groundwater systems have been increasingly threatened, directly and indirectly, by human activities. In addition to the challenges posed by land use / land cover (LULC) change, environmental pollution, and water diversion, groundwater systems are expected to be stressed by global climate change (Hamblin and Christiansen 2004).

Groundwater is typically not easily contaminated yet once this occurs, it is difficult to restore. Furthermore, groundwater pollution is not visible and is detected only when a well or spring becomes noticeably polluted or the pollutant is discharged into surface waters (Novotny 2003). Groundwater management is necessary to maintain clean groundwater. Groundwater management has usually been facilitated by either modeling aquifer vulnerability with computational models or monitoring aquifers using groundwater sampling.

Groundwater monitoring and sampling have benefits in analyzing groundwater because they can estimate groundwater quality and quantity directly in real time. A groundwater monitoring network can provide quantity and quality data necessary to make informed decisions regarding the state of the environment. A properly designed monitoring system provides a representative understanding of the state of the monitored area (Baalousha

2010). Groundwater monitoring and sampling, however, have disadvantages because they are complex, difficult to apply for broad areas, and are a costly undertaking. Also, improper distribution of monitoring sites or an insufficient number of sites may not provide a representative view of the state of the environment. On the other hand, if the sampled sites are too numerous, the information obtained is redundant and the monitoring network is costly and inefficient (Baalousha 2010).

Compared with groundwater monitoring and sampling, groundwater modeling is less complex and costly, and allows easy evaluation of broad areas. Groundwater modeling could help select the optimal number of monitoring locations and their spatial distribution for detecting pollution in groundwater aquifers and could be useful to assess groundwater quality and provide a guide to manage groundwater efficiently (Wu 2004; Chadalavada and Datta 2008; Baalousha 2010). However, if only modeling techniques are used, groundwater quality and quantity would be indirectly estimated and could not be calibrated and validated. Therefore, if monitoring is conducted after identifying the most vulnerable areas by modeling techniques as an initial screening tool, potential monitoring sites and areas where Best Management Practices (BMPs) for groundwater quality protection can be determined in an effective and economic manner (Fienen et al. 2011).

There are various groundwater models such as MODFLOW, GSFLOW, and GWM-2005 (McDonald et al. 1998; Ahlfeld et al. 2005; Harbaugh 2005; Markstrom et al. 2008) which have been widely used to evaluate groundwater quality. They require significant input data to run, and for most users it is not easy to use the models because they are

complicated. Moreover, they have limitations to simulate large areas. For these reasons, DRASTIC, which uses overlay / index methods, has been used by many researchers for efforts related to groundwater quality assessment because DRASTIC uses simple and straightforward methods (Pacheco and Fernandes 2013; Chen et al. 2013).

On the other hand, the DRASTIC method has received criticism due to limited validation. Holden et al. (1992) and Maas et al. (1995) reported little correlation between model results and field data. Navulur (1996) used DRASTIC and low resolution aquifer vulnerability maps were created using low resolution input data. In spite of these concerns, DRASTIC has been widely used to evaluate environmental impact associated with groundwater pollution with the use of different ratings criteria, and the strength of the vulnerability concept is that it is performed by classifying a geographical area regarding its susceptibility to groundwater contamination (Babiker et al. 2005; Akhavan et al. 2011). Advantages of the DRASTIC model include the method's low cost of application (Aller et al. 1987; Akhavan et al. 2011) and relative accuracy of model results for extensive regions with a complex geological structure (Kalinski et al. 1994; McLay et al. 2001). Moreover, DRASTIC requires limited input data and has small computational needs, because there is no complex numerical analysis which requires many parameters and there is no complicated simulation process (Barbash and Resek 1996). DRASTIC is a reconnaissance tool, but has proven its value as an indicator of areas deserving detailed hydrogeologic evaluation. It is useful as an initial screening tool to evaluate aquifer vulnerability in a broad area. Aquifer vulnerability mapping identifies areas with high pollution potential, and in turn, priorities for monitoring (Baalousha 2010). For these

reasons, DRASTIC was applied in this study for intrinsic aquifer vulnerability assessment and for identifying groundwater monitoring locations. Disadvantages of DRASTIC identified in previous studies were modified to improve estimation of aquifer vulnerability in this study.

The objectives of this study are to 1) conduct intrinsic aquifer vulnerability assessment with DRASTIC using high resolution data, 2) calibrate DRASTIC weights using a binary classifier calibration with a genetic algorithm (Bi-GA), and 3) identify areas that are potential high aquifer vulnerability and select potential aquifer monitoring and management sites for effective monitoring planning and areas where BMPs to prevent groundwater contamination might be considered.

2.3 Methodology

2.3.1 DRASTIC to Estimate Aquifer Vulnerability

Various groundwater vulnerability assessment approaches have been developed to evaluate aquifer vulnerability. These include process based methods, statistical methods, and overlay / index methods (Zhang et al. 1996; Tesoriero et al. 1998). The process based methods use simulation models (i.e., SWAT (Arnold et al. 1998), HSPF (Bicknell et al. 2001), and GLEAMS (Knisel 1999)) to simulate contaminant transport (Barbash and Resek 1996). Statistical methods identify relationships between simulated results or spatial variables and observed data in the aquifer. The overlay / index methods (i.e., AVI (van Stemproot et al. 1993), COP (Vías et al. 2006), DRASTIC (Aller et al. 1987), GOD (Foster 1987), and IRISH (Daly and Drew 1999)) are based on assembling information

on the most relevant characteristics affecting aquifer vulnerability. Using overlay / index methods, aquifer vulnerability is evaluated by scoring, integrating, or classifying the information to produce an index, rank, or class of vulnerability (Harter and Walker 2001). The overlay / index methods are easy to apply, especially on regional or larger areas. Therefore, these are the most popular methods used in aquifer vulnerability assessment for various spatial scales (from local to global scale).

DRASTIC is a conceptual model defined as a composite description of the most important geological and hydrological factors that could potentially affect aquifer pollution. DRASTIC yields a numerical index map that is derived from ratings and weights assigned to the seven map parameters (Aller et al. 1987; Akhavan et al. 2011).

DRASTIC has four assumptions (Al-Zabet 2002) as described below:

1) The pollutant is introduced at the ground surface

2) The pollutant is flushed into the groundwater by precipitation

3) The pollutant has the mobility of water

4) The area being evaluated using DRASTIC is 40 hectares (=0.4 km2) or larger

DRASTIC is a numerical ranking system, which uses weights, ranges, and ratings to provide groundwater vulnerability. The DRASTIC index is calculated using Equation 2.1. The higher the DRASTIC index score, the greater the groundwater vulnerability. The smallest possible DRASTIC index is 23 and the largest is 230, if the range of DRASTIC weights ranges from 1 to 5.

$$\text{DRASTIC} = \sum_{i=1}^{n1} D_{wi} D_{ri} + \sum_{i=1}^{n2} R_{wi} R_{ri} + \sum_{i=1}^{n3} A_{wi} A_{ri} + \sum_{i=1}^{n4} S_{wi} S_{ri} + \sum_{i=1}^{n5} T_{wi} T_{ri} + \sum_{i=1}^{n6} I_{wi} I_{ri} + \sum_{i=1}^{n7} C_{wi} C_{ri} \quad (2.1)$$

Where:

D_r = Ratings to the depth to water table

D_w = Weight assigned to the depth to water table

R_r = Ratings for ranges of aquifer recharge

R_w = Weight for aquifer recharge

A_r = Ratings assigned to aquifer media

A_w = Weight assigned to aquifer media

S_r = Ratings for soil media

S_w = Weight for soil media

T_r = Ratings for topography

T_w = Weight assigned to topography

I_r = Ratings assigned to vadose zone

I_w = Weight assigned to vadose zone

C_r = Ratings for rates of hydraulic conductivity

C_w = Weight given to hydraulic conductivity

With various DRASTIC weights, ranges, and ratings (Table 2.1), users can assign ratings and weights in determining D, R, A, S, T, I, and C maps. The variable rating allows users to select either a typical value or to modify the value based on users' experience and knowledge in a specific area. The DRASTIC model was designed to allow users to make a flexible modification so that the local hydrogeological characteristics could be reflected and its parameters could be weighted properly (Rahman 2008).

DRASTIC has also been applied to many regions around the world. Babiker et al. (2005) estimated aquifer vulnerability and demonstrated the combined use of DRASTIC and GIS in Kakamigahara Heights, Gifu Prefecture, Central Japan. They utilized sensitivity analyses to evaluate the relative importance of the model parameters for aquifer vulnerability. Navulur (1996) developed a technique for estimating groundwater vulnerability to nitrate contamination from non-point sources (NPS) on a regional scale. The technique was applied to evaluate vulnerability of groundwater systems in Indiana, United States using a GIS environment with 1:250,000 scale data. Vulnerability of Indiana aquifer systems to NPS of pollution was also evaluated using DRASTIC and SEEPAGE analyses.

Table 2.1 Typical DRASTIC ranges and ratings

Depth to water (m)	
Range	Rating
0 - 1.5	10
1.5 - 4.6	9
4.6 - 6.8	8
6.8 - 9.1	7
9.1 - 12.1	6
12.1 - 15.2	5
15.2 - 22.9	4
22.9 - 26.7	3
26.7 - 30.5	2
30.5+	1

Net recharge (mm/yr)	
Range	Rating
254+	10
235 - 254	9
216 - 235	8
178 - 216	7
147.6 - 178	6
117.2 - 147.6	5
91.8 - 117.2	4
71.4 - 91.8	3
51 - 71.4	2
0 - 51	1

Aquifer media	
Range	Rating
Karst limestone	10
Basalt	9
Sand and gravel	8
Massive sandstone	7
Massive limestone	7
Bedded sandstone	6
Limestone	6
Glacial till	5
Weathered metamorphic igneous	4
Metamorphic igneous	3
Massive shale	2

Soil media	
Range	Rating
Thin or absent / Gravel	10
Sand	9
Peat	8
Shrinking clay	7
Loamy sand	6
Sandy loam	6
Loam	5
Sandy clay	4
Sandy clay loam	4
Silt loam	4
Silty clay	3
Clay loam	3
Silty clay loam	3
Muck	2
Non-shrinking clay	1

Topography (%)	
Range	Rating
0 - 2	10
2 - 6	9
6 - 12	5
12 - 18	3
18+	1

Vadose zone media	
Range	Rating
Thin or absent / Gravel	10
Sand	9
Peat	8
Shrinking clay	7
Loamy sand	6
Sandy loam	6
Loam	5
Sandy clay	4
Sandy clay loam	4
Silt loam	4
Silty clay	3
Clay loam	3
Silty clay loam	3
Muck	2
Non-shrinking clay	1

Hydraulic conductivity (m/s)	
Range	Rating
0.00095 +	10
0.0005 - 0.00095	8
0.00033 - 0.0005	6
0.00015 - 0.00033	4
0.00005 - 0.00015	2
0.00000015 - 0.00005	1

2.3.2 Genetic Algorithm for Optimized Calibration

In hydrologic and water quality modeling, parameter calibration and uncertainty analysis are important steps to avoid over or under estimating modeling predictions (Wu and Liu 2012; Strauch et al. 2012). In recent years, hydrologic and water quality models have become more complicated. For this reason, manual calibration is almost infeasible due to model complexity. Instead of manual calibration, semi-automated or automated

calibration techniques have been widely used by many researchers who utilize hydrologic and water quality models (Thompson et al. 2013; Green and van Griensven 2008).

Among various optimization algorithms for calibration, GA has become increasingly popular for optimization problems (Wang 1991; Liu et al. 2013; Song et al. 2012; Atia et al. 2012). GA is based on heuristic combinatorial search techniques inspired by evolutionary biology of natural selection and genetics of Darwin's evolution principle. GA is designed to find a solution that maximizes the fitness function (objective function) which is a user supplied function that describes the fitness of a particular solution. With each succeeding generation, the genetic algorithm transforms a population into better performing individuals as defined by the objective function (Holland 1975; Goldberg 1989). The concept of GA is to simulate the natural evolution mechanisms of chromosomes (or strings) including selection, crossover, and mutation. GA operates on a population of decision variable sets which are called chromosomes or strings. The procedures of GA start with initializing a population generation of solutions. In this stage, GA generates random strings from the parameter space. Each string is a set of values of optimization parameters. In each generation, the individual strings are selected by survival of fittest among string structures based on the value of objective function. Next, GA generates a new generation by selection, crossover, and mutation operations to evolve the solutions in order to find the best one(s). GA runs for a maximum number of generations or until some stopping criterion is met (Zhang et al. 2009). Users can define and specify variables to suit each study's purpose, and GA can be applied to a wide range of areas ranging from science and engineering to business. For these reasons, GA will be

used for calibration of DRASTIC weights to obtain accurate estimation of nitrate leaching and its effect on aquifer vulnerability in this study.

2.3.3 Study Area

The study area (Figure 2.1) is the Upper White River Watershed (UWRW) (Latitude: 39°29'51"N, Longitude: 86°24'02"W) in Indiana. The UWRW is a Hydrologic Unit Code (HUC) 8 watershed (05120201) located in central Indiana and includes seventeen HUC 10 subwatersheds. UWRW is important for public drinking water supplies because UWRW includes more than 3,508 km of streams, numerous artificial lakes, and 4 reservoirs. Sixteen counties are located in the watershed, and the UWRW serves as a portion of the drinking water supply for the city of Indianapolis which is Indiana's largest city. The water sources in the rural areas of UWRW traditionally are individual wells to provide groundwater for residential, commercial, and industrial purposes (Tedesco et al., 2011; Fleming et al., 1993). The UWRW was selected to evaluate the improved DRASTIC (ver.2015).

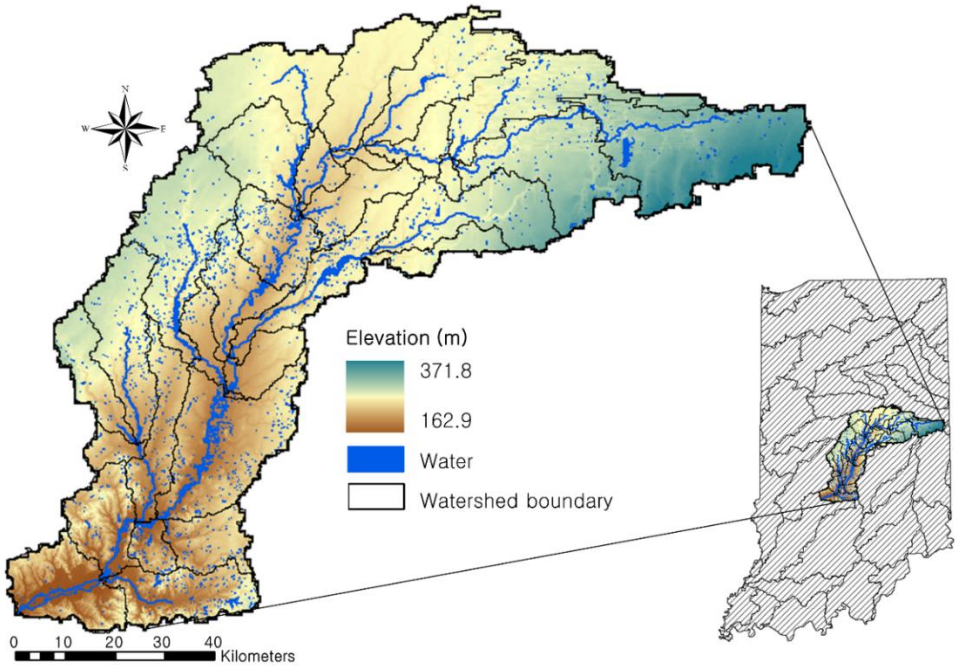


Figure 2.1 Location of the Upper White River Watershed

2.3.4 Sources of Data

In this study, improved data (Table 2.2) were applied in the UWRW as described below. Most data for DRASTIC ver.2015 are of 1:24,000 scale, unlike the 1:250,000 scale data for DRASTIC ver.1996. The data include water table depth, precipitation, evapotranspiration, LULC, aquifer systems, SSURGO used to produce recharge, soil media, and topography layers, lithology, and aquifer transmissivity data (Table 2.2).

Table 2.2 Data used for creating DRASTIC input data

Data type	Source	Format	Scale	Date	Used to produce
Water well	IDNR[1]	Point Shapefile	1:24,000	1959 - 2010	**D**epth to water
Annual Precipitation	NCDC[2]	Tabular data	-	1949 - 2013	**R**echarge
LULC	MRLC[3]	Raster	1:250,000	2006	**R**echarge
Aquifer Systems	USGS[4]	Polygon Shapefile Text	1:48,000	2003 - 2011	**A**quifer media
SSURGO[5]	NRCS[6]	Polygon Shapefile	1:12,000	2005	**R**echarge **S**oil media **T**opography
iLITH data	IGS[7]	Point Shapefile	1:24,000	2001	**I**mpact of vadose
Aquifer Transmissivity	IDNR[1]	Point Shapefile	1:24,000	2011	**C**onductivity

[1]IDNR: Indiana Department of Natural Resources

[2]NCDC: National Climate Data Center

[3]MRLC: Multi-Resolution Land Characteristics Consortium

[4]USGS: U.S. Geological Survey

[5]SSURGO: Soil Survey Geographic Database

[6]NRCS: Natural Resources Conservation Service

[7]IGS: Indiana Geological Survey

2.3.5 Nitrate Measurements

In order to calibrate the DRASTIC index map, nitrate concentration was selected as the contaminant parameter. Nitrate levels in groundwater under natural condition are typically less than 2 ppm in Indiana. Any nitrate detection > 2 ppm has been assumed to be caused by human activities. Thus, a threshold value of background concentration has been set at 2 ppm in this study (Navulur, 1996). 116 wells (total 678 wells) (Figure 2.2) with nitrate levels > 2 ppm (> 2 mg/l) were selected to calibrate and validate estimated high aquifer vulnerability areas. Nitrate levels vary from 0.1 to 18.3 mg/l with an average of 1.2 mg/l in the study area.

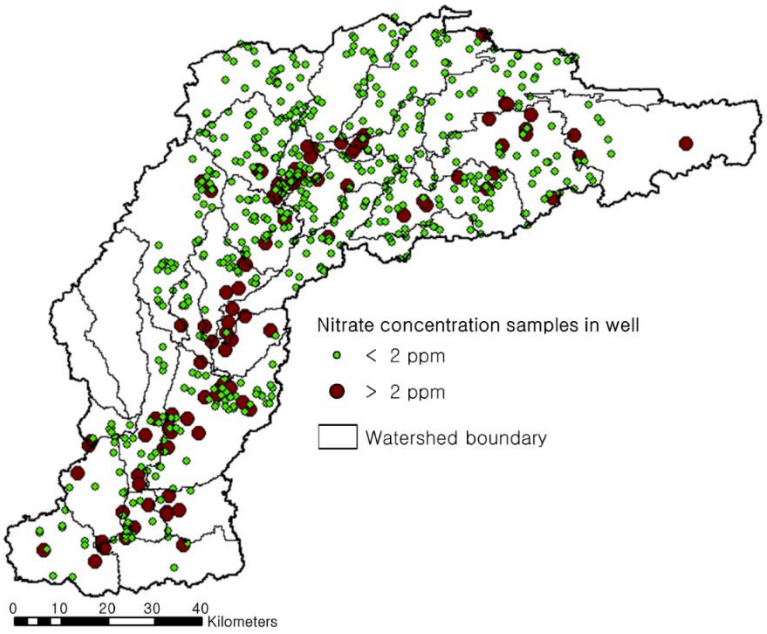


Figure 2.2 Nitrate concentration samples in wells in the UWRW

2.3.6 Intrinsic Aquifer Vulnerability Mapping using DRASTIC

Methods described below were used for DRASTIC to create an intrinsic aquifer vulnerability map with a high resolution. The seven map layers (Table 2.3), representing the seven parameters of DRASTIC, were prepared to create the intrinsic aquifer vulnerability mapping for the UWRW. DRASTIC ratings and weights were assigned to each map according to DRASTIC standards (Aller et al., 1985). Then, weights were modified to reflect local characteristics for aquifer vulnerability maps. Finally, model calibration was conducted for more accurate predictions (Figure 2.3).

Intrinsic aquifer vulnerability indices were divided into five classes (“Very low”, “Low”, “Moderate”, “High”, and “Very high” vulnerability classes) by normalization of DRASTIC indices using Equation 2.2. Feature scaling (data normalization) is a method

used to standardize the range of independent variables (min=0, max=1). It is generally utilized during the data preprocessing procedure. The map resolutions of previous DRASTIC aquifer vulnerability maps usually are crude. For more detailed evaluation of aquifer vulnerability, high resolution aquifer data would be required and innovative approaches which were used to produce DRASTIC input data were applied in this study as well. Also, DRASTIC parameters were optimized to adjust DRASTIC weights with Bi-GA.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2.2}$$

where x' is the normalized value and x is the original value.

Table 2.3 Description of DRASTIC parameters and DRASTIC original weights

DRASTIC parameters	Description	Original Weight
Depth to water (D)	Depth from the ground surface to the water table. Deeper water table levels imply lesser contamination chances	5
Recharge (R)	The amount of water which enters the aquifer. The amount of recharge is positively correlated with the vulnerability rating	4
Aquifer media (A)	Material property of the saturated zone, which controls the pollutant attenuation processes based on the permeability of each layer of media.	3
Soil media (S)	Soil media affects the transport of the contaminant and water from the soil surface to the aquifer	2
Topography (T)	Slope of the land surface. Topography of the land affects groundwater vulnerability. With a low slope, the contaminant is less likely to become runoff and therefore more likely to infiltrate the aquifer	1
Impact of vadose zone media (I)	Vadose zone is the typical soil horizon above the water table and below the ground surface. If the vadose zone is highly permeable then this will lead to a high vulnerability rating	5
Hydraulic conductivity (C)	Hydraulic conductivity represents the ability of the aquifer to transmit water. Hydraulic conductivity is positively correlated with the vulnerability rating	3

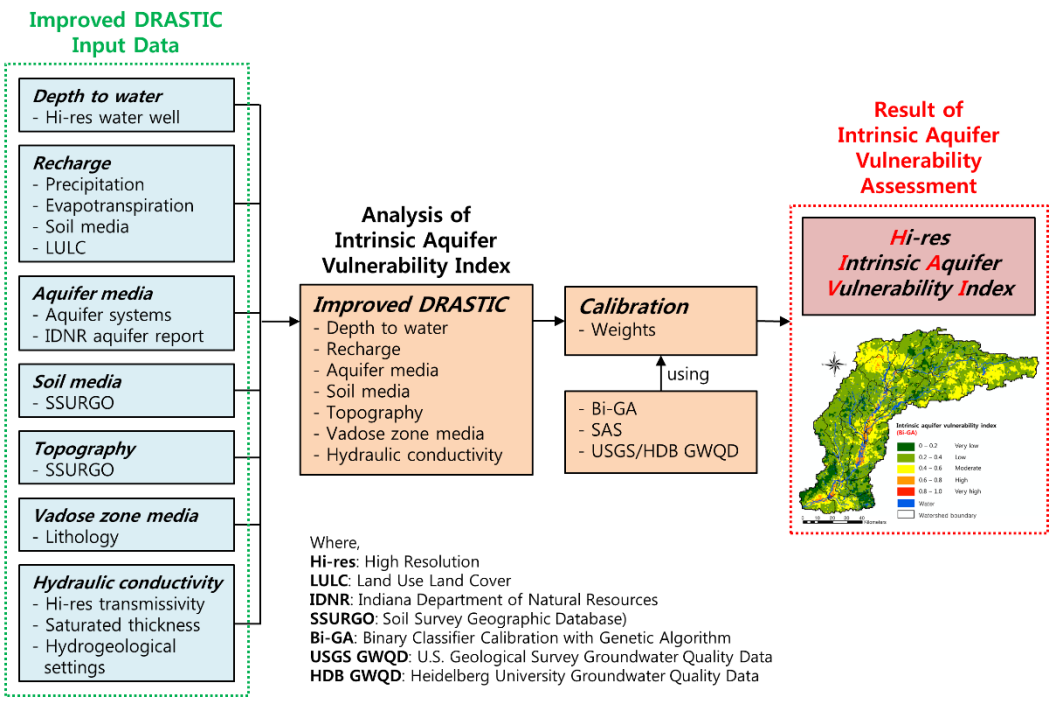


Figure 2.3 Flowchart for analysis of intrinsic aquifer vulnerability mapping

2.3.6.1 Depth to Water

High spatial resolution long-term static water level data in wells (1990 ~ 2008 years) were utilized to obtain the 'Depth to water' (D) map and interpolation was required. Minimum, maximum, and average depths to water are 3.1 m, 167.6 m, and 10.6 m, respectively. Kriging interpolation was used because this method is an effective way to interpolate a limited number of observations for hydrologic properties, such as rainfall, aquifer characteristics, effective recharge and to preserve the theoretical spatial correlation (de Marsily, 1984). Then, interpolated data were reclassified into ratings according to Table 2.1.

2.3.6.2 Recharge

Land uses / land covers (LULCs) and soils are sensitive parameters in calculating the amount of recharge. To consider regional LULC and soil characteristics, the Soil Conservation Service (SCS) runoff curve number (CN) method was used to estimate potential recharge with annual precipitation, soils and LULCs. Potential recharge in DRASTIC was computed by precipitation minus surface runoff which is determined by the SCS-CN method. Annual precipitation data from National Climate Data Center (NCDC), LULC data from National Land Cover Database (NLCD) and soil data (SSURGO) from Natural Resources Conservation Service (NRCS) were used to produce a potential recharge map using the SCS-CN method (Equation 2.3). Then, DRASTIC rating for annual potential recharge was determined using Table 2.1 (Yang and Wang, 2010; Nobre et al., 2007). Even though the calculation for potential recharge is straightforward, the results (Table 2.1) show various ranges of the amount of recharge because the calculation reflects the combination of soils and LULCs based on the SCS-CN method. When detailed recharge estimation is needed for small areas with DRASTIC, the approach to estimating recharge suggested in this study would be more useful than other studies (Babiker et al., 2005; Chen et al., 2013), which calculates the amount of recharge using only two LULC classes (e.g., urban and remaining areas).

Many studies have used the CN method or hydrologic soil group (HSG) to estimate potential recharge in DRASTIC (Yang and Wang, 2010; Nobre et al., 2007; Poiani, 1996; Zomorodi, 2004) because the concept of HSG includes soil permeability and available water capacity related to infiltration rate (i.e., Soil A having high infiltration rates, soil B

having moderate infiltration rates, soil C having slow infiltration rates, and soil D having very slow infiltration rates), and HSG indicates areas with different susceptibilities to infiltration. Potential recharge estimated in this study may not reflect the actual amount of recharge but rather indicates possible recharge rate. Estimation of potential recharge (potential infiltration rate) used the concept of HSG, while evapotranspiration (ET) was ignored when calculating potential recharge because ET occurs after infiltration. Thus, even though there is a limitation in ignoring ET when calculating potential recharge, this approach for estimating potential recharge has been used in DRASTIC (Yang and Wang, 2010; Nobre et al., 2007; Poiani, 1996; Zomorodi, 2004).

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S} \text{ (When } P > I_a \text{)} \quad (2.3)$$

$$I_a = 0.2S$$

$$S = \frac{25,400}{CN} - 254$$

Where,
Q: Depth of runoff (mm)
P: Depth of rainfall (mm)
I_a: Initial abstraction (mm)
S: Maximum potential retention (mm)
CN: Curve number (dimensionless)

2.3.6.3 Aquifer Media

An ‘Aquifer media’ (A) map was created using the aquifer systems map and report by U.S. Geological Survey (USGS) and Indiana Department of Natural Resources (IDNR). Most aquifer media of the study area were sand and gravel, but based on the INDR reports, aquifer media rating was assigned in more detail (IDNR, 2012). IDNR reports of

counties in Indiana described vulnerability of each aquifer system such as ‘very high susceptibility to surface contamination (very high)’, ‘highly susceptible to surface contamination (high)’, ‘moderately susceptible to surface contamination (moderate)’, ‘low susceptibility to surface contamination (low)’, and ‘very low susceptibility to surface contamination (very low)’. Vulnerability ratings were divided into five levels (very high=10, high=8, moderate=6, low=4, and very low=2). Then, modified reclassification of ratings was conducted as shown in Table 2.4.

Table 2.4 Typical and modified ranges and ratings of aquifer media (A)

Aquifer media			
Range	Rating (typical)	Vulnerability (IDNR report)	Rating (modified)
Karst limestone	10	Very high	10
Basalt	9	High	8
Sand and gravel	8	Moderate	6
Massive sandstone Massive limestone	7	Low	4
Bedded sandstone Limestone Shale	6	Very Low	2
Glacial till	5		
Weathered metamorphic	4		
Metamorphic Igneous	3		
Massive shale	2		

2.3.6.4 Soil Media and Topography

‘Soil media’ (S) and ‘Topography’ (T) maps were obtained through SSURGO data from USDA-NRCS instead of STATSGO data normally used. The map scale of SSURGO data is 1:12,000, whereas STATSGO is 1:250,000. Of many fields of the SSURGO

table, ‘MUNAME’ is needed to analyze DRASTIC S and required information is a soil type (e.g., loam, silt loam, and sandy loam). However, a MUNAME field in the original SSURGO table describes detailed soil type such as ‘Martinsville loam, 1 to 5 percent slopes’, ‘Jasper silt loam, 1 to 5 percent slopes’. This detailed information is unnecessary for DRASTIC ratings, because extracting only soil type of a number of fields is time-consuming. Therefore, an essential database to produce DRASTIC S was constructed using ArcGIS 10.2 and Python programming, and DRASTIC S and T maps were generated using a modified SSURGO data table. The S and T maps have ratings as described in Table 2.1.

2.3.6.5 Impact of Vadose Zone Media

‘Impact of vadose zone media’ (I) map was estimated using sand, silt, and clay thickness point data within lithology data from IDNR. Kriging interpolation was implemented to estimate unknown areas with known scattered data points, and DRASTIC ratings were assigned according to Table 2.1.

2.3.6.6 Hydraulic Conductivity

The ‘Hydraulic Conductivity’ (C) map (equation 2.4) was calculated with high resolution transmissivity (1:24,000) and saturated thickness data from IDNR based on hydrogeological settings. The C map was then reclassified into ranges and assigned ratings from 1 to 10 according to Table 2.1. Regions with higher hydraulic conductivity have a greater possibility of contamination.

Hydraulic conductivity (m/s) = Transmissivity (m^2/s) / Thickness of aquifer (m) (2.4)

2.3.7 Model Calibration

Probabilistic predictive or decision-making models (i.e., DRASTIC and SEEPAGE) needs post-processing prior to calibration with observed data because results of probabilistic predictive or decision-making models have different scale or format with observed data for calibration of the models (Naeini et al., 2015; Xu et al., 2016). Thus, in this study, a binary classifier calibration method was combined with a genetic algorithm (hereafter referred to as Bi-GA) and used for calibration of DRASTIC weights. In the binary classifier calibration, a result of model and observed data are classified as 0 or 1. DRASTIC produces five vulnerability class (i.e., very high, high, moderate, low, very low). This study classified that very high and high vulnerability class are 1 and other classes are 0 and over 2 ppm and below of observed nitrate concentrations in wells are 1 and 0, respectively because nitrate detections > 2 ppm are conducted with DRASTIC. Bi-GA was utilized for calibration of DRASTIC weights using Heidelberg University and USGS groundwater quality data which are mean nitrate concentration data from 116 wells with nitrate levels > 2 ppm which is the threshold value for the background concentration level of nitrate. Calibration with Bi-GA was conducted to improve the performance of the DRASTIC model. Original DRASTIC weights vary from 1 to 5 (Table 2.3). Based on original DRASTIC weights, Depth to water (D) and Impact of vadose zone (I) are the most sensitive parameters, and the second most sensitive parameter is Recharge (R) in assessing an aquifer vulnerability. However, in many studies using DRASTIC, original relationships of DRASTIC weights between the seven

maps layers have been ignored in the calibration process. For this study, even though calibrated DRASTIC weights were different from original values, the ratio of DRASTIC weights (5 (D), 4 (R), 3 (A), 2 (S), 1 (T), 5 (I), and 3 (C)) was maintained. The Bi-GA modified the ratio of DRASTIC weights based on calibrated weights with Bi-GA, and there are weight boundaries which are ±1 from original DRASTIC weights. For instance, maximum and minimum weight values of D are 6 and 4. Based on calibrated weights by the Bi-GA which ignored the original DRASTIC weight relationships between the seven map layers, new calibrated weights which consider the ratio of DRASTIC weights (5 (D), 4 (R), 3 (A), 2 (S), 1 (T), 5 (I), and 3 (C)) were generated by 1) no calibration and 2) calibration with Bi-GA. Root mean square error (RMSE) was used to evaluate the effectiveness of the Bi-GA and its ability to make predictions in the calibration procedure (Equation 2.5). The GA driving variables used in this study are shown in Table 2.5.

An accuracy assessment error matrix was computed to validate the results following calibration using Bi-GA. Using an accuracy assessment error matrix, spatial patterns in success (detections of nitrate concentration in wells over 2 ppm) and failure (detections of nitrate concentration in wells under 2 ppm) of DRASTIC prediction were analyzed with a total accuracy.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (S_i - O_i)^2}{n}} \tag{2.5}$$

Where

S_i : Simulated nitrate concentration DRASTIC binary value

O_i : Observed nitrate concentration binary value

Table 2.5 Driving variables in GA for DRASTIC parameter optimization

GA driving variables	Values
Population size	100
Max generation	10,000
Initial random value	1,000
Min. value of parameters	0
Max. value of parameters	6
Crossover probability	0.5
Mutation probability	0.02

2.3.8 Intrinsic Aquifer Vulnerability Mapping

The intrinsic aquifer vulnerability map was created by combining the seven map layers after multiplying each map layer with its theoretical ratings and weights (Equation 2.1). Then, calibration for DRASTIC weights was carried out using Bi-GA and the statistical methods. Two DRASTIC result maps with no calibration and calibration with Bi-GA had different ranges of DRASTIC vulnerability indices because different weight values were applied for each DRASTIC result map. Finally, optimized high resolution intrinsic aquifer vulnerability predictions with calibrated DRASTIC weights were generated using the Spatial Analyst tool in ArcGIS 10.2.

2.3.9 Evaluation for Potential Groundwater Monitoring Sites

Hotspot analysis using the Getis-Ord Gi* (Gi*) (Getis and Ord, 1992; Ord and Getis, 1995) was applied to select potential groundwater monitoring sites. This method works by examining each feature (each grid cell) within the context of neighboring features. If

a feature has a high value (high DRASTIC vulnerability index) and is surrounded by other features with high values, this feature is defined as a hotspot with statistical significance. The Gi* statistic returned for each feature in the dataset is a z-score. For statistically significant positive z-scores, the larger the z-score, the more intense the clustering of high values (hotspot) (Mitchell, 2005). Thus, potential groundwater monitoring sites (where aquifer may be the most vulnerable to contamination) would be found based on the z-score with statistical significance. The hotspot analysis using the Gi* statistic was conducted using Equations 2.6 - 2.8 with Spatial Analyst in ArcGIS 10.2.

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \quad (2.6)$$

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \quad (2.7)$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (\bar{X})^2} \quad (2.8)$$

Where

G_i^*: Getis-Ord local statistic

x_j: Attribute value for feature j

$w_{i,j}$: Spatial weight between feature i and j

n : Total number of features

2.4 Results and Discussion

2.4.1 Calibration of DRASTIC Weights

For accurate estimation of intrinsic aquifer vulnerability, calibration of DRASTIC weights was conducted before producing an intrinsic vulnerability map. 116 wells with nitrate levels > 2 ppm were used to calibrate DRASTIC weights for better prediction of intrinsic aquifer vulnerability. As shown in Table 2.6, RMSE for intrinsic aquifer vulnerability without calibration was 0.70. The RMSE for intrinsic aquifer vulnerability with calibrated DRASTIC parameters using Bi-GA was 0.57. RMSE for Bi-GA might be little decreased (the lower RMSE, the better performance) because calibrated DRASTIC weights using Bi-GA maintained the ratios of original DRASTIC weights. Previous studies did not maintain the ratios of original DRASTIC weights to improve just performance evaluation of DRASTIC. However, if the ratios of original DRASTIC weights are not maintained, the number of degrees of freedom of the DRASTIC index (result scores for aquifer vulnerability) would be increased by calibrating DRASTIC weights. Further, physical properties for intrinsic aquifer vulnerability could potentially be ignored.

For validation of the results by using Bi-GA, accuracy assessment was computed. As shown in the Tables 2.7 and 2.8, total accuracies of uncalibrated DRASTIC and calibrated DRASTIC were 35% and 42%, respectively. Thus, the results of accuracy

assessment indicate calibrated DRASTIC predicted intrinsic aquifer vulnerability areas contaminated by human activities more accurately than uncalibrated DRASTIC did.

DRASTIC, an overlay and index GIS model, does not compute nitrate concentrations in aquifers, rather it predicts intrinsic aquifer vulnerability classes from very high vulnerability to very low vulnerability. In this study, DRASTIC was used to predict locations most vulnerability to contamination by human activities. This study assumed nitrate concentrations greater than 2 ppm were caused by human activities and over 2 ppm of nitrate concentrations should typically be detected in "High" and "Very high" vulnerability classes because the purpose of this study was to identify intrinsic aquifer vulnerability. Thus, the greater the proportion of nitrate detections > 2 ppm in "High" and "Very high" vulnerability areas, the better the prediction of intrinsic aquifer vulnerability. If "High" and "Very high" vulnerability areas as a percentage are larger than number of nitrate detections > 2 ppm as a percentage, the model performance should be regarded as poor, which would be overestimated by DRASTIC. Thus, the concept of detection ratio (percent of nitrate detections > 2 ppm to percent of "Very high" and "High" vulnerability areas with larger detection ratio indicating better prediction) was used to evaluate model performance in this study.

Table 2.6 Calibrated DRASTIC weights using Bi-GA for better prediction of intrinsic aquifer vulnerability

Calibration Methods	**D**	**R**	**A**	**S**	**T**	**I**	**C**	**RMSE**
No calibration	5	4	3	2	1	5	3	0.70
Bi-GA[1]	5.7	4.3	3	1.6	0.7	5.4	2.8	0.57

[1]Bi-GA: Binary classifier calibration with genetic algorithm

Table 2.7. Error matrix for uncalibrated DRASTIC

		Uncalibrated DRASTIC		
		Success	Failure	Totals
Obs N[1]	Success	41	75	116
	Failure	75	41	116
	Totals	116	116	82

[1]Observed nitrate concentrations in wells

Table 2.8. Error matrix for calibrated DRASTIC

		Calibrated DRASTIC		
		Success	Failure	Totals
Obs N[1]	Success	49	75	116
	Failure	75	49	116
	Totals	116	116	98

[1]Observed nitrate concentrations in wells

2.4.2 Intrinsic Aquifer Vulnerability Mapping

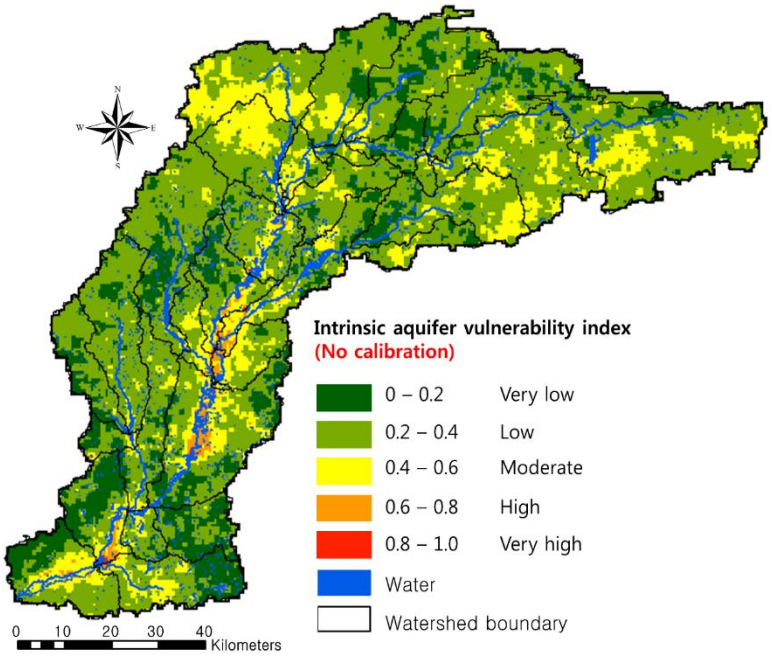
Intrinsic aquifer vulnerability maps were created without calibration of DRASTIC weights and calibrated weights using Bi-GA. An intrinsic aquifer vulnerability map without calibrating DRASTIC weights was created using DRASTIC (Figure 2.4(a)). Intrinsic aquifer vulnerability indices were classified into five classes: 0 - 0.2 (“Very low”), 0.2 - 0.4 (“Low”), 0.4 - 0.6 (“Moderate”), 0.6 - 0.8 (“High”), and 0.8 - 1.0 (“Very high”). As shown in Figure 2.4(a) and Table 2.7, 10.6% of the aquifer systems in the UWRW were within in “Very low” vulnerability class, and 60.4% of the area was estimated as “Low”, 25.8% within “Moderate” vulnerability class, 3.0% within “High” vulnerability class, and 0.2% within “Very high” vulnerability class.

The intrinsic aquifer vulnerability results (Tables 2.9) without calibration of DRASTIC weights were validated with the observed nitrate concentrations in wells. The results showed that approximately 35.3% of nitrate detections > 2 ppm are within "High" and "Very high" vulnerability areas (represent 3.2% of vulnerability area) as simulated by DRASTIC. Moreover, 60.3% of the nitrate detections were within the "Moderate" vulnerability class (25.8% of area), 3.4% of the nitrate detections were within the "Low" vulnerability class (60.4% of area), and 0.9% of the nitrate detections were within the "Low" vulnerability class (10.6% of area) (Table 2.9).

