# ArcNLET

## Technical Manual

Manuscript Completed: June 2011

Prepared by
J. Fernando Rios[1], Ming Ye[1], Liying Wang[1], and Paul Lee[2]

[1] Florida State University
   Tallahassee, FL

[2] Florida Department of Environmental Protection
   Tallahassee, FL

Rick Hicks, FDEP Contract Manager

**Revision Sheet**

| Release No. | Date | Revision Description |
|---|---|---|
| 1.0.0 | June 21, 2011 | Initial release. |
| 1.1.0 | June 28, 2011 | Revision to NLE module. Updated comparison to MT3DMS. |
| 1.3.0 | July 22, 2011 | Update to v1.1 of the software. |

# Contents

# List of Tables

# List of Figures

# Abstract

Estimating groundwater nitrate fate and transport is an important task in water resources and environmental management because excess nitrate loads may have negative impacts on human and environmental health. This work discusses the development of a simplified nitrate transport model and its implementation as a geographic information system (GIS)-based screening tool, whose purpose is to estimate nitrate loads to surface water bodies from onsite wastewater-treatment systems (OWTS). Key features of this project are the reduced data demands due to the use of a simplified model, as well as ease of use compared to traditional groundwater flow and transport models, achieved by embedding the model within a GIS. The simplified conceptual model consists of a simplified groundwater flow model in the surficial aquifer, and a simplified transport model that makes use of an analytical solution to the advection-dispersion equation, used for determining nitrate fate and transport. Denitrification is modeled using first order decay in the analytical solution with the decay constant obtained from literature and/or site-specific data.

The groundwater flow model uses readily available topographic data to approximate the hydraulic gradient, which is then used to calculate seepage velocity magnitude and direction. The flow model is evaluated by comparing the results to a previous numerical modeling study of the U.S. Naval Air Station, Jacksonville (NAS) performed by the USGS. The results show that for areas in the vicinity of the NAS, the model is capable of predicting groundwater travel times from a source to a surface water body to within ±20 years of the USGS model, 75% of the time.

The transport model uses an analytical solution based on the one by Domenico and Robbins (1985), the results of which are then further processed so that they may be applied to more general, real-world scenarios. The solution, as well as the processing steps are tested using artificially constructed scenarios, each meant to evaluate a certain aspect of the solution. For comparison purposes, each scenario is solved using a well known numerical contaminant transport model. The results show that the analytical solution provides a reasonable approximation to the numerical result. However, it generally underestimates the concentration distribution to varying degrees depending on choice of parameters, especially along the plume centerline. These results are in agreement with previous studies (Srinivasan et al., 2007; West et al., 2007). The adaptation of the analytical solution to more realistic scenarios results in an adequate approximation to the numerically calculated plume, except in areas near the advection front, where the model produces a plume whose shape differs noticeably from the numerical solution.

Load calculations are carried out using a mass balance approach where the system is considered to be in the steady state. The steady-state condition allows for a load estimate by subtracting the mass removal rate due to denitrification from the input mass rate. The input mass rate is calculated by taking into account advection and dispersion while the mass removal rate due to denitrification is calculated from the definition of a first order reaction. Comparison with the synthetic scenarios

of the transport model shows that for the test cases, when denitrification rates are low, the model agrees well with the load calculation from the numerical model. With higher denitrification rates, the model generally underestimates the mass removal rate due to denitrification, resulting in an overestimation of the output load. In the best case, the underestimation is 8% and in the worst case, it is 27% (all comparisons made against MT3DMS simulations).

# 1

## Introduction

Estimating nitrate fate and transport in groundwater is an important task in water resources and environmental management, because excess nitrate load on ground water and surface waters may cause negative impacts on human and environmental health. In humans, nitrate-nitrogen becomes harmful due to its conversion to the more toxic nitrite-nitrogen (U.S. Environmental Protection Agency, 1993). Although low levels of nitrates, on the order of micrograms per liter (ppb), are used to treat cardiovascular diseases such as coronary heart disease (Bode-Böger and Kojda, 2005), nitrate is a potential carcinogenic agent and elevated nitrate intake have been identified as a possible cause of gastric cancer (U.S. Environmental Protection Agency, 1993). High levels of nitrates can also cause fatal poisoning in infants due to a condition known as methemoglobinemia (blue baby syndrome) where cells are starved of oxygen (U.S. Environmental Protection Agency, 1993). Elevated levels also harm livestock, fish and aquatic ecosystems. In aquatic ecosystems, high levels of nitrate can cause eutrophication of the water body, leading to algae blooms and excessive plant growth, the decay of which can cause anoxic conditions (Art, 1993). The Gulf of Mexico is highly susceptible to nutrient contamination. The discharge of high levels of nutrients from the Mississippi River, including nitrates, are thought to be responsible for a large region of reduced oxygen conditions (hypoxia) in the northern part of the Gulf of Mexico, along the Texas-Louisiana coast (Mueller and Spahr, 2006).

Because of the damage nitrates can inflict on health and natural resources, it is important to be able to effectively manage and regulate the amount of nitrates entering the environment. The EPA has determined that the nitrate concentration in drinking water should be 10 mg/l or less. In order to achieve this goal, it is first necessary to have the ability to identify actual or potential sources of nitrate contamination. There are several sources of nitrogen contamination such as agriculture, wastewater treatment plants, industrial sources and natural sources, with the contribution of each source varying depending on the region of study. The contribution of nitrogen due to the treatment of wastewater by onsite wastewater treatment systems (OWTS) comprises a significant portion of the nitrogen load to groundwater (U.S. Environmental Protection Agency, 1993). This contribution from OWTS may be especially important in areas where OWTS are located in close proximity to a
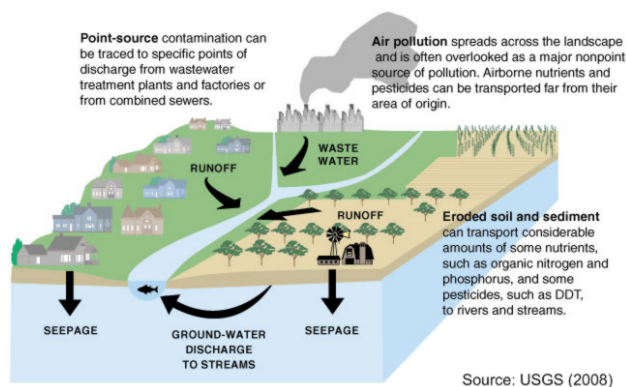
**Figure 1.1:** Sources of groundwater contamination.

surface water body. Because effluent from OWTS generally discharges into the soil, areas with a shallow water table are also vulnerable to nitrate contamination. This may be problematic if there is a dependence on domestic shallow wells for drinking water (Hitt and Nolan, 2005).

Approximately one third of Florida's population uses OWTS for treating wastewater (Brown et al., 2001; Ursin and Roeder, 2008). As a result, the negative impacts of nitrates in Florida are of wide concern. This prompted the Florida Department of Environmental Protection (FDEP) to initiate a project to develop an easy to use screening level model to estimate long-term nitrate loads from normally functioning (i.e. not failed) OWTS to surface water bodies, while at the same time taking into account nitrate attenuation mechanisms, such as dispersion and denitrification. The work presented here discusses the development of a simplified contaminant transport model and its implementation within a geographical information system (GIS).

The FDEP has identified several neighborhoods in Jacksonville, Florida that are of interest in regards to OWTS nitrate loading. For model validation purposes, a specific area, the Lakeshore neighborhood and surrounding area, was selected. The test cases presented in this discussion are either Lakeshore itself, or areas in close proximity.

## 1.1   Background

In this section, background information related to the simplified nitrate model is presented, starting with a brief discussion on the nitrogen cycle followed by a description of the functioning of OWTS and their relationship to nitrate contamination. Subsequently, a brief discussion on groundwater flow and contaminant transport modeling is presented, after which previous efforts related to the modeling of the transport and fate of contaminants (including nitrates) from OWTS are discussed, with an emphasis on simplified models.
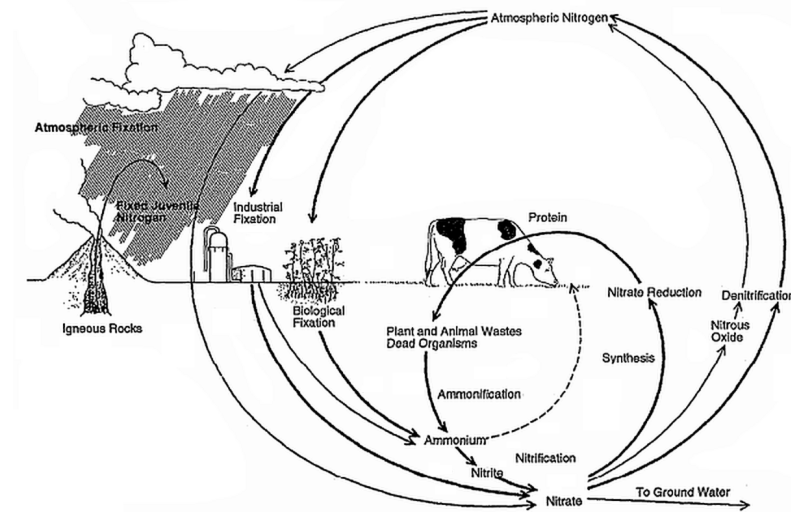
**Figure 1.2:** The nitrogen cycle. Source: U.S. Environmental Protection Agency (1993)

### 1.1.1   The Nitrogen Cycle

There are many sources of groundwater contamination. Fig. 1.1 shows several of the primary sources which include agricultural and industrial sources. When it comes to nitrate contamination, the largest contribution comes from agricultural sources with nitrates originating from the application of fertilizers as well as animal waste, however nitrates from OWTS may be of concern in localized areas (Canter, 1996).

**Table 1.1:** Forms of nitrogen of interest in groundwater

| Form of Nitrogen | Formula |
| --- | --- |
| Ammonia | $NH_3$ |
| Ammonium Ion | $NH_4^+$ |
| Nitrogen Gas | $N_2$ |
| Nitrite Ion | $NO_2^-$ |
| Nitrate Ion | $NO_3^+$ |

Nitrogen is present in the environment in several forms. In the context of soil and groundwater contamination, the forms given in Table 1.1 are of interest (U.S. Environmental Protection Agency, 1993). The transformation between these forms is given by the nitrogen cycle (Fig. 1.2). The segment starting with the process of ammonification is of relevance when dealing with nitrates from septic tanks. The process of ammonification transforms organic nitrogen to ammonium and ammonia via biological processes and generally occurs during the decomposition of organic matter such as plant material or fecal matter (Canter, 1996). Nitrification refers to the conversion of
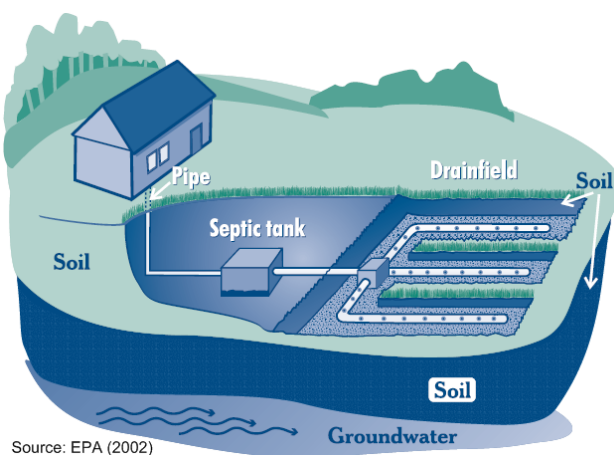
**Figure 1.3:** Schematic of an OWTS

ammonium to nitrite and then nitrate, however the conversion from nitrite to nitrate is considered to occur rapidly and therefore it is generally assumed that the transformation from ammonium to nitrate occurs in a single step (Canter, 1996; McCray et al., 2005).  An important removal process (and one that is considered in the model presented) is denitrification. Denitrification is a biological process that converts nitrate to nitrogen gas. The bacteria responsible for denitrification will, in the absence of oxygen, use nitrate to oxidize dissolved organic carbon (introduced to the system from organic matter in the soil), the end result of the process being the removal of nitrate and creation of nitrogen gas (U.S. Environmental Protection Agency, 1993). It is important to note that conditions must be anoxic in order for denitrification to occur, since denitrifying bacteria will preferentially use oxygen in the oxidation process because it yields more energy (U.S. Environmental Protection Agency, 1993). From Fig. 1.2, it is evident that there are other nitrogen input and removal mechanisms, however these are not considered in this simplified model.

### 1.1.2   Nitrates From Septic Tank Effluent

The functional portion of an OWTS consists of three major components: the septic tank, the drain field and the soil into which the drainfield discharges.  A schematic of a typical OWTS installation is shown in Fig. 1.3. When household wastewater enters the septic tank, solids will settle to the bottom, forming sludge while oils and grease move to the top (scum). After a period of time, the contents of the septic tank will degrade, after which the wastewater is moved (through gravity or with the assistance of a pump) from the tank to the drainfield.  Screens prevent the sludge and scum from exiting the tank into the drainfield. The septic tank must be pumped out periodically to prevent an excess of sludge and scum the build up of which may cause the OWTS to fail. Another reason for failure is the disposal of toxic substances, for example gasoline, into the septic system. These toxic substances will effectively stop the organic processes that are necessary for the degradation of household wastes (U.S. Environmental Protection Agency, 2002).
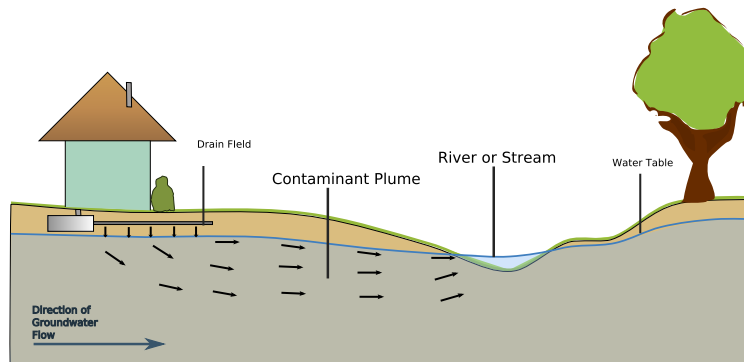
**Figure 1.4:** Flow of nitrates once they reach the water table

While nitrate concentrations are low in the septic tank and in the effluent discharged to the drain field (0 to 1.94 mg/L), once the septic tank effluent (STE) enters the soil, aerobic conditions cause the ammonium in the STE to undergo nitrification, causing the nitrate concentration at the water table to increase to between 25 and 80 mg/L (McCray et al., 2005). Once nitrates reach the water table, they become highly mobile and will migrate according to the flow patterns in the area. One possibility is that they will flow through the surficial aquifer and discharge, through groundwater discharge zones, to a surface water body (Fig. 1.4). The likelihood of this occurring is much higher if the aquifer is highly porous and the drainfield is located relatively close to the water body. Both of these conditions are satisfied in many locations in Florida.

### 1.1.3   Previous Modeling Efforts

Because nitrate contamination and contaminant transport in general is such an important component of environmental resource management, a significant amount of effort has been spent towards constructing models that can aid in the understanding of physical systems for the purpose of determining contaminant sources, their impact on the surrounding environment, mitigation and treatment schemes for the reduction of said impact, and for aiding in policy making and regulatory decisions for preventing contamination in the first place.

**Traditional Models**

Constructing a groundwater flow model is a necessary first step towards modeling the transport and fate of contaminants since in order to determine how contaminants move, it is necessary to first know how groundwater moves. Traditional groundwater models, such as the widely used MODFLOW (McDonald and Harbaugh, 1988) or FEFLOW (Diersch, 2006), tackle this problem by numerically solving the governing partial differential equation (PDE) of groundwater flow

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_{yy}\frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_{zz}\frac{\partial h}{\partial z}\right) + W = S_s\frac{\partial h}{\partial t} \tag{1.1}$$

using finite difference (MODFLOW) or finite element (FEFLOW) techniques. In Eq. (1.1), $h$ [L] is the hydraulic head, $K_{xx}$, $K_{yy}$ and $K_{zz}$ [L/T] are the hydraulic conductivities of the flow medium in the $x$, $y$ and $z$ directions, $W$ [1/T] is the source and sink term, $S_s$ [1/L] is the specific storage of the medium and $t$ is time (the units of each term are given in the square brackets; e.g., L/T means length per time). After the appropriate boundary conditions are determined, and the domain discretized, the model can be run and analyzed or the results passed to another model, such as a transport model. Contaminant transport modeling is generally much more involved than groundwater flow modeling due to the increased complexity of the model, such as the need to consider not only the movement of contaminants but their chemical interaction with other substances, as well as the environment. The PDE considered by the popular transport modeling software MT3DMS (Zheng and Wang, 1999) is

$$\frac{\partial\left(\theta C^k\right)}{\partial t} = \frac{\partial}{\partial x_i}\left(\theta D_{ij}\frac{\partial C^k}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(\theta v_i C^k\right) + q_s C_s^k + \sum R_n \qquad (1.2)$$

where $\theta$ [-] is the medium's porosity, $C^k$ [ML$^{-3}$] is the dissolved concentration of species $k$, $t$ is time, $x_{ij}$ [L] is the distance along the respective Cartesian coordinate, $D_{ij}$ [L$^2$T$^{-1}$] is the hydrodynamic dispersion tensor, $v_i$ [LT$^{-1}$] is the seepage velocity, $q_s$ [T$^{-1}$] is the flow rate per unit volume of sources and sinks, $C_s^k$ [ML$^{-1}$] is the concentration of the source or sink for species $k$, and $\sum R_n$ is the chemical reaction term. The chemical reaction term can take several forms. Considering only sorption and first order decay, the term becomes (Zheng and Wang, 1999)

$$\sum R_n = -\rho_b \frac{\partial \bar{C}^k}{\partial t} - \lambda_1 \theta C^k - \lambda_2 \rho_b \bar{C}^k \qquad (1.3)$$

where $\rho_b$ [ML$^{-1}$] is the bulk density of the medium, $\bar{C}^k$ [MM$^{-1}$] is the concentration of species $k$ sorbed on the subsurface solids, $\lambda_1$ [T$^{-1}$] is the first order reaction rate for the dissolved phase and $\lambda_2$ [T$^{-1}$] is the first order reaction rate for the sorbed phase. Like the flow equation, there are several numerical methods for solving the transport equations.

While traditional models are capable of producing results that agree well with field measurements, their setup can be quite complex, requiring a large amount of input data, and the execution time can be quite lengthy. Additionally, the user must take into account possible model inaccuracies that may not be immediately obvious from examining the output, such as inaccuracies arising from the various numerical schemes used. Additionally, and especially for transport modeling, depending on the reactions modeled in Eq. (1.3), a detailed knowledge (i.e., chemical reactions) of the system being studied is required. Furthermore, some powerful models (notably MODFLOW and MT3DMS) are difficult to use due to the fact that model inputs must be specified via text files and the model results must be analyzed using third party software.

In order to alleviate the problem of the models being hard to use, many graphical interfaces to these tools exist, both GIS-based and non GIS-based. These interfaces generally act as pre- or post-processors to the actual model. Some examples of GIS-based front-ends to flow models

include MODFLOW Analyst, which is part of the Arc Hydro Groundwater Tools (Aquaveo, LLC), or ArcAEM (Silavisesrith and Matott, 2005) which is a front end to the analytic element modeling software SPLIT (Bandilla et al., 2006). A specific application of a contaminant transport model integrated with a GIS is the one described by Becker and Jiang (2007) where groundwater flow was simulated using the analytic element method and Laplace transforms were used to calculate the mass flux at a specified boundary. Even though these tools simplify working with these numerical models, the issues of large data requirements and the inherent model complexity remain.

**Simplified Models**

The goal of simplified models is to reduce the complexity of the modeling process, providing approximate solutions to Eqs. (1.1) and (1.5) by making assumptions and simplifications to the system being modeled, so as to reduce model construction times and the computational burden in order to provide quick estimates of the quantity of interest. These simplifications mean that the models cannot be applied to more general problems like traditional numerical models can.

In the context of nitrate transport modeling, several tools consisting of simplified models for flow and transport have been developed, each having varying degrees of simplification. Some of these models were listed by McCray (2009) and include MANAGE (Kellogg et al., 1996), NLM (Valiela et al., 1997) and PLSM (Adamus and Bergman, 1995). Heinen (2006) provided a comparison and summary of a large number of simplified models that specifically consider denitrification as a component. The topography-based model TNT2 (Beaujouan et al., 2002) considered nitrogen and nitrate transport at the catchment scale due to agricultural activities.

Simplified GIS-based modeling tools exist either for flow modeling, transport modeling, or both. Ye et al. (1996) presented a map-based subsurface flow modeling tool that integrated directly with GIS. The models given by McCray (2009), are simplified GIS-based screening models for nitrogen transport and fate however they did not necessarily focus on the transport and fate of nitrates. Schilling and Wolter (2007) developed a GIS-based model that focused on nitrate pollution from agricultural sources, which also estimated groundwater travel time by using a DEM to approximate the hydraulic gradient.

Comparatively few tools have been developed for models specifically focusing on the transport and fate of contaminants originating from OWTS, The New Jersey Department of Environmental protection developed a tool to determine the effect of dilution on nitrate contamination from OWTS (Hoffman and Canace, 2004), the goal of which was to determine the average area required per OWTS in order to dilute the STE to acceptable levels. The model SepTTS (Schecher, 1997; Lee et al., 1998) is a simplified model that focused on the fate and transport of household chemicals from OWTS but it did not focus on nitrogen or nitrate.

Two simplified contaminant transport models that are important to mention due to their similarity with the model developed and presented here are BIOSCREEN (Newell et al., 1996) and BIOCHLOR (Aziz et al., 2000), both developed for the EPA. BIOSCREEN and BIOCHLOR are

both simplified contaminant transport screening-level models implemented in Microsoft Excel. BIOSCREEN aimed to model the biodegradation of petroleum hydrocarbons in groundwater. BIOCHLOR had a similar goal, however, the target compounds were chlorinated solvents. The simplified model presented in this work uses the analytical solution to the advection-dispersion equation (Eq. (1.5)) based on the one introduced by Domenico and Robbins (1985) and further developed by Domenico (1987) and Martin-Hayden and Robbins (1997), hereafter referred to the Domenico solution. BIOSCREEN and BIOCHLOR are both based on the Domenico solution and therefore serve a useful guides on how the Domenico solution can be applied. This is especially relevant for our model due to the similarities with BIOSCREEN and BIOCHLOR. For example, comparing our model to BIOSCREEN, our model considers contaminants from OWTS after they have reached the saturated zone. This can be considered similar to the contaminant source considered by BIOSCREEN. Another similarity is that both models consider a process of biodegradation (also known as natural attenuation). In BIOSCREEN, the contaminants of interest are hydrocarbons which, when introduced into groundwater, react with other compounds present in groundwater (such as nitrates) thereby becoming attenuated. In our model, the contaminant of interest is nitrate, and the attenuation takes place with background carbon already dissolved in groundwater. Because of the similarities, it would be possible, in principle, to use an alternative method of attenuation, used by BIOSCREEN in addition to first order decay (from the Domenico solution), known as instantaneous "biodegradation". In this type of attenuation, the system has a "biodegradation capacity" which is used to correct concentration values predicted by the Domenico solution.

## 1.2    Objectives

Although there are many modeling tools and software programs for simulating the transport and fate of contaminants, including nitrate, many times, those models and tools have a large number of parameters and require a large amount of input data, in addition to having a steep learning curve and not being user friendly. The net effect of these factors is the requirement of a large budget, both for data collection and for the recruitment of experienced modelers, as well as a lengthy time frame in which to complete the modeling study. While it is possible for the results generated by those models to agree well with field measurements, the models may be to difficult to use and generally unwieldy for the purposes of obtaining a quick estimate of the quantity of interest. As a result, simplified models have been developed to address those issues. The focus of these simplified modeling tools is not the generation of a highly detailed representation of the study area. Rather, they are meant as a screening tool to provide order-of-magnitude estimates while at the same time have reasonably quick and easy set-up as well as fast execution times compared to traditional numerical contaminant transport models. These order-of-magnitude estimates can then be used to determine whether a particular area warrants further investigation by means of additional data collection and through the use of more advanced models. The degree of simplification varies from model to model, from some models making extremely broad assumptions about the system, to models that consider details up to a point where it is not far from the details considered by

advanced numerical models. The approach taken with the model developed and presented here is the simplified model approach. The degree of simplification falls somewhere in between the simplest of models and the highly detailed simplified models. This approach is suitable for the purpose intended by the FDEP which is to provide quick, but scientifically based estimates of nitrate loading of OWTS to nearby surface water bodies such as rivers or lakes.

Nitrate attenuation mechanisms in groundwater play an important role in the effect of nitrate load on a surface water body. An important facet of this model is the consideration of nitrate attenuation due to denitrification. Without considering nitrate removal mechanisms, nitrate loads can be overestimated, resulting in unnecessary nitrate reduction initiatives. An overview of how denitrification is handled in the model is provided in Section 1.4 with full details given in Chapter 4. An introduction to the nitrogen cycle, of which denitrification is a part of, is given in Section 1.1.1.

An important limitation of other simplified contaminant transport models of the same complexity level (e.g., BIOSCREEN (Newell et al., 1996), BIOCHLOR (Aziz et al., 2000)) as the one presented is that they only consider a single contaminant source. This model considers multiple sources while keeping model complexity within the simplified modeling regime. The consideration of multiple sources introduces an additional dimension into the modeling process, namely, the dimension of spatial location. Because multiple OWTS can be located in arbitrary locations with respect to each other and nearby water bodies, the specific arrangement will affect contaminant concentration distributions. A GIS is an ideal tool to work with this new spatial dimension.

Because the modeling of nitrate contaminants from septic tanks contains a spatial component, a GIS becomes a natural choice for enabling the efficient management, quick and easy visualization, and analysis of the spatial nature of the model results. GIS are widely used tools in government and industry that enable users to efficiently work with any kind of data that embodies a spatial component. Because of the significant advantages that a GIS provides, the simplified model, along with its full graphical user interface (GUI), is implemented as a plug-in to the commercial software GIS package ArcGIS (ESRI Inc.) using the Visual Basic .NET (Microsoft Corp.) programming language. ArcGIS, a well known and widely used GIS, was selected mainly due to its wide industry acceptance and the fact that many users are already familiar and comfortable working within the ArcGIS environment.

The focus of this discussion is to present the development of the model from a scientifically and mathematically justifiable standpoint, including any assumptions and approximations made, and the potential impact on the accuracy of model results arising from the aforementioned approximations and assumptions as well as resulting model limitations. Additionally, a high level discussion on the implementation the model is presented. Although important for real-world scenarios, a detailed validation scenario, such as one that includes calibration with site specific data will not be undertaken here due to time and budgetary constraints. Instead, a verification and validation study using a combination of real and synthetic input data is provided for each component of the model, so as to evaluate the performance of each individual model component separately from one another. Although this tool may be used for resource management purposes, a discussion of

management practices and policy regarding septic tank regulation and contamination will not be undertaken.

## 1.3   Technology and Terminology

In order to more effectively understand the discussion presented in subsequent sections and chapters, it is instructive, especially for readers unfamiliar with GIS-related terminology, to present a brief introduction on geographical information systems. Additionally, for those unfamiliar with the .NET Framework (Microsoft Corp.) and how the interaction between separate pieces of software occurs in a Windows (Microsoft Corp.) environment, a short description is provided.

### 1.3.1   Geographical Information Systems

A GIS can be defined in several ways. A simplistic definition would be to state that GIS is a tool for creating and displaying maps. In the early history of GIS, this may have been true however geographical information systems have evolved considerably in the four-and-a-half decades since the first identifiable systems were developed by Environment Canada (CGIS) and the State of Minnesota, amongst others (Peuquet and Marble, 1990). Longley et al. (2005) provided a table listing various definitions of a GIS and one that is particularly relevant is the following: A GIS is

> "A tool for revealing what is otherwise invisible in geographic information." (Longley et al., 2005, Table 1.3)

In the context of modeling nitrate transport from septic tanks, the previous definition is relevant in that by using GIS, a spatial pattern (e.g. certain zones or "hot spots" of contamination) may be revealed by simple inspection and further analyzed. Martin et al. (2005) defined a GIS in terms of water resource models as

> "a popular spatial analysis, interpretation, and display method"

Martin et al. (2005) also linked GIS to the idea of decision support systems (DSS) in science and engineering. These DSS, which utilize GIS as a part of the problem solving procedure, enable the solution of more complex problems by facilitating tasks such as planning and design of projects with a strong location-oriented component.

In the context of the discussion in subsequent chapters, it is informational to define the two primary ways in which information can be represented within a GIS: the *vector* representation and the *raster* representation, corresponding to the vector data model and the raster data model respectively. The vector data model is useful for representing discrete objects or *features* that have a specifically identifiable location in space. These features can then be encoded using an appropriate

geometry or *feature class*. For example, the locations of the center of septic tanks have a uniquely defined point in space. Correspondingly, this location information can be encoded using a point feature class. The location of the banks of a river is a one dimensional entity whose location on the surface of the earth can be encoded using a line feature class. Similarly, a lake can be encoded using a polygon class where the area enclosed by the boundary of the polygon represents the lake. The raster data model is used to represent entities that do not have a discrete location in space, such as continuous fields. An example of a continuous field is topographical elevation information, which can be in the form of a digital terrain model or digital elevation model (DEM). In a DEM, although the land surface elevation is defined everywhere on the surface of the earth, it can only be sampled at discrete points. The raster data model divides up the domain into a grid of cells where the value of each cell represents the discretized version of the continuous field of interest, obtained by sampling the continuous field. A major difference between the vector and data models is that normally, the amount of disk space required to encode vector data is much smaller than raster data. This is because with the vector model, the information is defined mathematically while for the raster model, each cell value must be stored individually. For example, when storing the location of a river bank with the vector data model, only a few nodes need to be saved. The location of the river bank between the nodes is known from the mathematical definition of the line connecting them. The description of raster and vector data presented here is summarized to the parts relevant to the discussion in the following sections and chapters. Longley et al. (2005); Gogu et al. (2001) and many others provide more detailed definitions of raster and vector representations as well as additional data representations. Readers are directed to those resources for more in-depth information regarding those topics.

