The Development and Assessment of A Spatial Decision Support System for Watershed Management in the Niantic River Watershed: A Geodesign Approach

by

Melinda Shimizu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved August 2014 by the
Graduate Supervisory Committee:

Elizabeth Wentz, Chair
Arthur Gold
Robert Pahle
Craig Kirkwood

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

This dissertation advances spatial decision support system development theory by using a geodesign approach to evaluate design alternatives for such systems, including the impacts of the spatial model, technical spatial data, and user interface tools. These components are evaluated with a case study spatial decision support system for watershed management in the Niantic River watershed in Connecticut, USA. In addition to this case study, this dissertation provides a broader perspective on applying the approach to spatial decision support systems in general. The spatial model presented is validated, the impacts of the model are considered. The technical spatial data are evaluated using a new method developed to quantify data fitness for use in a spatial decision support system. Finally, the tools of the user interface are assessed by applying a conceptual framework and evaluating the resulting tools via user survey.

DEDICATION

This dissertation is dedicated to my family, because I certainly could not have done it without your support, understanding, encouragement, and patience.

Erik – I'm pretty sure I thank you for not letting me quit. Camden – I hope that summer session of Intro Geography you suffered through in utero does you some good one day. Audrey – I'm sorry that your first year of life was in competition with the final year of my dissertation. Both were Events that required great time and attention. I hope you weren't neglected too badly. If it helps, I'm sure that 10 years from now...I won't remember much of either.

47526.

ACKNOWLEDGMENTS

I would like to acknowledge the support of my committee; each of you provided invaluable support in different ways. My dissertation is assuredly better for your guidance. I would like to especially acknowledge the support of my chair, Libby Wentz – thank you for having high standards and your guidance to meet them.

I would like to also acknowledge the support of colleagues, without whom this dissertation could not have been completed. There are many of you, but most notably Joanna Merson, Jay Laura, Q Kellogg, Scott Kelley, and Jill Elder.

I would also like to acknowledge our office support staff – Kathleen, Laura, and Sue: you have been so helpful and it is truly appreciated.

Finally, I would like to acknowledge the financial support of ASU, and the USDA - CSREES - National Integrated Water Quality Program.

This material is based upon work supported in part by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under Agreement No. 2007-51130-03873. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

The goal of this dissertation is to advance spatial decision support system development theory by using a geodesign approach to consider the impacts of design alternatives for such systems. Geodesign is an emerging field with a focus on a design approach to development of spatial tools. This dissertation considers the impacts of three different design aspects for spatial decision support system development: the spatial model, the technical spatial data, and the user interface tools. The remainder of the introductory chapter provides a historical perspective on decision support system development, an overview of the general research questions addressed by this dissertation, and a description of the overall dissertation structure.

1.1 Historical Perspective

Decision support systems (DSS) have been developing and evolving since the mid-1960s. This evolution has been unstructured, and strongly influenced by technological advances (Power 1999). For example, the development of geographic information systems (GIS) and later the Internet provided new venues and new capabilities for decision support systems leading to the development of spatial decision support systems (SDSS) and web-based DSS, respectively.

In addition to technological advances, and perhaps more critically, new theoretical perspectives have had a strong impact on DSS design. Geodesign is a relatively new perspective for the development of spatial tools that focuses on the impacts of design alternatives and their impacts on the use of such tools (Steinitz 2012, Batty 2013).

This perspective is but the latest in a history of theoretical developments that influence decision support system design and evolution. Much of the early theory developed in the 1970s and 1980s focused on defining DSS and establishing a framework from which to understand them (Keen 1977, Sprague 1980, Power 1999). While some of these efforts focused on defining what functionalities are required for something to be called a DSS, others considered any data processing that provides information for decision making a form of decision support (Jankowski 2006). Nonetheless, DSS are typically defined as systems that improve the effectiveness of individual (Densham and Armstrong 1987) and/or groups of (Armstrong 1993) decision makers where the decision alternatives, outcomes, and evaluation criteria are not initially explicitly known (Arnott & Pervan 2005, Jankowski 2006).

The 1980s and 1990s followed with the introduction and integration of Geographic Information Systems (GIS) to generate *spatial* decision support systems (SDSS). GIS has a proven history of providing tools that facilitate spatial data management, visualization, and analysis (Goodchild 1992, Coppock and Rhind 1991, Andrienko et al., 2007). These capabilities, combined with DSSs, provide decision support for a wide range of subjects that include a spatial element.

The development of such systems depends upon the selection of several key components including the model, spatial technical data, and perceptual data, which includes probabilities and human judgments (Jankowski 2006). The system that results from different choices of components provides a different decision support experience for the end users. Therefore, careful consideration should be given in selecting each of these components when developing a SDSS.

An additional consideration is the choice of tools included in the SDSS. In the mid-1990s to 2000s, SDSS research began to consider the usability of decision support tools (Crossland 1995, Gregory and Keeney 2002, Jankowski 2006). Some of these efforts focused on human-computer interaction and trying to understand which software tools and interfaces are most usable (Nyerges 1993, Blaser et al 2000). Further, a concurrent paradigm shift in decision science promotes a change in focus that puts more emphasis on supporting better decision outcomes instead of describing decision alternatives (Keeney 1992, 1994, 1996). For SDSS to focus increased support on outcomes, adjustments in the types of tools and visualizations used in SDSSs are needed (Andrienko et al 2007).

The latter-1990s to present day have brought increased attention to public participation or participatory GIS and SDSSs, sometimes called collaborative GIS. This development pushes for collaboration with end users and decision makers in the SDSS design process (Couclelis 2005, Nyerges 2006, Ramsey 2009). Emphasis on a collaborative approach is especially relevant for decisions that are highly contentious (Gregory and Keeney 1994, 2002, Ramsey 2009).

Most recently, the emerging field of geodesign is promising for the development of even more responsive tools and decision support that consider spatial decisionskj (Steinitz 2012, Batty 2013). Esri defines geodesign as follows: “Geodesign combines geography with design by providing designers with robust tools that support rapid evaluation of design alternatives against the impacts of those designs.” The 2013 Geodesign Summit meeting emphasized geodesign as a way to bring science into decision-making, and collaboration and sharing into the design and adaptation of GIS tools in a variety of applications.

1.2 Goals and specific research questions

The geodesign approach to the design, implementation, and evaluation of a SDSS considers the impacts of design alternatives in the development of a SDSS. While most current research on SDSS design tends to be case-centric, this dissertation includes a broader perspective. The design, implementation, and evaluation are illustrated with a case study, but the techniques presented in each of the chapters are generic enough to be applied to other contexts.

The three overarching research questions in this dissertation are: 1) what impact does the spatial model have on a SDSS, 2) what impacts do technical spatial data have on a SDSS, and 3) what impacts do tools in the user interface have on a SDSS? Each of these questions in addressed in a separate chapter, and when considered together reflect a geodesign approach to considering the impacts of design alternatives in the selection of model, data, and tools for a SDSS.

1.3 Dissertation structure

This dissertation is organized as five chapters, including this introduction and a final conclusion chapter. The middle three chapters are written as distinct but related papers that consider respectively the design impacts of the spatial model, the spatial data, and the suite of tools on the case study SDSS developed for this project. Each paper uses the same case study of water management decision support for illustration purposes.

Chapter 2 presents the development and validation of a spatial model for nonpoint source nitrogen pollution. The model is applied to a case study in the Niantic River watershed in Connecticut, USA, but has the potential to be applied elsewhere. Model validation is achieved by comparing model estimates to measurements made by the United States Geological Survey (USGS). The potential sources for and implications from model variation are considered. This chapter addresses several research questions: 1) how to implement a spatial model of nonpoint source pollution and interaction within a watershed, 2) how valid such a model is based on measurements of pollution within the watershed, and 3) what impacts the model design and validation have upon a SDSS that incorporates the model.

Chapter 3 considers the technical spatial data used in the spatial model presented in Chapter 2. The model relies on spatial data for pollution sources, pollution sinks, and watershed morphology. Several different sources of data are freely available for each of these data requirements, and the research questions in this chapter include: 1) how to assess the fitness for use of different data sets in a SDSS design, 2) which data set for pollution sinks in the Niantic River case study is most fit for use in the case study SDSS, and 3) what impact does different data options have on a resulting SDSS? Most current considerations of data fitness for use tend to be very case-centric, and we present a new assessment method (AAA_Q) that is general enough to be applied to a variety of data types and uses. This method is then demonstrated using the spatial technical data driving the spatial model used to generate our case study SDSS.

Chapter 4 presents a conceptual framework for SDSS design that focuses on the tools implemented in the user interface. The design framework is based on a framework for decision-making presented by Howard (1968, 2007) that focuses on three different aspects of decision-making: what you know about the problem, the alternatives you want to consider, and the values placed on those alternatives. The research questions addressed in this chapter include: 1) what does a conceptual framework for classifying SDSS design look like, 2) how can this framework be applied to SDSS design, and 3) what are the impacts of using this framework on SDSS development.

The concluding chapter presents an overview of the work in the dissertation synthesized from each of the three research initiatives in the earlier chapters and how these studies integrate into applying a geodesign approach to SDSSs. It also addresses the project implications and limitations, and concludes with a section on future research directions.

CHAPTER 2

NSINK: A GEOSPATIAL INTERACTION MODEL OF NITROGEN LOADING AND WATERSHED DENITRIFICTION

2.0 Contribution to the dissertation

This paper assesses the modeling approach used to develop the water quality interactions in the Spatial Decision Support System (SDSS) outlined in this dissertation and is intended for publication in *Ecological Modeling or Environmental Modeling and Software.* It is co-authored by Melinda Shimizu, Elizabeth A. Wentz, Joanna Merson, Arthur J. Gold, and Dorothy Q. Kellogg. The modeling approach is validated by comparison to USGS stream quality data measured within the watershed. Both the approach and validation are an important part of the geodesign approach to SDSS development, and are described in this paper.

2.1 Introduction

The goal of this paper is to implement the geospatial approach to modeling watershed denitrification at the local level using: (a) widely available geospatial data, (b) current findings from peer-reviewed literature, and (c) USGS stream gauge data. We implement and test a full scale watershed model called NSink based on nitrogen modeling methods introduced by Kellogg et al (2010). In this paper, we describe the model equations, software environment, data needs, model output, and software validation. NSink use is demonstrated and evaluated on a case study in the Niantic River watershed in Connecticut.

Such a model is particularly important in coastal New England where eutrophication due to excess nitrogen in estuaries is of concern (Conley et al, 2009; Howarth et al, 2000). Eutrophication is increased growth of plankton and algae with the adverse effect of oxygen reduction in the affected waters. This oxygen reduction causes natural habitat degradation and death for native shellfish and fish (Ryther and Dunstan, 1971).

Nitrogen (N) occurs within the watershed both naturally and as a man-made pollutant. N pollution (sources) in the study area comes primarily from septic system based residential and farmland land uses. N is removed (sinks) from the watershed through natural micro biotic processes (microorganism based denitrification) within streams, wetlands, and ponds, or through best management practices (BMP) including septic system design (Oakley et al, 2010), stormwater retention basins (Collins et al, 2010), and denitrifying bioreactors (Schipper et al, 2010).

The interaction between the N sources and the N sinks is inherently spatial; the location of these sources and sinks plays a key role in effecting how much nitrogen input in the watershed is removed and how much is ultimately delivered to the sensitive estuaries at the outlet of the watershed. Consequently, it is important for decision makers to understand the impacts of land use change within their watershed in a spatial context. The software described herein is able to support this understanding with details at the local land use level.

Both the implementation and validation of the model are part of applying the geodesign process to spatial decision support system development. In this example, the NSink model is included as part of a spatial decision support system called NitroSim, which supports watershed management and decision making for the Niantic River watershed in Connecticut, USA (IN PREP). The geodesign process considers the development of such a spatial tool in terms of the impacts of design alternatives (Steinitz 2012, Batty 2013, Tulloch 2013).

2.2 The NSink approach

This section describes how NSink models the addition, transport, and retention of N within a watershed. The NSink approach to N modeling assumes that the interaction is isolated primarily to surface watershed processes, with the exception of riparian wetland interactions. NSink currently models sources as excess nitrogen from two non-point source types, residential and agriculture. At this time, other sources such as atmospheric deposition and forest biomass are not considered. N transport occurs along flowpaths from sources (or potential sources) to the watershed outlet, defined by elevation fluxes. Denitrification occurs as flowpaths intersect nitrogen sinks, such as wetlands, ponds, or streams. N sinks remove nitrogen from the watershed through biotic processes that convert inorganic N into organic N biomass and N gases (Gilliam, 1994; Hill, 1996, Gold et al, 2001; McClain et al, 2003, Kellogg et al, 2010). The specifics of each step relevant to NSink are described here.

2.2.1 Nitrogen sources

In the case study examined here, we have limited nitrogen loading to two nitrate-nitrogen (nitrate-N) source types, typical in Southern New England watersheds (Gold et al. 1990). One type is agricultural parcels with manure-fertilized crops. These areas are estimated to contribute nitrogen at a rate of 53.7 lbs/acre/year, due primarily to excess fertilization (Gold et al. 1990). This value will vary due to different farming practices and crop types, but our case study uses the 53.7 lbs/acre/year value. Residential parcels contribute to watershed nitrogen loading from septic systems at a rate of 24 lbs/acre/year (Gold et al. 1990, Dept. of the Environment, Maryland). This rate is based on roughly 1 acre lots with 3 persons per lot. Different population densities result in different loading values. In practice, an NSink user can designate many types of loading rates to fit their needs.

2.2.2 Nitrogen sinks

Nitrogen removal or denitrification occurs when a flowpath carrying nitrogen enters a nitrogen sink, such as lakes, ponds, streams, and or riparian zones. In this treatment, ponds and lakes are modeled based on their dimensions and expected water retention times. Nitrogen sinks remove the nitrate-nitrogen through a micro biotic process that converts dissolved N to N gas, which is then released to the atmosphere. Kellogg et al. (2010) provides the removal rates of nitrogen that are used in NSink, as described here. Nitrogen removal is calculated as a percentage from lakes and ponds through a relationship between reservoir and drainage area dimensions (Eq. 1). Nitrogen removal as a percentage from streams is calculated using stream and drainage area dimensions (Eq 2).

$$\text{Reservoir N removal (\%)} = 79.24 - 33.26 \times \log_{10} (D/T) \qquad (1)$$

Where: D = average reservoir depth = Volume (km^3) / A_r (km^2)
T = residence time (years) = Volume (km^3) / Q_{yr}
$Q_{yr} = A_d \times Q_{norm} \times 0.031536$
D/T [meters/year] = $Q_{yr}/Ar \times 1000 = Q_{norm} \times (A_d/A_r) \times 31.536$
$Q_{norm} = 0.0224$
A_d = Pond drainage area (km^2)
A_r = Pond surface area (km^2)

$$\text{Stream N removal (\%)} = (1-\exp(-\theta_{S1} D^{\theta_{S2}} T)) \times 100 \qquad (2)$$

Where: θ_{S1} = 0.0513 meters/day
θ_{S2} = -1.319
$D = 0.2612Qa^{0.3966}$
$D'a = Da^{1.25} \times \text{sqrt}(g))/Qa$
$Q = Qa = Q_{norm} \times \text{Drainage Area}$
T = reach length (m)/ V (meters/day)
$Q_{norm} = 0.0224$

Nitrogen removal as a percentage from riparian wetland areas is calculated using the width of the flowpath through the riparian wetland (Table 1). The removal percentage is based on the length of the path within the wetland as illustrated in Figure 1.

Table 1. Nitrogen removal rates in riparian wetlands.

Landcover	**Width (m)**	**% Removal**
Vegetated hydric soils	< 5	0
	5 -15	40
	15 – 30	60
	> 30	80

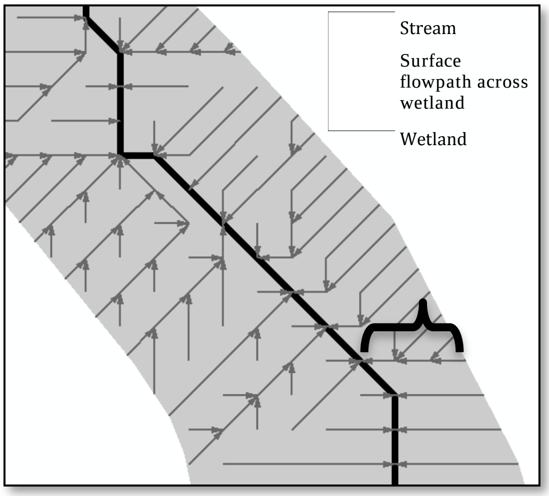


Figure 1. Wetland width.

2.2.3 Geospatial interaction

Geospatial interaction refers to the addition or subtraction of nitrogen along a flowpath from a source or potential source (non-sink location), which adds N, as it either flows over land where no nitrogen is exchanged, or through nitrogen sinks where nitrogen is subtracted, and ending at the watershed outlet. As the flowpath containing nitrogen intersects a sink, a percentage of nitrogen is removed using the calculations above based on the type of sink. Flowpaths are derived from elevation surfaces as least-costs paths. Each source location (as identified as an existing source or a potential source) in the watershed has a single and unique flowpath from the source to the outlet.

Two sample flowpaths and the associated calculations are illustrated in Figures 2 and 3. Figure 2 shows flow that moves from the land surface, through a wetland, into a stream, then to the outlet. The red boundary delineates the catchment area for the stream segment, needed to calculate the N removal occurring within the stream segment. The associated table identifies the removal rate as the flowpath intersects land, the wetland, and then the stream. Figure 3 and the associated table demonstrates flow through a pond, into a stream, and then to the outlet. The software manages the segments by maintaining a unique id per segment (hydroid), the nitrogen removal along that segment (nremoval), and the id for the next segment toward the outlet (nextid).

Figure 2. Sample flow path through a pond and associated calculations of nitrogen transport. The numbers on along the path are the HydroID for each segment.

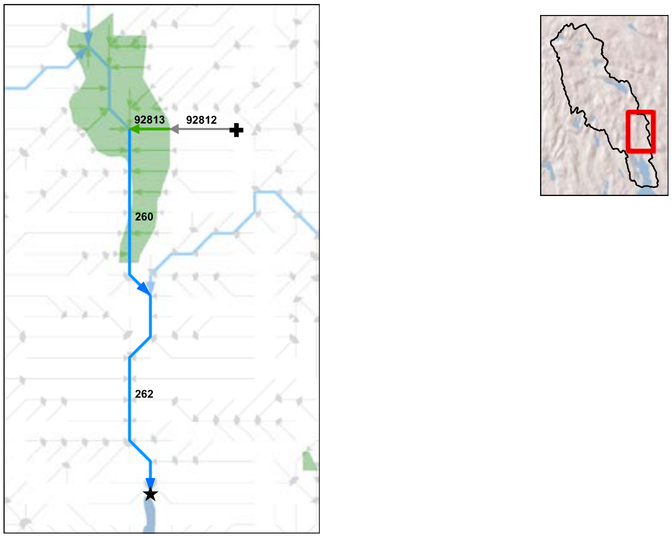


Figure 3. Sample flow path through a wetland and stream and the associated calculations for nitrogen transport. The numbers on along the path are the HydroID for each segment

Every 30 m x 30 m location in the watershed has a flowpath associated with it, each containing a unique id (hydroid). The model calculates the cumulative nitrogen removal from the start point of each flowpath to the watershed outlet. In other words, the model calculates how much of the nitrogen from a given point is removed before reaching the outlet. The remaining nitrogen ($N^{*}_{hydroid}$) delivered to the outlet is calculated as a percent with the following equation:

$$N^*_{hydroid\ j} = \prod_{j=1}^{n} (1 - r_j) \tag{3}$$

where:

n is the number of segments in flowpath *hydroid*
j is the index for a segment in a flowpath *hydroid*
r_j is the removal rate for segment j
hydroid is the index for the flowpaths

The result is a gridded map layer (30mx30m) with nitrogen remainder (%) that can be displayed visually to identify areas of high, medium, and low nitrogen contribution to the watershed outlet. The resulting data can also be used to calculate total nitrogen delivery to the outlet given the input load described in Section 3.2.1. To calculate total nitrogen delivery, we first classify cells based on the nitrogen load. The final nitrogen load calculation (in lbs/yr) is calculated with

$$NLoad = \sum_{n=1}^{hydroid} (41.7 * RES_n * N_n^*) + (53.7 * AG_n * N_n^*) \tag{4}$$

where

$$RES_n = \begin{cases} 1, & \text{residential} \\ 0, & \text{non residential} \end{cases} \tag{5}$$

$$AG_n = \begin{cases} 1, & \text{agricultural} \\ 0, & \text{non agricutural} \end{cases} \tag{6}$$

for all classified pixels (*n*) in the watershed boundary.

2.3 NSink implementation

This section describes the technical implementation of NSink including data requirements, software environment, module functions, and output. The specific implementation described here assumes nitrogen loading and removal were created for New England using the values described in Section 2.2. These could be customized for other regions based on empirical evidence.

2.3.1 Data requirements

Five spatial data layers are required for NSink, summarized in Table 2. The first is the watershed boundary, defined in our example by the NHD HUC-12. Watershed boundaries limit the extent of nitrogen sources and waterflow through the system. The second data layer is a 30-meter digital elevation model (DEM) acquired for our example from the USGS. The DEM is used to create least-cost paths from sources to the watershed outlet, identifying the flowpaths for geospatial nitrogen interaction. The third required data layer is land use to define the nitrogen sources. In our example we use the US National Land Cover Database (NLCD) from 2006, which contains seven categories that are redefined as nitrogen sources (Table 3). The fourth required layer defines the locations of reservoirs (lakes/ponds) and streams, for which our example uses the National Hydrology Dataset. These are reclassified as nitrogen sinks. The final dataset defines wetlands and in our example is from the United States Department of Agriculture (USDA) soil inventory. We extracted the hydric soils to delineate riparian wetlands as nitrogen sinks.

These specific data sources were selected because, compared to state-specific sources, they are available throughout the US northeast region. More detailed and potentially up-to-date data (e.g., land use for quantifying nitrogen sources) may be available at a local level.

Table 2. Technical spatial data used by the model are identified based on what part of the model the data contribute to, the source of the data, and the data type.

Description	Source	Role	Type
boundary	USGS National Hydrology Dataset (NHD) HUC 12	Delineation of watershed boundaries	Vector
flowpaths	USGS 30 m Digital Elevation Model (DEM)	Identification of least cost pathways through the watershed	Raster, 30 m cell
N sources	USGS National Land Cover Data (NLCD)	Residential and agricultural land use as nitrogen sources	Raster, 30 m cell
N sinks	National Hydrography Dataset for lakes and ponds and streams	Reservoirs including lakes and ponds; all streams	Vector 1:100,000 converted to raster 30 m cell
N sinks	Soil Survey Geographic (SSURGO) database	Hydric soils which are identified as riparian wetlands	Vector (1:12,000) converted to 30 m cell

2.3.2 Software environment and model parameters

NSink was developed as a customized extension to Esri's ArcGIS 10.0 software. NSink users are expected to have a working knowledge of the data manipulation and display functionalities of ArcGIS to effectively operate NSink. The individual modules are written as independent units comprised of a combination of model builder tools and python scripts. We utilize several of the ArcHydro data processing steps to prepare the digital elevation models (DEMs) for flowpath generation and calculation. Geospatial interaction between nitrogen sources and sinks are created in python scripts. Specific details on the modules are described in Appendix A.

2.3.3 Module functions

The four modules to generate NSink results are: 1) nitrogen sources and sinks identification, 2) flowpath generation, 3) source and sink geospatial interaction, and 4) nitrogen load calculation (Figure 4). This section explains the input, calculation, and output from each module.

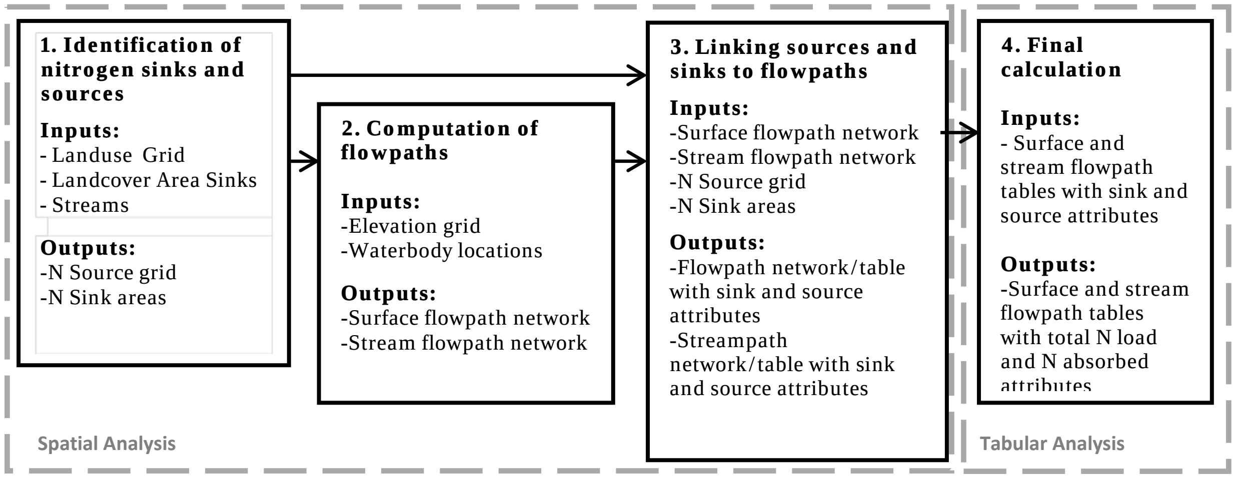


Figure 4. The model simulation is generated by a sequence of steps that are grouped into four separate models, identified here.

The first module creates layers identifying nitrogen sources and sinks. The inputs are land use, wetlands (hydric soils), and streams. The source layer is created by extracting residential and agricultural crops from the land use dataset and assigning the appropriate nitrogen load by parcel area (Table 3). The residential land use is further refined by selecting only those areas that use septic systems, instead of sewered area.

This is accomplished by overlay of a sewer layer from the State of Connecticut. Nitrogen load values (24 lbs N/acre/yr and 53.7 lbs N/acre/yr; Gold et al 1990, Dept. of the Environment, Maryland) are associated with each of the two source types. These values are converted to 2.42 kg N/30m cell/yr and 5.42 kg N/30m cell/yr respectively. A sinks layer is also generated that stores the locations and attributes for ponds/lakes and wetlands within the watershed.

Table 3. Nitrogen sources as classified in 2006 NLCD data set and the associated nitrogen load.

N Sources	**NLCD Classification 2001/2006**		**Nitrogen Load**[1](lbs N/acre/yr)	**Nitrogen Load**[1](kg N/30m cell/yr)
Developed Land	**22**	developed, low intensity	24	2.42
	23	developed, medium intensity		
Agricultural Land	**82**	row or cultivated crops	53.7	5.42
	82	small grains		

The second module generates the flowpaths throughout the watershed. The inputs are the DEM and the sinks layer generated in Module 1. This module uses several tools in ArcHydro to condition the DEM and generate stream paths based on the DEM. Cost path tools are used to generate flowpaths that occur overland. The outputs are the stream flowpath network and the overland flowpath network.

Module 3 links the flowpath networks generated in Module 2 with the source layer generated in Module 1. It also pre-calculates the nitrogen removal percentages on stream and pond segments of flowpaths and combines the two flowpath networks into one watershed flowpath network comprised of stream and overland flows. The inputs to this module are the stream flowpath network and overland flowpath network from

Module 2, the nitrogen source layer and the sinks layer generated in Module 1. The output is a single flowpath network with attribute information storing the type of flowpath (e.g., overland, stream, pond, wetland), the percent nitrogen removal for reservoirs and streams, and the associated source load for any paths originating at a source location.

The final module, Module 4, consists of a series of python scripts to process the flowpaths from origin to outlet and to calculate the removal due to wetland flow based on the width of the flowpath through the associated wetland (see equations in Table 1). The input to this module is the watershed flowpath network generated in Module 3. The outputs include an updated flowpath network for the watershed with complete nitrogen removal by percent for all flowpaths, the associated nitrogen load carried throughout the flowpaths, and a multipoint feature file and an associated grid file that delineate nitrogen delivery to the outlet by percent for each 30mx30m location within the watershed.

2.3.4 Output

NSink creates three spatial data layers of the geospatial interaction of nitrogen sources and sinks. The spatial data layers include: existing nitrogen sources and sinks; nitrogen contribution rates at all non-sink locations showing areas with high, medium, and low contribution to the estuary; and nitrogen load based on current land cover as delivered to the watershed outlet. Each of these can be manipulated and displayed as a watershed scale or local scale map.

In addition to visual displays of nitrogen, quantifiable results can be generated as tables for output to other statistical systems for manipulation or additional statistical analysis. For example, simple descriptive statistics on overall watershed nitrogen retention can be calculated.

2.4 Case study in the Niantic River watershed

NSink is demonstrated and evaluated through a case study in the Niantic River watershed (NRW) in coastal Connecticut. We selected this study area because of the impacts that nitrogen has on the region (Gold 1990), and the demonstrated local interest in nitrogen as a pollution source. In this section, we describe the specifics of the study area, the specific data sources we used, and the resulting output from NSink.