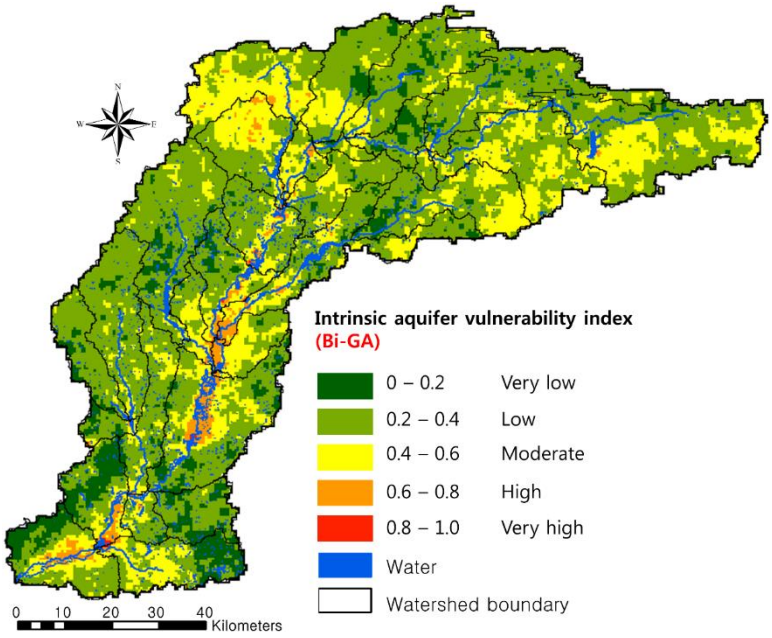
An intrinsic aquifer vulnerability map with calibrated DRASTIC using Bi-GA was produced (Figure 2.4(b)). As shown in Figure 2.4(b) and Table 2.10, 9.6% of the aquifer systems in the UWRW was within the "Very low" vulnerability class, and 60.1% of the area was estimated as "Low", 26.9% within the "Moderate" vulnerability class, 3.2% within the "High" vulnerability class, and 0.2% within the "Very high" vulnerability class.

The intrinsic aquifer vulnerability results (Table 2.10) from calibrated DRASTIC using Bi-GA were validated with the well database. The results showed that approximately 42.2% of nitrate detections > 2 ppm were within "High" and "Very high" vulnerability areas (represent 3.4% of vulnerability area) as simulated by DRASTIC. Moreover, 53.4% of the nitrate detections were within the "Moderate" vulnerability class (26.9% of area), and 4.3% of the nitrate detections were within the "Low" vulnerability class (60.1% of area). In intrinsic aquifer vulnerability assessment, nitrates in wells > 2 ppm

were not detected within the “Very low” vulnerability class (9.6% of area) (Table 2.10). These results indicated that intrinsic aquifer vulnerability assessment using DRASTIC with Bi-GA better predicted nitrate detections than DRASTIC without calibration.



(a)



(b)

Figure 2.4 Comparison of intrinsic aquifer vulnerability maps for the UWRW ((a) no calibration and (b) calibration using Bi-GA)

Table 2.9 Vulnerability areas (%) and number of nitrate detections > 2 ppm without calibration

Class	Area (%)	Number of nitrate detections > 2 ppm
Very low	10.6	1 (0.9%)
Low	60.4	4 (3.4%)
Moderate	25.8	70 (60.3%)
High	3.0	34 (29.3%)
Very high	0.2	7 (6%)

Table 2.10 Vulnerability areas (%) and number of nitrate detections > 2 ppm with calibration

Class	Area (%)	Number of nitrate detections > 2 ppm
Very low	9.6	0 (0%)
Low	60.1	5 (4.3%)
Moderate	26.9	62 (53.4%)
High	3.2	42 (36.2%)
Very high	0.2	7 (6%)

GIS-based overlay and index models such as DRASTIC can be affected by data resolution and accuracy (Woodrow et al., 2016). Navulur (1996) used three models (i.e., DRASTIC, SEEPAGE, and combined DRASTIC and NLEAP (Nitrate Leaching and Economic Analysis)) to estimate aquifer vulnerability of groundwater systems in Indiana using a GIS environment at a 1:250,000 scale. The data scale used in Navulur's (1996) study was coarse (1:250,000) for field scale simulations. However, in this study, high resolution data (1:24,000) were used by data preprocessing of recharge (R), aquifer media (A), soil media (S), topography (T), and impact of vadose zone media (I) maps.

As shown in Navulur's (1996) results for all of Indiana, the result of DRASTIC shows 80.7% of nitrate detections in wells > 2 ppm are within "High" and "Very high" vulnerability areas (represent 24.8% of area) as predicted by DRASTIC. For SEEPAGE, 60.5% of nitrate detections in wells > 2 ppm are within "High" and "Very high" vulnerability areas (28.6% of area). The result of the combined DRASTIC and NLEAP indicate 91.8% of nitrate detections in wells > 2 ppm are within "High" and "Very high" vulnerability areas (56.9% of area).

Compared with Navulur's (1996) study, the results presented herein had approximately 42.2% of nitrate detections in wells > 2 ppm within "High" and "Very high" (3.4% of area) vulnerability areas as predicted by DRASTIC with high resolution data. Detection ratio (% of nitrate detections to % of vulnerability areas with larger detection ratio indicating better prediction) for "High" and "Very high" areas from Navulur (2006) results in a value of 3.3 for DRASTIC, 2.1 for SEEPAGE, and 1.6 for the combined

DRASTIC and NLEAP, respectively. In contrast to the three models from Navulur's (1996) results, the results presented herein provide a value of 12.4. Thus, the results of detection ratio indicate that DRASTIC with high resolution data may estimate areas of "High" and "Very high" vulnerability classes more accurately than the models with coarse resolution data (Table 2.11).

Table 2.11 Comparison of detection ratio between previous and current study

	Navulur (1996)			
	DRASTIC	**SEEPAGE**	**Combined DL[1]**	**DRASTIC[2]**
HV-Area[3] (%)	24.8	28.6	56.9	3.4
N-Detections[4] (%)	80.7	60.5	91.8	42.2
D-Ratio[5]	3.3	2.1	1.6	12.4

[1]Combined DRASTIC and NLEAP

[2]Results from this study

[3]"High" and "Very high" vulnerability areas

[4]Number of nitrate detections

[5]Detection ratio

2.4.3 Potential Groundwater Monitoring and Management Sites

The Gi* statistic method was used to determine potential groundwater monitoring and management sites. Three ranges of z-scores (1.65-1.96, 1.96-2.58, and more than 2.58) indicate potential groundwater monitoring and management sites (hotspots). Hotspots were predicted based on the z-score with statistical significance using the Gi* statistic method. The Gi* statistic method identifies statistically significant spatial clusters of high values (high vulnerability areas) and low values (low vulnerability areas). The Gi* statistic method returns a z-score and the higher the z-score, the stronger the intensity of the clustering.

In Table 2.12 and Figure 2.5, z-scores of hotspot analysis maps to identify potential groundwater monitoring sites were estimated using calibrated DRASTIC by Bi-GA. Higher z-scores and red color (potential vulnerability areas) in the maps (Table 2.10 and Figure 2.5) indicate hotspots which suggest priority areas for groundwater monitoring and management. The portion of the study area with a z-score ≥ 1.65 for Bi-GA is 19.9% (percentage of study area, 6,944 km^2), suggesting areas where groundwater monitoring and BMPs for groundwater quality might be considered. In Figure 2.5, hotspot areas (z-score ≥ 1.65) were located along the stream and river because those areas include highly permeable alluvium, sand, and gravel. Further, depth to water is shallow. These areas would be priorities for groundwater protection.

Table 2.12 Results of hotspot analysis using Gi* statistic method

Calibration Methods	**Potential groundwater monitoring and management sites (%)**		
	Z-scores		
	1.65 - 1.96	**1.96 - 2.58**	**> 2.58**
Bi-GA	3.4	5.6	10.9

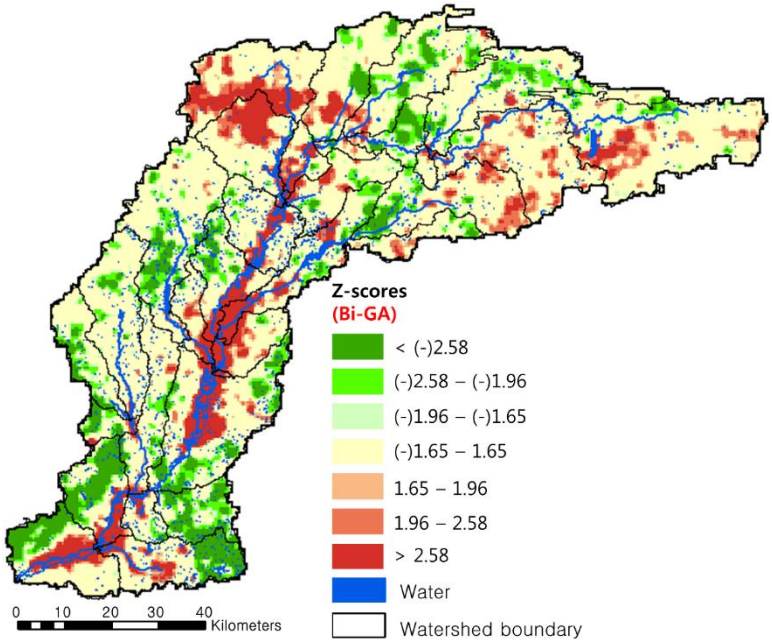


Figure 2.5 Hotspot analysis maps to determine the potential groundwater monitoring and management sites

2.5 Conclusions

Intrinsic aquifer vulnerability assessment was conducted with improved high resolution data and optimized DRASTIC parameters by modifying DRASTIC weights using Bi-GA. Simulated results to explore the most vulnerable aquifer areas estimated by DRASTIC methods were compared with long-term Heidelberg University and USGS groundwater quality data (nitrate concentrations in well) (1949 - 2010) in the UWRW, Indiana. Intrinsic aquifer vulnerability indices by improved DRASTIC were compared with observed groundwater quality data to explore how well simulated results predict observed nitrate data > 2 ppm. RMSE without calibration was 0.70, and RMSE with calibrating DRASTIC weights with Bi-GA was 0.57.

An accuracy assessment error matrix was computed for spatial validation of the calibrated DRASTIC by using Bi-GA. Total accuracies of uncalibrated DRASTIC and calibrated DRASTIC were 35% and 42%, respectively. Thus, the results of accuracy assessment indicate calibrated DRASTIC by using Bi-GA predicted intrinsic aquifer vulnerability areas more accurately than DRASTIC without calibration.

The intrinsic aquifer vulnerability results from DRASTIC using Bi-GA were validated with a well database. The results showed that approximately 42.2% of nitrate detections > 2 ppm are within "High" and "Very high" vulnerability areas (represent 3.4% of vulnerability area) as simulated by DRASTIC. Moreover, 53.4% of the nitrate detections were within the "Moderate" vulnerability class (26.9% of area), and 4.3% of the nitrate detections were within the "Low" vulnerability class (60.1% of area). In intrinsic aquifer vulnerability assessment, nitrates in wells > 2 ppm were not detected within the "Very low" vulnerability class (9.6% of area). Intrinsic aquifer vulnerability assessment using calibration with Bi-GA better predicted nitrate detections than DRASTIC without calibration.

The selection of potential monitoring locations and areas where groundwater protection should be focused was determined based on the Gi* statistic method. A portion of z-score over 1.65 by Bi-GA is 19.9% (represents percentage area of total study area, 6,944 km^2), indicating these are areas where groundwater monitoring and BMPs for groundwater quality protection should be focused. Hotspot areas (z-score ≥ 1.65) were

seen along the stream and river because those areas include high permeability of alluvium, sand, and gravel. Further, depth to water is very low. These areas would be priority for groundwater protection.

The results of this study are expected to be an efficient guide for managing groundwater resources for policy makers, natural resources protection practitioners, and groundwater-related researchers. It also could be used as a screening tool prior to applying complex numerical groundwater models for more detailed analysis. Moreover, it is expected that better parameterization of DRASTIC input data related to aquifer systems will improve aquifer vulnerability assessment and be applicable to other locations in the Midwestern, United States.

For integrated aquifer vulnerability assessment, pollutant transport should be considered. However, DRASTIC does not consider pollutant transport properties such as nitrate leaching but only focuses on intrinsic aquifer vulnerability. Thus, an additional model which can estimate pollutant transport should be combined for overall aquifer vulnerability assessment.

2.6 References

Ahlfeld, D.P., Barlow, P.M., Mulligan, A.E., 2005. GWM – A groundwater management process for the US Geological Survey modular ground-water model (MODFLOW-2000). US Geological Survey Open-File Report 2005-1072.

Akhavan, S., Mousavi, S., Abedi-Koupai, J., Abbaspour, K.C., 2011. Conditioning DRASTIC model to simulate nitrate pollution case study: Hamadan-Bahar plain. Environ. Earth Sci. 63, 1155-1167.

Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., 1985. DRASTIC—a standardized system for evaluating ground water pollution potential using hydrogeologic settings: U.S. Environmental Protection Agency, Robert S. Kerr Environmental Research Laboratory, Office of Research and Development, EPA/600/2–85/018, 163, 38-57.

Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., Hackett, G., 1987. DRASTIC: A standardized system for evaluating groundwater potential using hydrogeologic settings. USEPA Rep.600/287/035. Ada, Oklahoma.

Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and storage in groundwater systems. Science 296, 1985-1991.

Al-Zabet, T., 2002. Evaluation of aquifer vulnerability to contamination potential using the DRASTIC method. Environmental Geology 43, 203-208.

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

Atia, D.M., Fahmya, F.H., Ahmeda, N.M., Dorrah, H.T., 2012. Optimal sizing of a solar water heating system based on a genetic algorithm for an aquaculture system. Mathematical and Computer Modelling 55, 1436-1449.

Baalousha, H., 2010. Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. Agricultural Water Management 97, 240-246.

Babiker, I.S., Mohamed, M.A.A., Hiyama, T., Kato, K., 2005. A GIS based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. Science of the Total Environment 345, 127–140.

Barbash, J.E., Resek, E.A., 1996. Pesticides in ground water: distribution, trends, and governing factors. Chelsea, M17 Ann Arbor Press.

Bicknell, B.R., Imhoff, J.C., Kittle Jr., J.L., Jobes, T.H., Donigian Jr., A.S., 2001. Hydrological Simulation Program-Fortran, HSPF Version 12 User's Manual. AQUA TERRA Consultants, Mountain View, California.

Chadalavada, S., Datta, B., 2008. Dynamic Optimal Monitoring Network Design for Transient Transport of Pollutants in Groundwater Aquifers. Water Resour. Manage. 22, 651-670.

Chen, S., Jang, C., Peng, Y., 2013. Developing a probability-based model of aquifer vulnerability in an agricultural region. Journal of Hydrology 486, 494-504.

Chen, S.T., Yu, P.S., 2007. Real-time probabilistic forecasting of flood stages. Journal of Hydrology 340, 63-77.

Conservation Foundation, 1985. Groundwater - Saving the unseen resource: Proposed conclusions and recommendations, CF, Washington, DC.

Daly, D., Drew, D., 1999. Irish methodologies for karst aquifer protection. In: B. Beek, ed. Hydrogeology and engineering geology of sinkholes and karst. Rotterdam: Balkema 267-272.

EPA, 2009. National Primary Drinking Water Regulations, U.S. Environmental Protection Agency Report 816-F-09-004.

Fienen, M.N., Hunt, R.J., Doherty, J.E., Reeves, H.W., 2011. Using Models for the Optimization of Hydrologic Monitoring, U.S. Geological Survey Scientific Investigations Report 2011-3014.

Fleming, A. H., Brown, S. E., Ferguson, V. R., 1993. The hydrogeologic framework of Marion County, Indiana at Atlas illustrating hydrogeologic terrain and sequence, Indiana Geological Survey Open File Report 93-5, 67.

Foster, S.S., 1987. Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy. In: W. van Duijvenbooden and H.G. van Waegeningh, eds. Vulnerability of soil and groundwater to pollutants. Proceedings and information. The Hague: TNO Committee on Hydrological Research 38, 69-86.

Getis, A., Ord, J.K., 1992. The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24(3), 189-206.

Goldberg, D. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, Massachusetts, USA.

Hamblin, W.K., Christiansen, E.H., 2004. Earth's dynamic systems (10th ed.), Prentice Hall.

Harbaugh, A.W., 2005. MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model-the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16, Variously Paginated.

Harter, T., Walker, L.G., 2001. Assessing vulnerability of groundwater. California Department of Health Services, Report 1-11.

Holden, L.R., Graham, J.A., Whitmore, R.W., Alexander, W.J., Pratt, R.W., Liddle, S.F., Piper, L.L., 1992. Results of the national alachlor well water survey. Environmental Science and Technology 26(5), 936-943.

Holland, J. 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, Michigan, USA.

IDNR (Indiana Department of Natural Resources), 2012. Aquifer Systems Mapping (1:48,000). Retrieved from http://www.in.gov/dnr/water/4302.htm.

Javadi, S., Kavehkar, N., Mousavizadeh, M.H., Mohammadi, K., 2011. Modification of DRASTIC model to map groundwater vulnerability to pollution using nitrate measurements in agricultural areas. J. Agr. Sci. Tech. 13, 239-249.

Kalinski, R.J., Kelly, W.E., Bogardi, I., Ehrman, R.L., Yamamoto, P.O., 1994. Correlation between DRASTIC vulnerabilities and incidents of VOC contamination of municipal wells in Nebraska. Ground Water 32(1), 31–34.

Knisel, W.G., Davis, F.M., 1999. GLEAMS: Groundwater Loading Effects of Agricultural Management Systems. User Manual Version 3.0, Publication No. SEWRL-WGK/FMD-050199.

Liu, S., Tai, H., Ding, Q., Li, D., Xu, L., Wei, Y., 2013. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. Mathematical and Computer Modelling 58, 458-465.

Maas, R.P., Kucken, D.J., Patch, S.C., Peek, B.T., Van Engelen, D.L., 1995. Pesticides in eastern North Caroline rural supply wells: landuse factors and persistence. Journal of Environmental Quality 24, 426-431.

Markstrom, S.L., Niswonger, R.G., Regan, R.S., Prudic, D.E., Barlow, P.M., 2008. GSFLOW-Coupled Ground-Water and Surface-Water FLOW Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005). U.S Geological Survey, Techniques and Methods 6-D1.

McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water flow model. US Geological Survey Techniques of Water Resources Investigations Report Book 6, Chapter A1.

McLay, C.D.A., Dragden, R., Sparling, G., Selvarajah, N., 2001. Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches. Environ. Pollut. 115, 191-204.

Mitchell, A., 2005. The ESRI Guide to GIS Analysis, Volume 2. ESRI Press.

Naeini, M.P., Cooper, G.F., Hauskrecht, M., 2015. Binary Classifier Calibration Using a Bayesian Non-Parametric Approach. Proceedings of the SIAM International Conference on Data Mining. SIAM International Conference on Data Mining, 208–216.

Navulur, K.C.S., 1996. Groundwater vulnerability evaluation to nitrate pollution on a regional scale using GIS. PhD dissertation, Purdue University. West Lafayette: ProQuest/UMI.

Nobre, R.C.M., Rotunno Filho, O.C., Mansur, W.J., Nobre, M.M.M., Cosenza, C.A.N., 2007. Groundwater vulnerability and risk mapping using GIS, modeling and a fuzzy logic tool. Journal of Contaminant Hydrology 94, 277-292.

Novotny, V., 2003. Water quality (2nd Ed.) - Diffuse pollution and watershed management. Wiley.

Ord, J.K., Getis, A., 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 27(4), 286-306.

Poiani, K.A., Bedford, B.L., Merrill, M.D., 1996. A GIS-based index for relating landscape characteristics to potential nitrogen leaching to wetlands. Landscape Ecology 11(4), 237-255.

Rahman, A., 2008. A GIS based DRASTIC model for assessing groundwater vulnerability in shallow aquifer in Aligarh, India. Applied Geography 28, 32-53.

Solly, W.B., Pierce, R.R., Perlman H.A., 1998. Estimated use of water in the United States in 1995. U.S. Geological Survey Circular 1200.

Song, K., Lu, D., Li, L., Li, S., Wang, Z., Du, J., 2012. Remote sensing of chlorophyll-α concentration for drinking water source using genetic algorithms (GA)-partial least square (PLS) modeling. Ecological Informatics 10, 25-36.

Strauch, M., Bernhofer, C., Koidec, S., Volk, M., Lorz, C., Makeschina, F., 2012. Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. Journal of Hydrology 414-415, 413-424.

Tedesco, L.P., Hoffmann, J., Bihl, L., Hall, B.E., Barr, R.C., Stouder, M., 2011. Upper White River Watershed Regional Watershed Assessment and Planning Report.

Tesoriero, A.J., Inkpen, E.L., Voss, F.D., 1998. Assessing ground-water vulnerability using logistic regression. Proceedings for the Source Water Assessment and Protection 98 Conference, Dallas, TX, 157-165.

van Stemproot, D., Evert, L., Wassenaar, L. 1993. Aquifer vulnerability index: a GIS compatible method for groundwater vulnerability mapping. Canadian Water Resources Journal 18(1), 25–37.

Vías, J.M., Andreo, B., Perles, M.J., Carrasco, F., Vadillo, I., Jiménez, P., 2006. Proposed method for groundwater vulnerability mapping in carbonate (karstic) aquifers: the COP method. Hydrogeology Journal 14, 912-925.

Wang, Q.J., 1991. The genetic algorithm and its application to calibrating conceptual rainfall–runoff models. Water Resources Research 27(9), 2467-2471.

Woodrow, K., Lindsay, J.B., Berg, A.A., 2016. Evaluating DEM conditioning techniques, elevation source data, and grid resolution for field-scale hydrological parameter extraction. Journal of Hydrology 540, 1022–1029.

Wu, Y., 2004. Optimal design of a groundwater monitoring network in Daqing, China. Environmental Geology 45, 527-535.

Wu, Y., Liu, S., 2012. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. Environmental Modeling & Software 31, 99-109.

Xu, P., Davoine, F., Zha, H., Denœux, T., 2016. Evidential calibration of binary SVM classifiers. International Journal of Approximate Reasoning 72, 55–70.

Yang, Y. S., Wang, L., 2010. Catchment-scale vulnerability assessment of groundwater pollution from diffuse sources using the DRASTIC method: a case study. Hydrological Sciences Journal 55(7), 1206-1216.

Zhang, R., Hamerlinck, J.D., Gloss, S.P., Munn, L., 1996. Determination of nonpoint-source pollution using GIS and numerical models. Journal of Environmental Quality 25, 411-418.

Zhang, X., Srinivasan, R., Bosch, D., 2009. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. Journal of Hydrology 374, 307-317.

Zomorodi, K. 2004. Curve Number and Groundwater Recharge Credits for LID Facilities in New Jersey. Dewberry, 8401 Arlington Blvd., Fairfax, VA 22031.

CHAPTER 3. EFFICIENT FLOW CALIBRATION REGIME FOR ACCURATE ESTIMATION OF HYDROLOGIC AND WATER QUALITY COMPONENTS USING A WATERSHED SCALE HYDROLOGICAL MODEL

3.1 Abstract

Accurate analysis of water flow (streamflow and baseflow) is important to estimate nitrate contamination in aquifers. For accurate estimation of aquifer hazard, modeled surface and groundwater hydrology should be calibrated first, and then simulation of nitrate leaching to aquifers should be conducted because nitrate leaching to aquifers is interdependent with baseflow or vice versa. The objectives of this study are: 1) to develop an efficient flow calibration regime (EFCR) for accurate baseflow estimation by combining the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) and modified SWAT 2012 code, and 2) to evaluate the performance of the EFCR by streamflow and baseflow estimation. Both streamflow and baseflow estimated using the EFCR performed well based on three model evaluation methods (i.e., NSE, R^2, and PBIAS). Thus, the EFCR would be a practical method for aquifer hazard assessment by calibrating baseflow accurately as well as streamflow at a watershed scale. This study can be used as a data-driven model for in-depth groundwater modeling because the baseflow-related parameters (i.e., groundwater recharge and hydraulic conductivity) calibrated in this

study can be used as a set of input data (initial parameter values) in computer-based numerical groundwater models.

3.2 Introduction

Groundwater is the primary component of water resources for people worldwide (Alley et al., 2002). Groundwater provides humans water for drinking and irrigation as well as being a fundamental contributor to ecological productivity. However, groundwater has been increasingly contaminated resulting from human activities. Nitrate leaching has resulted in risk of groundwater contamination (Dahan et al., 2014; Babiker at al., 2004). Nitrate leaching is mainly generated from agricultural fields to groundwater in agricultural areas as well as urban areas and nitrate leaching is associated with water flow from precipitation to stream and aquifer (Thorosen, 2001). Accurate analysis of water flow (streamflow and baseflow) is important to estimate nitrate leaching to aquifers (Molenat, 2002). Sustainable groundwater management is required for sound ecosystems and quality of human life. For sustainable groundwater management, aquifer hazard assessment has been conducted by groundwater modeling or monitoring. Compared with groundwater monitoring, groundwater modeling is less complicated and costly and can be used to evaluate broad areas. There are many groundwater models, but most are complicated and need many input data.

The Soil and Water Assessment Tool (SWAT) (Arnold et al, 1998) has been widely used for water resources management in part because SWAT has a user friendly Graphical User Interface (GUI) and a well-organized database by worldwide users. SWAT

simulates surface flow and shallow groundwater dynamics based on hydrological response units (HRUs), which are the smallest computational units in SWAT. Surface runoff, shallow groundwater dynamics, soil water content, nutrient cycles, and sediment erosion are simulated for each HRU and then HRUs are combined and calculated for the subbasin by a weighted value (Neitsch et al., 2011). Thus, SWAT has been used to predict the risk of groundwater contamination (aquifer hazard) from non-point sources.

For accurate estimation of aquifer hazard, surface and groundwater hydrology should be calibrated first, and then simulation of nitrate leaching to aquifers should be conducted (Santhi et al., 2001; Arnold et al., 2012) because nitrate leaching to aquifers is interdependent with baseflow or vice versa. However, many studies have only considered streamflow (direct runoff + baseflow) calibration without conducting direct runoff and baseflow separately (Vilaysane et al., 2015; Zhang et al., 2015). When SWAT users only calibrate streamflow, even though streamflow meets statistical criteria, baseflow may not meet statistical criteria for the calibration process (Jang et al., 2011; Zhang et al., 2011). The results showed that direct runoff or baseflow would be underestimated or overestimated. However, when streamflow and baseflow calibrations are conducted simultaneously, both variables meet statistical criteria for the calibration process (Feyereisen et al., 2007). Thus, streamflow and baseflow calibrations should be conducted simultaneously for accurate estimation of aquifer hazard. Jang et al. (2010) studied evaluation of SWAT direct runoff and baseflow components using web-based K-means clustering EI estimation system. They reported that even though the Nash-Sutcliffe Efficiency (NSE) of total streamflow was high, the NSE values of hydrologic

components (i.e., direct runoff and baseflow) were not high. Zhang et al. (2011) studied simultaneous calibration of surface flow and baseflow simulations. The study showed that the uncertainty of low flows on baseflow were overestimated while uncertainty of high flows on surface flow were underestimated.

In hydrologic and water quality modeling, calibration and uncertainty analysis is required to produce reliable prediction by determining the appropriate values of parameters (Abbaspour, 2011). There are two methods of calibration: manual calibration and autocalibration. Many modelers have increasingly used autocalibration instead of manual calibration (trial and error calibration) because the autocalibration has several advantages over the manual calibration (Shi et al., 2013; Rathjens and Oppelt, 2012; Shrestha et al., 2016). The autocalibration can produce parameter estimation by uncertainty analysis in the modeling as well as provide minimal labor on the part of the user (Arnold et al., 2012). Moreover, autocalibration can minimize the difference between observed and simulated values (Van Liew at al., 2005). SWAT Calibration and Uncertainty Program (SWAT-CUP) (Abbaspour, 2011) is one of the most popular autocalibration programs, and SWAT users use the program to perform calibration, validation, sensitivity, uncertainty analysis of SWAT. Many SWAT modelers use SWAT-CUP because it is well organized and easy to use. The program includes five different calibration procedures such as SUFI-2, Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Bindly, 1992), Parameter Solution (ParaSol) (Van Griesven et al., 2006), Markov chain Monte Carlo (MCMC) (Kuczera and Parent, 1998; Vrugt et al., 2003), and

Particle Swarm Optimization (PSO) (Eberhart and Kennedy, 1995) for the SWAT calibration.

For those reasons, an efficient flow calibration regime for accurate aquifer hazard assessment was developed in this study. The efficient flow calibration regime (EFCR) developed in this study allows SWAT to simultaneously calibrate streamflow and baseflow by combining the autocalibration algorithm, SUFI-2 (Abbaspour et al., 2004) and modified SWAT 2012 code. Compared with other calibration regimes, the calibration regime presented in this study is easy to use because SUFI-2 and modified SWAT 2012 code can be used in SWAT-CUP. With the EFCR, calibration parameter space and ranges for Latin Hypercube Sampling (LHS) can be reduced by adjusting and resorting parameter range based on a previous calibration step. This process would improve calibration accuracy and reduce simulation time. The objectives of this study are: 1) to develop the EFCR for accurate baseflow estimation with SUFI-2 and modified SWAT 2012 code, and 2) to evaluate the performance of the EFCR by streamflow and baseflow estimation.

3.3 Materials and Methods

3.3.1 Study Area

The study area (Figure 3.1) is the Upper White River Watershed (UWRW) (Latitude: 39°29'51"N, Longitude: 86°24'02"W) in Indiana. The UWRW is a Hydrologic Unit Code (HUC) 8 watershed (05120201) located in central Indiana and includes seventeen HUC 10 subwatersheds. Drainage area of the study watershed is 6,944 km^2, and the most

dominant land use is agriculture (3,160 km^2). UWRW is important for public drinking water supplies because UWRW includes more than 3,508 km of streams, numerous artificial lakes, and 4 reservoirs. Sixteen counties are located in the watershed, and the UWRW serves as the drinking water supply for the city of Indianapolis which is Indiana's largest city. The water sources in the UWRW traditionally are individual wells to provide groundwater for residential, commercial, and industrial purposes (Tedesco et al., 2011; Fleming et al., 1993). The UWRW was selected to identify the EFCR for aquifer hazard assessment because this watershed has available streamflow (10 USGS streamflow stations) and water quality (5 EPA fixed stations) data (Figure 3.1 and Table 3.1).

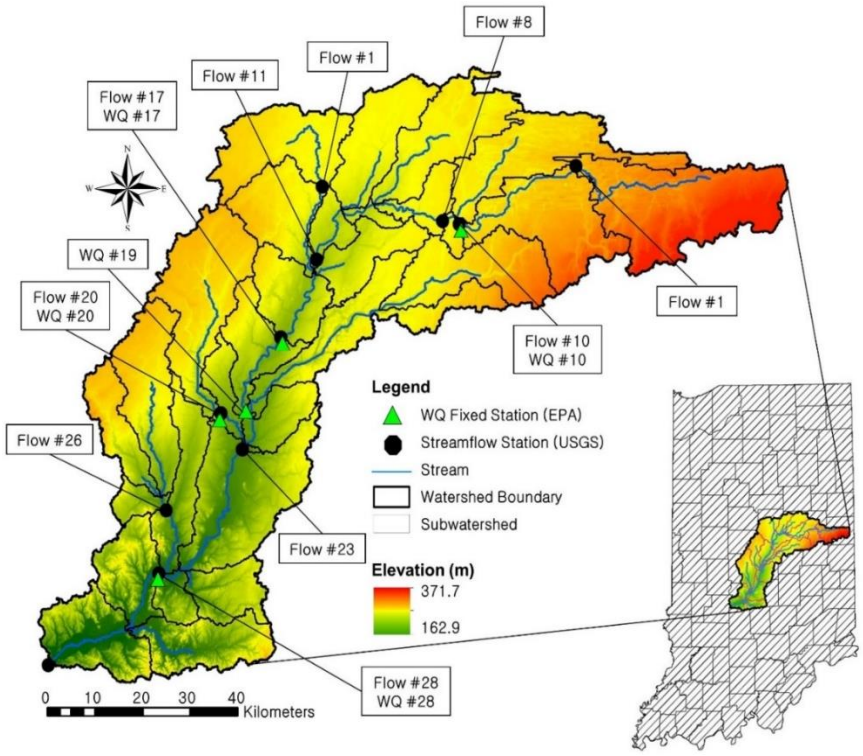


Figure 3.1 Location of the Upper White River Watershed, Indiana

Table 3.1 Monitoring stations for streamflow and water quality data in the UWRW

ID	Subbasin #	Monitoring station	Drainage area (km^2)	Type
Flow[1] #1	1	USGS 03349510	329	Streamflow
Flow #2	2	USGS 03347000	595	Streamflow
Flow #8	8	USGS 03348130	1411	Streamflow
Flow #10	10	USGS 03348000	997	Streamflow
Flow #11	11	USGS 03349000	2237	Streamflow
Flow #17	17	USGS 03351000	3201	Streamflow
Flow #20	20	USGS 03353500	448	Streamflow
Flow #23	23	USGS 03353611	4756	Streamflow
Flow #26	26	USGS 03353800	544	Streamflow
Flow #28	28	USGS 03354000	6227	Streamflow
WQ[2] #10	10	INSTOR WQX[3]-2398	997	Nitrate
WQ #17	17	INSTOR WQX-2434	3201	Nitrate
WQ #19	19	INSTOR WQX-2408	690	Nitrate
WQ #20	20	INSTOR WQX-2371	448	Nitrate
WQ #28	28	USGS 03354000	6227	Nitrate

[1]Flow: Streamflow data

[2]WQ: Water quality data

[3]INSTOR WQX: Indiana STORET (STOrage and RETrieval) data warehouse by EPA

3.3.2 Baseflow Separation using the WHAT System

Baseflow separation techniques have been used to separate direct runoff and baseflow from streamflow because it is difficult to measure baseflow in contrast with the measurement of streamflow. Among various baseflow separation techniques (Sloto and Crouse, 1996; Rutledge, 1998; Arnold and Allen, 1999), the Web GIS-based Hydrograph Analysis Tool (WHAT) (https://engineering.purdue.edu/~what/) (Lim et al., 2005) was used to perform baseflow separation from USGS streamflow (USGS 03354000 White River near Centerton, IN). A user friendly and fully automated WHAT system uses the maximum value of the baseflow index (BFImax) and filter parameter values proposed by Echhardt (2005). The WHAT system includes a genetic algorithm (GA) to determine

optimum BFImax (Eckhardt, 2005) and filter parameter values to reflect local hydrological and hydrogeological situations for accurate baseflow separation (Lim et al., 2010). A BFImax value of 0.80 and filter parameter value of 0.98 based on Lim et al. (2010) and Eckhardt (2005) were used because streams in UWRW are perennial streams with porous aquifers. So, a BFImax value of 0.80 and filter parameter value of 0.98 were used to separate baseflow from USGS streamflow. In the WHAT system, baseflow separation from streamflow is conducted with Equation 3.1 as shown below:

$$b_t = \frac{(1-\mathrm{BFI}_{\max})\alpha + b_{t-1}(1-\alpha)\mathrm{BFI}_{\max} Q_t}{1-\alpha \mathrm{BFI}_{\max}} \quad (3.1)$$

where b_t is the filtered baseflow at the *t* time step, b_{t-1} is the filtered baseflow at the *t*-1 time step, $\mathrm{BFI}_{\max}$ is the maximum value of long-term ratio of baseflow to total streamflow, α is the filter parameter, and Q_t is the total streamflow at *t* time step.

3.3.3 Hydrologic and Water Quality Modeling using SWAT

3.3.3.1 Overview of SWAT

SWAT is a physically based distributed, deterministic, and long-term continuous time model which is used to predict impact of management practices, LULC change, and climate change on hydrology and water quality on a watershed scale with a daily time step (Arnold et al., 1998; Neitsch et al., 2011). SWAT represents the large scale spatial heterogeneity of the study area by dividing a watershed into subbasins. The subbasin is the first level of subdivision of the watershed. Subbasins possess a geographic position in the watershed and are spatially related to one another. The land area in a subbasin may be divided into hydrologic response units (HRUs) which are the smallest computational

units in SWAT. HRUs are parts of a subbasin that possess unique land use, management, and soil attributes. HRUs are created by one or more unique land use and soil combinations for each subbasin. Surface runoff, soil water content, crop growth, nutrient cycles, and erosion are simulated for each HRU, and then HRUs are combined and calculated for the subbasin by a weighted value (Neitsch et al., 2011; Williams et al., 1984).

3.3.3.2 SWAT Input Data

Various spatial and temporal data are required for the SWAT simulation. The period of SWAT simulation for this study is from 1990 to 2010 because the SWAT format climate data (i.e., precipitation and temperature) were available in this period provided by National Climate Data Center (NCDC) and processed by the Agricultural Research Service (ARS). Also, there are sufficient water quality data from 1990 to 2010 for the UWRW. As described in Table 3.2, the primary input data for the SWAT simulation are topography, soil, Land Use and Land Cover (LULC), and weather data. Additionally, for nitrate simulation, the scheduled management operation data were prepared to consider application of fertilizer, pesticides, and manure (Table 3.3). A general strategy of management practices for corn-soybean rotation is described by Her et al. (2016). The Digital Elevation Model (DEM) (Figure 3.1) from U.S. Geological Survey (USGS) was used for watershed delineation, and soil and LULC data from Natural Resources Conservation Service (NRCS) and Natural Agricultural Statistics Service (NASS) were used with respect to hydrology and water quality in the watershed. Weather data include minimum and maximum temperature, daily precipitation, mean wind speed, relative

humidity, and solar radiation which were obtained from National Climate Data Center (NCDC) (Table 3.2). With the input data (Table 3.2), hydrology (surface water + groundwater hydrology) and water quality were simulated in the UWRW by SWAT (Figure 3.1). Then, observed streamflow from the USGS were used to calibrate and validate SWAT.

The spatial soil map for the UWRW (Figure 3.2(a)) was obtained by USGS. Twenty eight soil types are distributed in the study area such as 33.5% IN013 (Crosby), 24% IN040 (Miami), 6.3% IN029 (Sawmill), 6% IN026 (Fox), 5.2% IN054 (Miamian) and 24.9% other soil types. LULC in the UWRW (Figure 3.2(b)) includes 45.5% agricultural field (22.5% soybeans, 22.2% corn, and 0.8% others), 23.5% urban area, 14.9% forest, 12.6% pasture, and 3.5% other LULC types. The mean annual precipitation is 1,093 mm, and the highest, the lowest daily temperature, and mean daily temperatures are 36.1°C, -31.3°C, and 10.8°C, respectively.

Tile drainage was applied in areas where the land use is corn or soybean and the soil drainage condition is poorly drained (Boles et al., 2015; Jiang et al., 2014; Sui and Frankenberger, 2008). Based on the previous studies for Indiana watersheds (Boles et al., 2015; Jiang et al., 2014; Green et al., 2006) and the technical report about tile drainage (USDA-NRCS, 2011), parameters related to the tile drainage in SWAT were used as Table 3.4.

Table 3.2 SWAT input data for hydrologic and water quality modeling

Data	Source	Format	Scale	Date
DEM	USGS[1]	Raster	1:24,000	2010
Soil	NRCS[2]	Polygon	1:12,000	2010
LULC[3]	NASS[4]	Raster	1:250,000	2010
Weather	NCDC[5]	Tabular data	-	1986 - 2010
Streamflow[6]	USGS[1]	Tabular data	-	1986 - 2010

[1]USGS: U.S. Geological Survey

[2]NRCS: Natural Resources Conservation Service

[3]LULC: Land Use and Land Cover

[4]NASS: Natural Agricultural Statistics Service

[5]NCDC: National Climate Data Center

[6]Data for model calibration and validation

Table 3.3 Management practices for corn-soybean rotation in SWAT

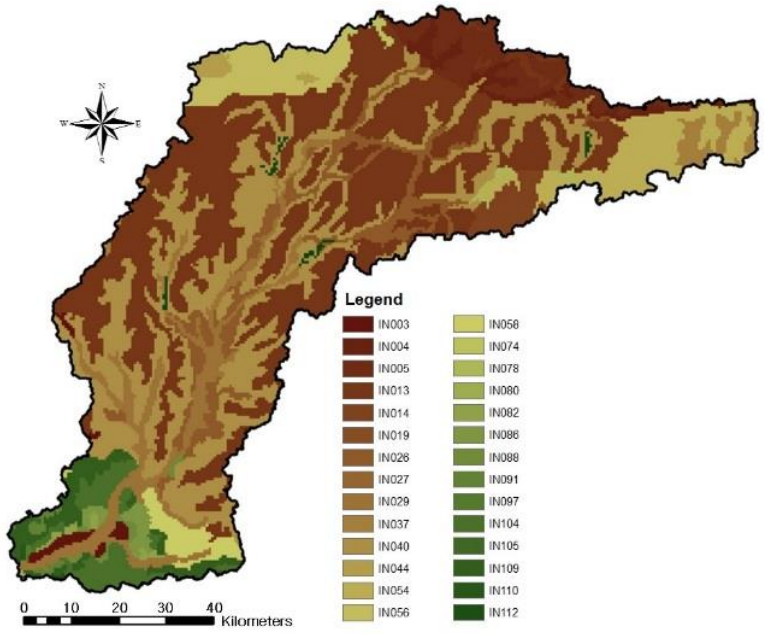
Year	Date	Management	ID
Corn Year	Apr-22	Fertilizer application	Anhydrous ammonia[1]
	Apr-22	Pesticide application	Atrazine[2]
	May-6	Tillage	Field cultivator
	May-6	Planting	Corn
	Jun-6	Fertilizer application	Urea[3]
	Oct-14	Harvest	-
	Oct-15	Kill / end of growing season	-
Soybean Year	May-24	Tillage	No-till
	May-24	Planting	Soybean
	Oct-7	Harvest	-
	Oct-8	Kill / end of growing season	-
	Oct-15	Fertilizer application	P_2O_5[4]
	Nov-1	Tillage	Chisel plow

[1]Anhydrous ammonia: 53 kg/ha (N of 43 kg/ha)

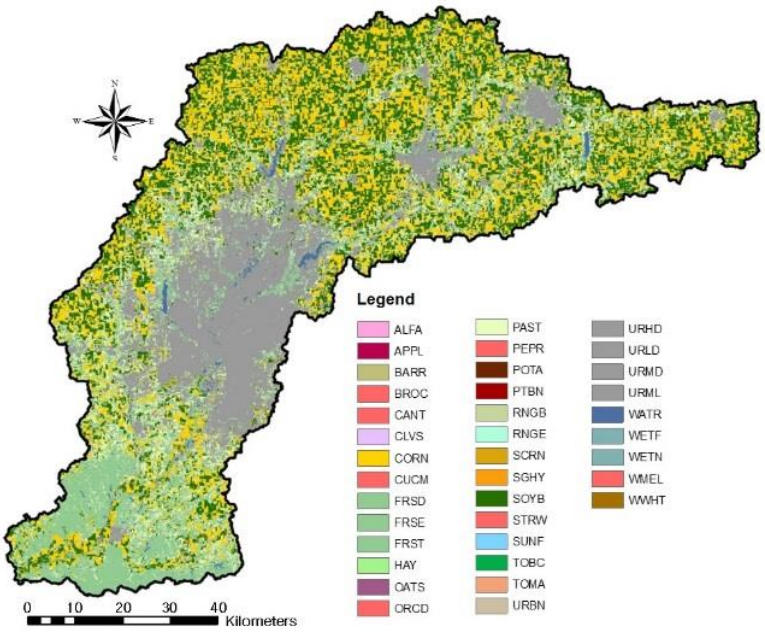
[2]Atrazine: 2.2 kg/ha

[3]Urea: 284 kg/ha (N of 131 kg/ha)

[4]P_2O_5: 123 kg/ha (P of 54 kg/ha)



(a)



(b)

Figure 3.2 SSURGO soil map (a) and NLCD land use map (b) of the UWRW

Table 3.4 Tile drainage parameters in SWAT

Parameter	Description	Value
DDRAIN	Depth to drains (mm)	1000
G_DRAIN	Drain tile lag time (h)	48
DRAIN_CO	Drainage coefficient (mm/d)	10
SDRAIN	Tile spacing (mm)	20000
LATKSATF	Multiplication factor to determine Ksat	1.2
RE	Effective radius of drains (mm)	20
ITDRN	Tile drainage equations flag/code	1 (new routine)

3.3.3.3 Hydrologic and Water Quality Modeling in SWAT

The land phase of the hydrological cycle controls the amount of water, sediment, nutrient and pesticide loadings to the main channel and to the aquifers in each subbasin. The land phase of the hydrological cycle simulated in SWAT is based on the water balance equation (Equation 3.2) (Neitsch et al., 2011).

$$SW_t = SW_0 + \sum_{i=1}^{t} (R_{day,i} - Q_{surf,i} - E_{a,i} - Q_{lat,i} - Q_{gw,i}) \quad (3.2)$$

where SW_t is the final soil water content (mm) at time *t*, *t* is the time (day), SW_0 is the initial soil water content (mm H_2O), $R_{day,i}$ is the amount of precipitation on day *i* (mm H_2O), $Q_{surf,i}$ is the amount of surface runoff on day *i* (mm H_2O), $E_{a,i}$ is the amount of evapotranspiration on day *i* (mm H_2O), $Q_{lat,i}$ is the amount of lateral flow released to the main channel on day *i* (mm H_2O), and $Q_{gw,i}$ is the amount of return flow on day *i* (mm H_2O).