Another expansive topic that will only be discussed briefly here is the concept of coordinate systems and map projections. In order to be able to define the location of a point on the surface of the earth, a point of reference, a datum, is needed. A widely adopted reference for North America is the North American Datum of 1983, NAD83 (Longley et al., 2005). NAD83 defines, among other things, the ellipsoid that best represents the surface of the earth. This reference is then used as a basis of the spherical coordinate system that gives the location of a point on the earth in terms of latitude and longitude. Because the surface of the earth is not flat, a projection of the curved coordinate system is needed so as to be able to represent the surface in two dimensions. Again there are many such projections and coordinate systems. The most convenient projection and coordinate system to use when it comes to the model developed here is the transverse Mercator projection and the Universal Transverse Mercator (UTM) coordinate system because this choice of coordinate system and projection facilitates accurate (<0.04% error) calculations involving short distances (Longley et al., 2005). In the UTM projected coordinate system, the surface of the earth is divided into zones where each zone has been selected such that by applying a specific transverse Mercator projection, the distortion is minimized. As a side note, most of Florida is located in UTM zone 17 (17N in north-south notation or 17R in latitude band notation).

In the context of this work, it is useful to define terminology specific to ArcGIS. In ArcGIS, there are many ways to encode information in the vector data model. One such method is the

"shapefile" method. A shapefile encodes the vector data model by using a simplified relational database management system (RDBMS) to store individual features. Each feature can then have additional attributes (called *fields* in RDBMS terminology) which can then be manipulated using a subset of SQL (structured query language) commands. Analogously, ArcGIS supports many different encodings of raster information. The encoding chosen here was the IMAGINE Image format, essentially a lossless image compression image format.

### 1.3.2  .NET Framework

The model was implemented using the Visual Basic .NET (Microsoft Corp.) programming language (a.k.a VB .NET) which makes use of the .NET framework. There are several versions of the framework, version 3.5 was used for this project. The purpose of the .NET framework is to simplify application development and deployment by providing a robust application environment which can be used by multiple programming languages across different hardware and software architectures (there is a .NET implementation for Linux) by introducing a common language infrastructure (CLI) that includes definitions of shared programming constructs and types via a common language specification (CLS) and a common type system (CTS) (Gittleman, 2003; Troelsen, 2008), while at the same time providing a comprehensive set of class libraries. The .NET framework is similar to the Java programming language in that both approaches seek to hide details of the underlying architecture by using the idea of a virtual machine; Java has the Java virtual machine (JVM) and .NET has the common language runtime (CLR). While the JVM and CLR are similar in terms of features (support for object oriented languages, garbage collection, platform independence, type safety, etc.) and performance, the CLR has the distinct advantage of being language independent (Singer, 2003).

Because the model presented in this work is implemented as a plug-in to ArcGIS, communication between the model and ArcGIS is required so as to make full use of the features provided by the software as well as to enhance the user experience. Communication is achieved through interfaces and class libraries (called *assemblies* in .NET terminology) provided by the ArcGIS software development kit (SDK). ArcGIS collectively refers to the set of libraries and interfaces provided by the ArcGIS SDK as "ArcObjects". These assemblies are loaded by the model and interaction can then take place. In reality, the .NET assemblies provided by ArcGIS are merely wrappers around C++ libraries that use a different inter-process communication method known as COM (component object model), however thanks to the .NET COM Interop libraries, a COM object can be handled like a .NET object via a runtime callable wrapper (RCW).

## 1.4   Overview

### 1.4.1   Conceptual Model and Design

Although the model can be treated as a single entity, it can be logically separated into three distinctly identifiable components: the groundwater flow model, the nitrate transport model, and a component dealing with computing the nitrate load. Of course each component is not completely independent of one another as the results from the flow model will be used by the transport model, whose results will then be utilized to estimate the nitrate load. Because of the distinct functionality of the three components, it is also reasonable to divide the project into three logical modules, each relating to a specific component and its implementation, as well as an additional fourth component consisting of functionality not related to the actual models themselves. It is now important to make the distinction between a component and a module. The Flow, Transport and Nitrate Load Estimation *modules* consist of the combination of the respective component of the model and associated functionality (e.g. GUI, ArcGIS interaction, etc.). The focus of the discussion contained in subsequent chapters will be of the Flow, Transport and Nitrate Load Estimation modules.

A brief overview of each model component is now presented. The first component is the groundwater flow model. The flow model simplifies groundwater flow by using a DEM and processing it with a smoothing filter, the result of which is then used to approximate the hydraulic gradient. Once the hydraulic gradient ($\nabla h$) is known, a simple application of Darcy's Law (Eq. (1.4)) yields the groundwater flow velocity (seepage velocity) given the soil porosity ($\theta$). The porosity of a volume of soil is the fraction of the total volume that is void space.

$$v = -\frac{K}{\theta} \cdot \nabla h \tag{1.4}$$

The flow model considers two-dimensional, steady-state flow within the saturated zone of the surficial aquifer. All flow occurring at or below the water table is defined as saturated flow. The flow model is discussed in detail in Chapter 2. While not a part of the flow model strictly speaking, a particle tracking routine is used to calculate flow lines from the generated flow field. This particle tracking function is the fourth component mentioned at the beginning of this section.

The next component is the transport model. The purpose of the transport model is to estimate the migration of nitrate from septic tanks once it enters the saturated zone. The model output is a contaminant plume (or plumes if there is more than one source) representing the concentration distribution of nitrate contaminants in the study domain. Data requirements are reduced, and model application is simplified by using an analytical solution to the governing equation for modeling contaminant transport and fate, known as the advection-dispersion equation (ADE) (also called the convection-diffusion or advection-diffusion equation). The analytical solution used is one based on the solution introduced by Domenico and Robbins (1985). For simplicity, and so as to be able to solve the advection-dispersion equation in closed form, the version of the advection-dispersion equation considered by Domenico and Robbins (1985) allows for advection

only in the $x$-direction however dispersion is considered to be in all three directions. The version of the ADE with decay (the one shown in Eq. (1.5)) that was considered by Domenico (1987) is

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + D_z \frac{\partial^2 C}{\partial z^2} - kC \tag{1.5}$$

where $C \equiv C(x, y, z, t)$ $[ML^{-3}]$ is the concentration of nitrate, $v$ $[LT^{-1}]$ is the seepage velocity, $D_x$, $D_z$ and $D_z$ $[L^2 T^{-1}]$ are the dispersion coefficients in the $x$, $y$ and $z$ directions and $k$ $[T^{-1}]$ is the first-order decay constant. The Domenico and Robbins solution was used as a base for the solution to Eq. (1.5), presented by Domenico (1987) (Eq. (3.4). Discussed further in Section 3.1.1). The Domenico solution to Eq. (1.5) is a fully three-dimensional, transient state solution incorporating first-order decay. Due to constraints regarding execution speed and memory limitations on desktop computers, as well as the inefficient handling of three-dimensional data within GIS, the two-dimensional steady-state form of the Domenico solution is used. Chapter 3 deals with the transport model and the Domenico solution is discussed in more detail in Section 3.1.1. It is important to note that it is the transport model that considers, via the decay term, the effect of denitrification on the plume's shape and size. According to McCray et al. (2005), first-order decay is the most common method of considering denitrification.

Finally, the nitrate load estimation component determines the amount of nitrate removed due to the process of denitrification and determines the nitrate load to the target water body. A mass balance approach is taken in order to calculate denitrification. Because the transport model produces a concentration distribution in the steady state, we know that the amount of mass flowing into the system must equal the total amount of mass being removed from the system. Because we can calculate the amount of mass being removed through denitrification by determining the denitrification rate using the relationship for first order decay, $rate = -kC$, and multiplying it by the volume of solute, we can then determine via subtraction, the amount that must be entering the water body (i.e. the nitrate load). As is to be expected, the Transport module and Nitrate Load Estimation module are not completely independent as the Transport module determines via the solution to the Domenico analytical equation, the removal behavior of the nitrate contaminant. This interaction, as well as the general procedure is discussed in the chapter dealing with the Nitrate Load Estimation module (Chapter 4).

Both the flow and transport models are steady state. The steady state condition is adopted because the quantity of interest is the long-term effect of OWTS on the nitrate load to a given water body. It is assumed that in the long-term, the system will have reached steady-state conditions. By using steady-state condition, it enables the calculation of the mass of nitrate removed due to denitrification using the mass balance approach as previously explained. Adopting steady-state conditions also has the additional effect of keeping the model simpler to interpret.

## 1.4.2   Implementation

This section includes a discussion of some of the more general considerations that were part of the decision-making process regarding the model's implementation, as well as giving a high-level overview of the model's structure and organization.

**Practical Considerations**

It was previously stated in Section 1.2 that the model was developed as a plug-in (or *extension* in ArcGIS terminology) to ArcGIS. ArcGIS was selected based on several factors. One important factor was the widespread industry support for ArcGIS. By adopting ArcGIS, it makes the model usable by a larger number of people, since the eventual intention is to release the model as a free and open source tool. Other less important criteria were the wide array of analysis functions made available to the user, reasonable developer documentation and resources, as well as the presence an active user and developer community.

One of the primary advantages of using a GIS is that it simplifies preparation of model inputs and the visualization of results by taking advantage of the graphical nature GIS to facilitate the intuitive interaction with model data. By using the GIS to prepare model inputs, a significant amount of effort will be saved on the part of the user, compared to using text files to prepare inputs and command line tools to run the model and process results. An additional advantage of using a GIS is the ability to easily and efficiently share modeling results with others. Often, modeling tools require inputs to be in a format specific to that model. Similarly, model results will be in a format that must be translated so that it can be used by external visualization applications (e.g. plotting tools) or by other models. By using GIS, model input and output formats are standardized to a format understood by all users of that particular GIS. If another format is required, the GIS will likely have the functionality to easily convert it to the required format without the need for external tools.

The plug-in interaction method between the model and the GIS is termed "embedded coupling" by Gogu et al. (2001), Longley et al. (2005), and Martin et al. (2005). Embedded coupling has an advantage over other interfaces by providing the user with a more familiar environment. In embedded coupling, the model runs entirely within the GIS environment where all the model's functionality is accessible via the GIS interface and any functions or libraries needed by the model are embedded within the model code. If such functions are located externally, they are directly accessible by the model. Embedded coupling is made possible by a GIS that supports an integrated programming or scripting environment(e.g., Python or VBScript) and/or provides access to its internal objects and data structures via some mechanism (e.g., COM). While embedded coupling provides significant advantages in terms of being able to utilize the functionality of the GIS for displaying and analyzing model results as well as providing a comfortable environment for the user, there may be significant drawbacks. One of the primary disadvantages to embedded coupling is that it often requires the use of simplified models. This disadvantage arises due to the limitations

of GIS data structures that are not well suited to handle rigorous or complex mathematical models, such as ones that require the solution of differential equations (Martin et al., 2005). In cases like the one presented here, where a simplified model is desirable for other reasons, an embedded model implementation is not disadvantageous. One drawback of concern is the fact that because GIS were not designed with a modeling purpose in mind, the performance (e.g. execution speed and memory utilization) of the model may be degraded compared to a model that has been developed with a loose coupling interface to the GIS. In loose coupling, the model runs separately from the GIS and the GIS is used mostly for pre- or post-processing of the model data. According to Gogu et al. (2001), another disadvantage to embedded coupling is that it requires a significant investment in programming compared to other types of coupling. This additional programming investment arises from the fact that the model must conform to the data structures and file formats supported by the GIS. Enabling this support requires additional programming so that the model creates and uses data structures in files that can then be directly used by the GIS. Furthermore, an additional investment in time is required to learn the application programming interface (API) supported by the GIS as well as the programming effort required to implement the additional, non-model related code (e.g., user interface and related code). Even with its disadvantages, embedded coupling is still worthwhile as it greatly improves the user experience compared with loosely coupled models.

In order to implement the model using the idea of embedded coupling, there were several choices of programming languages and environments that were considered. These were: C++, VBScript (an embedded programming environment within ArcGIS), Python (also embedded within ArcGIS) and the .NET framework languages (of which Visual Basic .NET and C# were considered). VBScript was quickly discarded due to its lack of support for object-oriented programming. As mentioned previously VB .NET was selected as the development environment, the reasons for which are discussed subsequently. Firstly, the .NET framework met the FDEP requirement of making the model easy to modify and maintain. Secondly, .NET was designed from the ground up to be able to interact with a GUI and as such, provides a rich toolset and tightly integrated development environment for their implementation. Although the other options provide methods for GUI creation and interaction, their development is much more laborious than with .NET. Thirdly, .NET provides a comprehensive set of libraries for many common functions ranging from file I/O to multi-threading to bitmap graphics manipulation. Additionally, it provides several useful language constructs such as generics (similar to templates in C++) and language integrated queries (LINQ) which allow for SQL-type queries to be run on ordinary lists, arrays and other enumerable collections. These constructs serve to produce clean and readable code. Finally, although more verbose than C# or Python, VB .NET has a syntax that arguably makes it easier for a developer to quickly understand and analyze another developer's code.

**Project Structure**

In keeping with the object oriented paradigm, the code project was structured in a modular fashion. The GUI elements were kept separate from the model elements, with further modularization within
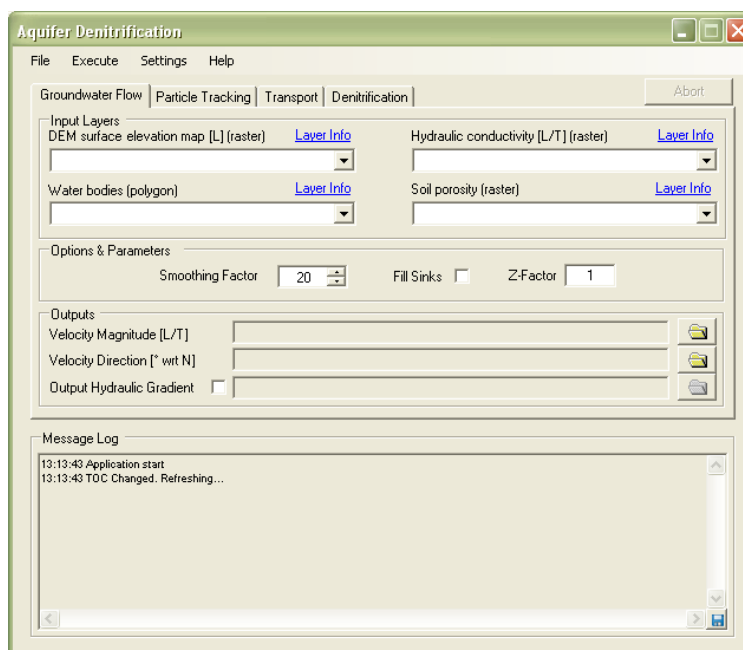
**Figure 1.5:** Screenshot of the main model GUI

the GUI and model submodules. The main panel of the model GUI (Fig. 1.5) serves to illustrate some of this modularization. The main component of the window is the tab panel where each tab represents a separate model component. Programmatically, each tab is a self-contained module that has been separated into its own class and is capable of executing individually of the other modules (tabs). Note that the stand-alone functionality of each module only extends to the practical implementation as conceptually speaking, the modules are dependent on one another to produce meaningful results. An advantageous effect of this functional separation into tabs furnishes the user the flexibility to selectively run parts of the model. A situation where this is useful is during model calibration when only one or two parameters can be changed without having to re-run the entire model.

The conceptual models behind each module as well as their algorithmic implementations are discussed in Chapters 2 to 4. Details of the programmatic implementation (the "nitty gritty") will be avoided as much as possible since programming details are only relevant to developers. Interested developers should examine the source code (for an overview of the program's code, see the project documentation help file).

# 2

## Groundwater Flow Module

This chapter presents the development and implementation of a two dimensional steady-state groundwater flow model capable of generating a map of groundwater flow velocity (seepage velocity) in the surficial aquifer. A key piece of information that is needed to calculate the seepage velocity is the hydraulic gradient. Once the hydraulic gradient is known, the seepage velocity can be approximated by a simple application of Darcy's Law. E.g., for the x-component of the seepage velocity, Darcy's Law reads

$$v_x = -\frac{K}{\phi}\frac{\partial h}{\partial x} \tag{2.1}$$

where $v$ [L/T] is the seepage velocity, $K$ [L/T] is the hydraulic conductivity of the porous medium, $\phi$ [-] is the soil porosity, $h$ [L] is the hydraulic head and $x$ [L] is the spatial coordinate. In this model, the shape of the water table is considered to be a subdued replica of the overlying topography and by using this approximation, the calculation of the hydraulic gradient and therefore seepage velocity becomes straightforward. This simplification eliminates the need to numerically solve differential equations (along with the associated stability and convergence considerations) and reduces the need to consider detailed boundary conditions, as is the case with traditional models. Additionally, a great deal of manpower is saved by not having to conduct traditional modeling (see Section 1.1.3).

Because this flow model has the potential to be used for many different study areas, the validity of the model must be examined and the conditions under which the model is valid determined. This is accomplished by constructing a semi-synthetic numerical model of one of the areas of interest (Lakeshore neighborhood). A further analysis of the model is carried out via the examination of the groundwater travel times predicted by the model. This analysis consists of comparing a calibrated numerical model (developed for the U.S. Naval Air Station, Jacksonville Florida) with this simplified model. These two analyses give a clearer indication of the conditions under which the model is valid and whether those conditions exist in the area of interest.
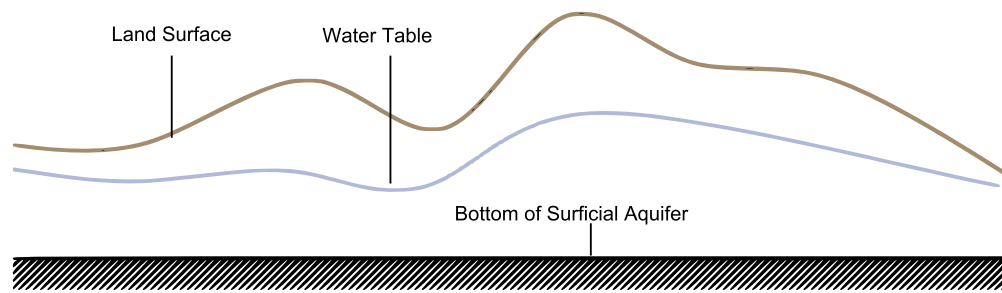
**Figure 2.1:** Conceptualization of the water table being a "subdued" replica of the topography.

## 2.1    Background

Maps of the phreatic surface are one of the most basic pieces of information required for virtually any investigation concerning groundwater flow. From a map of the water table, it is possible to derive many properties of the flow system, such as the direction of groundwater flow, the location of recharge or discharge zones and aquifer structure. Although this map of the water table is a very important piece of information, it is very difficult to obtain. A brute force approach would be to drill many wells in the area of interest; however the associated cost rapidly increases as the mapped area increases, rendering this method infeasible for large areas. One popular method of obtaining this map is to estimate the hydraulic head using numerical modeling tools such as MODFLOW (McDonald and Harbaugh, 1988). In this approach, a model of the area is constructed based on available information, including measured parameters such as the stage of nearby rivers, recharge amounts to the unconfined aquifer and general aquifer properties. The governing differential equation that defines the model is then numerically solved and a map of the water table obtained. While these traditional models can possibly produce simulated results that may agree well with field measurements, achieving such agreement generally requires extensive data collection of the study area, mainly due to the high data requirements of these types of models (Kalivarapu and Winer, 2008), as well as an experienced modeler to build and calibrate the simulation. In some circumstances, the sparsity of data necessitates a simpler model. In those circumstances, a simple model that has lower data requirements and is easier to build and run, is necessary. One such simplified model consists of assuming the shape of the water table generally follows the shape of the overlying topography (Fig. 2.1). As early as the beginning of the 20th century, King realized that the topography and the location of the water table were related (Desbarats et al., 2002). By assuming that the water table generally followed the topography, Desbarats et al. (2002) developed a method of interpolating sparse measurements of the water table elevation which were combined with information regarding the overlying topography thereby in principle, providing a better estimate of the water table elevation via interpolation alone.

Because topographic data is relatively easy to obtain compared with water table data, many hyrdologic models and groundwater models assume that the water table follows the topography to a certain degree. One example is the widely used TOPMODEL (Beven and Kirkby, 1979).

TOPMODEL is a physically based watershed rainfall runoff model which assumes that the water table is identical or very similar in shape to the overlying topography (Stieglitz et al., 1997). Another example in which the topography was assumed to be related to the shape of the water table, is Tóth's analysis of flow patterns in drainage basins. In his analysis, Tóth assumed that the water table was a subdued replica of the topography (Tóth, 1963). More recently, Sepúlveda assumed a relationship between the topography and the water table in order to construct a map of the water table elevation (Sepúlveda, 2002). His method was used by the Florida Aquifer Vulnerability Assessment (FAVA) study which examined the potential of contamination of Florida's aquifers (Arthur et al., 2005). The assumption that the topography controls the water table at local and regional scales has been used to derive analytical solutions to a number groundwater flow problems, such as the problem of surface-groundwater interaction solved by Wörman et al. (2006).

Even though the assumption that the shape of the water table can be approximated by the overlying topography is wide spread, relatively few authors have attempted to quantify under what circumstances the assumption is valid and to what degree the shape of the water table follows the topography. Anderson Jr. et al. (2000) presents data, obtained using direct water table measurements, ground penetrating radar measurements and numerical models, which suggests that the water table for a barrier island is a subdued replica of the topography. Haitjema and Mitchell-Bruker (2005) analyzed the water table-topography relationship via a synthetic model in order to develop a rule of thumb for determining whether a water table can be considered a subdued replica of the topography (see Section 2.4.4). An attempt to quantitatively investigate the shape of the water table in relation to the topography was conducted by Shahbazi et al. (1968). By comparing the results of a two-dimensional numerical model with the observations of a physical model, Shahbazi et al. (1968) concluded that the previously held assumption of prior studies, namely the assumption that the water table follows the topography everywhere, was false. They stated that the shape of the water table did not follow the topography in areas of high topographical elevation (i.e., the area between two valleys) however in areas of lower elevation and in discharge zones, the water table was coincident with the ground surface. They also concluded that for the case of a homogeneous and isotropic aquifer, the shape of the water table was determined by the following factors (for the given boundary conditions of their model: no-flow on the left, bottom and right boundaries): the relative elevations between the peaks (high points) and valleys (low points) of the topography, the average slope of the ground surface, the distance between peaks and valleys on the ground surface, the number of peaks and valleys on the ground surface and the amount of recharge. Other studies have investigated the relationship between the topography and the water table in the context of determining the accuracy of the so-called "topographic index", used as a proxy for water table depth in TOPMODEL and other hydrologic models. The topographic index can also be thought of as a measure of the "wetness" potential of any given point in a catchment. The topographic index is generally calculated as

$$index = ln\left(\frac{a}{tan\beta}\right) \tag{2.2}$$

where *a* is the upslope area drained by a given location per unit contour length and *tanβ* is the slope of the ground at the given location. Studies by Thompson and Moore (1997), Desbarats et al. (2002), and Siebert et al. (1997) have shown that the topographic index (whose calculation is based on the assumption that the slope of the water table follows the slope of the topography) inconsistently predicts water table depths.

Regardless of how the map of the water table is obtained, once it has been obtained, the hydraulic gradient can be calculated using Eq. (2.1).

## 2.2   Conceptual Model

As mentioned in this chapter's introduction, the focus of this model is the estimation of the hydraulic gradient and seepage velocity. In this simplified model, the only data required as input are: a digital elevation model (DEM) of the study area, a map of the spatial distribution of the aquifer hydraulic conductivity, a map of soil porosities, and a map of water bodies. By processing the DEM with a smoothing filter, it is possible to generate a subdued replica of the topography from which the hydraulic gradient can be estimated. This estimate is then combined with the location of water bodies and aquifer properties to generate maps of flow direction and magnitude. It is important to note that only saturated flow (i.e., flow in soil that is fully saturated with water) is considered in this model. As will be seen later, recharge to the surficial aquifer is an important parameter in determining the shape of the water table, however in this simplified model, recharge (both from rainfall and effluent from the OWTS) and its transient effect on the shape of the water table will not be considered due to the steady-state condition of the model.

Because the model deals only with the surficial aquifer, using an approximation to the hydraulic gradient based on the topographic gradient is justified since under the right circumstances, it can be considered reasonable that a relationship exists between the shape of the water table and the topography (see Section 2.4.4). This is not the case for any aquifer below the topmost one since the hydraulic head for these confined aquifers will likely have no relation to the topography. By approximating the hydraulic gradient with the slope of the topography, it means that the model is not capable of considering flow from, to or through deeper aquifers. As a result, the model assumes all flow is restricted to the surficial aquifer, which implies that any sinks in the hydraulic head distribution must correspond to an area of groundwater discharge; a lake or a river for example. Any sinks in the head distribution that do not correspond to actual water bodies must be filled so that flow does not become "stuck" in such a sink. This sink filling idea is common when using topographic data such as Digital Elevation Models (DEM) for surface hydrologic models (Martz and Garbrecht, 1998; Pan et al., 2004), and it has been done in the context of groundwater flow in the simplified groundwater travel time model of Schilling and Wolter (2007). In some circumstances however, not filling sinks may produce more realistic flow fields if the flow lines of the area of interest do not interact with a given sink. This is especially true if the sinks in question are within broad valleys or pits. In these cases, sink filling will produce large flat areas which may
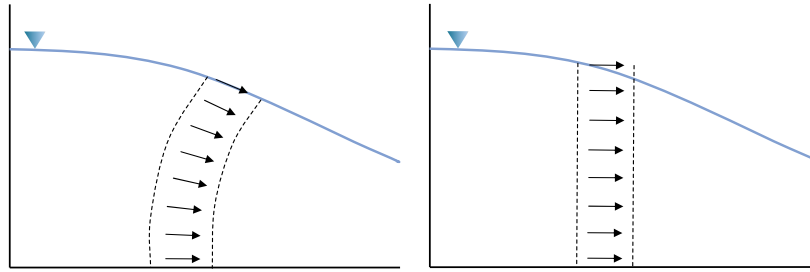
**Figure 2.2:** Illustration of the Dupuit approximation. The left figure shows flow lines with vertical gradients. The right shows flow under the Dupuit approximation. Dashed lines represent equipotentials and arrows represent flow direction. The solid line represents the water table.

unnecessarily distort the flow lines.

Flat areas in the input water table map cause problems with the flow calculation because in these areas, the flow magnitude will be calculated as zero due to the hydraulic gradient being zero. In this model, flat areas are considered to be spurious since it is unlikely that the hydraulic gradient in a topographically flat area is actually zero since it would mean that the water in the aquifer is stationary. In the case of using DEMs to calculate a map of the water table, as is done in this model, there can be several causes of flat areas. Two common causes are: limited vertical DEM resolution and rounding or truncation of elevation values. A more detailed discussion on the causes and treatment of flat areas is given in Martz and Garbrecht (1998). In this model, another cause of flat areas is the filling in of sinks described above (see Section 2.3.2 for more details). To the best of our knowledge, no other groundwater flow models that use topography to approximate the water table treat the issue of flat areas in the input topography.

The model makes use of the Dupuit approximation. Essentially, the Dupuit approximation allows us to approximate three-dimensional flow with a two dimensional gradient by ignoring the vertical flow gradients. Under the Dupuit conditions as illustrated in Fig. 2.2, flowlines are assumed to be horizontal and equipotential lines vertical. Additionally, the hydraulic gradient is assumed to be the same as the slope of the water table and invariant with depth (Freeze and Cherry, 1979). By making use of the Dupuit approximation, it allows for the use of a DEM to approximate the hydraulic gradient, the reason being that the DEM is inherently only capable of estimating the hydraulic gradient in two dimensions. The Dupuit approximation is reasonable when the slope of the water table is small and the average aquifer thickness is small compared to the aquifer extent (Freeze and Cherry, 1979).

Finally, the steady-state solution is used because it is deemed sufficient for the purpose of long-term environmental evaluation. Although the model is steady-state, different water table conditions for different times can be accounted for by controlling the amount of smoothing applied to the DEM.
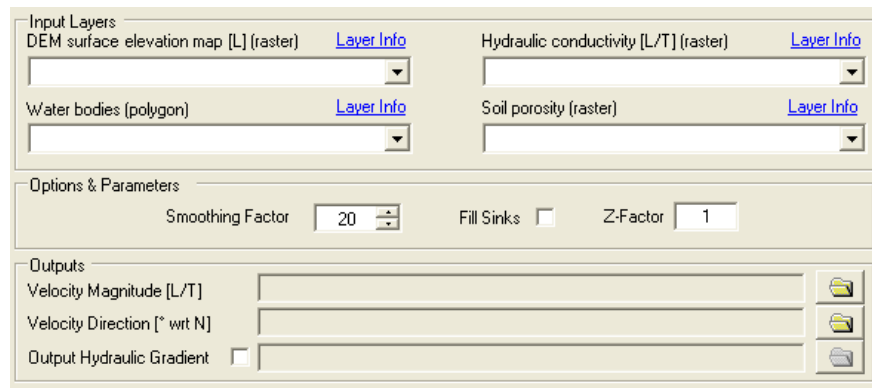
**Figure 2.3:** The user interface to the flow module.