2.4.1 Niantic River watershed study area

The Niantic River watershed is situated between Hartford, Connecticut and Providence, Rhode Island (Figure 5). The area encompasses four towns: East Lyme, Waterford, Salem, and Montville. The US Census 2010 reports that there are over 62,000 people living in the region. Increasing residential growth has mostly occurred in the towns of Salem and East Lyme, which saw a 25% increase from 1990 to 2010 (Table 4).

The area is home to resident and seasonal waterfowl, shellfish, crustaceans, and finfishes. As of 2009, the Niantic River did not meet state water quality standards due to the observed degradation of aquatic life (Eastern Connecticut Conservation District). Shellfishing and swimming are now prohibited after a one-inch rainfall (Connecticut Department of Environmental Protection Office of Long Island Sound Programs) due to the resulting runoff. Research links degradation of aquatic life to an overload of nutrients, such as nitrogen (Marshall, 1994).

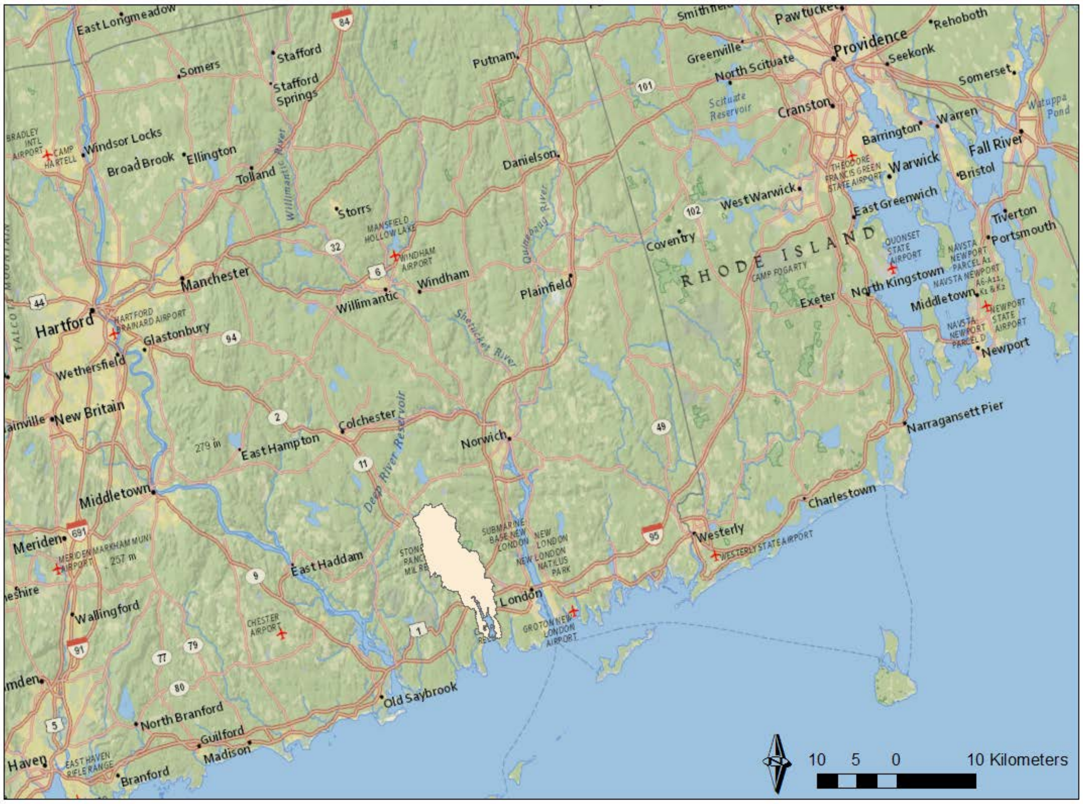


Figure 5. Location of the Niantic River Watershed, Connecticut, USA.

Table 4. Population of four towns that contain the Niantic River watershed.

Population for:	**2005**
East Lyme	18,459
Montville	19,612
Salem	4,094
Waterford	19,152
Total:	61,317

2.4.2 Data sources

The data sources for the Niantic River case study are summarized in Table 2 in Section 2.3.1. The first data source is NHD HUC-12, which defines the watershed boundary. The second data layer is a 30-meter digital elevation model (DEM) from the USGS. To define the nitrogen sources we use the 2006 US National Land Cover Database

(NLCD, which contains seven categories that are redefined as nitrogen sources (Table 3). The fourth required layer defines the locations of reservoirs (lakes/ponds) and streams, for which our example uses the National Hydrology Dataset. These are reclassified as nitrogen sinks. The final dataset defines wetlands and in our example is from the United States Department of Agriculture (USDA) soil inventory. We extracted the hydric soils to delineate riparian wetlands as nitrogen sinks.

2.4.3 NSink output

The NSink model generates a sequence of data layers and products culminating in the final flowpath network with nitrogen transport and removal. The input to Module 1 includes land use data, wetland and pond data, stream data, and sewered area data and the output of Module 1 includes layers identifying nitrogen sources and sinks (Figure 6). These areas are quantified in Table 5. Module 2 generates the flowpaths for the entire watershed, including overland flow and stream flow (Figure 7). Module 3 links these flowpaths with the sources and sinks identified in Module 1, in preparation for the final Nitrogen transport calculations generated by Module 4. Module 4 output includes flowpaths with information about nitrogen input and retention by sources and sinks, respectively. These flowpaths are converted into points corresponding with each 30mx30m cell for visual clarity (Figure 8). Figure 9 is a histogram showing the number of these cells that fall within each range of percent nitrogen delivery.

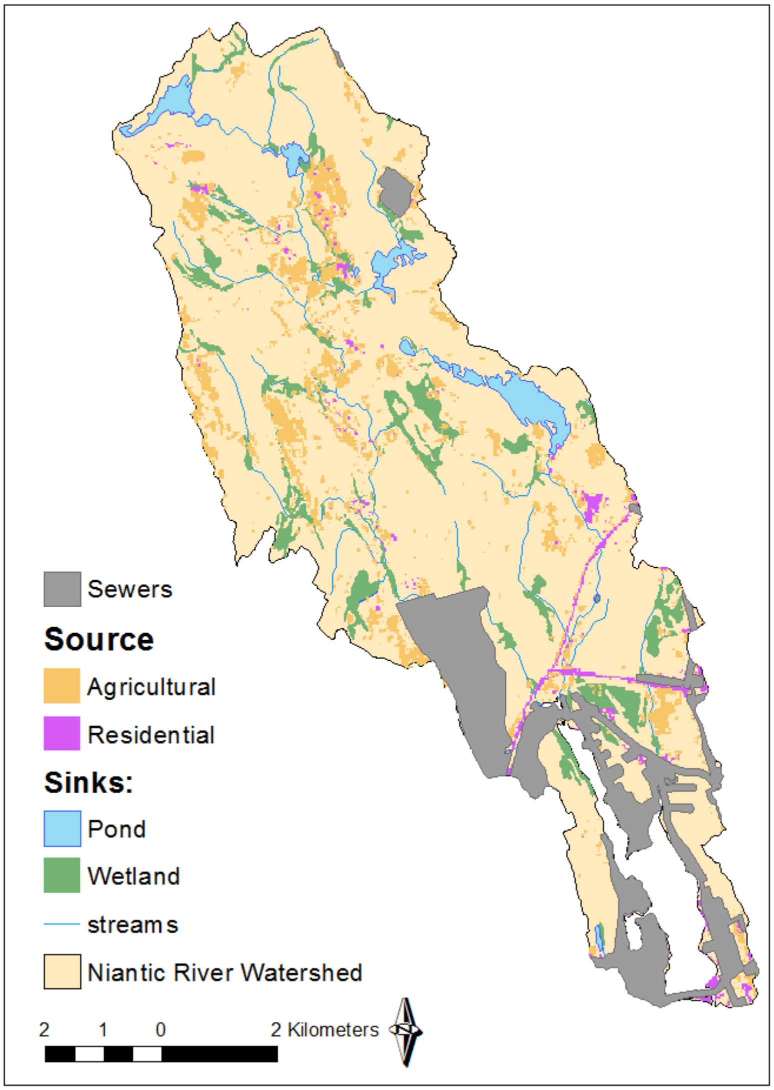


Figure 6. Nitrogen sources and sinks in the Niantic River watershed. Also the areas where sewers are present, precluding residential land use from being a nonpoint source of nitrogen for the NSink model.

Table 5. Areas of sources and sinks within the Niantic River watershed.

Residential (km^2)	Agriculture (km^2)	Ponds/Lakes (km^2)	Wetlands (km^2)	Total Stream Length (km)
2.1366	7.1226	1.824	5.107	82.862

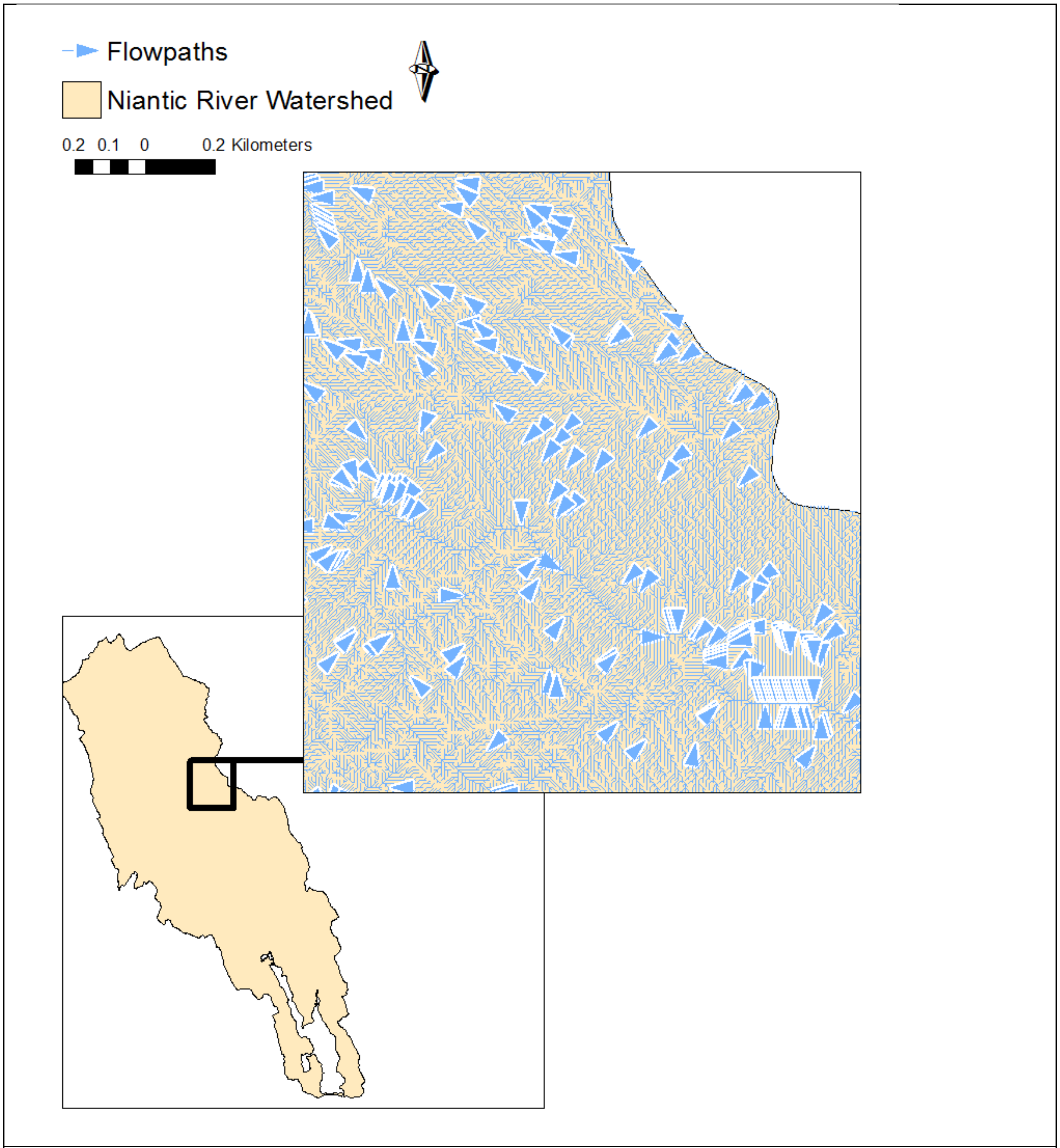


Figure 7. Inset shows close up view of flowpaths generated in Module 2 in the Niantic River watershed. The arrows indicate the direction of flow. A stream is visible flowing from left to right across the lower third of the inset.

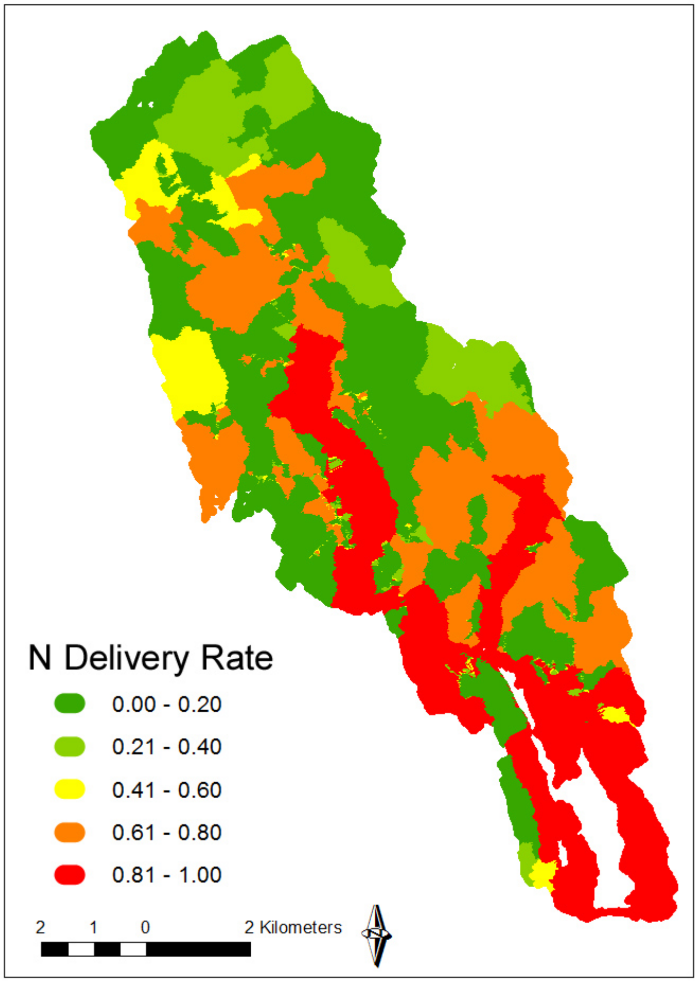


Figure 8. Nitrogen delivery to the estuary from each point within the Niantic River watershed – the color indicates the percent of input from that point that is delivered to the estuary. This estimate of delivery does not include sewers.

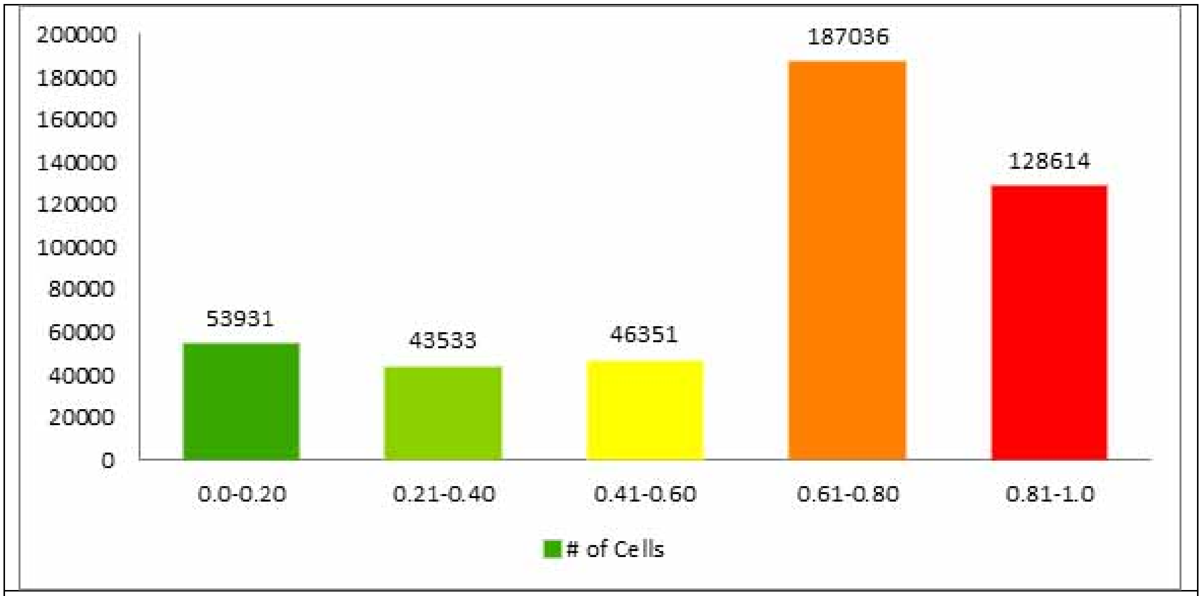


Figure 9. Number of 30mx30m cells that fall within each range of percent nitrogen delivery to the estuary.

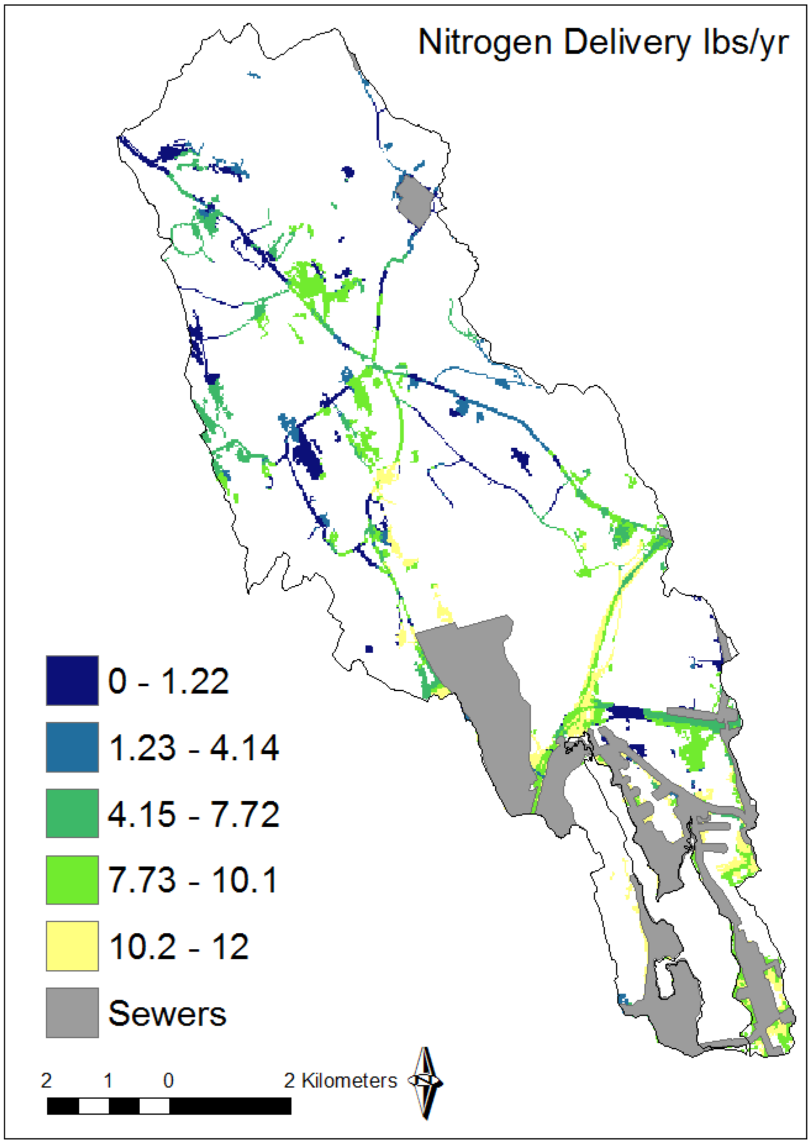


Figure 10. Current nitrogen delivery to the estuary in lbs/year for the current residential and agriculture land uses in the Niantic River watershed. This figure applies the nitrogen contribution map from Figure 8 to the current land use in the watershed, to calculate the current contribution from agriculture and residential land uses.

2.5 NSink Validation

NSink is validated using stream gauge data obtained from the USGS National Water Information System. The stream gauge data were collected in June 2005 during a 3-day dry period when it was estimated flows were approximately at mean conditions (USGS/EPA report 2013). A study in the nearby Pawcatuck River by Fulweiler and Nixon (2005) found that more nitrogen moves with high flows, so it’s likely this June dry period leads to an underestimate in measured loading rates. 11 sites within the watershed were sampled for instantaneous flow and nitrogen concentration during this period (Figure 11). One of the 11 sites (USGS Site # 12779165) does not correspond with a stream in the model, and consequently this site is omitted from the validation analysis. A second site (USGS Site # 11277875) happens to be located on a sub-catchment boundary within the NSink model. Because this significantly impacts the model estimate at that precise location and another site (USGS Site # 1127786) is quite close by in the watershed, # 11277875 is excluded from our analysis. The nitrogen concentration (mg/L) measured at each of the remaining 9 sites is converted to nitrogen load (kg/year) (Equations 7-9) for comparison to the nitrogen load estimate from NSink. We compare nitrogen load because the conversion between load and concentration relies on an estimate of stream flow, which requires a more detailed temporal view of streams that NSink does not provide. By using the instantaneous stream flow measured at the same time as the nitrogen concentration to convert the concentration into load, we are able to directly compare that load measurement to the load estimates of NSink for these 9 individual locations.

$$x\left(\frac{lbs}{yr}\right)\times\left(453592\frac{mg}{lbs}\right)\times\frac{1\,year}{3.156*10^{7}\,seconds}=y\left(\frac{mg}{s}\right) \tag{7}$$

$$Discharge\left(\frac{L}{s}\right)=1000\times\left(0.0295\left(\frac{\frac{m^3}{s}}{km^2}\right)\right)\times DrainageArea\ of Watershed\ (km^2) \tag{8}$$

$$Nitrogen\ Concentration\ \left(\frac{mg}{L}\right)=Total\ Nitrogen\ \left(\frac{mg}{s}\right)\div Discharge\left(\frac{L}{s}\right) \tag{9}$$

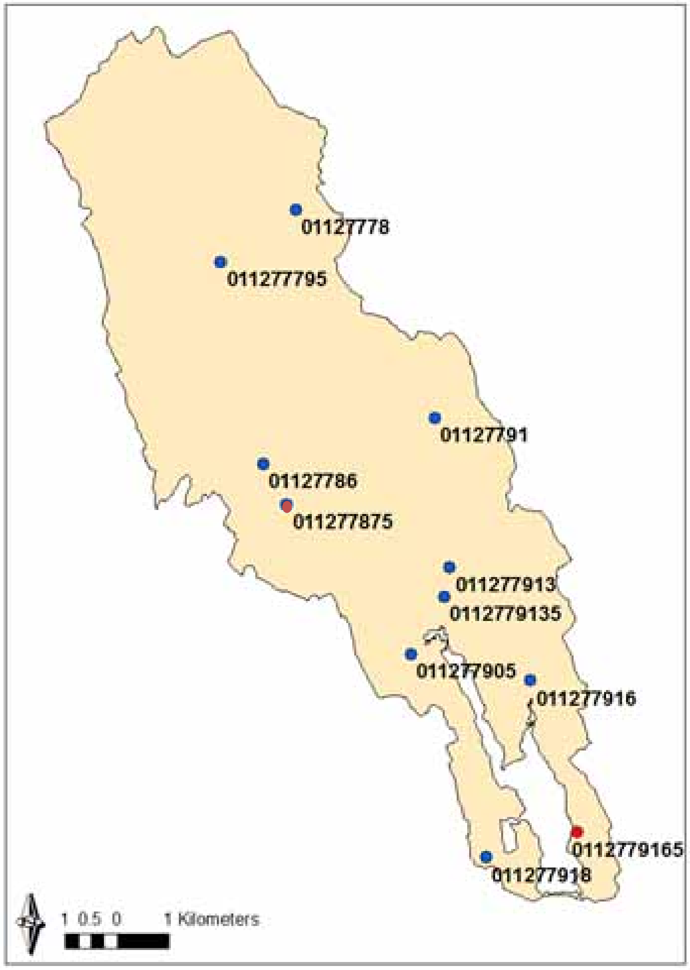


Figure 11. 11 USGS sites sampled in 2005 used to validate NSink model.

A simple linear regression between simulated nitrogen load and measured nitrogen load for the 9 USGS sites shows statistically significant ($P < 0.05$) correlation ($R^2 = 0.9683$) (Figure 12 and Table 6). Despite a high correlation overall between measured and simulated nitrogen loads, in each of the 9 sites, the NSink model overestimates the nitrogen load – in one case by as much as 11,782 kg/year.

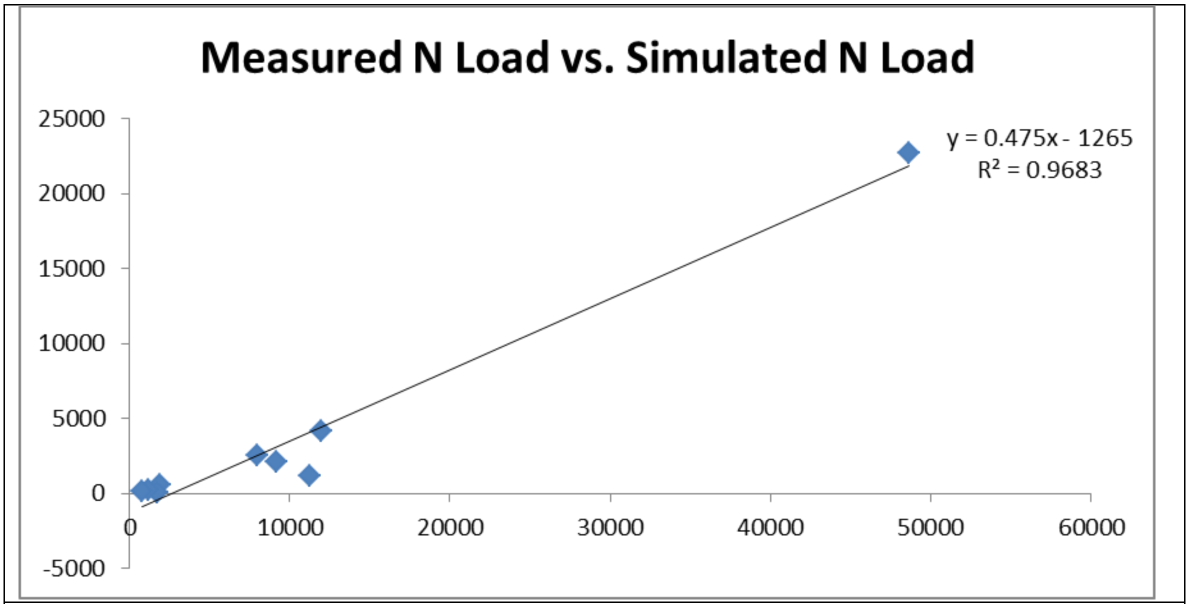


Figure 12. Simple linear regression comparing simulated nitrogen estimates from NSink to measured nitrogen at 9 USGS sites measured in 2005.

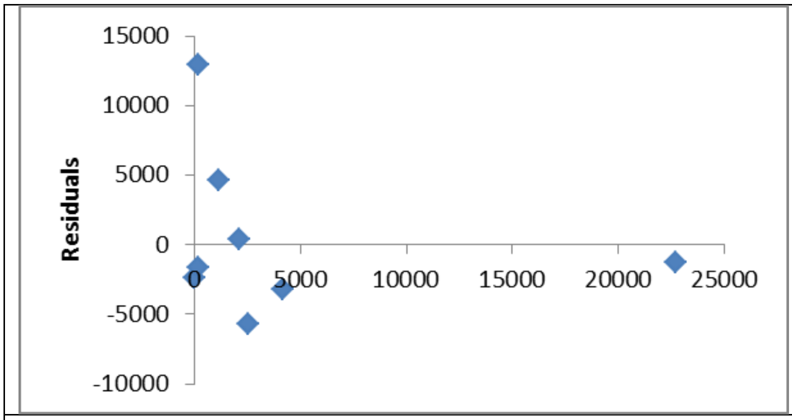


Figure 13. Residuals from the simple linear regression comparing simulated nitrogen estimates from NSink to nitrogen measured in 2005 at 10 USGS sites.