Water that moves the soil profile by percolation or bypass flow enters and flows through the vadose zone. The shallow aquifer (an unconfined aquifer) receives groundwater

recharge from the vadose zone, and a fraction of the groundwater recharge to the shallow aquifer can be routed to the deep aquifer (a confined aquifer) by percolation (Equations 3.3 and 3.4) (Neitsch et al., 2011).

$$Q_{gw,sh,i} = Q_{gw,sh,i-1} \cdot \exp(-\alpha_{gw,sh} \cdot \Delta t) + w_{rchrg,sh,i} \cdot [1 - \exp(-\alpha_{gw,sh} \cdot \Delta t)] \quad (3.3)$$

where $Q_{gw,sh,i}$ is the groundwater flow from the shallow aquifer on day *i* (mm H_2O), $Q_{gw,sh,i-1}$ is the groundwater flow from the shallow aquifer on day *i*-1 (mm H_2O), $-\alpha_{gw,sh}$ is the baseflow recession constant, Δt is the time step (day), and $w_{rchrg,sh,i}$ is the amount of recharge entering the shallow aquifer on a day *i* (mm H_2O).

$$Q_{gw,dp,i} = Q_{gw,dp,i-1} \cdot \exp(-\alpha_{gw,dp} \cdot \Delta t) + w_{rchrg,dp,i} \cdot [1 - \exp(-\alpha_{gw,dp} \cdot \Delta t)] \quad (3.4)$$

where $Q_{gw,dp,i}$ is the groundwater flow from the deep aquifer on day *i* (mm H_2O), $Q_{gw,dp,i-1}$ is the groundwater flow from the deep aquifer on day *i*-1 (mm H_2O), $-\alpha_{gw,dp}$ is the baseflow recession constant, Δt is the time step (day), and $w_{rchrg,dp,i}$ is the amount of recharge entering the deep aquifer on a day *i* (mm H_2O).

SWAT considers the shallow aquifer and deep aquifer as groundwater storage. Total baseflow (groundwater flow) is calculated by total amount of water in the shallow aquifer and deep aquifer (Equation 3.5) (Neitsch et al., 2011).

$$Q_{gw,i} = Q_{gw,sh,i} + Q_{gw,dp,i} \quad (3.5)$$

where $Q_{gw,i}$ is the total groundwater flow on day *i* (mm H_2O), $Q_{gw,sh,i}$ is the total groundwater flow in the shallow aquifer on day *i* (mm H_2O), and $Q_{gw,dp,i}$ is total groundwater flow in the deep aquifer on day *i* (mm H_2O).

Lateral flow is significant in areas with soils having high hydraulic conductivities in surface layers and an impermeable layer at a shallow depth. In such a stream, rainfall percolates vertically until it encounters the impermeable layer. The water then ponds above the impermeable layer forming a saturated zone of water (i.e., perched water table). This saturated zone is the source of water for lateral flow. Lateral flow is calculated with Equation 6 (Neitsch et al., 2011).

$$Q_{lat,i} = 0.0224 \cdot \left(\frac{2 \cdot SW_{ly,excess} \cdot K_{sat} \cdot slp}{\phi_d \cdot L_{hill}} \right) \tag{3.6}$$

where $Q_{lat,i}$ is the amount of lateral flow released to the main channel on day *i* (mm H_2O), $SW_{ly,excess}$ is the drainage volume of water stored in the saturated zone of the hillslope per unit area (mm H_2O), K_{sat} is the saturated hydraulic conductivity (mm/h), slp is the average slope of the subbasin (m/m), ϕ_d is the drainage porosity of soil (mm/mm), and L_{hill} is the hillslope length (m).

Groundwater flow entering the main channel from the shallow aquifer can contain nitrate. Nitrate in the shallow aquifer may remain in the aquifer, move with recharge to the deep aquifer, move with groundwater flow into the main channel, or be transported out of the shallow aquifer with water moving in the soil zone in response to water deficiencies.

SWAT can simulate nitrate in recharge to the shallow aquifer and deep aquifer on a given day with Equations 3.7 and 3.8 (Neitsch et al., 2011).

$$NO3_{rchrg,i} = (1 - \exp[-1/\delta_{gw}]) \cdot NO3_{perc} + \exp[-1/\delta_{gw}] \cdot NO3_{rchrg,i-1} \quad (3.7)$$

where $NO3_{rchrg,i}$ is the amount of nitrate in recharge entering the aquifers on day *i* (kg N/ha), δ_{gw} is the delay time (drainage time) of the overlying geologic formation (day), $NO3_{perc}$ is the total amount of nitrate exiting the bottom of the soil profile on day *i* (kg N/ha), $NO3_{rchrg,i-1}$ is the amount of nitrate in recharge entering the aquifers on day *i*-1 (mm H_2O).

$$NO3_{gw,i} = (NO3_{sh,i-1} + NO3_{rchrg,i}) \cdot Q_{gw,i} / (aq_{sh,i} + Q_{gw,i} + w_{revap,i} + w_{rchrg,dp,i}) \quad (3.8)$$

where $NO3_{gw,i}$ is the amount of nitrate in groundwater flow from the shallow aquifer on day *i* (kg N/ha), $NO3_{sh,i-1}$ is the amount of nitrate in the shallow aquifer at the end of day *i*-1 (kg N/ha), $Q_{gw,i}$ is the groundwater flow into the main channel on day *i* (mm H_2O), $aq_{sh,i}$ is the amount of water stored in the shallow aquifer at the end of day *i* (mm H_2O), $w_{revap,i}$ is the amount of water moving into the soil zone in response to water deficiencies on day *i* (mm H_2O), $w_{rchrg,dp,i}$ is the amount of recharge entering the deep aquifer on day *i* (mm H_2O).

Many studies about aquifer hydrology and water quality (especially surface and subsurface transport of nutrients (mainly nitrogen and phosphorus) in soil and water) have been carried out with SWAT and have demonstrated the robustness of SWAT in

simulating nutrient losses at the watershed scale (Moriasi et al., 2013; Sultan et al., 2011; Cerro et al., 2011; Akhavan et al., 2010; Vale and Holman, 2009). As shown in Figure 3.3, in this study, streamflow and baseflow were simulated and calibrated based on USGS streamflow data and baseflow data generated by the WHAT system.

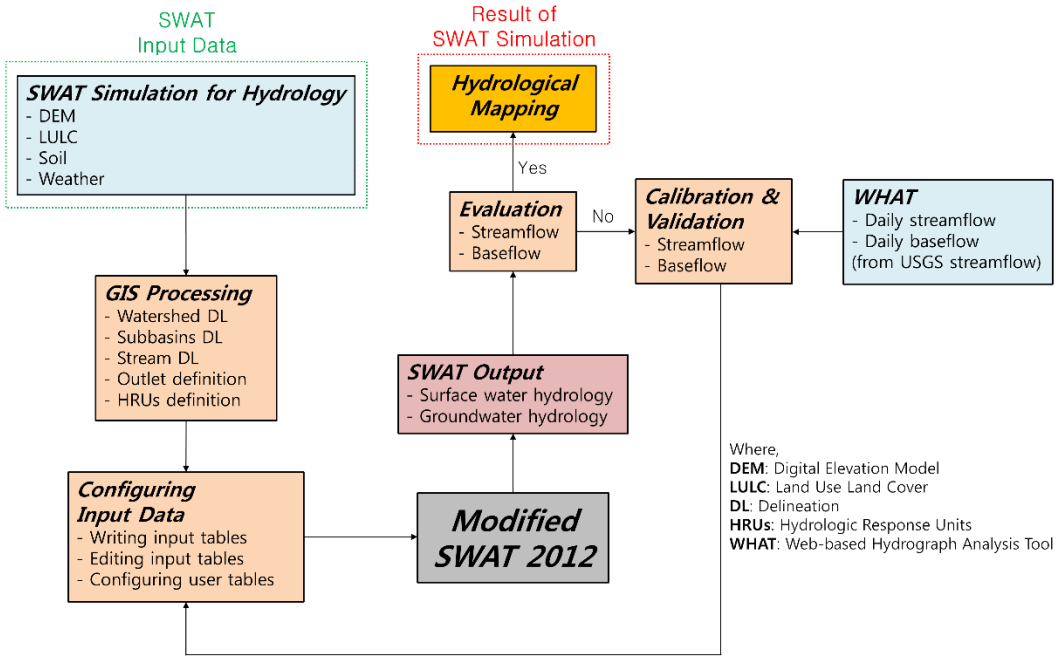


Figure 3.3 Flowchart of SWAT simulation for hydrology using simultaneous streamflow and baseflow calibration

3.3.4 Efficient Flow Calibration Regime using SUFI-2 and Modified SWAT 2012 Code

3.3.4.1 Overview of SUFI-2

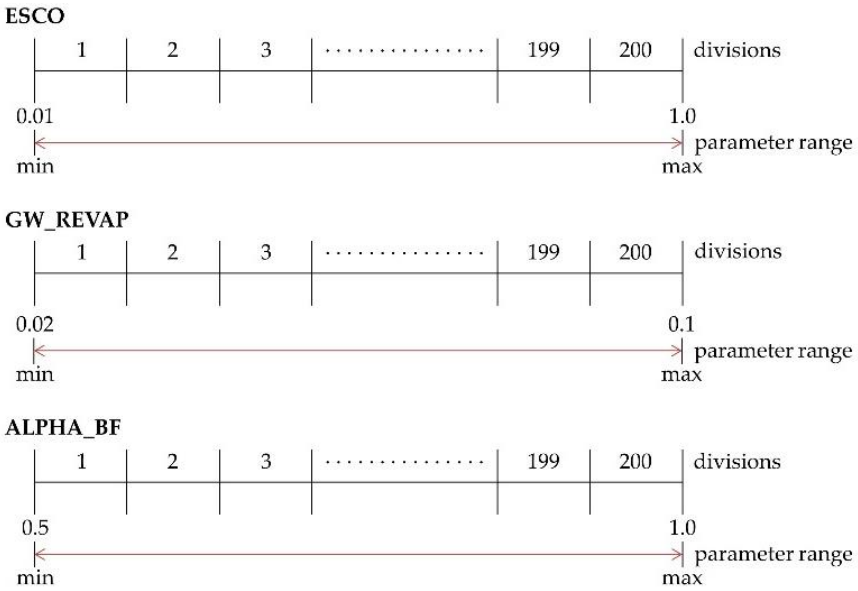
SUFI-2 represents all uncertainties such as uncertainty of input variables, conceptual model, parameters, and observed data. Uncertainties in the model output are determined by the 95% prediction uncertainty (95PPU) calculated at the 2.5% and 97.5% levels of the cumulative distribution of the output variables drawn by LHS which is a statistical

method to generate controlled random parameter sets from a multidimensional distribution (Abbaspour et al., 2007, 2011). If the number of parameters and simulations are 3 (ESCO, GW_REVAP, and ALPHA_BF) (Table 3.5) and 200, LHS is conducted as follows (modified from Abbaspour et al., 2011):

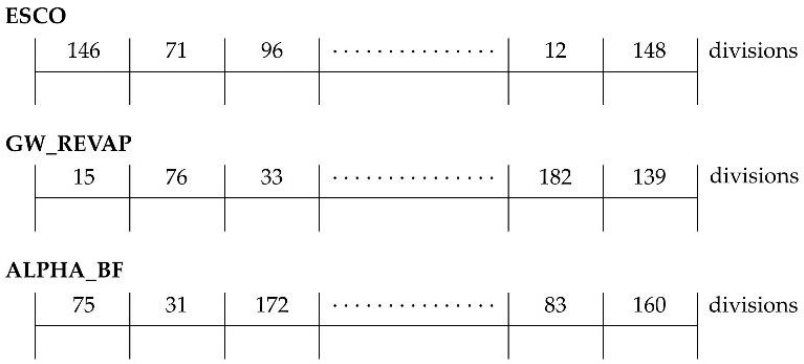
1. Three parameters are divided into the number of simulations that user defined (Figure 3.4 (a)).

2. Parameter segments are randomized (Figure 3.4 (b)).

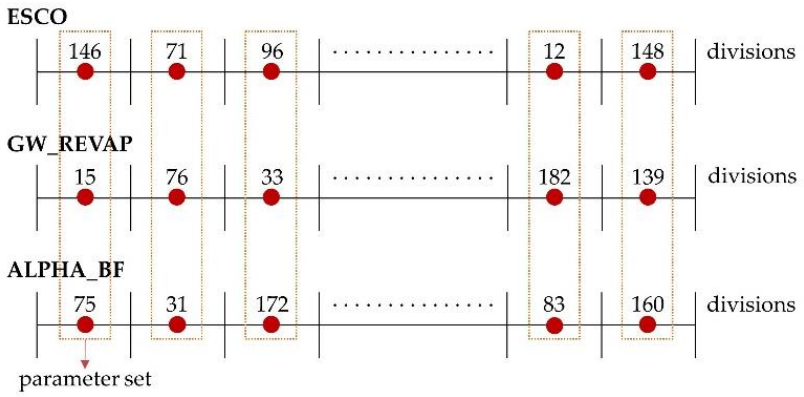
3. A sample is selected and every vertical combination is defined as a parameter set (Figure 3.4(c)).



(a)



(b)



(c)

Figure 3.4 Latin hypercube sampling procedure to generate controlled random parameter sets from a multidimensional distribution

Table 3.5 Example parameter set to describe Latin hypercube sampling

Parameter	Description	Ranges	
		LB[1]	UB[2]
ESCO	Soil evaporation compensation factor	0.01	1.00
GW_REVAP	Groundwater "revap" coefficient	0.02	0.1
ALPHA_BF	Baseflow alpha factor	0.5	1.0

[1]LB: Lower bound

[2]UB: Upper bound

In SUFI-2, two indices (P-factor and R-factor) are used to judge the goodness of calibration and validation. P-factor is the percentage of observed data bracketed by the 995PPU and ranges from 0 to 100%, where 100% represents a perfect model. R-factor is the average thickness of the 995PPU divided by the standard deviation of the observed data and varies from 0 to infinity. The closeness of R-factor to zero indicates that simulation in the model exactly corresponds to observed data. When acceptable values of P-factor and R-factor are reached, the parameter ranges are taken as the calibration parameters. For further evaluation of the model performance, Nash-Sutcliff Efficiency (NSE), coefficient of determination (R^2), and percent bias (PBIAS) were selected as the model evaluation methods for calibration and validation of streamflow and baseflow (Abbaspour et al., 2007, 2015). Simulated and observed values were compared using the three model evaluation methods, one at a time. NSE is a normalized statistic that explains the relative magnitude of the residual variance ("noise") compared to the measured data variance and it varies from minus infinity to 1. An NSE of 1 indicates a perfect match of simulated data to the observed data (Equation 3.9) (Nash and Sutcliffe, 1970). R^2 is the square of the correlation (r) between simulated and observed values and it ranges from 0 to 1. An R^2 of 1 means the simulated data are predicted perfectly without error (Equation 3.10) (Krause et al., 2005). PBIAS measures the average tendency of the simulated data to be larger or smaller than the observations. The optimal value of PBIAS is zero, where low magnitude values indicate better model simulations. Positive values indicate model underestimation and negative values indicate model overestimation (Equation 3.11) (Gupta et al., 1999).

$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs} - Y^{mean})^2} \right] \tag{3.9}$$

where Y_i^{obs} is the i^{th} observed data, Y_i^{sim} is the i^{th} simulated data, Y^{mean} is the mean of observed data, and n is the total number of observed data.

$$\mathrm{R}^2 = \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - \bar{Y}^{obs})(Y_i^{sim} - \bar{Y}^{sim})^2}{\sqrt{\sum_{i=1}^{n} (Y_i^{obs} - \bar{Y}^{obs})^2} \sqrt{\sum_{i=1}^{n} (Y_i^{sim} - \bar{Y}^{sim})^2}} \right]^2 \tag{3.10}$$

where Y_i^{obs} is the i^{th} observed data, $\bar{Y}^{obs}$ is the mean of observed data, Y_i^{sim} is the i^{th} simulated data, $\bar{Y}^{sim}$ is the mean of simulated data, and n is the total number of observed data.

$$\mathrm{PBIAS} = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim}) \times 100}{\sum_{i=1}^{n} (Y_i^{obs})} \right] \tag{3.11}$$

where Y_i^{obs} is the i^{th} observed data, Y_i^{sim} is the i^{th} simulated data, and n is the total number of observed data.

3.3.4.2 Modification of the SWAT 2012 Code for Baseflow Calibration

Among various SWAT output files, the ‘output.rch’ file (main channel output file) has been used to analyze streamflow at the outlet of each subbasin because USGS-provided flow is measured at the outlet of each subbasin. Also, SWAT users have used the

‘output.rch’ file for calibration, validation, sensitivity, and uncertainty analysis of streamflow at the outlet of each subbasin because the ‘output.rch’ file contains summary information for each routing reach in the watershed (Neitsch et al., 2011). However, only streamflow is calculated and printed in the ‘output.rch’ file while baseflow is not. Modified SWAT 2012 code for baseflow calibration allows users to calibrate automatically streamflow and baseflow simultaneously. Through graphical and numerical calibration, manual calibration for both streamflow and baseflow-related parameters is available as well as autocalibration. Moreover, if users use the SWAT-CUP interface, it would be easy to check the simulation results and summary statistics file which shows the statistics of comparing observed data with the simulation band through P-factor and R-factor and the optimal simulation of the current iteration by using various types of the model evaluation methods such as R^2, NSE, and PBIAS (Abbaspour et al., 2011).

SUFI-2 reads ‘output.rch’ in ‘TxtInOut’ folder which contains all output files of SWAT and ‘output.rch’ is used as an input file for the sequence of program execution in SUFI-2. In this study, SWAT 2012 code was modified with the FORTRAN programming to calculate the amount of baseflow for efficient baseflow calibration using SUFI-2. Baseflow (Equation 3.12) is defined as a summation of lateral flow and groundwater flow (Equations 3.1-3.6). If tile drainage is installed and simulated in SWAT, baseflow is calculated by adding Equation 3.13 because water entering tiles is considered lateral flow (Equation 3.14) (Neitsch et al., 2011).

$$Q_{bf,i} = Q_{lat,i} + Q_{gw,i} \tag{3.12}$$

where $Q_{bf,i}$ is the total baseflow on day *i* (mm H_2O), $Q_{lat,i}$ is the amount of lateral flow released to the main channel on day *i* (mm H_2O), and $Q_{gw,i}$ is the total groundwater flow on day *i* (mm H_2O).

$$Tile_{wtr,i} = \frac{h_{wtbl,i} - h_{drain,i}}{h_{wtbl,i}}(SW_i + FC_i)\cdot\left(1-\exp\left[\frac{-24}{t_{drain}}\right]\right) \qquad \text{if } h_{wtbl,i} > h_{drain,i} \tag{3.13}$$

where $Tile_{wtr,i}$ is the amount of water removed from the layer on day *i* by tile drainage (mm H_2O), $h_{wtbl,i}$ is the height of the water table above the impervious zone on day *i* (mm), $h_{drain,i}$ is the height of the tile drain above the impervious zone on day *i* (mm), SW_i is the water content of the profile on day *i* (mm H_2O), FC_i is the field capacity water content of the profile on day *i*, and t_{drain} is the time required to drain the soil to field capacity (hrs).

$$Q_{bf,tile,i} = Q_{lat,i} + Tile_{wtr,i} + Q_{gw,i} \tag{3.14}$$

where $Q_{bf,tile,i}$ is the total baseflow on day *i* when tile drainage is applied (mm H_2O).

Moreover, baseflow is printed in 'output.rch' to be able to calibrate baseflow for the main outlet of the watershed using SUFI-2 in SWAT-CUP. Thus, SWAT users can identify baseflow simulation results and the statistics of comparing observed data with the simulated data in the SWAT-CUP interface. As a result, SUFI-2 in SWAT-CUP has

become easy and efficient for SWAT users to calibrate baseflow using modified SWAT 2012 code.

3.3.4.3 Calibration of Streamflow and Baseflow with the EFCR

For accurate estimation of streamflow and baseflow and for better parameterization of them, the EFCR is proposed and evaluated. SUFI-2 uses LHS to generate controlled random samples for the calibration parameters. In the EFCR, the calibration parameters corresponding to the top 20% optimal NSE are updated three times for each 500 iterations based on the previous calibration results in order to reduce the size of LHS to the population size by adjusting and resorting parameter ranges for improving calibration accuracy. The following steps are suggested for the EFCR with SUFI-2 and modified SWAT 2012 code (Figure 3.5):

1. Conduct baseflow separation from USGS streamflow data using the WHAT system and use the baseflow as observed baseflow in the baseflow calibration procedure. If there is observed baseflow, this step can be omitted.
2. Determine initial parameters for the observed data (i.e., streamflow and baseflow) from the previous studies and run SWAT with the initial parameters.
3. Define calibration and validation periods from the entire simulation period.
4. Run the SUFI-2 in SWAT-CUP 500 iterations with the modified SWAT2012 executable file (swat2012.exe) using parallel processing with 4 CPUs, extract parameters corresponding to top 20% optimal NSE, and update parameter ranges (lower and upper bounds) based on the extracted parameters.

5. Perform the sensitivity analysis to identify the most sensitive parameters. Based on the results, remove insensitive parameters from the calibration process.
6. Repeat step 4 and evaluate the model performance until model evaluation statistics such as NSE, R^2, PBIAS meet the criteria. However, this iteration has stopping point. I set maximum iteration for updating parameter ranges is five.

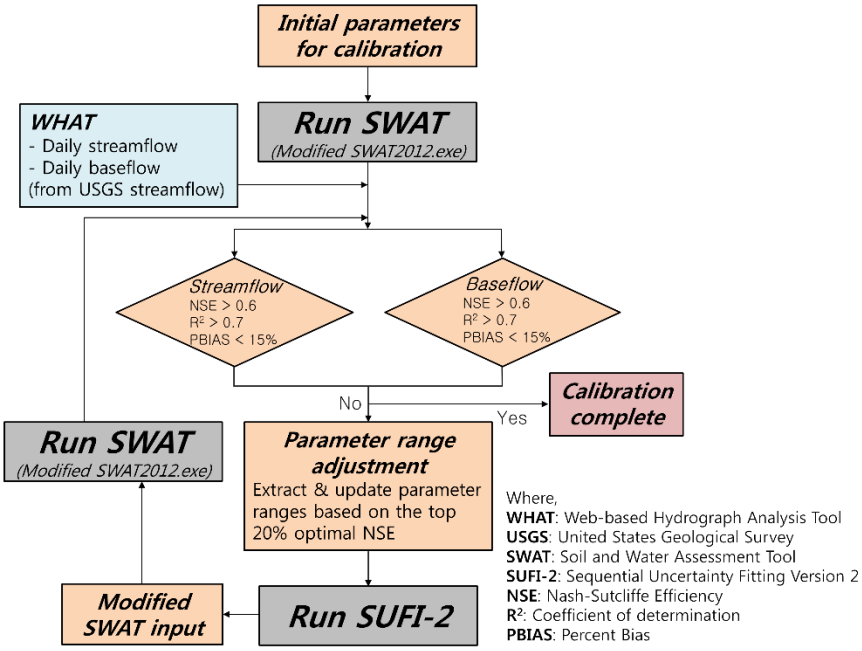


Figure 3.5 Flowchart of the procedures of the efficient flow calibration regime (EFCR)

There are four scenarios to evaluate streamflow and baseflow and determine which scenario is the most suitable calibration regime for accurate baseflow simulation. First is the default SWAT which is the model without streamflow and baseflow calibration (hereafter referred as C1). C1 is the baseline scenario (the default model) in this study to be compared with other scenarios to understand impact of baseflow calibration on the

levels of nitrates in aquifers. Second is the model with streamflow calibration alone (C2). The third scenario is the model for which streamflow and baseflow calibration were conducted simultaneously (C3). The last scenario is the model with calibration of streamflow and baseflow using the EFCR proposed in this study (C4) (Figure 3.5).

The total simulation period was from 1990 - 2010 (21 years) with the first 4 years as the model warm up period. Calibration (1990 - 2001) was carried out with 16 parameters and corresponding parameter ranges (Table 3.6) based on the result of sensitivity analysis and previous studies (Arnold et al., 2012; Zhang et al., 2011; Park et al., 2014). Sixteen parameters were selected based on sensitivity analysis to identify a set of key parameters for model calibration. These data were adjusted to minimize the differences between simulated and observed streamflow and baseflow during the calibration process. Among the 16 parameters, 10 parameters are related to streamflow calibration, and 6 parameters are more related to baseflow calibration. After the model calibration, validation was performed from 2002 - 2010 with the calibrated parameters. The model performance of each scenario was assessed for monthly streamflow and baseflow simulations with the three statistical metrics (i.e., NSE, R^2, and PBIAS) (Equation 3.9-3.11) and the reported model performance ratings (Table 3.7) (Moriasi et al., 2007; Van Liew et al., 2003; Singh et al., 2004; Engel et al., 2007). The performance ratings were utilized to evaluate the success rate of calibration results. Calibration results were categorized into four model performances: “Very good”, “Good”, “Satisfactory”, and “Unsatisfactory”. Then, percentage of model performances for streamflow and baseflow simulation were analyzed based on Table 3.7.

Table 3.6 SWAT parameters for calibration of streamflow and baseflow

Parameter	Description	Unit	Variation Method	Ranges	
				LB[1]	UB[2]
Water balance					
ESCO	Soil evaporation compensation factor	-	I[3]	0.01	1
SFTMP	Snowfall temperature	°C	I	-5	5
SMTMP	Snow melt base temperature	°C	I	-5	5
TIMP	Snow pack temperature lag factor	-	I	0.01	1
SMFMX	Melt factor for snow on June 21	mm/°C	I	0.01	10
SMFMN	Melt factor for snow on December 21	mm/°C	I	0.01	10
Subsurface water					
GW_REVAP	Groundwater evaporation coefficient	-	I	0.02	0.2
REVAPMN	Depth of water for evaporation	mm	I	0.01	250
GWQMN	Depth of water for return flow	mm	I	0.01	500
GW_DELAY	Groundwater delay time	day	I	0.1	50
ALPHA_BF	Baseflow alpha factor	1/day	I	0.1	1
RCHRG_DP	Deep aquifer percolation fraction	-	I	0.01	1
Surface runoff					
CN2	Initial SCS runoff curve number (CN II)	-	II[4]	-0.25	0.25
SURLAG	Surface runoff lag coefficient	-	I	0.1	10
Physical properties of the soil					
SOL_AWC	Available water capacity of the soil layer	mm/mm	II	-0.25	0.25
Physical properties of the channel					
CH_K2	Effective hydraulic conductivity	mm/hr	I	5	300

[1]LB: Lower bound / [2]UB: Upper bound

[3]I: Replace by value

[4]II: Multiply by value (%)

[5]III: Add to value

Table 3.7 SWAT performance evaluation criteria for NSE, R^2, and PBIAS

Measure	Output	Performance Evaluation Criteria			
		Very Good	Good	Satisfactory	Unsatisfactory
NSE	Flow	> 0.80	0.70 < NSE ≤ 0.80	0.50 < NSE ≤ 0.70	NSE ≤ 0.50
R^2	Flow	> 0.85	0.75 < R^2 ≤ 0.85	0.60 < R^2 ≤ 0.75	R^2 ≤ 0.60
PBIAS	Flow	< ±10	±10 ≤ PBIAS < ±15	±15 ≤ PBIAS < ±25	PBIAS ≥ ±25

Adapted by Van Liew et al. (2003), Singh et al. (2004), and Moriasi et al. (2015)

3.3.4.4 Calibration Effects on Monthly Hydrograph

After streamflow and baseflow were calibrated and validated, analysis of monthly hydrograph characteristics based on three hydrological conditions (wet, normal, and dry

years) was implemented for four scenarios (C1, C2, C3, and C4). Season-based evaluation is important for efficient hydrologic simulation and a sufficient range of hydrologic events based on wet, normal, and dry years would be required (Gan et al., 1997; Mueleta, 2012; Zhang, et al., 2012). In order to determine the hydrological conditions (wet, normal, and dry years), the USGS streamflow data (USGS 03354000 White River near Centerton, IN) were utilized from 1990 to 2010. Wet years correspond to the years with USGS streamflow above the 20^{th}-percentile exceedance level from April to July. Dry years are defined when the USGS streamflow is below the 80^{th}-percentile exceedance level from April to July. Normal years are the years between the 20^{th}-percentile exceedance level and 80^{th}-percentile exceedance level.

3.4 Results

3.4.1 Baseflow Separation using the WHAT System

The WHAT system (Lim et al., 2005) which has a user friendly and fully automated interface was used to implement baseflow separation from USGS streamflow to calibrate baseflow simulated by SWAT. For a total simulation period of 21 years (1990 - 2010), monthly mean streamflow was 85.7 m^3/s and monthly mean baseflow was 35.9 m^3/s, respectively. The amount of baseflow contributing to streamflow was 41.9% (monthly mean value). Min, max, and standard deviation of baseflow were 8.07, 231, and 28.8 m^3/s (Figure 3.6).

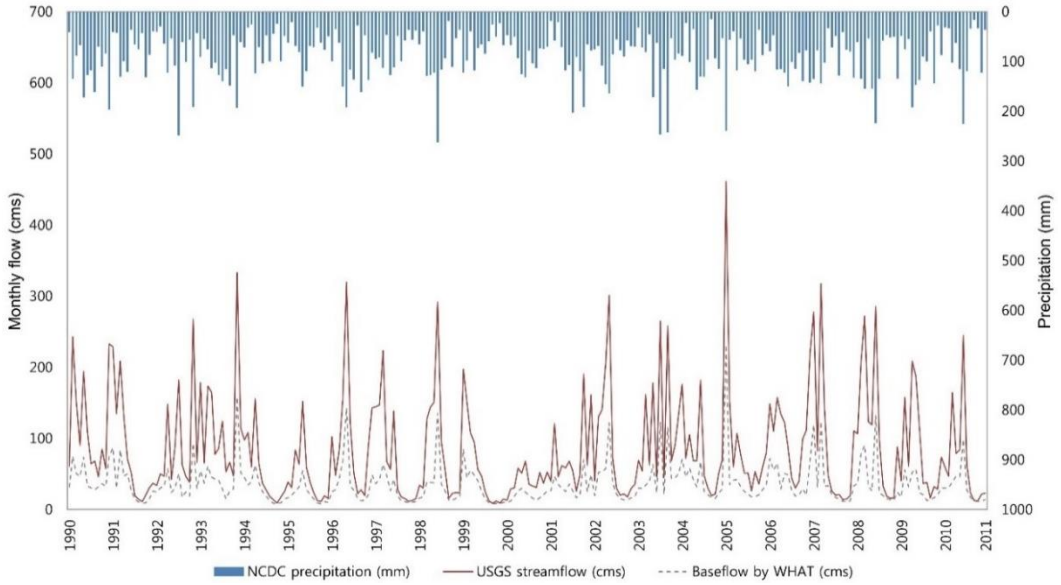


Figure 3.6 USGS streamflow and baseflow separated by the WHAT system (1990 - 2010) to calibrate streamflow and baseflow simulated by SWAT

3.4.2 Model Performance Evaluation

After streamflow and baseflow were simulated, model calibration was conducted with the four scenarios. Based on the calibration results, model performances were analyzed by the three model evaluation methods such as NSE, R^2, and PBIAS. As shown in Table 3.8, compared to default parameter values in C1 (uncalibrated model called the baseline model), adjusted ALPHA_BF and RCHRG_DP were considerably different. ALPHA_BF, the baseflow recession constant, varies from 0.1 to 0.3 for land with slow response to groundwater recharge and from 0.9 to 1.0 for land with a rapid recharge response. Thus, ALPHA_BFs (0.52 to 0.70) of C2, C3, and C4 indicate a slightly rapid recharge response. RCHRG_DP, deep aquifer percolation fraction, increase deep aquifer recharge which ranges from 0.0 - 1.0. This parameter is important for both streamflow and baseflow simulation because in SWAT water traveling to a deep aquifer from the

root zone is not redistributed into the soil, shallow aquifer or main channel but losses from the system boundary. Parameters of RCHRG_DPs (0.31 to 0.66) in C2, C3, and C4 were increased, which indicates water traveling to the deep aquifer was increased.

Table 3.8 Final values of the calibration parameters for each scenario

Parameter	Final values			
	C1[1]	C2[2]	C3[3]	C4[4]
ESCO	0.95	0.49	0.31	0.35
SFTMP	1.00	-4.11	-1.09	2.31
SMTMP	0.50	1.41	0.75	-0.77
TIMP	1.00	0.54	0.61	0.93
SMFMX	4.50	7.99	4.05	2.50
SMFMN	4.50	6.65	2.25	2.50
GW_REVAP	0.02	0.05	0.19	0.04
REVAPMN	750	165	238	210
GWQMN	1000	125	218	52.5
GW_DELAY	31.0	4.87	10.4	21.0
ALPHA_BF	0.048	0.53	0.52	0.70
RCHRG_DP	0.05	0.31	0.66	0.40
CN2[6]	va[5]	-0.13	-0.25	-0.23
SURLAG	4.00	6.72	4.37	7.42
SOL_AWC[6]	va	0.13	-0.12	-0.15
CH_K2	0.00	95.0	269	34.8

[1]C1: Default model without calibration (baseline)

[2]C2: Streamflow calibration alone

[3]C3: Streamflow and baseflow calibration

[4]C4: Streamflow and baseflow calibration with the EFCR

[5]va: Different value according to HRUs

[6]CN2, SOL_AWC: Percentage change (%)

Model performance evaluation was conducted according to NSE, R^2, and PBIAS (Table 3.9). For the monthly streamflow calibration and validation, NSE values of all scenarios range from 0.73 to 0.95. R^2 values vary from 0.74 to 0.95, and PBIAS values range from -6.42 to 11.2. Based on Table 3.7, all scenarios are within the “Very good”, “Good”, and

"Satisfactory" ranges. It indicates all scenarios are adequate. However, for the monthly baseflow calibration and validation, NSE values of all scenarios vary from -0.06 to 0.65.

R^2 values range from 0.25 to 0.95, and PBIAS values vary from -6.42 to 11.16. Compared with the streamflow calibration and validation, the results of the baseflow calibration and validation show "Very good", "Good", "Satisfactory", and "Unsatisfactory" ranges. Especially, C2 (streamflow calibration alone) shows very good performance for streamflow calibration and validation, but for baseflow calibration and validation, NSE values vary from 0.21 to 0.39. These values indicate that the model performance for baseflow simulation are in "Unsatisfactory" ranges. C3 (simultaneous streamflow and baseflow calibration) shows the model performances for streamflow evaluation are more than "Good". For baseflow evaluation, all values are within "Good" and "Satisfactory" aside from four indications (NSE: 0.48 and PBIAS: 33.6, 34.5, and 32.5) of "Unsatisfactory" model performance. C4 (simultaneous streamflow and baseflow calibration with the EFCR) indicates all values for streamflow calibration and validation are within "Very good" in terms of NSE, R^2, and PBIAS. Also, for baseflow calibration and validation, all values fall within "Good" and "Satisfactory". However, even though C4 provided good model performances for both streamflow and baseflow simulation, streamflow and baseflow in the model were still slightly underestimated or overestimated.

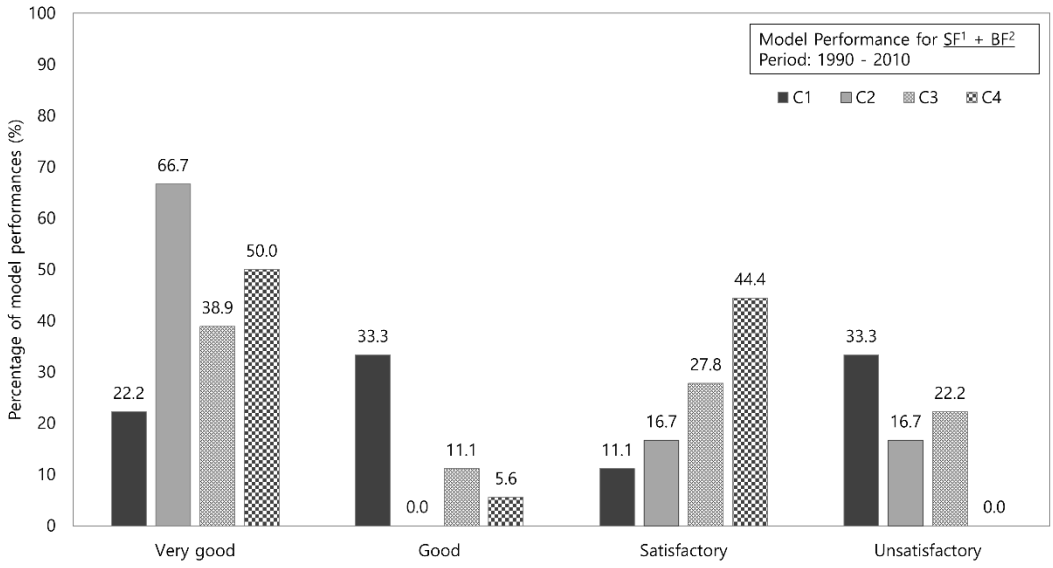
In all model performances (model performances of streamflow and baseflow) (Figure 3.7(a)) based on NSE, R^2, and PBIAS (Table 3.7), C1 which is the baseline model has

22.2 / 33.3 / 11.1 / 33.3% of “Very good” / “Good” / “Satisfactory” / “Unsatisfactory”. In C2, model performances show 66.7 / 0.0 / 16.7 / 16.7% of “Very good” / “Good” / “Satisfactory” / “Unsatisfactory”. C3 shows 38.9 / 11.1 / 27.8 / 22.2% of “Very good” / “Good” / “Satisfactory” / “Unsatisfactory”. In C4, model performances indicate 50.0 / 5.6 / 44.4 / 0.0% of “Very good” / “Good” / “Satisfactory” / “Unsatisfactory”. These values show C4 has fairly good model performances because C4 has 50.0% of “Very good” and 0.0% of “Unsatisfactory”. In model performances for streamflow calibration (Figure 3.7(b)), C2 and C4 have performed well for streamflow simulation because percentage of model performances indicate 100.0% of “Very good” for both C2 and C4. However, in model performances for baseflow calibration (Figure 3.7(c)), C2 show 33.3% of “Unsatisfactory” while C4 has still performed fairly well (11.1 / 88.9% of “Good” / “Satisfactory”) without “Unsatisfactory”. These values indicate that streamflow calibration alone (C2) would result in only good performance but not be good for baseflow simulation. Baseflow-related parameters would be overestimated or underestimated when streamflow calibration alone is conducted. Thus, streamflow and baseflow calibration should be carried out simultaneously for accurate baseflow-related simulation such as nitrate contamination in aquifers.