## 2.3  Algorithm and Implementation

The desired output of the flow model is a set of raster files representing the groundwater flow velocity magnitude and direction. The user interface to the flow module is shown in Fig. 2.3. The model described in this paper is similar in purpose to the Darcy Flow tool provided by the Spatial Analyst ArcGIS extension . Like our flow model, the Darcy Flow tool also outputs a map of seepage velocity and direction however unlike our model, it requires as input a map of the water table; it does not calculate such a map on its own. Additionally, because the Darcy Flow tool requires a water table map as input, it has no functionality to deal with flat areas or sinks. The consequences of this were described in Section 2.2. The purpose of this section is not to give a guide on how to use the model or give details on its programmatic implementation. Rather, it is meant to discuss details of the algorithmic implementation.

### 2.3.1  Input and Output

The inputs to the model are (as shown in Fig. 2.3):

1. A DEM (raster) of the area of interest

2. A raster representing the spatial distribution of hydraulic conductivity for the area of interest

3. A raster representing the spatial distribution of porosity for the area of interest

4. A set of polygons representing the locations of surface water bodies

The parameters for processing the DEM are:

1. Smoothing factor. This controls how much the gradient of the topography will resemble the gradient of the water table.
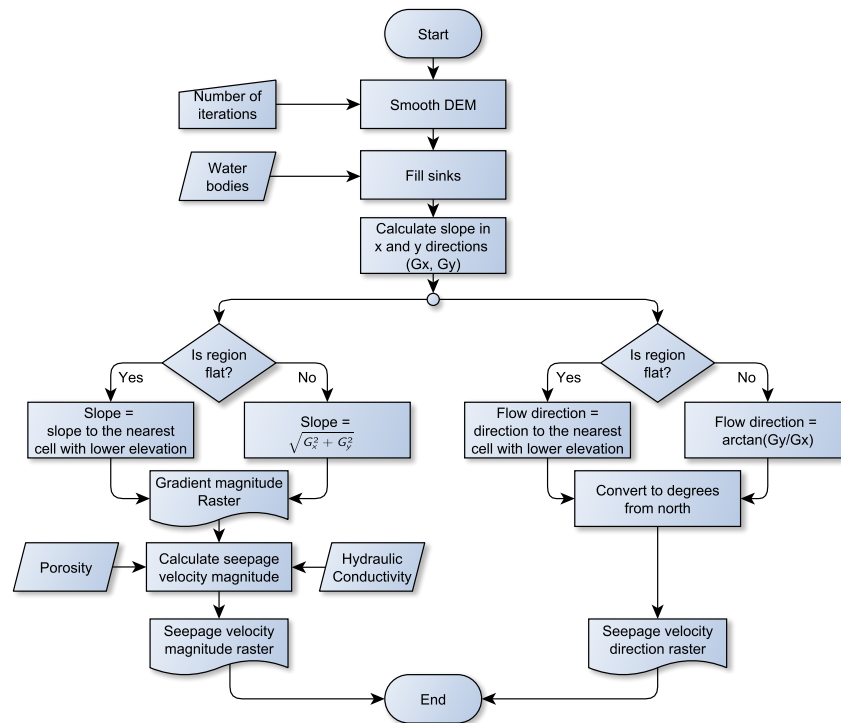
**Figure 2.4:** Logic flow for the groundwater flow model.

2. Z-Factor. This is the conversion factor between the vertical measurement units of the DEM and the horizontal measurement units.

3. Fill Sinks. Controls whether to fill sinks or not.

The model outputs are two rasters, one for the flow from each cell (in degrees clockwise from north), and the other for the magnitude of flow velocity from each cell in the direction of the flow direction (in the same units as the input hydraulic conductivity raster). Refer to the User's Manual for a more detailed explanation of the inputs.

### 2.3.2   Process

The model algorithm proceeds as shown in Fig. 2.4. First, the input DEM is smoothed by an amount specified by the user to generate the "subdued" replica of the topography. Smoothing is accomplished via a discrete cross-correlation of the input DEM with the 7 x 7 averaging kernel

(a.k.a moving window average)

$$\frac{1}{49}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \star R$$

where $R$ is a 7 x 7 block of pixels in the DEM and $\star$ is the cross-correlation (moving window or sliding window) operation. The net result is that each individual pixel in the output is the average of all the pixels surrounding it in a 7 x 7 rectangular window. The moving window operation can be thought of as a stencil placed over the pixels of the DEM covering an area of 7 x 7 elements. Each component of the kernel $K_{i,j}$ is multiplied with the value of the DEM directly underneath it, $R_{i,j}$. The sum of all such multiplications within the kernel window is then assigned to the cell in the center of the window (because of this, the dimensions of the kernel are necessarily odd). The window is then moved to the next DEM cell and the process repeated. The moving window or stencil operation is also sometimes referred to as the discrete convolution operation. This is strictly speaking, not true, although both operations are related. The cross-correlation operation is mathematically related to the convolution operator in that the kernel used for the cross-correlation must be rotated by 180° when used with the convolution operation. While other smoothing kernels were considered, such as Savitzky-Golay type filters (Savitzky and Golay, 1964), the simple averaging filter used in this study is capable of smoothing the DEM to a higher degree using fewer iterations.

Next, the smoothed DEM is processed to fill any sinks (if the option is enabled). Sinks are filled, using the ArcGIS "Fill" tool, which fills sinks up to their pour point, the minimum elevation along its boundary that has a neighboring downslope cell. A way to conceptualize this sink filling method is to imagine a bucket with a chip on its edge that is being filled with water. When the water reaches the top of the bucket, it will overflow through the chip, i.e. the pour point. Because the Fill tool fills all sinks in the domain, the elevations from the smoothed, unfilled DEM will be superimposed on the filled DEM in the locations of water bodies (note that the water bodies input layer is only used by the flow module when the Fill Sinks option is enabled). Effects due to this superposition may only be seen in certain cases where the flow direction in the raster cells immediately bordering a water body flow in the wrong direction (due to over-smoothing for example). Filling must be performed after smoothing because smoothing can re-introduce sinks in the sink-filled DEM. Filling sinks results in the creation of artificially flat areas where a sink has been filled (similar to the flat surface of the water in the overflowing bucket example). This is dealt with by the section of the algorithm dealing with flow direction calculations described later.

Eq. (2.1) is a 1-dimensional equation. In order to calculate two dimensional flow, it is necessary

to extend the gradient to two dimensions. The components of seepage velocity are then given by:

$$v_x = -\frac{K}{\phi}\frac{\partial h}{\partial x} = -\frac{K}{\phi}G_x \tag{2.3a}$$

$$v_y = -\frac{K}{\phi}\frac{\partial h}{\partial y} = -\frac{K}{\phi}G_y \tag{2.3b}$$

In this model, the calculation of the gradient is accomplished by using another stencil operation. In this case, the stencil operation used is the well known Sobel filter. The Sobel kernels are

$$G_x = \frac{1}{8h}\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \star R, \quad G_y = \frac{1}{8h}\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \star R$$

where $h$ is the cell size of $R$ in the $x$ and $y$ directions respectively. The Sobel kernels are essentially a combination of a triangle smoothing filter in one direction (Bartlett filter of order 3) and a central difference operator in the perpendicular direction. Given the above estimates of the partial derivatives in $x$ and $y$, it is possible to calculate the gradient magnitude and direction of the seepage velocity vector as

$$\begin{aligned} v &= \sqrt{v_x^2 + v_y^2} \\ &= \frac{K}{\phi}\sqrt{G_x^2 + G_y^2} \end{aligned} \tag{2.4}$$

and

$$\theta = \tan^{-1}\frac{G_y}{G_x} \tag{2.5}$$

Other kernels were tested for the calculation of $G_x$ and $G_y$ including ordinary central difference, one-sided difference, central differences of higher order and Sobel-type filters of higher order (higher order central differences with higher order smoothers). The ordinary Sobel filter yielded the best results, in that a) the gradient values did not vary too much in the local vicinity of a given cell (as expected since the input DEM was smoothed prior) and b) the computational cost was small (larger kernels take longer to compute).

Once $G_x$ and $G_y$ are calculated, the logic branches into two parts so as to generate two separate outputs. Starting with the velocity magnitude calculation (left branch in Fig. 2.4), the magnitude of the gradient can be easily calculated from $\sqrt{G_x^2 + G_y^2}$. A problem arises if a given area of the smoothed DEM is flat since in this case, the gradient is zero and water will not flow. According to the arguments given in the conceptual model, these areas must be eliminated. This is accomplished by making use of the ArcGIS FlowDirection tool. The tool implements the method of Jenson and Domingue (1988) to assign a slope to a flat area equal to the slope to the nearest cell with a lower elevation. After treating the flat areas, the gradient magnitude raster is then used to calculate the seepage velocity magnitude on a cell by cell basis using Eq. (2.4).

The calculation of seepage velocity direction is as follows: in non-flat regions, the direction is given by Eq. (2.5). In flat regions, the direction is calculated using the method of Jenson and Domingue (1988) (from the FlowDirection tool) where the direction is iteratively assigned starting from cells on the boundary of the flat area that have a downslope neighbor. Starting from a cell on the boundary, all cells in the flat area that do not have a defined flow direction and are immediate neighbors of the boundary cell are assigned a flow direction towards the boundary cell. In subsequent iterations, all cells in the flat area that do not already have a flow direction are assigned a direction to the immediate neighbor having the largest downslope value. The process is repeated until all cells have an assigned flow direction. In this way, the assignments of flow directions "grow" into the flat area starting from the boundary cells. The results from the flat and non-flat areas are then combined and converted to an angle between 0 and 360° clockwise w.r.t. north. Note that because the assignment of direction to a cell in a flat area occurs only from that cell to one of its eight immediate neighbors, the flow directions in flat areas are limited to one of eight directions, which is in contrast to the direction outside of flat areas which can be in any direction. When assigning flow directions to flat areas caused by sink filling, the procedure is the same however the results have a distinctive characteristic that all flow paths tend to travel towards the pour point of the filled sink.

### 2.3.3   Particle Tracking

An important part of the Groundwater Flow *module*, though not strictly part of the flow *model* itself, is the particle tracking functionality. Particle tracking consists of placing an imaginary particle at any given location within the flow field and tracing its movement through the flow field until it reaches a water body or exits the domain. This functionality is useful for two reasons: a) it allows for the visualization of the flow field and b), it is used by the Transport module to determine the path that the contaminants will take. The inputs to this sub module are shown in Fig. 2.5.



**Figure 2.5:** The inputs to the Particle Tracking Module

The algorithm used for particle tracking is quite simple but effective. Given a flow field (seepage velocity magnitude and direction rasters) and user defined parameters (the step size and maximum

number of steps to take), the particle begins by taking a step of size "Step Size", starting from the source location. The position and velocity of the particle at the end of a given step is calculated by using the cell value of the magnitude and direction rasters at the starting particle position for the step. Since the particle may be located at a point that does not correspond to the center of the raster cell, a nearest neighbor interpolation is used. For each point given by the Source Location input layer, the Particle Tracking module saves the sequence of segments that correspond to the flow path from that source. The segments are then linked to the point that they correspond to via the PathID attribute. The PathID attribute for each segment is set to the FID attribute of the input source point. In addition, each segment stores the current magnitude of the velocity flow field and the current value of the soil porosity. These values are used by the Transport and Nitrate Load Estimation modules. See the User's Manual for additional information regarding particle tracking.

## 2.4   Can the Water Table Be Considered a Subdued Replica of the Topography?

In this section, the validity of the assumption that the water table follows the topography is tested for the area of interest (Jacksonville, Florida). Specifically, two areas in Jacksonville are considered: the U.S Naval Air Station, Jacksonville (NAS) and the Lakeshore neighborhood, also in Jacksonville.

### 2.4.1   Description of Study Sites

The NAS and the Lakeshore neighborhood are located in close proximity. Lakeshore is approximately 6 km north-west of the NAS. The location of the NAS is shown in Fig. 2.6. The NAS itself lies on the banks of the St. Johns river and is bounded to the west by the Ortega river. Lakeshore is bounded to the east and west by the Cedar and Ortega rivers respectively. To the south, Lakeshore is bounded by the confluence of the Cedar and Ortega rivers and to the north, it is partially bounded by Big Fishweir Creek.

The general physiography of Duval county was studied by Fairchild (1972) while the NAS was studied by Davis et al. (1996). The composition of the surficial aquifer in the vicinity of the naval base consists of medium to fine-grained unconsolidated sands of Pleistocene and Holocene age rising to an elevation of approximately 9 m above sea level. Even though the surficial aquifer can be considered as a single unit and no base-wide impermeable zones have been identified, there are occurrences of local sandy clay beds (Davis et al., 1996). The base of the aquifer corresponds to the top of the Hawthorn group which is a confining clay layer which separates the surficial aquifer from the Upper Floridian aquifer. The depth to the Hawthorn group is variable but ranges between 3 to 30 m (10 to 100 ft) in the vicinity of the base. Recharge to the surficial aquifer was estimated to be approximately 255 mm/yr (10 in/yr) (Fairchild, 1972). On average, the hydraulic conductivity in the area is about 1.5-2.1 m/day (5-7 ft/day) as determined from slug tests (Davis et al., 1996). Vertical hydraulic gradients in the areas away from the river were determined to be small and the
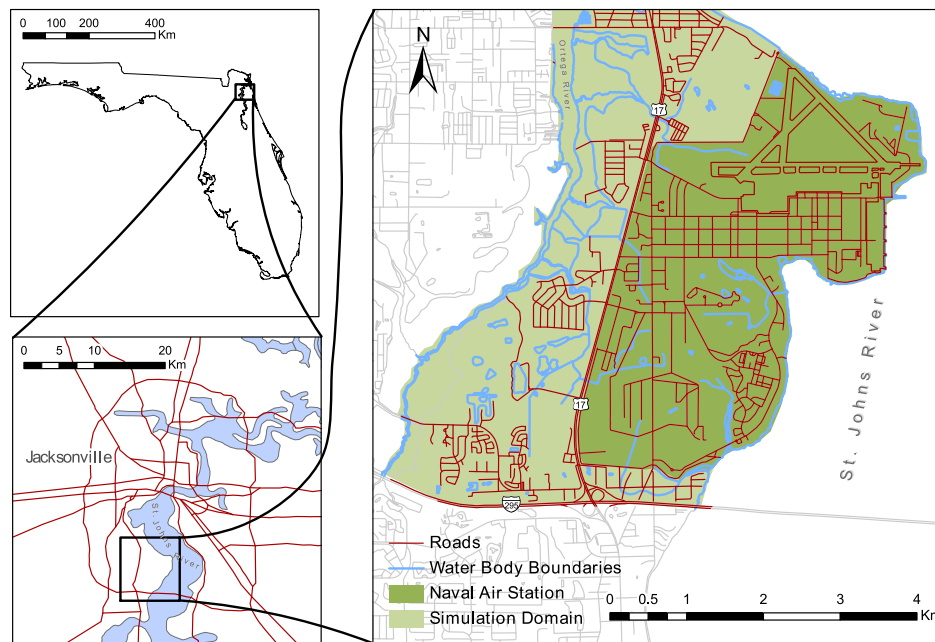
**Figure 2.6:** The U.S. Naval Air Station, Jacksonville Florida situated in Duval county

depth to the water table from the land surface at the NAS was measured to be between 0 and 1.5 m (0 and 5 ft) (Davis et al., 1996).

### 2.4.2   U.S. Naval Air Station, Jacksonville

Davis et al. (1996) of the USGS constructed and calibrated a detailed groundwater flow model of the NAS with the intent of determining the fate of contaminants at the base and evaluating the effects of remediation efforts on groundwater flow patterns. Because the model was calibrated with measurements of the hydraulic head in the surficial aquifer, this site is ideal for the study of the relationship between the shape of the water table and the overlying topography in the study area of Jacksonville, Florida. Additionally, because this model is a single layer model (essentially a 2D model), it makes it ideal for comparison to our simplified model. Finally, the NAS is a good study site due to its proximity to Lakeshore.

**Model Construction**

A numerical model based on the one by Davis et al. (1996) was set up as a steady-state, single layer model having a homogeneous hydraulic conductivity and recharge rate and was solved using MODFLOW (Harbaugh et al., 2000). The finite-difference cell length and width were both set to 30 m (100 ft). The bottom of the each cell extended to the top of the Hawthorn Group, whose elevation determined the cell thickness. The simulation domain is shown in Fig. 2.6 (indicated
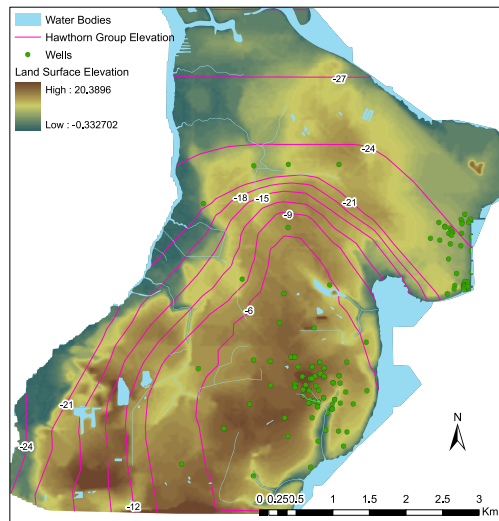
**Figure 2.7:** The locations of calibration wells in the USGS study along with a contour map of the elevation of the top of the Hawthorn group are shown on a USGS DEM with 10 by 10 m resolution. All elevation values are in meters.

by the areas shaded green) and is bounded to the north, east and west by the Ortega River and the St. Johns River. The stage for both of these rivers was set to sea level. Because the general flow direction in the area is approximately parallel to I-295, the southern boundary was set to a no-flow boundary. The bottom of the aquifer was also modeled as a no-flow boundary. The NAS is also drained by several creeks and ditches which were modeled as drains using the MODFLOW drain package. The elevation of the creek stages were assigned to be the same as the elevation of the creek bottoms, which were estimated from the land elevation in that cell. This removed the ability of the creeks to leak water to the aquifer but retained the capability of draining water from it. The model was calibrated using 128 well measurements (locations shown in Fig. 2.7) and 14 streamflow measurements. The calibrated hydraulic conductivity was 2.13 m/day (7 ft/day). The head measurements along with their (calibrated) simulated values are shown in Fig. 2.8. As a result of calibration, recharge was adjusted to the zones and values shown in Fig. 2.9. After calibration, the simulated head in 121 of the 128 wells was within the calibration criterion of 2.5 ft (0.762 m) of the measured value.

**Results and Discussion**

In order to evaluate the relationship between the shape of the water table and the overlying topography, the elevation of the water table was compared to a 1/3 arcsecond (approximately 10 m) resolution DEM obtained from the USGS National Map Seamless Server `http://seamless.usgs.gov`. Examining cross-sections of the topography and the water table from the USGS model suggests a strong relationship between the water table and topography although the relationship is stronger in certain areas and weaker in others. Fig. 2.10 shows two representative cross sections.
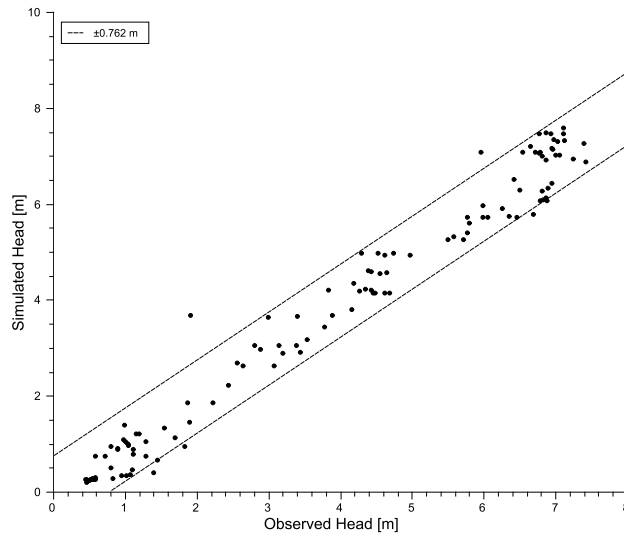
**Figure 2.8:** Simulated vs. observed head for the USGS model in the calibration wells of Figure Fig. 2.7. 121 out of the 128 well measurements fall within the 0.762 m (2.5 ft) calibration criterion (dashed lines)
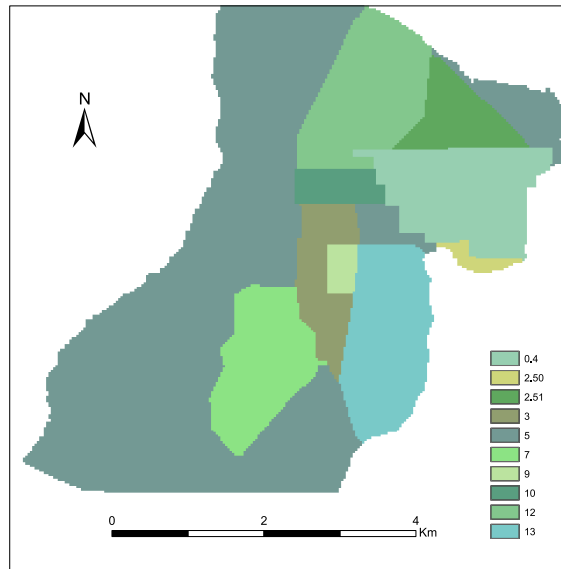


**Figure 2.9:** Map of recharge values for the USGS model. All values are in in/yr

The upper cross-section shows a case where the shape of the water table follows the topography only very generally. In the lower cross-section, the relationship is much closer. A more quantitative comparison can be obtained by plotting the land surface elevation against the water table elevation. If there exists a relationship, it can be quantified via the correlation coefficient. The scatter plot in Fig. 2.11 was created from every 1000th elevation value of the DEM (the water table data was re-sampled to the resolution of the DEM) and it clearly shows a linear relationship between the
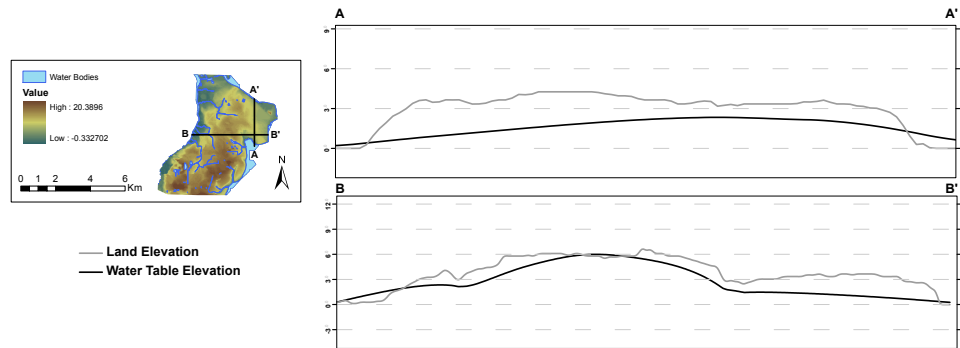
**Figure 2.10:** Selected cross-sections comparing the topography with the water table from the USGS study illustrating the fact that the water table can follow the topography (lower plot) but it is not always so (upper plot). Elevations are in meters.
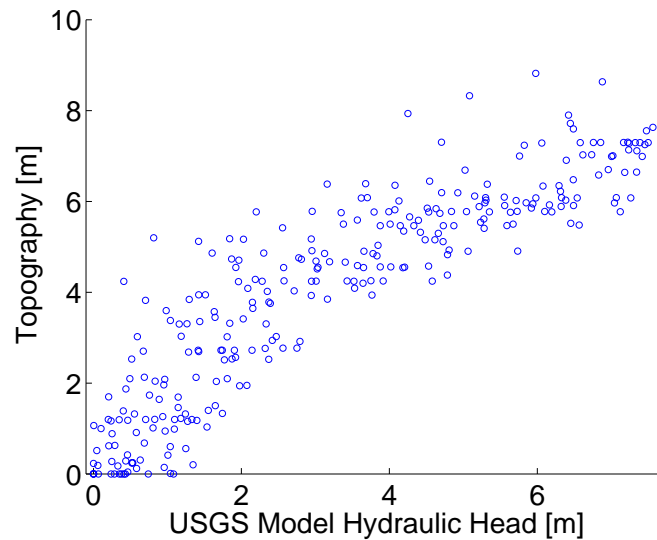


**Figure 2.11:** Correlation between the water table and the topography. The correlation coefficient is 0.87 (Pearson) or 0.91 (Spearman)

elevation of the water table and the topography. Pearsons's correlation coefficient was found to be 0.87 and Spearman's rank correlation was calculated to be 0.91. The scatter of the points in Fig. 2.11 shows that even though the relationship is strong, it is not a perfect correlation. This is to be expected since the water table is not equal to the topography everywhere and is consistent with the result presented in Fig. 2.10, showing that the relationship is stronger in certain areas and weaker in others. Although the elevation of the water table and the topography shows a good correlation, it does not immediately follow that the *slope* of the water table and the *slope* of the topography will be correlated to the same degree. After calculating the derivatives in the x- and y-direction (using the Sobel filter method described in Section 2.3.2), the correlation of the gradient for both cases is less

than 0.5 (Pearson and Spearman). This lack of correlation can generally be expected because the topography usually varies more significantly in space than a water table generated by a numerical model. This assertion can be demonstrated by applying the smoothing step in the simplified model. After smoothing the topography (with 50 smoothing iterations as described in Section 2.3.2), the correlation for the slope in the x-direction increases to 0.70 (Pearson) and 0.75 (Spearman). For the correlation in the y-direction, the values are 0.3 (Pearson) and 0.60 (Spearman). The correlation for the y-direction is much lower in this case, indicating that in this direction, the relationship between the water table and topography is weak. This weaker correlation in the y-direction is represented in the A-A' cross-section of Fig. 2.10.

From the results presented, it is now possible to say with a fair degree of certainty that overall, the water table is a subdued replica of the topography in the vicinity of the NAS although the degree to which the water table and topography are related may vary as shown in Fig. 2.10. The slope of the smoothed topography and the slope of the water table are generally related however the relationship is greater for the x-direction. Because of the close proximity of the NAS to the Lakeshore neighborhood, it stands to reason that the physiological and hydrological conditions will be similar, implying that the water table will also follow the topography in Lakeshore to the same degree. This assertion is examined in the following section.

### 2.4.3   Lakeshore Neighborhood

Because Lakeshore is one of the neighborhoods of interest, it is useful to determine to what degree the water table follows the topography in this area. From the previous section, the conclusion can be drawn that it is likely the water table will follow the topography to the same degree as in the NAS. To further analyze this, a MODFLOW 2005 (Harbaugh, 2005) model was constructed with a similar configuration e.g., single-layer and steady-state, as the model of Davis et al. (1996). Unlike the analysis of NAS where the model had already been constructed by Davis et al. (1996), the model in this section was constructed from scratch. Additionally, because no calibration data from wells are available, the model was not constructed to the same level of detail as the NAS. The MODFLOW GUI package ModelMuse v2.1.1.0 (Winston, 2009) was used to facilitate the set-up of the model and prepare the input files for MODFLOW.

**Model Construction**

The model was set up for the Lakeshore neighborhood by making use of the National Elevation Dataset (NED) DEM obtained from the seamless data server (`http://seamless.usgs.gov/`). The elevation information was used to set the elvation of the top of each simulation cell in MODFLOW. The dataset is the NED 1/3 arcsecond raster dataset which has a horizontal resolution of 1/3 arcseconds (approximately 10 meters). While a more detailed DEM based on lidar data and having a cell size of 5ft x 5ft and a vertical resolution of 0.6ft was available, this data was too detailed for calculation and comparison purposes. The effects of DEM resolution on the calculations of

topographic parameters (like the TOPMODEL topographic index) have been studied previously (Thompson and Moore, 1997; Thompson et al., 2001; Sørensen and Seibert, 2007) and some authors suggest that for approximating groundwater flow, a coarser DEM might better suited since the water table does not usually follow the topography exactly (Wolock and Price, 1994).

The DEM was then imported into ArcGIS 9.3 and re-sampled (using bi-linear interpolation) so that the size of each cell was exactly 10 x 10 meters. A 10 m cell size was chosen because it closely represents the original DEM resolution of 1/3 arcseconds. The total size of the domain based on the cell size was 371 columns and 373 rows. Based on this information, the MODFLOW grid size and spacing was set to 371 columns and 373 rows with spacing of 10 m between each row and column so that each MODFLOW simulation cell corresponded to exactly one cell in the DEM.

Data from the National Hydrography Dataset (NHD) obtained from the Florida Department of Environmental Protection (FDEP) was used to determine the geographical locations of rivers and other water bodies. In the NHD dataset (a vector dataset) water bodies are represented by lines and polygons, the extent of which give the location of the water body. Knowing the locations of water bodies is important to be able to set the constant-head boundary conditions. The NHD data was then imported into ArcGIS for further processing.

After importing the DEM and the water body dataset into ArcGIS, both datasets were projected to a common projection so that they could be exported without problems to ModelMuse. Without re-projecting, importing the data into ModelMuse caused the DEM and the water body dataset to not line up properly. The DEM was then exported into ASCII format and imported into ModelMuse by making use of the Import Gridded Data function. The DEM was used to set the elevation of the top of each cell as mentioned previously. The elevation of the bottom of each cell was set to the top of the Hawthorn Group, similar to the NAS model. The top of the Hawthorn Group was estimated to be -15 meters (-50 ft) relative to sea level, corresponding to approximately the mid point of the range quoted by Davis et al. (1996). The locations of the water bodies were imported into ModelMuse by making use of the Import Shapefile feature.