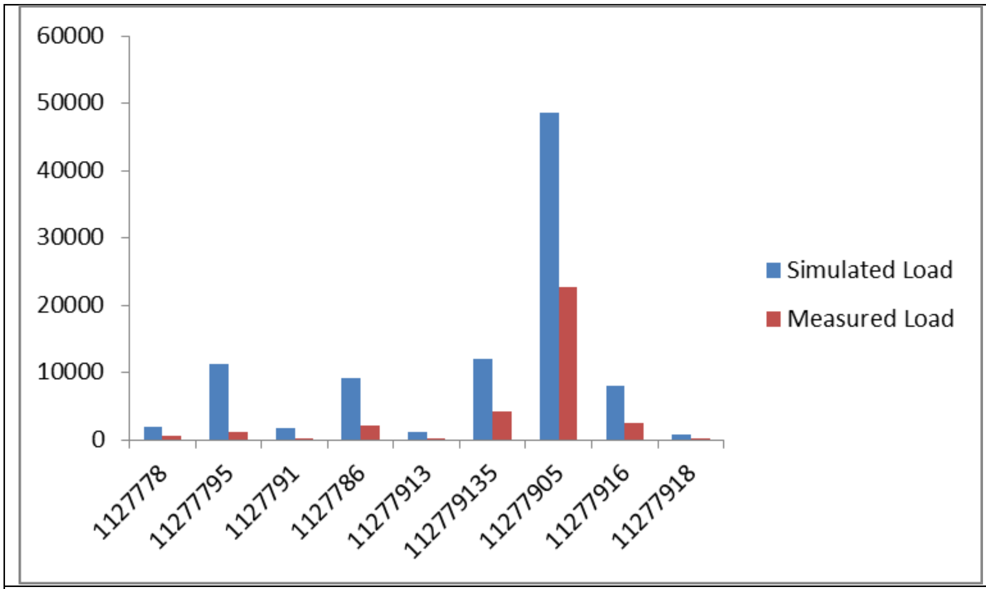


Figure 14. Comparison of simulated load versus measured load (in kilograms) at each of the 9 USGS sites.

Table 6. Univariate measures, regression values, and difference measures comparing NSink to the 10 measured USGS values.

Univariate Measures (kg/yr)					Regression values			(kg/yr)	(kg/yr)	Difference Measures (kg/yr)		
$\bar{O}$	$\bar{P}$	S_o	S_p	N	R Square	Adjusted R Square	Standard Error	$RMSE_s$	$RMSE_u$	MAE	RMSE	d
3,733	10,521	7,242	15,001	9	0.970	0.965	1,431	13,089	4,969	5,554	9,166	0.87

In addition to the linear regression, we calculate a series of univariate measures and difference measures to further evaluate the performance of the model (Table 6). The univariate measures include the mean of the mean of the observed values ($\bar{O}$), the mean of the predicted model values ($\bar{P}$), 3,733 kg/yr and 10,521 kg/yr respectively, and the standard deviations for each (S_o, S_p), 7,242 kg/yr and 15,001 kg/yr respectively. While the mean of the values simulated by NSink is noticeably lower than the mean of the measured values, it is within one standard deviation. The difference measures include the Mean Absolute Error (MAE), 5,554 kg/yr, the root mean square error (RMSE), 9,166 kg/yr, and an error measure (d) that estimates how closely the simulated load values approach the observed. The error measure (d) value of 0.87 suggests the model performs reasonably well. Further, the systematic error ($RMSE_s$) is higher than the unsystematic error ($RMSE_u$), thus we infer the model is performing reasonably well but the data input to the model may be a source of disagreement between the simulated and measured values.

The overall nitrogen retention of the nitrogen sinks is modeled to be 66%, which falls within the range of 60-90% established in the literature (Howarth et al., 1996, Jordan et al., 1997). Because the NSink model does not account for atmospheric deposition, this value is probably lower than the actual watershed retention of the Niantic River watershed (Fulweiler and Nixon 2005). Using a general flow value for coastal New England watersheds (Armstrong and Parker, 2003) we are able to convert NSink's estimated nitrogen load delivered to the Niantic Bay at the watershed outlet into a nitrogen concentration of 0.75 mg/L. This value is quite comparable (~3% difference) to the mean of 0.77 mg/L measured on the 6 USGS sites in the lower watershed around Niantic Bay (Figure 11). This analysis suggests that the NSink model performs reasonably well at the local level and even better at the watershed scale.

6 Discussion

The goal of NSink is to model the geospatial interaction of nitrogen sources and sinks at the local and watershed levels. Our model validation using 10 USGS sites showed a statistically significant correlation between the NSink model and the measured values. Additional consideration of the model at the watershed scale revealed a reasonable approximation at that scale as well.

Validating the model is an important step in the geodesign approach to SDSS development. To adequately consider the impacts of model performance as a design alternative, it is necessary to evaluate the model performance in light of impacts on decision-making. In our example, model performance must be validated to ensure decisions about land use change reasonably reflect what would happen in the event of such changes in the Niantic River watershed environment.

There is an inherent challenge in watershed modeling and validation, notably because of the problem of non-uniqueness (see for example, Yen et al. 2014). Nonetheless, our efforts to validate the model revealed some likely issues with the NLCD data used to model nitrogen sources – particularly the residential land use areas. We noted that between the years of 1992 and 2001 the land use classification scheme changed and starting in 2001, and including 2006, some areas that appear to be highways are actually classified as residential land use in the new classification scheme. It is likely that this inclusion of highways as residential source areas is contributing to the overestimation of the NSink model, however the USGS data for validation was not available prior to 2005, thus we chose to proceed with the 2006 NLCD data.

Further, the values used for residential land use loading are highly dependent upon actual occupancy of areas using septic systems. We masked out the areas known to be on sewer (Figure 6; based on 1998 data from CT DEEP), but the estimates of how many people and how many septic systems are employed in the remaining areas is not readily available. We used an estimate of population density from the 2010 Census, which showed an average population of approximately 3 people per acre leading to the residential load estimate of 24 lbs/acre, or 2.42 kg/30m cell.

Additionally, the agricultural land loading estimate is based on a single type of agriculture, corn row crop agriculture, and may not accurately reflect all types of agriculture present in the Niantic River watershed. Corn is a particularly high nitrogen load type of agriculture, so using this estimate may lead to overestimates in the entire model.

Three spatial layers of output are generated from the NSink application to the NWR (section 2.4.3), and each provides a different focus on the geospatial interaction within the watershed. The first layer shows the sinks and sources that play a part in the denitrification taking place in the watershed. In the current implementation of NSink, this includes residential and agricultural land use and ponds/lakes, streams, and riparian wetlands. The second layer represents a “heatmap” showing the potential for nitrogen delivery to the watershed outlet from any 30mx30m cell within the watershed. This layer is expected to be particularly useful to decision-makers as it highlights areas of the watershed where new nitrogen inputs will have a stronger or lesser impact. The final layer provides insight into the current nitrogen sources in the watershed, and is also expected to be of use to decision-makers. This layer can help identify areas of nitrogen source that are most contributing to nitrogen in the Niantic Bay.

One of the challenges of implementing NSink is the extensive data processing required to specify the geospatial interaction between sources and sinks. Our visual observation on the Niantic River case study suggests that a positive relationship exists between the area of sinks within the watershed and total N removal, which is supported by other studies (for example, Yang 2012). Future applications of NSink should consider if modeling the detailed interaction at the local level is necessary. Finally, the model does not presently account for atmospheric deposition of nitrogen.

7 Conclusions

As growth in the Niantic River watershed continues, it is important that growth be directed in ways that do not further contribute to environmental degradation from increased nitrogen loading. This research introduces a model that helps identify key areas within the watershed for consideration of future growth and helps identify areas important to denitrification within the watershed. The model is validated at both the local and watershed scale and estimates nitrogen retention via nitrogen sinks in the Niantic River watershed to be approximately 66% over the entire watershed. The three output layers generated with the NSink model can be used by local decision-makers to make informed land use decisions.

Future efforts with NSink include optimization of the model and incorporation of the model into an interactive SDSS. We also consider the sensitivity of the model to different data inputs in (IN PREP). Further future efforts could include measures of uncertainty for both the model and its incorporation into a SDSS. We can also consider improving source loading estimates with improved estimates of residential density and more explicit characterization of farming practices within the watershed. Additionally, NSink could be applied to more watersheds in the area to help support a concerted effort at reducing nonpoint source nitrogen pollution in coastal New England.

CHAPTER 3

DATA FITNESS FOR USE: AN EVALUATION OF ACCURACY, AGREEMENT, AND APTNESS (AAAQ) OF FOUR WETLAND DATASETS FOR USE IN AN ENVIRONMENTAL SPATIAL DECISION SUPPORT SYSTEM

3.0 Contribution to the dissertation

This paper assesses the data fitness-for-use of four different datasets used to develop the Spatial Decision Support System (SDSS) outlined in this dissertation. The manuscript is being written for possible publication in *The International Journal of Geographic Information Science*. It is co-authored by Melinda Shimizu, Elizabeth A. Wentz, Joanna Merson, and Arthur J. Gold. The manuscript compares four different data sets and assesses them based on accuracy, agreement, and aptness. These measures are defined in the paper. The analysis aims apply the geodesign approach to determine if different data sets alter the nitrogen interaction model and resulting output enough to potentially impact decisions made by decision makers.

3.1 Introduction

Measures of data quality and the related measures of data fitness for use have long strived to clearly and precisely convey information that allows data consumers to choose the data best suited to their needs (Chrisman 1983, Goodchild 1992, Guptill and Morrison 1995, Frank 1998, Veregin 1999, De Bruin et al., 2001). While measures of data quality focus primarily on the accuracy, resolution, consistency, and completeness of data (Veregin and Hargitai 1995, Veregin 1999), measures of data fitness for use focus on how closely the data meet data consumer needs (Chrisman 1983, Brassel et al. 1995, Hunter and Goodchild 1995, Veregin 1999, DeBruin et al., 2001, Frank 2004). In other words, measures of data quality can be assessed regardless of the specific application, while measures of data fitness for use require an understanding of the intended application. Almost certainly because of this difference, measures of data quality are more commonly evaluated than measures of data fitness for use (see for examples of data quality: Guptill and Morrison 1995, Veregin 1999; see for examples of data fitness for use: Frank 1993, Hunter and Goodchild 1995, Agumya and Hunter 1999, Klein 1999, Frank 2004, Devillers 2007). As more data covering similar topics become available, potentially with high levels of accuracy and reliability, mechanisms to compare data fitness for use are needed.

This paper presents a structure to evaluate data fitness for use in the context of a spatial decision support system (SDSS). SDSS are tools that utilize spatial data, models, analytics, and visualization tools to improve decision-making, policy formation, and dissemination of scientific information. We apply a geodesign approach by considering the implications of different design alternatives; specifically the design impacts from using different data sources.

We define a ranking structure for data based on a combination of quantifying accuracy, agreement, and aptness of data, which we call $\boldsymbol{AAA_Q}$. By including accuracy, we incorporate the conventional ideas of data quality by addressing whether the data measurements represent what is implied. Agreement relates to precision or the consistency of a measurement by measuring similarity or uniqueness between datasets. Lastly, we incorporate the aptness, or how well a dataset meets model or application context needs. It is our implementation of these last two measures that make the AAA_Q methodology especially suited to SDSS data evaluation. The AAA_Q data fitness for use ranking method is necessary for SDSS data evaluation, because absolute standards of quality (e.g., accuracy and precision) depend on scale, temporal consistency, data definitions, and ground truth options, which may not be sufficient measures when examining data for a specific application, like a SDSS.

We implement our assessment structure on a case study for a SDSS called NitroSim that supports watershed management in the Niantic River watershed of Connecticut, USA (Shimizu Dissertation 2014). In our example, the SDSS supports decisions about nonpoint source nitrogen pollution planning. The system relies on a model of nitrogen removal via streams, ponds, and wetlands within the watershed (Shimizu et al 2014). Four different data sources are available to delineate the wetlands within the watershed, and we address the question of which wetland data set is most fit for use in a nitrogen SDSS.

3.2 Background

All data have inherent error or uncertainty, and it is important to assess the effect this error has on the resulting product or application of data, especially in cases of decision support (Chrisman 1983, Guptill and Morrison 1995, Goodchild 1995, Veregin 1999, Fisher and Tate 2006, Devillers 2007). It logically follows that users should utilize the "best quality" data available, but what that means exactly is difficult to answer. Current data quality descriptions tend to focus on accuracy, resolution, consistency, and completeness (Guptill and Morrison 1995, Veregin and Hargitai 1995), but metadata are often found to be difficult to understand or apply to a specific data need (Devillers et al., 2007). Therefore, it is important to clarify what is meant by "best quality." Chrisman (1983) first used the term *fitness for use*, which describes the idea that different applications require different forms of data quality. Agumya and Hunter (1999) and DeBruin et al., (2001) compare fitness of use to quality control for data. Despite the fact that "most laboratories aim to produce data of the highest quality," Bedard and Barnes

(2010) point out that “not all scientific problems require the same data quality.” The *best quality* data, therefore, are not necessarily pushing detection limits or at a very fine resolution, but are instead the data most suitable to the application (Agumya and Hunter 1999, Fisher and Tate 2006, Devillers 2007, Bedard and Barnes 2010). That is to say, the issues of scale and resolution are commonly considered when characterizing data quality, but they may not be the most important considerations. The data should be evaluated in terms of their intended use.

Several fields of study consider data quality and accuracy, but the field of remote sensing in particular frequently considers data accuracy, user error and producer error of remotely sensed data. Consequently, several methods for measuring and communicating data quality and error have been developed, especially relating to classification of remotely sensed data. These measures of data quality include discrete multivariate techniques (Congalton et al. 1983), minimum accuracy value (Aronoff 1985), and applications of probability theory to line sampling methods (Skidmore and Turner 1989). In each of these techniques, it is necessary to have accurate reference data for comparison. Usually, users and producers of these data agree upon a reference data set that by necessity has high accuracy (Congalton 1991).

Typically, these site-specific accuracy assessments are represented as an error matrix (Congalton 1991). Such matrices describe the overall data accuracy and user and producer error, and communicate the data quality in a concise way that is useful to data users evaluating the classification and data accuracy. The error matrix is limited, however, when there is no reliable reference set against which to compare the data. This is particularly the case when considerations of data quality are in terms of how the data will be used, instead of an absolute measure of accuracy.

Because data fitness for use depends on the application, absolute standards of quality are not very useful and are quite limited (Agumya and Hunter 1999, Fisher and Tate 2006). Even measures of user and producer error are limited because they focus solely on the data and not on any models that use the data. To address this shortfall, many efforts have been put forth to assess data for a given purpose. For example, Devillers et al. 2007 puts forward a tool to support experts in data fitness for use assessments. This tool relies on metadata to quantify and aggregate data quality indicators based on expert user selections. While their approach has potential for broader application, the implementation was demonstrated as a “one-shot contractual activity” and was largely designed to support legal issues related to data quality. It is further limited by its focus on metadata instead of focusing on how the data are used. Another example is the use of decision trees by DeBruin et al. 2001. Their approach allows selection of a data set for decision making based on the reported data error. Consequently, this approach requires probabilistic accuracy measures, similar to the high accuracy reference data set required for measures of data quality in remote sensing, which is limited in settings where such reference data are not available.

Agumya and Hunter (1999) took a slightly different approach to assessing data fitness for use, with a focus on risk-based assessment. Their approach considers data quality in terms of the potential impacts of data error with an example of delineating flood zones based on a digital elevation model (DEM). Again, this approach relies on ground truth or an accurate reference data set to compare the DEM against. In short, there is no one way to assess data for usefulness in application, possibly because of the wide variety of applications data are used for.

Further, most efforts thus far rely on some form of reference data to assess data error. Most typically, data fitness for use assessments focus on DEM data (see for example, Hunter and Goodchild 1995, Agumya and Hunter 1999, DeBruin et al., 2001, Vasseur et al., 2003), which are more likely to have ground truth or a highly accurate reference data set for comparison. This ground truth option allows for estimates of probability and the application of statistics to estimate the data quality in terms of accuracy, resolution, completeness, and consistency. Methods for considering data quality for data that do not have a ground truth or reference data set for comparison are presently underdeveloped.

With this in mind, we put forth a new methodology (AAA_Q) that builds upon the idea of the remote sensing error matrix. Instead of using user and producer error to evaluate data quality, our method uses three aspects: accuracy, agreement, and aptness, which are described in Section 3.3 below. The AAA_Q methodology is especially developed with SDSS data evaluation in mind, which could potentially include both data for which reference data are available and data for which reference data are not available. SDSS design depends on three critical design choices: 1) the choice of model to simulate the decision environment, 2) the choice of data utilized by said model, and 3) the choice of tools included for interaction with the model and data. This paper addresses the second key choice: the choice of data.

3.3 Accuracy, agreement, and aptness ($\boldsymbol{AAA_Q}$)

Our method to assess data fitness for use requires that $\mathbf{Q > 1}$ datasets representing the same phenomenon are compared where each dataset is assigned an integer value for $\boldsymbol{AAA_Q}$. The values of $\boldsymbol{AAA_Q}$ are sorted from smallest to largest to rank and subsequently compare the fitness for use of the datasets. Smaller values of $\boldsymbol{AAA_Q}$

represent datasets that are more fit for use and larger integers less. Table 1 and Charts 1 and 2 (explained in this section) provide the generalized equations for each diagnostic required to calculate $\boldsymbol{AAA_Q}$. The $\boldsymbol{AAA_Q}$ methodology simultaneously considers three diagnostics of data: accuracy (**A1**), agreement (**A2**), and aptness (**A3**). The weighted ($\boldsymbol{w}_i$) values of the three diagnostics are then summed to calculate $\boldsymbol{AAA_Q}$ for an individual data set $\boldsymbol{q}$.

Accuracy (**A1**), quantified in Table 1, Column 1 and illustrated in Figure 1, aims to quantify how well the data abstraction represents the reality within a given application. Accuracy can be considered on the whole as Veregin (1999) did or as a combination of accuracy components ($\boldsymbol{b}$) such as positional, attribute, and temporal accuracy as Guptill and Morrison (1995) did. In a data quality context, accuracy simply addresses whether or not a measured value represents the observational version of that value. In a fitness for use case, accuracy considers that but also considers reality within a specific model or application context. The value **A1** is then calculated (Eqn 1) as the sum of the accuracy components ($\boldsymbol{a_{1j}}$) up to $\boldsymbol{b}$ components.

$$A_{11} = w_1 \sum_{j=1}^{b} a_{1j} \tag{1}$$

The value $\boldsymbol{a_{ij}}$=1 is assigned to components that are accurate. Larger integers are assigned to less accurate components as determined by the application context or model. In the simplest case when accuracy is considered as a whole, then $b=1$ and **A1** ranges from 1 to Q. When evaluated as components, b>1 and A1 ranges from b *to* $b \cdot Q$.

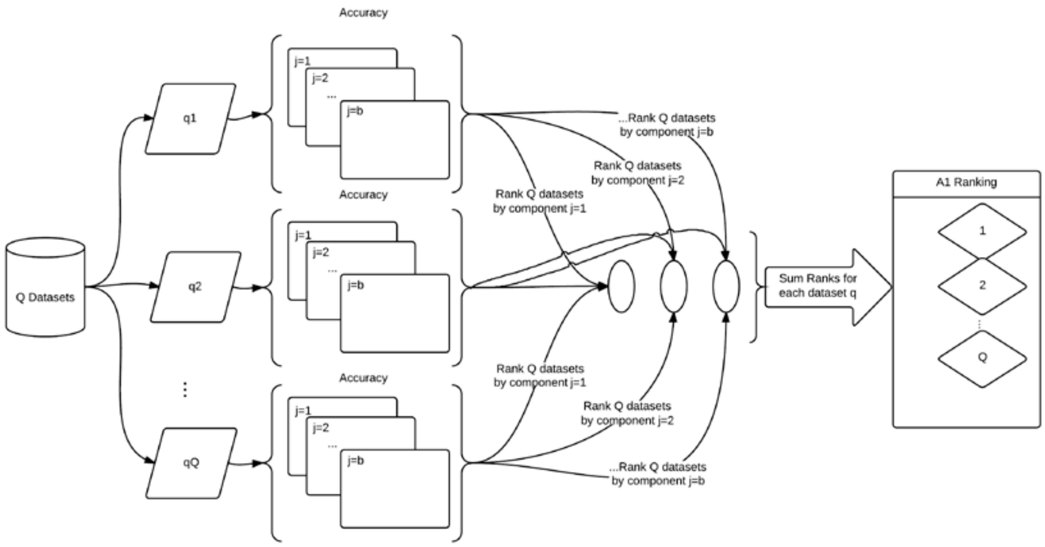


Figure 15. Illustration of A1 ranking process.

The second diagnostic of $\boldsymbol{AAA_Q}$ evaluates the agreement (**A2**) between the Q datasets. Agreement refers to the consistency or similarity between datasets and identifies datasets that have numerous differences as "poor." Because we are considering data without a highly accurate agreed upon reference data set for comparison, we instead compare each data set to each other. The assumption is that agreement between data sets most likely indicates those features reasonably represent reality. To quantify **A2**, the Q datasets are converted into raster layers for a cell-by-cell comparison between datasets and evaluated as sets.

The two cases considered for **A2** are nominal datasets and interval/ratio datasets. In the simplest nominal case, the number of attributes is m = 1. In this case datasets (**Q**) are coded according to the presence or absence of the attribute and are represented as sets (Equation 2) and then the resulting datasets are summed cell-by-cell (Equation 3):

$$\forall\, q \in Q\!:\ R_q = \begin{cases} r = 1 \textit{ when attribute is present} \\ r = 0 \textit{ when attribute is absent} \end{cases} \tag{2}$$

$$R = \sum_{q=1}^{Q} R_q \tag{3}$$

The resulting R raster data layer is a set that represents the similarity between the datasets such that the integer values represent the cell-by-cell agreement (**c**) between the datasets:

c=0: none of the datasets have an attribute assigned to that cell
c=1: one dataset assigned an attribute to that cell
c=2: two datasets assigned an attribute to that cell
.
.
.
c=Q: all datasets assigned an attribute to that cell

To create the $\mathbf{a_{2q}}$ ranking (Table 7, Column 2) for each dataset $\mathbf{q} \in \mathbf{Q}$, each is evaluated independently by calculating the percent of cells in total agreement. The result is Q datasets ($\boldsymbol{R_{qA2}}$), representing agreement values for dataset q (Equation 4):

$$\forall\, q \in Q\!: R_{qA2} = R_q \cdot R \tag{4}$$

From $\boldsymbol{R_{qA2}}$, we extract a subset, $\boldsymbol{P_{qA2}}$ that represents total agreement, such that only cases where $\boldsymbol{c=Q}$, are included:

$$P_{qA2} \subset R_{qA2} | \{p = 1 : c = Q\} \tag{5}$$

To calculate the percent of total agreement per dataset $\boldsymbol{q}$, the cardinality, or the number of elements within set $\boldsymbol{P_{qA2}}$ is compared to the total number of elements in the binary dataset $\boldsymbol{R_q}$:

$$\forall\, q \in Q : P = \{p_q = \frac{|P_{qA2}|}{|R_q|} \cdot 100\} \tag{6}$$

The elements of $\boldsymbol{P}$ are sorted into descending order. **A2** is the set of ranks for datasets Q such that the value of $\boldsymbol{a_{2q}}$ is assigned the rank position of element **p**.

$$\boldsymbol{A2} = \{a_{2q} = p_q \mid p_q = \text{the rank position of } p_q \in P\} \tag{7}$$

The procedure for calculating **A2** in cases where **m > 1** (e.g., the number of attributes in a nominal dataset is greater than one), is similar. The datasets Q are converted in **m** binary datasets (Equation 2). The operations are repeated for all **m** cases up to and including Equation 7. The final calculation of **A2** (Equation 5) depends on summing the ranks for each of the **m** cases of **P**:

$$P = \sum_{k=1}^{m} P_k \tag{8}$$

The elements of **P** are sorted as before (Equation 7) to assign $\mathbf{a_{2q}}$ to **A2**. Interval/ratio cases follows a same procedure except that the interval/ratio data are pre-processed into

m bins, which then becomes a nominal case of $\mathbf{m} > \mathbf{1}$. This process is illustrated along with the process for **A3** in Figure 16

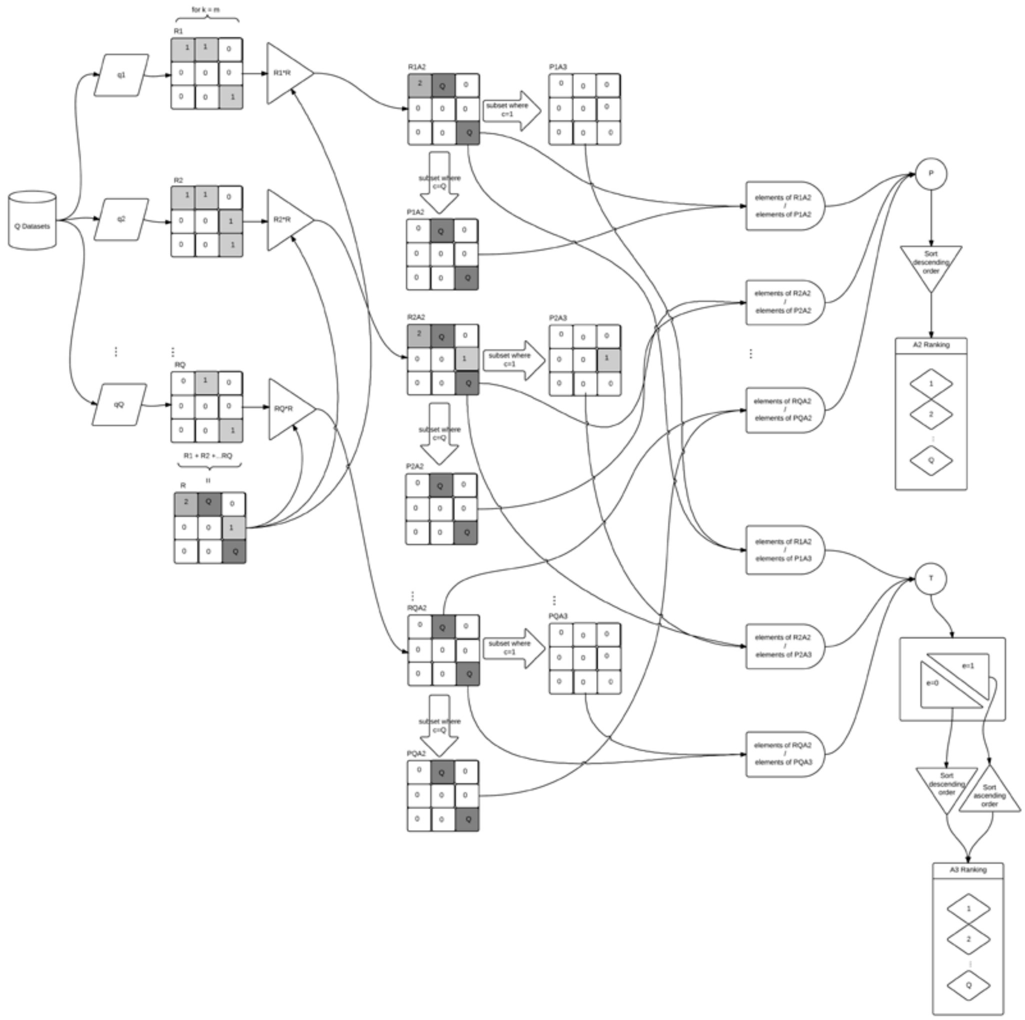


Figure 16. Illustration of ranking process for A2 and A3 diagnostics.

The final diagnostic of $\boldsymbol{AAA_Q}$ considers the aptness (**A3**) of the Q datasets within the context of the intended application. This diagnostic evaluates the data in terms of errors of omission or commission and is similar to the concept of data completeness (Guptill and Morrison 1995, Veregin 1999). To quantify **A3** it is necessary to identify which error type (**e**) is preferred by the consumers intending to use the data. Then the same raster sets used to quantify **A2** are used evaluate aptness by identifying areas unique to each dataset, **q**.

$$e = \begin{cases} 1 \; when\ error\ of\ comission\ is\ preferred \\ 0 \;\; when\ error\ of\ omission\ is\ preferred \end{cases} \tag{9}$$

As with **A2**, the two cases considered are nominal data sets and interval/ratio datasets and the **A3** diagnostic uses the same raster data sets described in equations 1-3. However, from $\boldsymbol{R_{qA2}}$, we extract a subset, $\boldsymbol{P_{qA3}}$ that represents aptness, such that only cases where $\boldsymbol{c=1}$, are included:

$$P_{qA3} \subset R_{qA2} | \, \{p = 1 \colon c = 1\} \tag{10}$$

To calculate the percent of unique area per dataset $\boldsymbol{q}$, the cardinality, or the number of elements within set $\boldsymbol{P_{qA3}}$ is compared to the total number of elements in the binary dataset $\boldsymbol{R_q}$:

$$\forall \, q \, \in Q \colon T = \{p_q = \frac{|P_{qA3}|}{|R_q|} \cdot 100\} \tag{11}$$

The elements of $\boldsymbol{T}$ are sorted depending on which error type (**e**) is preferred. When errors of commission are preferred (e=0), the elements of $\boldsymbol{T}$ are sorted into ascending order. When errors of omission (e=1) are preferred, the elements of $\boldsymbol{T}$ are sorted into descending order. **A3** is the set of ranks for datasets Q such that the value of $\boldsymbol{a_{3q}}$ is assigned the rank position of element **t**.

$$\boldsymbol{A3} = \{a_{3q} = t_q \mid t_q = \text{the rank position of } t_q \in T\} \tag{12}$$

The procedure for calculating **A3** in cases where **m > 1** (e.g., the number of attributes in a nominal dataset is greater than 1) is similar. The datasets Q are converted in **m** binary datasets (Equation 2). The operations are repeated for all **m** cases up to and including Equation 8. The final calculation of **A3** (Equation 12) depends on summing the ranks for each of the **m** cases of **T**:

$$T = \sum_{k=1}^{m} T_k \tag{13}$$

The elements of **T** are sorted as before (Equation 10), depending on the preferred error (e), to assign $\mathbf{a_{3q}}$ to **A3**. Interval/ratio cases follows a same procedure except that the interval/ratio data are pre-processed into **m** bins, which then becomes a nominal case of **m > 1**.