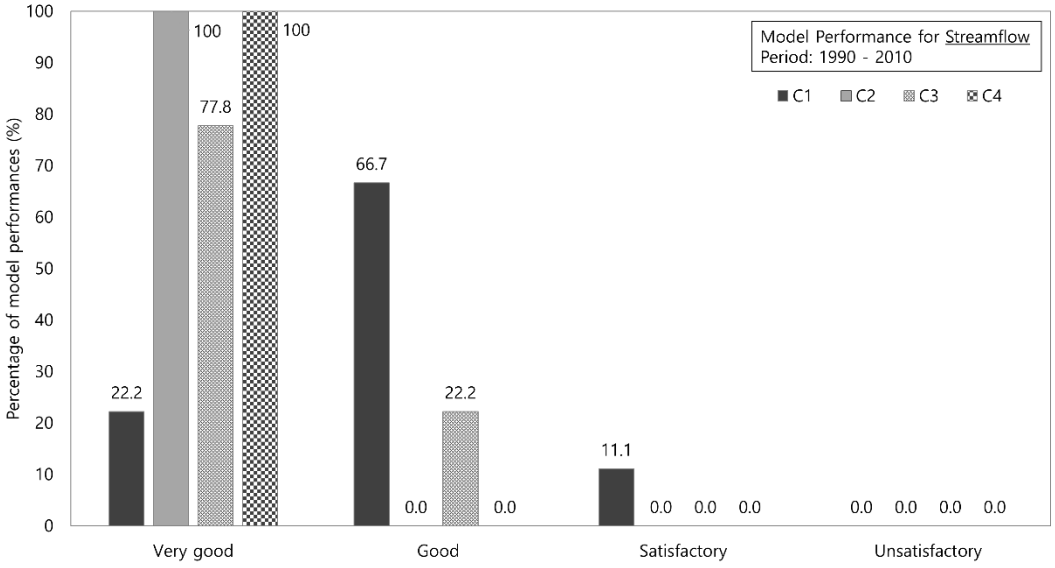
Table 3.9 Performance evaluation of streamflow and baseflow simulated from the four scenarios

	Scenario	Streamflow			Baseflow		
		NSE	R^2	PBIAS[1]	NSE	R^2	PBIAS
Total period (1990 - 2010)	C1	0.76	0.77	6.62	-0.01	0.27	8.29
	C2	0.93	0.93	4.69	0.32	0.70	-4.99
	C3	0.83	0.86	-2.49	0.54	0.71	33.6
	C4	0.87	0.89	1.36	0.64	0.71	15.3
Calibration (1990 - 2001)	C1	0.73	0.74	11.2	-0.06	0.28	13.5
	C2	0.91	0.92	8.55	0.21	0.69	-2.31
	C3	0.80	0.84	1.04	0.48	0.70	34.5
	C4	0.85	0.87	3.90	0.63	0.73	16.7
Validation (2002 - 2010)	C1	0.79	0.79	1.57	0.01	0.25	2.48
	C2	0.95	0.95	0.40	0.39	0.71	-7.98
	C3	0.85	0.90	-6.42	0.57	0.72	32.5
	C4	0.88	0.92	1.50	0.65	0.70	13.8

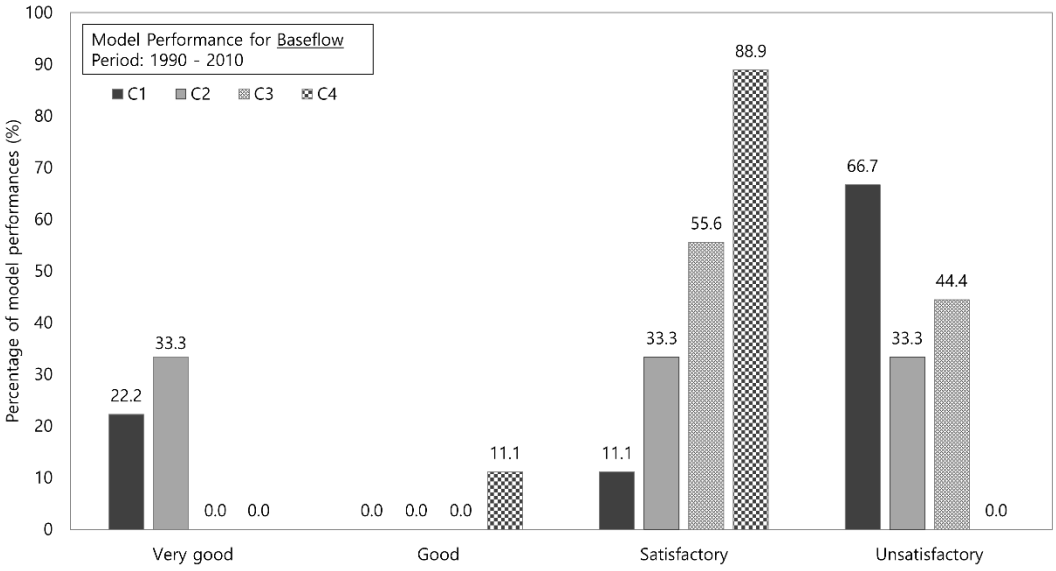
[1]PBIAS: Percent (%)



(a)



(b)



(c)

Figure 3.7 Percentage of model performances in four categories (very good, good, satisfactory, and unsatisfactory): (a) All model performances (streamflow + baseflow), (b) Model performances for streamflow, and (c) Model performances for baseflow

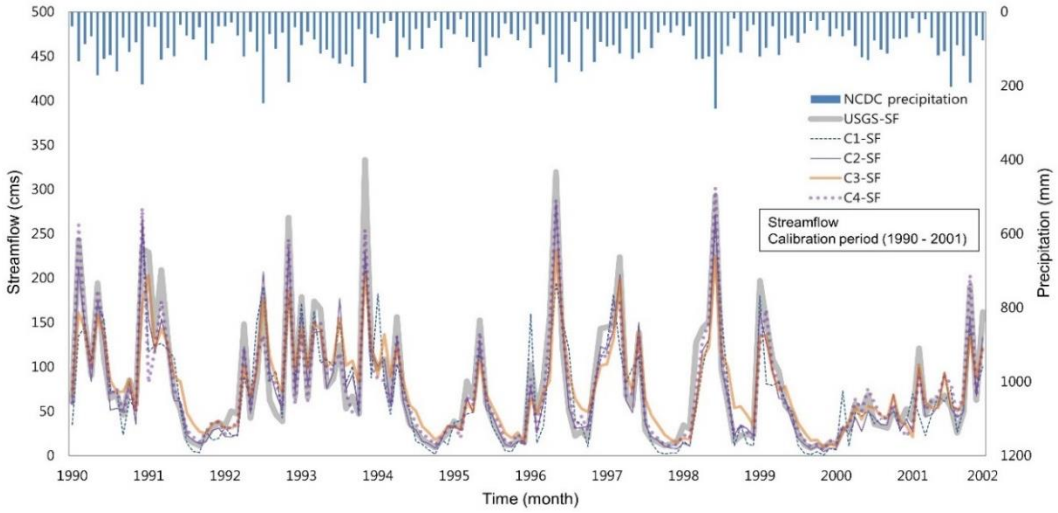
Note: [1]SF: Streamflow and [2]BF: Baseflow

3.4.3 Calibration Effects on Monthly Hydrograph

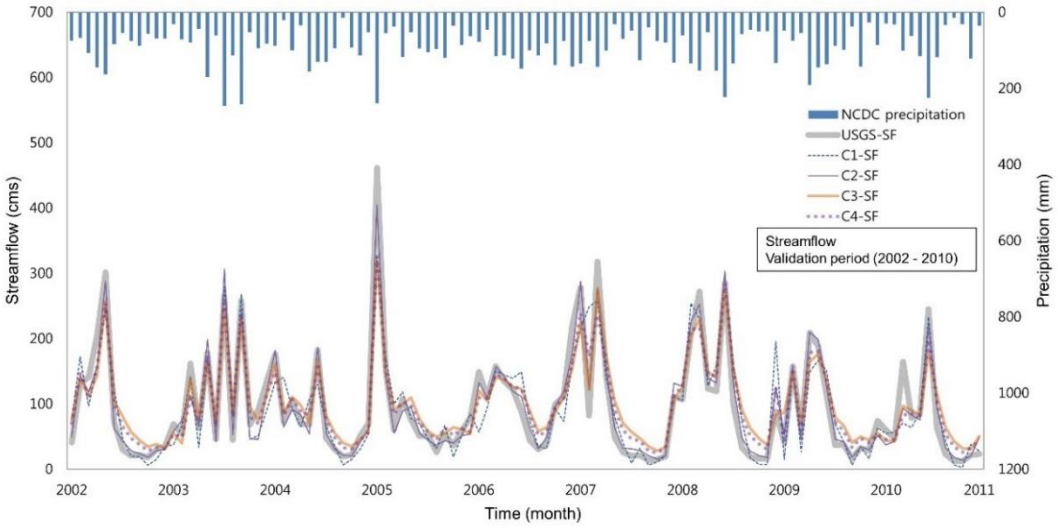
Graphical comparisons between the observed and simulated monthly streamflow during the calibration and validation periods are shown in Figure 3.8(a) and (b). As shown in Figure 3.8(a) and (b), even though some peak and low streamflow values were underestimated, three scenarios (C2, C3, and C4) performed well for monthly streamflow simulation. C1 mostly underestimated both peak and low streamflow. In C2, better peak and low streamflow simulation was achieved than for other scenarios. While C3 and C4 were underestimated during peak streamflow periods, they slightly overestimated during low streamflow periods. Generally, C2 shows the best performances (NSE: 0.93, R^2: 0.93, and PBIAS (%): 4.69) for streamflow simulation for both calibration and validation periods (Figure 3.9(a)), and streamflow for the four scenarios during the validation period was better estimated during the calibration period. Figure 3.8(c) and (d) show results of monthly baseflow simulation during the calibration and validation periods. In C1, simulated baseflow did not match observed baseflow well in most peak points because an obvious time lag existed during whole simulation periods. C2 strikingly overestimated baseflow during both calibration and validation periods. However, in C3 and C4, baseflow was simulated satisfactorily. Especially, C4 estimated baseflow better than C3 during all periods. Thus, it is noted that C4 simulates both streamflow and baseflow well. Table 3.10 shows summary statistics for the USGS observed flow (streamflow and baseflow) and simulated flow by various scenarios. As mentioned above, for streamflow simulation during both calibration and validation periods, the results of statistical analysis (i.e., mean, standard deviation, and max) of calibrated streamflow by C2, C3, and C4 are not significantly different from the observed streamflow. These values indicate all

calibrated streamflow replicated the observed streamflow well. However, for baseflow simulation during both calibration and validation periods, only C3 and C4 simulated satisfactory results based on mean, standard deviation, and maximum values. Among the four scenarios, C4 shows the best performances (NSE: 0.64, R^2: 0.71, and PBIAS (%): 15.32) for baseflow simulation (Figure 3.9(b)). Thus, C4 performs both streamflow and baseflow simulation well compared with other scenarios. While C2 shows the best performances for streamflow simulation, significant differences were found between the observed and simulated baseflow in C2 because C2 represents model calibration for streamflow alone. The results suggest simultaneous streamflow and baseflow calibration should be conducted for accurate flow estimation.

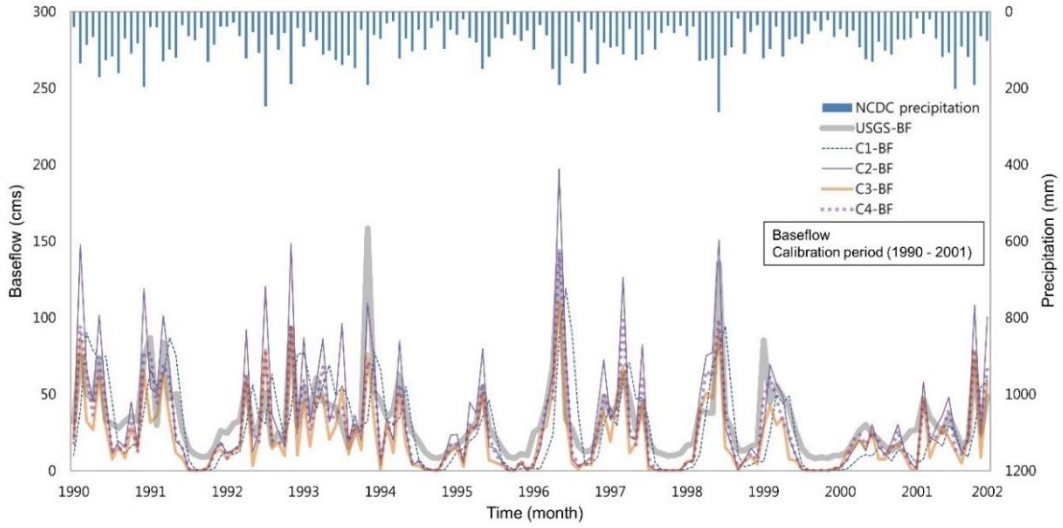
Using the calibrated parameters of C4, cross validation was conducted to examine whether calibrated parameters of C4 are suitable for other USGS streamflow stations. As shown in Table 3.11, simulated streamflow using calibrated parameters of C4 replicated the all observed streamflow (10 USGS streamflow stations) well.


Streamflow (cms)
Precipitation (mm)
Time (month)
NCDC precipitation
USGS-SF
C1-SF
C2-SF
C3-SF
C4-SF
Streamflow
Calibration period (1990 - 2001)

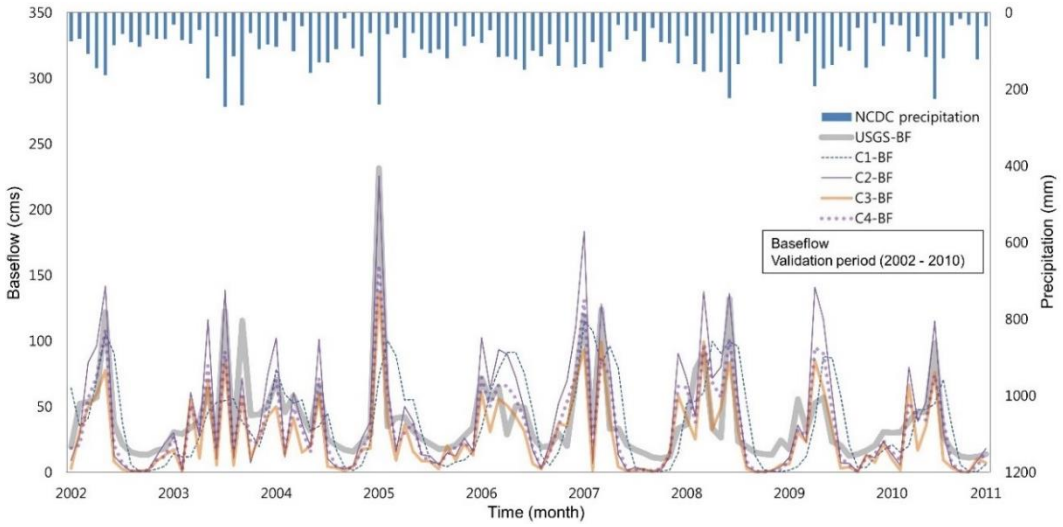
(a)


Streamflow (cms)
Precipitation (mm)
Time (month)
NCDC precipitation
USGS-SF
C1-SF
C2-SF
C3-SF
C4-SF
Streamflow
Validation period (2002 - 2010)

(b)

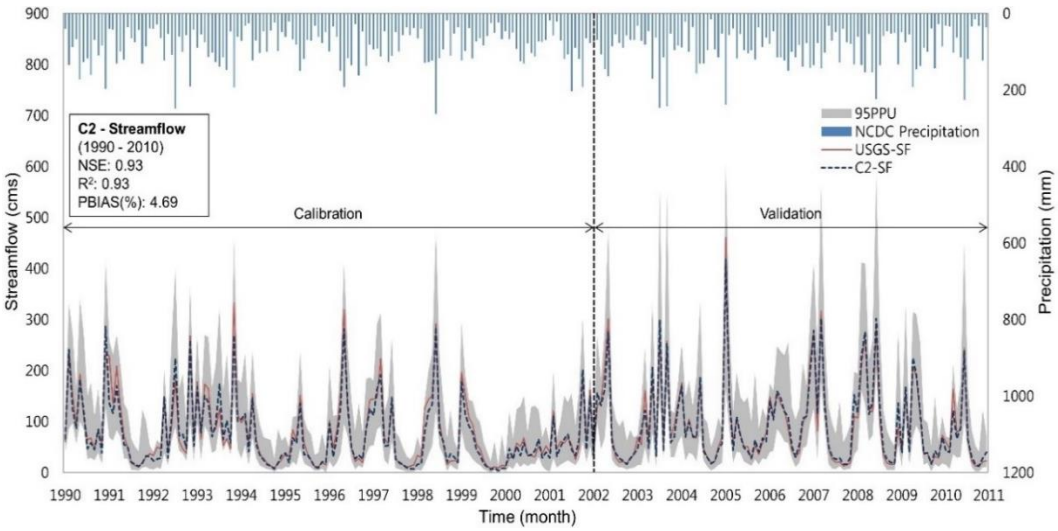


(c)

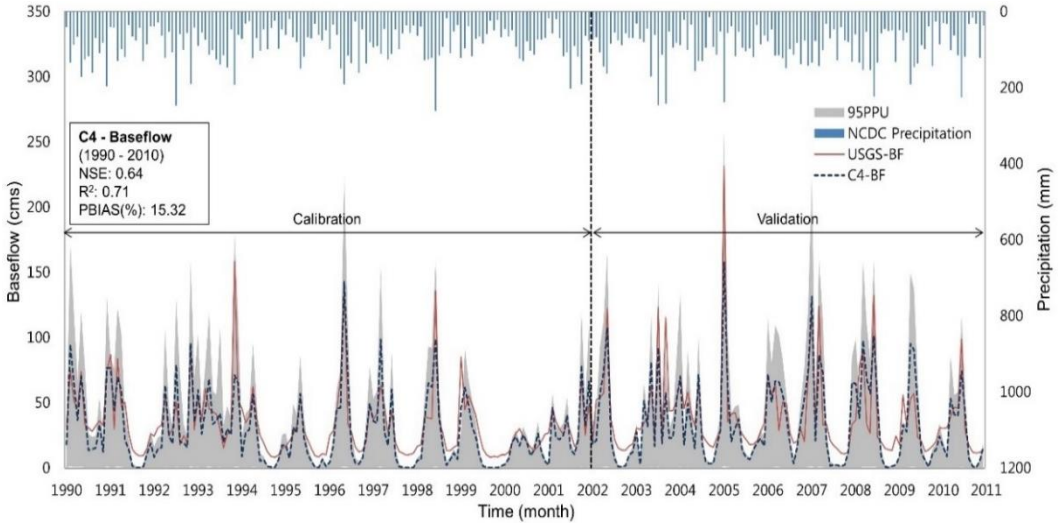


(d)

Figure 3.8 Comparisons of monthly streamflow in the UWRW between four scenarios: (a) Streamflow during the calibration period (1990 - 2001), (b) Streamflow during the validation period (2002 - 2010), (c) Baseflow during the calibration period (1990 - 2001), and (d) Baseflow during the validation period (2002 - 2010)



(a)



(b)

Figure 3.9 The best performances of streamflow and baseflow simulations in the UWRW: (a) Streamflow hydrograph of C2 (the model with streamflow calibration alone), and (b) Baseflow hydrograph of C4 (the model with simultaneous streamflow and baseflow calibration)

Table 3.10 Summary flow statistics for the USGS observed and simulated streamflow and baseflow from the four scenarios in the UWRW

	Scenario	Streamflow (cms)				Baseflow (cms)			
		Mean	Std[2]	Max	Min	Mean	Std	Max	Min
Calibration (1990 - 2001)	USGS[1]	78.9	68.1	333	8.67	33.2	24.8	158	8.06
	C1	70.2	57.3	259	0.71	28.7	26.9	119	0.17
	C2	72.2	60.0	285	4.32	33.9	37.8	197	0.15
	C3	78.1	48.8	231	9.83	21.7	22.9	110	0.10
	C4	76.4	51.2	248	8.48	27.6	26.5	143	0.12
Validation (2002 - 2010)	USGS	94.7	82.4	461	12.5	39.6	33.0	232	10.8
	C1	93.2	75.1	324	2.63	38.6	32.7	115	0.54
	C2	94.3	78.2	405	13.3	42.7	46.2	226	0.29
	C3	101	60.9	319	28.2	26.7	28.2	137	0.20
	C4	98.6	64.1	331	24.9	34.1	32.1	158	0.27

[1]USGS observed monthly flow data

[2]Standard deviation of monthly flow

Table 3.11 Results of multi-site validation using calibration parameters of C4 (the model with simultaneous streamflow and baseflow calibration) in the UWRW

ID	Monitoring station	R^2	NSE	PBIAS	Period
Flow #1	USGS 03349510	0.91	0.89	-14.3	2004-2010
Flow #2	USGS 03347000	0.77	0.75	10.7	1990-2010
Flow #8	USGS 03348130	0.90	0.90	-2.7	1999-2010
Flow #10	USGS 03348000	0.83	0.80	15.9	1990-1993
Flow #11	USGS 03349000	0.89	0.89	1.2	1990-2010
Flow #17	USGS 03351000	0.91	0.90	-2.2	1990-2010
Flow #20	USGS 03353500	0.88	0.86	-16.4	1990-2010
Flow #23	USGS 03353611	0.92	0.91	-3.9	1992-2010
Flow #26	USGS 03353800	0.89	0.88	-11.8	1990-2010
Flow #28	USGS 03354000	0.92	0.88	1.5	1990-2010

3.4.4 Calibration Effects on Streamflow and Baseflow

As shown in Table 3.12, wet years are the 4 years with streamflow above the 20th-percentile exceedance level and normal years occur for 13 years. Dry years are the 4 years for which streamflow is below the 80th-percentile exceedance level. In order to conduct season-based evaluation in streamflow and baseflow across all scenarios, visual analysis was implemented through the observed and simulated streamflow and baseflow

exceedance probability distributions based on the hydrologic conditions (Table 3.9) (Figures 3.10 and 11).

In wet years for streamflow (Figure 3.10(a)), C1 underestimated streamflow in dry conditions and low flow conditions, and C3 and C4 overestimated streamflow in dry conditions. However, C2 replicated the observed streamflow well. In normal years for streamflow (Figure 3.10(b)), C2 also simulated the observed streamflow well. However, C1 underestimated streamflow in low flow conditions and C3 overestimated streamflow in dry conditions. C4 also marginally overestimated streamflow in low flow conditions. In dry years (Figure 3.10(c)), all scenarios simulated streamflow well except for C1 which underestimated streamflow in dry and low flow conditions.

In wet years for baseflow (Figure 3.11(a)), all scenarios underestimated baseflow in dry and low flow conditions. C1 and C2 overestimated baseflow in moist and mid-range flow conditions. C3 simulated baseflow well in high flow, moist, and mid-range flow conditions. C4 slightly overestimated baseflow in moist and mid-range flow conditions. In normal years for baseflow (Figure 3.11(b)), C1 and C4 simulated baseflow well in high flow and moist conditions but underestimated baseflow in dry and low flow conditions. C2 slightly overestimated baseflow in high flow and moist conditions but underestimated baseflow in dry and low flow conditions. In mid-range, dry, and low flow conditions, C3 underestimated baseflow more than the other three scenarios. In dry years for baseflow (Figure 3.11(c)), C1, C3, and C4 simulated baseflow well in high flow and moist conditions but underestimated baseflow in dry and low flow conditions. C2

slightly overestimated baseflow in high flow conditions but underestimated baseflow in dry and low flow conditions.

As a result, all scenarios are likely to conduct better simulation of streamflow rather than baseflow. All scenarios simulated streamflow and baseflow well in high flow conditions but in low flow conditions, all models underestimated streamflow and baseflow. Even though all scenarios overestimated and/or underestimated baseflow, C4 simulated baseflow better than the other three scenarios.

Table 3.12 Classification of wet, normal, and dry years to conduct season-based evaluation in streamflow and baseflow in the UWRW

Year	USGS-SF[1] (cms) (Apr - Jul)	Exceedance Probability (%)	Hydrological Condition
1998	683.3	4.5	Wet year
1996	648.5	9.1	
2008	627.3	13.6	
2002	609.4	18.2	
2003	551.7	22.7	Normal year
2009	546.3	27.3	
2010	471.2	31.8	
1992	462.4	36.4	
1990	458.5	40.9	
1993	453.3	45.5	
2006	389.0	50.0	
2004	367.2	54.5	
1995	314.8	59.1	
1997	291.8	63.6	
2005	286.1	68.2	
1994	285.2	72.7	
1991	271.9	77.3	
2001	243.5	81.8	Dry year
2007	236.4	86.4	
1999	223.7	90.9	
2000	213.1	95.5	

[1]USGS streamflow

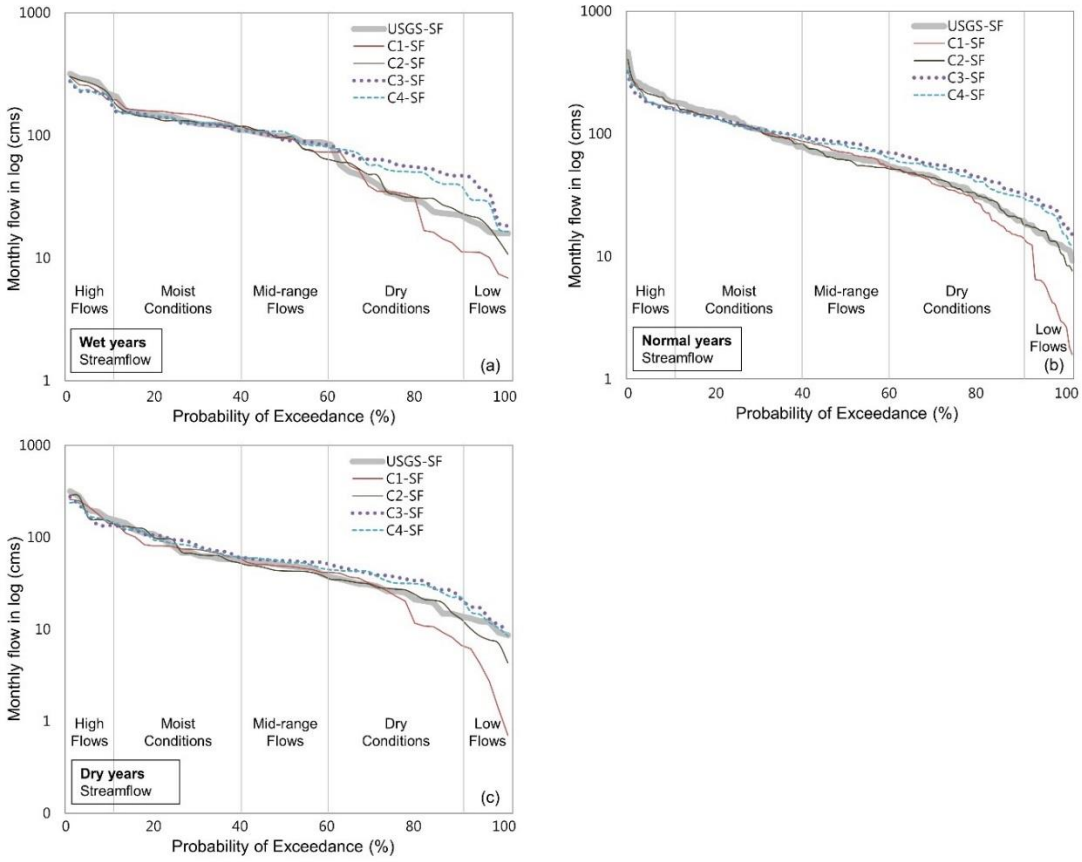


Figure 3.10 Exceedance probability distribution of the USGS observed and simulated streamflow based on the hydrologic conditions for the UWRW: (a) Wet years, (b) Normal years, and (c) Dry years

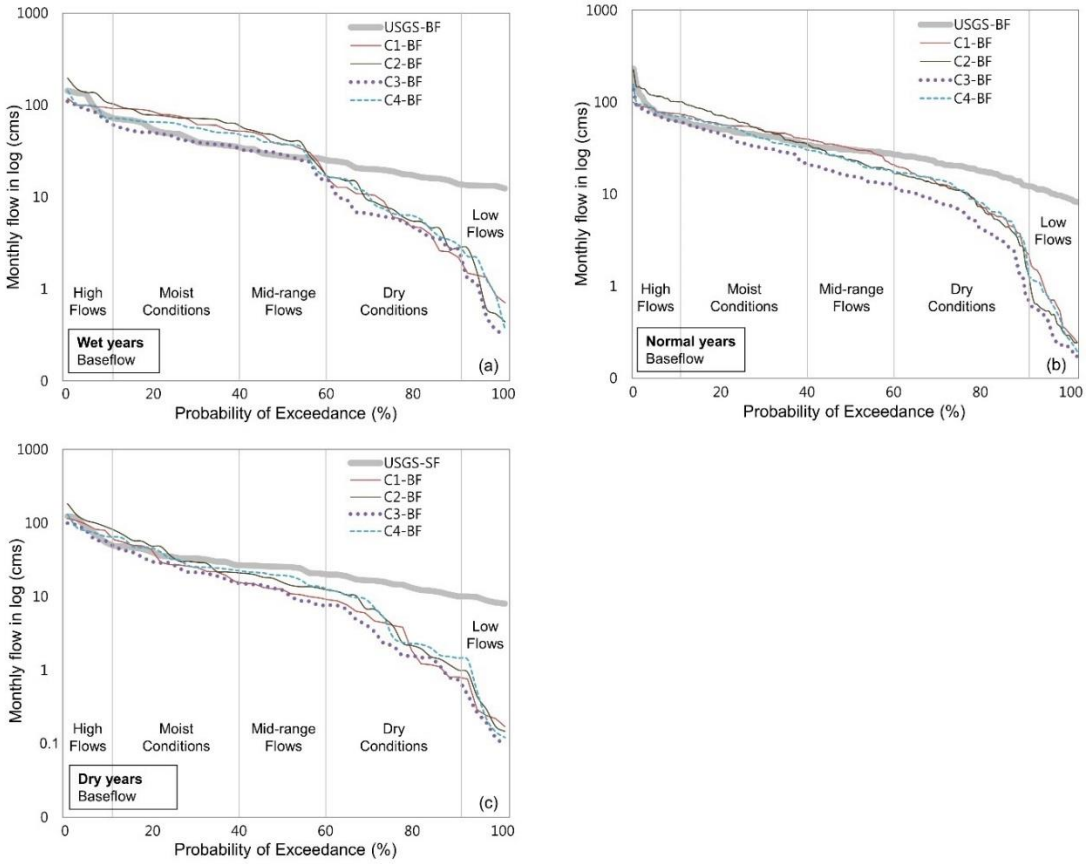


Figure 3.11 Exceedance probability distribution of the USGS observed and simulated baseflow based on the hydrologic conditions for the UWRW: (a) Wet years, (b) Normal years, and (c) Dry years

This study presented the efficient flow calibration regime (EFCR) for both accurate streamflow and baseflow estimation. The calibration methodology developed and suggested in this study was composed of the modified SWAT 2012 code and SUFI-2 in order to calibrate automatically streamflow and baseflow simultaneously. Measured streamflow and estimated baseflow (or observed baseflow if available) data are necessary to calibrate and validate the model performances. Measured streamflow was retrieved from USGS streamflow gauging station. However, it was difficult to obtain measured baseflow data. The filter-based model, WHAT system, was used to estimate baseflow

data by separating baseflow from a total streamflow hydrograph using USGS streamflow data.

Sixteen parameters were selected based on sensitivity analysis to identify a set of key parameters for the model calibration. Using these 16 parameters related to streamflow and baseflow, the model performance was evaluated based on Table 3.6 in the UWRW. The results of the calibration and validation showed that the simulated streamflow and baseflow agreed with the observed data well. In the case of C4 which is the model with simultaneous streamflow and baseflow calibration with the EFCR, for the total simulation period (1990 - 2010), NSE / R^2 / PBIAS for streamflow and baseflow were 0.87 / 0.89 / 1.36 (streamflow) and 0.64 / 0.71 / 15.3 (baseflow), respectively. For the calibration period (1990 - 2001), NSE / R^2 / PBIAS for streamflow were 0.85 / 0.87 / 3.90 and NSE / R^2 / PBIAS for baseflow were 0.63 / 0.73 / 16.7. For the validation period (2002 - 2010), NSE / R^2 / PBIAS for streamflow and baseflow showed 0.88 / 0.92 / 1.50 (streamflow) and 0.65 / 0.70 / 13.8 (baseflow), individually. These values indicate that the model is more than "Satisfactory" for all periods. In the results, it is noted that even though C2 (streamflow calibration alone) showed the best performances for both streamflow calibration and validation periods, C2 would not be appropriate for baseflow calibration and validation. Thus, this study indicates that calibrating streamflow and baseflow simultaneously would be important for t efficient hydrological cycle assessment because baseflow is the main part of the hydrological cycle for the study location. Moreover, for baseflow-related study, such as analysis of nitrate leaching for aquifer hazard assessment,

simultaneous streamflow and baseflow calibration should be necessary to achieve better results for baseflow-related simulation.

C3 and C4 performed streamflow and baseflow simulation well. However, C4 indicated better model performances than C3 because C4 reduced initial condition uncertainty by adjusting and resorting calibration parameter ranges based on the previous calibration process. It indicates a proper adjustment of parameter ranges will cause better convergence and next-better solutions as next iterations has a better region of the parameter space (Abbaspour et al., 2011). Using the validated model, streamflow and baseflow were evaluated based on the different hydrologic conditions (wet, normal and dry years). The results showed that all scenarios estimated streamflow well in high flow conditions, but in dry and low flow conditions, C3 and C4 slightly overestimated streamflow in dry and low flow conditions. In baseflow simulation, C1 and C2 overestimated baseflow in high flow conditions and underestimated it in dry and low flow conditions. Even though in high flow conditions baseflow values were simulated well, but C3 and C4 underestimated baseflow in dry and low flow conditions. It means all scenarios underestimated baseflow in dry and low flow conditions.

Dry and low flow conditions usually appear during the winter season. If the soil temperature module in SWAT underestimates soil temperature in the soil freezing/thawing area, less frozen soil would be thawed and it would cause the amount of soil moisture to be reduced at that time (Yang et al., 2014). Thus, proper soil temperature estimation would be important to enhance hydrological cycle representation by increasing

the amount of soil moisture in dry and low flow conditions. According to Benham et al. (2006), SWAT was not able to capture the conditions of a dry year in combination with flows and SWAT. Benaman et al. (2005) found that SWAT reasonably replicated streamflow but the model underestimated snowmelt-driven winter and spring streamflow. Tolston and Shoemaker (2007) found that lateral subsurface flow does not occur in frozen soils because if the soil temperature in a particular layer reaches less than or equal to 0 °C, no percolation is allowed from that layer (Neitsch et al., 2002). The uncertainty of meteorological input would exist because it would be difficult to measure small or accurate amount of precipitation. In addition, because NSE and R^2 were used as the model evaluation methods which are sensitive to high flow, low flow in dry and low flow conditions might be ignored. Thus, different model evaluation methods capturing low flow well should be utilized for better low flow estimation in dry and low flow conditions. Moreover, other parameters related to soil moisture and baseflow (not used in this study) would be required to conduct sensitivity and uncertainty analysis for more detailed description of baseflow simulation in dry and low flow conditions. For coarse-to-fine calibration, both autocalibration and manual calibration should be conducted together to capture flow well. Once autocalibration is implemented for tuning flows coarsely with recommended parameter ranges based on the watershed characteristics, fine tuning of flow-related parameters should be conducted by manual calibration. Wet, normal, and dry seasonal calibration would also be necessary to capture more detailed flow regimes.

As shown in Table 3.9, for baseflow calibration and validation, C4 performed baseflow simulation better than C3. However, the differences of NSE, R^2, and PBIAS between C3 and C4 are not significant. Even though the differences between C3 and C4 are insignificant in terms of the values computed for the three model evaluation methods, the differences would be very significant in terms of the model parameter values. For example, GW_REVAP (groundwater evaporation coefficient) / GWQMN (depth of water for return flow) / GW_DELAY (groundwater delay time) in C3 and C4 are 0.19 / 218 / 10.4 and 0.04 / 52.5 / 21, respectively. Compared with C3, C4 shows a decrease in GW_REVAP (-375%), and GWQMN (-315%) but an increase in GW_DELAY (51%).

3.5 Conclusions

A new calibration regime created by incorporating the modified SWAT 2012 code and SUFI-2 was developed for accurate streamflow and baseflow estimation by calibrating streamflow and baseflow simultaneously. This new calibration regime called the efficient flow calibration regime (EFCR) was tested for hydrologic and water quality modeling in the Upper White River Watershed, Indiana. The EFCR is a user-friendly baseflow calibration methodology for outlet-based calibration using SWAT. Both streamflow and baseflow using the EFCR performed well based on three model evaluation methods (i.e., NSE, R^2, and PBIAS). Calibration methodology should be flexible based on the purpose of research. The EFCR was developed for aquifer hazard assessment and this study focused on accurate baseflow calibration. Thus, the EFCR would be a practical method for aquifer hazard assessment by calibrating baseflow accurately as well as streamflow at a watershed scale. This study is expected to be used

as a data-driven model for in-depth groundwater modeling because the baseflow-related parameters (i.e., groundwater recharge, hydraulic conductivity, and so on) calibrated in this study can be used as a set of input data (initial parameter values) in computer-based numerical groundwater models. If calibrated baseflow-related data are used as initial parameter values in groundwater modeling, the uncertainty of groundwater modeling would be reduced by minimizing the initial parameter uncertainty.

This study has limitations that should be considered in future research. The EFCR should be applied in more watersheds and more uncertainty analysis of baseflow-related parameters should be conducted to determine if baseflow estimation can be improved further.

3.6 References

Abbaspour, K.C., 2011. Swat-Cup2: SWAT Calibration and Uncertainty Programs Manual Version 2, Department of Systems Analysis, Integrated Assessment and Modelling (SIAM), Eawag. Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland. 106 p.

Abbaspour, K.C., Johnson, C.A., Van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal 3, 1340–1352.

Abbaspour, K.C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., Kløve, B., 2015. A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. Journal of Hydrology 524, 733–752.

Abbaspour, K.C., Vejdani, M., Haghighat, S., 2007. SWAT-CUP calibration and uncertainty programs for SWAT, in: MODSIM 2007 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand.

Akhavan, S., Abedi-Koupai, J., Mousavi, S.-F., Afyuni, M., Eslamian, S.-S., Abbaspour, K.C., 2010. Application of SWAT model to investigate nitrate leaching in Hamadan–Bahar Watershed, Iran. Agriculture, Ecosystems & Environment 139, 675–688.

Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and storage in groundwater systems. Science 296, 1985-1991.

Arnold, J.G., Allen, P.M., 1999. Automated methods for estimating baseflow and ground water recharge from streamflow records. Journal of the American Water Resources Association 35 (2), 411-424.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M. K., 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE 55, 1491-1508.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: Model development. Journal of the American Water Resources Association 34 (1), 73-89.

Babiker, I.S., Mohamed, M.A.A., Terao, H., Kato, K., Ohta, K., 2004. Assessment of groundwater contamination by nitrate leaching from intensive vegeTable 3.cultivation using geographical information system. Environment International 29, 1009-1017.

Benham, B.L., Baffaut, C., Zeckoski, R.W., Mankin, K.R., Pachepsky, Y.A., Sadeghi, A.M., Brannan, K.M., Soupir, M.L., Habersack, M.J., 2006. Modeling bacteria fate and transport in watershed models to support TMDLs. Trans. ASABE 49 (4), 987-1002.

Beven, K., Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298.

Boles, C.M.W., Frankenberger, J.R., Moriasi, D.N., 2015. Tile Drainage Simulation in SWAT2012: Parameterization and Evaluation in an Indiana Watershed. Transactions of the ASABE 1201-1213.

Cerro, I., Antigüedad, I., Srinavasan, R., Sauvage, S., Volk, M., Sanchez-Perez, J.M., 2014. Simulating Land Management Options to Reduce Nitrate Pollution in an Agricultural Watershed Dominated by an Alluvial Aquifer. Journal of Environment Quality 43, 67.

Dahan, O., Babad, A., Lazarovitch, N., Russak, E.E., Kurtzman, D., 2013. Nitrate leaching from intensive organic farms to groundwater. Hydrology and Earth System Sciences Discussions 2013; 10 (7), 333-341.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19, 507–515.

Engel, B., Storm, D., White, M., Arnold, J., Arabi, M., 2007. A Hydrologic/Water Quality Model Applicati1. Journal of the American Water Resources Association 43, 1223–1236.

Feyereisen, G.W., Strickland, T.C., Bosch, D.D., Sullivan, D.G., 2007. Evaluation of SWAT manual calibration and input parameter sensitivity in the Little River watershed. Transactions of the ASABE 50, 843–855.

Fleming, A. H., Brown, S. E., Ferguson, V. R., 1993. The hydrogeologic framework of Marion County, Indiana at Atlas illustrating hydrogeologic terrain and sequence, Indiana Geological Survey Open File Report 93-5, 67.

Green, C.H., Tomer, M.D., Di Luzio, M., Arnold, J.G., 2006. Hydrologic evaluation of the soil and water assessment tool for a large tile-drained watershed in Iowa. Transactions of the ASABE 49, 413-422.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. Journal of Hydrologic Engineering 4, 135–143.

Her, Y., Chaubey, I., Frankenberger, J., Smith, D., 2016. Effect of conservation practices implemented by USDA programs at field and watershed scales. Journal of Soil and Water Conservation 71, 249-266.

Jang, W.S., Moon, J., Kim, N.W., Engel, B.A., Yoo, D.S., Kim, I.J., Mun, Y., Ok, Y.S., Lim, K.J., 2010. Evaluation of SWAT Direct Runoff and Baseflow Component using Web-based K-means Clustering EI Estimation System, in: 2010 Pittsburgh, Pennsylvania, June 20-June 23, 2010. American Society of Agricultural and Biological Engineers, p. 1.

Jiang, Y., Frankenberger, J.R., Sui, Y., Bowling, L.C., 2014. Estimation of Nonpoint Source Nitrate Concentrations in Indiana Rivers Based on Agricultural Drainage in the Watershed. J Am Water Resour Assoc 50, 1501-1514.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. IEEE International Conference on Neural Networks, 1995. Proceedings 4, 1942–1948.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences 5, 89–97.

Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. Journal of Hydrology 211, 69–85.

Li, C., Wang, H., Liu, J., Yan, D., Yu, F., Zhang, L., 2010. Effect of calibration data series length on performance and optimal parameters of hydrological model. Water Science and Engineering 3, 378–393.

Lim, K.J., Engel, B.A., Tang, Z., Choi, J., Kim, K.-S., Muthukrishnan, S., Tripathy, D., 2005. Automated web GIS based hydrograph analysis tool, WHAT. Journal of the American Water Resources Association 41 (6), 1407-1416.

Lim, K.J., Park, Y.S., Kim, J., Shin, Y.-C., Kim, N.W., Kim, S.J., Jeon, J.-H., Engel, B.A., 2010. Development of genetic algorithm-based optimization module in WHAT system for hydrograph analysis and model application. Computers & Geosciences 36, 936–944.

Molénat, J., Gascuel-Odoux, C., 2002. Modelling flow and nitrate transport in groundwater for the prediction of water travel times and of consequences of land use evolution on water quality. Hydrological Processes 16, 479-492.

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

Moriasi, D.N., Gowda, P.H., Arnold, J.G., Mulla, D.J., Ale, S., Steiner, J.L., Tomer, M.D., 2013. Evaluation of the Hooghoudt and Kirkham tile drain equations in the soil and water assessment tool to simulate tile flow and nitrate-nitrogen. J. Environ. Qual. 42, 1699–1710.

Muleta, M.K., 2011. Improving model performance using season-based evaluation. Journal of Hydrologic Engineering 17, 191–200.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology 10, 282–290.

Neitsch, S.L., Arnold, J.C., Kiniry, J.R., Williams, J.R., King, K.W., 2002. Soil and Water Assessment Tool Theoretical Documentation. Version 2000. Texas Water Resources Institute, College Station, Texas, USA.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool, Theoretical Documentation, Version 2009. Blackland Research Center, Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, TX.

Park, J.-Y., Ahn, S.-R., Hwang, S.-J., Jang, C.-H., Park, G.-A., Kim, S.-J., 2014. Evaluation of MODIS NDVI and LST for indicating soil moisture of forest areas based on SWAT modeling. Paddy Water Environ 12, 77–88.

Rathjens, H., Oppelt, N., 2012. SWAT model calibration of a grid-based setup. Advances in Geosciences 32, 55–61.

Rutledge, A.T., 1998. Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records: Update. US Department of the Interior, US Geological Survey.

Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M., 2001. Validation of the Swat Model on a Large River Basin with Point and Nonpoint Sources1. Journal of the American Water Resources Association 37, 1169-1188.

Scheidleder, A., Grath, J., Winkler, G., Stärk, U., Koreimann, C., Gmeiner, C., Nixo, S., Casillas, J., Gravesen, P., Leonard, J., Elvira, M., 1999. Groundwater quality and quantity in Europe. Copenhagen. European Environment Agency Report.

Shi, P., Hou, Y., Xie, Y., Chen, C., Chen, X., Li, Q., Qu, S., Fang, X., Srinivasan, R., 2013. Application of a SWAT Model for Hydrological Modeling in the Xixian Watershed, China. Journal of Hydrologic Engineering 18, 1522–1529.