As mentioned previously, the Lakeshore neighborhood is bounded to the east and west by the Cedar and Ortega rivers respectively and to the south by the confluence of the Cedar and Ortega rivers. To the north, it is partially bounded by Big Fishweir Creek. The locations of these rivers, for simulation purposes, were determined by making use of the NHD dataset previously mentioned. The boundary condition for the rivers was set to a constant head boundary and the value was set to the river stage. For the portion of the simulation domain not bounded by any rivers or creeks (the north-west corner), the boundary was set to a no flow boundary. Care was taken when analyzing results, to avoid analyzing results in close proximity to this corner because of the artificial boundary there. Finally, all cells outside of the neighborhood of interest were set to inactive. Fig. 2.12 shows the model simulation domain. Because data for calibration was not available, the values for hydraulic conductivity and recharge from Davis et al. (1996) were used as baseline values. The baseline values of 10 in/yr for recharge and 7.5 ft/day for hydraulic conductivity were assigned to all cells in the simulation domain. Because the baseline values caused
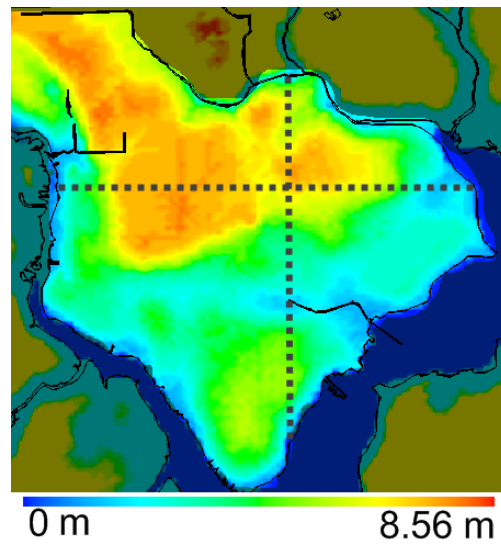
**Figure 2.12:** Model as constructed within ModelMuse. The colormap represents the elevation of the top of each cell. The highlighted area is the simulation area. The dotted lines indicate the selected cross sections of the water table used for determining recharge and conductivity parameters.

the elevation of the water table to exceed the land surface elevation, sometimes by a significant amount (up to 3m), in order to obtain results more tailored to this location, a series of runs were performed and values of hydraulic conductivity and recharge were selected so that the water table did not exceed the land surface elevation. The parameters tested are shown in Table 2.1. Because the aquifer geometry remained fixed in all runs, it is the ratio of recharge to conductivity, R/K, that controlled the shape of the water table.

**Table 2.1:** Parameter values used in the simulation. Baseline values are in italics.

| # | R | K | R/K |
|---|---|---|---|
| 1 | *6.954E-04 m/d (10 in/yr)* | *2.286 m/d (7.5 ft/day)* | *3.042E-04* |
| 2 | 5.563E-04 m/d (8 in/yr) | 2.286 m/d (7.5 ft/day) | 2.433E-04 |
| 3 | 5.563E-04 m/d (8 in/yr) | 3.048 m/d (10 ft/day) | 1.825E-04 |
| 4 | 3.477E-04 m/d (5 in/yr) | 3.048 m/d (10 ft/day) | 1.141E-04 |
| 5 | 3.477E-04 m/d (5 in/yr) | 4.267 m/d (14 ft/day) | 8.149E-05 |
| 6 | 2.086E-04 m/d (3 in/yr) | 4.267 m/d (14 ft/day) | 4.889E-05 |

In order to select parameters that a) do not cause the water table to exceed the land surface elevation, b) give water table depths in the ranges given by Davis et al. (1996) and c), are reasonably close in value to the baseline values, two representative cross sections are selected: one along a row (138) and another along a column (208). The specific row and column are chosen because they represent areas of the domain having markedly different topographical profiles. For the section
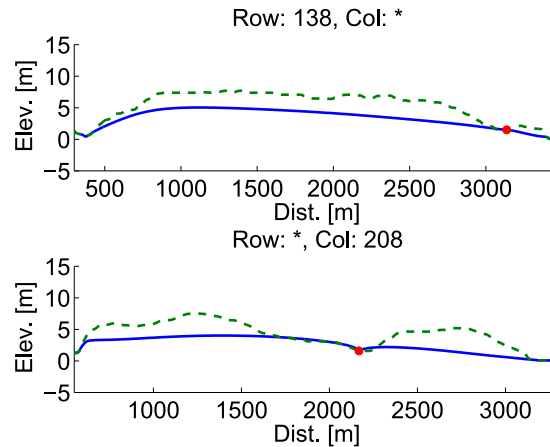
**Figure 2.13:** Water table and topography cross sections for the selected recharge and conductivity parameters (R=5 in/yr, K=10 ft/day). Dashed line is the topography, solid line, the water table. The dot represents the location where the water table intersects a small stream. Upper figure: east-west cross section. Lower figure: north-south cross section

along the row (call it the east-west cross section), the topography varies quite smoothly and is in the form of a low hill. The section along the column (north-south cross section) is more varied, meaning it has several bumps and valleys. The location of these sections is indicated in Fig. 2.12 by the dotted lines.

**Results and Discussion**

The model was executed for the values of recharge and conductivity given in Table 2.1. The recharge and conductivity values that best satisfied the selection criteria are located in row #4 in Table 2.1. The resulting water table is compared to the overlying topography in Fig. 2.13. For the selected parameters, the average and maximum depth to the water table are calculated for both cross sections. For the east-west cross section, the average depth is 1.9 m and the maximum 3.5 m. For the north south cross section, the average is 1.5 m and the maximum 3.9 m. Note that the water table still exceeds the land surface by a small amount in the lower cross section, a situation which cannot be remedied without site specific calibration data.

The simulation shows that in the Lakeshore area, the water table follows the topography only generally, given the most plausible values of recharge and conductivity. The most plausible parameter values are the ones that give a plausible water table and result in a water table depth similar to the one reported by Davis et al. (1996) (0 to 1.5 m). From this simulation, it is likely that the water table only generally follows the topography. In order to more definitively answer this question, site specific calibration data is required.

### 2.4.4   Rule of Thumb

Up to this point, the evidence suggests that the water table in the vicinity of the NAS and the Lakeshore neighborhood is a subdued replica of the topography, to varying degrees. A final piece of evidence to support or disprove this hypothesis can be obtained from a rule of thumb developed by Haitjema and Mitchell-Bruker (2005) that tests whether a water table can be considered to be topography controlled. The rule of thumb is defined as

$$\phi = \frac{RL^2}{mKHd} \begin{cases} > 1 & \text{topography controlled} \\ < 1 & \text{recharge controlled} \end{cases} \tag{2.6}$$

where $R$ [m/day] is recharge, $L$ [m] is the average distance between surface waters, $m$ is a dimensionless factor accounting for the aquifer geometry, and is between 8 and 16 for aquifers that are strip-like or circular in shape, $K$ [m/day] is the hydraulic conductivity, $H$ [m] the average aquifer thickness, and $d$ [m] is the maximum distance between the average water level in surface water bodies and the elevation of the terrain.

Results from Section 2.4.2 indicate that the water table at the NAS should follow the topography to a reasonable level in most areas. The rule of thumb is now applied to the NAS assuming limited a priori knowledge. Recharge is taken to be the estimated average recharge for the area: 6.95E-4 m/day (10 in/yr or 255 mm/day) as mentioned in Section 2.4.1. The parameter $L$ is calculated by using ArcGIS to measure the distance between the two major rivers bounding the domain (the Ortega River and the St. Johns) by drawing cross-sections from the bank of one river to the bank of the other. This is done at several locations in the east-west and north-south directions and the average length, 4330 m, is used for $L$. The hydraulic conductivity $K$ is set to 1.52 m/day; the value obtained from slug tests mentioned in Section 2.4.1. $H$ is estimated by assuming the water table is 0.762 m below the land surface everywhere (0.762 m obtained by averaging the estimated range of depths given in Section 2.4.1). The elevation of the Hawthorn Group is then subtracted from this rough water table estimate and the average value of the result, 19 m, is taken as $H$. The parameter $d$ is estimated by using the same cross sections that were used in the calculation of $L$ and finding the maximum difference between the elevation along each line and the water level (assumed to be sea level) in the Ortega and St. Johns rivers. Once the maximum has been found for each cross section, the average value of the maximum values, 6.1 m, is used. The parameter $m$ is estimated to be 12 due to the semi-elongated shape of the study area. The rule of thumb gives

$$\phi = 6.2$$

indicating a water table that is controlled by the topography, though not strongly so. In their paper, Haitjema and Mitchell-Bruker (2005) gave an example that used the aquifer studied by Tóth (1963). After some calculations, a value of $\phi = 40$ was given for Tóth's aquifer, with a statement indicating it should be considered topography controlled. An example of a recharge controlled aquifer was also given. In that recharge controlled case, the value of $\phi$ was 0.0088. Therefore, for the case at

hand, a resulting value of $\phi = 6.2$ means that although the aquifer is topography controlled, it is not strongly so.

For the Lakeshore area, suitable recharge and conductivity values were selected in Section 2.4.3 for recharge, 3.477E-04 m/day (5 in/yr), and conductivity, 3.048 m/day (10ft/day). Although these values were estimated from the MODFLOW model, they can essentially be considered to be obtained without a priori knowledge since the MODFLOW model itself was constructed using only literature values for the model parameters. Using ArcGIS, the average distance between water bodies, $L$, is estimated to be approximately 2700 m. The value of $m$ is set to 13 since the geometry of the Lakeshore area is deemed to be more circular than strip-like. $H$ is estimated to be about 17 m, given that the base of the aquifer is set to 15 m below sea level and the depth to the water table is approximately 2 m. From Fig. 2.13, the parameter $d$ is estimated to be about 6 m. The resulting value of $\phi$ is

$$\phi = 0.63$$

This would indicate that the water table is in fact recharge controlled. However, because the value is close to 1, this is not a clean-cut conclusion. Taking into account the values given by Haitjema and Mitchell-Bruker (2005) for recharge and topography controlled water tables, for the case at hand, a resulting value of $\phi = 0.63$ means that the aquifer is not strongly topography controlled but neither strongly recharge controlled; it is somewhere in between. Therefore for Lakeshore, the rule of thumb indicates a water table that is not topography controlled however this result conflicts with the results of the nearby NAS. This conflict yields an inconclusive result and a topography controlled water table in Lakeshore cannot be ruled out.

### 2.4.5   Conclusion

The rule of thumb indicates that for the NAS, the water table is controlled by the topography although the degree to which it is controlled may be strong in some areas and weak in others. Overall, the water table is topography controlled, although not strongly so. For the case of Lakeshore, the rule of thumb yields an inconclusive result. Taking into account all the evidence (MODFLOW models of the NAS and Lakeshore as well as the rule of thumb results), it can be concluded that the water table at the NAS and Lakeshore areas follow the topography to a reasonable degree however the strength of the relationship varies. In order to reach this conclusion, a higher weight is given to the MODFLOW simulations, especially the one by Davis et al. (1996) since that result is obtained from a calibrated model. To conclude this analysis, a statement by Haitjema and Mitchell-Bruker (2005) regarding the general relationship between the water table and topography is presented:

> "...in a very general sense the water table and the topography are always related. After all, water table lows occur at surface water, which in turn occur in topographically low areas."

## 2.5   Analysis of Model Performance

After evaluating the validity of the approximations made by the simplified model in the area of interest, it is possible to analyze the performance of the simplified model. A quantity of interest is the travel time (and to a lesser extent the travel distance) of an imaginary particle placed at a given starting point to its final destination at a water body. In fact, this was also the purpose of the simplified GIS-based model of Schilling and Wolter (2007). As shown in Section 2.4.3, the values of recharge and hydraulic conductivity play important roles in determining the shape of the water table and therefore the hydraulic gradient and groundwater flow velocity. What remains to be seen is the effect that the smoothing factor has on the calculated water table and how it affects the groundwater travel times. In order to investigate this relationship, several values of the smoothing factor were selected based on preliminary tests to find values that yielded plausible water tables. From these preliminary tests, it was found that a smoothing value of 50 yielded an acceptable approximation to the water table, using the water table from the calibrated USGS MODFLOW model of Davis et al. (1996) (Section 2.4.2) as reference. Based on this number and the observation that the change in the shape of the calculated water table decreased as the value of the smoothing factor increased, eight representative values of the smoothing parameter were selected for testing:

$$2, 5, 10, 20, 40, 50, 60, 80$$

### 2.5.1   Comparison with NAS Model

In order to quantify how the value of the smoothing factor affects the agreement between the estimated travel times and travel distances compared to the USGS model, the model was run with the previously mentioned values of the smoothing parameter for both the entire domain and a small sub area located near the north-east corner of the NAS. For the case of the entire domain, points were uniformly placed 100 m apart and for the sub area, the points were uniformly placed 70 m apart. The model was then run using the 10 m DEM (from the USGS National Map) and particle tracking conducted for both areas using the previously given smoothing factor values, a uniform hydraulic conductivity of 2.113 m/day, and a porosity of 0.25. The values of the conductivity and porosity were the same as those used in the USGS model. Fig. 2.14 shows the flow paths generated by the model, for the case using a smoothing factor value of 50 for both the entire NAS domain and for the smaller sub area. Additionally, for comparison purposes, flow paths are generated for the sub area using data generated by the USGS MODFLOW model. In order to generate the flow paths for the USGS model, particle tracking was conducted (with the method of Section 2.3.3) using the velocity flow field obtained from the MODFLOW cell-by-cell flow output file.

In order to quantify the quality of the match between the estimated travel times and travel distances of our model compared with those of the USGS model, scatter plots of the estimated travel times and distances were constructed for both the entire domain and the smaller sub area using different values of the smoothing factor. In such a plot, an ideal result would be perfect
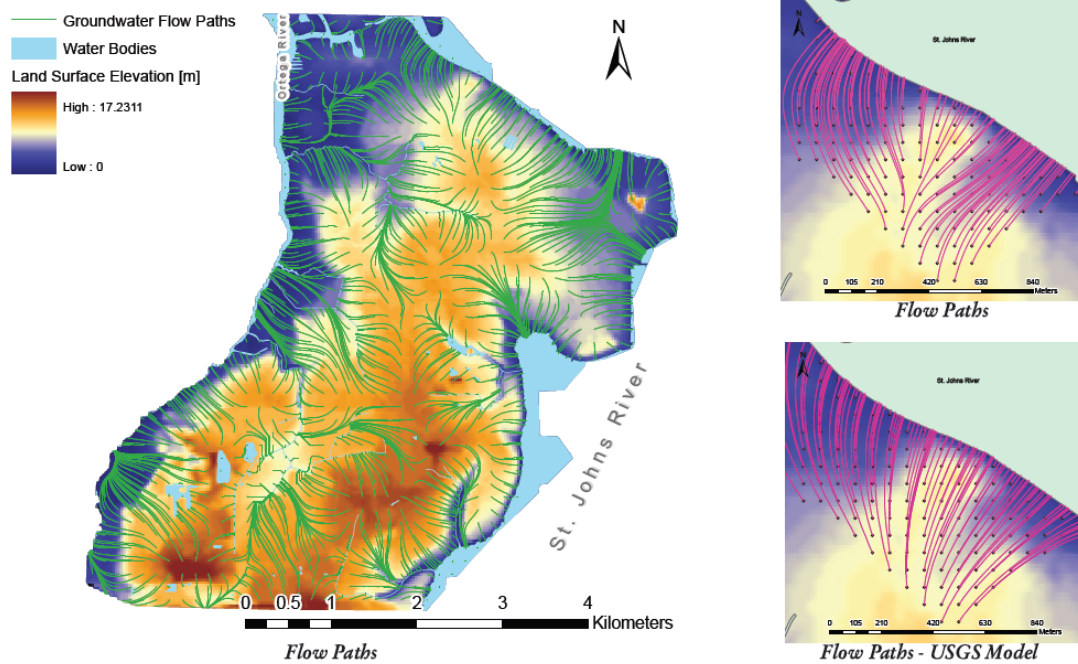
**Figure 2.14:** Results of running the model (with particle tracking), smoothing parameter of 50 is shown. Left: entire domain. Right: sub area located near the north east corner of the domain

positive Pearson correlation as well as having a best fit line of slope one and intercept zero. The Spearman and Pearson correlations were calculated for each plot to measure the agreement, the results of which are summarized for the domain in Fig. 2.15. Fig. 2.15 shows the relationship between the correlation of this simplified model and the USGS model plotted against the smoothing factor for the entire domain for both travel time and travel distance. Fig. 2.15(a) indicates that the optimum value for the smoothing factor for travel distance is around 40 (taking into account both Pearson and Spearman correlation). Fig. 2.15(b) indicates the optimum value should be 20, however 40 is the next best value. Fig. 2.16 shows the difference between the simplified model and the USGS model (simplified minus USGS) for both the travel distance and travel time for the entire domain. Like Fig. 2.15(a), Fig. 2.16(a) suggests the best value for the smoothing factor is 40. In Fig. 2.16(b), the minimum occurs at the smoothing factor value of 60, however 40 and 50 (tie) have the next smallest difference. Figs. 2.17 and 2.18 show the same information as Figs. 2.15 and 2.16 except corresponding to the sub area. For the sub area, all plots indicate the best value for the smoothing factor, out of the ones tested, is 80. Fig. 2.19 shows the number of flow paths (in percent) having an estimated travel time and distance that fall within a window of ±20 years of the travel time and ±20 m of the travel distance predicted by the USGS model for both the entire domain and the sub area. Fig. 2.19 suggests the best value of the smoothing factor for the entire NAS is approximately 40 while for the sub area, a value of 80 is suggested (considering both travel time and distance).

In order to ensure that the data are correctly compared, flow paths that ended up in different
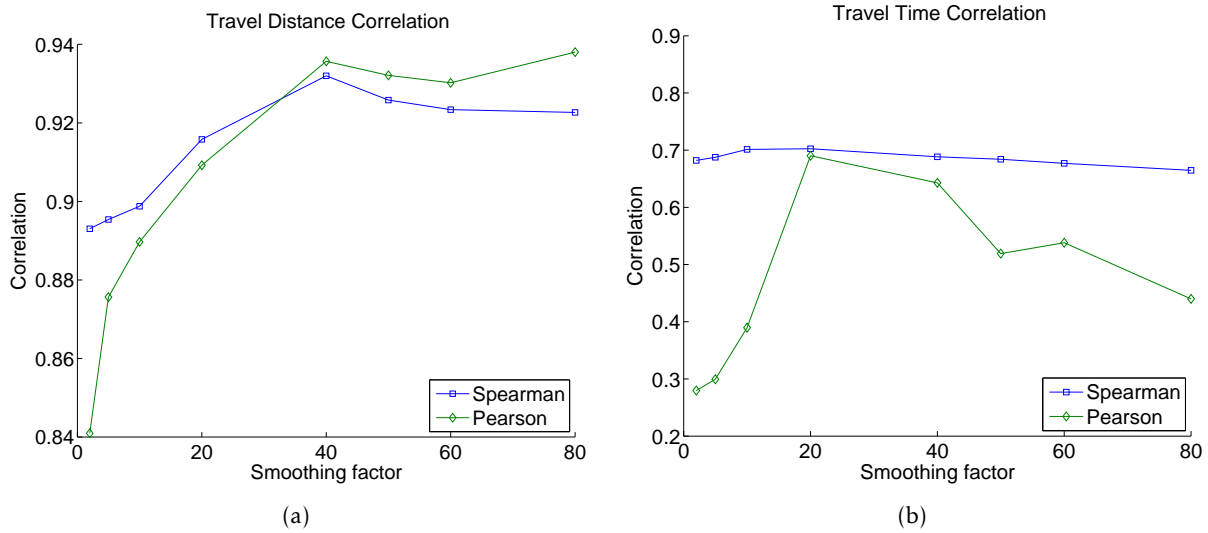
**Figure 2.15:** Correlation of estimated travel distance and travel time between the simplified model and the USGS model. Estimation area: entire domain
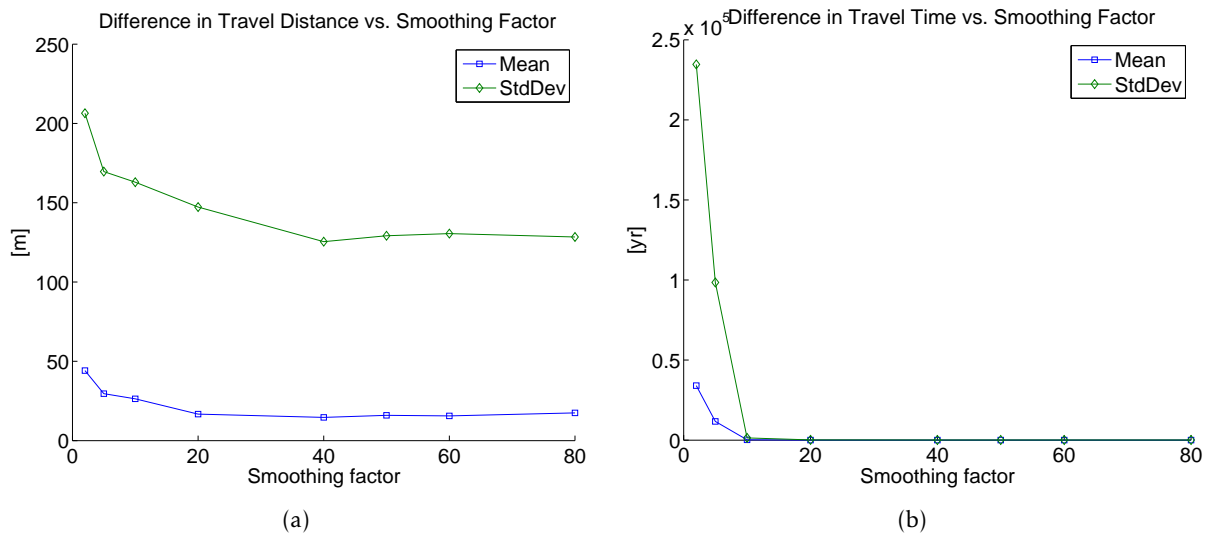


**Figure 2.16:** Difference between the travel distance and travel time compared with USGS model plotted against smoothing factor. Estimation area: entire domain.
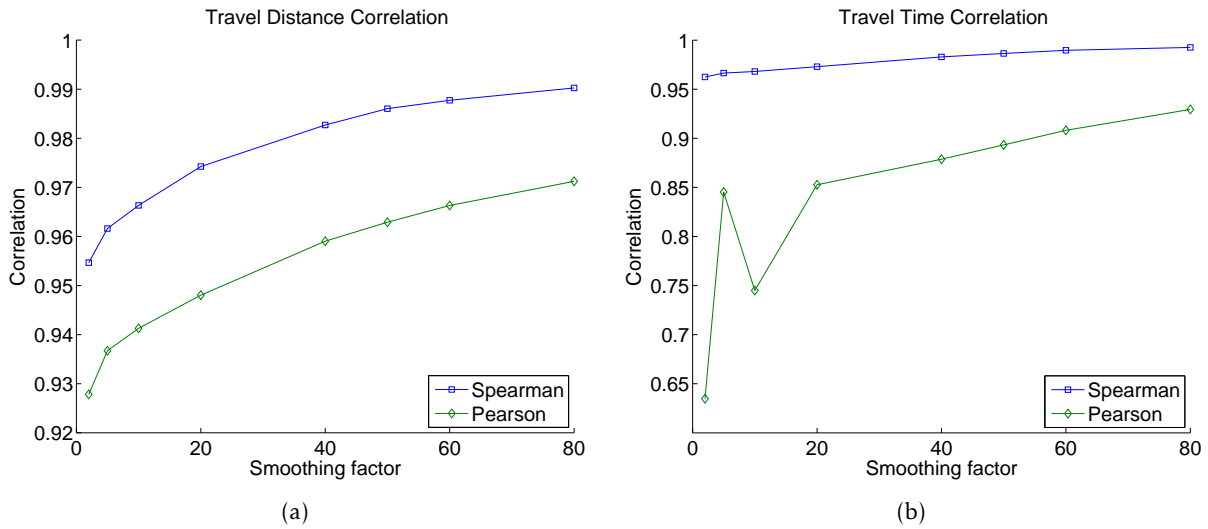
**Figure 2.17:** Correlation of estimated travel distance and travel time between of the model compared with the USGS model. Estimation area: sub area
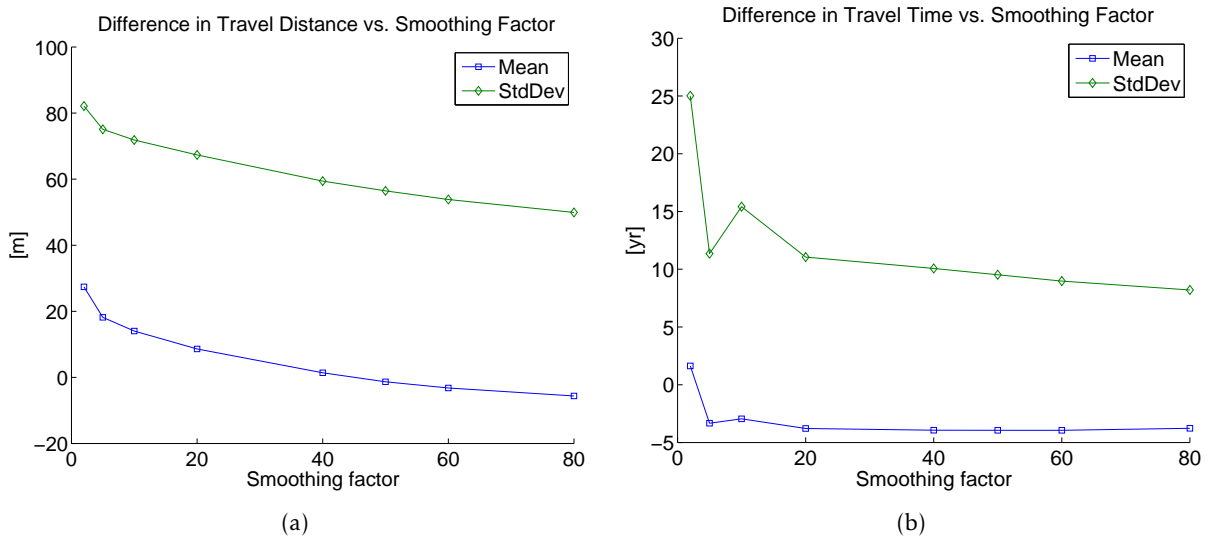


**Figure 2.18:** Difference between the travel distance and travel time compared with USGS model plotted against smoothing factor. Estimation area: sub area.

(a) Entire domain

(b) Sub area
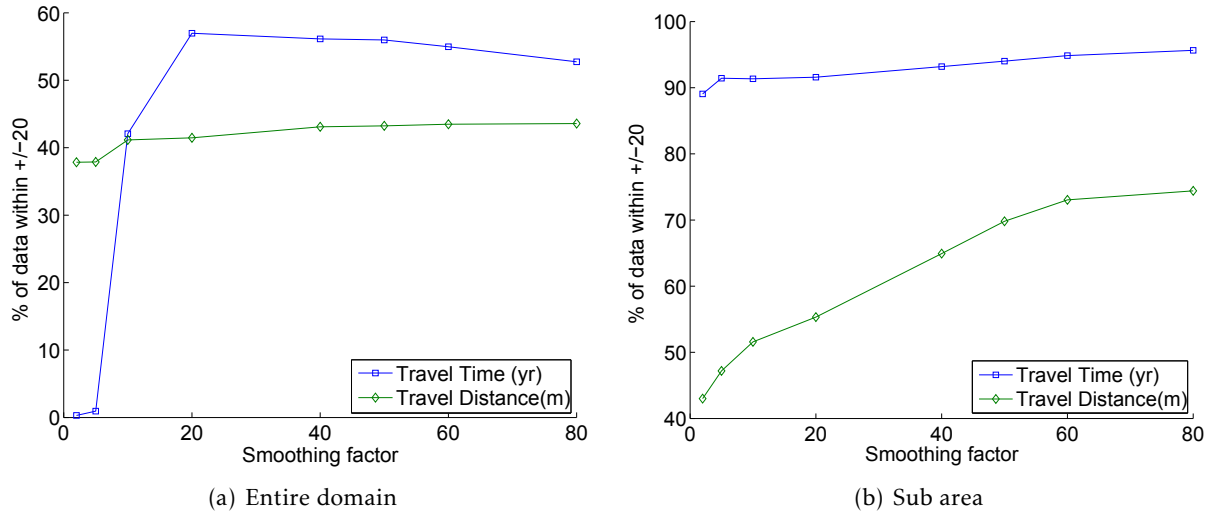
**Figure 2.19:** The amount of data within ±20 years and ±20 m of the USGS model.

water bodies were discarded. Additionally, flow paths were discarded when the travel time in the USGS model was greater than 5000 years. This occurred when there was a sink in the water table in the USGS model. When a sink is present flow becomes trapped and never leaves. Because the particle tracking routine terminates after a set number of iterations (usually a high number), flow into a sink results in unrealistic travel times for those cases. For the case of the entire domain, out of the 1175 flow paths available, on average, approximately 39 % were discarded for each run. In the sub area, out of the 129 available flow paths none were discarded.

It is important to note that although the two correlation coefficients (Spearman and Pearson) are calculated, it is the Pearson's correlation that is most useful for examination since Pearson's correlation coefficient measures the linearity of the relationship while the rank correlation will also measure any non-linear effects. The ideal relationship between the values estimated by our model and those estimated by the USGS model is linear.