For each dataset, the accuracy (**A1**), agreement (**A2**), and aptness (**A3**) scores are summed resulting in an $\boldsymbol{AAA_Q}$ (Table 7, Column 5). With $\boldsymbol{AAA_{Q\_min}}$ = 3 and $\boldsymbol{AAA_{Q\_max}}$ = $[b \cdot Q + 2m \cdot Q]$, the ascending order represents the most to least fit for use of the Q datasets compared.

Table 7: Procedure for creating dataset rankings on data fitness for use diagnostics where $\boldsymbol{w}_a$ is an optional weight applied to diagnostics

	Accuracy (A1)	**Agreement (A2)**	**Aptness (A3)**	**AAA$_Q$**
Dataset 1	$A_{11} = w_1 \sum_{j=1}^{b} a_{1j}$	$A_{21} = w_2 \cdot a_{2q}$	$A_{31} = w_3 \cdot a_{3q}$	$AAA_1 = \sum_{i=1}^{3} A_{i1}$
Dataset 2	$A_{12} = w_1 \sum_{j=1}^{b} a_{1j}$	$A_{22} = w_2 \cdot a_{2q}$	$A_{32} = w_3 \cdot a_{3q}$	$AAA_2 = \sum_{i=1}^{3} A_{i2}$
Dataset 3	$A_{13} = w_1 \sum_{j=1}^{b} a_{1j}$	$A_{23} = w_2 \cdot a_{2q}$	$A_{33} = w_3 \cdot a_{3q}$	$AAA_3 = \sum_{i=1}^{3} A_{i3}$

3.4 Case study on watershed nitrogen modeling

Our case study SDSS, NitroSim, implements the data fitness for use $\boldsymbol{AAA_Q}$ methodology described in the previous section to consider data input for the SDSS model. The SDSS supports decision making for considering the location of nitrogen loading land use (sources) and potential needs for best management practices (BMP). In the nitrogen interaction model, the locations of wetlands within the watershed act as ‘sinks’ for nitrogen (Kellogg et al 2010). That is, wetlands supply an important ecosystem service by performing a biochemical process that converts the nitrogen in the water into harmless gas (Gilliam, 1994; Hill, 1996, Gold et al, 2001; McClain et al, 2003, Kellogg et al, 2010). Therefore, the locations of the wetlands within the watershed play an important role in determining exactly how much harmful nitrogen reaches the coastal estuaries where it contributes to the eutrophication process.

Wetland data as a case study are difficult to evaluate for data quality. The wetland data sets considered in this paper are not as clearly defined in reality as, for example, DEMs. Historically, there has been controversy over what defines a wetland (Sader et al., 1995), which is exacerbated by lack of clarity in terms used to describe areas commonly considered to be wetlands, such as marsh, bog, or peat (Lehner and Döll 2004). The 1971 Ramsar Convention was supposed to set a global definition and determine practices for "wise use" of wetlands, but there continues to be great variety in wetland definition and policy evidenced by the large body of literature since 1971. Further, a review of available data sources for wetland identification by Lehner and Döll in 2004 found 13 different data sources for wetland identification, most of which overlap spatially.

Not only are there different definitions of wetlands, but there are different methods for identifying wetlands, even when a definition can be agreed upon. Wetland identification techniques fall into two categories: field mapping or remote sensing. Field mapping relies on trained experts to identify wetlands in situ (Tiner 1984), while remote sensing techniques rely on aerial or satellite imagery and include a variety of techniques (see for example, Sader et al., 1995, Finlayson et al., 1999, Rebelo et al., 2008, Qamer et al., 2009, Zomer et al., 2009). Different classification techniques may yield different data representations, and different experts will almost certainly map the extent of features at least slightly differently.

Rosenblatt et al (2001) evaluated the accuracy of the USDA Soil Survey (SSURGO), which is a nationally compiled wetland data set. Their evaluation showed 93 of 100 sites were accurately identified, but this example of accuracy evaluation is not available for all wetland datasets, nor is it available for all areas delineated in such sets. And it does not answer the challenge of different field experts having slight variations for in situ wetland identification.

So which resulting data product is right? Or more importantly, at what point does this unavoidable variation affect the fitness for use of the data produced? Specifically in our example, at what point does this variation affect the decision-making supported by a

SDSS developed with these data? Klein (1999) notes “A person will consider a decision to be poor if the knowledge gained would lead to a different decision if a similar situation arose.” It is with this in mind that we conduct our own assessment of data fitness for our SDSS.

3.4.1 Study area

The study area is the watershed of the Niantic River in Connecticut USA (Figure 2). The topography in the Niantic River watershed is mild and the terrain is formerly glaciated with several naturally formed lakes and wetland areas. This watershed includes four municipalities: Salem, Montville, East Lyme, and Waterford and a number of lakes, ponds, and reservoirs. The drainage area is approximately 78 km^2. The Niantic River ends in an estuary, Niantic Bay, which drains into Long Island Sound.

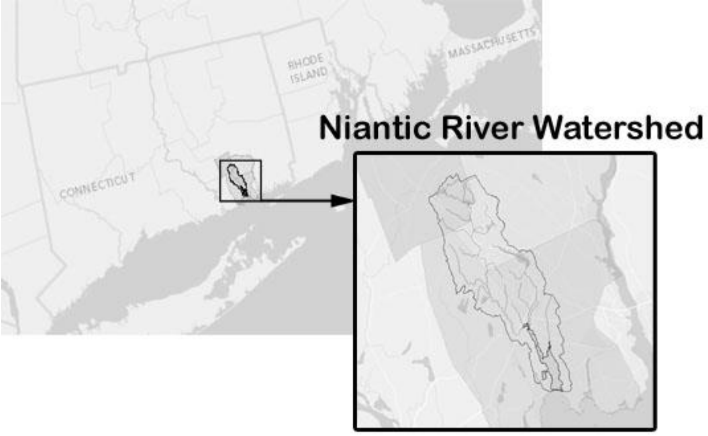


Figure 17: The study area for the case study is the Niantic River watershed in Connecticut USA.

Due to the long residence time and proximity to agricultural runoff and septic systems, the estuary at the end of the Niantic River watershed is particularly sensitive to eutrophication. Native eelgrass beds and local fish populations are affected by increased anthropogenic nitrogen in the watershed (CT, DEP, 2006). Further effort to mitigate this nitrogen is needed, and local communities are interested in a SDSS to support these efforts.

3.4.2 Data sources

The nitrogen interaction model in the SDSS requires data for elevation, streams/rivers, ponds, wetlands and nitrogen sources to calculate the base nitrogen-load level and nitrogen watershed retention that is the foundation for water quality decision support in the SDSS. While all of the data have relevance in the nitrogen model, we focused our data fitness for use case study on the wetland data. Wetlands, which are used to calculate nitrogen retention of surface flow in the watershed, have the most variability and least clear definition of the sink data layers. This variability in available data sources and wetland definition makes the wetland datasets a prime example for implementing the $\boldsymbol{AAA_Q}$ methodology. The four independent wetland data sources we compare are: 1) the National Wetland Inventory (NWI), 2) the USDA Soil Survey (SSURGO), 3) the National Landcover Dataset (NLCD), and 4) the Connecticut state land use data (CLEAR). These data provide the spatial location (e.g., placement and area) of the riparian wetlands in the study area. These data also vary notably in timeliness, with the oldest data collected in 1980 and the most recent data collected in 2011. While the wetlands themselves are not expected to change much with time, the potential for land development over time does exist. It is therefore possible that the older data include wetlands that have since been converted to other land uses. Because the model for our case study SDSS also includes current land use data, this potential impact of timeliness of the data is negligible for our purpose.

The NWI is part of the U.S. Fish and Wildlife Service and has been a source of wetland maps and geospatial wetland data since the mid-1970s (www.fws.gov/wetlands/). These data are generated primarily from the analysis of high altitude imagery in combination with collateral data sources and field work. Wetlands are classified based on the 5-class classification scheme presented by Cowardin et al. (1979), which considers soil saturation as the key defining characteristic of wetlands. The NWI data we consider in this study was generated from imagery taken in 1980 with a scale of 1:80,000, and is delivered as a polygon shapefile.

The SSURGO data set defines wetlands based on field observations and laboratory testing of soil samples. These data have been collected by the National Cooperative Soil Survey as part of the United States Department of Agriculture for about 100 years and are collected at scales ranging from 1:12,000 to 1:63,360 (www.nrcs.usda/gov). Wetlands in SSURGO are determined by soil type, which in turn influences vegetation type. The SSURGO data set we consider in this study was generated in 2011 for the entire state of Connecticut. We used the soil polygons designated as hydric to delineate wetlands.

The third data set we evaluate is the NLCD 2006, a land cover classification with 16 classes that has been applied consistently across the United States with a resolution of 30 meters. This data set is generated from decision-tree classification of Landsat satellite imagery (www.mrlc.gov/nlcd2006.php). Wetland in this case is determined spectral reflectance from the surface and would likely be a vegetation surface. We use the land areas classified as Woody Wetlands (90) or Emergent Herbaceous Wetlands (95) to delineate wetlands from this data set.

Similarly, the CLEAR data set is generated from land classification of 2006 Landsat data with a resolution of 30 meters. This data set was generated using a cross-correlation analysis that results in 12 land cover classes (clear.uconn.edu). We use the land areas classified as "Non-forested Wetland" (8) and "Forested Wetland" (9) to delineate wetlands for our assessment. So while the NLCD and CLEAR datasets both use satellite observations to classify wetlands, the distinction here provides a comparison between a national versus a local data source.

Table 8: Summary of data properties for the four data sets considered.

	NWI	**SSURGO**	**NLCD**	**CLEAR**
Type	National	National	National	State
Scale	1:80,000	1:12,000 to 1:63,360	30 meter	30 meter
Date	1980	2011	2006	2006
Minimum Mapping Size (ha):	0.32	0.121	0.0434	0.0429
Maximum Mapping Size (ha):	84.1	14.6	29.1	19.1
Average Mapping Size (ha):	11.3	2.77	27.7	17.8
Total Area Riparian Wetland sinks (ha):	609	148	219	153

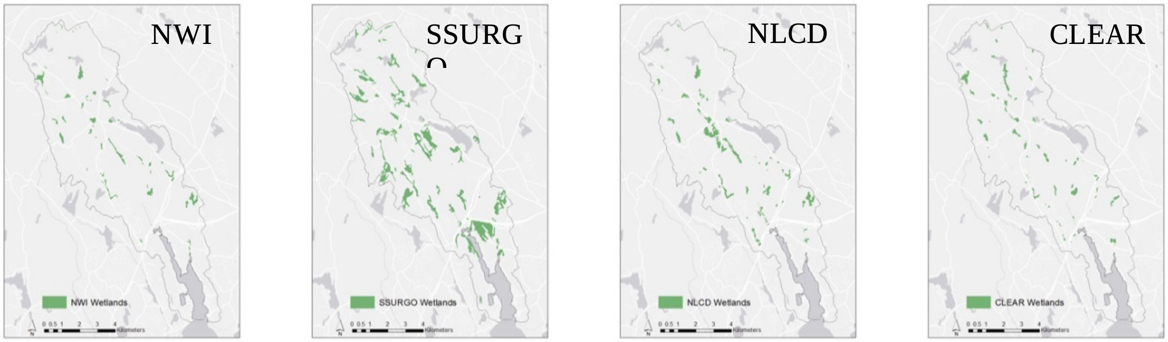


Figure 18. Locations of wetlands for the 4 wetland data sets.

3.4.3 Calculating accuracy, assessment and aptness (AAA_Q)

We calculated accuracy (**A1**), agreement (**A2**), and aptness (**A3**) and subsequently $\boldsymbol{AAA_Q}$ to rank and compare the four datasets to determine fitness for use for the NitroSim SDSS. This section describes how the three diagnostics were calculated in our case study for our Q=4 datasets. For our case study, we used unweighted ranks ($\boldsymbol{w}_i$=1) but acknowledge the possibility of allowing SDSS developers or decision-makers to customize rankings for a specific application.

Accuracy A1:

For our case study, we calculate **A1** using two component parts, resulting in b=2 components for our equation 14 derived from Table 1, Column 1:

$$A_{1q} = \sum_{j=1}^{2} a_{1j} \tag{14}$$

for all q in Q=4.

The two component parts, $\boldsymbol{a_{11}}$ and $\boldsymbol{a_{12}}$, represent how accurately the model calculates the total nitrogen removed by percent (N_r) and the nitrogen load in mg/L (N_l). The environmental science literature reports that nitrogen retention for watersheds in New England is within the range of 60-90% (Howarth et al 1996, Jordan et al 1997). We therefore calculate $\boldsymbol{a_{11}}$ with consideration for range ($\boldsymbol{d}$):

$$\boldsymbol{a_{11}} = \begin{cases} 1 & when \ \min(d) < N_r < \max(d) \\ s & when \min(d) > \ N_r \ or \max(d) < \ N_r \end{cases} \tag{15}$$

Where:

$$\boldsymbol{S} = abs(N_r - median(d)) \tag{16}$$

and the elements of S are in ascending order, s is:

$$s = \text{ the rank position of } s_q \ \in \boldsymbol{S} \tag{17}$$

To calculate $\boldsymbol{a_{12}}$, we found that for the Niantic River watershed, the USGS measured nitrogen concentration at 10 sites during a 3-day dry period in 2005. Of these 10 sites, 2 are near the watershed outlet used in the NSink model. The mean measured nitrogen load at these two sites is 0.56 mg/L, with actual measurements of 0.40 mg/L and 0.72 mg/L.

We use the mean of these two measurements to assess accuracy of the model because both sites are equidistant from the point used to generate the NSink estimate. We then find the absolute difference ($\boldsymbol{g}$) between the model estimate (N_{lq}) to the mean measured load ($\overline{N_l}$), (Equation 18), and put $\boldsymbol{g_q}$ in ascending order such that the value of $\boldsymbol{a_{12}}$ is assigned the rank position of element $\boldsymbol{g}$ (Equation 19).

$$\boldsymbol{g} = |\overline{N_l} - N_{lq}| \tag{18}$$

$$\boldsymbol{a}_{\mathbf{12}} = N_{lq} \mid N_{lq} \;=\; \text{the rank position of } n_{lq} \in N_l \tag{19}$$

Table 9. Two components of accuracy are assessed for each of the four wetland data sets: percent N retention by the entire watershed, and estimated N load delivered to the estuary (mg/L). Each of the four data sets is ranked for each of the two components, and these ranks are summed for a final A1 ranking for each dataset. The minimum possible ranking is 2, and the maximum is 8.

Data Combination (identified by the data source for the wetlands)	**Estimated percent Nitrogen retained by watershed (N_r)**	**Estimated Nitrogen Delivered to estuary (mg/L) (N_{lq})**	**g** $= \lvert \overline{N_l} - N_{lq} \rvert$ **N_l = 0.56**	**A1**		
				a_{11} min(d)=60% max(d)=90%	**a_{12}** $\overline{N_l} = 0.56$	$A_{1q} = \sum_{j=1}^{2} a_{1j}$
				Retention Estimate (N_r)	Delivery Estimate (N_l)	**Total**
SSURGO	66%	0.75	0.19	1	1	**2**
NWI	63%	0.80	0.24	1	4	**5**
NLCD	65%	0.76	0.20	1	2	**3**
CLEAR	61%	0.79	0.23	1	3	**4**

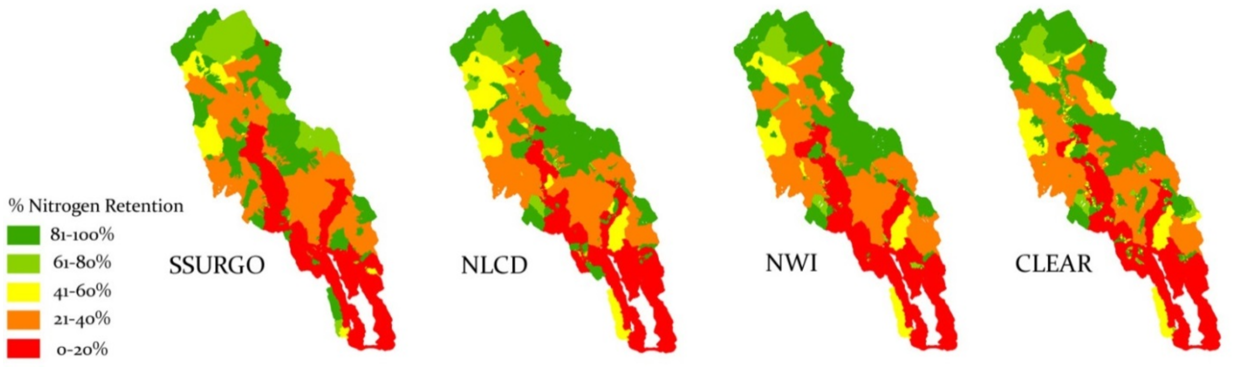


Figure 19. Nitrogen retention in the Niantic River watershed estimated using each of the four considered wetland datasets. This is how much nitrogen is removed from each location, so a value of 81-100% indicates most of the nitrogen input will not make it to the estuary at the end of the watershed. Conversely, the 0-20% areas will contribute most of the input to the estuary.

Table 10. Percent Nitrogen removed by nitrogen sink type for each of the four wetland datasets considered. The minimum and maximum percent removal is also reported. The count indicates the number of 30mx30m cells actively removing nitrogen in each category.

	Removal For:	**Mean**	**Minimum**	**Maximum**	**Count**
SSURGO	Ponds	28.7	0	73.4	28
	Streams	8.57	0	37.8	81
	Wetlands	70.9	40	80	2495
	Entire Watershed	68.7	0.132	80	2596
NLCD	Ponds	52.7	6.98	94.6	22
	Streams	8.91	0.078	28.8	75
	Wetlands	69.7	40	80	998
	Entire Watershed	65.2	0.078	94.6	1095
CLEAR	Ponds	52.9	6.96	91.9	22
	Streams	8.21	0.079	29.2	83
	Wetlands	67.6	40	80	714
	Entire Watershed	61.2	0.079	91.9	819
NWI	Ponds	52.3	6.96	91.9	21
	Streams	8.21	0.079	29.2	81
	Wetlands	97.2	40	80	1025
	Entire Watershed	62.7	0.079	91.9	1127

Agreement A2:

To calculate **A2** in our case study, we consider a single attribute (m=1): wetlands. Using Equation 20, from Table 7, Column 2, we calculate agreement for each of our Q = 4 datasets:

$$A_{2q} = a_{2q} \tag{20}$$

for all q in Q.

The a_{2q} component represents the rank position of each dataset in order of most agreement by area. Table 5 reports the **A2** ranking for our datasets.

Table 11. The Agreement diagnostic considers the amount of each dataset that overlaps with the other datasets and ranks them accordingly. A rank of 1 indicates the most agreements and a rank of 4 indicates the least.

Data Set	Wetland Area in Total Agreement (km^2)	Percent of data set in Total Agreement	A2 Ranking
SSURGO	0.094	1.54%	3
NWI	0.094	0.05%	4
NLCD	0.094	6.4%	1
CLEAR	0.094	6.11%	2

Aptness A3:

A3 is evaluated by using the same raster sets created for **A2**, again with only the wetland attribute (m=1). In our case study, a typical decision supported by the SDSS is deciding where to approve developments. In general, the upper part of the watershed will have the most nitrogen retention and therefore lend itself to the most development with the least impact on environmental degradation. Conversely, the lower part of the watershed closest to the outlet will have the least retention and the most impact on environmental degradation. There are, however, exceptions to this rule when wetlands are located downstream of new developments – thus allowing development lower in the watershed without the otherwise expected environmental degradation. Development lower in the watershed is advantageous because it has more economic value.

The risk, in our example, can be described with two possible situations. First, if the wetland data set used to develop the SDSS does not include a wetland that does, in fact, exist – an error of omission, then the decision-maker might not approve a specific development. The cost of this decision is economic – the decision-maker has foregone economic advantage on a false premise. On the other hand, if the wetland data set includes a wetland that does not, in fact, exist – an error of commission, then the decision-maker might approve a specific development. The cost in this case is environmental – unexpected degradation will result. The decision-makers have to decide which risk they prefer – an error of omission (e=0) or an error of commission (e=1). This decision determines the ranking order of the elements of $\boldsymbol{T}$ (Equation 11). In our example, the decision-makers indicated a preference for errors of commission, so e=1 and the elements of $\boldsymbol{T}$ are ranked in descending order.

A3, for our case study, is calculated with Equation 21 from Table 1, Column 3:

$$A_{3q} = a_{3q} \tag{21}$$

where $\mathbf{a_{3q}}$ represents the rank position of each dataset ordered by most to least unique area.

The estimated nitrogen retention for the entire watershed and the estimated nitrogen concentration for each for the four data sets being evaluated are described in Table 3. A map view summary of nitrogen retention by location is also generated for each data set in Figure 4. These maps indicate how much nitrogen is estimated to be removed at a given location, in other words, if a source is placed at a location how much of that nitrogen input would be removed before reaching the watershed outlet.

Agreement is assessed by the results of an overlay analysis, presented with the associated rankings in Table 6. The overlay maps that are generated as part of the overlay analysis are included in Figures 5 and 6. The area in total agreement between all four data sets, that is the area delineated as wetlands in all four data sets, is 0.094 km^2. This amounts to 6.11% of the CLEAR data set, which is ranked “1” for agreement; 6.4% of the NLCD data set, which is ranked “2”; 1.54% of the SSURGO data set, which is ranked “3”; and 0.05% of the NWI data set which is ranked “4”.

Aptness is assessed by considering a different output from the overlay analysis used to evaluate the area of each data set that is not overlapped with wetland area from any other data set. The results of this assessment are recorded in Table 7 and in map view in Figure 6. SSURGO has the most unique area with 2.82% of the data set unique, which is ranked “1”; CLEAR has the next most unique area of wetlands, with 0.25% of the data set unique, which is ranked “2”; NWI and NLCD are equally ranked “3,” with 0.16% of each data set unique from any other.

Table 12. The Aptness diagnostic considers the amount of each dataset that does not overlap with any of the other datasets and ranks them accordingly. The ranking order depends on decision maker preference for errors of commission or omission. In our example, error of commission is preferred so the dataset with the most unique area is ranked 1.

Data Set	**Unique Wetland Area (km^2)**	**Percent of data set that is Unique**	**A_{3q} (e=1)**
SSURGO	5.35	2.82%	1
NWI	0.294	0.16%	3
NLCD	0.294	0.16%	3
CLEAR	0.477	0.25%	2

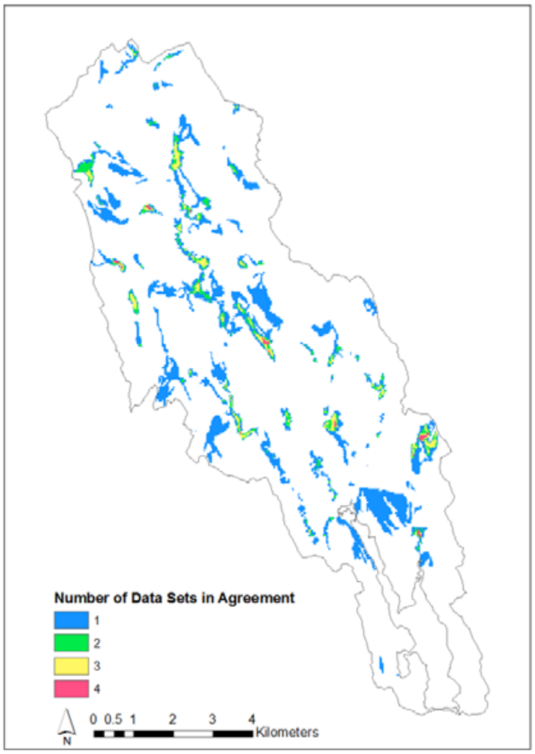


Figure 20. Overlay analysis of four wetland datasets in the Niantic River watershed considered in this study. Areas in blue indicate area identified as wetland by only one dataset; green indicates two datasets identify the area as wetland; yellow indicates three datasets; pink indicates the area identified by all four datasets as wetland.

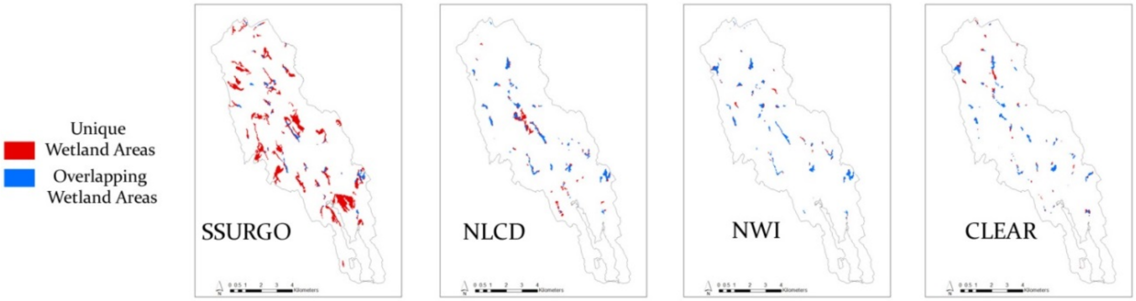


Figure 21. Each of the four wetland datasets are mapped with red areas indicating areas of each dataset that are unique to that dataset. Areas in blue overlap with at least one other dataset.

The results of each of the three diagnostic rankings are tallied in Table 13 and each dataset is assigned a final AAA_Q ranking, which is the sum of the other 3 rankings. The lowest ranking value indicates the dataset that is most fit for use. In our example, the lowest possible ranking value is 4, and the highest possible ranking value is 16.

Table 13. Final calculation of AAA_Q on the four wetland datasets.

	Accuracy (A1)	**Agreement (A2)**	**Aptness (A3)**	**AAA_Q**
SSURGO (Q=1)	2	3	1	**6**
NWI (Q=2)	5	4	3	**12**
NLCD (Q=3)	3	2	3	**8**
CLEAR (Q=4)	4	1	2	**7**

3.4.4 Case study results

With AAA_1 = 6, the SSURGO data set is the most fit for use in our case study SDDS. SSURGO is the data set that: 1) produces estimates of nitrogen retention and load that are closest to the established range of retention and measured nitrogen load values, 2) has the most area of wetland delineation that overlaps with the other data sets, and 3) additionally best satisfies the identified preference for error of commission. Therefore, SSURGO is the data set that best meets user needs in the NitroSim SDSS example.

In this case study, we considered four different data sets for wetland delineation to be used in the generation of a SDSS for water quality planning. The accuracy diagnostic assesses the closeness of the resulting estimates for mean watershed nitrogen retention and nitrogen concentration at the watershed outlet to published ranges of watershed retention and a mean measured nitrogen concentration value from the outlet. The published range of retention is 60-90% (Howarth et al., 1996, Jordan et al., 1997) and the mean measured nitrogen load value is 0.62 (USGS, 2005-2012).

The SSURGO data set has the most accuracy, with an estimated mean retention of 66% and estimated nitrogen concentration of 0.75 mg/L. It is interesting that this data set is the most accurate as it also has the most area delineated as wetlands. This suggests that the other data sets may underestimate the actual presence of wetlands in the watershed. By contrast, the NWI data set ranked least accurate with a mean estimated retention of 63% and an estimated nitrogen concentration of 0.80 mg/L. This is not altogether surprising given that the NWI dataset is the smallest scale of the four datasets considered.

The agreement diagnostic considers how much the four data sets overlap, or agree, on watershed delineation. The agreement between data sets indicates a high likelihood that those areas are wetlands in reality and gives confidence in these parts of the wetland abstractions. The CLEAR data set has the most agreement by area, while NWI has the least agreement. This is interesting because it follows the same ranking as if we order the data sets by scale; the largest scale data set (NLCD) has the most agreement, while the smallest scale data set (NWI) has the least. Further, the two data sets with the most agreement (NLCD and CLEAR), are also the two data sets with highest spatial resolution. These data sets also are both derived from the same Landsat imagery, so perhaps high agreement is to be expected. We would, therefore, expect these data sets to have the least error of commission, measured with **A3**, but that is not the case.

The dataset with the most aptness (highest ranking **A3)** is SSURGO, followed by CLEAR, and then NWI and NLCD, which interestingly have the same area by percent of unique wetland area. Perhaps SSURGO ranks highest because it is based on soil types, while the other three data sets define wetlands based on vegetation and water saturation. SSURGO also has considerably larger total wetland area than the other three data sets, so it makes sense that much of the data set is unique from the other three.