Shrestha, M.K., Recknagel, F., Frizenschaf, J., Meyer, W., 2016. Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga catchment in South Australia. Agricultural Water Management. doi:10.1016/j.agwat.2016.02.009.

Singh, J., H. V. Knapp, and M. Demissie. 2004. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, Ill.: Illinois State Water Survey. Available at: www.sws.uiuc.edu/pubdoc/CR/ISWSCR2004-08.pdf. Accessed 2 April 2016.

Sloto, R.A., Crouse, M.Y., 1996. HYSEP, a computer program for streamflow hydrograph separation and analysis. US Department of the Interior, US Geological Survey.

Sui, Y., Frankenberger, J.R., 2008. Nitrate loss from subsurface drains in an agricultural watershed using SWAT2005. Transactions of the ASABE 51, 1263–1272.

Sultan, M., Metwally, S., Milewski, A., Becker, D., Ahmed, M., Sauck, W., Soliman, F., Sturchio, N., Yan, E., Rashed, M., Wagdy, A., Becker, R., Welton, B., 2011. Modern recharge to fossil aquifers: Geochemical, geophysical, and modeling constraints. Journal of Hydrology 403, 14–24.

Tedesco, L.P., Hoffmann, J., Bihl, L., Hall, B.E., Barr, R.C., Stouder, M., 2011. Upper White River Watershed Regional Watershed Assessment and Planning Report.

Thorsena, M., Refsgaarda, J.C., Hansenb, S., Pebesmac, E., Jensena, J.B. Kleeschulte, S., 2001. Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. Journal of Hydrology 242, 210-227.

Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour. Res. 43 (1), W01413.

USDA-NRCS, 2011. Scope and effect equations. Wisconsin supplement to Chapter 19: Hydrology tools for wetland determination, in Engineering Field Handbook, Part 650. EFH Notice 210-WI-121. Madison, WI: USDA-NRCS Wisconsin State Office.

Vale, M., Holman, I.P., 2009. Understanding the hydrological functioning of a shallow lake system within a coastal karstic aquifer in Wales, UK. Journal of Hydrology 376, 285–294.

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

Van Liew, M.W., Arnold, J.G., Bosch, D.D., 2005. Problems and potential of autocalibrating a hydrologic model. Transactions of the ASAE 48, 1025–1040.

Van Liew, M.W., Arnold, J.G., Garbrecht, J.D., 2003. Hydrologic simulation on agricultural watersheds: Choosing between two models. Transactions of the ASAE 46, 1539.

Vilaysane, B., Takara, K., Luo, P., Akkharath, I., Duan, W., 2015. Hydrological Stream Flow Modelling for Calibration and Uncertainty Analysis Using SWAT Model in the Xedone River Basin, Lao PDR. Procedia Environmental Sciences 28, 380–390.

Williams, J.R., Jones, C.A., Dyke, Pt., 1984. A modeling approach to determining the relationship between erosion and soil productivity. Transactions of the ASAE 27, 129–144.

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

Yang, Y., Onishi, T., Hiramatsu, K., 2014. Improving the Performance of Temperature Index Snowmelt Model of SWAT by Using MODIS Land Surface Temperature Data. The Scientific World Journal 2014, 1–20.

Zhang, D., Chen, X., Yao, H., Lin, B., 2015. Improved calibration scheme of SWAT by separating wet and dry seasons. Ecological Modelling 301, 54–61.

Zhang, Q., Singh, V.P., Peng, J., Chen, Y.D., Li, J., 2012. Spatial–temporal changes of precipitation structure across the Pearl River basin, China. Journal of Hydrology 440–441, 113–122.

Zhang, X., Srinivasan, R., Arnold, J., Izaurralde, R.C., Bosch, D., 2011. Simultaneous calibration of surface flow and baseflow simulations: a revisit of the SWAT model calibration framework. Hydrol. Process. 25, 2313–2320.

CHAPTER 4. EFFECTIVE INTEGRATED AQUIFER VULNERABILITY ASSESSMENT: A CASE STUDY OF THE UPPER WHITE RIVER WATERSHED

4.1 Abstract

Agriculture can be a major cause of groundwater degradation due to movement of chemicals applied for agricultural production. High nitrate concentrations in groundwater are detected in some areas within Midwest states. In this study, nitrate contamination in groundwater was evaluated using the concept of integrated aquifer assessment (by combining an intrinsic aquifer vulnerability assessment and aquifer hazard assessment) in the Upper White River Watershed (UWRW) in Indiana. The intrinsic aquifer vulnerability map was created by DRASTIC, and aquifer hazard assessment was conducted using a distributed watershed model (Soil and Water Assessment Tool (SWAT)) and a machine learning technique (Geospatial-Artificial Neural Network (Geo-ANN)). Finally, integrated aquifer vulnerability assessment was conducted by combining intrinsic aquifer vulnerability assessment and aquifer hazard assessment. Approximately 81.0% of nitrate detections > 2 ppm were within "High" and "Very high" vulnerability classes (represent 5.8% of area) as predicted by incorporating DRASTIC, SWAT, and Geo-ANN. Moreover, 12.1% of the nitrate detections were within the "Moderate" vulnerability class (30.7% of area), and 6.9% of the nitrate detections were within the "Low" vulnerability class (50.7% of area). This approach did not result in observed nitrates in wells > 2 ppm within the "Very low" vulnerability class

(12.8% of area). The results indicate that integrated aquifer vulnerability assessment performed better than only DRASTIC or SWAT/Geo-ANN. Thus, overall assessment of aquifer vulnerability can be performed using the integrated aquifer vulnerability assessment technique provided in this study. Moreover, this approach is expected to be an efficient guide for managing groundwater resources for policy makers and groundwater-related researchers.

4.2 Introduction

Groundwater is an important water resource for many people in the world and is also a primary resource in Indiana with 60% of the state's drinking water coming from groundwater (Alley et al., 2002; Solley et al., 1998). Moreover, groundwater is a vital component of the local, regional and global environment with groundwater feeding ecosystems as well as providing baseflow in rivers (Morris et al., 2003; NSW Department of Land and Water Conservation, 1998). Anthropogenic activities affect the quantity and quality of water resources including groundwater which offers human populations a number of services such as water for drinking and irrigation (Winter et al., 1998). Contaminated groundwater has been identified in both urban and rural areas. Agriculture can cause groundwater degradation due to application of chemicals in agriculture and chemicals (e.g., fertilizers, pesticides, and herbicides) that are spread across wide agricultural areas. If chemicals used in agriculture are slightly soluble in water, they are less likely to result in groundwater contamination. However, ammonium which is a major fertilizer and manure which mainly consists of nitrogen and phosphorus is highly soluble and mobile in water, and nitrate concentrations generated by nitrification of

ammonium are widespread in both surface water and groundwater. High nitrate concentrations are detected in some areas of Midwest states (Puckett et al., 1994; Kellogg et al., 1994; Winter et al., 1998).

To alleviate the negative effects of agriculture on groundwater resources and to maintain clean groundwater, groundwater management has been facilitated by groundwater monitoring and modeling (Petheram et al., 2003; Nourani et al., 2015; Unland et al., 2015; Grimmeisen et al., 2016). Groundwater monitoring has a benefit to identify groundwater quality and quantity directly in real time, and groundwater monitoring data can help enhance the planning, sustainable development, and management of groundwater resources. However, groundwater monitoring is complicated and an expensive process. Therefore, many areas would be data sparse areas with regard to groundwater monitoring data (Alcalá et al., 2005).

Compared with groundwater monitoring, groundwater modeling is less complicated and less expensive. Also, groundwater modeling allows assessment of broad areas. Both groundwater monitoring and modeling should be mutual and complementary for efficient evaluation of groundwater quality and quantity in broad areas (i.e., regional scale or continental) if the two systems are utilized simultaneously.

Various hydrologic and water quality models have been developed based on conceptual, statistical, stochastic, analytic, physical and numerical models of surface water and groundwater systems, which provide a means for predicting surface water and

groundwater system responses to future conditions. Thus, each model has its own purpose and characteristics. A particular model is usually selected and used as a suitable tool depending on a research or project goal. However, various research efforts related to water management require interdisciplinary fields. For example, if a research effort or project is considered for the analysis, planning and management of a wide range of water resources and environmental problems related to surface water and groundwater, the model or models not only should address surface water but should also deal with groundwater. Thus, this often results in the demand for an integrated approach and coupled models for different systems (Kamp and Savenije, 2007; Koudstaal et al., 1992). Therefore, sometimes when a model cannot address a research or project problem, two or more models can be used, or a combined model can be utilized to enhance the physical representation of hydrologic process for better estimation. Many studies related to hydrologic and water quality have been conducted to determine efficient water management using a coupled model (Noori and Kalin, 2016; Chen and Wu, 2012; Maxwell et al., 2015).

For efficient groundwater resources management, integrated aquifer vulnerability assessments are required. Integrated aquifer vulnerability assessments are incorporated into a groundwater characterization and risk analysis with tiered approaches for aquifer intrinsic vulnerability (aquifer intrinsic properties) and aquifer hazard (pollutant transport properties) assessment. Thus, overlay and index GIS model (DRASTIC) provided in Chapter 2, Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998), and Geo-

spatial ANN (Geo-ANN) were used to evaluate both aquifer intrinsic properties and aquifer hazard potential.

Aquifer hazard is regarded as potential sources of aquifer contaminations transported from the land surface to aquifers (Brouyére et al., 2001). Output variables from SWAT are used as input variables of ANN to conduct aquifer hazard assessment by forecasting nitrate concentration in aquifers.

The machine learning technique, ANN, is a data-driven model based on data experienced in the real world phenomena of a specific system. In contrast to analytical or numerical models, data-driven models can be used to solve problems in the field of hydrology and water resources engineering where knowledge of the physical behavior of the system and data are limited (Solomatine and Ostfeld, 2008). Thus, for the simulation of complex systems, data-driven models are useful to define the patterns within the behavior of the system (Araghinejad, 2014). ANN is also called a black box model that does not require detailed knowledge of the internal functions of a system to identify the complex, dynamic and non-linear relationships from given patterns by input and output (Ha and Stenstrom, 2003). However, as a black box model, users have no control except providing input data and initial parameters such as learning rate, maximum number of training cycle, and target error (Ali et al., 2015).

Machine learning is a study of artificial intelligence for improving the computer's ability to learn without being explicitly programmed. The main goal of machine learning is to

find relationships between the system state variables and to replace time-consuming manual processes with automatic techniques that improve accuracy or efficiency by exploring and determining regularities in training data (Simon and Langley, 1995; Solomatine and Shrestha, 2009).

As time passes nonstationary trends of hydrology and water resources time series has been exhibited more frequently (Yu and Lin, 2015; Coulibaly and Baldwin, 2005). Thus, for accurate estimation of hydrology and water resources systems, techniques which can simulate the nonstationary patterns of variables of hydrology and water resources are required. Many studies have already proven that machine learning is a suitable technique in predicting nonstationary behavior of hydrology and water resources systems (Pulido Calvo et al., 2003; Nourani et al., 2009; Nourani et al., 2015). For those reasons, ANN was used to conduct aquifer hazard assessment using data related to aquifer hazard (i.e., nitrate leached from the soil profile and nitrate transported into main stream in the groundwater loading) by SWAT and the nitrate concentration data in wells which are utilized for the training and validation processes in ANN.

If ANN manipulates polygon and/or raster files, ANN would be more useful in hydrology and water quality modeling. Thus, in this study Graphic User Interface (GUI)-based Geospatial-ANN (Geo-ANN) was developed to be compatible with GIS/RS data format (e.g., polygon and raster).

The objectives of this study were: 1) to develop Geo-ANN which is compatible with GIS/RS data formats for flexible hydrology and water quality modeling, and 2) to evaluate integrated aquifer vulnerability by incorporating intrinsic aquifer vulnerability and aquifer hazard assessment.

4.3 Methodology

4.3.1 Study Area

The Upper White River Watershed (UWRW) (Latitude: 39°29'51"N, Longitude: 86°24'02"W) is located in central Indiana (Figure 4.1). The drainage area of the UWRW is 6,944 km^2, and the most dominant land use is agriculture (3,160 km^2). The UWRW is important for public drinking water supply because the UWRW includes more than 3,508 km of streams, numerous artificial lakes, and 4 reservoirs. The UWRW serves as a drinking water supply for part of the city of Indianapolis which is Indiana's largest city. The water sources in the rural areas of UWRW traditionally are individual wells to provide groundwater for residential, commercial, and industrial purposes (Tedesco et al., 2011; Tedesco, 2005; Fleming et al., 1993). The UWRW was selected to identify hydrologic, water quality and aquifer risk assessment because the UWRW has available streamflow (U.S. Geological Survey (USGS) streamflow stations), water quality (Environmental Protection Agency (EPA) fixed stations and USGS monitoring well stations) data for both streamflow and groundwater (Figure 4.1 and Table 4.1).

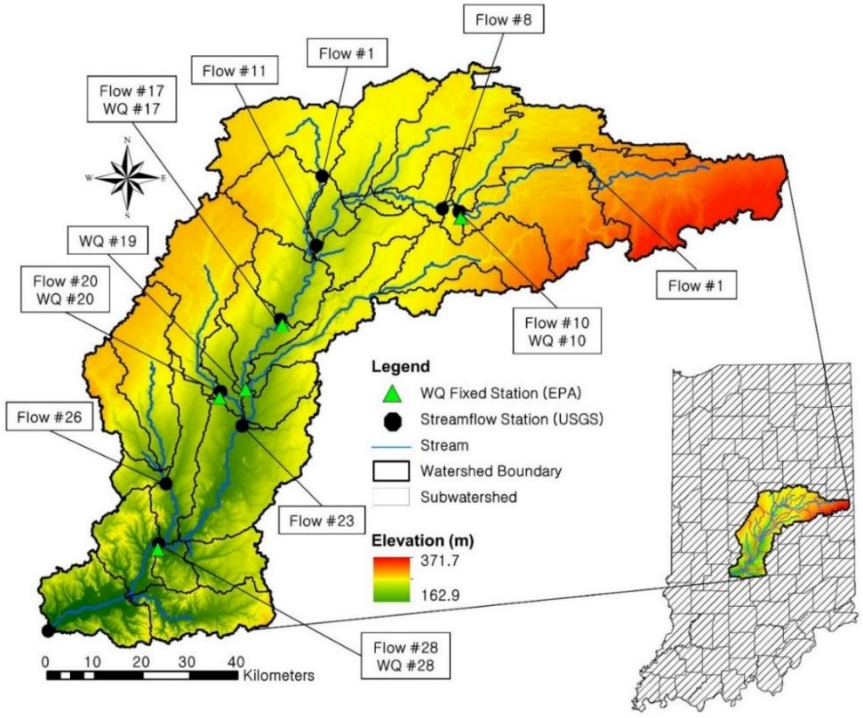


Figure 4.1 Location of the Upper White River Watershed, Indiana

Table 4.1 Monitoring stations for streamflow and water quality data in the UWRW

ID	Subbasin #	Monitoring station	Drainage area (km^2)	Type
Flow[1] #1	1	USGS 03349510	329	Streamflow
Flow #2	2	USGS 03347000	595	Streamflow
Flow #8	8	USGS 03348130	1411	Streamflow
Flow #10	10	USGS 03348000	997	Streamflow
Flow #11	11	USGS 03349000	2237	Streamflow
Flow #17	17	USGS 03351000	3201	Streamflow
Flow #20	20	USGS 03353500	448	Streamflow
Flow #23	23	USGS 03353611	4756	Streamflow
Flow #26	26	USGS 03353800	544	Streamflow
Flow #28	28	USGS 03354000	6227	Streamflow
WQ[2] #10	10	INSTOR WQX[3]-2398	997	Nitrate
WQ #17	17	INSTOR WQX-2434	3201	Nitrate
WQ #19	19	INSTOR WQX-2408	690	Nitrate
WQ #20	20	INSTOR WQX-2371	448	Nitrate
WQ #28	28	USGS 03354000	6227	Nitrate

[1]Flow: Streamflow data

[2]WQ: Water quality data

[3]INSTOR WQX: Indiana STORET (STOrage and RETrieval) data warehouse by EPA

4.3.2 Integrated Aquifer Vulnerability Assessment

4.3.2.1 Overview of Integrated Aquifer Vulnerability Assessment

For overall aquifer vulnerability assessment, intrinsic and hazard potential assessment should be carried out together. In this study, overall aquifer vulnerability assessment is called integrated aquifer vulnerability assessment which includes intrinsic vulnerability and hazard potential assessment. Integrated aquifer vulnerability assessment would be an effective method for estimation of groundwater contamination for drinking water with regard to aquifer intrinsic and transport properties (Brouyére et al., 2001). Aquifer intrinsic and transport properties are analyzed by aquifer vulnerability assessment and aquifer hazard assessment, respectively. Aquifer vulnerability (intrinsic properties) is defined as natural susceptibility to contamination based on the properties of the land, and subsurface and aquifer hazard (transport properties) is regarded as pollution potential with respect to nitrate transporting from the land surface to aquifers (Brouyére et al., 2001) (Figure 4.2). Aquifer vulnerability assessment is usually conducted by overlay and index GIS model (DRASTIC), and aquifer hazard assessment is implemented by numerical models (Pacheco and Sanches Fernandes, 2013; Chen et al., 2013; Akhavan et al., 2011; Akhavan et al., 2010). In this study, intrinsic aquifer vulnerability was conducted with DRASTIC, and aquifer hazard assessment was implemented using a distributed watershed model (SWAT) and machine learning technique (Geo-ANN).

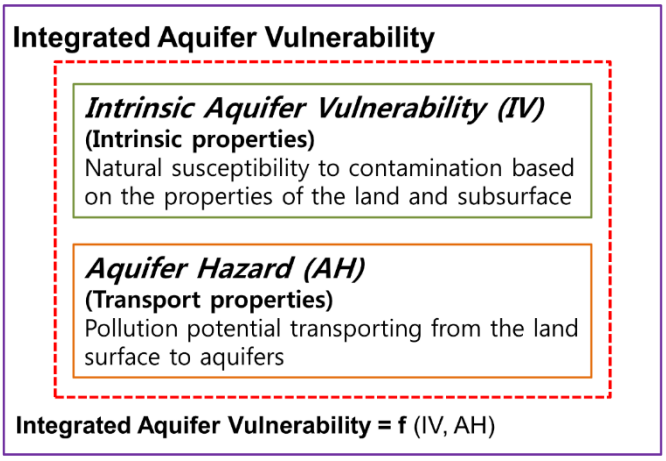


Figure 4.2 Concept of integrated aquifer vulnerability

4.3.2.2 Intrinsic Aquifer Vulnerability Assessment

4.3.2.2.1 Overview of DRASTIC

For intrinsic aquifer vulnerability assessment, DRASTIC, overlay and index GIS model, was used in this chapter. DRASTIC is a conceptual model defined as a composite description of the most important hydrogeological factors with regard to natural susceptibility to aquifer contamination based on the properties of the land and subsurface (Brouyére et al., 2001). DRASTIC yields a numerical index map that is derived from ratings and weights assigned to the seven map parameters (Aller et al. 1987; Akhavan et al. 2011). DRASTIC is a numerical ranking system, which uses weights, ranges, and ratings to provide groundwater vulnerability. The DRASTIC index is calculated using Equation 4.1. The higher the DRASTIC index score, the greater the groundwater vulnerability. In this chapter, DRASTIC is not described in detail because DRASTIC was already described in chapter 2.

$$DRASTIC = \sum_{i=1}^{n1} D_{wi} D_{ri} + \sum_{i=1}^{n2} R_{wi} R_{ri} + \sum_{i=1}^{n3} A_{wi} A_{ri} + \sum_{i=1}^{n4} S_{wi} S_{ri} + \sum_{i=1}^{n5} T_{wi} T_{ri} + \sum_{i=1}^{n6} I_{wi} I_{ri} + \sum_{i=1}^{n7} C_{wi} C_{ri} \qquad (4.1)$$

Where

D_r = Ratings to the depth to water table

D_w = Weight assigned to the depth to water table

R_r = Ratings for ranges of aquifer recharge

R_w = Weight for aquifer recharge

A_r = Ratings assigned to aquifer media

A_w = Weight assigned to aquifer media

S_r = Ratings for soil media

S_w = Weight for soil media

T_r = Ratings for topography

T_w = Weight assigned to topography

I_r = Ratings assigned to vadose zone

I_w = Weight assigned to vadose zone

C_r = Ratings for rates of hydraulic conductivity

C_w = Weight given to hydraulic conductivity

4.3.2.2.2 Intrinsic aquifer vulnerability mapping using DRASTIC

Detailed methods for creating intrinsic aquifer vulnerability mapping for the UWRW using DRASTIC are described in the methods of chapter 2. With hydrogeology and weather data used for creating DRASTIC input data shown in Table 4.2, the seven map layers (Depth to water, Recharge, Aquifer media, Soil media, Topography, Impact of

vadose zone media, and Hydraulic conductivity), representing the seven variables of DRASTIC, were prepared to create the intrinsic aquifer vulnerability map for the UWRW. After the intrinsic aquifer vulnerability map was created by Equation 4.1, model calibration was conducted using a binary classifier calibration with a genetic algorithm (Bi-GA) and manual calibration to adjust DRASTIC weights. A more detailed description used to create the intrinsic aquifer vulnerability map is provided in the methodology part of chapter 2.

Table 4.2 Data used for creating DRASTIC inputs for the UWRW

Data type	Source	Scale	Date	Used to produce
Water well	IDNR[1]	1:24,000	1959 - 2010	**D**epth to water
Annual precipitation	NCDC[2]	-	1949 - 2010	**R**echarge
LULC	MRLC[3]	1:250,000	2006	**R**echarge
Aquifer systems	USGS[4]	1:48,000	2003 - 2010	**A**quifer media
SSURGO[5]	NRCS[6]	1:12,000	2005	**R**echarge **S**oil media **T**opography
iLITH[7]	IGS[8]	1:24,000	2001	**I**mpact of vadose
Aquifer transmissivity	IDNR[1]	1:24,000	2011	**C**onductivity

[1]IDNR: Indiana Department of Natural Resources

[2]NCDC: National Climate Data Center

[3]MRLC: Multi-Resolution Land Characteristics Consortium

[4]USGS: U.S. Geological Survey

[5]SSURGO: Soil Survey Geographic Database

[6]NRCS: Natural Resources Conservation Service

[7]iLITH: Lithology data for Indiana wells edited by Indiana Geological Survey

[8]IGS: Indiana Geological Survey

4.3.2.3 Configuration of Aquifer Hazard Assessment

4.3.2.3.1 Overview of Hydrologic and Water Quality Model (SWAT)

SWAT is a physically based distributed, deterministic, and long-term continuous time model which is used to predict impact of management practices, Land Use and Land Cover (LULC) change, and climate change on hydrology and water quality on a watershed scale at a daily time step (Arnold et al., 1998; Neitsch et al., 2011). SWAT represents the large scale spatial heterogeneity of the study area by dividing a watershed into subbasins. The subbasin is the first level of subdivision of the watershed. Subbasins possess a geographic position in the watershed and are spatially related to one another. The land area in a subbasin may be divided into hydrologic response units (HRUs) which are the smallest computational units in SWAT. HRUs are parts of a subbasin that possess unique land use, management, and soil attributes. HRUs are created by one or more unique land use and soil combinations for each subbasin. Surface runoff, soil water content, crop growth, nutrient cycles, and erosion are simulated for each HRU, and then HRUs are combined and calculated for the subbasin by a weighted value (Neitsch et al., 2011; Williams et al., 1984). In this study, variables related to nitrate movement were extracted for each HRU, and HRU maps were created as input data for aquifer risk assessment using Geo-ANN.

4.3.2.3.2 SWAT Input Data

The period of SWAT simulation is from 1990 to 2010 because the SWAT format climate data (i.e., precipitation and temperature) were available in this period provided by

National Climate Data Center (NCDC) and processed by the Agricultural Research Service (ARS). Also, there are sufficient water quality data from 1990 to 2010 for the UWRW. As described in Table 4.3, the primary input data for the SWAT simulation are Digital Elevation Model (DEM), soil, LULC, and weather data. Additionally, for nitrate simulation, the scheduled management operation data were prepared to consider application of fertilizer, pesticide, and manure (Table 4.4). A general strategy of management practices for corn-soybean rotation is described in Table 4.4 (Her et al., 2016). DEM from U.S. Geological Survey (USGS) was used for watershed delineation, and soil and LULC data from Natural Resources Conservation Service (NRCS) and Natural Agricultural Statistics Service (NASS) were used with respect to the hydrological response in the watershed. Weather data include minimum and maximum temperature, daily precipitation, mean wind speed, relative humidity, and solar radiation which were obtained from National Climate Data Center (NCDC) (Table 4.3).

The spatial DEM (Figure 4.3(a)) and soil maps in the UWRW (Figure 4.3(b)) were obtained from USGS. Elevation of the UWRW varies from 162.9 to 371.7 m. Twenty eight soil types are distributed in the study area such as 33.5% IN013 (Crosby), 24% IN040 (Miami), 6.3% IN029 (Sawmill), 6% IN026 (Fox), 5.2% IN054 (Miamian) and 24.9% other soil types. LULC in the UWRW (Figure 4.3(c)) includes 45.5% agricultural fields (22.5% soybeans, 22.2% corn, and 0.8% others). The mean annual precipitation is 1,093 mm, and the highest, the lowest daily temperature, and mean daily temperatures are 36.1°C, -31.3°C, and 10.8°C, respectively. Figure 4.3(d) shows well locations in the UWRW. 678 well nitrate data were utilized for the training, validation, and testing of

Geo-ANN using SWAT outputs (Appendix A). The well nitrate data were obtained from Heidelberg University and the Water Quality Portal (WQP) sponsored by the U.S. Geological Survey (USGS), the U.S. Environmental Protection Agency (EPA), and the National Water Quality Monitoring Council (NWQMC). Min/max/mean values of well nitrate data are 0.1/18.29/1.22 mg/L, respectively.

With the input data (Table 4.3), hydrology (surface water + groundwater hydrology) and water quality were simulated in the UWRW by SWAT. Then, observed streamflow and estimated baseflow (or observed baseflow if available) from the USGS were used to calibrate and validate SWAT.

Tile drainage was applied in areas where the land use is corn or soybean and the soil drainage condition is poorly drained (Boles et al., 2015; Jiang et al., 2014; Sui and Frankenberger, 2008). Based on the previous studies for Indiana watersheds (Boles et al., 2015; Jiang et al., 2014; Green et al., 2006) and the technical report about tile drainage (USDA-NRCS, 2011), parameters related to the tile drainage in SWAT were used as described in Table 4.5.

Annual average (2008 - 2015) mass of point source pollutants related to NO3 showed less than 1 ton/year from the National Pollutant Discharge Elimination System (NPDES) in the UWRW. Observed NO3 loads (estimated NO3 loads using LOADEST) were approximately 2,333 ton/year. The small amount of point source pollutants (0.04% of observed NO3 loads) would not significantly affect results of this study. Further, point

source data years were not matched for the periods of this study (1990 - 2010). Thus, point sources were not considered in the UWRW.

Table 4.3 SWAT input data for hydrologic and water quality modeling

Data	Source	Format	Date
DEM	USGS[1]	Raster	2010
Soil	NRCS[2]	Polygon	2010
LULC[3]	NASS[4]	Raster	2010
Weather	NCDC[5]	Tabular data	1986 - 2010
Streamflow	USGS	Tabular data	1986 - 2010
Baseflow	WHAT[6]	Tabular data	1986 - 2010
Stream nitrate	USGS	Tabular data	1991 - 2010

[1]USGS: U.S. Geological Survey

[2]NRCS: Natural Resources Conservation Service

[3]LULC: Land Use and Land Cover

[4]NASS: Natural Agricultural Statistics Service

[5]NCDC: National Climate Data Center

[6]WHAT: Web-based Hydrograph Analysis Tool

Table 4.4 Management practices for corn-soybean rotation in SWAT

Year	Date	Management	ID
Corn Year	Apr-22	Fertilizer application	Anhydrous ammonia[1]
	Apr-22	Pesticide application	Atrazine[2]
	May-6	Tillage	Field cultivator
	May-6	Planting	Corn
	Jun-6	Fertilizer application	Urea[3]
	Oct-14	Harvest	-
	Oct-15	Kill / end of growing season	-
Soybean Year	May-24	Tillage	No-till
	May-24	Planting	Soybean
	Oct-7	Harvest	-
	Oct-8	Kill / end of growing season	-
	Oct-15	Fertilizer application	P_2O_5[4]
	Nov-1	Tillage	Chisel plow

[1]Anhydrous ammonia: 53 kg/ha (N of 43 kg/ha)

[2]Atrazine: 2.2 kg/ha

[3]Urea: 284 kg/ha (N of 131 kg/ha)

[4]P_2O_5: 123 kg/ha (P of 54 kg/ha)

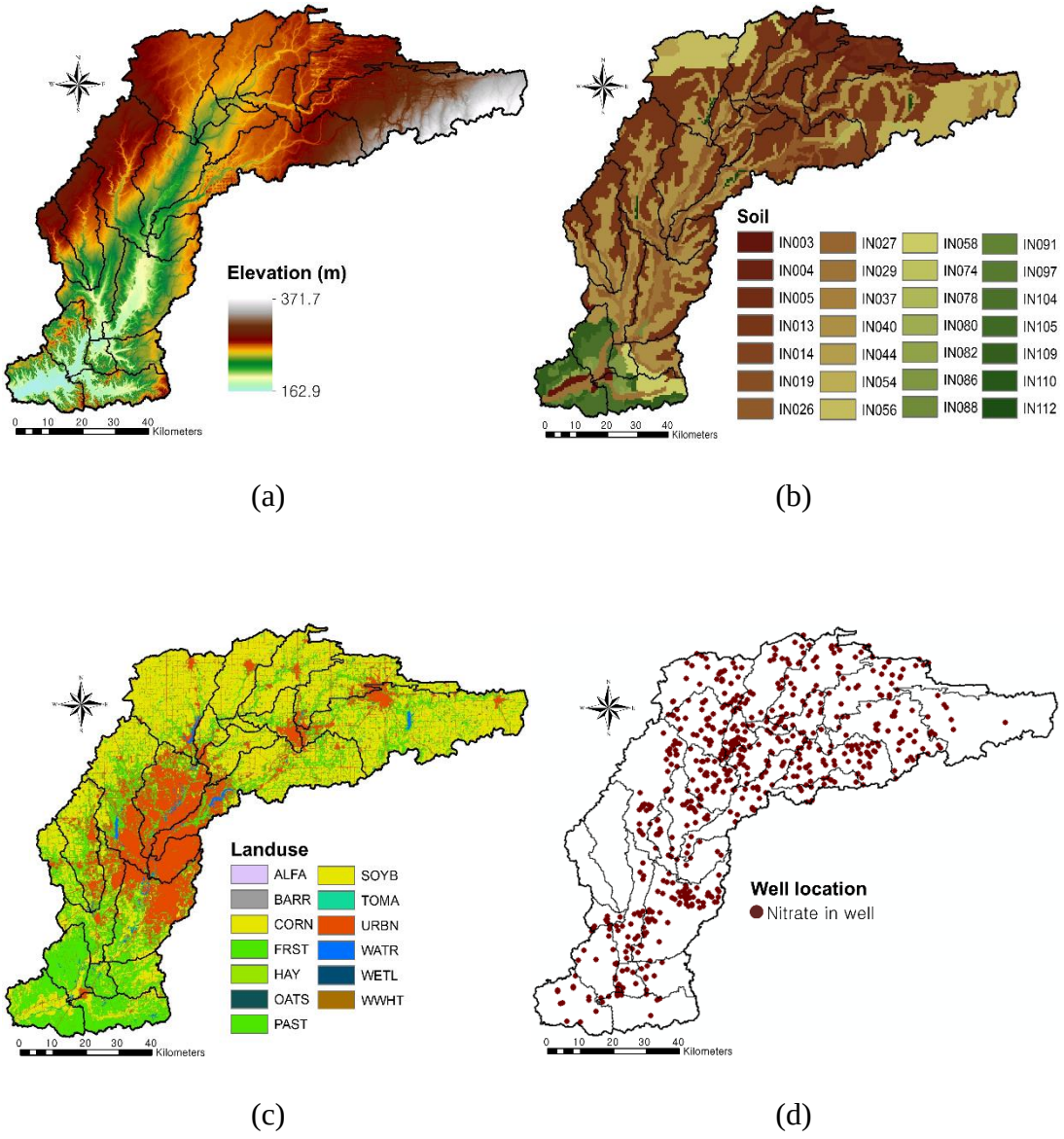


Figure 4.3 DEM (a), soil (b), land use (c), and sampled well locations (d) of the UWRW

Table 4.5 Tile drainage parameters in SWAT

Parameter	Description	Value
DDRAIN	Depth to drains (mm)	1000
G_DRAIN	Drain tile lag time (h)	48
DRAIN_CO	Drainage coefficient (mm/d)	10
SDRAIN	Tile spacing (mm)	20000
LATKSATF	Multiplication factor to determine Ksat	1.2
RE	Effective radius of drains (mm)	20
ITDRN	Tile drainage equations flag/code	1 (new routine)

4.3.2.3.3 Hydrologic and Water Quality Modeling using SWAT

The land phase of the hydrologic cycle controls the amount of water, sediment, nutrient and pesticide loadings to the main channel and to the aquifers in each subbasin. The hydrologic cycle simulated in SWAT is based on the water balance equation (Equation 4.2). Also, nitrate leaching from the soil profile to the aquifer and nitrate loading to streamflow contributed by baseflow are estimated in SWAT (Figure 4.4) (Neitsch et al., 2009).

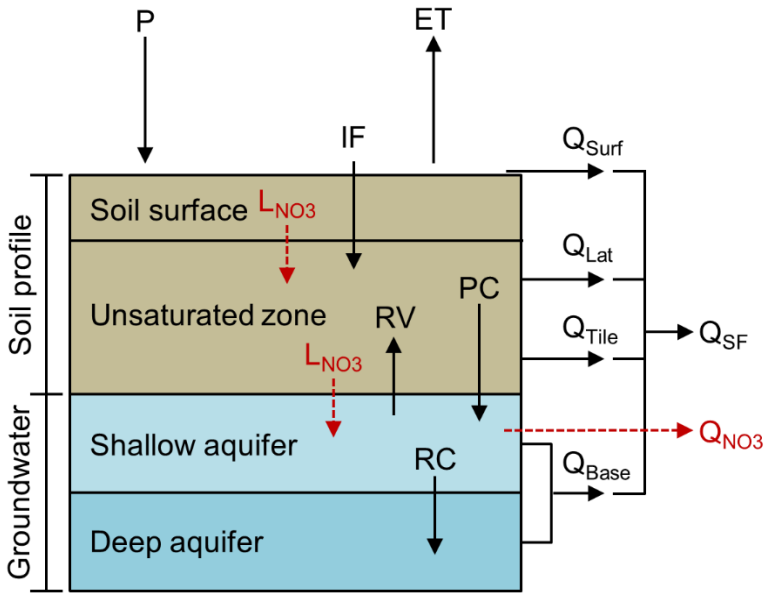


Figure 4.4 Schematic diagram of the hydrologic cycle and nitrate transport

Note: P is the precipitation, ET is the evapotranspiration, IF is the infiltration, RV is the water moving from the shallow aquifer into the overlaying unsaturated zone, PC is the percolation, RC is the groundwater recharge, Q_{Surf} is the surface runoff, Q_{Lat} is the lateral flow, Q_{Tile} is the drainage from tiles, Q_{Base} is the baseflow contribution to streamflow, Q_{SF} is the streamflow, L_{NO3} is the leachate nitrate, and Q_{NO3} is the nitrate loading into the streamflow

$$SW_t = SW_0 + \sum_{i=1}^{t} (R_{day,i} - Q_{surf,i} - ET_{a,i} - Q_{tile,i} - Q_{lat,i} - Q_{gw,i}) \quad (4.2)$$

where SW_t is the final soil water content (mm) at time *t*, *t* is the time (day), SW_0 is the initial soil water content (mm H_2O), $R_{day,i}$ is the amount of precipitation on day *i* (mm H_2O), $Q_{surf,i}$ is the amount of surface runoff on day *i* (mm H_2O), $ET_{a,i}$ is the amount of evapotranspiration on day *i* (mm H_2O), $Q_{tile,i}$ is the amount of water removed from the layer by tile drainage on day *i*, $Q_{lat,i}$ is the amount of lateral flow released to the main channel on day *i* (mm H_2O), and $Q_{gw,i}$ is the amount of return flow on day *i* (mm H_2O)

Surface runoff ($Q_{surf,i}$) was estimated using the Soil Conservation Service (SCS) curve number (CN) method (Equation 4.3) (SCS, 1972).

$$Q_{surf,i} = \frac{(R_{day,i} - 0.2S)^2}{R_{day,i} + 0.8S} \tag{4.3}$$

$$I_a = 0.2S$$

$$S = \frac{25,400}{CN} - 254$$

Where I_a is the initial abstraction (mm H_2O), S is the retention parameter (mm H_2O), and CN is the curve number (dimensionless)

The Penman-Monteith method was used to calculate the rate of evapotranspiration ($ET_{a,i}$) (Equation 4.4) (Monteith, 1965).

$$ET_{a,i} = \frac{\Delta \cdot (H_{net} - G) + \rho_{air} \cdot c_p \cdot (e_z^0 - e_z) / r_a}{\Delta + \gamma \cdot (1 + r_c / r_a) \cdot \lambda} \tag{4.4}$$

where Δ is the slope of the saturation vapor pressure-temperature curve (*de*/*dT*, kPa °C), H_{net} is the net radiation (MJ $m^{-2}d^{-1}$), G is the heat flux density to the ground (MJ $m^{-2}d^{-1}$), ρ_{air} is the air density (kg m^{-3}), c_p is the specific heat at constant pressure (MJ kg^{-1} $°C^{-1}$), e_z^0 is the saturation

vapor pressure of air height z (kPa), e_z is the water vapor pressure of air at height z (kPa), γ is the psychrometric constant (kPa °C^{-1}), r_c is the plant canopy resistance (s m^{-1}), r_a is the diffusion resistance of the air layer (aerodynamic resistance) (s m^{-1}), and λ is the volumetric latent heat of vaporization (Energy required per water volume vaporized) (MJ m^{-3})

Lateral flow was calculated with Equation 4.5 (Neitsch et al., 2009).

$$Q_{lat,i} = 0.0224 \cdot \left(\frac{2 \cdot SW_{ly,excess} \cdot K_{sat} \cdot slp}{\phi_d \cdot L_{hill}} \right) \quad (4.5)$$

where $SW_{ly,excess}$ is the drainage volume of water stored in the saturated zone of the hillslope per unit area (mm H_2O), K_{sat} is the saturated hydraulic conductivity (mm/h), slp is the average slope of the subbasin (m/m), ϕ_d is the drainage porosity of soil (mm/mm), and L_{hill} is the hillslope length (m)

Both shallow aquifer and deep aquifers are considered as groundwater storage. Total baseflow (groundwater flow) was calculated by total amount of water in the shallow aquifer and deep aquifer (Equation 4.6 - 4.11) (Neitsch et al., 2009).

$$Q_{gw,i} = Q_{gw,sh,i} + Q_{gw,dp,i} \quad (4.6)$$

where $Q_{gw,i}$ is the total groundwater flow on day *i* (mm H_2O), $Q_{gw,sh,i}$ is the total groundwater flow in the shallow aquifer on day *i* (mm H_2O), and $Q_{gw,dp,i}$ is total groundwater flow in the deep aquifer on day *i* (mm H_2O)

$$Q_{gw,sh,i} = Q_{gw,sh,i-1} \cdot \exp(-\alpha_{gw,sh} \cdot \Delta t) + w_{rchrg,sh,i} \cdot [1 - \exp(-\alpha_{gw,sh} \cdot \Delta t)] \tag{4.7}$$

where $Q_{gw,sh,i-1}$ is the groundwater flow from the shallow aquifer on day i-1 (mm H_2O), $-\alpha_{gw,sh}$ is the baseflow recession constant, Δt is the time step (day), and $w_{rchrg,sh,i}$ is the amount of recharge entering the shallow aquifer on a day i (mm H_2O)

$$Q_{gw,dp,i} = Q_{gw,dp,i-1} \cdot \exp(-\alpha_{gw,dp} \cdot \Delta t) + w_{rchrg,dp,i} \cdot [1 - \exp(-\alpha_{gw,dp} \cdot \Delta t)] \tag{4.8}$$

where $Q_{gw,dp,i-1}$ is the groundwater flow from the deep aquifer on day i-1 (mm H_2O), $-\alpha_{gw,dp}$ is the baseflow recession constant, Δt is the time step (day), and $w_{rchrg,dp,i}$ is the amount of recharge entering the deep aquifer on a day i (mm H_2O)

$$w_{rchrg,i} = (1 - \exp(-1/\delta_{gw}) \cdot w_{seep,i} + \exp(-1/\delta_{gw}) \cdot w_{rchrg,i-1} \tag{4.9}$$

$$w_{rchrg,dp,i} = \beta_{dp} \cdot w_{rchrg,i} \tag{4.10}$$

$$w_{rchrg,sh,i} = w_{rchrg,i} - w_{rchrg,dp,i} \tag{4.11}$$

where $w_{rchrg,i}$ is the amount of recharge entering the both aquifers on day i (mm H_2O), δ_{gw} is the drainage time of the overlaying geologic formations (days), $w_{seep,i}$ is the total amount of water exiting the bottom of the soil profile on day i (mm H_2O), $w_{rchrg,i-1}$ is the amount of recharge

entering the aquifers on day *i*-1 (mm H_2O), β_{dp} is the aquifer percolation coefficient (dimensionless)

Nitrate in the shallow aquifer may remain in the aquifer, move with recharge to the deep aquifer, move with groundwater flow into the main channel, or be transported out of the shallow aquifer with water moving in the soil zone in response to water deficiencies. Nitrate in recharge to the shallow aquifer and deep aquifer on a given day was calculated with Equations 4.12 and 4.13 (Neitsch et al., 2009).

$$NO3_{rchrg,i} = (1 - \exp[-1/\delta_{gw}]) \cdot NO3_{perc} + \exp[-1/\delta_{gw}] \cdot NO3_{rchrg,i-1} \quad (4.12)$$

where $NO3_{rchrg,i}$ is the amount of nitrate in recharge entering the aquifers on day *i* (kg N/ha), δ_{gw} is the delay time (drainage time) of the overlying geologic formation (day), $NO3_{perc}$ is the total amount of nitrate exiting the bottom of the soil profile on day *i* (kg N/ha), $NO3_{rchrg,i-1}$ is the amount of nitrate in recharge entering the aquifers on day *i*-1 (mm H_2O)

$$NO3_{gw,i} = (NO3_{sh,i-1} + NO3_{rchrg,i}) \cdot Q_{gw,i} / (aq_{sh,i} + Q_{gw,i} + w_{revap,i} + w_{rchrg,dp,i}) \quad (4.13)$$

where $NO3_{gw,i}$ is the amount of nitrate in groundwater flow from the shallow aquifer on day *i* (kg N/ha), $NO3_{sh,i-1}$ is the amount of nitrate in the shallow aquifer at the end of day *i*-1 (kg N/ha), $Q_{gw,i}$ is the groundwater flow into the main channel on day *i* (mm H_2O), $aq_{sh,i}$ is the amount of

water stored in the shallow aquifer at the end of day *i* (mm H_2O), $w_{revap,i}$ is the amount of water moving into the soil zone in response to water deficiencies on day *i* (mm H_2O), $w_{rchrg,dp,i}$ is the amount of recharge entering the deep aquifer on day *i* (mm H_2O).