### 2.5.2  Discussion

The only adjustable parameter in this simplified model is the smoothing factor, which indicates to what degree the topography will be smoothed in order to generate the shape of the water table (the physical aquifer parameters were fixed to the same values as the USGS model). The results indicate that there is an optimum value of the smoothing parameter for both the entire NAS and the sub area, however the value itself is different in each case. Additionally, the optimum value varies depending on the desired quantity to estimate (travel time or travel distance). For estimating the travel time, the optimum value of the smoothing factor for the entire NAS was approximately 40, determined by simultaneously considering Figs. 2.15, 2.16 and 2.19(a). For this value of the smoothing factor, statistics of interest are extracted and shown in Table 2.2. The correlation statistic

**Table 2.2:** Statistics of travel distance and time for the entire domain for smoothing factor 40

|  | Correlation | Percent of Data within +/- 20 | Mean difference | StdDev of Difference |
|---|---|---|---|---|
| Travel Time | Spearman: 0.68 Pearson: 0.64 | 56 | 16.4 yr | 77.2 yr |
| Travel Distance | Spearman: 0.93 Pearson: 0.93 | 42 | 14.6 m | 125 m |

**Table 2.3:** Statistics of travel distance and time for the sub area for smoothing factor 80

|  | Correlation | Percent of Data within +/- 20 | Mean difference | StdDev of Difference |
|---|---|---|---|---|
| Travel Time | Spearman: 0.99 Pearson: 0.92 | 92 | -3.8 yr | 8.2 yr |
| Travel Distance | Spearman: 0.99 Pearson: 0.97 | 73 | -5.6 m | 49 m |

represents the correlation between the value estimated by this model and the value estimated from using the data from the model by Davis et al. (1996). The "percent of data within ±20" column represents the number of model estimates that fall within a window of ±20 (years or meters depending on whether the comparison is between travel time or travel distance). The remaining two columns are self explanatory. A similar analysis can be carried out for the sub area (Figs. 2.17, 2.18 and 2.19(b)). For the sub area, the optimum value of the smoothing parameter out of the ones tested was 80 and the corresponding statistics are shown in Table 2.3. For comparison purposes, the statistics are also shown for a smoothing value of 40 in Table 2.4 From Tables 2.2 and 2.3, it is immediately apparent that a single value of the smoothing parameter cannot capture the varying degrees to which the water table follows the topography for a domain as large as the entire NAS. For a smaller sub domain however, the model is capable of yielding acceptable results.

**Table 2.4:** Statistics of travel distance and time for the sub area for smoothing factor 40. Compare with Table 2.2

|  | Correlation | Percent of Data within +/- 20 | Mean difference | StdDev of Difference |
|---|---|---|---|---|
| Travel Time | Spearman: 0.98 Pearson: 0.86 | 91 | -4.7 yr | 10 yr |
| Travel Distance | Spearman: 0.98 Pearson: 0.95 | 65 | 5 m | 61 m |
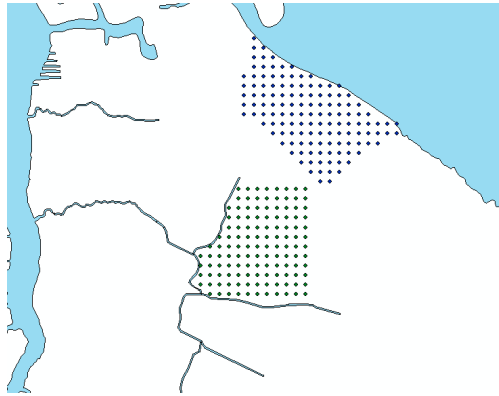
**Figure 2.20:** Locations of the sub areas within the NAS. Each point represents a contaminant source.

From Fig. 2.19 and Tables 2.2 to 2.4, it is clear that the effect of the smoothing factor on the estimates of travel time and distance is location- and scale-dependent. When the model is run for the entire domain, there is clearly a value of the smoothing parameter that optimizes the model agreement with the USGS model (the value is around 40). For the case when the model was applied to the sub area, the behavior is markedly different. The tested values of the smoothing factor suggested that the optimum value is 80 or above. From this, it is clear that a single value of the smoothing parameter is not adequate to model an area as large as the entire NAS. It is more appropriate to define a parameter for smaller areas such as the sub area tested here. The reason for this is simple: the analysis between the water table and the topography done in Section 2.4.4 suggests that although in general, the water table in the vicinity of the NAS can be considered a subdued replica of the topography, the relationship varies with space, i.e., in some areas the relationship may be stronger than others and a smaller smoothing factor may be needed.

Because of the scale and location dependent nature of the optimal smoothing parameter, in order to better evaluate the effectiveness of the model, more data points were required. A second sub area was selected, immediately to the south of the first sub area, having comparable dimensions and a comparable number of sources to the first. Fig. 2.20 shows the location of both sub areas. Due to the increased area of the combined sub areas (compared to a single sub area), the smoothing factor was set to 50, a compromise between the factors determined for the entire domain and the first sub area in the previous section. The travel time analysis from the previous section was repeated for the combined sub areas. The estimates for travel times are compared to the results from the USGS model of Section 2.4.2 and the results are summarized by calculating the ±20 year travel time statistic from Table 2.3. The summarized results (both as a probability density function and a cumulative density) are shown in Fig. 2.21. The plot shows that the model yields an acceptable performance by estimating the travel time to within ±20 years of the Davis et al. (1996) 75% of the time, with a mean difference of 13 years. Additionally, the results from Fig. 2.21 can be used to estimate the probability that any given point will fall within ±20 years of the "true" value (the value from Davis et al. (1996) in this case). To facilitate this estimate, the CDF and PDF were fitted using a kernel smoothing density estimate, shown as the "Fit" line in the plots.
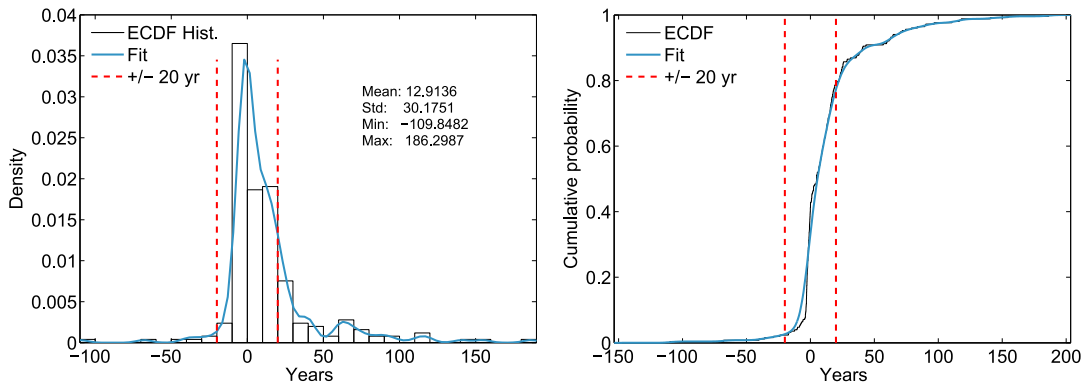
**Figure 2.21:** The number of points falling within a window of ±20 years of the values estimated from the model by Davis et al. (1996). The left plot shows the data as a PDF while the right plot shows it as a CDF
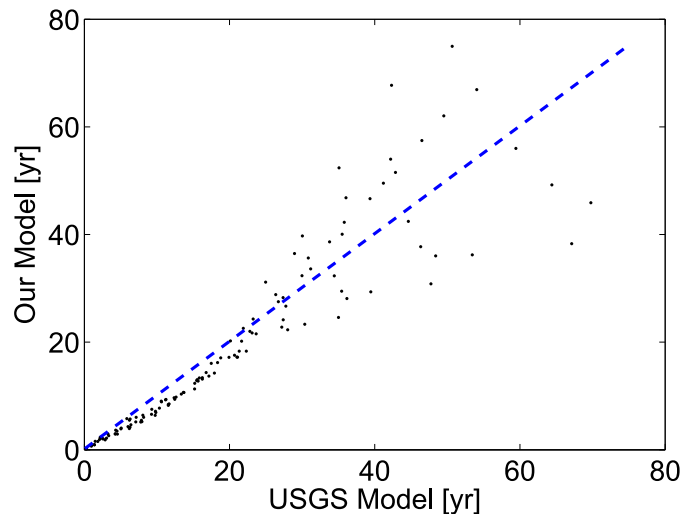


**Figure 2.22:** The correlation of travel time estimates between the model estimated times and the times from the USGS model of Davis et al. (1996).

A final analysis that can be used to measure the effectiveness of the model is a plot showing how well the model travel time estimates correlate with the travel times calculated from the USGS model. Such a plot is shown in Fig. 2.22 for the sub area; the dashed line represents the 1:1 line. Each data point represents the travel time from an individual source. From the figure, it can be seen that for low travel times, the correlation is excellent, indicating that the simplified model predicts the groundwater travel time quite well compared to the more sophisticated MODFLOW model. As time increases, the points become more scattered. The points with a higher travel time correspond to sources that are farther away from the water body. This indicates that longer flow paths have a longer travel time as expected. More importantly, it reveals that as the travel

distance increases, the simplified model results become less and less reliable compared to the USGS model. The reason behind this is the accumulation of error along the flow path. This error does not arise from numerical issues, rather it comes from the approximations made by the flow model (mainly from the approximation that the water table is a subdued replica of the topography). From the figure, it can also be seen that the points begin to scatter significantly at approximately 30 years, which corresponds to a distance of approximately 500 m. It can therefore be said that for this site, travel time estimates from this model can be considered acceptable up to a distance of approximately 500 m.

As a final note, from Tables 2.2 and 2.3 (2nd column) it may initially appear that the model provides a better estimate for the travel time rather than the travel distance because a greater percentage of data that falls within the specified range of 20 for both the entire NAS and the sub area. This is not the case however because even though it is true that the percentage is greater for the travel time compared to the travel distance, the ranges themselves are not comparable. That is, for the travel time estimate, the range represents +/- 20 years and the value represents the amount of data points that are +/- 20 years from the USGS model while for the travel distance estimate, the value represents the percentage of points within +/- 20 meters from the USGS model. By examining the correlation values however, it is clear that the model provides a better estimate of the travel distance rather than the travel time.

*3*

<div style="background:#d9d9d9">

**Transport Module**

</div>

In this chapter, the development and implementation of the Transport module is discussed in detail. First, introductory material is presented giving a discussion of the motivation behind using the Domenico analytical solution to solve the advection-dispersion equation as well as a review of the solution itself. Subsequently, the implementation of the transport model within ArcGIS will be discussed, followed by an analysis of model performance using a combination of synthetic test cases and real world data.

## 3.1 Background

Modeling the fate and transport of contaminants normally involves the solution of the governing equation of contaminant transport, the advection-dispersion equation:

$$\frac{\partial\left(\theta C^k\right)}{\partial t} = \frac{\partial}{\partial x_i}\left(\theta D_{ij}\frac{\partial C^k}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(\theta v_i C^k\right) + q_s C_s^k + \sum R_n \tag{3.1}$$

where $\theta$ [-] is the medium's porosity, $C^k$ [ML$^{-3}$] is the dissolved concentration of species $k$, $t$ is time, $x_{ij}$ [L] is the distance along the respective Cartesian coordinate, $D_{ij}$ [L$^2$T$^{-1}$] is the hydrodynamic dispersion tensor, $v_i$ [LT$^{-1}$] is the seepage velocity, $q_s$ [T$^{-1}$] is the flow rate per unit volume of sources and sinks, $C_s^k$ [ML$^{-1}$] is the concentration of the source or sink for species $k$, and $\sum R_n$ is the chemical reaction term (Zheng and Wang, 1999). Depending on the types of chemical interactions considered, the chemical reaction term can take several forms.

Several methods have been developed to solve Eq. (3.1) including analytical, semi-analytical, and numerical methods. In the context of numerical modeling, several different methods have been developed for different circumstances although no single method can solve Eq. (3.4) adequately (Zheng and Wang, 1999). These methods can be divided into Eulerian, Lagrangian and Eulerian-Lagrangian schemes. Eulerian methods can be thought of as characterizing the movement

of contaminant by observing a point at a fixed location in the domain and by measuring individual contaminant particles as they pass. Lagrangian schemes can be thought of characterizing the movement by following the individual particles as they move through the domain. Eulerian schemes include traditional explicit and implicit finite-difference schemes, finite element schemes, and finite volume methods. Lagrangian methods are composed of particle tracking and random walk methods. Examples of mixed Eulerian-Lagrangian methods include the method of characteristics (MOC) and hybrid method of characteristics (HMOC). In these methods, the advection term of Eq. (3.1) is solved via particle tracking while the dispersion and reaction terms are solved via Eulerian methods.

When a full numerical solution is not feasible and/or required, as is the case for the GIS-based model presented in this work, an analytical solution to Eq. (3.1) is highly desirable because it allows for the relatively straightforward calculation of a solution. Unfortunately, because Eq. (3.1) is complex, simplifications are needed so as to make it solvable via analytic means. These simplifications can be made in various ways, including assuming advection along only one coordinate and assuming simplified boundary and initial conditions. Without reducing the complexity of the equation, the problem is simply not tractable. Unfortunately, because deriving an analytical solution requires the specification of initial and boundary conditions, the resulting solution is only applicable to systems satisfying those conditions. In the literature, many analytical and semi-analytical solutions to Eq. (3.1) have been derived under various assumptions and boundary conditions. Exploring these alternate analytical solutions is given as a suggestion for future work.

### 3.1.1   The Domenico Solution

The Domenico solution with decay (Domenico, 1987) considers a simplified form of Eq. (3.1)

$$\frac{\partial C}{\partial t} = D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + D_z \frac{\partial^2 C}{\partial z^2} - v \frac{\partial C}{\partial x} - kC \tag{3.2}$$

where $C$ is the contaminant concentration $[M/L^3]$, $D_x$, $D_y$ and $D_z$ are the homogeneous dispersion coefficients in $x$, $y$ and $z$ respectively $[L^2 T^{-1}]$ $v$ is the constant seepage velocity in the $x$ direction
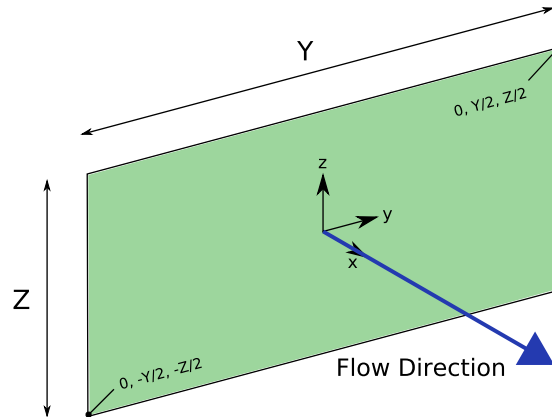
**Figure 3.1:** The geometry of the Domenico solution source plane.

$[L]$ and $k$ is the first order decay constant $[T^{-1}]$. The boundary and initial conditions are

$$C(x,y,z,0) = 0 \qquad \forall\; 0 < x < \infty, \quad -\infty < y < \infty, \quad -\infty < z < \infty$$

$$C(0,y,z,t) = C_0 \qquad -\frac{Y}{2} < y < \frac{Y}{2}, \quad \frac{Z}{2} < z < \frac{Z}{2} \qquad \forall\; t > 0$$

$$\lim_{x \to +\infty} \frac{\partial C(x,y,z,t)}{\partial x} = 0 \tag{3.3}$$

$$\lim_{y \to \pm\infty} \frac{\partial C(x,y,z,t)}{\partial y} = 0$$

$$\lim_{z \to \pm\infty} \frac{\partial C(x,y,z,t)}{\partial z} = 0$$

These conditions essentially correspond to considering a single plume, having a source plane centered at $(0,0,0)$, with dimensions $Y$ and $Z$ and a constant concentration $C_0$ while only considering groundwater flow in the $x$ direction but dispersion in all three directions (Fig. 3.1). Additional constraints assume that the system evolves only in the positive half of the x coordinate space and that the system is initially free of contaminant.

The general form of the Domenico solution used in this model is the three-dimensional transient solution of Martin-Hayden and Robbins (1997), which is based on the work by Domenico and Robbins (1985) and Domenico (1987),

$$C(x,y,z,t) = \frac{C_0}{8} F_1(x,t) F_2(y,x) F_3(z,x) \tag{3.4}$$

with

$$F_1 = exp\left[\frac{x}{2\alpha_x}\left(1 - \sqrt{1 + \frac{4k\alpha_x}{v}}\right)\right] \times erfc\left[\frac{x - vt\sqrt{1 + \frac{4k\alpha_x}{v}}}{2\sqrt{\alpha_x vt}}\right] +$$

$$exp\left[\frac{x}{2\alpha_x}\left(1 + \sqrt{1 + \frac{4k\alpha_x}{v}}\right)\right] \times erfc\left[\frac{x + vt\sqrt{1 + \frac{4k\alpha_x}{v}}}{2\sqrt{\alpha_x vt}}\right] \tag{3.5a}$$

$$F_2 = erf\left(\frac{y + Y/2}{2\sqrt{\alpha_y x}}\right) - erf\left(\frac{y - Y/2}{2\sqrt{\alpha_y x}}\right) \tag{3.5b}$$

$$F_3 = erf\left(\frac{z + Z/2}{2\sqrt{\alpha_z x}}\right) - erf\left(\frac{z - Z/2}{2\sqrt{\alpha_z x}}\right) \tag{3.5c}$$

where $\alpha_x$, $\alpha_y$ and $\alpha_z$ are the longitudinal, horizontal transverse and vertical transverse dispersivities $[L]$, $k$ is the first order decay coefficient $[T^{-1}]$, $v$ is the groundwater seepage velocity in the longitudinal direction $[LT^{-1}]$, $Y$ and $Z$ are the width and height of the source plane respectively $[L]$ and $t$ is time $[T]$.

The actual form of the Domenico solution used in this model is the steady-state, two-dimensional version of Eq. (3.4) as follows

$$C(x, y) = \frac{C_0}{2} F_1(x) F_2(y, x) \tag{3.6}$$

where

$$F_1 = exp\left[\frac{x}{2\alpha_x}\left(1 - \sqrt{1 + \frac{4k\alpha_x}{v}}\right)\right] \tag{3.7a}$$

$$F_2 = erf\left(\frac{y + Y/2}{2\sqrt{\alpha_y x}}\right) - erf\left(\frac{y - Y/2}{2\sqrt{\alpha_y x}}\right) \tag{3.7b}$$

Eq. (3.6) (along with Eq. (3.7)) is obtained by ignoring vertical dispersion in Eq. (3.4) by setting the transverse vertical dispersivity, $\alpha_z$, in Eq. (3.5c) equal to zero (alternatively, the same result can be obtained by ignoring the $z$ terms when deriving the Domenico solution using Eq. (3.8) and Eq. (3.9)). The error function tends to $\pm 1$ as the argument tends to $\pm$infinity. Therefore, Eq. (3.5c) becomes 2 when $-Z/2 < z < Z/2$, and becomes zero otherwise. This makes the solution two-dimensional. To impose a steady-state condition, $t$ is taken to infinity. The complementary error function is defined as $erfc(z) = 1 - erf(z)$ therefore as the argument tends to infinity, $erfc(z)$ tends to zero and as it tends to negative infinity, it tends to 2. Using these properties, as $t$ goes to infinity, the terms after the addition sign in Eq. (3.5a) become zero and the first $erfc$ term becomes 2. The other terms do not depend on $t$ and therefore remain unchanged.

**Origins**

The original analytical solution, assuming an isolated plume originating from a constant concentration plane presented by Domenico and Robbins (1985), was derived based on a superposition of 1D analytical solutions of a continuous source which were derived by Crank (1975). The 1D solution for advection and dispersion given by Crank (1975) is

$$C(x,t) = \frac{C_0}{2} erfc\left(\frac{x-vt}{2\sqrt{D_x t}}\right) \tag{3.8}$$

Note that Eq. (3.8) is a truncated version of the original solution for 1D transport with advection and dispersion derived by Ogata and Banks (1961). The solution given by Crank (1975) for 1D transverse spreading in the $y$ and $z$ directions is

$$C(y,t) = \frac{C_0}{2} erf\left(\frac{y+Y/2}{2\sqrt{D_y t}}\right) - erf\left(\frac{y-Y/2}{2\sqrt{D_y t}}\right)$$
$$C(z,t) = \frac{C_0}{2} erf\left(\frac{z+Z/2}{2\sqrt{D_z t}}\right) - erf\left(\frac{z-Z/2}{2\sqrt{D_z t}}\right) \tag{3.9}$$

Domenico and Robbins (1985) then used a superposition argument to multiply Eq. (3.8) and the individual 1D spreading terms from Eq. (3.9) to get the final solution (Eq. (3.10)), however the superposition step was not justified.

$$C(x,y,z,t) = \frac{C_0}{8} erfc\left[\frac{x-vt}{2\sqrt{D_x t}}\right] \times$$
$$\left\{ erf\left(\frac{y+Y/2}{2\sqrt{D_y \frac{x}{v}}}\right) - erf\left(\frac{y-Y/2}{2\sqrt{D_y \frac{x}{v}}}\right) \right\} \times$$
$$\left\{ erf\left(\frac{z+Z/2}{2\sqrt{D_z \frac{x}{v}}}\right) - erf\left(\frac{z-Z/2}{2\sqrt{D_z \frac{x}{v}}}\right) \right\} \times \tag{3.10}$$

Domenico (1987) later extended Eq. (3.10) to include first order decay by incorporating the 1D solution for the infinite pulse source with decay given by Bear (1979) (the same solution is also given in Bear (1972)) by replacing the first $erfc$ term in Eq. (3.10) with

$$exp\left[\frac{x}{2\alpha_x}\left(1-\sqrt{1+\frac{4k\alpha_x}{v}}\right)\right] \times erfc\left[\frac{x-vt\sqrt{1+\frac{4k\alpha_x}{v}}}{2\sqrt{\alpha_x vt}}\right] \tag{3.11}$$

Later, Martin-Hayden and Robbins (1997) replaced the $erfc$ term in Eq. (3.10) with the full Ogata and Banks (1961) term with decay given by Bear (1972) (as well as Bear (1979)) shown in Eq. (3.12),

thereby giving the solution as presented in Eq. (3.5a).

$$exp\left[\frac{x}{2\alpha_x}\left(1-\sqrt{1+\frac{4k\alpha_x}{v}}\right)\right]\times erfc\left[\frac{x-vt\sqrt{1+\frac{4k\alpha_x}{v}}}{2\sqrt{\alpha_x vt}}\right]+$$

$$exp\left[\frac{x}{2\alpha_x}\left(1+\sqrt{1+\frac{4k\alpha_x}{v}}\right)\right]\times erfc\left[\frac{x+vt\sqrt{1+\frac{4k\alpha_x}{v}}}{2\sqrt{\alpha_x vt}}\right]$$

(3.12)

**Empirical Nature**

The Domenico solution is empirical in nature due to the unjustified superposition of the 1D solutions provided by Crank (1975) and as such, it is only an approximate solution to Eq. (3.2). Several problems with the Domenico solution have been identified by several authors and an overview of the major issues is now presented.

Note that in Eq. (3.10), the time $t$ in the transverse spreading terms (Eq. (3.9)) has been substituted by $x/v$. Domenico and Robbins (1985) justify this by saying that is is common in other transverse spreading models as well as it serves to keep the source dimensions constant the same as time increases. This substitution is still a subject of debate (Srinivasan et al., 2007; West et al., 2007) however it allows for the derivation of the Domenico solution from the Wexler (1992) solution as shown in Srinivasan et al. (2007). Regarding mass balance, West et al. (2007) (as well as in a subsequent reply by Ungs et al. (2009) to a comment by Prakash (2009)) state that because the Domenico solution is not a solution to Eq. (3.2) (as can be verified by substitution), it does not conserve mass. Finally, West et al. (2007) noted that the superposition of the 1D solutions given by Crank (1975) in Eq. (3.8) and Eq. (3.9) to get Eq. (3.10) would result in the concentration $C_0^3$ instead of $C_0$. According to West et al. (2007), Domenico and Schwartz (1998) sidestepped this issue by treating $C$ as the relative concentration, $C/C_0$, and is equal to the product of the relative 1D solutions (Eqs. (3.8) and (3.9)). Batu (2008) provides further insight and discussion regarding the analysis carried out by West et al. (2007) and Srinivasan et al. (2007).

These problems and inconsistencies with the Domenico solution serve to cement the fact that the Domenico solution can only be considered an approximate solution to Eq. (3.2), as recognized in the literature. Despite its flaws, the Domenico solution still proves popular in simplified models as discussed in Section 3.1.2.

### 3.1.2 Domenico Solution in Other Transport Models

The Domenico solution is a popular analytical solution and is used in many other simplified transport models. The U.S. Environmental Protection Agency (EPA) has developed the models BIO-SCREEN, BIOCHLOR, FOOTPRINT and RemCHLOR which all make use of the Domenico solution

(these models are available from the EPA's web site, along with a discussion on the limitations of the Domenico solution at `http://www.epa.gov/ada/csmos/domenico.html`). Additionally, it is used by simple models developed by the Pennsylvania Department of Environmental Protection such as the spreadsheet-based tools Quick_Domenico.xls and SWLOAD.xls (Pennsylvania Department of Environmental Protection, 2002).

A model that warrants a closer look, due to the types of contaminants being simulated is BIOSCREEN (Newell et al., 1996). BIOSCREEN is a spreadsheet-based tool that models the transport and fate of hydrocarbons in groundwater. The attenuation of hydrocarbons and nitrates in groundwater are closely related via the process of biological degradation, or biodegradation. In groundwater, both hydrocarbons and nitrates are subject to biodegradation due to bacteria living in the soil that make use of these substances to extract energy. In the case of a hydrocarbon contaminant plume, hydrocarbons act as elctron donors in the biodegradation reaction while dissolved oxygen acts as the electron acceptor (in aerobic, i.e. oxygen present, conditions). In absence of oxygen however, bacteria will use nitrate (among other substances) as the electron acceptor. The link between BIOSCREEN and our model is as follows: BIOSCREEN is concerned with the attenuation of hydrocarbons via biodegradation reactions with background substances that include nitrate, while our model is concerned with the reverse, namely the attenuation of nitrate plumes due to the presence of dissolved organic carbon. This means that the concepts developed and applied by BIOSCREEN should be applicable to modeling nitrate as well. One such concept utilized by BIOSCREEN is the concept of biodegradation capacity and the consideration of contaminant decay as an "instantaneous reaction". Since our model simulates the process of denitrification as a first order decay process and not an instantaneous reaction, these concepts are not implemented in the current version of our model however, the exploration of biodegradation as a model of denitrification is given as a possible avenue for future research.

### 3.1.3   Dispersivity Considerations

The scale-dependent nature of the dispersivity parameter has been widely recognized (Freeze and Cherry, 1979). In order to account for this dependency, several authors have attempted to relate the field scale with the dispersivity via regression of measured values (Pickens and Grisak, 1981; Arya, 1986; Neuman, 1990; Gelhar et al., 1992; Xu and Eckstein, 1995). Some authors and models have attempted to incorporate some of these regression formulas into the Domenico solution. For example, in the analysis by West et al. (2007), the formula provided by Pickens and Grisak (1981) and the one provided by Xu and Eckstein (1995) were used to improve the agreement of the Domenico solution to the solution provided by Wexler (1992). Jackson et al. (2004) also used the Xu and Eckstein (1995) regression with the Domenico solution. The EPA models BIOSCREEN (Newell et al., 1996) and BIOCHLOR (Aziz et al., 2000) both have the option to use the Xu and Eckstein (1995) formula to calculate dispersivity. In BIOCHLOR and in the investigations by Jackson et al. (2004) and West et al. (2007), the dispersivity was considered to be a function of the distance along the plume. The consideration of dispersivity in this way meant that the dispersivity
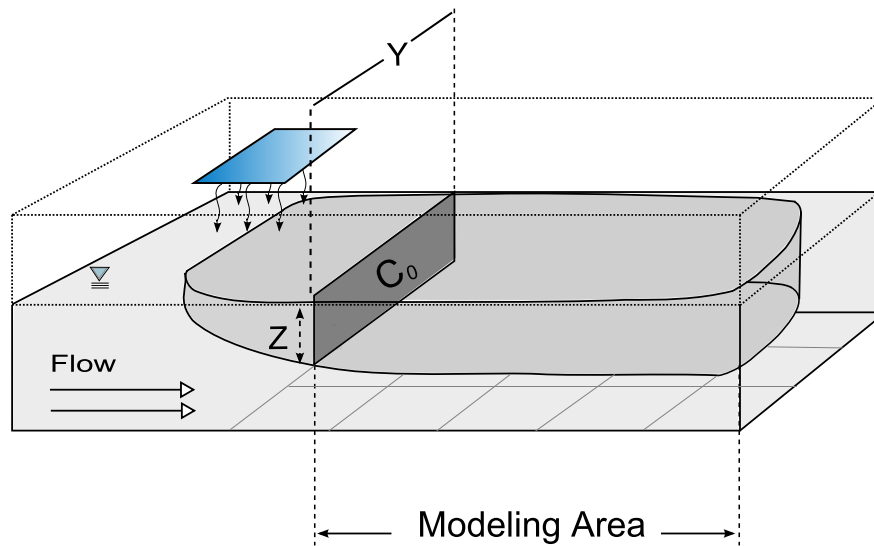
**Figure 3.2:** Conceptualization of the application of the Domenico solution in the model. Adapted from Aziz et al. (2000)

changed along the length of the plume, i.e. it was a function of $x$. This violates the assumption made by the PDE on which the Domenico solution is based one (Eq. (3.2)) which assumes the dispersivity is homogeneous. For comparison, the unsimplified governing equation (Eq. (3.1)) does not make this assumption. As a result, using a variable dispersivity in this way (i.e., in conjunction with the Domenico solution) may not be well justified. This may well be the reason that in a subsequent revision of BIOCHLOR, the use of the Xu and Eckstein (1995) formula in this way was discontinued. The reason for moving away from a variable dispersivity was not explicitly stated in the BIOCHLOR manual and the only explanation provided was that the model switched to basing the dispersivity on the estimated plume length so that dispersivity remained constant throughout the entire plume.