3.5 Implications for data fitness for use

The $\boldsymbol{AAA_Q}$ method provides a means to consider data fitness for use by focusing on data characteristics that are most impactful for the data consumer. This is a useful contribution to a geodesign approach to SDSS development, because it provides a means to consider the impacts of different data sources. While measures of data quality contribute, the methods we present here are also necessary to fully evaluate data options for a specific use as is the case with an SDSS. This is similar to the error matrix concept in remote sensing, that allows data consumers to evaluate data in terms of overall accuracy, user error, and producer error, but our method allows such comparison for data lacking an agreed-upon highly-accurate reference data set. The strength of our method is that it both quantifies data characteristics in a way that is comparable and it is flexible for a range of data uses.

Measurements of data accuracy often focus on probabilities or other statistical measures that compare a dataset to ground truth. When ground truth options are limited, for example in cases of unclear phenomena definition or inaccessible study areas, these measurements of accuracy are similarly limited. In such cases, a methodology for comparing datasets based on a ranking system is quite useful.

Beyond just accuracy, other components of a dataset influence its fitness for use. In the $\boldsymbol{AAA_Q}$ methodology we also present the concepts of Agreement and Aptness as two additional dataset diagnostics to quantify data fitness for use. The concept of Agreement is similar to internal consistency, but instead of focusing on consistency within a single dataset Agreement focuses on consistency between the datasets being compared. The assumption is that areas of agreement between data sets most likely indicate areas that accurately represent reality. This diagnostic is particularly useful for dataset attributes that are difficult to define with singular clarity.

Additionally, the Aptness measure brings the context of how the data will be used directly into the quantification of data fitness. By considering how the data will be used in determining the ranking of $\mathbf{A3}$, the $\boldsymbol{AAA_Q}$ methodology identifies which dataset is most aligned to data consumers' needs.

The $\boldsymbol{AAA_Q}$ methodology is indirectly impacted by factors such as scale, minimum mapping unit, and the timeliness of data. Scale and minimum mapping unit in particular have the potential to impact data fitness for use, because these aspects of the data directly affect the level of detail available to users. While it is likely the $\boldsymbol{AAA_Q}$ diagnostics are consequently impacted by issues of scale, it is not something explicitly considered at this time. In our example, timeliness of data was not a notable factor due to the nature of the case study model foundation for the SDSS. In other examples, timeliness of data could be greatly important and should also be considered.

3.6 Conclusions

As more data from different sources become more readily available, considerations of data quality and data fitness for use become ever more important. Significant contributions have already been made to develop rigorous evaluation of data quality, typically relying on measures of accuracy, resolution, consistency, and completeness. Measures of data fitness for use are less standardized, and necessarily so as different uses require different measures and standards. Most assessments of data fitness for use focus on evaluation of DEMs or data that have clear measures in reality. Assessments of data similar to wetland delineation are less common, perhaps because there is more ambiguity about the "right" answer.

In cases where such ambiguity exists, a method for quantifying and comparing datasets within the context of their intended use is ever more important. While some applications may not be sensitive to variations in data, other applications are. By using an assessment methodology that quantifies datasets both in terms of accuracy and in terms of intended use, data consumers can better evaluate their data choices and understand the potential impacts each of their data choices may have.

The $\boldsymbol{AAA_Q}$ methodology presented here addresses this need by considering three diagnostics of measurement: accuracy, agreement, and aptness of the data. By considering all three diagnostics of the data, our assessment can accommodate datasets with inherent variability, such as the wetland datasets in our case study, and it can provide a quantitative measurement to help data consumers choose the data set best suited to their needs. At present the methodology does not explicitly consider impacts of data scale, minimum mapping unit, or timeliness.

Future directions include applying the $\boldsymbol{AAA_Q}$ methodology to other dataset options, and considering additional diagnostics to add to the methodology. For example, the case study presented in this paper also has several data options for nitrogen source locations and elevation models. It might also be beneficial to include a diagnostic that can address the scale or timeliness of datasets.

CHAPTER 4

CONCEPTUAL DESIGN FRAMEWORK FOR DEVELOPING SPATIAL DECISION SUPPORT SYSTEMS, WITH A ILLUSTRATIVE APPLICATION TO SYSTEM DEVELOPMENT FOR WATERSHED MANAGEMENT IN THE NIANTIC RIVER WATERSHED IN CONNECTICUT USA

4.0 Contribution to the Dissertation

This paper puts forth a new conceptual framework to facilitate the design of spatial decision support systems, which is illustrated with a demonstration applied to watershed management in the Niantic River watershed in Connecticut. The system demonstrated is built with the model and data considered in chapters 2 and 3 of this dissertation, and the use of this framework fits within a geodesign approach to SDSS development by considering the usefulness and impacts of different tool options. It is co-authored by Melinda Shimizu, Elizabeth A. Wentz, Craig W. Kirkwood, Robert Pahle, and Stephanie Deitrick and is intended for publication in *Environment and Planning B.*

4.1 Introduction

This paper presents a decision-focused conceptual framework for developing spatial decision support systems (SDSS) and illustrates the use of this framework by developing and evaluating a demonstration SDSS supporting watershed management addressing nonpoint source nitrogen pollution planning in the Niantic River watershed in Connecticut, USA. This framework helps system designers develop a geographic information system (GIS) for a specific decision situation that supports a wide range of decision-making scenarios in contrast to current system design approaches, which tend to be more rigid and support only a specific, limited, set of decision-making scenarios. We illustrate the usefulness of supporting a wider range of scenarios, and show that this framework is especially useful for decisions that involve a variety of stakeholders, such as technical specialists, public and private sector managers, public interest groups, and the general public. Specifically, the framework provides the flexibility to support a wider range of decision-maker needs within a single system. Finally, the approach is evaluated with a survey of geography and urban planning graduates and current graduate students that considers different SDSS tool designs.

Our framework combines elements from *input-oriented* SDSS with *goal-oriented* approaches from decision science theories that have not typically been applied within GIS, where input-oriented means the focus is on the decision alternatives or choices, and goal-oriented means the focus is on the decision outcomes. This framework is also in line with the emerging field of geodesign, which encourages the development of technologies like a SDSS with a consideration for design impacts. The nitrogen pollution planning SDSS illustrates that applying this framework for SDSS development can result in systems that are more useful in complex decision situations.

The remainder of this paper is organized as follows: Section 2 reviews related work from decision science and GIScience; Section 3 provides a functional design framework for development of flexible SDSS and applies the framework to a SDSS that supports watershed management of nonpoint source nitrogen pollution in the Niantic River watershed in Connecticut, USA; and Section 4 evaluates the SDSS tools developed with the framework via user survey. Section 5 considers the implications of the framework and survey.

4.2 Current SDSS Design

There is a large literature in both the decision sciences and GIScience, and we consider the most closely related previous work in those two fields in this section. We pay special attention to literature regarding SDSS. This section makes clear what has been done previously and how our decision frame moves SDSS design forward.

4.2.1 Decision Sciences

Decision science theory explains decision making from two perspectives: descriptive and prescriptive. Descriptive decision theory explains how unaided decision makers *actually make* decisions. Extensive experimental studies over the past four decades have established that unaided decision makers use a variety of heuristics and biases in their assessment of decision-making situations that lead to less than optimal decisions. See for example, Mellers and Locke 2007, and Tversky and Kahneman 1974; 1981. Prescriptive decision theory explains how decision makers *should* make decisions. The methods of prescriptive decision sciences, usually called *decision analysis*, have been developed to address these heuristics and biases to facilitate better decision-making.

Decision analysis consists of a theory for a good decision-making process, and technology (e.g., computer-based decision support systems) to implement that theory. (See Howard 1968 for an early description of decision analysis.) Howard (2007) notes that a logical decision making process should address three elements: What you can do (your alternatives), what you know (your information about the state of the world), and what you want (your goals, objectives, or values). He states that these three elements should be operated on by a logic that systematically addresses all three elements to produce the best alternative.

Often a decision analysis includes a quantitative or computer model that provides the logic that operates on the three elements, and in this situation Figure 1, adapted from Howard (1968), shows the framework of the analysis. In this figure, $d_1, d_2, \ldots, d_m$ represent the *decision variables* that you have control over (what you can do), $s_1, s_2, \ldots, s_n$ represent the state of the world (what you know), and $v_1, v_2, \ldots, v_r$ represent the variables that measure achievement of your goals (what you want). The arrows show the direction of influence for the variables, and the box represents that quantitative or computer model that calculates the values of the v_i for any specified d_i and s_i .

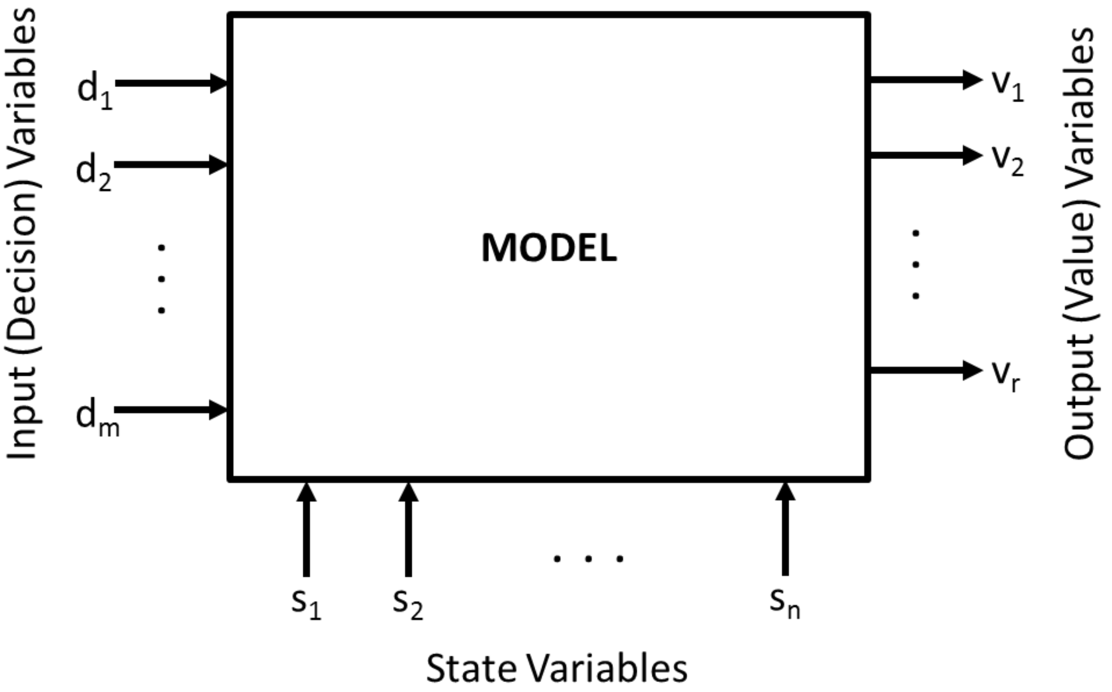


Figure 22. Decision Analysis Modeling Framework.

Spreadsheet methods (Kirkwood 1997) and specialized software (Logical Decisions 2013, Syncopation Software 2013) have been used to implement decision analysis methods. These methods have been widely applied to both public and private sector decisions. (See Corner and Kirkwood 1992, and Keefer et al, 2007 for reviews of published applications.)

It is generally accepted within the decision analysis field that all three of the decision elements discussed above (alternatives, information, and values) should be taken into account in a logical decision making process, but the relative importance of the three elements may vary depending on the decision context. For example, in some private sector business decisions, the main value criterion is financial, and therefore most of the effort in the decision analysis will focus on developing good alternatives and accurately modeling the factors that will impact the financial outcome.

On the other hand, in complex public/private decisions, such as the watershed management decision considered later in this paper, the dimensions of value that are relevant to the decision may be different for different stakeholders, and therefore understanding and appropriately modeling values may be important. In such situations, Keeney (1992; 1994; 1996) argues that focusing first on values can lead to better decision outcomes. That is, in terms of Figure 1, he proposes that focusing first on the output (value) variables $v_1, v_2, \ldots, v_r$ will be useful. One of his rationales for this conclusion is that by first focusing on values, it may be possible to develop decision alternatives (that is, $d_1, d_2, \ldots, d_m$) that better meet the varying goals of the various stakeholders for the decision. In a similar way, focusing first on values may clarify what state variables $s_1, s_2, \ldots, s_n$ should be considered, as well as the key relations among the v_i, d_i and s_i that need to be considered in the model. This approach is successfully applied by Keeney and others, as reported in the applications review articles cited above. Despite this success, decision tool development has not followed suit and presently least often focuses on the values aspect of the three-element decision frame.

4.2.2 GIScience

The field of GIScience literature advances decision support by providing data, methodologies, and visualization tools and techniques for decision situations with a spatial component. In this section, we apply the three-element decision analysis framework reviewed above to previous published approaches for the design of a spatial decision support system (SDSS) in the GIScience literature. Decision support systems (DSS) are data- or model-oriented tools designed to “improve the effectiveness of the decision maker,” (Arnott and Pervan, 2005). *Spatial* decision support systems (SDSS) integrate spatial or geographic elements into the DSS, including data, modeling, and visualization. For example, Dymond et al, (2004) integrate hydrologic, economic, and fish health models into one system to support environmental decisions at different spatial scales and temporal resolutions. They demonstrate that their SDSS provides tools that help citizens and decision-makers evaluate the impacts on economics, water quality and fish health of different development scenarios. Their system also allows the users to evaluate the assumptions used to generate different scenarios. In terms of the three-element decision analysis framework, their system allows users to evaluate the alternatives and information elements of the decision, but does not directly address values.

The most common approach to SDSS design relies on input-oriented functions (see for example, Crosetto and Tarntola 2001, Ballas et al 2007, Meyer and Grabaum 2008, Arentze et al 2011). That is, these systems depend on the user to provide data and model parameters, from which the SDSS predicts an outcome. In other words, these functions focus on the alternatives identified in the decision frame described in Section 2.1. This design approach provides a SDSS that gives decision-makers a forecasted outcome for any specified alternative, which is a useful feature when decision-makers want to address specific what-if questions. For example, a decision maker may want to evaluate the impact on water qualify if a 5-hectare industrial park is placed at a particular parcel within a watershed. A specific example where this design approach proved useful is a study by Robinson et al, (2002). In this study the authors provided decision support for trypanosomiasis control based on six input environmental factors controlled by the user to predict the likeliness of disease and identify important areas for disease control. Again, the focus of this system is on the user input to the system and the different decision alternatives, but the tools do not provide help to the user to directly focus on desired decision outcomes or the values in the Figure 1 three-element framework.

Thus, an input-oriented SDSS design provides limited support to analyze how a large range of plausible input scenarios will impact outcomes or values for decision makers who are most interested in outcomes. For such decision makers, it would be useful to have a SDSS where the user can input a goal future-state or desired decision outcome, and the system then returns the range of input scenarios and states of the world that would achieve that goal. Consider, for example, if the question in the previous paragraph were turned around to ask which parcel in the watershed would have the least impact on water quality from a 5-hectare development? With an input-oriented SDSS, the user would have to create a large number of scenarios and eventually encompass all possible alternatives to arrive at an answer to this question. Producing a large range of alternatives, however, requires an equally large number of inputs to the system. Creating these would be time consuming and would produce a large suite of output information that is difficult to efficiently sort through, assess, and understand. Thus, while the input-oriented approach in Robinson et al, (2002) is useful for addressing precise questions about specified combinations of input factors, this type of decision support is less useful for assisting decision-makers with broader questions about what input alternatives will help them achieve their goals. For such users, a more useful system design would allow users to specify their input resource constraints and their goal regarding disease control, and based on these the system would return the possible scenarios that achieve that goal.

Effectively addressing this kind of question becomes possible with goal-oriented design of SDSSs. Using the terminology from the Figure 1 three-element decision analysis framework, such a system focuses more on the values than the alternatives or information described in the decision frame. The need for a goal-oriented approach is widely recognized in planning and decision science (Couclelis 2005; Hitch, 1955; Lempert et al, 2006; Rittell and Webber, 1973). Hitch (1955) insisted that it is not enough to only look critically at models and their inputs – we must also look critically at our objectives. For example, many public policy issues by their very nature often lack a single clear objective or goal that all stakeholders agree on, For such issues, it is valuable to have SDSSs that can be "re-solved" continually, especially as goals and criteria change (Rittell and Webber, 1973). Hence a SDSS design framework that explicitly addresses goals and values would be particularly useful for developing SDSS in these situations.

Theory in the decision sciences informs the development of SDSSs that can support goal-oriented questions (Couclelis 2005; Hitch 1955; Lempert et al, 2003; Marchau et al, 2010; Rittell and Webber, 1973; Walker et al, 2010). This theory supports a shift away from the input-oriented approach that focuses on identifying the most likely outcome for a user-specified set of input parameters, toward a goal-oriented approach that focuses on supporting a full range of possible outcomes and determining the input scenarios that will yield a desired set of output goals or values (Lempert et al, 2003; Marchau et al, 2010; Walker et al, 2010). Couclelis (2005) refers to the idea of "visioning" which "works by back-casting desirable futures to current conditions and trying to identify robust paths leading from the present to these desired future states." The key to implementing such a process in a SDSS is having a process for identifying possible input scenarios that will lead to the desired future state. Similarly, Lempert et al. (2006) make the case for robust decision-making illustrated with an application of pollution-control strategies. Here "robust" refers to policies that yield acceptable outputs across a wide range of potential future conditions. Their solution is to take a more goal-oriented approach, in their words "anticipatory", and focus on the desired outcomes with an iterative process to design strategies that yield desirable outcomes across a wide range of possible futures.

Participatory GIS especially can benefit from this goal-oriented approach (Goosen et al, 2007; Ramsey 2009). While Couclelis (2005) suggested that "planning with the people" is too complex, Buchecker et al, 2003 found including local residents in planning results in more sustainable development and only requires a somewhat more lengthy initial process. Further, Goosen et al, (2007) and Ramsey (2009) discovered including local residents early on in the process of decision-making helps reduce unnecessary conflicts and clarify the values and preferences of the parties involved. Additionally, Nyerges (2006) found that a low-technology configuration of their decision support system promoted more deliberation, while a medium-technology configuration promoted more analysis. The greatest challenges to developing comprehensive decision support are the processing requirements for generating a wide range of possibilities, finding effective ways to visualize and communicate the multiplicity of scenarios, and, most critically, carefully defining the problem space to be considered. A goal-oriented GDSS provides useful support in addressing all three of these challenges.

The SDSS framework described in the following section combines the strengths of the input-oriented and goal-oriented approaches reviewed above, by applying the three-element decision frame shown in Figure 1 to SDSS design. While the goal-oriented approach supports robust decision-making unmet with the input-oriented approach, it does not replace the strength of the input-oriented approach – that is, answering very specific decision-maker questions focused on alternatives. Further, the decision framework presented in Section 2.1 makes clear the importance of including functions that support user interaction with all three decision elements (alternatives, information, values).

4.3 Functional Design Framework for a SDSS

In this section, we identify common functions present in current designs and implementations of SDSS, based on a review of the literature describing systems that have been implemented. Functions are treated here as an organizing concept that describes a grouping of tools that have a common purpose. For example, a function could be “collect information” and some of the tools within this function would include zoom, pan, and identify. We consider these functions as described in the SDSS literature and classify them by which of the three decision elements in Figure 1 they address: alternatives (inputs), state, or outputs/values. In addition, there exists a fourth class of functions that is inherent in all SDSS – data management functions. Functions that support basic data management functions are an integral part of an SDSS, but are not prescribed in detail as they are inherent in any data management system. Additionally, we address how these function classes fit into the alternatives, information, values decision frame outlined in the previous section. Finally, we present a case study illustration of the SDSS development approach recommended in this paper.

4.3.1 Analysis of Current SDSS Functions

From a content analysis of current SDSSs conducted initially in June 2012 and updated in July 2014 we compiled a list of eleven common functions currently found in SDSS design. These functions are identified by first listing and describing all functions found in the systems considered, and then grouping these functions into common categories. This grouping yielded eleven common categories or types of function. We searched the term "spatial decision support system" as a topic in Web of Science – a respected literature search engine, which yielded roughly 1300 resulting papers. We narrowed these papers to those in Geography – approximately 169 papers. These 169 papers were sorted by highest number of citations first, and then the first 50 were selected for the analysis. This initial sample size was determined because it was large enough to be representative, and the papers beyond the first 50 were becoming less and less related to the topic, with few of them describing actual SDSSs. Of the 50 papers sampled, 42 described a specific spatial decision support system in sufficient detail to categorize the functions included in the system – these 42 papers are the sample used for this content analysis. The results of the general function analysis are outlined in Table 1.

The first three general functions listed in Table 1 relate to the state variables in the three-element decision analysis framework. These functions either support collecting information about the problem space or exploring and describing this information. The fourth function listed allows users to input criteria weights for optimization and/or preference analysis, which relates to the alternatives element of the decision framework.

The fifth, sixth, and seventh general functions listed relate to decision outcomes, or the values part of our decision framework. These functions support users in identifying, generating via exploration and/or discussion, and evaluating desired goals or decision outcomes. The eighth general function was one of the more common functionalities, and simply supports generating an output map. This function supported all three elements of our decision framework in different systems. The last three functions listed support the generation and evaluation of decision scenarios – the generation and evaluation of scenarios were by far the most commonly included functions. Less common was a function to generate statistics on these scenarios, to assist the evaluation.

Table 14. Classification of tool functions found in the literature for spatial decision support systems.

Class		**Functions**	**Number of Times Functions Found in 42 Systems**
State	1	Collect Information About Problem	8
	2	Describe/Explore Problem (construction of problem space)	16
	3	Explore patterns in complex data	5
Alternatives	4	Set Criteria Weights	7
Values	5	Input Goal preferences	6
	6	Generate Desired Goals	9
	7	Evaluate Desired Goals	3
All 3 Parts of Frame	8	Output Map	20
	9	Generate scenarios	28
	10	Evaluate scenarios	14
	11	Generate Statistics on Scenarios	2

By considering these functions in light of our three-element decision framework, three distinct broad classes of functions become apparent. The first class of function is referred to as Exploratory – these functions support the exploration and description of state variables, or in other words, the current world conditions surrounding the problem. The second broad class of function we refer to as Input-Oriented functions. These functions focus on the alternatives side of the decision frame. The alternatives are the more intuitive input into the system, and fit into the predict-and-plan approach that has dominated decision-making until recently. The final broad class of function is described as Goal-Oriented. These functions focus on the values of the three-element decision frame, and speak to the value-focused approach that decision science recommends for decision-making. It is our recommendation that SDSS include at least one function from each of these three classes to fully support decision-making. By including functions from each of these general classes, the resulting system can address each of the three elements in our decision framework and support a wider range of decision maker needs.

By considering the 42 systems of the analysis in terms of these three function classes, we can delineate the current configurations of SDSS. This delineation is outlined in Table 2. In sum, the analysis shows that the most common (43%) configuration of functions included in current spatial decision support systems belongs to a combination of exploratory and input-oriented classes, and 70% of the systems sampled include at least one function that fits within the input-oriented class. Only a fifth of the systems sampled (36%) include value-oriented tools at all, and only 5 of the 42 papers included at least one function from all three classes as is suggested in this paper. The results of this literature analysis illustrate the under-representation of systems with the SDSS framework recommended in this paper.

Table 15. Classification of systems based on the tool types included in the system.

Function Class Combinations	Function Numbers from Table 14	Number of systems (out of 42)	% of sample
Exploratory Only	1, 2, 3	3	7%
Input-oriented Only	4	5	12%
Goal-oriented Only	5, 6, 7	2	5%
Exploratory + Input-oriented	1, 2, 3, 4	18	43%
Exploratory + Goal-oriented	1, 2, 3, 5, 6, 7	7	17%
Input-oriented + Goal-oriented	4, 5, 6, 7	1	2%
Exploratory + Input-oriented+ Goal-oriented	8, 9, 10, 11	5	12%

4.3.2 Case Study Illustration: Nitrogen Decision Support (NitroSim)

Our case study decision support system is NitroSim. It represents an illustration of a SDSS developed following the three-element approach recommended in this paper by consider a case study supporting watershed management with regards to nonpoint source nitrogen pollution in the Niantic River watershed in Connecticut. This particular case study is but one of any number that could be well served by the three-element SDSS configuration, and was chosen as a suitable illustration for several reasons. First, there are known negative impacts from nitrogen pollution in the area, and decision-makers in the area are concerned with these impacts. While there is concern, different stakeholders have different views or goals, which makes the decision process more complicated and necessitates decision support tools. Further, the watershed selected includes a number of lakes and ponds, and is coastal, but not immediately adjacent to the coastline. These attributes are important for the nitrogen modeling and the specific case study. Finally, this case study demonstrates that leaving out any one of the three classes limits the types and range of scenarios that users can explore; thus the importance of including the complete suite of options.

Increased nitrogen degrades the water quality of coastal estuaries by causing eutrophication (Conley et al, 2009; Howarth et al, 2000). Eutrophication is increased growth of plankton and algae with the adverse effect of oxygen reduction in the affected waters. This oxygen reduction causes natural habitat degradation and death for native shellfish and fish (Ryther and Dunstan, 1971). This is of obvious concern for decision makers whose primary concern is environmental quality or preservation, but is also of concern for those focused on economic development – both because the loss of native shellfish and fish has a negative impact on the local fishing and tourism economy, but also because measures to mitigate the pollution can make development more costly. Regulatory issues are also driving watershed protection. The CT DEP classifies the Niantic River as impaired (303(d) list) due to nutrients and bacteria, the first step in a TMDL process which will document and structure clean-up efforts (CT DEP 2006). Decision makers concerned with economic development need a tool that helps them navigate those constraints to be able to develop within the watershed in a way that maximizes economic gains, without undue costs incurred by increased nitrogen pollution damage.

Nitrogen occurs both naturally and as a man-made pollutant. Nitrogen pollution (sources) in the study area comes primarily from septic system based residential and farmland land uses. Nitrogen is removed (sinks) from the watershed through natural biotic processes (microorganism based denitrification) within streams, wetlands, and ponds, or best management practices (BMP) including septic system design (Oakley et al, 2010), stormwater retention basins (Collins et al, 2010), and denitrifying bioreactors (Schipper et al, 2010)

The main waterbody of the Niantic River watershed is the Niantic River. The topography in the Niantic River watershed is mild and the terrain is formerly glaciated with several naturally formed lakes and wetland areas. This watershed includes four municipalities: Salem, Montville, East Lyme, and Waterford and a number of lakes, ponds, and reservoirs (Figure 2). It has a drainage area of approximately 20,000 acres (nianticriverwatershed.org). The Niantic River ends in an estuary that drains into Long Island Sound.

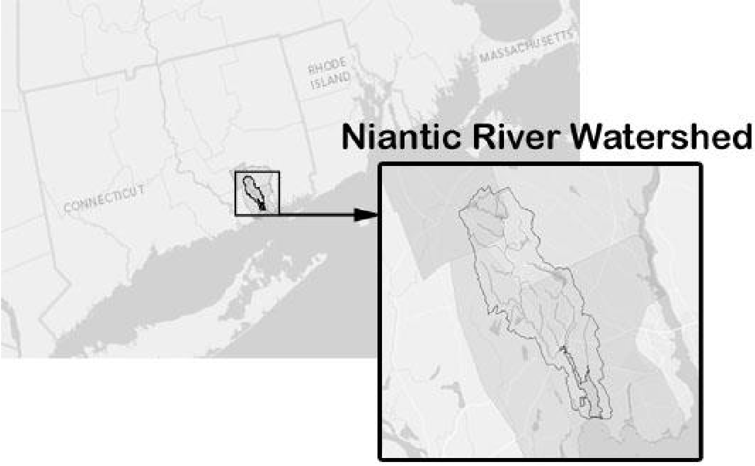


Figure 23. Niantic River Watershed.