4.3.2.3.4 Streamflow, baseflow, and nitrate calibration/validation

Streamflow and baseflow were calibrated and validated simultaneously based on USGS streamflow and simulated baseflow by modified SWAT 2012 code and the Web GIS-based Hydrograph Analysis Tool (WHAT) system (Lim et al., 2005) which can separate baseflow from streamflow. For calibration and validation of nitrate loads in streamflow, daily or monthly nitrate loads are necessary. However, nitrate concentration data were available only for a few days each year. Thus, LOADESTimator (LOADEST) (Runkel et al., 2004) was used to estimate mean monthly nitrate loads. Using estimated nitrate loads by LOADEST, calibration and validation of nitrate loads at the main outlet were conducted. For accurate estimation of nitrate concentrations in the aquifer, calibration and validation of nitrate loads in streamflow are necessary because nitrate transport to streamflow and nitrate leaching to the aquifer are interdependent (Phillips et al., 1999).

Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) (Abbaspour et al., 2004; Abbaspour, 2011) was used to perform sensitivity and uncertainty analysis as well as calibration and validation (Grusson et al., 2015; Yang et al., 2008). For evaluation of the model performance, Nash-Sutcliff Efficiency (NSE) (Equation 4.14), coefficient of determination (R^2) (Equation 4.15), and Percent Bias (PBIAS) (Equation 4.16) were

selected as the objective function for calibration and validation of streamflow, baseflow, and nitrate loads (Nash and Sutcliffe, 1970; Krause et al., 2005; Gupta et al., 1999). The total simulation period was from 1990 - 2010 (21 years) with the first 4 years as the model warm up period. Calibration for streamflow and baseflow (streamflow calibration: 10 outlets and baseflow calibration: 1 main outlet) was implemented with the parameters related to water balance, subsurface water, surface runoff, physical properties of soil, and physical properties of channel. Calibration for nitrate loads (nitrate loads calibration: 5 outlets) was conducted with the parameters associated with the nitrogen cycle (Table 4.6). Calibration parameter ranges for streamflow, baseflow, and nitrate loads (Table 4.6) were defined based on the results of sensitivity analysis and previous studies (Arnold et al., 2012; Zhang et al., 2011; Du et al., 2006; Lam et al., 2010; Yeo et al., 2014). After model calibration, validation for streamflow, baseflow, and nitrate loads were performed with the calibrated parameters. Model performance ratings for streamflow, baseflow, and nitrate loads were evaluated with three quantitative statistics shown in Table 4.7 (Moriasi at al., 2007; Van Liew et al., 2003; Singh et al., 2004; Engel et al., 2007).

$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs} - Y^{mean})^2} \right] \tag{4.14}$$

where Y_i^{obs} is the i^{th} observed data, Y_i^{sim} is the i^{th} simulated data, Y^{mean} is the mean of observed data, and n is the total number of observed data

$$R^2 = \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - \bar{Y}^{obs})(Y_i^{sim} - \bar{Y}^{sim})^2}{\sqrt{\sum_{i=1}^{n} (Y_i^{obs} - \bar{Y}^{obs})^2} \sqrt{\sum_{i=1}^{n} (Y_i^{sim} - \bar{Y}^{sim})^2}} \right]^2 \quad (4.15)$$

where Y_i^{obs} is the i^{th} observed data, $\bar{Y}^{obs}$ is the mean of observed data, Y_i^{sim} is the i^{th} simulated data, $\bar{Y}^{sim}$ is the mean of simulated data, and n is the total number of observed data

$$PBIAS(\%) = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim}) \times 100}{\sum_{i=1}^{n} (Y_i^{obs})} \right] \quad (4.16)$$

where Y_i^{obs} is the i^{th} observed data, Y_i^{sim} is the i^{th} simulated data, and n is the total number of observed data.

Table 4.6 SWAT parameters for calibration of streamflow, baseflow, and nitrate loads

Parameter	Description	Unit	Ranges	
			LB[1]	UB[2]
Water balance				
ESCO	Soil evaporation compensation factor	-	0.01	1
SFTMP	Snowfall temperature	°C	-5	5
SMTMP	Snow melt base temperature	°C	-5	5
TIMP	Snow pack temperature lag factor	-	0.01	1
SMFMX	Melt factor for snow on June 21	mm/°C	0.01	10
SMFMN	Melt factor for snow on December 21	mm/°C	0.01	10
Subsurface water				
GW_REVAP	Groundwater evaporation coefficient	-	0.02	0.2
REVAPMN	Depth of water for evaporation	mm	0.01	250
GWQMN	Depth of water for return flow	mm	0.01	500
GW_DELAY	Groundwater delay time	day	0.1	20
ALPHA_BF	Baseflow alpha factor	1/day	0.1	1
RCHRG_DP	Deep aquifer percolation fraction	-	0.01	1
Surface runoff				
CN2	Initial SCS runoff curve number	-	-0.25	0.25
SURLAG	Surface runoff lag coefficient	-	0.1	10
Physical properties of the soil				
SOL_AWC	Available water capacity of soil layer	mm/mm	-0.25	0.25
Physical properties of the channel				
CH_K2	Effective hydraulic conductivity	mm/hr	5	300
Nitrogen cycle				
ANION_EXCL	Fraction of porosity[3]	-	0	1
SDNCO	Denitrification threshold water content	-	0	1
CDN	Denitrification exponential rate	-	0	3
NPERCO	Nitrogen percolation coefficient	-	0	1
BIOMIX	Biological mixing efficiency	-	0	1
N_UPDIS	Denitrification coefficient	-	1	50
AI1	Fraction of algal biomass (nitrogen)	mg/mg	0.07	0.09
RCN	Concentration of nitrogen in rainfall	mg/L	0	10
SHALLST_N	Initial concentration of nitrate in SA[4]	mg/L	0	50
SOL_ORGN	Initial organic N concentration in the SL[5]	mg/kg	0	1500

[1]LB: Lower bound

[2]UB: Upper bound

[3]ANION_EXCL: Fraction of porosity from which anion are excluded

[4]SA: Shallow aquifer

[5]SL: Soil layer

Table 4.7 SWAT performance evaluation criteria for NSE, R^2, and PBIAS

	Output	Performance Evaluation Criteria			
		Very Good	**Good**	**Satisfactory**	**Unsatisfactory**
NSE	Flow	> 0.80	0.70 < N ≤ 0.80	0.50 < N ≤ 0.70	N ≤ 0.50
	Nitrate	> 0.65	0.50 < N ≤ 0.65	0.30 < N ≤ 0.50	N ≤ 0.30
R^2	Flow	> 0.85	0.75 < R^2 ≤ 0.85	0.60 < R^2 ≤ 0.75	R^2 ≤ 0.60
	Nitrate	> 0.70	0.60 < R^2 ≤ 0.70	0.35 < R^2 ≤ 0.60	R^2 ≤ 0.35
PBIAS (%)	Flow	< ±10	±10 ≤ P < ±15	±15 ≤ P < ±25	P ≥ ±25
	Nitrate	< ±25	±25 ≤ P < ±40	±40 ≤ P < ±70	P ≥ ±70

Adapted by Van Liew et al. (2003), Singh et al. (2004), and Moriasi et al. (2015)

4.3.2.3.5 Retrieval of variables of aquifer hazard using SWAT

As mentioned earlier, two variables in Equations 4.12 and 4.13 are related to nitrate contamination in aquifers in SWAT. The first variable (NO3L, $NO3_{rchrg,i}$) is the nitrate leached from the soil profile, and the second variable (NO3GW, $NO3_{gw,i}$) is the nitrate transported into the main stream from the groundwater loading. Those two variables were selected from the SWAT outputs to estimate aquifer hazard. The two variables were retrieved at HRU levels by an HRU extractor developed for this study. Then, spatial input variable maps for Geo-ANN were created to identify potential aquifer hazard areas.

4.3.2.4 Aquifer Hazard Assessment

4.3.2.4.1 Development of Geo-ANN

ANN is composed of a network architecture, activation function, and learning rule. The network architecture has a number of neurons and layers of neurons. Sigmoid,

hyperbolic tangent, and Gaussian functions have been widely used as the activation function. Basic learning rules are supervised, unsupervised, and reinforcement learning.

ANN can be categorized based on the direction of information flow and processing. In a feed forward network, the neurons (nodes) are generally arranged in layers, starting from a first input layer and ending at the final output layer. Information passes from the input to the output side. A synaptic weight is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input and output relationship (ASCE Task Committee, 2000). ANN requires a large number of examples of input and output for training and validation. The primary goal of training is to minimize the error function by searching for a set of connection strengths and threshold values defined by users. Then, ANN can predict outputs that are equal or close to target error (ASCE Task Committee, 2000). The following Equations 4.17-4.22 provide a brief description of the neural network operation.

Forward calculation:

$$\mathrm{U}_j = \sum (X_i \cdot w_{ij}) \tag{4.17}$$

where U_j is the internal value of the neural network operation, j is the every neuron in a layer, X_i is the each of the *i* input, and w_{ij} is the weight associated with input i of neuron

$$AF = \frac{1-e^{-2\lambda x}}{1+e^{2\lambda x}}$$

(4.18)

where AF is the activation function, and λ is the gain parameter

$$\mathrm{Y}_j = AF(U_j + thr_j) \tag{4.19}$$

where Y_j is the output of neuron j in the current layer, and thr_j is the threshold value in terms of neuron j

Backpropagation calculation:

$$\mathrm{e}_j = Y_j \cdot (1 - Y_j) \cdot (T_j - Y_j) \tag{4.20}$$

where e_j is the error signal for the output layer, Y_j is the actual output value, and T_j is the scaled target value

$$\mathrm{e}_j = Y_j \cdot (1 - Y_j) \cdot \sum (e_k \cdot w'_{jk}) \tag{4.21}$$

where e_k is the error signal for the hidden layer, and w'_{jk} is the prior weights of k-th neuron in the immediately succeeding layer

$$w_{ij} = w'_{jk} + (1 - M) \cdot LR \cdot e_j \cdot X_j + M \cdot (w'_{ij} - w''_{ij}) \tag{4.22}$$

where M is the momentum parameter, LR is the learning rate parameter w'_{ij} is the previous weight value, and w''_{ij} is the next previous weight value

Geo-ANN was developed to train, validate, and predict geospatial data as well as tabular data using Artificial Neural Networks (ANN). Geo-ANN would be suitable for studies of hydrology and water quality modeling because those studies usually use geospatial data as an input or output to conduct spatial analysis and prediction. Geo-ANN provides a user-friendly graphical user interface (GUI) for optimizing the parameters of the ANN architecture and training the network for the prediction. The shapefile and tabular format

can be directly used as input without converting to a new format and can be mapped with predicted values by ANN.

Various metrics for water resources and environmental management studies such as NSE, R^2, and PBIAS are implemented to evaluate the performance of the user-configured ANN models. Cross validation techniques are included in the Geo-ANN such as k-fold and leave-one-out cross validation (LOOCV). Cross validation is an effective validation method when the amount of calibration and validation data are limited (Shao and Er, 2016; Wong, 2015). Because there are limited data for groundwater hydrology and water quality, cross validation would be a supportive technique in groundwater hydrology studies.

The optimal neural network parameters were designed based on the nitrate observation data used in this study. Because the nitrate observation data are limited in most cases, training and validation have to proceed with limited data. However, validation using limited data often fails to accurately estimate the performance of the designed prediction model, causing overfitting or underfitting problems. Thus, cross validation was used, which partitions a sample data into different subsets of training, validation, and testing data. A k-fold cross validation was used as the cross validation technique which randomly divides the k-number of subset of sample data in each run (Heaton, 2008). In a rule of the thumb, the division ratios of training and validation are set as 70% and 30%, individually (Guyon, 1997). Once the sample data are split into these three sets, the

network performance is evaluated. The performance of the Geo-ANN is evaluated using the three metrics NSE, R^2, and PBIAS.

For the network design, the size of hidden layers and number of neurons (nodes) are one of the most important factors. No unified theory exists for determining such an optimal ANN architecture but, in a rule of thumb, it is known to set the number of neurons similar to the number of inputs and outputs. Before assigning the input and output data into the Geo-ANN, all data are normalized to fall in [0.1 0.9] using Equation 4.23 because the normalization improves accuracy, performance, and speed of Geo-ANN (Kalin et al., 2010; Sethi et al., 2010). In this study, the number of inputs and output are 2 and 1 and the number of neurons ranges from 2 to 4. For determining the number of layers, 1 or 2 is/are a typical number in a small or moderate size of the network for this study. The whole application including training/validation algorithms and GUI for Geo-ANN was developed using neural network toolbox in MATLAB. Many MATLAB users mainly use MATLAB version 2013, 2014, 2015, and 2016. Therefore, compatibility test for Geo-ANN was conducted and successfully completed.

$$x_{norm} = 0.1 - \left(\frac{x_0 - x_{\min}}{x_{\max} - x_{\min}} \right) \tag{4.23}$$

where x_{norm} is the normalized value, x_0 is the observed value, $x_{\min}$ is the minimum value, and $x_{\max}$ is the maximum value

The functionalities of the Geo-ANN are summarized as follows:

(1) Design of the ANN layers: number of hidden layers and number of neurons.

(2) Selection of training (optimization) algorithm: Levenberg-Marquardt, gradient descent, or Bayesian regularization (Beale et al., 2016).

(3) Selection of normalization methods (pre-conditioning): normalization of the user-defined range (e.g. [0.1,0.9], [0,1], or [-1, 1]).

(4) Functions for importing and exporting a raw shapefile or tabular format for training, validation, and prediction.

(5) Various metrics for performance measurements for hydrology and water quality modeling: NSE, R^2, and PBIAS (Moriasi et al., 2007).

(6) Two cross-validation techniques: LOOCV and k-fold cross validation (Heaton, 2008).

(7) Stopping criteria design: maximum number of validation increases, minimum performance value, and maximum number of training epochs (iterations).

(8) Enable GPU and/or parallel computing capability for large ANN training.

4.3.2.4.2 Analysis of Aquifer Hazard Using Geo- ANN

For aquifer hazard assessment (Figure 4.5), two predefined input variables (NO3L and NO3GW) were retrieved from the SWAT simulation. Maps for nitrate leached from the soil profile and nitrate in groundwater were obtained through surface and groundwater hydrology and water quality simulation by SWAT.

NO3L and NO3GW were utilized as input data to Geo-ANN that was developed in this study. Nitrate concentration data from 678 monitoring wells (Appendix A) (Figure 4.6) were used for training/validation/testing of Geo-ANN. 562 data points were less than 2 ppm and 116 data points were greater than 2 ppm. Nitrate levels over 2 ppm were assumed to be caused by human activities because nitrate levels in aquifers under natural conditions are typically less than 2 ppm in Indiana (Navulur, 1996). Thus, a threshold value of background concentration was set at 2 ppm in this study. Integrated vulnerability assessment was conducted using nitrate detections > 2 ppm as elevated N levels. In order to fill the data gap and validate spatially distributed nitrate concentrations, k-fold cross validation was used for training and validation of Geo-ANN. Many studies revealed that k-fold cross validation is a reliable method when the number of training/validation data are small (Wong, 2015; Shao and Er, 2016). Kalin et al. (2010) recommended model performance criteria for ANN with two metrics (i.e., NSE and PBIAS) in watershed modeling at a monthly time scale (Noori et al., 2016; Kalin et al., 2010). Model performance criteria for Geo-ANN were modified by adding one more metric (i.e., coefficient of determination, R^2) based on Moriasi et al. (2015) and Kalin et al. (2010). As shown in Table 4.8, Geo-ANN performance was evaluated using the three metrics (i.e., NSE, R^2, and PBIAS) (Equations 4.14 - 4.16).

Table 4.8 Geo-ANN performance evaluation criteria for NSE, R^2, and PBIAS

	Output	Performance Evaluation Criteria			
		Very Good	Good	Satisfactory	Unsatisfactory
NSE	Flow WQ*	> 0.70	0.50 < N ≤ 0.70	0.30 < N ≤ 0.50	N ≤ 0.30
R^2	Flow WQ	> 0.75	0.55 < R^2 ≤ 0.75	0.35 < R^2 ≤ 0.55	R^2 ≤ 0.35
PBIAS (%)	Flow WQ	< ±25	±25 ≤ P < ±50	±50 ≤ P < ±70	P ≥ ±70

Adapted by Kalin et al. (2010) and Moriasi et al. (2015)

*WQ: Water quality

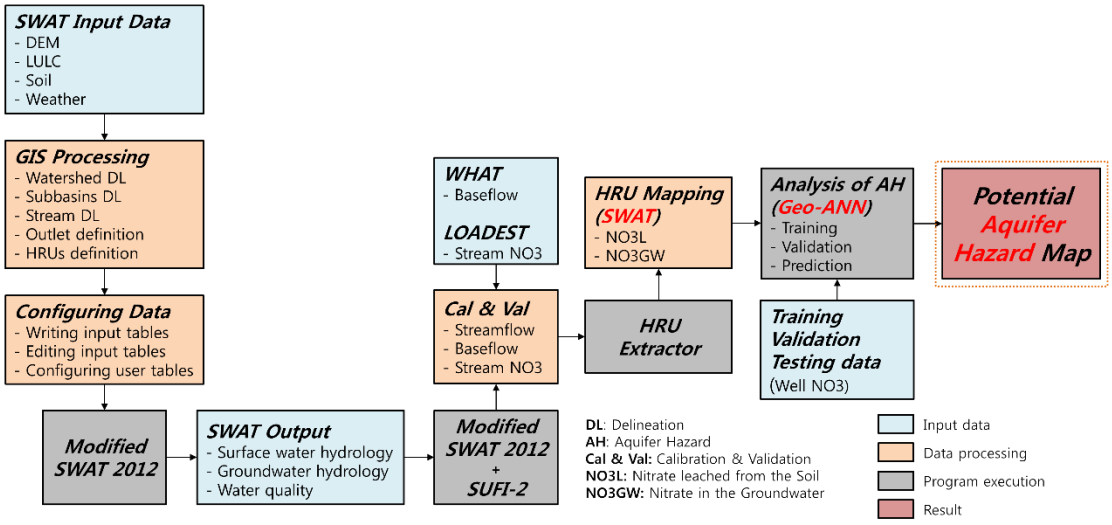


Figure 4.5 Flowchart of the aquifer hazard assessment procedure

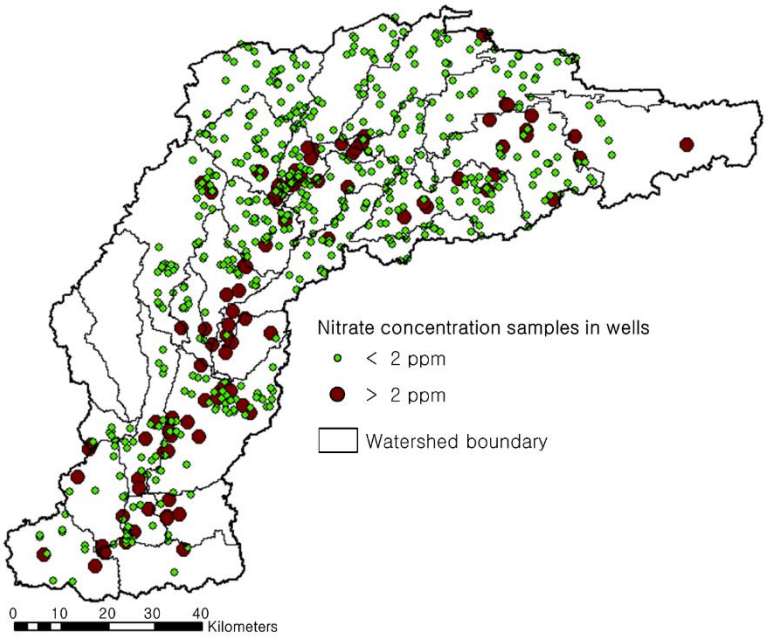


Figure 4.6 Nitrate concentration samples in wells in the UWRW

4.3.2.5 Analysis of Integrated Aquifer Vulnerability

As mentioned above, integrated aquifer vulnerability map was created by combining the intrinsic aquifer vulnerability map using DRASTIC and aquifer hazard map using SWAT and Geo-ANN. Each map was standardized by feature scaling (Equation 4.24) and both maps have the same range of values (min = 0 and max = 1) which means the intrinsic aquifer vulnerability map and aquifer hazard map have the same weight. This study assumed that two maps have the same impact on aquifer vulnerability.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{4.24}$$

where x' is the normalized value and x is the original value.

4.4 Results and Discussion

4.4.1 Calibration and Validation of Streamflow, Baseflow, and Nitrate Loads

For the simulation period of 21 years (1990 - 2010), monthly mean streamflow was 85.7 m^3/s, and monthly mean baseflow was 35.9 m^3/s, according to the WHAT system (Lim et al., 2005). Monthly mean baseflow accounted for 41.9 percent of monthly mean streamflow.

The comparison between the LOADEST-estimated and USGS observed nitrate loads indicate NSE of 0.84, R^2 of 0.89, and PBIAS of 0.92. These values indicate that LOADEST estimated nitrate loads quite well.

Model calibration and validation of streamflow, baseflow, and nitrate loads at the main outlet were conducted by adjusting twenty six parameters (Table 4.9). As shown in Table 4.9, calibrated values were estimated with simultaneous streamflow and baseflow calibration and nitrate loads calibration with multi-site flow and water quality stations.

Table 4.9 Final values of the SWAT calibration parameters for each scenario in the UWRW

Parameter	Initial value	Calibrated value	
	Default	Final range	Final value
ESCO	0.95	0.3 - 0.9	0.35
SFTMP	1	-5 - 5	2.31
SMTMP	0.5	-5 - 5	-0.77
TIMP	1	0 - 5	0.93
SMFMX	4.5	0 - 10	2.5
SMFMN	4.5	0 - 10	2.5
GW_REVAP	0.02	0.02 - 0.2	0.04
REVAPMN	750	0 - 250	210
GWQMN	1000	0 - 500	52.5
GW_DELAY	31	0 - 50	21
ALPHA_BF	0.048	0.1 - 1	0.7
RCHRG_DP	0.05	0.01 - 1	0.4
CN2[2]	va[1]	-25 - 25	-23
SURLAG	4	1 - 10	7.42
SOL_AWC[2]	va[1]	-0.2 - 0.2	-0.15
CH_K2	0	5 - 300	34.8
ANION_EXCL	0.5	0.01 - 1	0.15
SDNCO	1.1	0 - 1	0.32
CDN	1.4	0 - 3	0.3
NPERCO	0.2	0 - 1	0.19
BIOMIX	0.2	0.01 - 1	0.69
N_UPDIS	20	1 - 50	21
AI1	0.08	0.07 - 0.09	0.07
RCN	0	0 - 10	1.9
SHALLST_N	0	1 - 50	10.9
SOL_ORGN	0	500 - 1500	1200

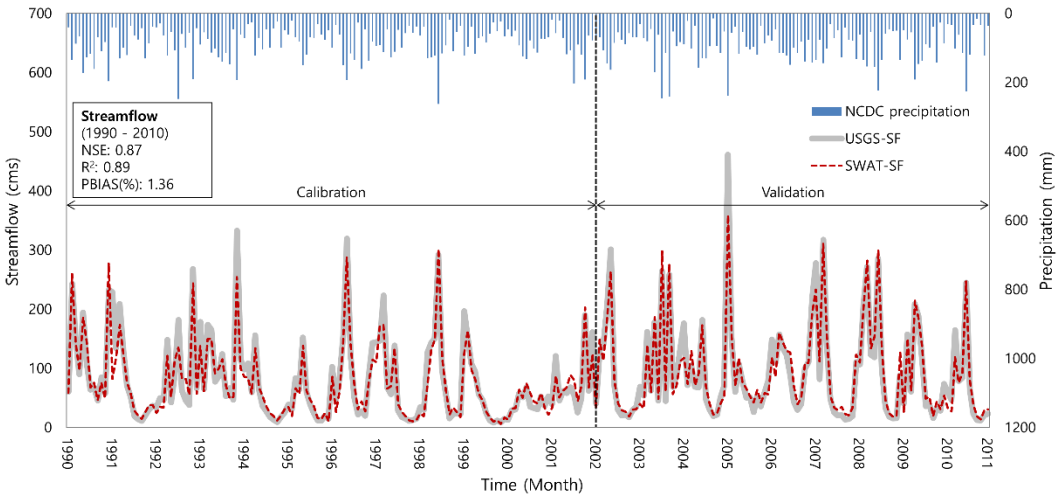
[1]va: Different value according to HRUs

[2]CN2, SOL_AWC: Percentage change (%)

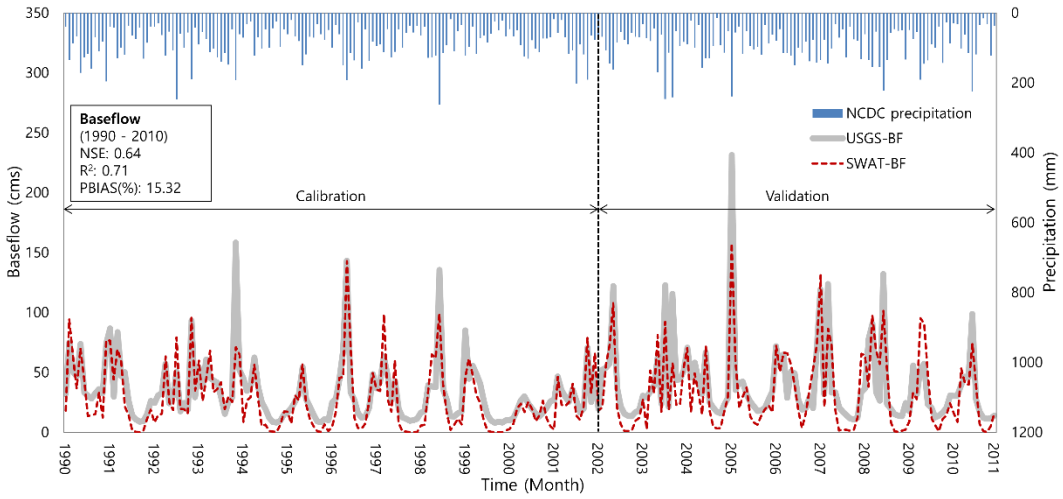
Hydrographs reproduced by SWAT between the observed and simulated monthly streamflow and baseflow at the main outlet (flow #28 and WQ #28) during the calibration and validation periods are shown in Figure 4.7(a) and (b). As shown in Figure 4.7(a), even though some peak and low streamflow values were underestimated, the calibrated model performed well for monthly streamflow simulation. The calibrated model underestimated during peak streamflow periods, and the model slightly overestimated

during some low streamflow periods. Streamflow during the validation period was better estimated than during the calibration period. Figure 4.7(b) shows results of monthly baseflow simulation during the calibration and validation periods. Baseflow simulated with the calibrated model was estimated satisfactorily. Thus, the result shows that the calibrated model simulates both streamflow and baseflow well. It is noted that simultaneous streamflow and baseflow calibration are necessary for robust estimation of hydrological parameters.

Based on the calibrated parameters from the main outlet of the UWRW, cross validation for streamflow was conducted at other 9 additional USGS streamflow stations. Even though some peak flows were underestimated, all simulations replicated the observed streamflow well. The calibrated parameters explained hydrology characteristics for the entire watershed (Figure 4.8).

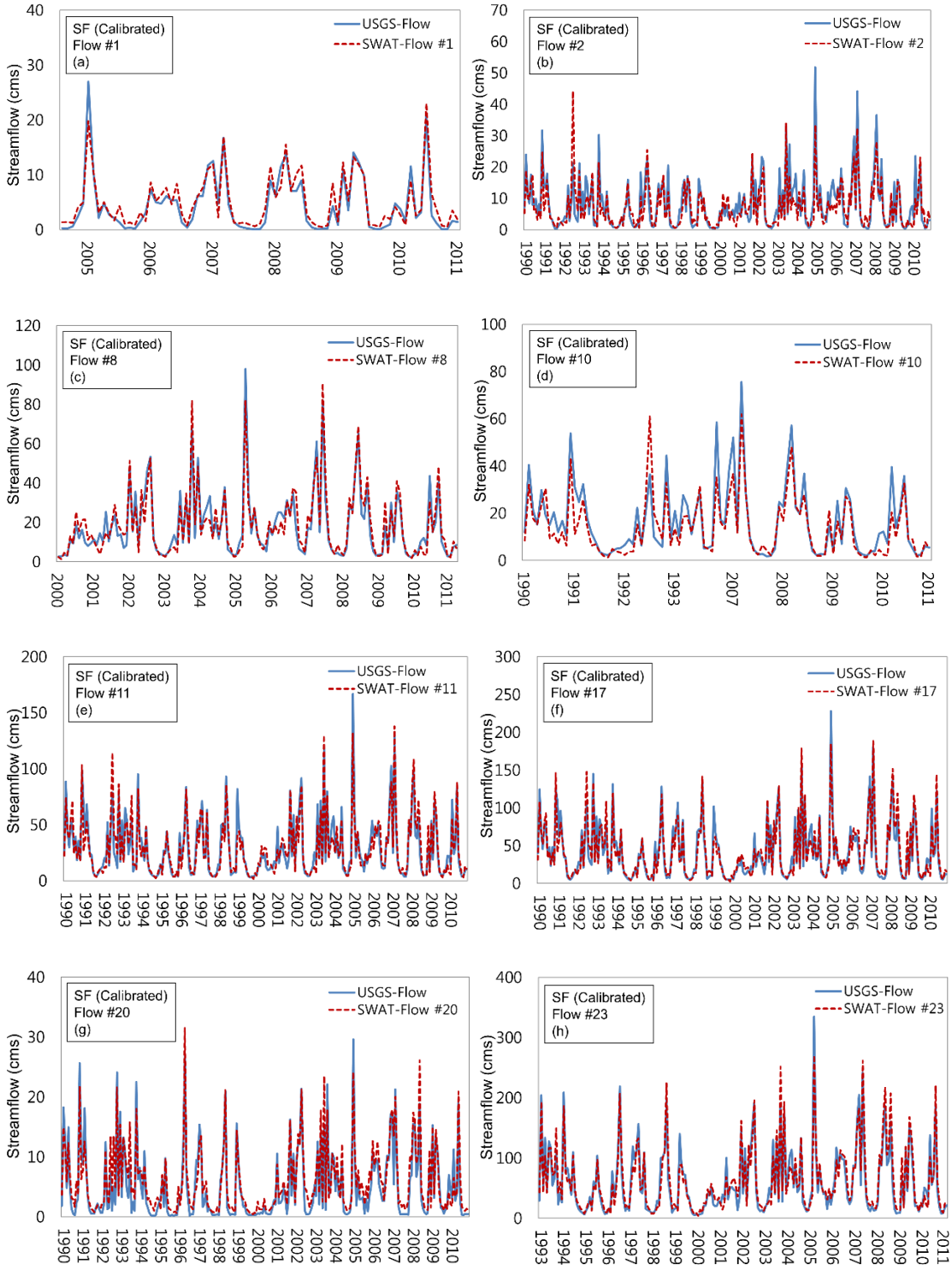


(a)



(b)

Figure 4.7 Comparison of monthly flow between USGS observed and SWAT calibrated flow at the main outlet of the UWRW (1990 - 2010): (a) Streamflow and (b) Baseflow


SF (Calibrated)
Flow #1
(a)
USGS-Flow
SWAT-Flow #1
Streamflow (cms)
SF (Calibrated)
Flow #2
(b)
USGS-Flow
SWAT-Flow #2
Streamflow (cms)
SF (Calibrated)
Flow #8
(c)
USGS-Flow
SWAT-Flow #8
Streamflow (cms)
SF (Calibrated)
Flow #10
(d)
USGS-Flow
SWAT-Flow #10
Streamflow (cms)
SF (Calibrated)
Flow #11
(e)
USGS-Flow
SWAT-Flow #11
Streamflow (cms)
SF (Calibrated)
Flow #17
(f)
USGS-Flow
SWAT-Flow #17
Streamflow (cms)
SF (Calibrated)
Flow #20
(g)
USGS-Flow
SWAT-Flow #20
Streamflow (cms)
SF (Calibrated)
Flow #23
(h)
USGS-Flow
SWAT-Flow #23
Streamflow (cms)

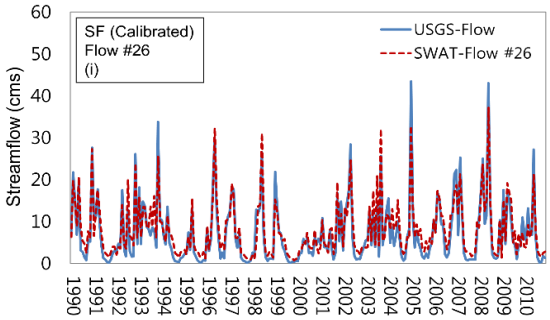


Figure 4.8 Comparison of monthly USGS observed and SWAT calibrated streamflow at the 9 USGS streamflow stations in the UWRW

After hydrology calibration and validation using 10 stations, calibration and validation for nitrate loads were conducted. Figure 4.9 shows that even though the calibrated model simulated nitrate loads reasonably well, the model greatly underestimated nitrate loads for some points. At WQ #10 for 1993, WQ #17 for 1998, and WQ #28 for 1992 and 1993, there was underestimation because timing of fertilizer application was unknown and simulated streamflow was underestimated by SWAT for those periods. The other reason for underestimated nitrate loads could be the uncertainties of nitrate monitoring data and LOADEST.

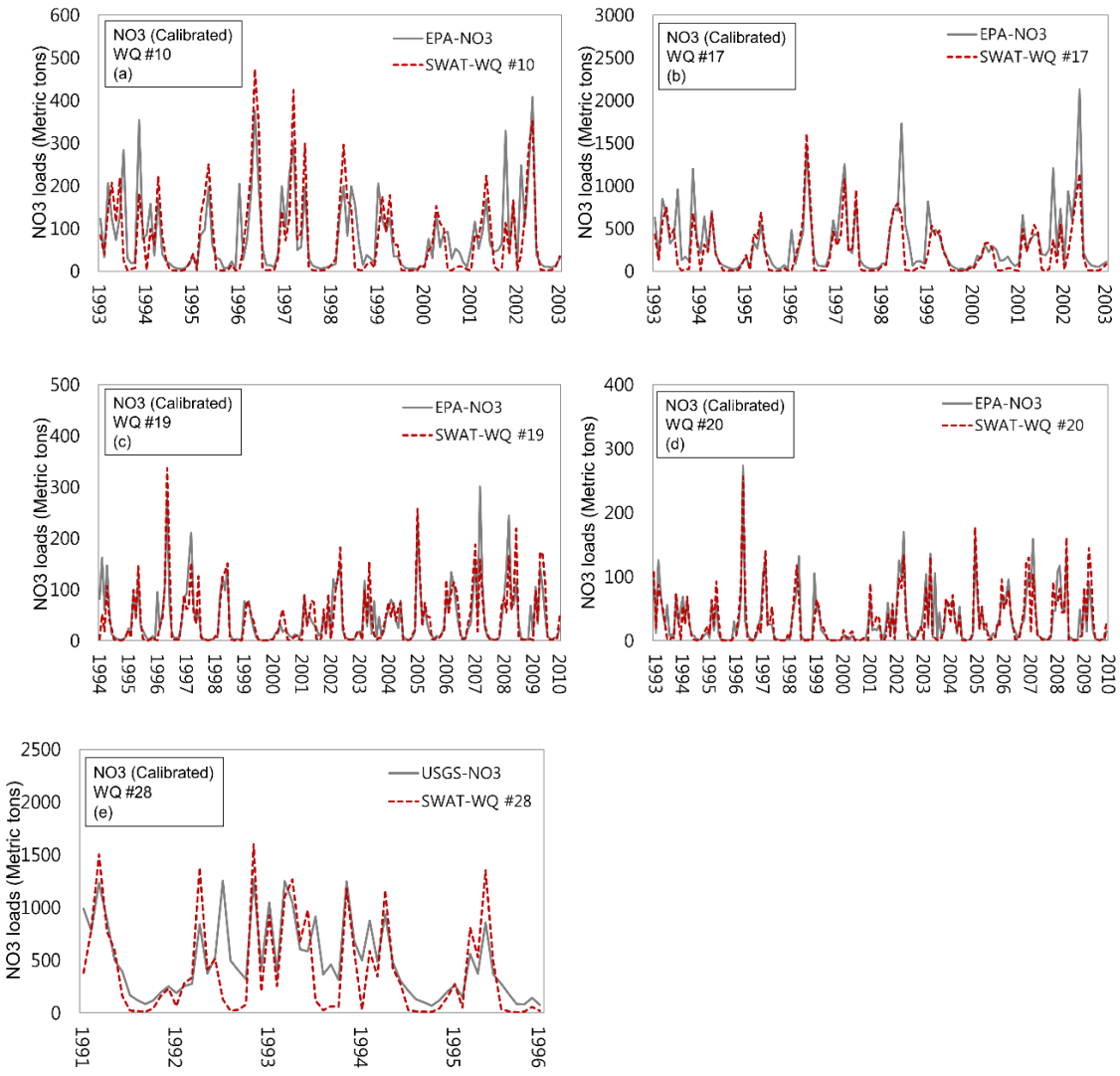


Figure 4.9 Comparison of monthly observed and SWAT calibrated NO3 loads at the 5 water quality stations in the UWRW

Model performance for streamflow, baseflow, and nitrate load simulation at the main outlet was evaluated by NSE, R^2, and PBIAS (Table 4.10). For monthly streamflow calibration and validation at the main outlet, NSE values range from 0.85 to 0.88. R^2 values vary from 0.87 to 0.92, and PBIAS values range from 1.36 to 3.90. Based on Table 4.7, all simulation periods for streamflow (total period, calibration, and validation)

are within the “Very good” range, indicating all simulation periods are acceptable. For the monthly baseflow calibration and validation, NSE values (total period, calibration, and validation) vary from 0.63 to 0.65. R^2 values range from 0.70 to 0.73, and PBIAS values vary from 13.8 to 16.7. The results of the baseflow calibration and validation show “Very good”, “Good”, and “Satisfactory” ranges which mean all simulation periods for baseflow (total period, calibration, and validation) are acceptable. The results of simultaneous streamflow and baseflow calibration indicate good performance for baseflow calibration as well as streamflow calibration. Even though the calibrated model provided good model performances for both streamflow and baseflow simulation, streamflow and baseflow in the model were still slightly underestimated or overestimated. For monthly nitrate load calibration and validation at the main outlet, NSE values of all simulation periods range from 0.51 to 0.72. R^2 values vary from 0.58 to 0.77 and PBIAS values ranges from 13.9 to 20.1. Based on Table 4.8, all simulation periods are within the “Very good”, “Good”, and “Satisfactory” ranges. The calibrated model is acceptable for all simulation periods of streamflow, baseflow, and nitrate loads.

As shown in Table 4.10, simulated streamflow using calibrated parameters at the main outlet replicated observed streamflow at the 9 USGS streamflow stations well. Based on Table 4.7, all simulated streamflow at the 9 USGS streamflow stations are within the “Very good”, “Good” and “Satisfactory” ranges which mean all 9 flow simulations are acceptable. NSE values vary from 0.75 to 0.91 and 90% of NSE are “Very good”. R^2 values range from 0.77 to 0.92 and 90 % of R^2 are also “Very good”. PBIAS values vary from -16.4 to 15.9 and 50% of PBIAS are “Very good” (Table 4.11).

For accurate nitrate load estimation, multi-site calibration was performed to satisfy all nitrate load simulations at the 4 EPA fixed stations and 1 USGS water quality station. All nitrate load simulations are better than "Satisfactory". NSE values vary from 0.58 to 0.77 and 40% of NSE are "Very good". R^2 values range from 0.60 to 0.79 and 40% of R^2 are also "Very good". PBIAS values vary from 3.5 to 29.7 and 80% of PBIAS are "Very good" (Table 4.12). Even though all simulations for nitrate loads are within acceptable ranges based on Table 4.7, most PBIAS values indicate most simulated nitrate were underestimated.

Table 4.10 Model performance for streamflow at the main outlet (Flow #28) in the UWRW

	Streamflow			**Baseflow**		
	NSE	**R^2**	**PBIAS**	**NSE**	**R^2**	**PBIAS**
Total period (1990 - 2010)	0.87	0.89	1.36	0.64	0.71	15.3
Calibration (1990 - 2001)	0.85	0.87	3.90	0.63	0.73	16.7
Validation (2002 - 2010)	0.88	0.92	1.50	0.65	0.70	13.8

Table 4.11 Model performance for streamflow at the 9 USGS streamflow stations in the UWRW

ID	Evaluation	NSE	R^2	PBIAS	Period
Flow #1	Calibration	0.89	0.92	-13.3	2004-2007
	Validation	0.89	0.91	-15.4	2008-2010
Flow #2	Calibration	0.67	0.69	4.2	1990-2001
	Validation	0.80	0.85	17.3	2002-2010
Flow #8	Calibration	0.87	0.88	0.5	1999-2006
	Validation	0.94	0.94	-8.6	2007-2010
Flow #10	Calibration	0.65	0.71	16.3	1990-1993
	Validation	0.89	0.94	15.5	2006-2010
Flow #11	Calibration	0.85	0.85	1.1	1990-2001
	Validation	0.92	0.92	1.4	2002-2010
Flow #17	Calibration	0.88	0.88	0.2	1990-2001
	Validation	0.93	0.93	-4.8	2002-2010
Flow #20	Calibration	0.87	0.89	-16.0	1990-2001
	Validation	0.85	0.88	-16.8	2002-2010
Flow #23	Calibration	0.91	0.91	-0.8	1992-2003
	Validation	0.92	0.92	-8.1	2004-2010
Flow #26	Calibration	0.87	0.89	-13.7	1990-2001
	Validation	0.88	0.90	-9.9	2002-2010

Table 4.12 Model performance for nitrate loads at the 5 water quality stations in the UWRW

ID	Evaluation	NSE	R^2	PBIAS	Period
WQ #10	Calibration	0.43	0.60	3.1	1993-1998
	Validation	0.56	0.60	15.7	1999-2002
WQ #17	Calibration	0.61	0.67	24.8	1993-1998
	Validation	0.57	0.71	35.3	1999-2002
WQ #19	Calibration	0.72	0.73	7.9	1994-2003
	Validation	0.7	0.73	-2.3	2004-2009
WQ #20	Calibration	0.77	0.81	10.3	1993-2003
	Validation	0.76	0.76	-1.6	2004-2009
WQ #28	Calibration	0.51	0.58	20.1	1991-1993
	Validation	0.72	0.77	13.9	1994-1995

4.4.2 Development and Application of Geo-ANN

The interface of Geo-ANN developed in this study is shown as Figure 4.10. The network performance was validated using different combinations of the number of neurons and hidden layers within the ranges (i.e., neuron ranges: 2 to 4 and layer ranges: 1 to 2). In all simulations, the Levenberg-Marquardt algorithm was used for training the network and a hyperbolic tangent sigmoid was implemented as the transfer function for hidden layers and output layer. The training process stops if the maximum number of incremental validation reaches more than 30. With two input variables (NO3L (kg/ha), NO3GW (kg/ha)) calculated by Equations 4.12-13 and observed well nitrate data, the two input variables and one observed variable were trained/validated/tested with the parameters and structures of the Geo-ANN to create the new model which can predict nitrate concentrations in wells. Then, three performance metrics were produced as shown in Table 4.13. Final outcomes of model performance are calculated using the median value of all outcomes by the number of k-fold cross validation. Among different combinations of the number of neurons and hidden layers, 2 hidden layers and 2 neurons produced the optimal solutions of NSE, R^2, and PBIAS during the testing. NSE/R^2/PBIAS for the testing shows 0.66/0.70/0.07 (each value is median out of 1000 simulation results). According to Table 4.8, all model performance for the testing of the Geo-ANN indicates better than “Satisfactory”. Nitrate prediction for HRUs which do not have observed nitrate data was conducted using the structure of 2 hidden layers and 2 neurons for each hidden layer (HL-NR-NR: 2-2-2).