## 3.2   Conceptual Model

In applying the Domenico solution to the problem solved by this model, the system can be conceptualized as shown in Fig. 3.2. As shown in the figure, contaminants exit the OWTS drainfield, percolate through the unsaturated zone (bounded by the rectangular box indicated by the dotted lines), and enter the saturated zone (bounded by the box indicated by the solid lines) which has a uniform and steady flow in the direction indicated. Once the nitrate reaches the saturated zone, the evolution of the contaminant plume is approximated using the Domenico solution with the first order decay term, which models the effect of denitrification on the concentration distribution. As mentioned in Section 1.4.1, the transport model has been simplified to a two-dimensional, steady-state form. In Fig. 3.2, the 2D grid below the plume represents the cells in which the Domenico solution is evaluated. The simplification to 2D improves the efficiency of the software
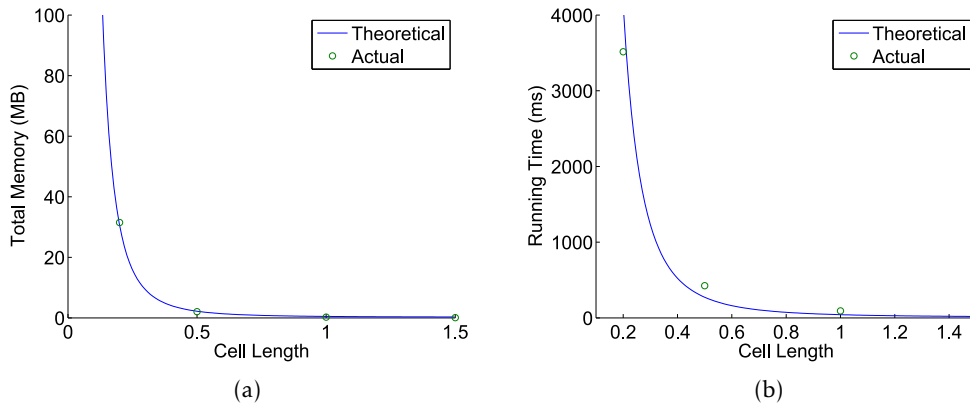
**Figure 3.3:** For a 3D test hemisphere, both the running time and memory requirements scale as $O(\Delta\ell^{-3})$, where $\ell$ is the length of the side of the evaluation cell. This shows the benefits of evaluating the Domenico solution in 2D

resulting in a large increase in execution speed and decrease in memory requirements. For a test hemisphere evaluated in 3D with a cubic mesh of size $\Delta\ell$, evaluating in two dimensions only, the savings in computational time and memory can be significant since in 3D, the running time and memory consumption both scale as $O(\Delta\ell^{-3})$ as shown in Fig. 3.3. As shown Fig. 3.2, the Domenico source plane is assumed to be at a location where the plume can essentially be considered 2D. Note that this arrangement actually corresponds to a pseudo-three dimensional, or 2.5D plume because although the plume is calculated in two dimensions, it also has a depth, obtained by simply extending the plume downwards by a depth $Z$ for the area labeled "Modeling Area" in Fig. 3.2. $Z$ corresponds to the estimated depth that the plume travels in and is necessary to be able to calculate the mass contained in the plume (see Chapter 4). This reduction of the 3D plume to 2D is also used by BIOCHLOR (Aziz et al., 2000) to calculate plume mass. With this pseudo-3D plume, the $Y - Z$ plane can still be thought of as the constant concentration plane of the 3D Domenico solution however the plume concentration at any point in the domain becomes invariant in the $z$ direction. This 2.5D approximation is considered to be valid when the vertical dispersivity of the aquifer is small or the seepage velocity is relatively high. Under most circumstances, the vertical dispersivity can be considered small since it is normally two orders of magnitude smaller than the longitudinal dispersivity and an order of magnitude smaller than the horizontal transverse dispersivity (Gelhar et al., 1992). Additionally, very low vertical dispersivities have been observed in field measurements (Gelhar et al., 1992). For the application of the model to OWTS nitrate loads, the OWTS will likely be relatively close to the water body, yielding a shallow plume. For comparison purposes, Fig. 3.4 shows a vertical $x - z$ view of a fully 3D Domenico plume where the vertical dispersivity has been set to a value two orders of magnitude lower than the longitudinal dispersivity. The figure shows a consequence of considering only a 2.5D plume versus a 3D plume since the plume depth. The 3D case is able to more accurately consider denitrification occurring only in the upper portion of the plume, where organic carbon is actually available. The reduced
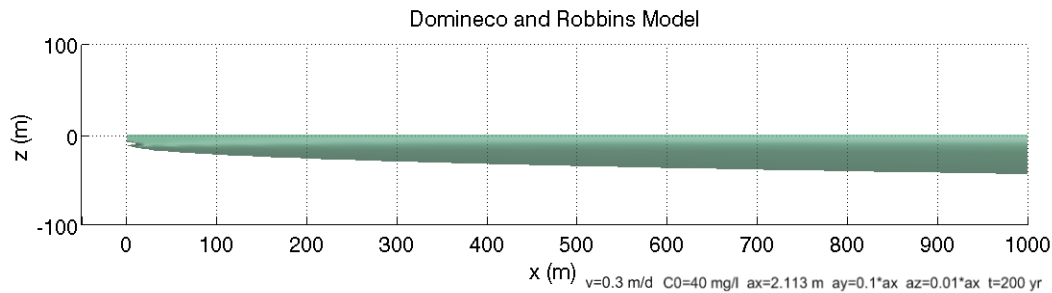
**Figure 3.4:** An $x-z$ view of the full Domenico solution with a realistic vertical dispersivity. The 0.01 mg/l isosurface is shown.

accuracy of the 2.5D approximation in considering denitrification remains a limitation of the model. In Fig. 3.4, the location of the water table (at $z = 0$) is taken into account by changing the $Z/2$ terms in Eq. (3.5c) to $Z$, as described in Domenico and Robbins (1985). In addition to assuming a 2.5D plume, the steady-state simplification is required so that the mass balance method of Chapter 4 may be used to calculate the nitrate load. Without the steady state simplification, a flux-based approach would be needed, such as the one employed in the simplified model developed by Andres (1992), which requires a longer computation time.

Eq. (3.6) shows that the flow velocity is treated as a constant for the entire plume. Recalling the simplified groundwater flow model from Chapter 2, the groundwater flow velocity is not considered as a constant in the domain. In fact, it varies spatially following the hydraulic gradient and its value is calculated via Darcy's Law. It is apparent that with the simplified transport model using the Domenico solution, it cannot fully consider heterogeneity of the flow field. However, despite the fact that the velocity for a single plume is fixed, it is possible to consider a different velocity for each plume. This is accomplished by taking the average velocity of the flow path corresponding to each source and using that value in the Domenico solution. This enables the partial consideration of heterogeneity in the flow velocity magnitude. The specifics of this averaging procedure are discussed in Section 3.3.2. Recall that the flow velocity is a vector. The velocity averaging routine only handles the heterogeneity in the flow *magnitude*, not the flow *direction*. The handling of different flow directions is important because the Domenico solution assumes flow is only in the x-direction in the entire domain. In the real world, the flow may be in any direction as determined by the hydraulic gradient. Like the heterogeneity in the flow direction, it is impossible for the Domenico solution, as given in Eq. (3.6), to directly consider a varying flow direction. Like the flow magnitude however, it is possible to approximate the effect of varying flow direction. In this model the heterogeneity in flow direction is incorporated via warping algorithms. Unlike the method for dealing with velocity magnitude heterogeneity, the warping method can more thoroughly (though not completely) deal with heterogeneity in the flow direction of a single plume as is seen in the following discussion.

The idea behind the warping of plumes to deal with heterogeneity in the groundwater flow
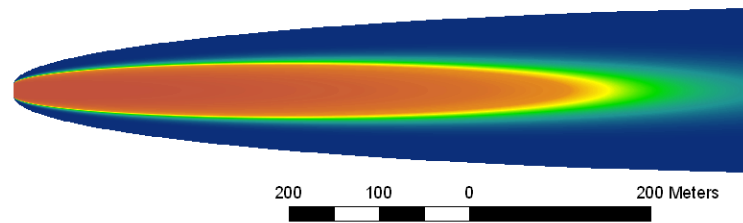
**Figure 3.5:** An $x - y$ (i.e. bird's eye view) cross section of a generic steady-state Domenico plume with decay. The source plane corresponds to the left-most vertical edge of the plume
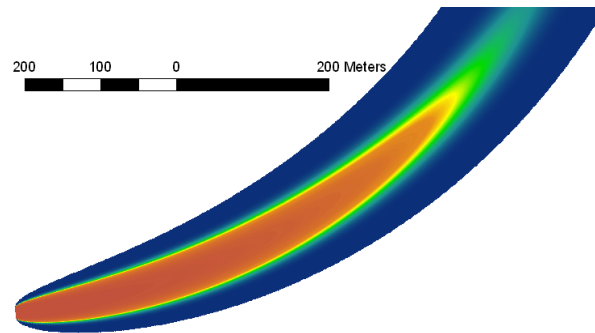


**Figure 3.6:** The desired transformation of Fig. 3.5 so as to take into account a curved flow path.

direction is illustrated in Fig. 3.5, which shows a generic (i.e. the parameters used are not important for this illustration) steady-state Domenico plume with decay. The plume shown in the figure shows that flow is indeed in the +x direction only. If the actual flow direction is also in this direction, then warping is unnecessary. Likely however, the flow direction will not be in the +x direction. Even more, it is highly probable that the flow path will not be straight, but an arbitrary shape as determined by the hydraulic gradient. For the sake of illustration, assume that the flow path is a simple arc, whose starting point corresponds to the location of the source and extends in the north-east direction. Then, the objective of the warping algorithm should be to transform the straight plume from Fig. 3.5 to the form shown in Fig. 3.6. This method of dealing with curved flow paths does not completely handle heterogeneity in the flow field as alluded to previously. In Fig. 3.6, it has been implicitly assumed that the flow direction has the same shape along the length of the plume. In other words, the flow direction at a points away from the plume's centerline is assumed to be the same as the flow direction at the centerline, which may not be the case. This is a limitation of this warping method of dealing with heterogeneity. The transformation from Fig. 3.5 to Fig. 3.6 can be obtained via a variety of geometric transforms. The model uses a combination of 1st order (affine) and 2nd order polynomial transforms as well as thin-plate spline warps. These are described in further detail in Section 3.3.2.

One of the distinct characteristics of this model compared to other simplified models is the ability to consider multiple contaminant sources. For example, if there is a neighborhood with
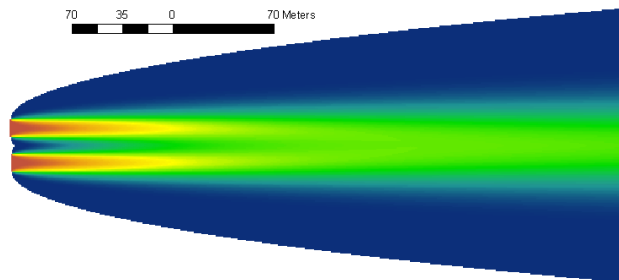
**Figure 3.7:** Illustration of the consideration of the overlap of two Domenico plumes.

multiple septic tanks, the model is able to handle them by merging the individual plumes into a single plume. Source locations are specified through an input GIS layer containing points representing the source locations. An example of this combination of plumes is shown in Fig. 3.7. Because denitrification is modeled as a first order decay process (as is common in the literature (McCray et al., 2005)), the amount of mass removed from a given volume of solute depends on the concentration of solute. Therefore, it is important to consider the influence of all plumes combined when calculating denitrification. Again, the limitations of the Domenico solution come into play in that the Domenico solution can only consider the evolution of a single plume in isolation. The influences on the plume evolution due to the presence of concentration gradients from other plumes cannot be considered directly. This may introduce some error when considering multiple plumes. Some authors (e.g., Ibaraki, 2001; Jackson et al., 2004) invoked the superposition principle to combine multiple plumes calculated via the Domenico solution. The superposition principle states that for a linear differential equation (Eq. (3.2) is linear), the linear combination of two separate solutions will also be a solution. Applying this principle to the Domenico solution is not, strictly speaking, correct. The reason being that as stated in Section 3.1.1, the Domenico solution is only an approximate solution to the governing equation. Because the Domenico solution does not consider the effect of other plumes, the concentration calculated by the Domenico solution may be slightly higher than it would be in the presence of another plume (due to the effect of higher concentration gradients in the overlap zone). Due to the lack of alternatives, our model also handles the superposition of plumes by taking the sum of each plume in the overlap area. Because the mass removed due to denitrification depends on the concentration, this method may result in errors in mass balance. An alternative method of considering overlap is to take the maximum of the overlapping concentration values. The effects of considering plume overlap using the maximum or summation on the concentration distribution are discussed further in Section 3.4.3.

As a summary of the transport conceptual model, the model considers the evolution of the nitrate contaminant plume by using the Domenico analytical solution to the ADE (Eq. (3.2)) where the input of nitrate contaminant is from a vertical plane source. The plume is assumed to be 2.5D (i.e., quasi-3D, see Fig. 3.2). Denitrification is modeled using a first order decay process. Heterogeneity in the flow field can be partially taken into account by averaging the velocity along individual flow paths and warping of the plume. The effect of multiple overlapping plumes is
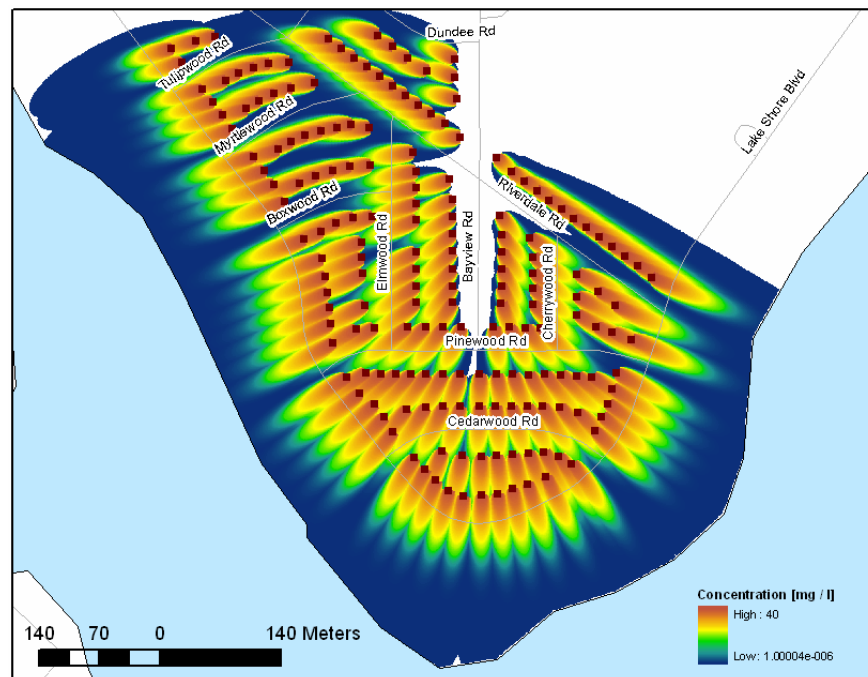
**Figure 3.8:** Illustration of the combined effects of heterogeneity in the flow field and the effect of multiple overlapping plumes. The effect of heterogeneity in the flow velocity magnitude can be seen by the differing plume lengths, especially the ones near the top. Heterogeneity in the flow velocity direction can be seen by the curved shape of some plumes (e.g. ones in the upper left))

taken into account by taking the superposition of all plumes in a given cell. Fig. 3.8 shows how the model handles the combined effects of all of the previously discussed conditions. In the figure, the squares represent septic tank locations in an area located in the Lakeshore neighborhood near the banks of the Cedar and Ortega rivers. Although flow lines are not shown for clarity, groundwater flows from each septic tank to the river, with a varying length and path taken.

## 3.3   Algorithm and Implementation

The desired output of the Transport module is a raster containing the combined nitrate concentration plumes of all sources in the input shapefile (Fig. 3.8 as an example). This is achieved by using the conceptual model described in Section 3.2. The user interface to the Transport model is shown in Fig. 3.9. Because the Transport module has a larger number of parameters than would otherwise comfortably fit on the main panel, a secondary panel (Fig. 3.9(b)), accessible to the user by clicking "more" on the main interface, was added.
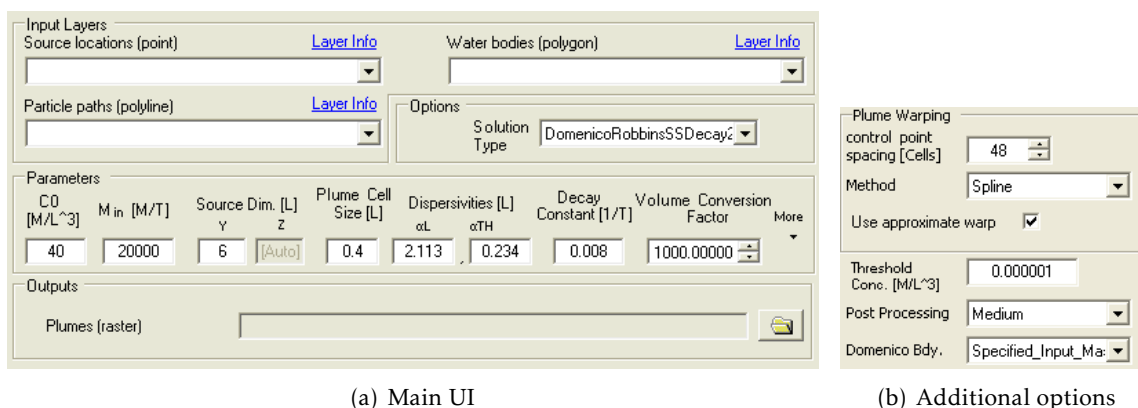
(a) Main UI                    (b) Additional options

**Figure 3.9:** Screenshot of the transport model UI

### 3.3.1   Input and Output

The inputs to the model (using the nomenclature from Fig. 3.9) are:

1. **Source Locations** This parameter is used to specify the layer containing point features that represent the locations of the nitrate sources. Additionally, each point may optionally have an associated set of attributes that can be used to specify transport parameters (e.g., $C_0$) on a plume by plume basis. Refer to the software user's manual for more details.

2. **Water bodies** Specifies the polygon feature class that contains the locations of water bodies. This information is used for determining load and for post-processing the plumes (described later).

3. **Particle Paths** This is a polyline feature class that specifies the flow path from each source to a water body. This input corresponds to the output of the Particle Tracking module. Note that every value of PathID present in the input Particle Paths feature class must have a corresponding FID in the Source Locations input feature class.

Model options and parameters are:

1. **Solution Type** The form of the Domenico solution to use. Currently, the Domenico solution is limited to two forms: Eq. (3.6) and a version where the decay rate, $k$, has been set to zero.

2. **$C_0$** The concentration of the source plane.

3. **$M_{in}$**. If the mass input rate of nitrate from the constant concentration plane is known, it can be specified using this parameter. To use this parameter, the Domenico Bdy. parameter must be set to "Specified_Input_Mass_Rate". When this parameter is specified, the $Z$ value of the plane source dimensions is automatically calculated (see Section 4.3.2 for details of the calculation).

4. **Source Dim.** The source dimensions specify the values of $Y$ and $Z$ for the source plane illustrated in Fig. 3.2. The source dimensions should be in the same units as the DEM (same distance units for $Y$ and same units as elevation for $Z$). If the value of the Domenico Bdy. parameter is "Specified_Z", then it is possible to manually specify the value of the $Z$ source dimension. If $Z$ is specified, $M_{in}$ is calculated automatically (see Section 4.3.2 for details of the calculation))

5. **Plume Cell Size** The size of the grid over which the Domenico solution is evaluated. Generally speaking, the finer the grid size, the more accurate the results are, at the expense of a decrease in execution speed and increase in memory consumption.

6. **Dispersivities** The dispersivities for the entire domain are assumed constant. This parameter sets the longitudinal and transverse dispersivities used by the Domenico solution.

7. **Decay Constant** The value of the first order decay constant, $k$.

8. **Volume Conversion Factor** This value is used to convert concentration values to mass values for use by the Nitrate Load Estimation module. It corresponds to the conversion factor between the volume units of the DEM and the volume units of concentration. For example, if the DEM has units of meters (for both distance and elevation) and the concentration $C_0$ is given in mg/l, then the conversion factor is 1000 since there are 1000 liters in one cubic meter.

9. **Warping Control Point Spacing** This value sets the spacing of the control points used for transforming the plumes via warping (see the following section for a detailed discussion of control points).

10. **Warping Method** The method to use for the warp (affine (1$^{\text{st}}$ order), 2$^{\text{nd}}$ order, or spline. See the following section for a detailed discussion of warping methods).

11. **Use Approximate Warp** Uses an approximate transformation by relaxing tolerances. Results in increased execution speed with only slightly less accurate warps

12. **Threshold Conc.** The threshold concentration below which the Domenico solution will not be calculated. All concentration values below this value will be implicitly considered zero. This value essentially defines the plume boundary, needed for calculating denitrification mass. Small values should be used to improve mass balance.

13. **Post-Processing** This option sets whether to use a quick, medium or slow processing algorithm to ensure that the boundary of plumes that enter a water body are accurately calculated. Details of this parameter are given later.

14. **Domenico Bdy.** Allows for specifying either the value of the $Z$ source dimension or the value of the input load from the source plane, $M_{in}$. When one is specified, the other is calculated automatically.

### 3.3.2 Process

Fig. 3.10 shows the algorithm for generating the output of the Transport module. In the same manner of describing the algorithm for the flow model, programmatic details are omitted. The algorithms for the sub-processes are described as they appear in the sequence of the main algorithm.

At the beginning of the main algorithm (Fig. 3.10(a)), the shapefile specifying the locations of the sources (given as point features) is loaded. First, a single source is read. The first task is to calculate the average velocity along the flow path corresponding to this source. The average velocity, $\bar{v}$, is calculated by taking the harmonic mean of the velocity corresponding to each segment (recall the segmented nature of the flow path as described in Section 2.3.3)

$$\bar{v} = \frac{(n-1) + w_n}{\left(\sum\limits_{i=1}^{n-1} \frac{1}{v_i}\right) + \frac{w_n}{v_n}} \tag{3.13}$$

where $v_i$ is the (constant) velocity of segment $i$ of the flow path and $n$ is the total number of segments in the path. The harmonic mean is used instead of the arithmetic mean so that the total travel time for a given travel distance along the path remains the same when calculated using the mean velocity and when using the sum of the travel times of the individual segments. In the steady-state case, the calculation of this average velocity is straightforward because the advection front of the plume (the location of the $x = \bar{v}t$ front) is at infinity therefore the weight given to the final segment of the flow path, $w_n$ in Eq. (3.13), is equal to one. However, in the transient state case, it is more complicated because the advection front may be at a location along the flow path that is shorter than the path itself. In the transient case, the weight corresponds to the location of the advection front, expressed as the ratio of the location of the advection front along the final flow segment, to the length of the final segment itself, thereby giving a $w_n$ between 0 and 1. Although the model is only for steady-state cases, the transport code is capable of handling transient state cases as well, and this functionality is used in the analysis contained in Section 3.4.

After $\bar{v}$ is calculated, the Domenico solution is evaluated. The evaluation of the Domenico solution is an iterative process. The evaluation begins at the location of $(0,0)$ and the routine proceeds to evaluate the positive $y$ half of the plane with a block-centered approach, in increments of $\Delta \ell$ (the plume cell size) while keeping $x$ constant. That is, the evaluation points correspond to the center of an array of blocks whose centers are separated by $\Delta \ell$. To speed up evaluation of the multiple error function terms contained in the Domenico solution, evaluation of the error function is carried out via a lookup table generated by using the polynomial approximation of the error function given in Press et al. (1996). Once the condition $C(x,y) < C_{thresh}$ is reached (discussed below), the value of the $y$ coordinate is reset to zero and $x$ is incremented by $\Delta \ell$. The process is repeated until the evaluation of the next value of $(x,0)$ triggers a stopping condition. The stopping conditions are: $C(x,0) < C_{thresh}$, or $x >$ flow path length (this corresponds to the 'None' post-processing option). The evaluation routine uses two properties of the Domenico solution in order to speed up computation and reduce memory requirements. The first property is that of
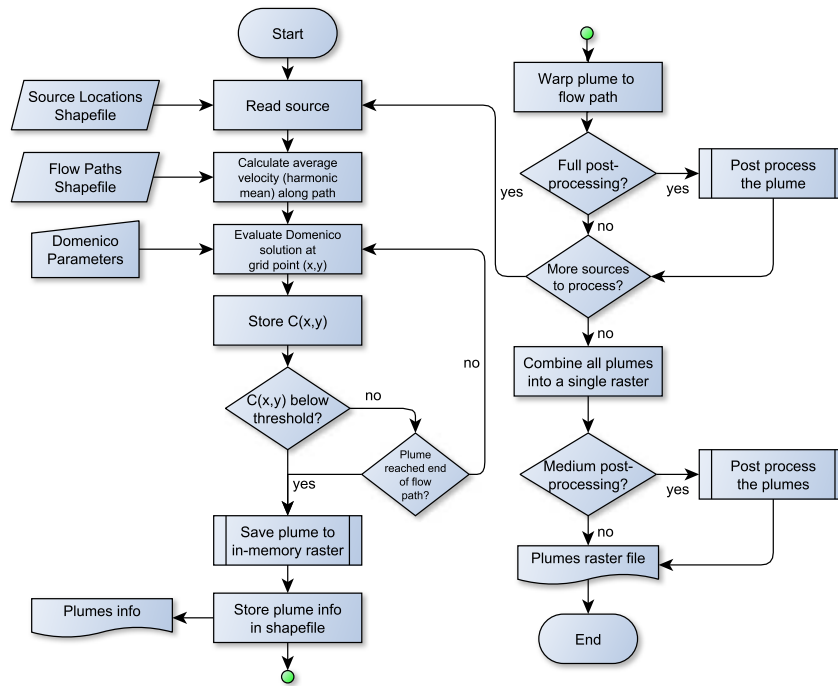
(a) Main algorithm
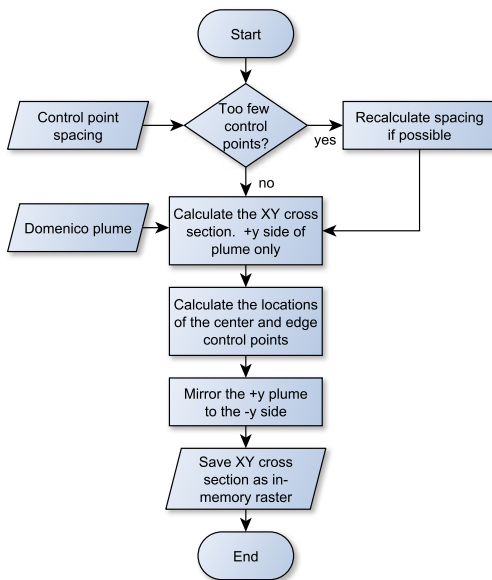


(b) Convert plume to an in-memory raster
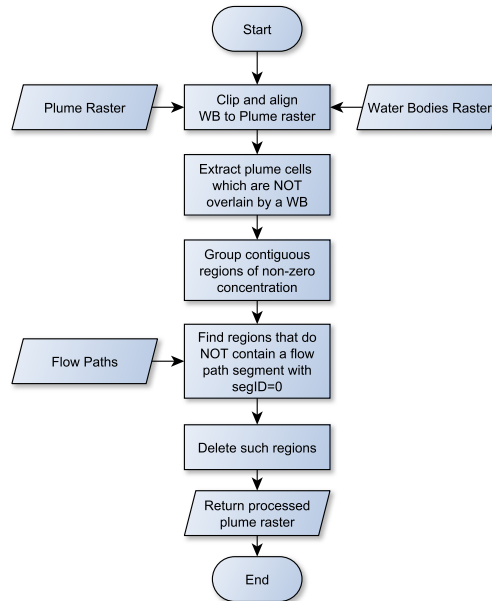


(c) Plume post-processing

**Figure 3.10:** Logic flow for the transport model
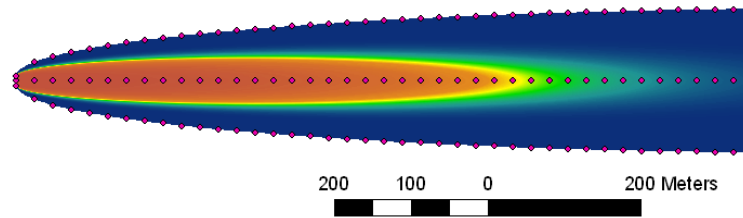
**Figure 3.11:** Plume of Fig. 3.5 with control points (circles) superimposed.

symmetry. Since the plume is always symmetric about the $y$ axis, only half of the plume needs to be evaluated resulting in significant computational savings. The second property is that of convexity. Because isosurfaces of the Domenico solution will always be convex, it is guaranteed that the concentration on the plume centerline will always have the largest values. In other words, the "tip" of the plume will always be the furthest point away from the source. This guarantees that, upon triggering the stopping condition $C(x, 0) < C_{thresh}$, there are no further areas of the plume having $C \geq C_{thresh}$.