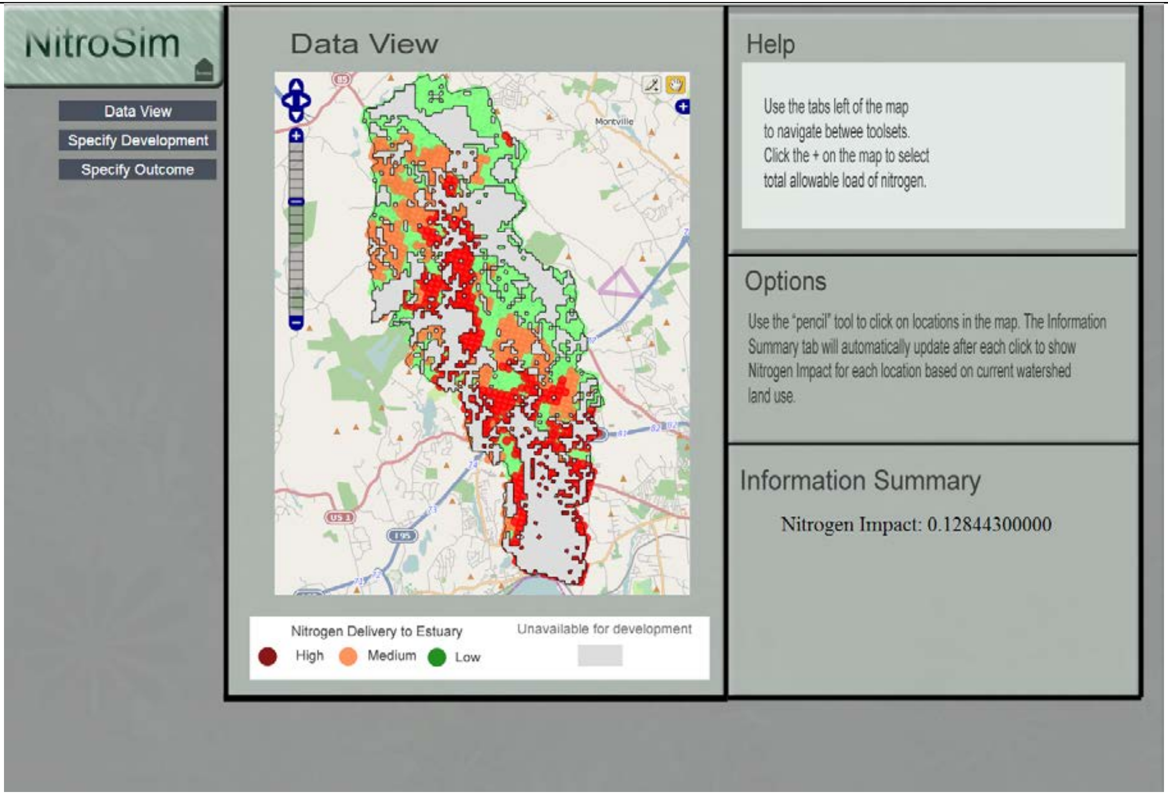
Due to the long residence time and proximity to agricultural runoff and septic systems the estuary at the end of the Niantic River watershed is especially sensitive to eutrophication, and native eelgrass beds and local fish populations are affected by increased anthropogenic nitrogen in the watershed (CT, DEP, 2006). Because further efforts to mitigate this nitrogen are needed, the local communities are supportive, and the needs of the decision-makers are varied, this watershed makes a good case study location for illustrating the usefulness of the three-element SDSS for decision-making with regards to non-point source nitrogen pollution.

4.3.3 Case Study Illustration: NitroSim functionality

It is our recommendation that a SDSS include at least one function from each of the three classes identified previously to fully support decision maker needs. We developed NitroSim with this in mind. While the illustrated functions from each of the three classes are described sequentially, there is no implied order of operation. The advantage to having a single comprehensive system over three distinct software systems is primarily the flexibility in the system’s ability to answer a range of different questions and to meet a wider range of user needs within one setting. This means that users only learn one software package and have the potential to ask a full range of questions within that single package. More functions are possible in each of the three classes than are shown here; the functions selected for this case study were chosen to be representative of functions in each class, and for their usefulness to decision makers in the New England case study area.

The user objective of the exploratory function class of a SDSS is to understand the breadth and depth of the problem including the problem context, data sources, and model parameters. The purpose of this class is primarily informative and allows decision-makers to familiarize themselves with the state variables (Figure 1) of the decision being addressed and to explore the available data. In the NitroSim illustration, the exploratory function allows users to toggle between different map views that highlight different information within the watershed. (Figure 3) More specifically, the

Data View highlights the sinks within the watershed and visualizes the amount of nitrogen removal for each 30mx30m location. Users can click on each location and see the amount of nitrogen that will be delivered to the watershed outlet from each location.



Use the tabs left of the map
to navigate betwee toolsets.
Click the + on the map to select
total allowable load of nitrogen.

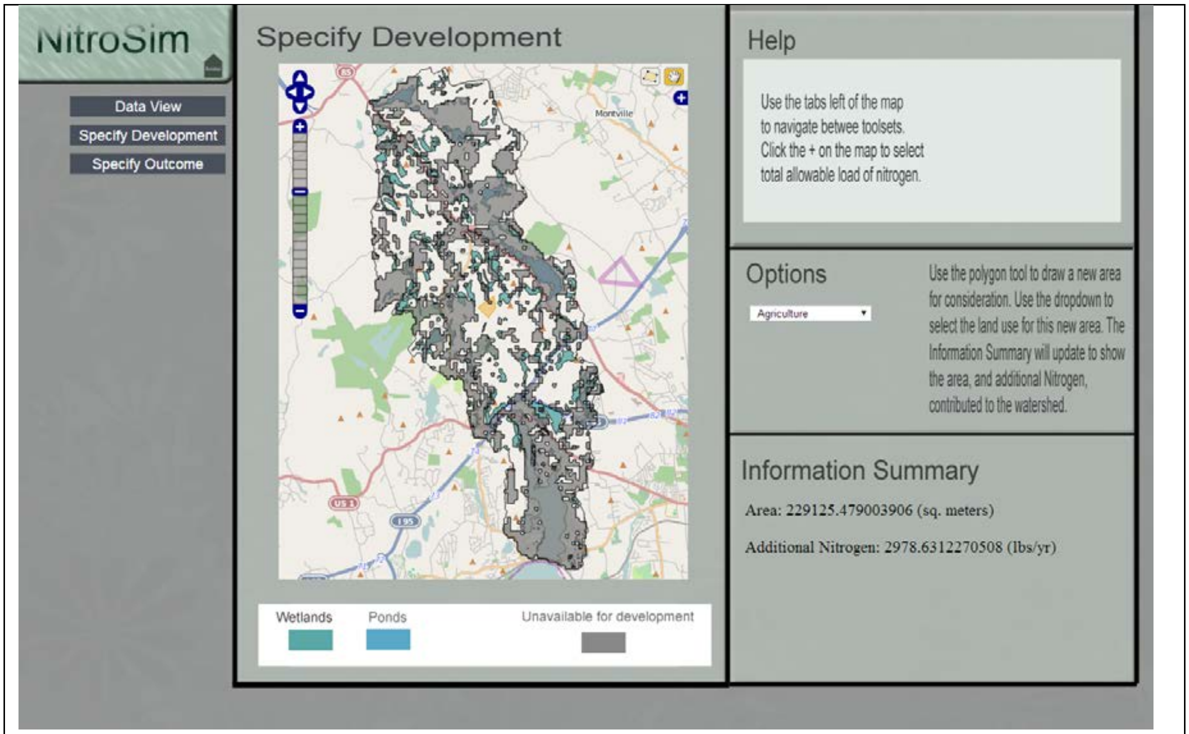
Use the "pencil" tool to click on locations in the map. The Information Summary tab will automatically update after each click to show Nitrogen Impact for each location based on current watershed land use.

Nitrogen Impact: 0.12844300000

Figure 24. Data View of the NitroSim SDSS. Users can click on points to find out the percent nitrogen contributed to the estuary from that point. The colored dots indicate high, medium, and low delivery to the estuary.

The input-oriented class of functions for a SDSS provides support for specific what-if questions and exploring the alternatives of the decision frame. This class is useful for supporting decisions regarding the likely outcome of a specific dataset and a set of model parameters within a specific geographic location. With this feature decision-makers specify a set of criteria, and learn the predicted outcome based on them. The user objective is to compare the results of model output based on varying inputs including data sources, policy choices, and model parameters.

NitroSim includes an input-oriented function that provides users with a map-based interface that allows them to specify land use type, location, and area for a specific land-use change (Figure 4). The user begins with selecting the land-use type and then uses a drawing tool to outline the desired area for the development. This function then provides the user with an output map that highlights the selected area and visualizes the nitrogen impacts. The information is also provided in a summary screen beside the map.



Use the tabs left of the map
to navigate betwee toolsets.
Click the + on the map to select
total allowable load of nitrogen.

Use the polygon tool to draw a new area for consideration. Use the dropdown to select the land use for this new area. The Information Summary will update to show the area, and additional Nitrogen, contributed to the watershed.

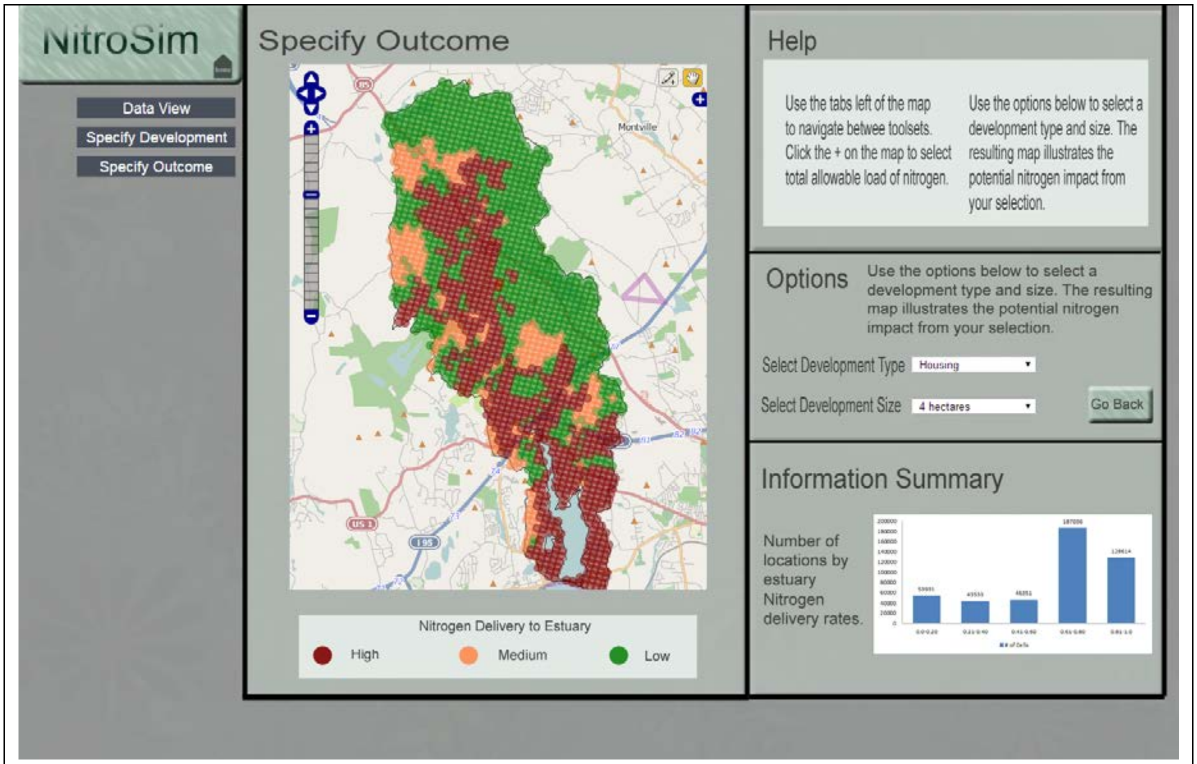
Area: 229125.479003906 (sq. meters)

Additional Nitrogen: 2978.6312270508 (lbs/yr)

Figure 25. The Specify Development view of NitroSim provides users with input-oriented functionality. Users can specify land use change type, area, and location, and see the resulting nitrogen impact.

The goal-oriented class of functions focuses on the values of the decision frame – in other words, the desired decision outcomes. In this class, decision-makers focus on a desired outcome, and find out what options are likely to lead to that outcome. The user objective is to understand the range of outcomes the model and data report, and to provide insights into the variety of alternatives that meet a certain goal.

In NitroSim, the desired outcomes are based on overall nitrogen impact. The goal-oriented function demonstrated asks the user to first select a desired land use size and type (Figure 5). This generates a histogram that shows the full range of possible parcel combinations delineated by nitrogen impacts. This histogram is linked to a map that shows the potential locations within the watershed. By selecting a specific nitrogen impact, or goal, on the histogram – all possible locations that satisfy that goal are correspondingly highlighted on the map.



Use the tabs left of the map to navigate betwee toolsets. Click the + on the map to select total allowable load of nitrogen.

Use the options below to select a development type and size. The resulting map illustrates the potential nitrogen impact from your selection.

Use the options below to select a development type and size. The resulting map illustrates the potential nitrogen impact from your selection.

Select Development Type: Housing

Select Development Size: 4 hectares

Number of locations by estuary Nitrogen delivery rates.

Figure 26. The goal-oriented function of NitroSim allows users to indicate a desired land use change type and size and to see the results in map view and binned in a histogram. The map and histogram are linked so that the user can select only those locations on the map that satisfy a desired outcome.

4.4 Survey Evaluation

To evaluate NitroSim with respect to the three-function framework, a user survey was deployed that presented participants with three different decision scenarios. Each scenario is accompanied by the output from the suite of tools developed for NitroSim. The results of the survey are presented here and a concluding section reiterates the importance of including functions from all three function classes.

4.4.1 Survey Participants

Survey participants were recruited using the email listservs for graduate students in the School of Geographical Sciences & Urban Planning at Arizona State University. This included both current and recently graduated masters, Ph.D. and professional masters students. This was a reasonable population to sample for the survey because these are people who are exposed to ideas of geospatial technology and planning and who may well become decision makers in their career. To test the tools we wanted to examine the response from people who would be familiar enough to be comfortable with spatial technology, but who are not experts in SDSS design. 31 participants completed the survey for this study. The average age of participants was 26.5 years with a standard deviation of 6.3 years. 53.3% of participants had already earned a Master's degree, 33.3% had a 4-year degree, and 13.3% had already earned a Doctoral degree. All participants had some experience with GIS, and most had at least 1 year of experience. Specifically, 30% of participants had 5+ years experience, 33% had 2-5 years experience, 23% had 1-2 years experience, and 14% had less than 1 year experience. Experience with planning was more varied, with 36.7% of participants having no experience with city planning. 33.3% had less than 1 year experience, 16.7% had 1-2 years, 10% had 2-5 years, and only 1 participant (3%) had 5+ years experience.

4.4.2 Scenarios

Survey participants were presented with each of the following three scenarios, but in random order. When considering the survey results, participants are grouped by the first scenario presented, which is viewed as a training scenario wherein the participants familiarize themselves with the available options.

The Incentive Scenario informs participants they are making a decision about water policy with consideration for the economic potential of development and the potential negative environmental impact from nitrogen. The specific policy decision they are considering is where to establish an incentive for development within specific zones in the planning area. They are told the goal is to maximize economic growth, while minimizing the potential damage to coastal estuaries. The participant's decision is to identify one of three potential zones as the one to receive the incentive.

The Development Scenario informs participants that a housing developer is seeking approval of a 4-hectare housing subdivision. The developer wants to avoid the possibility of extra fees by developing in areas of estuary sensitivity and the participants are tasked with deciding where to approve the development. Their decision options are: Blanket approval – the developer can build anywhere in the planning area; Approve one specific location that you know to not be sensitive to nearby estuaries; Approve all locations in the city that are not sensitive to nearby estuaries; Do not approve the development.

The Zoning Scenario presents the participants with a decision about a 2-square-mile piece of land. In this scenario, two prominent local groups have been vocal about the zoning for this land. One group wants the land designated a nature preserve to help preserve the nearby estuary, and the other group wants to build luxury housing. Participants are reminded that they are tasked with maximizing the economic potential in their planning area, while minimizing the impact on nearby estuaries. They are given two decision options: Approve the area for luxury housing; Approve the area for a nature preserve.

4.4.3 Survey design

The survey was deployed using a website called surveygizmo.com. Participants accessed the survey using a link from a recruitment email and completed the survey on their own using an internet browser. There was no time limit for completing the survey. The survey questions in their entirety, along with summaries of user responses, are included in Appendix B. Participants are presented with a welcome screen, asked demographics questions, and then read an overview that explains their role in the survey. Specifically, participants were told:

> You are a planner in a coastal area that is trying to balance economic growth with negative impacts on nearby coastal estuaries. Several parts of the area have a stronger negative impact on the nearby estuaries than other parts due to their location.
>
> You'll be asked to complete three different decision tasks using the tools provided, and then you'll be asked to reflect on the tools and your experience. You'll be presented with the option to use any combination of three tools with each task. It is up to you which tools you look at and the order you do so (similar to a choose-your-own-adventure if you're familiar with that concept). The tools you will be presented with are a basic overview tool, a more specific tool that looks at one or a few particular decision outcomes, and a tool that looks at a broader range of decision outcomes. You will see output from each

> tool you select - the survey is not designed for you to actually interact with the tools.
>
> The survey should take no more than 45 minutes to complete. It is a bit more thought-intensive than your typical survey.

Participants were then presented with the first of three scenarios, in random order. Each scenario is presented in the same fashion starting with an overview that explains the scenario and the decision to be made. From that point the participants can select any of the three SDSS tools to see their output, or to make their decision. The three tools included are an Exploratory Tool (state variables), a Specific Outcome Tool (alternatives), and a Broad Outcome Tool (values). The survey participant is in control of the survey flow from this point and is able to view any combination of the three SDSS tool outputs provided, in any order, and as many times as they like. They may view the same tool multiple times, and this is not considered going “backwards” or deviating from the intent of the survey. When they are prepared to make a decision they select “make a decision” and are taken to the decision screen. This screen presents the scenario overview again and prompts the participant to select their decision. Once the selection is made the participant is asked to answer a series of questions about the decision process they just completed. These questions include ranking the tools in order of most to least helpful, rating the importance of each tool, and rating their comfort making the decision. After completing these reflection questions, the participants are presented with the next scenario and repeat the process.

4.4.4 Survey results and implications

The survey results are analyzed by comparing participants divided into three groups delineated by the first scenario each group was presented with, which was assigned randomly. Group 1 includes 9 participants and started with the Incentive Scenario, Group 2 includes 12 participants and started with the Zoning Scenario, and Group 3 includes 7 participants and started with the Development Scenario. 4 participants' results were not included in the analysis because they went backwards in the survey after making an initial decision and changed their decision. The survey is not designed to accommodate this behavior and it is not clear what caused these 4 participants to change their decisions. 2 other participants did not use at least one tool to make their decision in at least one of the scenarios. Because they did not use the tools, it is not clear what they used to inform their decision. For consistency, we have excluded these 6 participants from the analysis.

The first scenario each group is presented is considered a training scenario, within which the participant familiarizes themselves with the survey and the SDSS tools. Therefore, we evaluate the results for the latter two scenarios in each grouping.

First, we assess the use of tools in each scenario. It was found that all groups, regardless of the initial scenario, tend to use all three of the tools available for each scenario (Table 3). This finding might support our hypothesis that including tools that speak to each of the three elements in our framework is valuable and desired by decision makers. It is also possible the participants surveyed are confused by the survey and therefore gathering as much information as possible.

To address this latter possibility, each participant is asked to rate their comfort making the decision for each scenario using a 4 option scale: very comfortable, somewhat comfortable, somewhat uncomfortable, and very uncomfortable. Using the independent samples T-test by converting these to a 1-4 scale, there is found to be no statistical difference between the groups in rating their comfort making the decisions. Indeed, the majority of participants in each grouping rated their comfort either somewhat or very comfortable. The one exception to this is Group 3, which was presented with the Zoning Scenario first. In that scenario, slightly more of the participants reported being somewhat or very uncomfortable, which is not altogether surprising since that scenario is the most ambiguous of the three and it was the first this group encountered. Because the majority of participants reported feeling comfortable with their decision making, and there was no statistical difference in the reported rating between groups, it is deemed likely to the authors that the participants' tendency to use of all of the tools was not due to confusion within the survey, but instead because the participants found benefit from using all three.

When asked to rank the three tools in order of importance, a slight difference between the groups is found. Group 3, which started with the Zoning Scenario, is notably different in rankings from Groups 1 and 2. This is interesting because their initial training scenario was the most ambiguous of the three scenarios. The rankings are summarized in Table 4, where a rank of 1 is most important and rank of 3 is least important. The noted difference is in the ranking of the tools in the Zoning Scenario. The group that encountered this scenario first rated the Exploratory tool (state) the most important, while the other two groups rated the Specific Outcome tool (alternatives) the most important. Besides this exception, the Specific Outcome tool (alternatives) was rated the most important by all of the groups. This type of tool is the most commonly developed tool in SDSS today; perhaps this familiarity affected the participants' ranking, or perhaps these tools really are the most important type to provide.

Table 16. Results of user survey of different tool classes for SDSS usability.

Group	Initial		Incentive Scenario			Development Scenario			Zoning Scenario		
	Tool:		Exploratory	Specific Outcom	Broad Outcom	Exploratory	Specific Outcom	Broad Outcom	Exploratory	Specific Outcom	Broad Outcom
1	Incentive	Rank	2	1	3	2	1	3	3	1	2
2	Development		2-tie	1	2-tie	3	1	2	2-tie	1	2-tie
3	Zoning		3	1	2	2	1	3	1	2	3

4.5 Conclusions

The implications of this study extend beyond the illustration demonstrated in this paper. The conceptual framework we have presented can be used to evaluate and develop SDSS that support all three elements of decision-making. While decision theory has long called for attention to the values of decision-making and goal-oriented functions and approaches to SDSS design, these functions remain the least developed. Our user survey showed that participants are likely to use these tools when they are provided, though they are likely to put them second in importance to the more commonly developed input-oriented functions. This information is particularly useful for applying a geodesign approach to SDSS when considering the impacts of using different tool configurations. Geodesign as a field draws on the pivotal work by McHarg (1969) by applying the concept of overlay to design options in the development of spatial technologies. The framework presented in this paper furthers these endeavors by more clearly delineating design options and the potential design impacts from the use or absence of such options.

Neglecting to include functions from all three categories limits the decision scenarios a system can support, and it is especially important to include goal-oriented functions for decisions where the values are the part of the decision that carries the most concern for decision makers. Future research possibilities include developing more solutions to the computational and visualization challenges of goal-oriented tools and more detailed user surveys to learn more about decision maker needs and preferences.

CHAPTER 5
CONCLUSIONS

5.0 Summary

This dissertation advances spatial decision support system development theory by using a geodesign approach to consider the impacts of design alternatives for such systems. The development of a SDSS is a multi-faceted process, which requires special attention to the choice of system components such as the spatial model, technical data, and tools employed to generate the system. These components are closely linked and any variation in these components creates the potential for a very different decision support experience for the end-users of the resulting system. The emerging field of geodesign offers a new approach to the design, implementation, and evaluation of SDSSs with emphasis on the impacts of design alternatives. This dissertation has applied this approach to the design, implementation, and evaluation of a case study SDSS, but in such a way that the techniques used are not specific to the case study demonstrated, and can be more generally applied.

5.1 Discussion

This dissertation uses a water management case study to demonstrate the design, implementation, and evaluation of three components of SDSS design: the spatial model, the technical data, and the tools. The aim of this project was to develop a broader approach to SDSS design while considering design alternatives, such that the design and evaluation of each of the three components presented herein can be applied to other SDSS design cases.

In general, the spatial model used in a SDSS has a significant impact on the resulting system because the model is responsible for simulating and predicting decision outcomes. Consequently, it is important to validate the model and consider any errors in terms of how the model is used in the SDSS.

The spatial model designed for this project simulates nonpoint source pollution loading, transport, and removal through a watershed. The loading values can be modified to fit different cases and in this study are set to values for housing and agriculture in the New England case study area. The removal values relate to biogeochemical processes in the watershed, which are also customizable to different cases as needed. For the illustration case study, the spatial model presented in **Chapter 2** simulates nonpoint source nitrogen loading, transport, removal, and ultimate delivery to an estuary at the watershed outlet. The simulated values were validated using actual measurements of nitrogen taken by the USGS in 2005.

This validation revealed that the model was reasonably accurate, but could be improved. The most likely options for improvement are the nitrogen loading values, improved accuracy of sink representation, improved accuracy of source representation, and increasing the precision of flow estimates for simulated flow within the watershed. With the current configuration, the model overestimates nitrogen loading. This error impacts the SDSS because any decision alternatives modeled with the SDSS will also overestimate nitrogen loading. These overestimated decision outcomes used by decision makers could result in unnecessarily conservative decisions with consideration for water quality. Consequently, it is important to communicate this to decision makers.

The technical spatial data used in developing a SDSS also have a significant impact on the resulting system. The data combined with the model are responsible for simulating and predicting the decision outcomes of different decision alternatives. As such, it is important to assess the fitness of the data used in developing SDSSs and in geodesign.

For this project, the technical spatial data include the DEM, delineation of the stream network, the spatial attributes of nitrogen sources including housing and agriculture, and the spatial attributes of nitrogen sinks including streams, wetlands, and ponds. **Chapter 3** of this dissertation presents a new method for assessing the fitness for use of such data, with an illustration using the wetland data of our case study.

This assessment showed the SSURGO wetland data are the most fit for use in the SDSS. The impacts of using the other options for wetland delineation on the SDSS are similar to the impacts of the model limitations. That is, since the simulated decision outcomes based on these data are what decision makers are using to make decisions, using a wetland data set that omits wetlands that actually exist could result in decisions that are overly conservative. Conversely, using a wetland data set that includes wetlands that do not exist could result in decisions that are not conservative enough, with regards to water quality and environmental degradation.

The final design component considered in this dissertation is the types of tools included in the SDSS interface, where the term tools is used to describe the interface options that allow users to store, manipulate, analyze, and display decision alternatives, decision outcomes, and decision values.

Chapter 4 evaluates the tools included in current SDSSs. This assessment classifies current tools into one of three categories, depending on which part of the decision framework the tool supports: those that support state variables, or exploring the current status of the problem, tools that support decision alternatives, and tools that support decision values. Chapter 4 also describes implementation of our case study SDSS with tools from each of the three categories. The usability of each of the included tools is tested with a user survey, with the tools supporting alternatives being ranked the most important by the majority of survey participants. This finding is consistent with current SDSS designs, as most systems today focus on tools in this category.

Including or omitting tools from a SDSS implementation has the potential to impact how decision makers think about and approach a problem, the number of decision alternatives that can be realistically considered, and how useful the system is perceived. While decision science theory supports systems with a more value-centric approach, our user survey showed participants prefer system tools that follow the alternatives-focused approach. The former is more useful for considering a wider range of decision alternatives and focuses on decision outcomes, while the latter is probably more familiar and intuitive. The challenge in this design alternative is including tools in the SDSS that are both useful and will actually be used by decision makers.

5.2 Implications

This dissertation contributes to the emerging field of geodesign by considering specific design alternatives in the development and evaluation of SDSSs. Specifically, this project considers three design alternatives: spatial model, technical spatial data, and tools. The findings of this project are general enough to include all SDSSs, regardless of the case they are built for, which is valuable in a field where much of the work is case-focused.

This dissertation presents a spatial model that can be customized to other cases of nonpoint source pollution, a methodology for assessing data fitness for use in a SDSS setting, and a conceptual framework for classifying and evaluating tools included in SDSS interfaces. More specifically, this dissertation also presents an illustration SDSS with a specific case study of watershed management in the Niantic River watershed of Connecticut, USA. The dissertation develops a geodesign approach to SDSS design and evaluation of spatial models, technical spatial data, and tools, as well as a spatial model that could be applied to other nonpoint source pollution scenarios.

5.3 Directions for future research

This study applies a geodesign approach to three aspects of SDSS design: the spatial model, the technical spatial data, and the tools included in the interface. There are more aspects in SDSS design, however; and this study does not address the platform used, user experience, the inclusion of uncertainty, or a public participation component.

The platform options can be divided into two categories: web-based and desktop systems. Web-based systems have the advantage of versatility of access, while desktop systems can be more computationally intensive. Each platform has advantages and disadvantages, which bear consideration in a geodesign approach to SDSS development.

User experience is also an important consideration – especially in concert with the selection of tools included within a SDSS. While it should be fairly obvious that more advanced users allow for the inclusion of more advanced tools, it is less clear what tool combinations is most beneficial for users with a mix of experience.

Uncertainty should also be addressed. Recent studies have shown the value of included uncertainty in visualizations, especially in the area of decision making, and SDSS development will benefit from the consideration of uncertainty.

Each of these aspects of design is important, but the public participation component is probably the most important component for future consideration. It is this component that speaks to the geodesign imperative for collaboration and sharing in the development of SDSSs. Including end users in the design process also better ensures usability of the resulting system, perhaps gives the most positive user assessment of their experience relevant to SDSS design, and generally results in the SDSS more closely reflecting user needs. Several benefits of public participation have been shown with specific cases, and it is an important avenue for continued investigation.

Beyond the suggestions highlighted above, an additional direction for future research is with generic tools. For example, there is current research in tools that allow the integration of disparate models, simulations, and visualizations. Such tools could be applied to SDSS design to allow more flexibility and range in the resulting system.

Additional consideration should be paid to understanding the usability of value-centric tools and why users are not rating their value as highly as theory suggests they should be. While decision theory suggests value-centric tools are important to supporting a wide range of decisions, and especially decisions that involve uncertainty, users and current SDSS design appear to be resistant to value-centric tools. Understanding why could allow designs to include such tools in a way that users find beneficial.

REFERENCES

Chapter 1 References

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, A., et al. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science, 21*(8), 839-858.

Armstrong, M. P. 1993. Perspectives on the Development of Group *Decision Support Systems* for Locational Problem-Solving. *Geographical Systems*, 1:69-82.

Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of Information Technology, 20*(2), 67-87.

Blaser, A. D., Sester, M., & Egenhofer, M. J. (2000). Visualization in an early stage of the problem-solving process in GIS. *Computers & Geosciences, 26*(1), 57-66.

Coppock, J. T., & Rhind, D. W. (1991). The history of GIS. In D. J. Maguire, M. F. Goodchild & D. W. Rhind (Eds.), (pp. 21)

Couclelis, H. (2005). "Where has the future gone?" rethinking the role of integrated land-use models in spatial planning. *Environment and Planning A, 37*(8), 1353-1371.

Crossland, M., Wynne, B., & Perkins, W. (1995). Spatial decision-support systems - an overview of technology and a test of efficacy. *Decision Support Systems, 14*(3), 219-235.