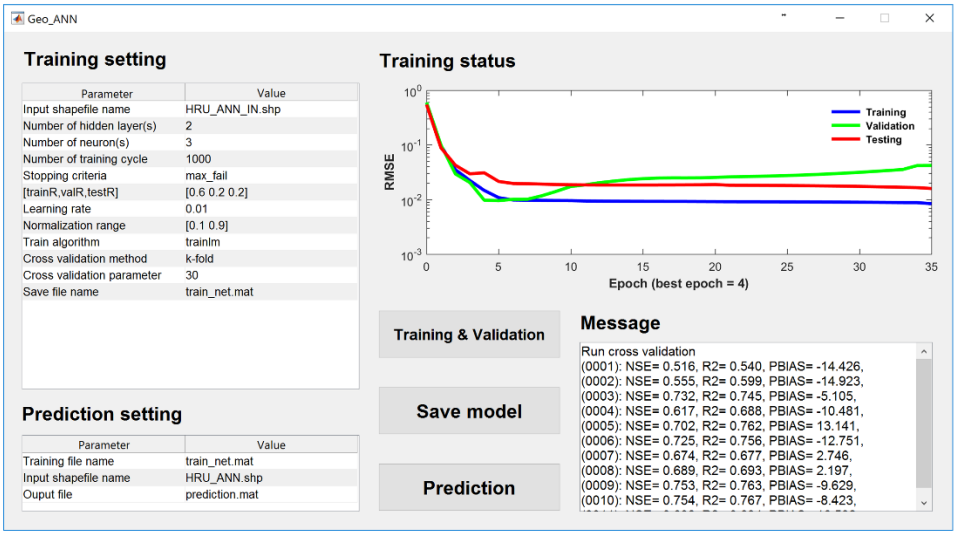


Figure 4.10 The interface of Geo-ANN developed in this study

Table 4.13 Model performance evaluation for the testing of the Geo-ANN in the UWRW

	Testing		
HL-NR-NR[1]	**NSE**	**R^2**	**PBIAS**
1-2-2	0.65	0.68	-0.03
1-3-3	0.63	0.67	0.13
1-4-4	0.62	0.66	0.50
2-2-2	0.66	0.70	0.07
2-3-3	0.63	0.67	-0.87
2-4-4	0.61	0.66	-0.29

[1]HL-NR-NR: # of hidden layer and neurons

4.4.3 Analysis of Integrated Aquifer Vulnerability

According to Pohlert et al. (2007), De Paz et al. (2009) and Akhavan et al. (2010), it is not easy to find correlation between the nitrate leached from the soil profile and the nitrate concentrations in aquifers because other factors such as groundwater age, groundwater depth, lateral flow and denitrification in the unsaturated zone also play a role. In this study, in order to reduce complexities in estimation of nitrate concentrations in aquifers, DRASTIC, SWAT, and Geo-ANN were utilized to find a relationship

between the nitrates leached from the soil profile and nitrate concentrations in aquifers and to identify intrinsic aquifer vulnerability and aquifer hazard areas.

An intrinsic aquifer vulnerability map was created using DRASTIC (Figure 4.11). DRASTIC scores range from 76 to 196, and they were normalized (from 0 to 1) by Equation 4.24 to have an equal weight when combining the intrinsic aquifer vulnerability map with the aquifer hazard map to produce the integrated aquifer vulnerability map. Intrinsic aquifer vulnerability indices were classified into five classes: 0 - 0.2 (“Very low”), 0.2 - 0.4 (“Low”), 0.4 - 0.6 (“Moderate”), 0.6 - 0.8 (“High”), and 0.8 - 1.0 (“Very high”). As shown in Figure 4.11 and Table 4.14, 9.6% of the aquifer systems in the UWRW was within in “Very low” vulnerability class, and 60.1% of the area was estimated as “Low”, 26.9% within “Moderate” vulnerability class, 3.2% within “High” vulnerability class, and 0.2% within “Very high” vulnerability class.

The intrinsic vulnerability results (Table 4.14) from DRASTIC were validated with the observed nitrate concentrations in wells (hereinafter, "well database") (Appendix A). The well database has 678 data points, including 116 nitrate data > 2 ppm. The results showed that approximately 42.2% of nitrate detections > 2 ppm are within “High” and “Very high” vulnerability areas (represent 3.4% of vulnerability area) as simulated by DRASTIC. Moreover, 53.4% of the nitrate detections were within “Moderate” vulnerability class (26.9% of area), and 4.3% of the nitrate detections were within “Low” vulnerability class (60.1% of area). In intrinsic aquifer vulnerability assessment, nitrates

> 2 ppm were not detected within the “Very low” vulnerability class (9.6% of area) (Table 4.14 and 4.17).

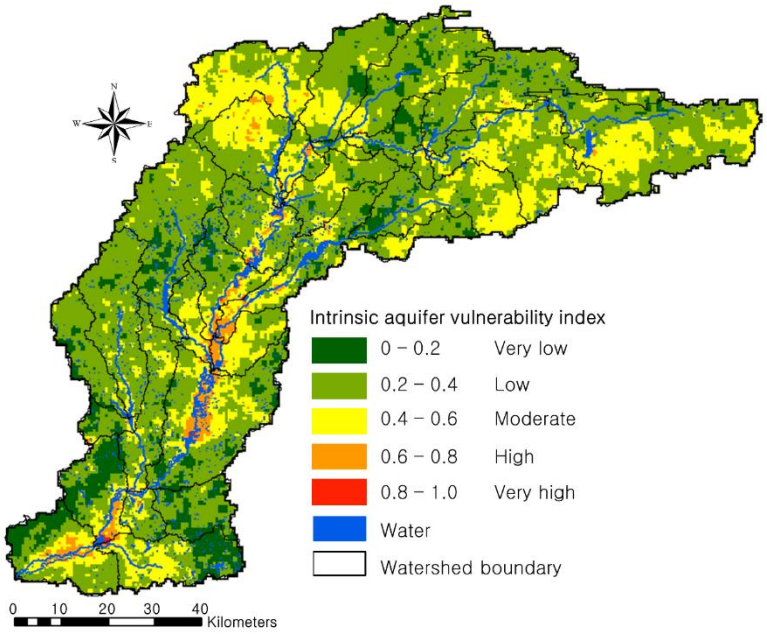


Figure 4.11 Intrinsic aquifer vulnerability map (DRASTIC) of the UWRW

Table 4.14 Vulnerability areas (%) and # of nitrate detections > 2 ppm within intrinsic aquifer vulnerability classes in the UWRW

Class	Areas (%)	Number of nitrate detections > 2 ppm
Very low	9.6	0 (0%)
Low	60.1	5 (4.3%)
Moderate	26.9	62 (53.4%)
High	3.2	42 (36.2%)
Very high	0.2	7 (6%)

NO3L (the nitrate leached from the soil profile) (mean value: 3.86 kg/ha/yr) and NO3GW (the nitrate transported into main stream from the groundwater loading) (mean value: 0.18 kg/ha/yr) were estimated for HRUs by SWAT simulation. Then, average (1990-2010) estimated annual nitrate concentrations in aquifers were predicted after

training/validation/testing using the Geo-ANN. Estimated nitrate concentration ranges from 0 to 12.25 mg/L, and these nitrate concentration values were also normalized (from 0 to 1) by Equation 4.24 for creating the integrated aquifer vulnerability map. Aquifer hazard indices were divided into five classes: 0 - 0.2 ("Very low"), 0.2 - 0.4 ("Low"), 0.4 - 0.6 ("Moderate"), 0.6 - 0.8 ("High"), and 0.8 - 1.0 ("Very high"). As shown in Figure 4.12 and Table 4.15, 25.0% of the aquifer systems in the UWRW was within in "Very low" vulnerability class, and 41.2% of the area was estimated as "Low", 22.7% within "Moderate" vulnerability class, 8.8% within "High" vulnerability class, and 2.4% within "Very high" vulnerability class.

The aquifer hazard results (Table 4.15) from SWAT and Geo-ANN were validated with the well database (Appendix A). The results indicated that approximately 79.3% of nitrate detections > 2 ppm are within "High" and "Very high" vulnerability areas (represent 11.2% of area) as predicted by SWAT and Geo-ANN. Moreover, 9.5% of the nitrate detections were within "Moderate" vulnerability class (22.7% of area), 8.6% of the nitrate detections were within "Low" vulnerability class (41.2% of area), and 2.6% of the nitrate detections were within "Very low" vulnerability class (25% of area) (Table 4.15 and 4.17).

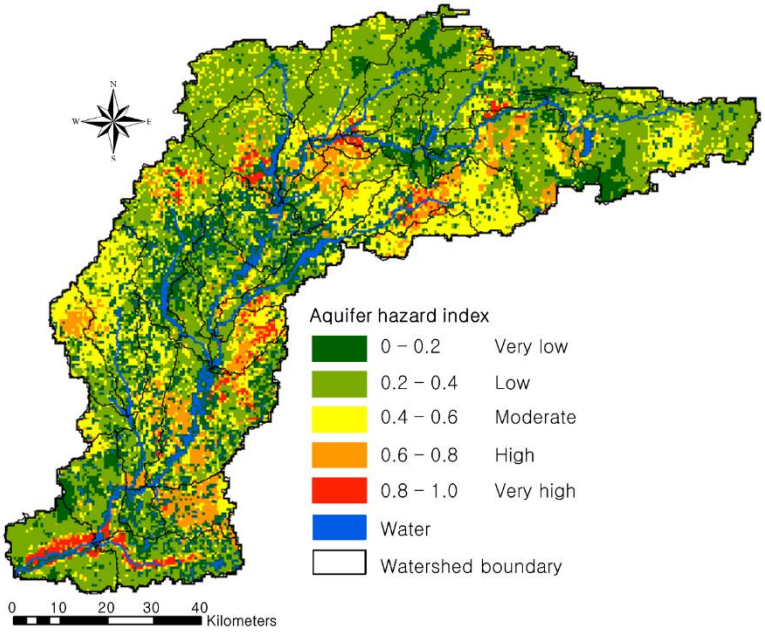


Figure 4.12 Aquifer hazard map (SWAT and Geo-ANN) of the UWRW

Table 4.15 Vulnerability areas (%) and # of nitrate detections > 2 ppm within aquifer hazard classes in the UWRW

Class	Area (%)	Number of nitrate detections > 2 ppm
Very low	25.0	3 (2.6%)
Low	41.2	10 (8.6%)
Moderate	22.7	11 (9.5%)
High	8.8	61 (52.6%)
Very high	2.4	31 (26.7%)

Finally, an integrated aquifer vulnerability map (Figure 4.12(a) and (b)) was generated by combining the normalized intrinsic aquifer vulnerability and the aquifer hazard maps. Then, integrated aquifer vulnerability indices were normalized again and its indices were also classified into five classes: 0 - 0.2 (“Very low”), 0.2 - 0.4 (“Low”), 0.4 - 0.6 (“Moderate”), 0.6 - 0.8 (“High”), and 0.8 - 1.0 (“Very high”). As shown in Figure 4.13 and Table 4.16, 12.8% of the aquifer system areas in the UWRW were within the “Very low” vulnerability class, and 50.7% of the area was estimated as “Low”, 30.7% within

"Moderate" vulnerability class, 5.4% within "High" vulnerability class, and 0.4% within "Very high" vulnerability class.

The results (Table 4.16) of integrated aquifer vulnerability assessment were validated with the well database (Appendix A). The results indicated that approximately 81.0% of nitrate detections > 2 ppm are within "High" and "Very high" (represent 5.8% of area) vulnerability areas as predicted by incorporating DRASTIC, SWAT, and Geo-ANN. Moreover, 12.1% of the nitrate detections were within "Moderate" vulnerability class (30.7% of area), and 4.3% of the nitrate detections were within "Low" vulnerability class (50.7% of area). Nitrates > 2 ppm were not detected within the "Very low" vulnerability class (12.8% of area) (Table 4.16 and 4.17). As shown in Tables 4.17 and 4.18, integrated aquifer vulnerability assessment better predicted nitrate detections than DRASTIC or SWAT/Geo-ANN by themselves. Table 4.18 indicates that detection ratio of integrated aquifer vulnerability assessment is the largest value among the three methods of aquifer vulnerability assessment (intrinsic aquifer vulnerability assessment by DRASTIC, aquifer hazard assessment by SWAT/Geo-ANN, and integrated aquifer vulnerability assessment by combining DRASTIC and SWAT/Geo-ANN).

For integrated aquifer vulnerability assessment, 19% of nitrate detections > 2 ppm were within the "Low", and "Moderate" vulnerability areas, which should have fallen in the "High", "Very High", or "Moderate" vulnerability areas. The nitrates detected in the "Low" and "Moderate" vulnerability areas might be caused by point sources, application timing, and excessive fertilizer application among other reasons which were not

considered in this study. These factors (i.e., point sources, application timing, and excessive fertilizer application) should be considered by adding additional data and modifying the models used in this study.

There were more nitrate detections in high classes ("High" + "Very high") in aquifer hazard and integrated aquifer vulnerability maps and fewer in the intrinsic vulnerability map. Areas of high classes in each map were different and the area of high classes might influence the number of nitrate detections. Machine learning may predict better than overlay and index GIS models, such as DRASTIC, because machine learning creates a model based on observed data – in this case nitrate concentrations in wells. Moreover, DRASTIC has limited ranges of parameters (i.e., DRASTIC ratings and weights). Further, the number of nitrate observations > 2 ppm are 116 (17.1%) out of 678. The small number of well samples > 2 ppm would increase uncertainties when models detect nitrates > 2 ppm in high classes ("High" + "Very high"). Also, if there are more well observations with nitrate > 2 ppm, better prediction would be expected.

This study assumed that the intrinsic aquifer vulnerability and aquifer hazard contribute equally to groundwater contamination in the UWRW. So, the intrinsic aquifer vulnerability and aquifer hazard indices were assigned equal weights when the integrated aquifer vulnerability map was produced. Researchers or policy makers can adjust each index's weight based on additional information or scientific experience.

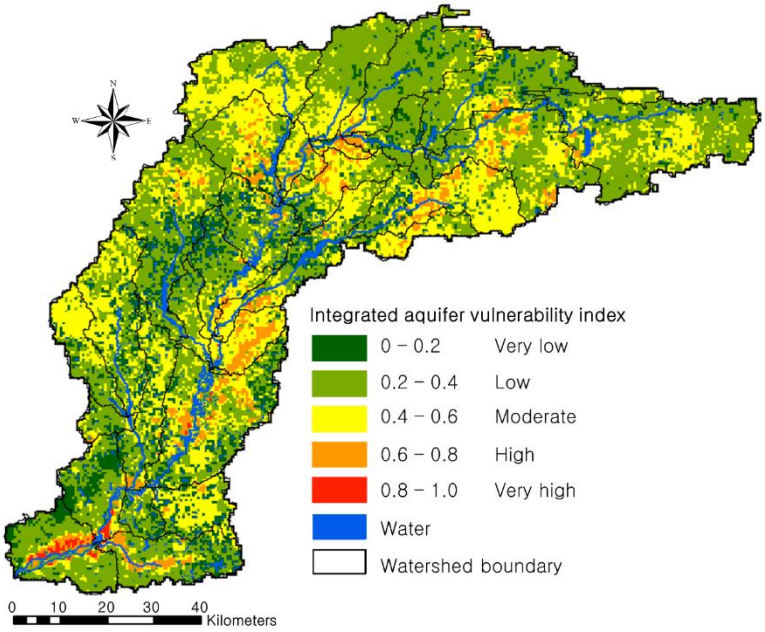


Figure 4.13 Integrated aquifer vulnerability map (combined intrinsic aquifer vulnerability map and aquifer hazard map) of the UWRW

Table 4.16 Vulnerability areas (%) and # of nitrate detections > 2 ppm within integrated aquifer vulnerability classes in the UWRW

Class	**Area (%)**	**Number of nitrate detections > 2 ppm**
Very low	12.8	0 (0%)
Low	50.7	8 (6.9%)
Moderate	30.7	14 (12.1%)
High	5.4	82 (70.7%)
Very high	0.4	12 (10.3%)

Table 4.17 Comparison of nitrate detections in wells with three types of aquifer assessment in the UWRW

	Number of nitrate detections > 2ppm		
Vulnerability class	IV[1]	AH[2]	IT[3]
Very low	0 (0%)	3 (2.6%)	0 (0%)
Low	5 (4.3%)	10 (8.6%)	8 (6.9%)
Moderate	62 (53.4%)	11 (9.5%)	14 (12.1%)
High	42 (36.2%)	61 (52.6%)	82 (70.7%)
Very high	7 (6.0 %)	31 (26.7%)	12 (10.3%)

[1]IV: Intrinsic aquifer vulnerability assessment using DRASTIC

[2]AH: Aquifer hazard assessment using SWAT and Geo-ANN

[3]IT: Integrated aquifer vulnerability assessment by combining IV and AH

Table 4.18 Comparison of detection ratio with three aquifer assessments in the UWRW

Category	Result (2016)		
	IV[1]	AH[2]	IT[3]
HV-Area[4] (%)	3.4	11.2	5.8
N-Detections[5] (%)	42.2	79.3	81.0
D-Ratio[6]	12.4	7.1	14.0

[1]IV: Intrinsic aquifer vulnerability assessment using DRASTIC

[2]AH: Aquifer hazard assessment using SWAT and Geo-ANN

[3]IT: Integrated aquifer vulnerability assessment by combining IV and AH

[4]HV-Area: "High" and "Very high" vulnerability areas (%)

[5]N-Detection: Nitrate detections > 2 ppm

[6]D-Ratio: Detection ratio

"High" and "Very high" vulnerability classes in the integrated aquifer vulnerability map (Figure 4.13) include crop areas (mainly corn/soybean areas). A primary cause of nitrate contamination in aquifers results from anthropogenic fertilization (Lin et al., 2001; Behm, 1989). Because most farmers consider nitrogen fertilizer to be "cheap insurance" against a crop failure, farmers would obviously rather add too much nitrogen to their crops for increasing profits (Looker, 1991). According to the study by Burkart and Kolpin (1993), they found that water samples from wells surrounded by corn/soybean have a dramatically larger frequency of excess nitrate than wells with other crops. Similarly, corn/soybean acreage is responsible for 11 times more nitrate contamination than acreage used as rangeland (Department of Commerce, 1993; Puckett, 1994). Therefore, Figure 4.13 would explain the previous studies and this study for nitrate contamination in aquifers.

In terms of the nitrate contamination in aquifers, the most effective approach to avoid health risks (e.g., methemoglobinemia (blue baby syndrome)) is to reduce application rate of fertilization and to check wells frequently. Further, fertilizer application timing is a critical factor to cause groundwater degradation (Sullivan et al., 2000). Various Best Management Practices (BMPs) to prevent aquifers from nitrate contamination should be implemented and they would help reduce nitrate leaching from the soil profile into aquifers. However, BMPs such as restriction of fertilizer application rate would help alleviate nitrate concentrations to biologically safe levels.

In previous studies, various models were combined to overcome model limitations in an attempt to improve prediction. Lim (2001) used two models to predict pollutant losses to shallow groundwater in the White River Basin from crop land, pasture, urban, and forest. The results of NAPRA (National Agricultural Pesticide Risk Analysis) for pasture and crop land and L-THIA for urban, forest, and water were combined to reduce limitations of each model. However, there were still limitations for watershed scale modeling. NAPRA uses the GLEAMS (Groundwater Loading Effects of Agricultural Management Systems) model, a field scale model, and thus does not represent some important watershed processes (Lim, 2001). In the work presented herein, the watershed model SWAT was used to estimate nitrate leaching to shallow aquifers to provide input data for Geo-ANN which was utilized to predict nitrate concentrations in wells.

Navulur (1996) used two models to estimate aquifer vulnerability of groundwater systems in Indiana using a GIS environment at a 1:250,000 scale. Navulur (1996)

combined DRASTIC with NLEAP (Nitrate Leaching and Economic Analysis) to predict well nitrate detections for aquifer vulnerability assessment. The data scale used in Navulur's (1996) study was coarse (1:250,000) for field scale simulations. The use of a single rainfall event and coarse data scale may result in underestimated "Low" vulnerability areas and overestimated "High" vulnerability areas. In this study, high resolution data (1:24,000) were used. Further, estimated nitrate leaching results from SWAT, which uses long-term and continuous rainfall data, were used to predict nitrate concentrations in wells using Geo-ANN.

As shown in Navulur's (1996) results, approximately 91.8% of nitrate detections in wells > 2 ppm are within "High" and "Very high" vulnerability areas (represent 56.9% of area) as predicted by the combined model. Compared with Navulur's (1996) study, the results presented herein had approximately 81.0% of nitrate detections in wells > 2 ppm within "High" and "Very high" (represent 5.8% of area) vulnerability areas as predicted by incorporating DRASTIC, SWAT, and Geo-ANN. Detection ratio (% of nitrate detections to % of vulnerability areas with the larger the detection ratio, the better the prediction) for "High" and "Very high" areas from Navulur (2006) results in a value of 1.6, and results presented herein provide a value of 14.0. Thus, Navulur's (1996) combined model may overestimate areas of "High" and "Very high" vulnerability classes.

4.5 Conclusions

Integrated aquifer vulnerability assessment was conducted using DRASTIC, SWAT, and Geo-ANN in the Upper White River Watershed (UWRW) located in central Indiana.

Integrated aquifer vulnerability was performed by combining intrinsic aquifer vulnerability assessment and aquifer hazard assessment in a GIS environment. The intrinsic aquifer vulnerability map was created using DRASTIC. Then, the integrated aquifer vulnerability map was produced by combining the intrinsic aquifer vulnerability map and aquifer hazard map. Each map was standardized by feature scaling, and both maps have a same range of values (min = 0 and max = 1) which mean the intrinsic aquifer vulnerability map and aquifer hazard map have the same weight. This study assumed both maps have the same impact to aquifer vulnerability.

In the analysis of integrated aquifer vulnerability, 12.8% of the aquifer systems in the UWRW was within in "Very low" vulnerability class, and 50.7% of the area was estimated as "Low", 30.7% within "Moderate" vulnerability class, 5.4% within "High" vulnerability class, and 0.4% within "Very high" vulnerability class.

Approximately 81.0% of nitrate detections in wells > 2 ppm were within "High" and "Very high" vulnerability areas (represent 5.8% of area) as predicted by incorporating DRASTIC, SWAT, and Geo-ANN. Moreover, 12.1% of the nitrate detections were within "Moderate" vulnerability class (30.7% of area), and 6.9% of the nitrate detections were within "Low" vulnerability class (50.7% of area). Nitrates in wells > 2 ppm were not detected within the "Very low" vulnerability class (12.8% of area). Those results show that integrated aquifer vulnerability assessment performed well. Integrated aquifer vulnerability assessment better predicted nitrate detections than DRASTIC or SWAT/Geo-ANN by themselves according to the results of detection ratio. Detection

ratio (percent of nitrate detections > 2 ppm to percent of "Very high" and "High" vulnerability areas with larger detection ratio indicating better prediction) indicated a value of 12.4 for intrinsic vulnerability assessment using DRASTIC, 7.1 for aquifer hazard assessment using SWAT/Geo-ANN, and 14.0 for integrated aquifer vulnerability assessment using the combined DRASTIC and SWAT/Geo-ANN.

For the simulation of complex systems such as groundwater dynamics in aquifers, this study indicates that machine learning is a suitable technique in predicting nonstationary behavior of groundwater quality. Also, with lack of detailed knowledge of the internal functions of complex systems and insufficient data for calibration and validation, machine learning techniques would be an efficient method to identify the nonstationary patterns of variables of groundwater quality.

The integrated aquifer vulnerability assessment considers both intrinsic aquifer properties and pollutant transport properties. Thus, the overall assessment of aquifer vulnerability can be performed using the integrated aquifer vulnerability assessment technique provided in this study. Moreover, this approach is expected to be an efficient guide for managing groundwater resources for policy makers and groundwater-related researchers.

The models used in this study are data-driven models. Therefore, if more data (i.e., nitrate concentration data in well, point sources, application timing, and fertilizer application) are available, the approach suggested in this study would be improved and more accurate.

Potential next steps in extending this work are to: 1) conduct comparison of various machine learning algorithms (e.g., convolution neural network, Bayesian linear regression, and decision forest regression) to better predict nitrate contamination in aquifers and 2) evaluate the application of BMPs to reduce nitrate leaching from the soil profile into aquifers.

4.6 References

Abbaspour, K.C., 2011. Swat-Cup2: SWAT Calibration and Uncertainty Programs Manual Version 2, Department of Systems Analysis, Integrated Assessment and Modelling (SIAM), Eawag. Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.

Abbaspour, K.C., Johnson, C.A., Van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal 3, 1340–1352.

Akhavan, S., Abedi-Koupai, J., Mousavi, S.F., Afyuni, M., Eslamian, S.S., Abbaspour, K.C., 2010. Application of SWAT model to investigate nitrate leaching in Hamadan–Bahar Watershed, Iran. Agriculture, Ecosystems & Environment 139, 675–688.

Akhavan, S., Mousavi, S.F., Abedi-Koupai, J., Abbaspour, K.C., 2011. Conditioning DRASTIC model to simulate nitrate pollution case study: Hamadan–Bahar plain. Environmental Earth Sciences 63, 1155–1167.

Alcalá, F.J., Martínez-Valderrama, J., Robles-Marín, P., Guerrera, F., Martín-Martín, M., Raffaelli, G., de León, J.T., Asebriy, L., 2015. A hydrological–economic model for sustainable groundwater use in sparse-data drylands: Application to the Amtoudi Oasis in southern Morocco, northern Sahara. Science of the Total Environment 537, 309–322.

Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and storage in groundwater systems. Science 296, 1985-1991.

Araghinejad, S., 2014. Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering, Springer 67, 1-5.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M. K., 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE 55, 1491-1508.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: Model development. Journal of the American Water Resources Association 34 (1), 73-89.

ASCE Task Committee, 2000. Artificial neural network in hydrology: I. Preliminary concepts. J. Hydrol. Eng. 5:115–123.

Beale, M.H., Hagan, M.T., Demuth, H.B., 2016. Neural Network Toolbox User's Guide, The MathWorks, Inc. Natick, MA.

Behm, D., 1989. Ill Waters: The Fouling of Wisconsin's Lakes and Streams (Special Report), The Milwaukee Journal 2.

Boles, C.M.W., Frankenberger, J.R., Moriasi, D.N., 2015. Tile Drainage Simulation in SWAT2012: Parameterization and Evaluation in an Indiana Watershed, 2015. . Transactions of the ASABE 1201–1213.

Brouyére, S., Jeannin, P.Y., Dassargues, A., Goldscheider, N., Popescu, C., Sauter, M., Vadillo, I., Zwahlen, F., 2001. Evaluation and validation of vulnerability concepts using a physically based approach. 7th Conference on Limestone Hydrology and Fissured Media, Besnacon 20–22 Sep. 2001, Sci. Tech. Envir., Mém. H.S. No. 13, 67–72.

Burkart, M.R., Kolpin, D.W., 1993. Hydrologic and land-use factors associated with herbicides and nitrate in near-surface aquifers, J. Environ. Qual. 22, 646-656.

Chen, J., Wu, Y., 2012. Advancing representation of hydrologic processes in the Soil and Water Assessment Tool (SWAT) through integration of the TOPographic MODEL (TOPMODEL) features, Journal of Hydrology 420-421, 319-328.

Chen, S.K., Jang, C.S., Peng, Y.H., 2013. Developing a probability-based model of aquifer vulnerability in an agricultural region. Journal of Hydrology 486, 494–504.

Coulibaly, P., Baldwin, C.K., 2005. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. Journal of Hydrology 307, 164–174.

De Paz, J.M., Delgado, J.A., Ramos, C., Shaffer, M.J., Barbarick, K.K., 2009. Use of a new GIS nitrogen index assessment tool for evaluation of nitrate leaching across a Mediterranean region. Journal of Hydrology 365, 183–194.

Department of Commerce, 1993. Statistical Abstract of the United States. Bureau of the Census. Washington, DC.

Department of Land and Water Conservation, 1998. Stressed Rivers Assessment Report, NSW State Summary, NSW Department of Land and Water Conservation, Parramatta, NSW.

Du, B., Saleh, A. Jaynes, D.B., Arnold, J.G., 2006. Evaluation of SWAT in simulating nitrate nitrogen and atrazine fates in a watershed with tiles and potholes. Trans. ASABE 49(4), 949-959.

Engel, B., Storm, D., White, M., Arnold, J., Arabi, M., 2007. A Hydrologic/Water Quality Model Application Protocol. Journal of the American Water Resources Association 43(5), 1223–1236.

Fleming, A.H., Brown, S.E., and Ferguson, V.R., 1993, Hydrogeologic Framework of Marion County, Indiana: Indiana Geological Survey Open-File Study 93-05, 67, 3.

Getis, A., Ord, J.K., 1992. The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24(3), 189-206.

Green, C.H., Tomer, M.D., Di Luzio, M., Arnold, J.G., 2006. Hydrologic evaluation of the soil and water assessment tool for a large tile-drained watershed in Iowa. Transactions of the ASABE 49, 413–422.

Grimmeisen, F., Zemann, M., Goeppert, N., Goldscheider, N., 2016. Weekly variations of discharge and groundwater quality caused by intermittent water supply in an urbanized karst catchment. Journal of Hydrology 537, 157–170.

Grusson, Y., Sun, X., Gascoin, S., Sauvage, S., Raghavan, S., Anctil, F., Sáchez-Pérez, J.M., 2015. Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed. Journal of Hydrology 531, 574–588.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. Journal of Hydrologic Engineering 4, 135–143.

Guyon, I., 1997. A scaling law for the validation-set training-set size ratio. AT&T Bell Laboratories 1–11.

Ha, H., Stenstrom, M.K., 2003. Identification of land use with water quality data in stormwater using a neural network. Water Research 37, 4222–4230.

Heaton, J., 2008. Introduction to Neural Networks for Java. Heaton Research, Inc.

Her, Y., Chaubey, I., Frankenberger, J., Smith, D., 2016. Effect of conservation practices implemented by USDA programs at field and watershed scales. Journal of Soil and Water Conservation 71, 249–266.

Jiang, Y., Frankenberger, J.R., Sui, Y., Bowling, L.C., 2014. Estimation of nonpoint-source nitrate concentrations in Indiana rivers based on agricultural drainage in the watershed. J. American Water Resources Assoc. 50(6), 1501-1514.

Kalin, L., Isik, S., Schoonover, J.E., Lockaby, B.G., 2010. Predicting Water Quality in Unmonitored Watersheds Using Artificial Neural Networks. Journal of Environment Quality 39, 1429-1440.

Kamp, R.G., Savenije, H.H.G., 2007. Hydrological model coupling with ANNs. Hydrology and Earth System Sciences Discussions 11, 1869–1881.

Kellogg, R.L., M.S. Maizel, and D.W. Goss; 1994. The potential for leaching of agrichemicals used in crop production: A national perspective. Journal of Soil and Water Conservation 49, 294-298.

Koudstaal, R., Rijsberman, F. R., Savenije, H. H. G., 1992. Water and Sustainable Development, Natural Resources Forum, 16(4), 277–290.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences 5, 89–97.

Lam, Q.D., Schmalz, B., Fohrer, N., 2010. Modelling point and diffuse source pollution of nitrate in a rural lowland catchment using the SWAT model, Agr. Water Manage. 97, 317–325.

Lim, K.J., Engel, B.A., Tang, Z., Choi, J., Kim, K.-S., Muthukrishnan, S., Tripathy, D., 2005. Automated web GIS based hydrograph analysis tool, WHAT. Journal of the American Water Resources Association 41(6), 1407-1416.

Lin, B-L., Sakoda, A., Shibasaki, R., Suzuki, M., 2001. A modeling approach to global nitrate leaching caused by anthropogenic fertilization. Wat. Res. 35(8), 1961-1968.

Looker, D., 1991. Nitrogen use still too high, Experts Say, Des Moines (Iowa) Register.

Maxwell, R.M., Condon, L.E., Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci. Model Dev. 8, 1-15.

Mohd Ali, J., Hussain, M.A., Tade, M.O., Zhang, J., 2015. Artificial Intelligence techniques applied as estimator in chemical process systems – A literature survey. Expert Systems with Applications 42, 5915–5931.

Monteith, J. L., 1965. Evaporation and environment. Symposia of the Society for Experimental Biology 19, 205-234.

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

Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Transactions of the ASABE 58, 1763–1785.

Morris, B.L., Lawrence, A.R., Chilton, P.J., Adams, B., Caylow, R.C., Klinck, B.A., 2003. Groundwater and its susceptibility to degradation: a global assessment of the problems and options for management. UNEP Early Warning & Assessment Report Series RS. 03-3, Nairobi, Kenya.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology 10, 282–290.

Navulur, K.C.S., 1996. Groundwater vulnerability evaluation to nitrate pollution on a regional scale using GIS. PhD dissertation, Purdue University. West Lafayette: ProQuest/UMI.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool, Theoretical Documentation, Version 2009. Blackland Research Center, Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, TX.

Noori, N., Kalin, L., 2016. Coupling SWAT and ANN models for enhanced daily streamflow prediction. Journal of Hydrology 533, 141–151.

Nourani, V., Alami, M.T., Vousoughi, F.D., 2015. Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling. Journal of Hydrology 524, 255–269.

Nourani, V., Mehdi, K., Akira, M., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resources Management 23, 2877–2894.

Nourani, V., Mehdi, K., Akira, M., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resour Manage 23, 2877–2894.

Ord J.K., Getis, A., 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 27(4), 286-306.

Pacheco, F.A.L., Sanches Fernandes, L.F., 2013. The multivariate statistical structure of DRASTIC model. Journal of Hydrology 476, 442–459.

Petheram, C., Dawes, W., Grayson, R., Bradford, A., Walker, G., 2003. A sub-grid representation of groundwater discharge using a one-dimensional groundwater model. Hydrological Processes 17, 2279–2295.

Phillips, S.W., Focazio, M.J., Bachman, L.J., 1999. Discharge, nitrate load, and residence time of ground water in the Chesapeake Bay watershed. US Department of the Interior, US Geological Survey. USGS Fact Sheet FS-150-99.

Pohlert, T., Breuer, L., Huisman, J.A., Frede, H.G., 2006. Assessing the model performance of an integrated hydrological and biogeochemical model for discharge and nitrate load predictions. Hydrology and Earth System Sciences Discussions 3, 2813–2851.

Puckett, L.J., 1994. Nonpoint and point sources of nitrogen in major watersheds of the United States. US Geological Survey.

Puckett, L.J., 1994. Nonpoint and Point Sources of Nitrogen in Major Watersheds in the United States. U.S. Geological Survey Water Resources Investigations Report 94-4001. Washington, DC.

Pulido-Calvo, I., Roldan, J., Lopez-Luque, R., Gutierrez-Estrada, J.C., 2003. Demand forecasting for irrigation water distribution systems. Irr Drain Eng 129(6), 422-431.

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

SCS, 1972. National Engineering Handbook Section 4 Hydrology, Chapters 4-10, U.S. Department of Agriculture, Soil conservation Service.

Sethi, R.R., Kumar, A., Sharma, S.P., Verma, H.C., 2010. Prediction of water table depth in a hard rock basin by using artificial neural network. International Journal of Water Resources and Environmental Engineering 4, 95–102.

Shao, Z., Er, M.J., 2016. Efficient Leave-One-Out Cross-Validation-based Regularized Extreme Learning Machine. Neurocomputing 194, 260–270.

Simon, H. A., Langley, P., 1995. Applications of machine learning and rule induction. Communications of the ACM 38(11), 54-64.

Singh, J., Knapp, H.V., Demissie, M., 2004. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, Ill.: Illinois State Water Survey.

Solley, W.B., R.R. Pierce, and H.A. Perlman. 1998. Estimated use of water in the United States in 1995. U.S. Geological Survey, Reston, Virginia.

Solomatine, D. P., Shrestha, D. L., 2009. A novel method to estimate model uncertainty using machine learning techniques, Water Resour. Res., 45, W00B11.

Solomatine, D.P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. J of Hydroinformatics 10(1), 3-22.

Sui, Y., Frankenberger, J.R., 2008. Nitrate loss from subsurface drains in an agricultural watershed using SWAT2005. Trans. ASABE, 51(4), 1263-1272.

Sullivan, D.M., Cogger, C.G., Bary, A.I. Fransen, S.C., 2000. Timing of dairy manure applications to perennial grass on well drained and poorly drained soils. J. Soil Water Conserv. 55, 147.

Tedesco, L.P., Hoffmann, J., Bihl, L., Hall, B.E., Barr, R.C., Stouder, M., 2011. Upper White River Watershed Regional Watershed Assessment and Planning Report.

Tedesco, L.P., Pascual, D.L., Shrake, L.K., Hall, R.E., Casey, L.R., Vidon, P.G.F., Hernly, F.V., Salazar, K.A., Barr, R.C., 2005. Eagle Creek Watershed Management Plan: An Integrated Approach to Improved Water Quality. Eagle Creek Watershed Alliance, CEES Publication 2005-07, IUPUI, Indianapolis, 182.

Unland, N.P., Cartwright, I., Daly, E., Gilfedder, B.S., Atkinson, A.P., 2015. Dynamic river-groundwater exchange in the presence of a saline, semi-confined aquifer: The Dynamics of River-Groundwater Exchange in River Banks. Hydrological Processes 29, 4817–4829.

USDA-NRCS, 2011. Scope and effect equations. Wisconsin supplement to Chapter 19: Hydrology tools for wetland determination, in Engineering Field Handbook, Part 650. EFH Notice 210-WI-121. Madison, Wisc.: USDA-NRCS Wisconsin State Office.

Van Liew, M.W., Arnold, J.G., Garbrecht, J.D., 2003. Hydrologic simulation on agricultural watersheds: Choosing between two models. Transactions of the ASAE 46, 1539.

Williams, J.R., Jones, C.A., Dyke, Pt., 1984. A modeling approach to determining the relationship between erosion and soil productivity. Transactions of the ASAE 27, 129–144.

Winter, T.C., Harvey, J.W., Franke, O.L., Alley, W.M., 1998. Ground Water and Surface Water: A Single Resource, U.S. Geological Survey Circular 1139, Denver, CO.

Wong, T.T., 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognition 48, 2839–2846.

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

Yeo, I.Y., Lee, S., Sadeghi, A.M., Beeson, P.C., Hively, W.D., McCarty, G.W., Lang, M.W., 2014. Assessing winter cover crop nutrient uptake efficiency using a water quality simulation model. Hydrology and Earth System Sciences 18, 5239–5253.

Yoon, T., Rhodes, C., Shah, F.A., 2015. Upstream water resource management to address downstream pollution concerns: A policy framework with application to the Nakdong River basin in South Korea. Water Resour. Res. 51, 787–805.

Yu, H.L., Lin, Y.C., 2015. Analysis of space–time non-stationary patterns of rainfall–groundwater interactions by integrating empirical orthogonal function and cross wavelet transform methods. Journal of Hydrology 525, 585–597.

Zhang, X., Srinivasan, R., Arnold, J., Izaurralde, R.C., Bosch, D., 2011. Simultaneous calibration of surface flow and baseflow simulations: a revisit of the SWAT model calibration framework. Hydrol. Process. 25, 2313–2320.

CHAPTER 5. CONCLUSIONS

5.1 Summary and Conclusions

Integrated aquifer vulnerability assessment was conducted using DRASTIC, SWAT, and Geo-ANN in the Upper White River Watershed (UWRW) located in central Indiana. Integrated aquifer vulnerability was performed by combining intrinsic aquifer vulnerability assessment and aquifer hazard assessment. DRASTIC was used for intrinsic aquifer vulnerability assessment, and SWAT and Geo-ANN were utilized to evaluate aquifer hazard in the UWRW. In an attempt to improve estimation of integrated aquifer vulnerability including intrinsic aquifer vulnerability and aquifer hazard, methodologies were modified and developed in a GIS environment.

The three objectives of the dissertation were to:

1. Assess intrinsic aquifer vulnerability for sustainable groundwater management using DRASTIC.
2. Develop efficient flow calibration regime (EFCR) for accurate estimation of hydrologic and water quality components using a watershed scale hydrological model (SWAT).
3. Evaluate integrated aquifer vulnerability using DRASTIC, a watershed scale hydrological model (SWAT), and Geo-ANN.

In the first objective, intrinsic aquifer vulnerability assessment was implemented using DRASTIC in the UWRW. Approaches for intrinsic aquifer vulnerability assessment were improved with high resolution data by data preprocessing. The Soil Conservation Service (SCS) runoff curve number (CN) method was used to produce a recharge (R) map with annual rainfall, SSURGO data and National Land Cover Database (NLCD), providing a wide range of recharge values compared with previous studies because high resolution SSURGO data generated various CN and recharge values. An aquifer media (A) map was created using the aquifer systems map and report from the U.S. Geological Survey (USGS) and Indiana Department of Natural Resources (IDNR). Most aquifer media of the study area were sand and gravel, but based on the INDR reports, aquifer media ratings were assigned in more detail. IDNR reports described vulnerability of each aquifer system such as 'very high susceptibility to surface contamination (very high)', 'highly susceptible to surface contamination (high)', 'moderately susceptible to surface contamination (moderate)', 'low susceptibility to surface contamination (low)', and 'very low susceptibility to surface contamination (very low)'. Soil media (S) and topography (T) maps were obtained through SSURGO data from USDA-NRCS instead of STATSGO data often used. The map scale of SSURGO data is 1:12,000, whereas STATSGO is 1:250,000. The impact of vadose zone media (I) map was estimated using sand, silt, and clay thickness point data within lithology data from IDNR.