Once the plume has been calculated, it is necessary to translate the internal array structure used for its storage into a raster format suitable for use by ArcGIS. A high level overview of this algorithm is shown in Fig. 3.10(b). This subroutine performs two functions. The first function calculates the positions of the control points along the plume used for warping. For this function, a check is first performed to ensure that given the plume length and user-specified control point spacing, there are enough control points for warping. For $1^{st}$-order transformations, a minimum of three points are required. For $2^{nd}$-order transformations, six points are required and for spline transformations, ten points are needed. If there are not enough control points, the spacing is automatically reduced so that the minimum is obtained such that there is a maximum of one control point per raster cell. If the plume is too short to contain the minimum number of points, an error is generated. After the appropriate spacing is calculated, the assignment of control points proceeds in the following fashion: first, a series of points are assigned along the plume centerline at regular intervals according to the specified control point spacing. Then, control points are assigned to the boundaries of the plume, again using the control point spacing interval. Using the plume from Fig. 3.5 as an example, the control points calculated with a spacing of 48 cells are shown in Fig. 3.11. Warping is discussed in further detail later. While calculating the control points, the plume is copied from the internal data structure to an array which will subsequently be copied to an ArcGIS in-memory raster. The final step of this subroutine is to generate the full plume by mirroring the $+y$ half of the plane to the $-y$ half. After the plume has been evaluated and the cross-section generated, the parameters used in its calculation as well as some extra information used by the Nitrate Load Estimation module are stored in a separate auxiliary shapefile having a suffix "_info". This shapefile contains point features corresponding to the location of the source of the plume in question and the associated plume information is stored in the attribute table of the feature class.
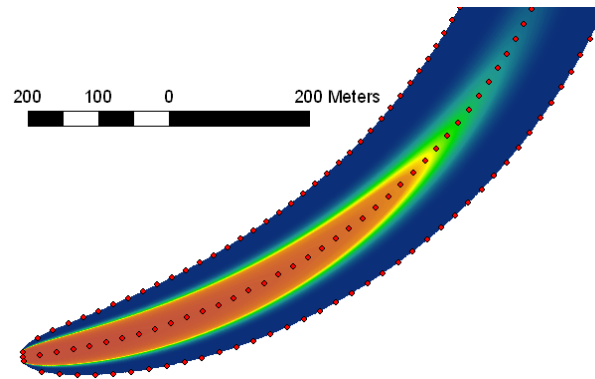
**Figure 3.12:** Plume of Fig. 3.6 with target points (circles) superimposed.

The plume raster is then transformed, via a warping procedure, from the originally straight flow path, to the actual flow path. The individual warping algorithms used in the model are described below. The general idea is that given the plume raster and a set of control points and target points, deform the plume so that the plume centerline conforms to the flow path. Additionally, the width of the plume (defined as the twice the perpendicular distance from the plume centerline to the plume boundary at any given point along the centerline) must remain the same after the deformation. The relationship between pairs of control points in the original plume, $(u_i, v_i)$, and target points in the warped plume, $(x_i, y_i)$ defines a transformation

$$(x_i, y_i) = (X(u_i, v_i), Y(u_i, v_i)) \tag{3.14}$$

where $X$ and $Y$ are mapping functions determined by the warping method. The mapping is then used, in conjunction with interpolation, to transform the values from the original coordinates, $(u, v)$ to the warped coordinates, $(x, y)$. In the algorithm, the target points are obtained based on the control points by mapping the centerline points to points along the flow path and constraining the points along the boundary of the plume so that the plume width is maintained. Fig. 3.12 shows the target points superimposed on the plume of Fig. 3.6. In fact, the plume shown in Fig. 3.6 was generated by the $2^{nd}$ order polynomial warp of Fig. 3.5. In the discussion that follows, only a brief overview of the mathematics behind warping is provided. This is done because the topics of transformations and interpolation techniques are expansive and have been widely studied and as a result, there is a large body of literature regarding them. The warping techniques described below are implemented by a set of ArcGIS libraries, which our modeling tool makes use of.

The transformation that maps cells in the original raster to new locations is given by Eq. (3.14). For practical purposes, it is usually the inverse transformation that is of interest (Schowengerdt, 2006):

$$(u_i, v_i) = (U(x_i, y_i), V(x_i, y_i)) \tag{3.15}$$

To avoid problems with "overlapping cells" and "holes", the output space is partitioned into a grid after which each point $(x, y)$ in the output domain is traversed in sequence and mapped back to

the original space $(u, v)$ where interpolation is used to obtain a value for the cell at $(x, y)$ from the value in the cells surrounding $(u, v)$ (Schowengerdt, 2006). Interpolation in $u, v$ space is necessary because the transformation from the regularly spaced $x, y$ grid will: a) likely not yield a 1 to 1 mapping of cells and b), the mapping from the cell center at $(x, y)$ will likely not correspond to the center of the cell in $u, v$ space. That is, one cell centered at $(x, y)$ in the output space may correspond to none, one, or many cells in the input space depending on the particular mapping function, with the additional possibility that the cell center in the output space will likely not correspond to the center of a cell in the original space.

For polynomial transformations, the mapping functions are of the form (Schowengerdt, 2006)

$$U(x, y) = \sum_{i=0}^{N} \sum_{j=0}^{N-i} a_{ij} x^i y^j, \quad V(x, y) = \sum_{i=0}^{N} \sum_{j=0}^{N-i} b_{ij} x^i y^j \tag{3.16}$$

where $N$ is the order of the polynomial and $a_{ij}$, $b_{ij}$ are coefficients that are to be determined. The case of $N = 1$ corresponds to an affine transform (translation, scaling, rotation and shear)

$$U(x, y) \equiv u = a_{00} + a_{10} x + a_{01} y$$
$$V(x, y) \equiv v = b_{00} + b_{10} x + b_{01} y \tag{3.17}$$

Eq. (3.17) is the general form of the affine transformation where each coordinate, $x$ or $y$, is back-transformed to the original coordinate system $u$ or $v$. Taking $u$ as an example, to solve the three unknowns (the $a_{ij}$), at least three equations are needed. These equations are obtained from the control and target points. Each pair of control-target points defines an equation that can be used to solve for the unknown $a_{ij}$. For example, if there are three control-target point pairs, designate the known control points (in the unwarped domain) as $(u_1, v_1)$, $(u_2, v_2)$ and $(u_3, v_3)$ while the target points in the warped domain are $(x_1, y_1)$, $(x_2, y_2)$ and $(x_3, y_3)$. Then, the system of equations can be written as

$$u_1 = a_{00} + a_{10} x_1 + a_{01} y_1$$
$$u_2 = a_{00} + a_{10} x_2 + a_{01} y_2 \tag{3.18}$$
$$u_3 = a_{00} + a_{10} x_3 + a_{01} y_3$$

or in matrix notation,

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \end{bmatrix} \tag{3.19}$$
$$\mathbf{U} \quad = \qquad \mathbf{L} \qquad \qquad \mathbf{A}$$

therefore,

$$\mathbf{A} = \mathbf{L}^{-1} \mathbf{U} \tag{3.20}$$

The system can then be solved exactly, resulting in the mapping function $U(x, y)$. A similar procedure is carried out for $V(x, y)$. In the general case, the minimum number of control points

required for polynomial transformations is given as (Schowengerdt, 2006):

$$K = \frac{(N+1)(N+2)}{2} \qquad (3.21)$$

Generally speaking however, it is advantageous to have more control points than are required in order to obtain a more accurate transformation over the entire raster (Schowengerdt, 2006). When there are more control-target points than required, the system is over determined, and the transformation parameters that give the smallest squared error in the target point locations can be found by the least-squares method.

Spline transforms, called thin-plate spline (TPS) warps, are somewhat more complicated. The idea is still to find transformation functions $U(x,y)$, and $V(x,y)$ from Eq. (3.15). The TPS can be thought of as the two-dimensional counterpart of the 1D cubic spline except that instead of using the polynomial basis $|x|^3$, the basis $U(r) = r^2 \log r^2$ is used where $r = \sqrt{x^2 + y^2}$ is the distance of the point $(x,y)$ from the origin (Bookstein, 1989). In 1D, the spline interpolant can be thought of as a thin, flexible rod that has been constrained at several points, which generate a smooth curve between the points. In 2D, it is a thin, flexible plate that has been fixed at several points above or below the z-axis and the plate smoothly interpolates the points by passing a surface of minimum bending through them. When applying the TPS to image warping instead of surface interpolation, the displacement in $z$ is instead applied to each of the image coordinates. Two splines are needed for the deformation in each coordinate, $U(x,y)$, and $V(x,y)$ (Bookstein, 1989) (note the function $U(r)$ is distinct from $U(x,y)$). Bookstein (1989) gives the transformation as

$$u = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U(\|(x_i, y_i) - (x,y)\|) \qquad (3.22)$$

where $a_1$, $a_x$, $a_y$ and $w_i$ are the unknowns and $n$ is the number of control-target point pairs. The vertical bars indicate the Euclidian norm. The unknown coefficients are determined by writing Eq. (3.22) in matrix form and solving the system of equations,

$$\mathbf{L}^{-1} \begin{bmatrix} \mathbf{U} \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{W} \\ a_1 \\ a_x \\ a_y \end{bmatrix} \qquad (3.23)$$

where $\mathbf{W} = (w_1 \cdots w_n)$ and $\mathbf{U} = (u_1 \cdots u_n)$, giving the mapping $U(x,y)$. A similar calculation can be carried out for $V(x,y)$. See Bookstein (1989) for the exact form of $\mathbf{L}$. An additional condition given by Donato and Belongie (2002) is that the points $(u_i, v_i)$ should not be collinear. Additionally, from Eq. (3.22), it can be seen that the TPS contains an affine component. When the number of specified control points is three, the $w_i$ terms become zero and the TPS warp reduces to an affine transform (Bookstein, 1989). As a result, the minimum number of control points required for a TPS warp is three, however the implementation in ArcGIS requires a minimum of 10 control points (ESRI,

(a) None                    (b) Medium/Full

**Figure 3.13:** Example of a simple case of post-processing on a single plume. The thin blue line represents the flow path. The light blue polygon is the water body.

2009). An important difference between the TPS warp and the previously discussed polynomial warps is that TPS optimizes local accuracy while polynomial warps optimize global accuracy (ESRI, 2009). This gives TPS an advantage over polynomial warps when there are many control points. When there are many control points that are well distributed, TPS will provide a more accurate transformation, especially when the amount of bending required is more extreme, the reason being that TPS warp maps the control points to the target points exactly. Because of this behavior, TPS is the default transform in the model.

After the plume is warped, the model algorithm continues with "Full" post-processing (if the user selected it). Note that the algorithm for "Full" and "Medium" post-processing is the same. The only difference between the two is the point in the algorithm at which the post-processing is applied: when the "Full" post-processing option is selected, processing is applied on a plume by plume basis while if "Medium" is selected, processing is done at the end after all the plumes have been combined into a single raster. When there are many plumes, "Medium" results in a significant increase in speed, however there are special cases where it does not give the expected result (e.g., Fig. 3.15(a)). In such cases, the user must select "Full" post-processing. The special cases are described later when "Medium" post-processing is discussed. Recall that the purpose of post-processing is to ensure that the plumes that intersect a water body terminate at the water body. The algorithm for post-processing is shown in Fig. 3.10(c) and proceeds as follows. First, the plume raster and the water bodies raster are clipped and aligned (via resampling) so that the cells in each raster line up. This is important so that the plumes are correctly clipped. The next step is to use the water bodies raster as a mask to delete any portions of the plume that fall in the water body. For simple cases where the water body is large (e.g., a large lake or river), this step is sufficient to correctly process the plume. The behavior of this simple case is shown in Fig. 3.13 along with the case of no post-processing. Note that for a single plume, "Medium" and "Full" post-processing are equivalent. Also note that in Fig. 3.13(a), the upper left corner of the plume does not reach the river while in Fig. 3.13(b) it does. This is because when post-processing is enabled, the plume is

(a) None                    (b) Incomplete                    (c) Medium/Full
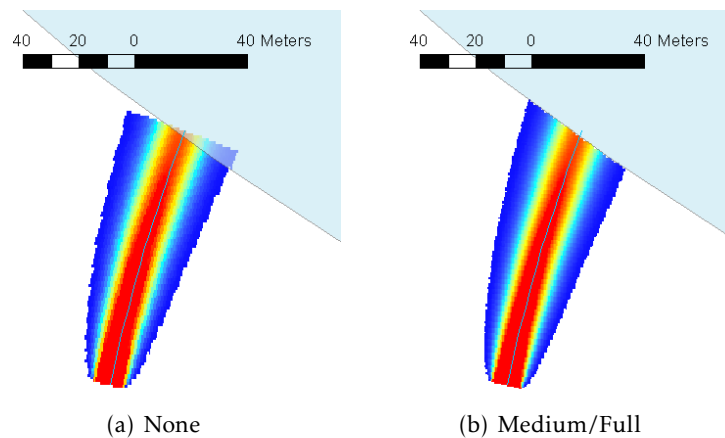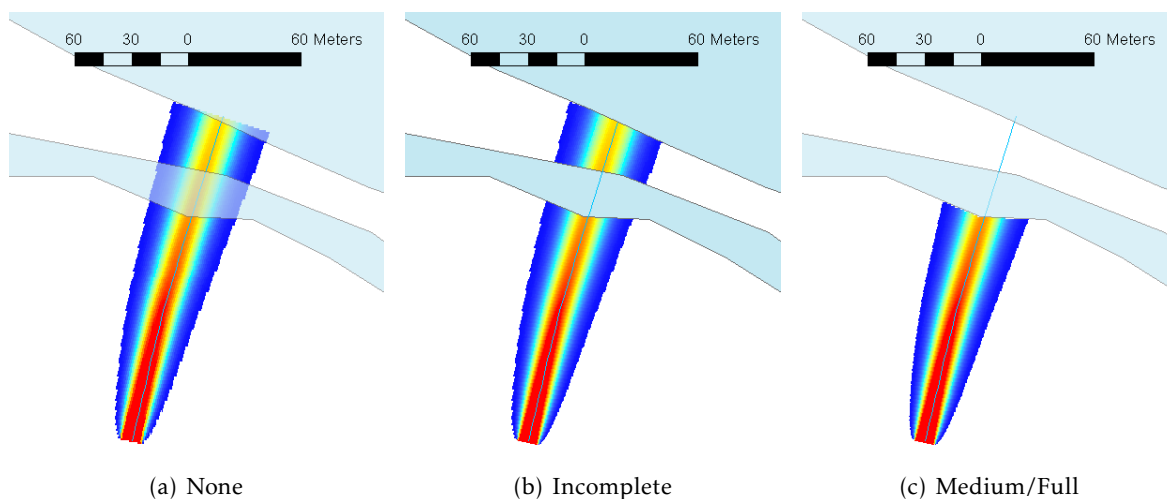
**Figure 3.14:** Example of a more complex case of post-processing on a single plume. The thin blue line represents the flow path. The light blue polygons are water bodies.

evaluated by the Domenico evaluation routine to a length 10% longer than the flow path so as to mitigate this issue.

For more complicated cases, further consideration is required, for example when the plume is long enough to flow *through* a small creek, the portion that is on the other side of the creek must be eliminated. This is achieved by subsequently grouping contiguous regions of non-zero concentration values. Then a table is built to identify which regions contain a contaminant source. The regions that do NOT contain a contaminant source are then deleted. This algorithm works because once the water bodies are masked out of the plumes raster, if a creek intersects a water body, the portion that is on the other side of the creek is isolated from the main plume. This portion can then be identified and deleted. An artificial case of this scenario is illustrated for a single plume in Fig. 3.14. Again, for a single plume, "Medium" and "Full" post-processing are equivalent. Fig. 3.14(b) shows the result if the algorithm only considers the simple case described previously. Fig. 3.14(c) shows the correct result, after the algorithm considers this special case using "Full" post-processing. The result with no post-processing is shown for comparison (Fig. 3.14(a)). Now, consider a water body geometry like the one shown in Fig. 3.13 except now, instead of just a single plume, there are many overlapping plumes, both behind the small creek and in the gap between the creek and the lake as shown in Fig. 3.15. In this case, selecting the "Medium" post-processing option would produce incorrect results as once the plumes have been combined into a single raster, it becomes impossible to separate them. Thus it is impossible to remove just the portion of plumes that cross the small creek. Fig. 3.15(a) shows the case where "Medium" post-processing was selected. Plumes in the gap between the creek and the lake cannot be deleted since the plumes were combined into a single raster. Fig. 3.15(b) is generated using the "Full" post-processing option. Because each plume is operated on individually, "Full" post-processing produces the expected behavior. Note that the "None" post-processing option (if selected) is actually carried out during

(a) Medium                                    (b) Full
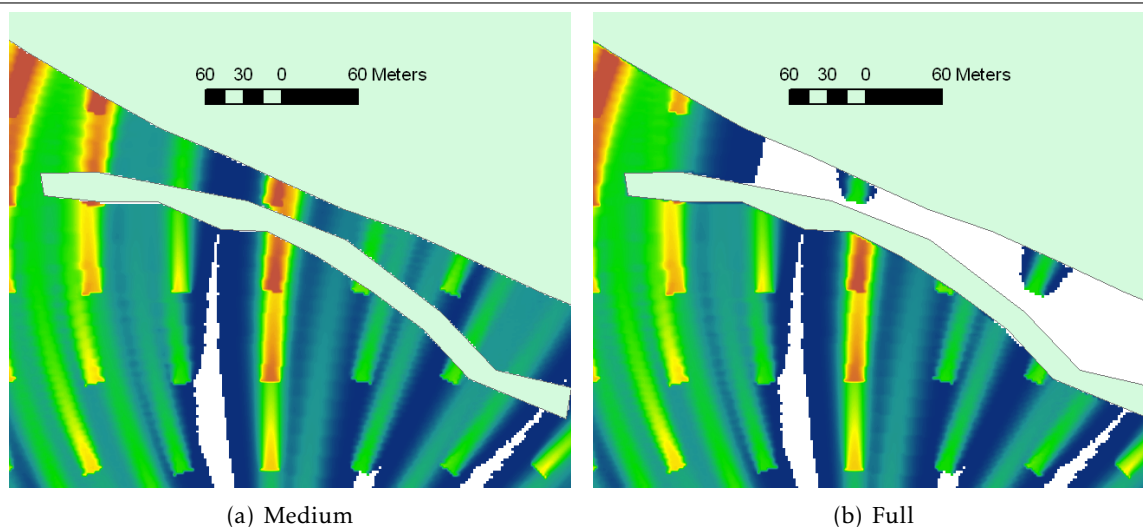
**Figure 3.15:** Example of a more complex case of post-processing with multiple plumes. For clarity, flow paths are not shown however the direction of flow is the same as in Fig. 3.13 (approx. south to north).

the evaluation of the Domenico solution. It is also important to note that "Medium" and "Full" post-processing only affects the concentration distribution. For the purposes of calculating nitrate loads, the "None" option is always used for performance reasons.

The algorithm now proceeds to save the individual plume to memory. If there are more sources to calculate, the entire plume calculation algorithm is repeated until all sources in the source locations shapefile have been processed. Once all plumes have been calculated, they are combined into a single raster by the ArcGIS mosaic command. As discussed in the conceptual model, in areas where plumes overlap, the sum of the concentrations is taken. Then, if "Medium" post-processing was selected, the plumes are processed according to the algorithm described previously. Finally, the plume raster is saved and output, along with the associated plume information shapefile.

## 3.4   Analysis of Model Performance

In this section, the behavior and performance of this simplified Domenico-based model will be analyzed using synthetic test cases. The model is then applied to a real-world test case using parameter values obtained from literature. The synthetic tests were constructed to evaluate the effectiveness of specific aspects of the approximations in the model including the treatment of heterogeneities in groundwater flow and the treatment of overlapping plumes. The comparison is done against a numerical model constructed with the combination of MODFLOW (Harbaugh et al., 2000) for the groundwater flow portion and MT3DMS (Zheng and Wang, 1999) for contaminant transport. In regards to the real-world test case, the model is compared to a limited amount of measurements conducted at several homes in the Hillcrest neighborhood of Jacksonville, Florida, located approximately 2.5 km (1.5 mi) to the north-west of the Lakeshore neighborhood studied in

Section 2.4.3.

### 3.4.1   Ideal Case

Before studying the effects of different aspects of the model, it is important to understand the performance of the model under ideal conditions (i.e, all of the Domenico conditions from Section 3.1.1 are satisfied). A steady-state MODFLOW model was constructed, consisting of a rectangular domain of width 2000 m and length 2000 m having a single layer of depth 2 m. The grid size was set to 4 m x 4 m. The boundary conditions of the MODFLOW model were set so that a flow field would have a constant magnitude and direction, thereby matching the constraints of the Domenico solution. At $x = 8$ m, the boundary was set to a constant head value of 10 m and at $x = 2000$ m, the boundary was set to a constant head of 5 m. The other boundaries were set to no flow. The hydraulic conductivity was set to a uniform 2.113 m/day and porosity was set to 0.25. Using Darcy's Law, one obtains, via manual calculation, a seepage velocity of 0.02121 m/day. The value calculated by MODFLOW was 0.02124 m/day with flow exclusively in the easterly direction. The Domenico plane was simulated in MT3DMS by a series of three constant concentration blocks located at the $(x, y)$ coordinates $(118, 1070)$, $(118, 1066)$, $(118, 1062)$. The source cells were placed at this location to eliminate effects from the domain boundary located at $x = 8$ m. In order to eliminate mass transfer in the -x direction, the dispersivities for all $x < 118$ and for $y < 1062$, $y > 1070$ when $x = 118$ m, were set to zero. Additionally, a second set of constant concentration cells at $(114, 1070)$, $(114, 1066)$, $(114, 1062)$ were added to prevent dispersive mass leakage from the first set of constant concentration cells into the zone of zero dispersivity. The presence of a water body was simulated in MT3DMS by setting the column located at $x = 500$ m to a constant concentration boundary with a value of zero. Because a true steady-state solution is not available in the version of MT3DMS of Zheng and Wang (1999), the MT3DMS plume was simulated for a period of 200 years using the TVD solver. It was determined, by examining concentration distribution changes from simulations using shorter simulation times, that the plume after 200 years could be considered to be in the steady state. The decay rate was set to zero.

Before examining the results of the ideal case, it is informative to state the conclusions of Srinivasan et al. (2007) and West et al. (2007). Srinivasan et al. (2007) and West et al. (2007) both examined the error associated with the Domenico solution by comparing it to the exact solution presented by Wexler (1992). The main conclusion stated by both investigations is that Domenico plumes will generally under-predict the concentration along the plume centerline, resulting in the under-prediction of the plume extent. Additionally, Srinivasan et al. (2007) stated that Domenico plumes are always wider than the actual plumes. These results are briefly examined here by comparing the Domenico solution as used in this model to the numerical solution generated by MT3DMS. Fig. 3.16 shows the results of the simulation. Fig. 3.16(b) shows an enlarged version of Fig. 3.16(a) near the source. It can clearly be seen from the figure that near the plume centerline, the concentrations are underestimated compared to the MT3DMS result. This seems to support the conclusions of Srinivasan et al. (2007) and West et al. (2007). Along the plume centerline, the
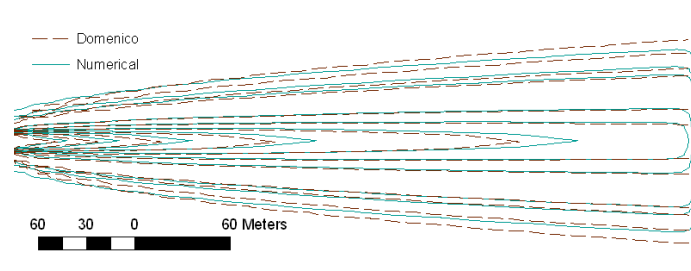
(a) Plume contours
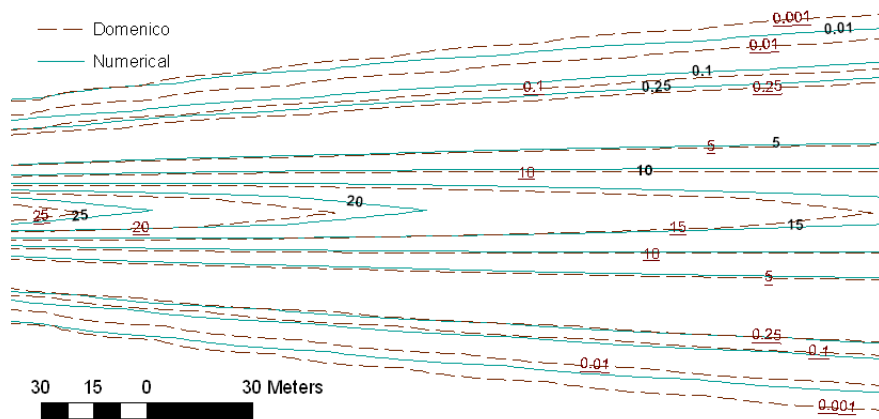


(b) Enlarged area near source

**Figure 3.16:** Comparison between the Domenico solution with no decay and the numerical MT3DMS solution for the same case. Underlined labels correspond to Domenico contours

concentration is underestimated by a maximum of 6.5%. In this case, the conclusion that Domenico plumes are always wider than the actual plume is not supported by the result, since examining the position of the outer contour levels in Fig. 3.16(b) (e.g. the 0.01 mg/l contour), it can be seen that the Domenico plume is in fact narrower than the MT3DMS plume. A case that supports the conclusion of Srinivasan et al. (2007) is given in Fig. 3.17. The only difference between Fig. 3.16(b) and Fig. 3.17 is that for Fig. 3.17, the decay constant was not zero but was set to $k = 0.008 \, day^{-1}$. It can be seen from Fig. 3.17 that for this set of parameters, the Domenico plume is wider than the numerical solution near the middle of the plume and narrower near the tip and near the source. This would indicate that in fact Domenico plumes are not *always* wider than Domenico plumes as indicated by Srinivasan et al. (2007), but in fact it may or may not be wider, depending on the chosen parameters, the selected contour level and location along the plume. The assertion that Domenico plumes under-predict the concentration along the centerline is confirmed again by Fig. 3.17. For this case, the underestimation ranges from 3% at the source and increases with distance to over 60%. Srinivasan et al. (2007) presents several more cases with varying values of velocity and dispersivity, each generating varying widths of Domenico plumes, some wider than the exact solution near the source, others wider near the advection front.
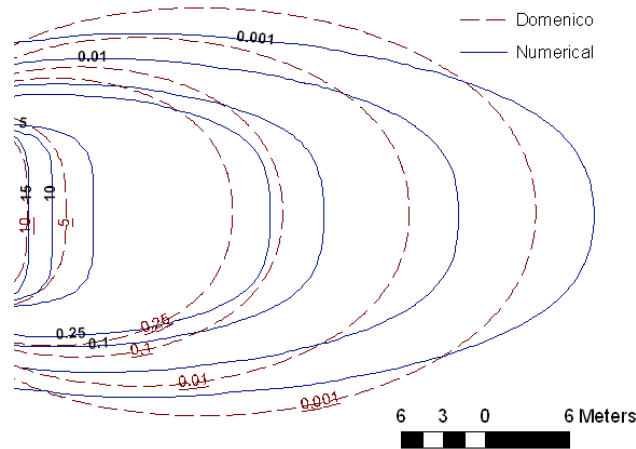
**Figure 3.17:** Comparison between the Domenico solution with decay, $k = 0.008 \, day^{-1}$ and the MT3DMS solution.

### 3.4.2   Effects of Flow Heterogeneity on the Domenico Solution

Now that the behavior of the Domenico solution in the ideal case has been studied, it is possible to examine the behavior when the Domenico parameters are no longer constant in space, as would be the case in a real application scenario. The cases presented below only test a variable flow velocity magnitude or direction. The effects of a variable dispersivity, porosity and decay rate are not tested. In the following sub-sections, effects are examined under controlled conditions by adhering to all Domenico assumptions except the one being tested.

**Varying Flow Velocity**

The performance of the Domenico model under a steady flow field having a spatially non-uniform flow magnitude but a constant flow direction (east), is evaluated. Specifically, this test examines the behavior of the velocity averaging algorithm, described in Section 3.3.2, on the shape of the plume. A simple case was devised in which the flow velocity varied linearly in the direction of flow. In order to obtain a linearly varying velocity magnitude in the numerical model, a constant amount of recharge was added to the domain. Additionally, the constant head boundary on the left side of the domain was adjusted from 10 m to 100 m to compensate for flow in the opposite (westerly) direction due to recharge. The velocity profile obtained is shown in Fig. 3.18. All parameters for the transport simulation remain the same as those of Section 3.4.1 except the decay rate has been set back to zero. In theory, using the Domenico solution, the profile given in Fig. 3.18 should produce a plume that is narrower than the numerical plume near the source where flow magnitude is slower and a plume that is wider near the end of the plume when the flow is faster. Intuitively speaking, given similar source conditions, a faster flow will generally produce narrower and longer plumes than slower flow. Because there is an important effect (discussed later) near the plume
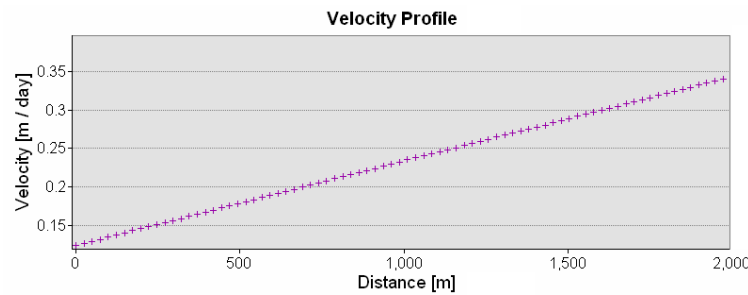
**Figure 3.18:** Velocity profile obtained using MODFLOW for testing the effect of varying velocity on the Domenico solution.