Dangermond, J., (2013) Keynote, 2013 Geodesign Summit (Esri, Redlands, CA) http://www.geodesignsummit.com/agenda/agenda.html

Densham, P.J. and Armstrong, M.P., 1987. A spatial decision support system for locational planning: design, implementation and operation. Proceedings, Auto Carto 8 (ACSM/ASPRS, Bethesda) pp. 112-121.

Goodchild, M. F. (1992). Geographical information-science. *International Journal of Geographical Information Systems, 6*(1), 31-45.

Gregory, R. S., & Keeney, R. L. (2002). Making smarter environmental management decisions. *Journal of the American Water Resources Association, 38*(6), 1601-1612.

Howard RA, 1968, "Foundations of Decision Analysis" Ieee Transactions on Systems Science and Cybernetics **SSC4** 211

Howard RA, 2007, Advances in Decision Analysis (Cambridge University Press, Cambridge)

Jankowski, P., Robischon, S., Tuthill, D., Nyerges, T., & Ramsey, K. (2006) Design considerations and evaluation of a collaborative, spatio-temporal decision support system. *Transactions in GIS*, Vol. 10, Issue 3.

Keen, P.G.W. (1977) Implementation Research in OR/MS and MIS: Description versus Prescription. *Operations, Information &* Technology, Paper No. 390

Keeney, R. L. (1994). Using values in operations-research. *Operations Research, 42*(5), 793-813.

Keeney, R. L. (1992). *Value-focused thinking :A path to creative decisionmaking*. Cambridge, Mass.: Harvard University Press.

Keeney, R. L. (1996). Value-focused thinking: Identifying decision opportunities and creating alternatives. *European Journal of Operational Research, 92*(3), 537-549.

Nyerges, T. (1993) *How Do People Use Geographical Information Systems?,* in Human Factors in Geographical Information Systems, edited by David Medyckyj-Scott and Hilary Hearnshaw, London: Belhaven Press, pp. 37-50.

Nyerges, T., Jankowski, P., Tuthill, D., & Ramsey, K. (2006). Collaborative water resource decision support: Results of a field experiment. *Annals of the Association of American Geographers, 96*(4), 699-725.

Power, D. J. (1999). A brief history of decision support systems. *World Wide Web, Http://dssresources.com/history/dsshistory.Html, Version, 1*

Ramsey, K. (2009). GIS, modeling, and politics: On the tensions of collaborative decision support. *Journal of Environmental Management, 90*(6), 1972-1980.

Sprague Jr, R. H., & Watson, H. J. (1996). *Decision support for management* Prentice-Hall, Inc. Upper Saddle River, NJ, USA.

Steinitz C, 2012 *A Framework for Geodesign: Changing Geography By Design* (ESRI Press, Redlands, CA)

Chapter 2 References

Armstrong, D.S. and G.W. Parker. (2003). Assessment of habitat and streamflow requirements for habitat protection, Usquepaug-Queen River, Rhode Island, 1999-2000. U.S. Geological Survery Open-File Report 02-438. 78 p.44

Batty M, 2013, "Defining geodesign (= GIS + design ?)" *Environment and Planning B: Planning and Design* **40**(1) 1 – 2

Collins, K. A., Lawrence, T. J., Stander, E. K., Jontos, R. J., Kaushal, S. S., Newcomer, T. A., et al. (2010). Opportunities and challenges for managing nitrogen in urban stormwater: A review and synthesis. *Ecological Engineering, 36*(11), 1507-1519.

Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., et al. (2009). Controlling eutrophication: Nitrogen and phosphorus. *Science, 323*(5917), 1014-1015.

Gilliam, J. W. (1994). Riparian wetlands and water quality. *Journal of Environmental Quality, 23*, 896-900.

Gold, A., DeRagon, W., Sullivan, W., & Lemunyon, J. (1990). Nitrate-nitrogen losses to groundwater from rural and suburban land uses. *Journal of Soil and Water Conservation, 45*(2), 305-310.

Gold, A. J., Groffman, P. M., Addy, K., Kellogg, D. Q., Stolt, M., & Rosenblatt, A. E. (2001). Landscape attributes as controls on ground water nitrate removal capacity of riparian zones. *Journal of the American Water Resources Association, 37*(6), 1457-1464.

Hill, A. R. (1979). Denitrification in the nitrogen budget of a river ecosystem.

Howarth, R. W., Anderson, D. M., Church, T. M., Greening, H., Hopkinson, C. S., Huber, W. C., et al. (2000). *Clean coastal waters : Understanding and reducing the effects of nutrient pollution*. Washington, DC, USA: National Academies Press.

Kellogg, D. Q., Gold, A. J., Cox, S., Addy, K., & August, P. V. (2010). A geospatial approach for assessing denitrification sinks within lower-order catchments. *Ecological Engineering, 36*(11), 1596-1606.

Marshall, S. S. (1994). Proposed management strategy for seychelles marine national parks. *Division of Environment,*

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., et al. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems, 6*(4), 301-312.

McHarg, I. 1969. Design with Nature. Garden City, N.Y.: The Natural History Press.

Oakley, S. M., Gold, A. J., & Oczkowski, A. J. (2010). Nitrogen control through decentralized wastewater treatment: Process performance and alternative management strategies. *Ecological Engineering, 36*(11), 1520-1531.

Ryther, J., & Dunstan, W. (1971). Nitrogen,phosphorus, and eutrophication in coastal marine environment. *Science, 171*(3975), 1008-&.

Schipper, L. A., Robertson, W. D., Gold, A. J., Jaynes, D. B., & Cameron, S. C. (2010). Denitrifying bioreactors-an approach for reducing nitrate loads to receiving waters. *Ecological Engineering*, *36*(11), 1532-1543.

Steinitz C, 2012 *A Framework for Geodesign: Changing Geography By Design* (ESRI Press,
Redlands, CA)

The Department of the Environment, Maryland. *Maryland's nitrogen-reducing septic upgrade program*. Retrieved 08/06, 2014, from http://www.mde.state.md.us/programs/Water/BayRestorationFund/OnsiteDisposalSystems/Pages/Water/cbwrf/index.aspx

Tulloch, D.L. (2013). Learning from students: geodesign lessons from the regional design studio, Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 6:3, 256-273, DOI: 10.1080/17549175.2013.765903

Yang, X. J. (2012). An assessment of landscape characteristics affecting estuarine nitrogen loading in an urban watershed. *Journal of Environmental Management, 94*(1), 50-60.

Chapter 3 References

Agumya, A., & Hunter, G. J. (1999). A risk-based approach to assessing the‘Fitness for use’ of spatial data. *URISA Journal 11*(No. 1)

Armstrong, D.S. and G.W. Parker. 2003. Assessment of habitat and streamflow requirements for habitat
protection, Usquepaug-Queen River, Rhode Island, 1999-2000. U.S. Geological Survery Open-File Report 02-438. 78 p.44

Aronoff, S. (1985). The minimum accuracy value as an index of classification accuracy. *Photogrammetric Engineering and Remote Sensing, 51*(1), 99-111.

Bedard, L. P., & Barnes, S. (2010). How fit are your data? *Geostandards and Geoanalytical Research, 34*(3), 275-280.

Chrisman, N. R. *Issues in digital cartographi quality standards: A progress report*

Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), 35-46.

Congalton, R., Oderwald, R., & Mead, R. (1983). Assessing landsat classification accuracy using discrete multivariate-analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing, 49*(12), 1671-1678.

Cowardin, L. M., Carter, V., Golet, F. C., and LaRoe, E. T. (1979), A Classification of Wetlands and Deepwater Habitats of the United States. Washington, DC: U.S. Government Printing Office, Office of Biological Services, Fish and Wildlife Service, FWS/OBS-79/31, stock number GPO 024-010-00524-6, 103 pp.

De Bruin, S., Bregt, A., & Van de Ven, M. (2001). Assessing fitness for use: The expected value of spatial data sets. *International Journal of Geographical Information Science, 15*(5), 457-471.

Devillers, R., Bedard, Y., Jeansoulin, R., & Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science, 21*(3), 261-282.

Finlayson, C., Davidson, N., Spiers, A., & Stevenson, N. (1999). Global wetland inventory - current status and future priorities. *Marine and Freshwater Research, 50*(8), 717-727.

Fisher, P. E., & Tate, N. J. (2006). Causes and consequences of error in digital elevation models. *Progress in Physical Geography, 30*(4), 467-489.

Gilliam, J. W. (1994). Riparian wetlands and water quality. *Journal of Environmental Quality, 23*, 896-900.

Gold, A. J., Groffman, P. M., Addy, K., Kellogg, D. Q., Stolt, M., & Rosenblatt, A. E. (2001). Landscape attributes as controls on ground water nitrate removal capacity of riparian zones. *Journal of the American Water Resources Association, 37*(6), 1457-1464.

Goodchild, M. F. (1992). Sharing imperfect data. *NCGIA Initiative 9: Institutions Sharing Spatial Information,* San Diego.

Guptill, S. C., & Morrison, J. L. A. (1995). *Elements of spatial data quality* Elsevier Science Limited.

Hill, A. R. (1979). Denitrification in the nitrogen budget of a river ecosystem.

Howarth, R. W., Anderson, D. M., Church, T. M., Greening, H., Hopkinson, C. S., Huber, W. C., et al. (2000). *Clean coastal waters : Understanding and reducing the effects of nutrient pollution*. Washington, DC, USA: National Academies Press.

Hunter, G., & Goodchild, M. (1995). Dealing with error in spatial databases - a simple case-study. *Photogrammetric Engineering and Remote Sensing, 61*(5), 529-537.

Jordan, N. R., Slotterback, C. S., Cadieux, K. V., Mulla, D. J., Pitt, D. G., Olabisi, L. S., et al. (2011). TMDL implementation in agricultural landscapes: A communicative and systemic approach. *Environmental Management, 48*(1), 1-12.

Kellogg, D. Q., Gold, A. J., Cox, S., Addy, K., & August, P. V. (2010). A geospatial approach for assessing denitrification sinks within lower-order catchments. *Ecological Engineering, 36*(11), 1596-1606.

Klein, G. A. (1999). *Sources of power: How people make decisions* MIT Press.

Lehner, B., & Doll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology, 296*(1-4), 1-22.

Lowry, J.B.C., 2006. The Application of Low-cost GIS Software and Data for Wetland Inventory, Assessment and Monitoring. Ramsar Convention Secretariat, Gland, Switzerland. http://www.ramsar.org/lib/lib_rtr_index.htm
Ramsar Technical Report No. 2.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., et al. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems, 6*(4), 301-312.

Qamer, F. M., Ashraf, M. S., Hussain, N., Saleem, R., Ali, H., Mirza, H., et al. (2009). Pakistan wetlands GIS - a multi-scale national wetlands inventory. *Proceedings of 33rd International Symposium on Remote Sensing of Environment,* Stresa, Italy.

Quinn, H., J. Tolson, J. Klein, S. P. Orlando, and C. Alexander. 1989. Strategic assessment of near coastal waters: Susceptibility of east coast estuaries to nutrient discharges: Albemarle/Pamlico Sound to Biscayne Bay. NOAA/EPA summary report. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Silver Spring, MD.

Rebelo, L. -., Finlayson, C. M., & Nagabhatla, N. (2009). Remote sensing and GIS for wetland inventory, mapping and change analysis. *Journal of Environmental Management, 90*(7), 2144-2153.

Sader, S., Ahl, D., & Liou, W. (1995). Accuracy of landsat-tm and gis rule-based methods for forest wetland classification in maine. *Remote Sensing of Environment, 53*(3), 133-144.

Skidmore, A. K., & Turner, B. J. (1989). Why areal accuracy is not correctly estimated using lowell's modification of the supervise nonparametric classifier. *Photogrammetric Engineering and Remote Sensing, 55*(6), 899-900.

Tiner, R. W. 1984. Wetlands of the United States: current status, recent trends. National wetlands Inventory, U. S. Fish and Wildlife Service, Washington, D. C.

Vasseur, B., Devillers, R., & Jeansoulin, R. (2003). Ontological approach of the fitness of use of geospatial datasets. *Proceedings of the 6th AGILE,* Lyon, France.

Veregin, H., 1999, Data quality parameters. In *Geographical Information Systems—Volume 1, Principles and Technical Issues*, 2nd edition, edited by P. A. Longley, M. F. Goodchild, D. J. Maguire and D. W. Rhind (New York: John Wiley & Sons), pp. 177–189.

Zomer, R. J., Trabucco, A., & Ustin, S. L. (2009). Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of Environmental Management, 90*(7), 2170-2177.

Chapter 4 References

Arnott D and Pervan G, 2005, "A critical analysis of decision support systems research" J.Inf.Technol. **20** 67-87

Barrett C, Eubank SG and Smith JP, 2005, "If Smallpox Strikes Portland..." Sci.Am. **292** 54-61

Buchecker M, Hunziker M and Kienast F, 2003, "Participatory landscape development: overcoming social barriers to public involvement RID B-8870-2008" Landscape Urban Plann. **64** 29-46

Clark NN, Gajendran P and Kern JM, 2003, "A predictive tool for emissions from heavy-duty diesel vehicles" Environ.Sci.Technol. **37** 7-15

Collins KA, Lawrence TJ, Stander EK, Jontos RJ, Kaushal SS, Newcomer TA, Grimm NB and Ekberg MLC, 2010, "Opportunities and challenges for managing nitrogen in urban stormwater: A review and synthesis" Ecol.Eng. **36** 1507-1519

Conley DJ, Paerl HW, Howarth RW, Boesch DF, Seitzinger SP, Havens KE, Lancelot C and Likens GE, 2009, "Controlling Eutrophication: Nitrogen and Phosphorus" Science **323** 1014-1015

Couclelis H, 2005, ""Where has the future gone?" Rethinking the role of integrated land-use models in spatial planning" Environ.Plann.A **37** 1353-1371

Dymond RL, Regmi B and Dietz R, 2004, "Interdisciplinary web-enabled spatial decision support system for watershed management" J.Water Resour.Plann.Manage. **130** 290

Gober P, Kirkwood CW, Balling RC, Ellis AW and Deitrick S, 2010, "Water Planning Under Climatic Uncertainty in Phoenix: Why We Need a New Paradigm" Annals of the Association of American Geographers

Goosen H, Janssen R and Vermaat JE, 2007, "Decision support for participatory wetland Decision making" Ecol.Eng. **30** 187-199

Hitch, C, 1955, "An Appreciation of Systems Analysis" Journal of the Operations Research Society of America **3** 466-481

Howard RA, 1968, "Foundations of Decision Analysis" Ieee Transactions on Systems Science and Cybernetics **SSC4** 211-&

Howard RA, 2007, Advances in Decision Analysis (Cambridge University Press, Cambridge)

Howarth RW, Anderson DM, Church TM, Greening H, Hopkinson CS, Huber WC, Marcus N, Naiman RJ, Segerson K, Sharpley A and Wiseman WJ, 2000, Clean Coastal Waters : Understanding and Reducing the Effects of Nutrient Pollution (National Academies Press, Washington, DC, USA)

Jamieson DG and Fedra K, 1996, "The 'WaterWare'decision-support system for river-basin planning. 1. Conceptual design" Journal of Hydrology **177** 163-175

Keefer DL, Kirkwood CW and Corner JL, 2007, "Perspective on decision analysis applications", in Advances in Decision AnalysisAnonymous (Cambridge University Press, Cambridge)

Keeney RL, 1994, "Using Values in Operations-Research" Oper.Res. **42** 793-813

Keeney RL, 1992, Value-Focused Thinking :A Path to Creative Decisionmaking (Harvard University Press, Cambridge, Mass.)

Keeney RL, 1996, "Value-focused thinking: Identifying decision opportunities and creating alternatives" Eur.J.Oper.Res. **92** 537-549

Kirkwood CW, 1997, Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets (Duxbury Press, Belmont, CA)

Klosterman RE, 1999, "The What if? Collaborative planning support system" Environ.Plann.B **26** 393-408

Lempert R, Groves D, Popper S and Bankes S, 2006, "A general analytic method for generating robust strategies and narrative scenarios" Management Science **52** 514-528

Leung Y, Lee Y, Lam KC, Lin K and Zeng FT, 2005, "An environmental decision-support system for the management of water pollution in a tidal river network" Int.J.Geogr.Inf.Sci. **19** 483-500

Logical Decisions, accessed at http://www.logicaldecisions.com/ **2013**

Marchau VAWJ, Walker WE and van Wee GP, 2010, "Dynamic adaptive transport policies for handling deep uncertainty" Technological Forecasting and Social Change **77** 940-950

Mellers B and Locke C, 2007, Advances in Decision Analysis (Cambridge University Press, Cambridge)

Nyerges T, Jankowski P, Tuthill D and Ramsey K, 2006, "Collaborative water resource decision support: Results of a field experiment" Ann.Assoc.Am.Geogr. **96** 699-725

Oakley SM, Gold AJ and Oczkowski AJ, 2010, "Nitrogen control through decentralized wastewater treatment: Process performance and alternative management strategies" Ecol.Eng. **36** 1520-1531

Pestotnik SL, Classen DC, Evans RS and Burke JP, 1996, "Implementing antibiotic practice guidelines through computer-assisted decision support: clinical and financial outcomes" Ann.Intern.Med. **124** 884-890

Ramsey K, 2009, "GIS, modeling, and politics: On the tensions of collaborative decision support" J.Environ.Manage. **90** 1972-1980

Rittel H and Webber M, 1973, "Dilemmas in a General Theory of Planning" Policy Sci. **4** 155-169

Robinson T, Harris R, Hopkins J and Williams B, 2002, "An example of decision support For trypanosomiasis control using a geographical information system in eastern Zambia." Int.J.Geogr.Inf.Sci. **16** 345-360

Ryther J and Dunstan W, 1971, "Nitrogen,phosphorus, and Eutrophication in Coastal MarineEnvironment" Science **171** 1008-&

Schipper LA, Robertson WD, Gold AJ, Jaynes DB and Cameron SC, 2010, "Denitrifying bioreactors-An approach for reducing nitrate loads to receiving waters" Ecol.Eng. **36** 1532-1543

Tversky A and Kahneman D, 1974, "Judgment Under Uncertainty - Heuristics and Biases" Science **185** 1124-1131

Tversky A and Kahneman D, 1981, "The Framing of Decisions and the Psychology of Choice" Science **211** 453-458

Waddell P, 2002, "Modeling Urban Development for Land Use, Transportation, and Environmental Planning" Journal öf the American Planning Association **68**

Walker WE, Marchau VAWJ and Swanson D, 2010, "Addressing deep uncertainty using adaptive policies: Introduction to section 2" Technological Forecasting and Social Change **77** 917-923

APPENDIX A

SUMSTREAMLOADSEASON.PY

This is a Python script to sum nitrogen transport through the network.

```
print "Start imports..."
import numpy as np # for math and for pysal
import pysal as ps # for importing a writing to the DBF
import gc # for cleanup
import time # for start and end messages
import sys, traceback #for error logging
import shutil


###column values: ##
HYDROID=0
NEXTID=0
SINKID=0
SINKTYPE=0
SINK_FAC=0
INLD=0
CURLD=0
TLBS=0
TLAS=0
TYPE=0
WIDTH = 0

## Other variables ##
i=0 #the column # when searching the header for required attributes
foundVars=0  #the number of required attributes (variables needed) identifed
j=1.0 #looping. how many total large loops have been executed
curSink=-1 #a global record to store the current sinkID. -1 if not in a sink.
cur_val=-1 # row of the record being evaluated

change = 0 #how much load not consumed by a sink
wetlandWidth = 0.0 #total distance accumulated while traversing reaches in a wetland
pourPoints = 0 # tracks number of pour points found
reachCount = 0 # for tracking number of segments per flowpath
num_array_rows = 0.0 # for calculating completion percent
percent_complete = 0.0 # for display

###Exceptions##
#class MyError(Exception):
  #def __init__(self, value):
  #  self.parameter = value
  #def __str__(self):
  #  return repr(self.parameter)
#  pass

##Functions ##
# Function to print messages to console
def addMessage(message): #@message: a string to print
```

```
    print message #prints to console.
    #insert print calls for other environments here. (ie. arcpy.AddMessage(message))

#Note: this function has been replaced by a dictionary lookup.
def findIndex(array, value): #@array: a column of Hydro IDs, @value: the ID of inquiry
    val_index_tupple = np.where(array==value)
    index = val_index_tupple[0] #(the row_ID is in the first (0th) position of the index
array)
    #print "downstream reach ID is at index" + str(index)
    if len(index)<>0: #check if the list is empty (length of the array is 0) (ie couldn't find a
next row with this ID)
        #print "returning " + str(index[0])
        return index[0]

    else:
        #print "Clean return. No HydroID found for value " + str(value)
        return -1

def processWetland(thisWidth): #determine the wetland removal rate based on the
given length @thisWidth: a float width value
    if thisWidth < 5:
        #print "Nremoval is 0"
        return 0
    elif thisWidth >= 5 and thisWidth < 15:
        #print "Nremoval is 40"
        return 40
    elif thisWidth >= 15 and thisWidth < 30:
        ##print "Nremoval is 60"
        return 60
    elif thisWidth >= 30:
        ##print "Nremoval is 80"
        return 80

def calculateNRemoved(thisRow):
    # calculate load that makes it through the sink. Set the TLAS and reset the CURLD
    if thisRow[SINK_FAC] > 0: # check if load should be removed
      change = thisRow[CURLD] * (1.0-(thisRow[SINK_FAC]/100)) #calculate how much
load gets through
      NEW_TLAS= change + thisRow[TLAS] #find the new TLAS by adding the amount
that makes it through the sink on this itteration
      thisRow[TLAS] = NEW_TLAS #set the new TLAS
      thisRow[CURLD] = change #set the curld as the amount of load that makes it
through

      return thisRow #return the updated row
    else:
      thisRow[TLAS] = thisRow[TLBS]
      return thisRow
```

```
##Main code ##

# Print start time
addMessage(time.strftime('%X %x') + " " +"Began")

#set source and destination files
#source =
'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\Module4\Inputs\Inp
utStreamCopy06'
#source =
'S:\Projects\Nitrogen\ModelTransfer\NSinkLaunch1\Data\Module4\Intermediate\Inp
utStreamCopy'
#source =
'C:\Users\joanna\Documents\RR\NianticProcessing\Landuseprocessing\Layers\Mod4
Input06'
#source
='C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\Module4\Inputs\St
reamSampleCopy'
#source =
'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\SDrive\Module4\Int
ermediate\InputStreamhalf'
source =
'C:\Users\Shimizu\Documents\Melindas\PhD\ModelPaper\SSURGO_sourcetweak\he
24lbres\sewered24lbresCopy'


# destination =
'C:\Users\joanna\Documents\RR\NianticProcessing\Landuseprocessing\Layers\Mod4
Output06d2'
#destination =
'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\Module4\Outputs\St
reamSampleSum2'
#'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\SDrive\Module4\O
utputs\FlowpathsOutputSample'
#destination =
'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\SDrive\Module4\Out
puts\DictionaryStreamSample'
#destination =
'C:\Users\joanna\Documents\RR\NianticProcessing\RRMethods\Module4\Outputs\Di
ctionarySample06'
#destination =
'S:\Projects\Nitrogen\ModelTransfer\NSinkLaunch1\Data\Module4\Outputs\Dictionar
ySample06'
destination =
'C:\Users\Shimizu\Documents\Melindas\PhD\ModelPaper\SSURGO_sourcetweak\he
24lbres\sewered24lbresOut'

if source == destination:
  sys.exit("Error:\nSource is the same as destination. Don't overwrite your input!!")
#jump to the Except, and Finnally
```

```
#read file
addMessage(time.strftime('%X %x') + " " +"Reading file " + source)
dbf_link = source + '.dbf' # open source file
db = ps.open(dbf_link)
my_array = np.squeeze(np.array(db)) # save content of file into a local array
num_array_rows = float(my_array.shape[0]) # find number of rows in the input file
if num_array_rows == 0:
  db.close()
  sys.exit("\n***\nError:\nNo records found in the input file.\n***")


addMessage(time.strftime('%X %x') + " " +"Creating output dbf file " + destination
+".dbf")
#create output file
dbo = ps.open(destination + '.dbf', 'w')
dbo.header = db.header
dbo.field_spec = db.field_spec

try:
  #find array index (column number) for each variable
  addMessage(time.strftime('%X %x') + " " +"Identifying attributes...")
  for item in dbo.header:
      #print "testing: " + item
      if item == 'HydroID':
         HYDROID = i
         foundVars= foundVars+1
         print "HYDROID is " + str(i)
      elif item =='NextDownID':
         NEXTID = i
         foundVars= foundVars+1
         print "NEXTID is " + str(i)
      elif item =='Id':
         SINKID = i
         foundVars= foundVars+1
         print "SINKID is " + str(i)
      elif item =='InputLd':
      #elif item =='SpLd2':
      #elif item =='SumLd2':
      #elif item =='FallLd2':
         INLD = i
         foundVars= foundVars+1
         print "INLD is " + str(i)
      elif item =='CurLd':
         CURLD = i
         foundVars= foundVars+1
         print "CURLD is " + str(i)
      elif item =='TLBS':
```

```
     TLBS = i
     foundVars= foundVars+1
     print "TLBS is " + str(i)
   elif item =='TLAS':
     TLAS = i
     foundVars= foundVars+1
     print "TLAS is " + str(i)
   elif item == 'ROW_ID':
     ROWID = i
     foundVars= foundVars+1
     print "ROWID is " + str(i)
   elif item == 'NRemoval':
     SINK_FAC = i
     foundVars= foundVars+1
     print "SINK_FAC is " + str(i)
   elif item == 'SnkTypeVal':
     SINKTYPE = i
     foundVars= foundVars+1
     print "SINKTYPE is " + str(i)
   elif item == "ReachLengt":
   #elif item == "Shape_Leng":
     WIDTH = i
     foundVars= foundVars+1
     print "WIDTH is " + str(i)

   i=i+1

 #check if all 11 required attributes were found
 if foundVars < 11:
  sys.exit("Error:\nMissing variables. Check documentation to see what is missing.")
#jump to the Except, and Finnally

 addMessage("")
 addMessage( time.strftime('%X %x') + " " + "Calculating Load...")

 #create dictionary
 addMessage("Creating dictionary...")
 HydroIDDict = {} #[HYDROID,INDEX]
 k = 0
 for vals in my_array:
  HydroIDDict[vals[HYDROID]] = k
  k=k+1


 #find which rows have input
 input_index_list = np.where(my_array[:,INLD]>0)
 #print input_index_list[0]
 numInputRows= len(input_index_list[0])
 print str(numInputRows) + " to itterate through"
```

```
 #loop through rows in the list
 for inputItem in input_index_list[0]:
  #print "input_item: " + str(inputItem)
  row=my_array[inputItem,:]

  #print row

  #percent every 2% complete
  percent_complete= j/numInputRows
  percentComplete2 = int(percent_complete*1000)#get the percentage to three decimal
ie 0.4563 => 456
  if (percentComplete2%20) == 0: #if the remainder of dividing by 20 is 0, print. ie 450
and 456 (aka 45.0% and 45.6%) will not print, but 460 (aka 46.0%) will
   print percentComplete2/10,"% complete        \r",
   if percent_complete > 0.999:
    addMessage("                    ")
    addMessage("100% complete")
  #print"" #debug - to print debug statements, use this blank line to stop overwriting the
percent complete.

  #check if this is a terminal segment
  if row[NEXTID] == -1:
      pourPoints += 1
      #print "Stop at pour point " + str(pourPoints)
      if reachCount == 0:
        #print "Single reach path found"
        row[TLAS]=row[TLBS]
        row[CURLD]=0
      j+=1 #incriment row count for percent completion calc
      continue

 #current value is the row's HYDROID
 #print "looking at HYDROID: "  + str(row[HYDROID])
 ##print row

  if row[CURLD] > 0:
    #print "enough current load. RowID is: " + str(row[ROWID])
    cur_val = row[ROWID] #note rows are 0-indexed on FID

    #while NEXTID is not the stop value
    while True: # while the NEXTID is not -1, there are still downstream reaches (-1
evaluated at the within the loop, and breaks out)
      reachCount = reachCount + 1
      #print "curent row: " + str(my_array[cur_valROWID]) + " load: " +
str(my_array[cur_valCURLD]) + " TLAS " + str(my_array[cur_valTLAS])

      #check for and deal with sink
      if not (my_array[cur_valSINKTYPE] == 0):
        #Sink found
        #print "sink of type " + str(my_array[cur_valSINKTYPE]) + " found"
```

```
            ##print "TLBS " + str( my_array[cur_valTLBS])

            #check type of sink:
            if my_array[cur_valSINKTYPE] == 1 or my_array[cur_valSINKTYPE] == 2:
#Pond or Stream, so use N removal by %
               ##print "In pond or stream"
               #clear wetland width and value because this ain't a wetland
               wetlandWidth = 0.0

               #check sink ID:
               if curSink <> my_array[cur_valSINKID]: #we are not in the same stream or
pond as before
                  #print "different sink " + str(my_array[cur_valSINKID])
                  curSink = my_array[cur_valSINKID] # store current sink variable

                  #print "row before change "
                  #print my_array[cur_val:]
                  #calculate N Removal
                  my_array[cur_val:]=calculateNRemoved(my_array[cur_val:])
                  #print "row after change "
                  #print my_array[cur_val:]

               else: # same sink so simply move the load without removal
                  ##print "same sink " + str(my_array[cur_valSINKID])
                  my_array[cur_valTLAS] = my_array[cur_valCURLD] +
my_array[cur_valTLAS]

            elif my_array[cur_valSINKTYPE] == 3: # Wetland. So use N removal by width
                #print "In wetland with width ID "  + str(my_array[cur_valSINKID]) + "
and width " +str(my_array[cur_valWIDTH]) + " which is stored in row " +
str(my_array[cur_valROWID])

                #check sink ID:
                if curSink <> my_array[cur_valSINKID]: #we are not in the same stream or
pond as before
                  #print "different sink " + str(my_array[cur_valSINKID])
                  curSink = my_array[cur_valSINKID] # store current sink variable

                wetlandWidth = wetlandWidth + my_array[cur_valWIDTH]
                ##print "width is now " + str(wetlandWidth)

                ##print "next index is " + str(next_index) + " and sink sype val is " +
str(my_array[next_index,SINKTYPE])
                #find the ROWID of the HYDROID that has the value of the NEXTID
                ###USE DICTIONARY! next_index ==
findIndex(my_array[:,HYDROID],my_array[cur_val NEXTID])
                next_index = HydroIDDict.get(my_array[cur_val NEXTID], -1)
```

```
                #if there is a not following reach or the reach is not a wetland, set removal
rate
                if my_array[cur_val NEXTID] == -1 or next_index == -1 or
my_array[next_index,SINKTYPE] <> 3:
                    ##print "no following wetlands"
                    my_array[cur_valSINK_FAC] = processWetland(wetlandWidth)
                    ##print "row before"
                    ##print my_array[cur_val:]

                    #calculate N Removal
                    my_array[cur_val:]=calculateNRemoved(my_array[cur_val:])
                    ##print "row after"
                    ##print my_array[cur_val:]

                else: #the wetland continues into the next reach
                     #set the Nremoval rate so far (but)
                     my_array[cur_valSINK_FAC] = processWetland(wetlandWidth)
                     # treat reach as if it had no sink, and put the TLBS into the TLAS to be
passed along
                     my_array[cur_valTLAS] = my_array[cur_valCURLD] +
my_array[cur_valTLAS]

            ##print "TLAS after sink evaluation:"
            ##print my_array[cur_valTLAS]

        else: #No Sink, so TLAS and TLBS are the same
            #print "No Sink"
            curSink = -1 #not in a sink, so clear the curent sink variable
            my_array[cur_valTLAS] = my_array[cur_valCURLD] +
my_array[cur_valTLAS]


        ### Dictionary next_index =
findIndex(my_array[:,HYDROID],my_array[cur_val NEXTID])
        next_index = HydroIDDict.get(my_array[cur_val NEXTID], -1)
        if my_array[cur_val NEXTID] == -1 or next_index ==-1: #if this is the (a) last
reach in the system return to the table
            curSink = -1
            my_array[cur_valCURLD] = 0
            reachCount = 0
            #print "breaking"
            break
        else:
            #update next row's total load before sink and current load, by giving them this
reaches CURLD
            #print "current next row:"
            #print my_array[next_index,:]
            my_array[next_index,TLBS] = my_array[next_index,TLBS] +
my_array[cur_val CURLD]
```

```
        my_array[next_index,CURLD] = my_array[next_index,CURLD] +
my_array[cur_val CURLD]
        #print "next row should now have additional curLd:"
        #print my_array[next_index,:]

      my_array[cur_valCURLD] = 0 #since the current load has been passed along, set
it to 0!

      #set new current row
      cur_val = next_index
      #print "looping to index " + str(next_index)

  #else: #debug - use this else with the next comment
  #    print str(row[ROWID])+" No current load, move along..."
  j+=1 #incriment row count for percent completion calc

 #copy duplicates over for destination file
 print time.strftime('%X %x') + " " +"writing the associated files"
 shutil.copy(source+'.sbn', destination+'.sbn')
 shutil.copy(source+'.shx', destination+'.shx')
 shutil.copy(source+'.shp', destination+'.shp')
 shutil.copy(source+'.sbn', destination+'.sbn')
 shutil.copy(source+'.sbx', destination+'.sbx')
 shutil.copy(source+'.sbn', destination+'.sbn')
 shutil.copy(source+'.prj', destination+'.prj')
 shutil.copy(source+'.shp.xml', destination+'.shp.xml')


except SystemExit as e:
 print "\n***"
 print e
 print "***"

except:  #pulled from the net. Not all lines may be needed
 #print "Unexpected error:", sys.exc_info()[0]

 exc_type, exc_value, exc_traceback = sys.exc_info()
 print "*** print_tb:"
 traceback.print_tb(exc_traceback, limit=1, file=sys.stdout)
 print "*** print_exception:"
 traceback.print_exception(exc_type, exc_value, exc_traceback,
              limit=2, file=sys.stdout)
 print "*** print_exc:"
 traceback.print_exc()
 print "*** format_exc, first and last line:"
 formatted_lines = traceback.format_exc().splitlines()
 print formatted_lines[0]
 print formatted_lines[-1]
 print "*** format_exception:"
 print repr(traceback.format_exception(exc_type, exc_value,
```

```
                                    exc_traceback))
  print "*** extract_tb:"
  print repr(traceback.extract_tb(exc_traceback))
  print "*** format_tb:"
  print repr(traceback.format_tb(exc_traceback))
  print "*** tb_lineno:", exc_traceback.tb_lineno


finally:

  #write final rows to table:
  addMessage(time.strftime('%X %x') + " Writing output dbf... ")
  for row in my_array:
      dbo.write(row)

  # Print End time
  addMessage(time.strftime('%X %x')+ " Completed ")
  addMessage("")

  ##Save/close files
  dbo.close()
  db.close()
  ##clean up
  gc.collect()
```