116 wells (total 678 wells) with nitrate levels > 2 ppm were selected to calibrate and validate estimated aquifer vulnerability areas. Nitrate levels vary from 0.1 to 18.3 mg/l (ppm) with an average of 1.2 mg/l in this area. DRASTIC parameters were optimized by

calibrating DRASTIC weights using a binary classifier calibration method with a genetic algorithm (Bi-GA). In the binary classifier calibration process, results of the model and observed data are classified as 0 or 1. DRASTIC produced five vulnerability classes (i.e., very high, high, moderate, low, very low) using data normalization (feature scaling), which is a method used to standardize the range of independent variables (min=0, max=1). Very high and high vulnerability classes were classified as 1, and other classes were classified as 0. Observed nitrate concentrations in wells > 2 ppm were classified as 1 because the purpose of this study is to identify high or low vulnerability areas using nitrate concentrations in wells > 2 ppm with DRASTIC. Root mean square error (RMSE) was used to evaluate the effectiveness of the Bi-GA and its ability to make predictions in the calibration procedure. Intrinsic aquifer vulnerability indices from the improved DRASTIC were compared with observed groundwater quality data to explore how well simulated results match observed nitrate data > 2 ppm. RMSE without calibration was 0.70, and RMSE following calibration of DRASTIC weights with Bi-GA was 0.57 (the lower RMSE, the better the performance). An accuracy assessment error matrix was computed for spatial validation of the calibrated DRASTIC by using Bi-GA. Total accuracies of uncalibrated DRASTIC and calibrated DRASTIC were 35% and 42%, respectively. Thus, the results of accuracy assessment indicate calibrated DRASTIC using Bi-GA predicted intrinsic aquifer vulnerability areas more accurately than DRASTIC without calibration.

The intrinsic aquifer vulnerability results from DRASTIC calibrated by using Bi-GA were validated with a well database of nitrate concentrations > 2 ppm. The results

showed that approximately 42.2% of nitrate detections in wells > 2 ppm are within the "High" and "Very high" vulnerability areas (represent 3.4% of area) as simulated by DRASTIC. Moreover, 53.4% of the nitrate detections were within "Moderate" vulnerability class (26.9% of area), and 4.3% of the nitrate detections were within "Low" vulnerability class (60.1% of area). In intrinsic aquifer vulnerability assessment, nitrates in wells > 2 ppm were not detected within the "Very low" vulnerability class (9.6% of area). Intrinsic aquifer vulnerability assessment with DRASTIC calibrated by using Bi-GA better predicted nitrate detections than DRASTIC without calibration.

Potential monitoring locations and areas where groundwater protection and management should be focused were determined based on the Gi* statistic method. The Gi* statistic method identifies statistically significant spatial clusters of high values (high vulnerability areas) and low values (low vulnerability areas). The Gi* statistic method returns z-score and the higher the z-score, the stronger the intensity of the clustering. The portion of z-score over 1.65 is 19.9% of the study area, indicating these are areas where groundwater monitoring and BMPs for groundwater quality protection should be focused. Hotspot areas were found along the stream and river because those areas include high permeability of alluvium, sand, and gravel. Further, depth to water is shallow. These areas should be priority areas for groundwater protection.

The second objective developed the EFCR for both accurate streamflow and baseflow estimation and for better prediction of aquifer hazard. The calibration methodology developed and suggested in the second objective was composed of the modified SWAT

2012 code and SUFI-2 in order to automatically calibrate streamflow and baseflow simultaneously.

Simultaneous streamflow and baseflow calibration with the EFCR (C4) for the calibration period (1990 - 2001) showed that NSE / R^2 / PBIAS for streamflow were 0.85 / 0.87 / 3.90, and NSE / R^2 / PBIAS for baseflow were 0.63 / 0.73 / 16.7. For the validation period (2002 - 2010), NSE / R^2 / PBIAS for streamflow and baseflow showed 0.88 / 0.92 / 1.50 (streamflow) and 0.65 / 0.70 / 13.8 (baseflow), respectively. These values indicate that the calibrated model (C4) is more than "Satisfactory" for both streamflow and baseflow estimation.

The results of the second objective indicated that even though the model with streamflow calibration alone (C2) showed the best performances for both streamflow calibration and validation periods, C2 would not be appropriate for baseflow calibration and validation. Calibrating streamflow and baseflow simultaneously would be needed for efficient hydrological cycle assessment because baseflow is the main part of the hydrological cycle for aquifer hazard assessment. Thus, the EFCR would be a practical step in the process for aquifer hazard assessment by calibrating baseflow accurately as well as streamflow at a watershed scale.

The second objective is expected to be used in data-driven models for in-depth groundwater modeling because the baseflow-related parameters (i.e., groundwater recharge, hydraulic conductivity, and so on) calibrated in the second objective can be

used as a set of input data (initial parameter values) in computer-based numerical groundwater models. If calibrated baseflow-related data is used as initial parameter values in groundwater modeling, the uncertainty of groundwater modeling would be reduced by minimizing the initial parameter uncertainty.

In the third objective, integrated aquifer vulnerability assessment was conducted using improved methodologies from the first and second objectives. For aquifer hazard assessment, Geo-ANN was developed for training, validation, testing, and predicting well nitrate concentrations with the shapefile format as well as tabular format using ANN. Geo-ANN provides a user-friendly GUI for optimizing the parameters of the ANN architecture and training the network for the prediction. Two predefined input variables (NO3L: the nitrate leached from the soil profile and NO3GW: the nitrate transported into streams from groundwater loading) which are concerned with aquifer hazard that reflects potential pollution transport from the land surface to aquifers were retrieved from the calibrated SWAT. SWAT and Geo-ANN were utilized to find the relationship between the nitrates leached from the soil profile and nitrate concentrations in aquifers and to identify potential aquifer hazard areas. Through these simulation processes, NO3L and NO3GW were estimated at the HRU level using SWAT. Then, with those two variables from SWAT and observed nitrate concentration data in wells after training, validation, and testing (NSE/R^2/PBIAS: 0.66/0.70/0.7) using the Geo-ANN, an aquifer hazard map was created.

An integrated aquifer vulnerability map was produced by combining the intrinsic aquifer vulnerability map from the first objective and aquifer hazard map. In the analysis of integrated aquifer vulnerability, 12.8% of the aquifer systems in the UWRW were within the "Very low" vulnerability class, and 50.7% of the area was estimated as "Low", 30.7% within the "Moderate" vulnerability class, 5.4% within the "High" vulnerability class, and 0.4% within the "Very high" vulnerability class. Approximately 81.0% of nitrate detections in wells > 2 ppm were within "High" and "Very high" vulnerability areas (represent 5.8% of area) as predicted by incorporating DRASTIC, SWAT, and Geo-ANN. An additional 12.1% of the nitrate detections were within the "Moderate" vulnerability class (30.7% of area), and 6.9% of the nitrate detections were within the "Low" vulnerability class (50.7% of area). Nitrates in wells > 2 ppm were not detected within the "Very low" vulnerability class (12.8% of area). Detection ratio (percent of nitrate detections > 2 ppm to percent of "Very high" and "High" vulnerability areas with larger detection ratio indicating better prediction) indicated a value of 12.4 for intrinsic vulnerability assessment using DRASTIC, 7.1 for aquifer hazard assessment using SWAT/Geo-ANN, and 14.0 for integrated aquifer vulnerability assessment using the combined DRASTIC and SWAT/Geo-ANN. Based on the results of detection ratio, integrated aquifer vulnerability assessment better predicted nitrate detections than DRASTIC or SWAT/Geo-ANN by themselves.

For the simulation of complex systems such as groundwater dynamics in aquifers, the third objective indicated that machine learning is a suitable technique in predicting nonstationary behavior of groundwater quality. Also, with lack of detailed knowledge of

the internal functions of complex systems and insufficient data for calibration and validation, machine learning techniques would be an efficient method to identify the nonstationary patterns of variables of groundwater quality.

The integrated aquifer vulnerability assessment considers both intrinsic aquifer properties and pollutant transport properties. Thus, the overall assessment of aquifer vulnerability can be performed using the integrated aquifer vulnerability assessment technique provided in the third objective. Moreover, results from this approach are expected to be an efficient guide for managing groundwater resources for policy makers and groundwater-related researchers.

5.2 Recommendations for Future Research

Even though the methodologies modified and developed in this dissertation show promise for efficient integrated aquifer vulnerability assessment, this study also indicates that further study is needed. Specific recommendations include the following:

1. Bi-GA was used as a new calibration framework for overlay and index modeling (DRASTIC). However, in this study only two groups were classified with over (high and very high vulnerability classes) or under (other vulnerability classes) 2 ppm of nitrate concentrations in wells. More classification is required for more accurate intrinsic aquifer vulnerability assessment. Thus, an improved calibration framework for DRASTIC should be developed and suggested for spatial optimization.

2. Optimal grid cell size for intrinsic aquifer assessment and aquifer hazard assessment should be determined based on the resolution of input data. Using different resolution data (various resolution data) in various study areas, optimal grid cell size should be suggested for each of aquifer systems.

3. Estimated nitrate loads and baseflow data from LOADEST and WHAT were used as observed nitrate loads and baseflow data, respectively, because of limited observed nitrate loads and baseflow data when calibrating nitrate loads and baseflow at each outlet. For more accurate calibration of nitrate loads and baseflow, additional observed nitrate loads and baseflow are needed.

4. Even though the model (C4: the model with simultaneous streamflow and baseflow calibration) performance with baseflow calibration was acceptable, baseflow calibration in low flow conditions needs to be improved because simulated baseflow did not replicate observed baseflow in the lowest flow conditions. Thus, baseflow calibration should be conducted by separately clustering high flow and low flow conditions for more accurate baseflow estimation.

5. The EFCR was only applied in the UWRW. The EFCR should be applied in more watersheds and more uncertainty analysis of baseflow-related parameters should be conducted for better baseflow estimation.

6. More observed nitrate concentrations in wells > 2 ppm (17.1% of total nitrate database) are required for better prediction of nitrate detections > 2 ppm. The models used in this dissertation are data-driven models (i.e., DRASTIC and Geo-ANN). Therefore, if more well nitrate concentration data are available, the approaches suggested in this study could potentially be improved, and better results would likely be provided because calibration and validation of DRASTIC or Geo-ANN using more observed nitrate concentrations in wells > 2 ppm could be performed more accurately.

7. Various machine learning algorithms (e.g., convolution neural network, Bayesian linear regression, and decision forest regression) for aquifer hazard assessment should be tested to better predict elevated nitrate levels in aquifers.

APPENDIX

APPENDIX

Table A Nitrate concentrations in wells in the UWRW

ID	Station	NO3 (mg/L)	Lat	Long
1	HDB	0.10	39.903	-86.324
2	HDB	0.21	39.588	-86.429
3	NWQMC	4.95	40.066	-85.939
4	NWQMC	1.79	40.164	-85.261
5	HDB	0.10	39.534	-86.351
6	HDB	0.10	39.916	-86.292
7	HDB	0.10	40.228	-85.563
8	HDB	0.10	40.086	-85.960
9	HDB	0.10	40.035	-86.086
10	HDB	0.26	39.955	-85.996
11	HDB	10.42	40.048	-86.196
12	HDB	0.12	40.077	-85.733
13	HDB	0.10	39.686	-86.263
14	HDB	0.10	40.116	-86.001
15	HDB	0.10	40.268	-86.039
16	HDB	0.10	40.160	-85.730
17	HDB	0.10	39.987	-85.902
18	HDB	1.78	39.933	-86.118
19	HDB	0.10	40.006	-86.013
20	HDB	0.25	40.094	-85.982
21	HDB	0.10	40.005	-86.225
22	HDB	0.10	40.081	-85.986
23	HDB	0.10	40.142	-85.830
24	HDB	0.10	39.958	-86.226
25	HDB	0.10	39.441	-86.398
26	NWQMC	6.00	39.952	-86.067
27	HDB	0.10	40.174	-85.831
28	HDB	0.10	40.000	-85.887
29	HDB	0.10	40.145	-86.080
30	HDB	0.10	40.046	-85.534
31	HDB	0.10	40.050	-85.863
32	HDB	0.10	40.044	-86.026
33	HDB	0.10	39.962	-85.775
34	HDB	5.53	40.127	-85.964

35	HDB	0.10	40.093	-85.318
36	HDB	0.10	39.916	-86.174
37	HDB	0.10	40.018	-86.192
38	HDB	0.10	40.040	-85.968
39	HDB	0.34	40.057	-85.345
40	HDB	0.10	39.670	-86.152
41	HDB	0.12	40.278	-85.628
42	NWQMC	10.00	39.788	-86.161
43	HDB	0.10	40.196	-85.289
44	HDB	0.10	40.036	-85.362
45	HDB	1.50	39.614	-86.436
46	NWQMC	0.29	40.066	-85.939
47	HDB	0.12	40.116	-85.633
48	HDB	0.10	40.068	-85.963
49	HDB	0.11	39.917	-86.311
50	HDB	0.10	40.065	-85.737
51	NWQMC	5.36	39.952	-86.067
52	HDB	1.72	39.570	-86.311
53	HDB	0.13	40.019	-85.531
54	HDB	0.10	40.091	-86.093
55	HDB	4.23	40.067	-86.000
56	HDB	0.10	40.333	-85.525
57	HDB	0.10	39.692	-86.058
58	HDB	0.10	40.214	-85.429
59	HDB	0.11	40.081	-85.753
60	HDB	0.10	40.031	-85.688
61	HDB	0.60	39.393	-86.592
62	HDB	0.10	40.165	-85.861
63	HDB	0.10	40.125	-85.930
64	HDB	0.10	40.094	-85.301
65	HDB	0.10	39.967	-86.060
66	HDB	1.19	39.494	-86.406
67	HDB	12.84	39.998	-85.735
68	HDB	0.10	39.846	-86.323
69	HDB	0.10	39.940	-86.114
70	HDB	0.10	40.270	-86.024
71	HDB	0.10	39.499	-86.251
72	HDB	0.10	40.083	-85.811
73	HDB	6.64	40.124	-85.848
74	HDB	0.11	40.105	-85.754

75	HDB	0.10	40.224	-85.828
76	HDB	1.86	40.145	-85.448
77	HDB	3.47	40.061	-86.004
78	HDB	0.17	39.596	-86.404
79	HDB	0.11	40.080	-85.834
80	HDB	0.10	40.158	-85.378
81	HDB	7.13	39.636	-86.292
82	HDB	0.10	40.291	-85.765
83	HDB	11.98	40.027	-85.380
84	HDB	0.10	40.166	-85.994
85	HDB	1.22	39.592	-86.313
86	HDB	0.10	40.030	-86.106
87	HDB	0.10	40.245	-86.034
88	HDB	0.10	40.275	-85.472
89	HDB	0.10	39.672	-86.142
90	HDB	0.10	39.977	-85.665
91	HDB	0.10	40.110	-85.504
92	HDB	0.79	40.086	-86.077
93	HDB	4.39	39.639	-86.291
94	HDB	0.10	40.269	-85.407
95	HDB	5.21	40.101	-85.313
96	HDB	0.10	39.657	-86.188
97	HDB	0.10	39.966	-85.866
98	HDB	0.16	40.229	-85.708
99	HDB	0.11	39.966	-85.748
100	HDB	0.10	39.932	-85.969
101	HDB	0.10	40.049	-86.041
102	HDB	0.73	39.960	-85.931
103	HDB	0.48	39.896	-86.010
104	HDB	2.31	39.583	-86.489
105	HDB	6.71	39.630	-86.331
106	HDB	0.10	39.932	-86.213
107	HDB	0.10	40.312	-86.103
108	HDB	0.13	40.246	-85.614
109	HDB	0.10	39.953	-86.122
110	HDB	0.10	39.993	-85.901
111	HDB	0.10	39.974	-85.733
112	HDB	3.14	39.961	-85.917
113	HDB	0.10	40.285	-86.138
114	HDB	0.10	40.059	-85.959

115	HDB	0.11	40.286	-85.800
116	HDB	0.12	40.326	-85.652
117	HDB	7.56	40.123	-85.949
118	HDB	0.33	40.183	-85.956
119	HDB	0.10	39.846	-86.251
120	HDB	0.10	40.302	-85.573
121	HDB	0.10	39.837	-86.266
122	HDB	0.10	40.017	-85.738
123	HDB	0.10	40.065	-85.290
124	HDB	0.10	40.132	-85.243
125	HDB	7.01	40.076	-85.980
126	HDB	0.10	40.156	-85.696
127	HDB	0.10	39.596	-86.386
128	HDB	0.10	40.113	-86.026
129	HDB	0.10	40.055	-86.220
130	NWQMC	0.42	39.866	-86.287
131	HDB	0.10	40.172	-85.266
132	HDB	0.10	40.008	-86.112
133	HDB	0.10	39.613	-86.283
134	HDB	0.10	39.967	-86.058
135	HDB	0.10	40.333	-85.634
136	HDB	3.37	39.997	-86.019
137	NWQMC	2.19	39.911	-86.113
138	HDB	5.00	40.037	-86.046
139	HDB	0.11	39.983	-85.673
140	HDB	0.10	40.061	-86.196
141	HDB	0.10	40.076	-86.007
142	HDB	0.10	40.074	-85.911
143	HDB	0.10	40.238	-86.003
144	HDB	0.10	40.073	-85.424
145	HDB	0.10	39.653	-86.147
146	HDB	0.10	40.141	-86.087
147	HDB	0.10	39.434	-86.403
148	HDB	0.11	40.299	-85.856
149	HDB	0.10	40.033	-85.769
150	HDB	0.98	39.999	-86.067
151	HDB	0.15	39.907	-86.138
152	HDB	0.80	40.200	-85.457
153	HDB	2.77	39.601	-86.356
154	HDB	0.16	39.570	-86.431

155	HDB	0.10	40.100	-86.124
156	HDB	16.52	40.085	-86.082
157	HDB	0.10	40.013	-85.950
158	HDB	0.10	40.127	-86.085
159	HDB	0.10	39.907	-86.288
160	NWQMC	4.40	39.832	-86.145
161	HDB	0.70	40.121	-85.819
162	HDB	0.10	40.156	-86.221
163	NWQMC	9.43	40.066	-85.939
164	HDB	4.88	39.647	-86.106
165	HDB	0.10	40.086	-86.200
166	HDB	0.10	39.633	-86.295
167	HDB	0.10	39.980	-86.049
168	HDB	0.18	39.901	-85.926
169	HDB	0.10	39.696	-86.128
170	HDB	0.10	39.957	-85.769
171	HDB	0.10	40.178	-85.748
172	HDB	0.12	40.150	-85.737
173	HDB	4.56	39.490	-86.302
174	HDB	0.10	40.078	-85.988
175	HDB	0.11	39.951	-85.766
176	HDB	0.10	40.193	-86.092
177	HDB	0.11	40.252	-85.812
178	HDB	0.10	40.292	-85.994
179	HDB	0.10	39.876	-86.263
180	HDB	0.10	40.071	-85.942
181	HDB	5.22	40.201	-85.487
182	HDB	0.10	40.158	-85.901
183	HDB	0.67	39.495	-86.410
184	HDB	0.11	40.294	-85.717
185	HDB	0.10	40.060	-86.031
186	HDB	6.48	40.067	-85.604
187	HDB	0.10	39.674	-86.176
188	HDB	1.00	40.016	-85.649
189	HDB	0.11	40.011	-85.747
190	HDB	1.70	39.959	-85.915
191	HDB	0.10	39.971	-85.650
192	HDB	0.10	40.182	-85.694
193	HDB	0.10	40.254	-85.445
194	NWQMC	3.60	39.405	-86.460

195	HDB	0.10	39.422	-86.607
196	HDB	0.14	40.144	-85.602
197	HDB	0.63	39.879	-86.098
198	HDB	0.11	40.068	-85.661
199	HDB	0.16	40.165	-85.693
200	HDB	0.10	40.248	-86.051
201	HDB	0.11	40.102	-85.765
202	HDB	0.10	40.013	-85.750
203	HDB	0.10	40.247	-86.034
204	HDB	0.10	39.903	-85.935
205	HDB	0.10	39.908	-86.151
206	HDB	0.80	39.357	-86.290
207	NWQMC	0.29	39.866	-86.287
208	HDB	0.10	40.010	-85.528
209	HDB	0.10	39.420	-86.389
210	HDB	0.10	40.251	-86.159
211	HDB	0.10	40.015	-85.993
212	HDB	0.11	40.204	-85.790
213	HDB	0.49	40.078	-85.415
214	HDB	0.10	39.671	-86.140
215	HDB	0.71	39.533	-86.417
216	HDB	0.10	39.635	-86.315
217	HDB	0.10	40.220	-85.498
218	HDB	0.10	40.301	-85.824
219	HDB	0.14	40.268	-86.035
220	HDB	0.10	40.108	-86.041
221	HDB	0.72	39.439	-86.397
222	HDB	0.10	40.208	-85.927
223	NWQMC	3.30	39.405	-86.460
224	HDB	0.10	40.189	-86.205
225	HDB	0.21	40.067	-85.532
226	HDB	0.10	40.106	-86.239
227	HDB	0.10	40.207	-85.664
228	HDB	0.10	39.619	-86.393
229	HDB	0.10	39.628	-86.343
230	HDB	0.10	39.543	-86.320
231	HDB	0.10	39.503	-86.290
232	NWQMC	8.20	39.604	-86.229
233	HDB	3.48	39.532	-86.518
234	HDB	0.10	40.010	-85.772

235	HDB	0.10	39.738	-86.322
236	HDB	0.10	40.107	-86.040
237	HDB	2.26	39.686	-86.153
238	HDB	0.10	40.078	-86.182
239	HDB	0.18	39.613	-86.440
240	HDB	0.10	39.945	-85.888
241	HDB	2.05	40.144	-85.440
242	HDB	0.10	39.567	-86.367
243	HDB	0.10	40.249	-86.034
244	HDB	5.38	39.472	-86.351
245	HDB	10.90	40.065	-86.216
246	HDB	0.10	40.192	-86.151
247	HDB	0.10	40.293	-86.155
248	HDB	0.35	40.178	-85.623
249	HDB	0.10	40.111	-85.419
250	HDB	0.47	39.414	-86.303
251	HDB	0.10	40.163	-85.998
252	HDB	0.10	40.106	-85.237
253	HDB	0.10	39.840	-86.264
254	NWQMC	9.18	40.121	-85.058
255	HDB	4.65	40.154	-85.440
256	HDB	0.10	39.978	-86.022
257	HDB	0.11	40.215	-85.717
258	NWQMC	1.20	40.145	-86.161
259	HDB	0.10	39.993	-85.902
260	HDB	0.10	39.924	-85.729
261	HDB	0.10	40.084	-86.196
262	NWQMC	1.60	40.111	-85.694
263	HDB	0.16	40.016	-85.831
264	HDB	1.02	39.953	-86.226
265	HDB	0.10	40.318	-86.121
266	HDB	0.10	40.001	-86.093
267	HDB	0.10	39.642	-86.120
268	HDB	4.30	40.045	-85.536
269	HDB	0.10	39.887	-86.089
270	HDB	0.10	39.413	-86.389
271	HDB	3.10	39.674	-86.185
272	HDB	0.10	39.632	-86.283
273	NWQMC	2.40	39.734	-86.223
274	HDB	0.10	39.443	-86.341

275	HDB	0.52	39.599	-86.430
276	HDB	0.11	40.064	-85.772
277	HDB	0.11	40.193	-85.947
278	HDB	0.10	39.995	-86.097
279	NWQMC	3.60	39.861	-86.160
280	HDB	0.10	40.236	-86.113
281	NWQMC	0.21	39.866	-86.287
282	HDB	0.10	39.982	-86.140
283	HDB	0.10	39.927	-86.216
284	HDB	0.10	39.715	-86.244
285	HDB	0.10	40.171	-85.988
286	HDB	0.38	39.919	-86.307
287	HDB	0.22	39.482	-86.320
288	HDB	0.10	39.912	-86.312
289	HDB	0.10	39.399	-86.501
290	HDB	0.10	39.999	-85.444
291	HDB	0.10	40.055	-85.588
292	HDB	0.15	39.506	-86.537
293	HDB	0.10	40.187	-86.004
294	HDB	2.37	40.180	-85.428
295	HDB	0.10	39.781	-86.043
296	NWQMC	8.22	40.329	-85.538
297	HDB	1.83	39.664	-86.164
298	HDB	0.10	40.213	-85.938
299	HDB	0.10	40.023	-85.976
300	NWQMC	5.79	39.771	-86.195
301	HDB	0.10	39.845	-86.323
302	HDB	0.10	40.047	-85.980
303	HDB	3.61	39.459	-86.410
304	HDB	3.21	39.578	-86.302
305	HDB	0.10	40.047	-85.536
306	NWQMC	2.25	40.066	-85.939
307	HDB	0.10	39.919	-85.745
308	HDB	0.10	39.956	-85.995
309	HDB	15.23	39.390	-86.599
310	HDB	0.10	40.300	-85.808
311	NWQMC	1.77	40.066	-85.939
312	HDB	0.11	40.145	-85.619
313	HDB	0.10	39.687	-86.172
314	HDB	0.10	40.104	-85.366

315	HDB	1.63	40.066	-86.000
316	HDB	0.10	39.992	-85.571
317	HDB	0.10	40.161	-85.993
318	HDB	0.10	40.221	-85.784
319	HDB	0.10	40.073	-85.501
320	HDB	0.10	40.144	-85.638
321	HDB	1.64	39.408	-86.500
322	HDB	0.10	40.050	-85.538
323	NWQMC	2.70	39.817	-86.117
324	HDB	0.10	40.126	-86.084
325	HDB	9.31	40.054	-85.869
326	HDB	0.12	40.333	-85.653
327	HDB	0.10	40.081	-86.024
328	HDB	0.87	39.869	-85.993
329	NWQMC	1.60	40.145	-86.161
330	HDB	0.10	39.409	-86.432
331	NWQMC	3.29	39.755	-86.165
332	HDB	0.12	40.235	-85.798
333	HDB	0.15	39.410	-86.403
334	HDB	0.10	40.287	-85.768
335	HDB	0.10	40.020	-85.962
336	HDB	0.10	39.959	-86.065
337	HDB	0.11	40.026	-85.601
338	HDB	4.81	39.591	-86.485
339	HDB	0.11	40.096	-85.792
340	HDB	17.96	40.133	-85.838
341	NWQMC	0.48	39.782	-86.224
342	HDB	0.10	40.314	-85.646
343	HDB	9.18	40.024	-85.689
344	NWQMC	6.04	40.066	-85.939
345	HDB	18.29	40.083	-86.074
346	HDB	0.10	39.923	-86.158
347	HDB	0.10	39.900	-86.134
348	HDB	0.10	39.849	-86.324
349	HDB	0.10	40.206	-85.691
350	HDB	0.10	40.010	-86.028
351	HDB	2.17	39.431	-86.383
352	HDB	0.10	40.114	-86.108
353	HDB	11.21	40.116	-85.855
354	HDB	0.18	39.476	-86.386

355	HDB	4.88	39.696	-86.168
356	HDB	0.10	40.154	-86.203
357	HDB	7.01	39.399	-86.267
358	HDB	0.10	39.762	-86.313
359	HDB	0.10	39.670	-86.156
360	HDB	7.53	39.528	-86.374
361	HDB	15.05	40.084	-86.073
362	HDB	0.11	40.223	-85.837
363	HDB	0.10	40.066	-85.497
364	HDB	0.10	40.122	-85.504
365	HDB	0.10	40.321	-85.625
366	HDB	0.18	39.436	-86.554
367	HDB	0.66	39.957	-86.090
368	HDB	0.18	40.282	-86.085
369	HDB	0.10	39.978	-86.049
370	HDB	1.50	39.433	-86.554
371	HDB	0.10	40.075	-85.987
372	HDB	0.10	39.981	-86.009
373	HDB	0.10	40.093	-85.391
374	HDB	0.30	39.584	-86.399
375	HDB	0.10	39.672	-86.080
376	HDB	0.10	40.301	-85.512
377	HDB	0.10	40.303	-85.512
378	HDB	0.10	39.341	-86.530
379	HDB	0.10	40.035	-85.386
380	HDB	0.12	40.090	-85.844
381	HDB	0.10	40.069	-85.574
382	HDB	0.10	40.187	-86.099
383	HDB	0.10	39.545	-86.325
384	NWQMC	4.80	39.869	-86.132
385	HDB	0.10	40.200	-85.968
386	NWQMC	0.37	39.866	-86.287
387	HDB	3.40	39.995	-86.019
388	HDB	0.10	40.064	-86.145
389	NWQMC	2.67	39.801	-86.269
390	HDB	0.10	39.561	-86.341
391	HDB	3.38	40.199	-85.482
392	HDB	0.10	39.897	-86.130
393	HDB	0.14	39.426	-86.303
394	HDB	0.10	40.156	-86.066

395	HDB	0.10	40.060	-85.566
396	HDB	0.11	39.985	-85.821
397	HDB	2.31	39.463	-86.276
398	HDB	0.10	40.278	-85.336
399	HDB	0.10	39.658	-86.061
400	HDB	0.10	40.029	-85.652
401	HDB	0.10	40.029	-85.878
402	HDB	0.10	40.088	-85.969
403	HDB	0.10	40.210	-86.027
404	HDB	0.13	40.276	-85.623
405	HDB	0.10	40.105	-85.462
406	HDB	4.57	40.200	-85.484
407	HDB	0.10	39.902	-86.299
408	HDB	1.18	39.498	-86.406
409	HDB	0.10	40.064	-86.003
410	HDB	0.10	40.231	-85.448
411	HDB	0.15	39.621	-86.436
412	HDB	0.17	40.011	-85.548
413	HDB	0.10	40.330	-85.655
414	HDB	0.10	39.845	-86.319
415	HDB	0.10	39.963	-86.068
416	HDB	0.10	39.837	-86.258
417	HDB	0.10	40.060	-86.224
418	HDB	0.39	39.710	-86.212
419	HDB	0.10	40.042	-85.326
420	HDB	0.10	39.852	-86.254
421	NWQMC	2.60	39.800	-86.213
422	HDB	0.10	40.046	-85.771
423	NWQMC	3.00	39.807	-86.155
424	HDB	0.10	40.326	-85.847
425	HDB	0.10	40.053	-85.514
426	HDB	10.24	40.017	-85.681
427	HDB	0.10	39.671	-86.056
428	HDB	3.74	40.060	-86.003
429	HDB	0.10	39.963	-85.553
430	HDB	0.14	39.542	-86.365
431	HDB	0.10	40.001	-86.092
432	HDB	0.10	40.030	-85.883
433	HDB	0.10	40.333	-85.755
434	HDB	0.10	39.678	-86.162

435	HDB	0.14	39.416	-86.403
436	HDB	0.10	39.653	-86.190
437	HDB	0.10	39.423	-86.401
438	HDB	0.10	39.979	-86.021
439	HDB	2.86	40.059	-86.032
440	NWQMC	5.20	39.511	-86.372
441	HDB	3.41	39.996	-86.018
442	NWQMC	10.00	40.135	-85.883
443	HDB	0.10	39.455	-86.411
444	HDB	0.11	39.988	-85.791
445	HDB	0.10	40.192	-85.989
446	HDB	0.10	40.273	-85.750
447	HDB	0.43	40.293	-85.351
448	HDB	0.10	40.304	-86.089
449	HDB	0.10	39.829	-86.211
450	NWQMC	0.89	39.788	-86.161
451	HDB	0.10	39.553	-86.258
452	HDB	0.10	40.256	-85.482
453	HDB	6.14	40.109	-85.957
454	HDB	0.10	40.288	-85.456
455	HDB	0.11	40.188	-85.755
456	HDB	0.10	40.045	-85.561
457	HDB	0.10	40.278	-85.965
458	HDB	0.10	40.345	-85.781
459	HDB	0.10	39.729	-86.051
460	HDB	0.10	39.526	-86.333
461	HDB	0.30	39.613	-86.298
462	HDB	0.14	39.606	-86.436
463	NWQMC	7.40	39.370	-86.477
464	HDB	0.10	40.104	-85.867
465	HDB	0.10	40.124	-85.992
466	HDB	0.10	39.866	-86.323
467	HDB	0.14	40.058	-85.283
468	HDB	0.10	40.074	-86.205
469	HDB	0.10	39.426	-86.607
470	HDB	0.10	40.123	-85.791
471	HDB	0.10	39.720	-86.314
472	HDB	0.10	39.682	-86.089
473	HDB	0.10	40.276	-85.364
474	HDB	0.10	40.186	-85.973

475	HDB	0.10	40.196	-85.457
476	HDB	0.10	40.290	-85.457
477	HDB	0.10	40.042	-86.024
478	HDB	0.10	40.053	-85.428
479	HDB	6.05	39.686	-86.176
480	HDB	4.09	39.631	-86.255
481	HDB	0.10	39.846	-86.319
482	HDB	2.87	39.914	-86.117
483	HDB	0.11	40.050	-85.666
484	HDB	12.73	39.576	-86.312
485	HDB	0.10	40.110	-85.983
486	HDB	0.12	40.025	-85.826
487	HDB	0.10	40.205	-85.981
488	HDB	8.69	39.456	-86.305
489	HDB	0.10	40.191	-86.203
490	HDB	7.04	39.607	-86.294
491	HDB	0.10	39.900	-85.928
492	HDB	0.10	40.144	-86.188
493	HDB	0.10	39.931	-85.910
494	HDB	0.10	39.949	-86.034
495	HDB	0.10	40.117	-86.067
496	HDB	0.10	39.878	-86.260
497	HDB	0.10	40.329	-85.852
498	HDB	0.10	39.960	-85.731
499	HDB	0.10	40.051	-86.196
500	HDB	0.10	40.261	-86.044
501	HDB	8.24	40.146	-85.829
502	HDB	0.10	39.568	-86.332
503	HDB	0.10	39.914	-86.314
504	HDB	0.21	39.465	-86.547
505	HDB	0.10	39.838	-86.312
506	HDB	0.10	39.738	-86.323
507	HDB	0.10	40.168	-86.106
508	HDB	0.10	40.278	-85.632
509	HDB	0.13	39.343	-86.576
510	NWQMC	2.69	40.066	-85.939
511	HDB	0.17	40.053	-86.016
512	HDB	0.10	40.092	-85.388
513	HDB	0.10	40.245	-85.418
514	HDB	0.11	40.067	-85.588

515	NWQMC	7.16	40.066	-85.939
516	HDB	0.10	39.992	-86.089
517	HDB	0.10	40.174	-86.034
518	HDB	0.10	40.243	-86.036
519	HDB	0.21	39.712	-86.231
520	HDB	0.10	39.913	-86.283
521	HDB	0.10	40.017	-85.877
522	HDB	0.10	40.066	-86.216
523	HDB	3.60	40.054	-86.036
524	HDB	0.10	40.134	-86.025
525	HDB	7.65	40.140	-85.324
526	HDB	6.96	39.413	-86.405
527	HDB	0.55	39.878	-86.090
528	HDB	0.10	40.070	-85.554
529	NWQMC	15.90	39.791	-86.056
530	HDB	0.14	40.056	-85.701
531	HDB	0.10	40.024	-86.212
532	HDB	0.10	40.081	-86.192
533	HDB	0.11	40.336	-85.731
534	HDB	0.10	40.108	-85.500
535	HDB	0.10	40.051	-85.508
536	HDB	0.12	39.982	-85.688
537	HDB	0.10	40.012	-85.549
538	HDB	0.10	39.906	-85.935
539	HDB	0.10	39.988	-85.780
540	HDB	0.10	39.995	-85.950
541	HDB	0.10	39.966	-86.064
542	HDB	10.19	39.459	-86.305
543	HDB	0.10	40.034	-85.973
544	HDB	0.10	39.910	-86.148
545	HDB	0.10	40.032	-86.126
546	HDB	0.12	40.154	-85.812
547	HDB	0.10	39.655	-86.106
548	HDB	0.10	40.284	-85.318
549	HDB	0.10	39.953	-85.914
550	HDB	0.12	40.102	-85.720
551	HDB	0.10	40.027	-85.825
552	HDB	0.11	40.050	-86.194
553	HDB	0.10	39.954	-86.233
554	NWQMC	8.10	39.405	-86.460

555	HDB	0.10	40.193	-85.992
556	HDB	3.28	40.073	-85.517
557	HDB	0.11	40.047	-85.613
558	HDB	0.10	40.143	-86.088
559	HDB	0.17	39.967	-85.686
560	HDB	0.12	40.203	-85.737
561	HDB	0.10	40.187	-86.102
562	HDB	0.19	40.311	-85.849
563	HDB	0.10	39.956	-85.959
564	HDB	0.10	40.058	-86.022
565	HDB	0.10	39.955	-85.999
566	HDB	0.10	40.074	-86.079
567	HDB	0.10	39.703	-86.167
568	HDB	0.10	40.220	-85.599
569	HDB	0.10	40.025	-85.563
570	HDB	9.20	40.124	-85.497
571	HDB	0.13	39.977	-85.661
572	HDB	0.60	40.088	-85.748
573	HDB	0.28	39.595	-86.479
574	HDB	0.10	39.675	-86.091
575	HDB	0.10	40.050	-85.779
576	HDB	0.10	40.082	-86.003
577	HDB	0.13	40.254	-85.978
578	HDB	0.10	40.063	-85.762
579	HDB	0.10	40.296	-85.866
580	NWQMC	9.78	40.174	-85.529
581	HDB	0.10	39.686	-86.053
582	HDB	0.10	40.117	-86.180
583	HDB	7.01	40.109	-85.955
584	HDB	0.10	39.971	-85.575
585	HDB	0.10	40.250	-85.415
586	HDB	0.14	40.031	-86.062
587	HDB	0.10	39.916	-86.168
588	NWQMC	2.60	39.869	-86.132
589	HDB	0.10	40.038	-86.229
590	HDB	0.10	39.647	-86.157
591	HDB	0.10	40.171	-85.976
592	HDB	0.16	39.877	-86.091
593	HDB	0.10	40.168	-86.190
594	HDB	0.10	39.957	-85.769

595	HDB	0.10	40.120	-86.070
596	HDB	0.10	39.671	-86.077
597	HDB	0.10	39.889	-85.987
598	HDB	0.10	40.181	-86.025
599	HDB	0.10	40.146	-85.880
600	HDB	0.10	40.062	-85.972
601	HDB	0.84	39.623	-86.324
602	HDB	0.10	40.316	-85.551
603	HDB	0.16	40.096	-85.607
604	HDB	0.12	40.245	-85.827
605	HDB	3.23	40.069	-86.000
606	HDB	5.87	40.061	-85.995
607	HDB	0.10	39.652	-86.166
608	NWQMC	0.39	39.866	-86.287
609	HDB	0.10	40.031	-86.101
610	HDB	0.10	39.953	-86.156
611	HDB	0.10	39.728	-86.310
612	HDB	0.10	39.507	-86.477
613	HDB	5.22	39.614	-86.300
614	HDB	0.10	40.001	-86.091
615	HDB	0.12	40.304	-85.706
616	HDB	0.10	39.978	-86.048
617	HDB	0.11	40.315	-85.752
618	HDB	0.10	39.953	-86.122
619	NWQMC	4.00	39.604	-86.229
620	HDB	0.10	39.970	-85.954
621	HDB	0.10	39.996	-86.021
622	HDB	0.10	40.221	-85.571
623	HDB	0.13	40.102	-86.136
624	HDB	0.10	39.994	-85.568
625	HDB	0.12	40.276	-85.625
626	HDB	0.10	39.900	-85.928
627	HDB	0.34	39.481	-86.370
628	HDB	0.10	40.265	-85.511
629	HDB	0.10	40.260	-85.632
630	HDB	0.10	40.158	-85.378
631	HDB	0.10	39.978	-85.438
632	HDB	7.98	39.696	-86.166
633	HDB	0.10	39.649	-86.145
634	HDB	4.67	39.660	-86.124

635	HDB	0.10	40.020	-85.880
636	HDB	0.10	40.212	-85.938
637	NWQMC	0.59	39.949	-86.321
638	HDB	0.12	40.236	-85.841
639	HDB	0.10	39.931	-86.111
640	HDB	2.14	40.086	-85.993
641	HDB	0.10	40.153	-86.205
642	HDB	0.47	40.111	-85.880
643	HDB	0.10	39.994	-86.060
644	HDB	1.71	39.608	-86.273
645	HDB	0.10	40.196	-85.993
646	HDB	2.43	39.394	-86.456
647	HDB	0.10	39.657	-86.068
648	HDB	0.10	40.050	-86.193
649	HDB	0.10	40.153	-85.792
650	HDB	0.10	40.362	-85.750
651	HDB	0.11	40.290	-85.803
652	HDB	0.10	40.277	-85.705
653	HDB	17.02	40.135	-85.834
654	NWQMC	5.00	39.670	-86.213
655	NWQMC	0.59	39.779	-86.216
656	HDB	0.10	39.925	-86.146
657	HDB	0.10	40.180	-86.015
658	HDB	0.10	40.360	-85.753
659	HDB	0.10	40.156	-85.437
660	HDB	0.10	40.146	-85.933
661	NWQMC	5.01	39.952	-86.067
662	HDB	3.23	39.994	-86.019
663	HDB	0.11	40.024	-85.798
664	HDB	0.10	39.666	-86.203
665	NWQMC	8.00	39.405	-86.460
666	HDB	5.91	40.034	-86.041
667	NWQMC	4.40	39.775	-86.149
668	HDB	0.10	39.982	-85.556
669	HDB	0.35	40.279	-85.987
670	HDB	0.10	40.075	-85.373
671	HDB	0.10	40.005	-86.008
672	HDB	0.10	40.025	-85.398
673	HDB	1.46	39.684	-86.191
674	HDB	0.10	40.038	-85.525

675	HDB	0.34	39.681	-86.246
676	HDB	0.10	40.059	-86.024
677	HDB	0.10	39.403	-86.258
678	HDB	0.10	40.050	-86.217

[1]HDB: Heidelberg University, OH

[2]NWQMC: National Water Quality Monitoring Council (http://www.waterqualitydata.us/)

VITA

VITA

Won Seok Jang was born in Seoul, the Republic of Korea (South Korea). He graduated with a Bachelor of Engineering (B.Eng.) degree in Mechanical Engineering from Sogang University, Seoul, South Korea in 2008.

He received his Master of Science (M.Sc.) degree in Agricultural Engineering from Kangwon National University, Chuncheon, South Korea in 2010.

He joined the graduate program in Agricultural and Biological Engineering at Purdue University in August, 2011 and received Doctor of Philosophy (Ph.D.) degree in Dec, 2016.