APPENDIX B

[SURVEY INSTRUMENT AND RESULTS]

Survey: Water Planning Survey - Pilot Study
Value Count Percent %
18-24 8 26.7%
25-29 8 26.7%
30-34 9 30.0%
35-39 4 13.3%
40-44 1 3.3%
45-59 0 0.0%
60+ 0 0.0%
Statistics
Total Responses 30
Sum 794.0
Avg. 26.5
StdDev 6.3
Max 40.0
Summary Report - Sep 22, 2013

1. What is your age?

1. What is your age?
18-24 26.7%
25-29 26.7%
30-34 30%
35-39 13.3%
40-44 3.3%

2. What is your highest completed level of education?
4-year degree 33.3%
Master's degree 53.3%
Doctoral degree 13.3%
Value Count Percent %
High School diploma 0 0.0%
Some college 0 0.0%
4-year degree 10 33.3%
Master's degree 16 53.3%
Doctoral degree 4 13.3%
Other 0 0.0%
Other 0 0.0%
Statistics
Total Responses 30
Sum 40.0
Avg. 4.0
Max 4.0
Value Count Percent %
<1 year 4 13.3%
1-2 years 7 23.3%
2-5 years 10 33.3%
5+ years 9 30.0%
No experience 0 0.0%
Statistics
Total Responses 30
Sum 72.0
Avg. 2.8

StdDev 1.7

Max 5.0

2. What is your highest completed level of education?

3. Please indicate your experience with GIS

3. Please indicate your experience with GIS

<1 year 13.3%

1-2 years 23.3%

2-5 years 33.3%

5+ years 30%

Value Count Percent %

< 1 year 10 33.3%

1-2 years 5 16.7%

2-5 years 3 10.0%

5+ years 1 3.3%

No experience 11 36.7%

Statistics

Total Responses 30

Sum 16.0

Avg. 1.8

StdDev 1.2

Max 5.0

Value Count Percent %

< 1 year 10 33.3%

1-2 years 5 16.7%

2-5 years 4 13.3%

5+ years 3 10.0%

No experience 8 26.7%

Statistics

Total Responses 30

Sum 28.0

Avg. 2.3

StdDev 1.6

Max 5.0

4. Please indicate your experience with city planning

5. Please indicate your experience with environmental planning

4. Please indicate your experience with city planning

< 1 year 33.3%

1-2 years 16.7%

2-5 years 10%

5+ years 3.3%

No experience 36.7%

5. Please indicate your experience with environmental planning

< 1 year 33.3%

1-2 years 16.7%

2-5 years 13.3%

5+ years 10%

No experience 26.7%

Value Count Percent %

Student 22 73.3%

Researcher 13 43.3%

Instructor 7 23.3%
Professional 6 20.0%
Other 2 6.7%
Statistics
Total Responses 30
Value Count Percent % Statistics

6. Are you a...(check as many as apply)

8. Please select a tool from the three options below. You may choose any of the three tools. Once you select the tool, the next screen will show you the results from using that tool. You may then select another tool, in any order, if you wish. You may continue selecting tools until you are ready to make your decision. The tool options are presented in random order.

6. Are you a...(check as many as apply)

73.3%
43.3%
23.3%
20%
6.7%
Student Researcher Instructor Professional Other
0
100
25
50
75

8. Please select a tool from the three options below. You may choose any of the three tools. Once you select the tool, the next screen will show you the results from using that tool. You may then select another tool, in any order, if you wish. You may continue selecting tools until you are ready to make your decision. The tool options are presented in random order.

Broad Outcome Tool. This tool allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map. 26.7%

Overview Tool. This tool shows you basic information about the current state of the planning area. Information is presented in both maps and tables. 53.3%

Specific Outcome Tool. This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table and a map for each location. 20%

Broad Outcome Tool.

This tool allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map.
8 26.7%
Overview Tool.
This tool shows you basic information about the current state of the planning area. Information is presented in both maps and tables.
16 53.3%
Specific Outcome Tool.
This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table and a map for each location.
6 20.0%
Make Decision 0 0.0%
Total Responses 30
Value Count Percent %
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map.
12 52.2%
Specific Outcome Tool.
This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table and a map for each location.
6 26.1%
Make Decision 5 21.7%
Statistics
Total Responses 23

9. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. You will be able to return to this tool's output later if you wish. (options presented in random order)

9. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. You will be able to return to this tool's output later if you wish. (options presented in random order)

Broad Outcome Tool. This tool allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map. 52.2%

Specific Outcome Tool. This tool allows you to
implement the incentive at specific locations, and
see what the outcome of that specific decision is.
The results will be presented as a table and a map
for each location. 26.1%
Make D ecision 21.7%
Value Count Percent %
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome (eg maximizing
economic potential while minimizing estuary impacts), and see what
possible decision alternatives (eg where to put the incentive) will result in
that outcome. The results will be presented as a chart and a map.
7 30.4%
Overview Tool.
This tool allows shows you basic information about the current state of
the planning area. Information is presented in both maps and a table.
2 8.7%
Make Decision 14 60.9%
Statistics
Total Responses 23

10. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order, even if you already selected it before. (options presented in random order)

11. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any

10. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order, even if you already selected it before. (options presented in random order)

Broad Outcome Tool. This tool allows you to
enter a desired decision outcome (eg maximizing
economic potential while minimizing estuary
impacts), and see what possible decision
alternatives (eg where to put the incentive) will
result in that outcome. The results will be
presented as a chart and a map. 30.4%
Overview Tool. This tool allows shows you basic
information about the current state of the
planning area. Information is presented in both
maps and a table. 8.7%
Make D ecision 60.9%

11. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Overview Tool. This tool allows shows you basic information about the current state of the planning area. Information is presented in both maps and a table. 16.7%

Specific Outcome Tool. This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table with a map for each location. 37.5%

Make D ecision 45.8%

Value	Count	Percent %
Overview Tool. This tool allows shows you basic information about the current state of the planning area. Information is presented in both maps and a table.	4	16.7%
Specific Outcome Tool. This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table with a map for each location.	9	37.5%
Make Decision	11	45.8%

Statistics

Total Responses 24

Value	Count	Percent %
Zone A	21	70.0%
Zone B	7	23.3%
Zone C	2	6.7%
Not enough information to decide	0	0.0%
Impose a fee in areas of Medium Estuary Impact	0	0.0%
Impose a fee in areas of Low Estuary Impact	0	0.0%
Impose a fee in both areas of High and Medium Estuary Impact	0	0.0%

Statistics

Total Responses 30

of the other tools in any order. (options presented in random order)

12. Please indicate your decision on the proposed incentive:

13. Please rank the tools in order of most helpful for making the decision (1 is the most helpful and 3 is the least helpful):

Item	Total Score1	Overall Rank
Specific Outcome Tool. This tool allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table with a map for each location.	68	1

Broad Outcome Tool.
This tool allows you to enter a desired decision outcome (eg maximizing economic potential while
minimizing estuary impacts), and see what possible decision alternatives (eg where to put the
incentive) will result in that outcome. The results will be presented as a chart and a map.
49 2
12. Please indicate your decision on the proposed incentive:
Zone A 70%
Zone B 23.3%
Zone C 6.7%
Value Count Percent %
Tool 1: This tool allows you to explore the current developments and the
current state of the groundwater.
0 0.0%
Tool 2: This tool allows you to impose the fee at specific locations, and
see what the outcome of that specific decision is
0 0.0%
Tool 3: This tool allows you to enter a desired decision outcome, and
see what possible decision alternatives will result in that outcome.
0 0.0%
Statistics
Total Responses 0
Value Count Percent %
Very important 5 16.7%
Somewhat important 11 36.7%
Somewhat unimportant 2 6.7%
Not important 7 23.3%
Did not use 5 16.7%
Statistics
Total Responses 30
Overview Tool.
This tool allows shows you basic information about the current state of the planning area.
Information is presented in both maps and a table.
48 3

Total Respondents: 28

1 Score is a weighted calculation. Items ranked first are valued higher than the following ranks, the score is the sum of all weighted rank counts.

Please rank the tools in order of importance:

14. Please rate the importance of the Overview Tool. This was the tool that shows you basic information about the current state of the planning area. Information is presented in both maps and a table. How important was the Overview Tool for your decision?

14. Please rate the importance of the Overview Tool. This was the tool that shows you basic information about the current state of the planning area. Information is presented in both maps and a table. How important was the Overview Tool for your decision?

Very important 16.7%
Somewhat important 36.7%
Somewhat unimportant 6.7%
Not important 23.3%
Did not use 16.7%
Value Count Percent %
Very important 18 60.0%
Somewhat important 6 20.0%
Somewhat unimportant 3 10.0%
Not important 0 0.0%
Did not use 3 10.0%
Statistics
Total Responses 30

15. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table with a map for each location. How important was this tool in your decision making?

16. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map. How important was this tool in

15. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to implement the incentive at specific locations, and see what the outcome of that specific decision is. The results will be presented as a table with a map for each location. How important was this tool in your decision making?

Somewhat important 20% Very important 60%
Somewhat unimportant 10%
Did not use 10%

16. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to enter a desired decision outcome (eg maximizing economic potential while minimizing estuary impacts), and see what possible decision alternatives (eg where to put the incentive) will result in that outcome. The results will be presented as a chart and a map. How important was this tool in your decision making?

Very important 50%
Somewhat important 20%
Somewhat unimportant 16.7%
Not important 6.7%
Did not use 6.7%

Value Count Percent %
Very important 15 50.0%
Somewhat important 6 20.0%
Somewhat unimportant 5 16.7%
Not important 2 6.7%
Did not use 2 6.7%
Statistics
Total Responses 30
Value Count Percent %
Very comfortable 19 63.3%
Somewhat comfortable 10 33.3%
Somewhat uncomfortable 0 0.0%
Very uncomfortable 1 3.3%
Statistics
Total Responses 30

your decision making?

17. How comfortable were you in making the decision?

17. How comfortable were you in making the decision?
Very comfortable 63.3%
Somewhat comfortable 33.3%
Very uncomfortable 3.3%
Value Count Percent %
Yes 28 93.3%
No 2 6.7%
Statistics
Total Responses 30
Value Count Percent %
1 7 23.3%
2 10 33.3%
3 13 43.3%
0 0 0.0%
Statistics
Total Responses 30
Sum 66.0
Avg. 2.2
StdDev 0.8
Max 3.0

18. Did you use at least one tool to make your decision?

19. How many tools did you use to make your decision?

18. Did you use at least one tool to make your decision?
Yes 93.3%
No 6.7%
19. How many tools did you use to make your decision?
1 23.3%
2 33.3%
3 43.3%
Value Count Percent %
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. The results are

presented in a table and a map.
11 36.7%
Overview Tool.
This tool allows you to view the current developments and the current and future state of the estuaries. The results are presented as a table and a map.
15 50.0%
Specific Outcome Tool.
This tool allows you to put the development at specific locations, and see what the outcome of that specific decision is. The results are presented in a table and a map.
2 6.7%
Make Decision 2 6.7%
Statistics
Total Responses 30

21. You may select any of the tool options below, in any order. After selecting one tool, you may select another, in any order. You may continue selecting tools until you are ready to make your decision. Options presented in random order.

21. You may select any of the tool options below, in any order. After selecting one tool, you may selectanother, in any order. You may continue selecting tools until you are ready to make your decision. Options presented in random order.

Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. The results are presented in a table and a map. 36.7%
Overview Tool. This tool allows you to view the current developments and the current and future state of the estuaries. The results are presented as a table and a map. 50%
Specific Outcome Tool. This tool allows you to put the development at specific locations, and see what the outcome of that specific decision is. The results are presented in a table and a map. 6.7%
Make Decision 6.7%

22. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Specific Outcome Tool. This tool allows you to put the development at specific locations, and see what the outcome of that specific decision is. 39.1%
Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that

outcome. 30.4%
Make Decision 30.4%
Value Count Percent %
Specific Outcome Tool.
This tool allows you to put the development at specific locations, and see what the outcome of that specific decision is.
9 39.1%
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
7 30.4%
Make Decision 7 30.4%
Statistics
Total Responses 23
Value Count Percent %
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
2 10.0%
Overview Tool.
This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table.
4 20.0%
Make Decision 14 70.0%
Statistics
Total Responses 20

22. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

23. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

23. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. 10%
Overview Tool. This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table. 20%
Make Decision 70%

Value Count Percent %
Overview Tool.
This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table.
3 14.3%
Specific Outcome Tool.
This tool allows you to put the development at specific locations, and see what is the outcome of that specific decision.
11 52.4%
Make Decision 7 33.3%
Statistics
Total Responses 21
Value Count Percent %
Blanket approval - they can build anywhere in the city 0 0.0%
Approve one specific location that you know to not be sensitive to
21 70.0%
Statistics
Total Responses 30

24. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

25. Select your decision from the options below:

24. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Overview Tool. This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table. 14.3%
Specific Outcome Tool. This tool allows you to put the development at specific locations, and see what is the outcome of that specific decision. 52.4%
Make Decision 33.3%

25. Select your decision from the options below:

Approve one specific location that you know to not be sensitive to nearby estuaries 70%
Approve all locations in the city that are not sensitive to nearby estuaries 30%

nearby estuaries
21 70.0%
Approve all locations in the city that are not sensitive to nearby estuaries 9 30.0%
Do not approve development 0 0.0%
Not enough information to decide 0 0.0%
Value Count Percent %
Tool 1: This tool allows you to explore the current developments and the current state of the groundwater.

0 0.0%
Tool 2: This tool allows you to impose the fee at specific locations, and see what the outcome of that specific decision is
0 0.0%
Tool 3: This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
0 0.0%
Statistics
Total Responses 0

26. Please rank the tools in order of how helpful they were to making your decision (1 is most helpful; 3 is least helpful):

Item Total Score[1] Overall Rank
Specific Outcome Tool.
This tool allows you to see the outcome of 3 specific development locations
66 1
Overview Tool.
This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table.
52 2
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives (or which locations) will result in that outcome.
47 3

Total Respondents: 28

[1] Score is a weighted calculation. Items ranked first are valued higher than the following ranks, the score is the sum of all weighted rank counts.

Please rank the tools in order of importance:

27. Please rate the importance of the Overview Tool. This was the tool that shows you basic information about the current state of the planning area. Information is presented in both maps and

27. Please rate the importance of the Overview Tool. This was the tool that shows you basic information about the current state of the planning area. Information is presented in both maps and a table. How important was this tool in your decision making?

Very important 23.3%
Somewhat important 33.3%
Somewhat unimportant 26.7%
Not important 10%
D id not use 6.7%

Value Count Percent %
Very important 7 23.3%
Somewhat important 10 33.3%
Somewhat unimportant 8 26.7%
Not important 3 10.0%
Did not use 2 6.7%
Statistics

Total Responses 30
Value Count Percent %
Very important 16 53.3%
Somewhat important 6 20.0%
Somewhat unimportant 3 10.0%
Not important 3 10.0%
Did not use 2 6.7%
Statistics
Total Responses 30

a table. How important was this tool in your decision making?

28. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to
see the outcome of 3 specific development locations. How important was this tool in your decision
making?

28. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to
see the outcome of 3 specific development locations. How important was this tool in your decision
making?

Very important 53.3%
Somewhat important 20%
Somewhat unimportant 10%
Not important 10%
Did not use 6.7%
Value Count Percent %
Very important 10 33.3%
Somewhat important 13 43.3%
Somewhat unimportant 2 6.7%
Not important 1 3.3%
Did not use 4 13.3%
Statistics
Total Responses 30
Value Count Percent %
Very comfortable 6 20.7%
Somewhat comfortable 21 72.4%
Somewhat uncomfortable 1 3.5%
Statistics
Total Responses 29

29. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to
enter a desired decision outcome, and see what possible decision alternatives will result in that
outcome. How important was this tool in your decision making?

30. How comfortable were you in making the decision?

29. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to
enter a desired decision outcome, and see what possible decision alternatives will result in that
outcome. How important was this tool in your decision making?

Very important 33.3%
Somewhat important 43.3%
Somewhat unimportant 6.7%
Not important 3.3%
Did not use 13.3%
30. How comfortable were you in making the decision?
Very comfortable 20.7%
Somewhat comfortable 72.4%
Somewhat uncomfortable 3.5%
Very uncomfortable 3.5%
Very uncomfortable 1 3.5%
Value Count Percent %
Yes 28 93.3%
No 2 6.7%
Statistics
Total Responses 30
Value Count Percent %
1 8 26.7%
2 10 33.3%
3 11 36.7%
0 1 3.3%
Statistics
Total Responses 30
Sum 61.0
Avg. 2.0
StdDev 0.9
Max 3.0
31. Did you use at least one tool to make your decision?
32. How many tools did you use to make your decision?
31. Did you use at least one tool to make your decision?
Yes 93.3%
No 6.7%
32. How many tools did you use to make your decision?
1 26.7%
2 33.3%
3 36.7%
0 3.3%
Value Count Percent %
Overview Tool.
This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table.
10 33.3%
Specific Outcome Tool.
This tool allows you to allocate different zoning at a specific location, and see what the outcome of that specific decision is.
12 40.0%
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
8 26.7%

Statistics

Total Responses 30

34. You may select any of the tool options below, in any order. After selecting one tool, you may select another, in any order. You may continue selecting tools until you are ready to make your decision . (options presented in random order)

35. If you are ready to make your decision, select "Make Decision." If you would like to see

34. You may select any of the tool options below, in any order. After selecting one tool, you may selectanother, in any order. You may continue selecting tools until you are ready to make your decision . (options presented in random order)

Overview Tool. This tool shows you basic information about the current state of the planning area. Information is presented in both maps and a table. 33.3%

Specific Outcome Tool. This tool allows you to allocate different zoning at a specific location, and see what the outcome of that specific decision is. 40%

Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. 26.7%

35. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Specific Outcome Tool. This tool allows you to allocate different zoning at a specific location, and see what the outcome of that specific decision is. 30.8%

Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. 23.1%

Make Decision 46.2%

Value	Count	Percent %
Specific Outcome Tool. This tool allows you to allocate different zoning at a specific location, and see what the outcome of that specific decision is.	8	30.8%
Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.	6	23.1%
Make Decision	12	46.2%

Statistics
Total Responses 26
Value Count Percent %
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
10 40.0%
Overview Tool.
This tool allows you to view the current developments and the current and future impacts on the estuaries.
8 32.0%
Make Decision 7 28.0%
Statistics
Total Responses 25

information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

36. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

36. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Broad Outcome Tool. This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. 40%
Overview Tool. This tool allows you to view the current developments and the current and future impacts on the estuaries. 32%
Make D ecision 28%
Value Count Percent %
Overview Tool.
This tool allows you to view the current developments and the current and future state of the estuaries.
5 20.8%
Specific Outcome Tool.
This tool allows you to allocate different zoning at a specific location, and see what is the outcome of that specific decision.
8 33.3%
Make Decision 11 45.8%
Statistics
Total Responses 24
Value Count Percent %
Approve the area for luxury housing 6 20.0%
Approve the area for a nature preserve 24 80.0%

Statistics
Total Responses 30

37. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

38. Please select your decision from the following options:

37. If you are ready to make your decision, select "Make Decision." If you would like to see information from other tools, please select the tool you would like to see next. You may select any of the other tools in any order. (options presented in random order)

Overview Tool. This tool allows you to view the current developments and the current and future state of the estuaries.	20.8%
Specific Outcome Tool. This tool allows you to allocate different zoning at a specific location, and see what is the outcome of that specific decision.	33.3%
Make D ecision	45.8%

38. Please select your decision from the following options:

Approve the area for luxury housing	20%
Approve the area for a nature preserve	80%

Value	Count	Percent %
Tool 1: This tool allows you to explore the current developments and the current state of the groundwater.	0	0.0%
Tool 2: This tool allows you to impose the fee at specific locations, and see what the outcome of that specific decision is	0	0.0%
Tool 3: This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.	0	0.0%

Statistics
Total Responses 0

Value	Count	Percent %
Very important	8	26.7%
Somewhat important	13	43.3%

Statistics
Total Responses 30

39. Please rank the tools in order of most helpful for making your decision (1 is most helpful; 3 is least helpful):

Item	Total Score1	Overall Rank
Specific Outcome Tool. This tool allows you to input the different zoning options and see the outcome of each.	67	1
Overview Tool.		

This tool shows you basic information about the current state of the planning area. Information is
presented in both maps and a table.
58 2
Broad Outcome Tool.
This tool allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome.
52 3

Total Respondents: 30

1 Score is a weighted calculation. Items ranked first are valued higher than the following ranks, the score is the sum of all weighted rank counts.

Please rank the tools in order of importance:

40. Please rate the importance of the Overview Tool. This was the tool that shows you basic
information about the current state of the planning area. Information is presented in both maps and
a table. How important was the Overview Tool to your decision making?

40. Please rate the importance of the Overview Tool. This was the tool that shows you basic
information about the current state of the planning area. Information is presented in both maps and
a table. How important was the Overview Tool to your decision making?

Very important 26.7%
Somewhat important 43.3%
Somewhat unimportant 6.7%
Not important 13.3%
D id not use 10%
Somewhat unimportant 2 6.7%
Not important 4 13.3%
Did not use 3 10.0%
Value Count Percent %
Very important 12 40.0%
Somewhat important 8 26.7%
Somewhat unimportant 7 23.3%
Not important 1 3.3%
Did not use 2 6.7%
Statistics
Total Responses 30

41. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to
input the different zoning options and see the outcome of each. How important was the Specific
Outcome Tool to your decision making?

41. Please rate the importance of the Specific Outcome Tool. This was the tool that allows you to
input the different zoning options and see the outcome of each. How important was the Specific
Outcome Tool to your decision making?

Very important 40%
Somewhat important 26.7%

Somewhat unimportant 23.3%
Not important 3.3%
Did not use 6.7%

42. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. How important was the Broad Outcome Tool to your decision making?

Very important 30%
Somewhat important 30%
Somewhat unimportant 16.7%
Not important 10%
Did not use 13.3%

Value	Count	Percent %
Very important	9	30.0%
Somewhat important	9	30.0%
Somewhat unimportant	5	16.7%
Not important	3	10.0%
Did not use	4	13.3%

Statistics
Total Responses 30

Value	Count	Percent %
Very comfortable	5	16.7%
Somewhat comfortable	10	33.3%
Somewhat uncomfortable	11	36.7%
Very uncomfortable	4	13.3%

Statistics
Total Responses 30

42. Please rate the importance of the Broad Outcome Tool. This was the tool that allows you to enter a desired decision outcome, and see what possible decision alternatives will result in that outcome. How important was the Broad Outcome Tool to your decision making?

43. How comfortable were you in making the decision?

43. How comfortable were you in making the decision?

Very comfortable 16.7%
Somewhat comfortable 33.3%
Somewhat uncomfortable 36.7%
Very uncomfortable 13.3%

Value	Count	Percent %
Yes	29	96.7%
No	1	3.3%

Statistics
Total Responses 30

Value	Count	Percent %
1	5	16.7%
2	8	26.7%
3	16	53.3%
0	1	3.3%

Statistics
Total Responses 30
Sum 69.0
Avg. 2.3
StdDev 0.9
Max 3.0

44. Did you use at least one tool to make your decision?

45. How many tools did you use?

URL Variable: snc

44. Did you use at least one tool to make your decision?

Yes 96.7%
No 3.3%

45. How many tools did you use?

1 16.7%
3 53.3% 2 26.7%
0 3.3%

Count	Response
1	1379031003_523257dbc73fd