(a) 0.1 and 0.01 mg/l contours



(b) Enlarged area near source

(c) Enlarged area near advection front
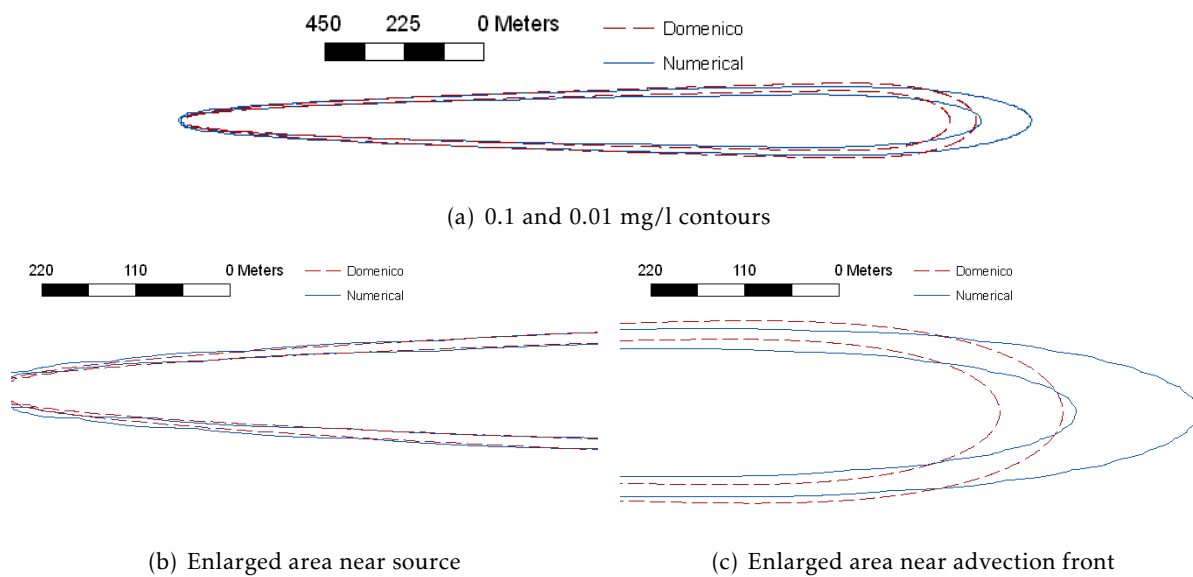
**Figure 3.19:** Comparison of Domenico and MT3DMS plumes for a spatially varying velocity.

advection front, this analysis will be carried out with the transient Domenico solution (2D version of Eq. (3.4)) instead of the steady-state solution used in the model (Eq. (3.6)). Since the plume can be considered to be in the steady-state far behind the advection front (Domenico, 1987), the results are still transferable to the steady-state case in the zone behind the advection front. It has been previously determined from Section 3.4.1, that the Domenico solution consistently underestimates plume concentrations at and near the centerline. Therefore, for clarity of the figures, only the outer contours, 0.1 and 0.01 mg/l, are plotted. The analysis was carried out with a plume with $t = 25$ years and is shown in Fig. 3.19. Note that although the transport simulation is transient, the flow simulation remains a steady-state simulation. Fig. 3.19 shows the expected result regarding plume width; near the source, where the flow magnitude is lower than the average value, the Domenico plume is slightly narrower than the numerical plume (Fig. 3.19(b)). As the flow velocity increases downstream, the MT3DMS plume responds to the increase in flow velocity and becomes narrower.

Because the Domenico plume is fixed to the average velocity, it cannot respond to the changing flow, and because the flow magnitude in this zone is higher than the average value (i.e. the value used by the Domenico solution), the Domenico plume is wider in this zone compared to the numerical result (Fig. 3.19(c)). In Fig. 3.19(c), the end of the numerically calculated plume is elongated compared to the Domenico solution. This is the combination of two effects: the first being the result of Section 3.4.1 that Domenico plumes are generally shorter than the actual plume, and the second, the previously described effect resulting from the faster flow velocity in this zone. This second effect highlights a limitation of the velocity averaging algorithm. Because the algorithm averages the velocity along the flow path up to the advection front, $x = vt$, it cannot take into account the "extra" contaminant that has entered a zone of higher velocity due to longitudinal dispersion. Therefore, for this configuration of the flow field, the higher the longitudinal dispersion, the more mass enters the zone of higher velocity and the difference in length between the Domenico plume and the numerical plume becomes larger and larger.

It must be noted that the preceding analysis is only valid for the flow field from Fig. 3.18. However one can imagine that if the slope of the linear flow field increases, the larger the difference in plume widths. As a result, it is reasonable to expect that as the flow field becomes more and more heterogeneous, the difference between the numerical MT3DMS solution will continue to increase. From the analysis it can be concluded that if the flow velocity magnitude does not vary too much, then the Domenico solution will provide an acceptable approximation to the numerical solution.

**Curved Flow Paths**

The effect of warping on the accuracy of the Domenico solution is now examined. A simple case was constructed so that all model parameters remained constant except for the flow direction. In MT3DMS, this was achieved by using the model of the ideal case and modifying the constant head boundary condition. In the ideal model, the boundary condition at the eastern domain boundary was set to 5 m. Now, this boundary changed to a no-flow boundary and the northern boundary is set to a constant head of 5 m. The flow field generated by this setup will result in flow from the western boundary curving upwards towards the northern boundary. The flow velocity magnitude along individual groundwater flow lines is constant. Like the case of varying flow velocity, an important difference between the Domenico solution and the numerical solution occurs near the advection front. The comparison is carried out on a transient state simulation of 50 years instead of the 25 years of the analysis of varying flow velocity magnitude. The transient simulation time was increased to better show the effects of the curving flow path. Fig. 3.20 shows the results of applying a $2^{nd}$-order polynomial warp to the Domenico solution. In the figure, the flow path is shown as the gray line and the location of the source is shown as the small circle. Additionally, equipotentials of the hydraulic head are shown as the dotted lines. For these boundary conditions, the resulting velocity along the flow line shown in the figure is approximately 0.0285 m/day with a standard deviation of 0.000356 m/day (there is some slight variation along the flow line due to the use of the
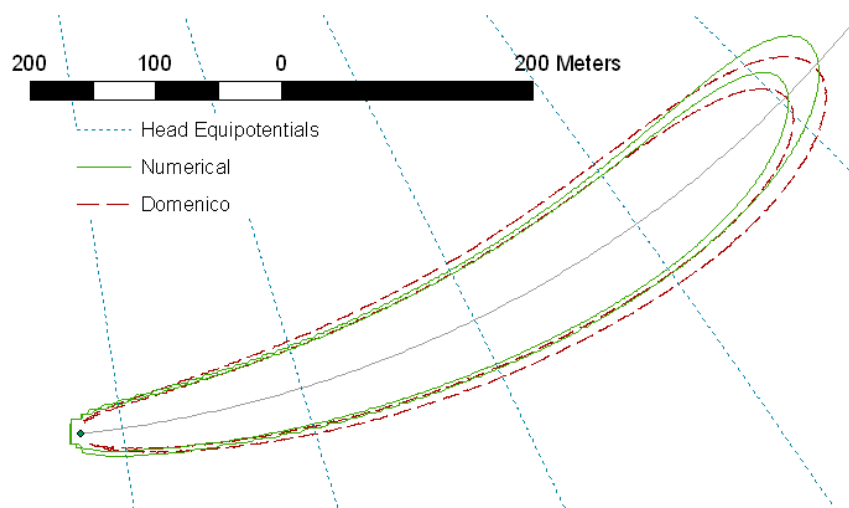
**Figure 3.20:** 0.1 and 0.01 mg/l contours for the simple warping scenario

numerical solution of MODFLOW). For this combination of parameters and the selected contours, the warped Domenico solution yields a plume that is generally slightly wider than the numerical solution except very near the source where it is narrower. An important effect occurs in the vicinity of the plume tip. It can be seen from Fig. 3.20 that the plume generated by the numerical solution is not perfectly symmetrical about the flow path. Instead, it is skewed slightly to the left. This is due to the varying flow velocity in the transverse plume direction. On the left boundary of the numerical plume, the flow velocity is slightly higher than on the right (this can be seen from the spacing of the equipotential lines), the result being the left side of the plume moves slightly faster than the right, thereby skewing the shape towards the faster flowing side. The Domenico solution on the other hand, cannot take into account this variation in flow velocity along the transverse flow direction, which can be seen from the fact that the Domenico plume is not skewed. The skewing effect occurs throughout the plume and becomes less pronounced towards the source.

An important side effect of warping that is not shown in Fig. 3.20 is the slight change in the concentration value of the Domenico solution at a given location in the plume, compared to the value at the corresponding location before warping. This is an unavoidable consequence of the warping algorithms. Recall that the warping algorithms essentially build an interpolation function to transform from warped coordinates to unwarped coordinates, which is then used in conjunction with raster resampling and value interpolation to obtain concentration values at the transformed locations. This procedure necessarily changes the concentration value of a given reference point in the plume. Fig. 3.21 shows a plot of a concentration profile along a transverse section, located at 500 m from the source, of the plume shown in Fig. 3.12. In Fig. 3.21, the concentration values of the warped plume vs. the unwarped plume are compared. As can be seen, the largest differences occur near the plume center for this particular cross section. The average percent difference over the whole cross section is 4%. Generally speaking, although the change in concentration values is small, it may become problematic in the load calculations of the Nitrate Load Estimation module,
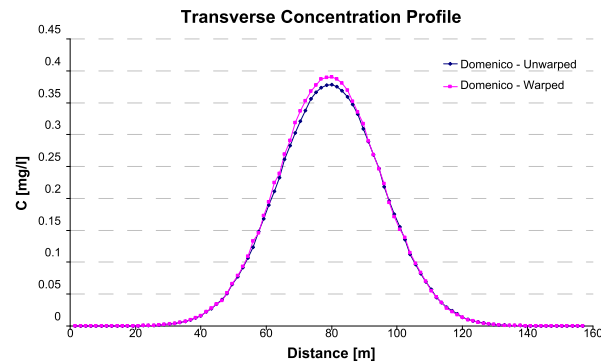
**Figure 3.21:** Comparison of the Domenico solution concentration values at a transverse cross section at 500 m before and after warping
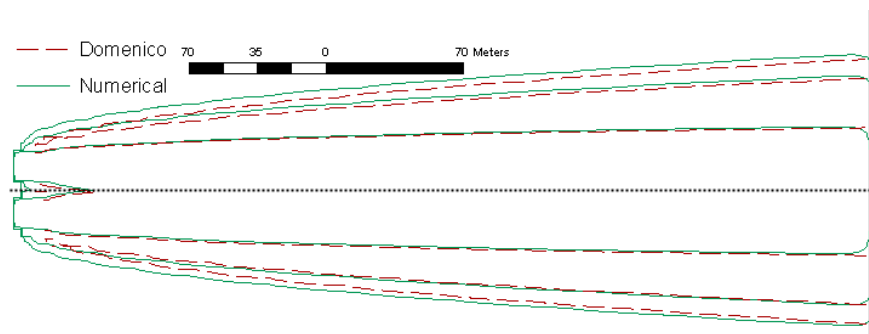


**Figure 3.22:** 0.1 and 0.01 mg/l contours for the overlapping case. No decay. The dotted line represents the cross-section of Fig. 3.23

especially if the shape of the flow path is too complex and extreme deformation of the original plume is required, or if the plume resolution is too low.

### 3.4.3   Effect of Overlapping Plumes

The effectiveness of the plume overlapping algorithm was evaluated by modifying the MT3DMS model that was used in the ideal case comparison. The model was modified by adding a second source with dimensions and initial concentration the same as the initial source. This new source was placed 24 m to the north of the first source and the region on zero dispersivity was adjusted accordingly. As in the ideal case, the Domenico plume is narrower than the MT3DMS plume at the outer contours, shown in Fig. 3.22. As before only the 0.1 and 0.01 mg/l contours are shown for clarity. In order to analyze the effect of overlapping, a concentration cross-section was constructed which was parallel to the direction of flow (east), and bisected the two plumes. The concentration profiles for the Domenico and numerical solutions are shown in Fig. 3.23. The concentration profile when the maximum concentration is taken in the overlapping region (instead of the sum) is also shown for comparison purposes. From the figure, because the curve using the sum of concentrations is a much better approximation to the numerical solution than taking the maximum,

**Concentration Cross-Section, No Decay**
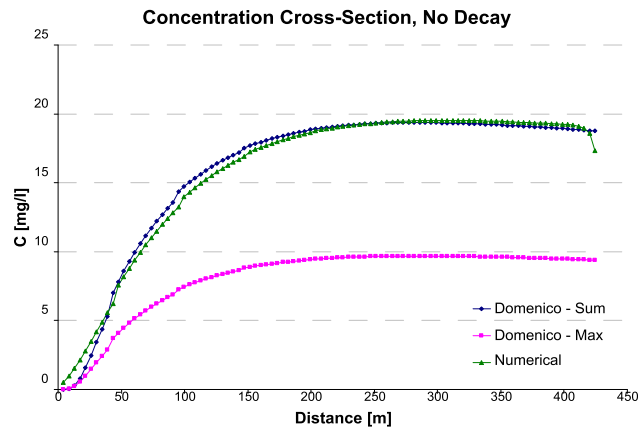


**Figure 3.23:** Concentration profile of bisecting line.

it is evident that the sum of the concentrations of the individual Domenico plumes should be taken in the overlap regions. All further discussion references the result calculated using the sum. Additionally, it can be seen that the model overestimates the concentration in a large portion of the overlapping area for the case when the concentrations are summed. The maximum overestimation is about 16%. Because the outer Domenico concentration contours are narrower than the numerical ones in the single plume case (Fig. 3.16), and also in the non-overlapping areas of Fig. 3.22, this overestimation must be due to the overestimation of the concentration in the overlapping region resulting from the inability of the Domenico solution to take into account the presence of the other plume. Very close to the source, the Domenico solution underestimates the concentrations. This is because that very close to the source, there is not much overlap of the Domenico plumes, hence the effect of summing the overlapping concentration values does not dominate over the general effect of Domenico underestimation. Towards the end of the cross-section, the behavior of the concentration profile seems to revert to the general effect of underestimation. This suggests that far from their respective sources, the two plumes merge and behave like a single plume. Examining the contours plotted in Fig. 3.22 and comparing like contours with Fig. 3.16(a), it can be seen that this is in fact the case. The effect of plume merging was also observed by Ibaraki (2001).

Intuitively, if the Domenico plumes become wider, say by increasing the decay rate (as in Fig. 3.17 for example), the amount of overestimation of concentration will increase due to the combined effects of wider Domenico plumes and increased overlapping area. This is in fact the case. For a decay rate of 0.008 day$^{-1}$, the concentration profile is shown in Fig. 3.24. The result shows that due to the wider plumes, as well as the overestimation of concentration due to the inability of the Domenico solution to take into account other plumes, the concentration in the overlapping areas is overestimated. Outside of the overlapping areas however, the concentration contours resemble the concentration contours of the single plume case (Fig. 3.17) as shown in Fig. 3.25.
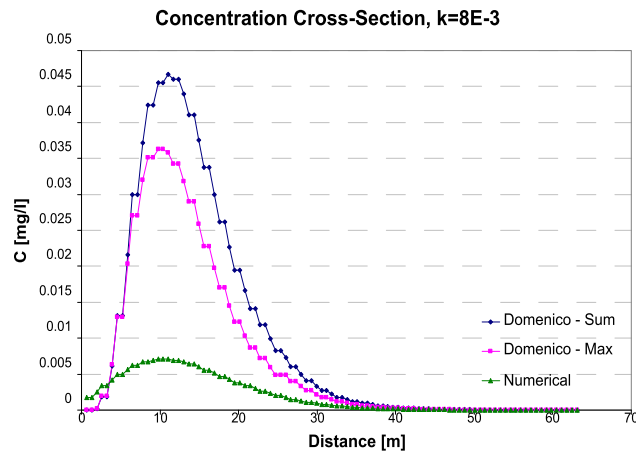
**Figure 3.24:** Concentration profile of plume bisector. Decay rate is 0.008 day$^{-1}$
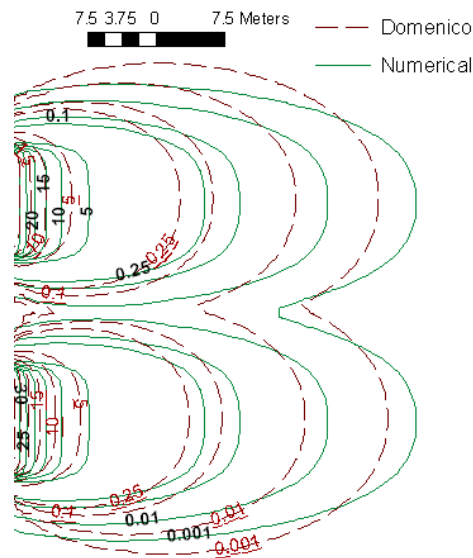


**Figure 3.25:** Concentration contours of overlapping plumes. Decay rate is 0.008 day$^{-1}$

### 3.4.4   Real-World Scenario

The model is now tested using preliminary data provided by the FDEP for two homes in the Hillcrest neighborhood in Jacksonville, Florida. Specifically, the two homes (5436 and 5476) are both located on Waterside Dr. which is in close proximity to the Lakeshore neighborhood, shown in Fig. 3.26. The model was executed for the two sites with the parameters shown in Table 3.1. While the groundwater flow velocity varied for each source, it was approximately 0.2 m/day for all the cases. The source locations were obtained from a shapefile provided by the FDEP, which file assumes that the drainfield is located in the center of the property. The parameters for the flow model were the ones used in the analysis of the NAS from Chapter 2. The value of $Y$ was estimated

**Figure 3.26:** Location of Waterside Dr. with respect to Lakeshore
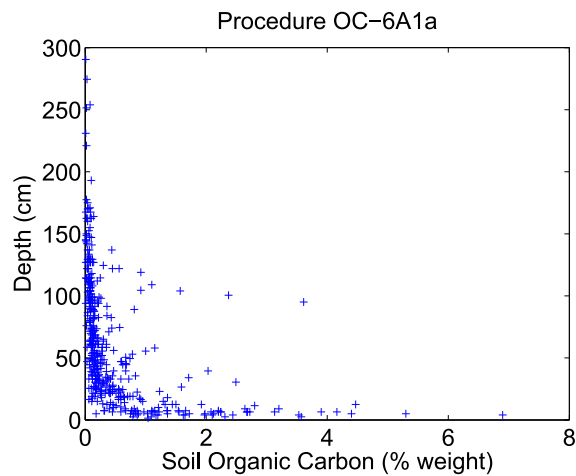


**Figure 3.27:** Soil depth plotted against percent organic carbon at that depth for soils in Florida.  Data Source:  USDA Natural Resources Conservation Service (http://ssldata.nrcs.usda.gov). Procedure 6A1a is described in detail in the NRCS Soil Survey Laboratory Methods Manual (U.S. Department of Agriculture, 2004)

**Table 3.1:** Parameters used for the Waterside Dr. test case

| | Flow Model | | |
|---|---|---|---|
| DEM | K | Porosity | Smoothing Factor |
| USGS 1/3 arcsec | 2.1336 m/day | 0.25 | 40 |

| | | | Transport Model | | | | |
|---|---|---|---|---|---|---|---|
| Y | Z | Cell Size | $\alpha_x$ | $\alpha_y$ | k | $C_0$ | $C_{thresh}$ |
| 6 m | 1.5 m | 0.4 m | 2.113 m | 0.234 m | 0.2 day$^{-1}$ | 40 mg/l | 0.0001 mg/l |



**Figure 3.28:** Modeling results superimposed on a satellite image (scale is approximate). Street numbers from top to bottom: 5476, 5436, 5428. 5428 Waterside Dr. was calculated for comparison purposes.

from the drainfield dimensions of Fig. 3.29. Ideally, the value of $Z$ should be estimated from the measured plume thickness however this information was not available in the data provided by the FDEP. If the water table is assumed to be shallow (since the study sites are so close to the river), $Z$ can be estimated from the depth of organic carbon because denitrification can only occur in the presence of organic carbon. The value of $Z$ in Table 3.1 corresponds to the depth such that approximately 95% of the data points in Fig. 3.27 are shallower than that depth. The longitudinal dispersivity, $\alpha_x$ is approximately the same value used by Davis (2000) (~7 ft).  The transverse dispersivity, $\alpha_y$, is approximately an order of magnitude lower than the longitudinal dispersivity (Gelhar et al., 1992). The value of the decay constant was based on the range of values provided by McCray et al. (2005) and adjusted to provide better agreement with field data, the value itself corresponds to the 70$^{\text{th}}$ percentile of the rates given by McCray et al. (2005), a relatively high value. The $C_0$ value corresponds to the median value of the typical ranges given by McCray et al. (2005). A cell size of 0.4 m was deemed adequate to represent each plume in sufficient detail. The results of the test run are shown superimposed on a satellite image showing the property boundaries (Fig. 3.28). In addition to the plumes, the concentration contours and flow lines are shown. The light blue polygon represents the modeled location of the river, as obtained from the GIS data. In order to make a more direct comparison, the plumes were superimposed on a schematic of each property, provided by the FDEP. The location of the plume was then adjusted so that the center of the source overlaid the center of the drainfield (assumes the flow and transport conditions remain the same at the center of the drainfield compared to the center of the property). The results are shown in Fig. 3.29.  Additionally, Fig. 3.29 shows the locations of the nitrate sample points, whose field-measured values are given in Tables 3.2 and 3.3.  From Fig. 3.29(a), it can be seen that the model can predict the value of NJ1 to within a reasonable amount as shown by the contours, however NJ3, NJ4, NJ5, and NJ6 have been overestimated. Since this home appears to have two septic tanks, it is not known what the influence of the other septic tank, located in the north part of the property, would be on the nitrate concentrations, since each septic tank appears to be for a different type of effluent. In Fig. 3.29(b), the model does not accurately predict the non-BDL concentration value.  Examining Table 3.3, it can be seen that the measurements immediately adjacent to DE2 are BDL, even though DE2 itself has a relatively high nitrate concentration. In this case, it may be necessary to better understand the plume behavior (e.g., could it be a very narrow plume?) before being able to analyze the model further.

This real-world test scenario serves to highlight the capabilities and shortcomings of the model. Although this model is a screening model, and as such can only provide ballpark estimates of nitrate concentrations, it still may be able to reproduce real-world measurements if the real-world conditions are better understood.  In the analysis above, the model was applied to the test area without site-specific model parameter values.  The results indicate that model calibration would be required to improve the model accuracy.  It should be noted however that the intended use of this modeling tool does not entail extensive data gathering for model calibration. Instead, it is meant to provide quick results that are of comparable magnitude to measured values.  The data used for a typical usage scenario would be quite similar to the data used here, i.e., physical parameter values obtained from the literature as well as approximate locations of septic tank drainfields,

(a)  5436 Waterside Dr. (to scale)          (b)  5476 Waterside Dr. (to scale)
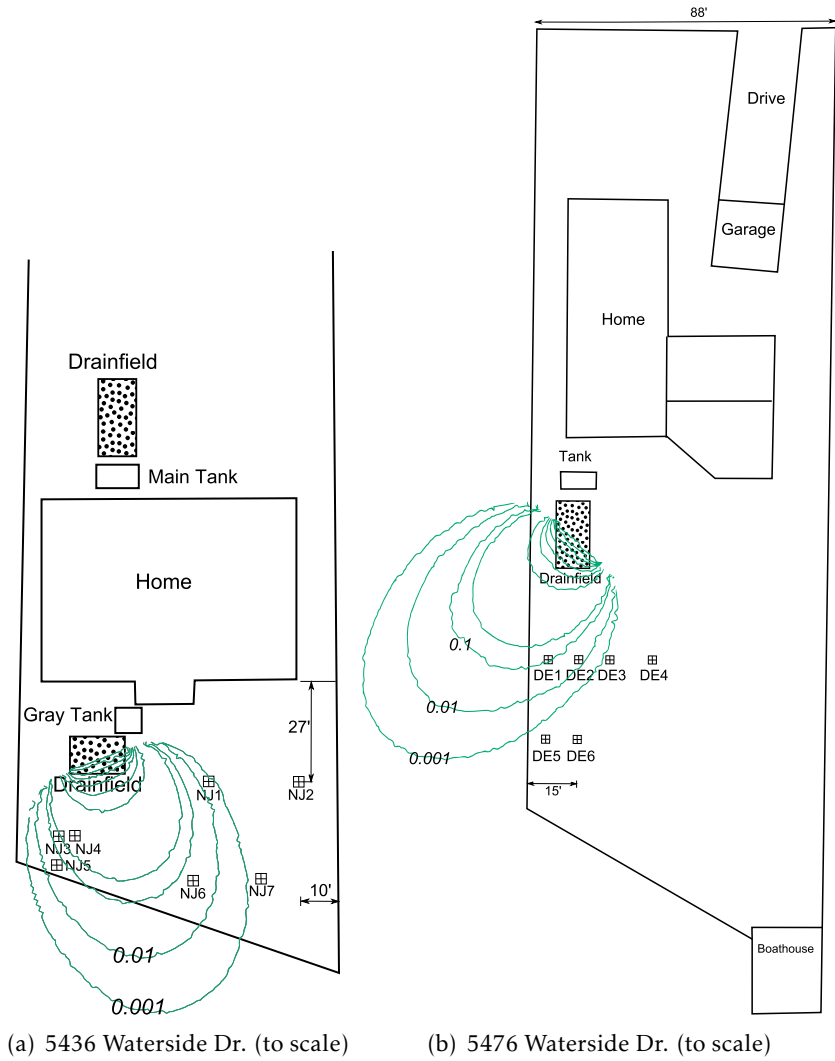
**Figure 3.29:** Test locations at Waterside Dr.

**Table 3.2:** $NO_3$ concentrations (mg/l) for 5436 Waterside Dr. BDL=Below detection limit.

| NJ1 | NJ2 | NJ3 | NJ4 | NJ5 | NJ6 | NJ7 |
|-----|-----|-----|-----|-----|-----|-----|
| 0.005 | BDL | BDL | BDL | BDL | BDL | BDL |

**Table 3.3:** $NO_3$ (mg/l) concentrations for 5476 Waterside Dr. BDL=Below detection limit.

| DE1 | DE2 | DE3 | DE4 | DE5 | DE6 |
|-----|-----|-----|-----|-----|-----|
| BDL | 1.9 | BDL | BDL | BDL | BDL |

and water body shorelines. As a result of these rough estimates, it is not advisable to apply the model to individual cases. Instead, the model should be applied to a larger number of sources simultaneously, and the results should be taken as a bulk estimate, thereby bringing expected model performance in line with the concept of a screening model.

### 3.4.5   Conclusions

From the analysis given in the previous sections, several conclusions are reached regarding the accuracy of the Domenico solution under ideal and non-ideal conditions. Additionally, the effects of the various approximations used in the model for dealing with scenarios that the Domenico solution cannot handle, are described. First, from Section 3.4.1, it can be said that in general, this Domenico-based model underestimates concentration values near the plume centerline, the degree depending on the specific case. For the tested cases, the underestimation can be over 60% in the worst case. Srinivasan et al. (2007) and West et al. (2007) also reached similar conclusions. The results also show that the conclusion reached by Srinivasan et al. (2007)

> "...plumes predicted by the modified-Domenico solution are always wider than the actual plumes."

is not always true as it depends on: a), the specific parameters (dispersivity, flow velocity, decay rate, etc.), b), the contour level selected for comparison and c), the location along the plume at which the comparison is made. In addition to the controlling factors, as determined by Srinivasan et al. (2007), which affect plume width, it was found that the decay rate can significantly affect the width of the plume, with higher decay rates generating plumes that are wider than the equivalent MT3DMS plumes.

Regarding the performance of the model under non-ideal conditions, a general conclusion on overestimation or underestimation of the Domenico solution cannot be reached, as the prediction depends, to a large degree, on the nature and extent of the heterogeneities. In the case of spatially heterogeneous groundwater flow velocity which increases linearly downstream, the Domenico solution behaves as expected. That is, when the groundwater flow velocity is slower than the average velocity, the Domenico plume is narrower than a plume calculated via numerical methods for the same conditions. Similarly, when the actual flow velocity is faster than the average, the Domenico plume is wider.

For a simple curved flow path, warping serves to map the straight plume of the Domenico solution to a curved flow path. The warping of the plume compares favorably to a numerical simulation of a curved plume for the simple case of a smooth arc. There is a noticeable difference in the shape near the tip of the plume due to the differing flow velocities across the plume transverse. The warping algorithms, due to their nature, inevitably change the concentration values of the plume at any given reference point within the plume. The changes are generally small, around 4%

on average, and should not be very problematic if the plume resolution is high enough and the amount of warping is not too extreme.

When dealing with multiple Domenico plumes that overlap, summing the concentration values in the overlapping region is more appropriate than taking the maximum value. Taking the sum of the concentrations slightly overestimates the concentration in the overlapping area and compares favorably with the numerical solution. However in the test case with decay, the amount of overestimation was much greater due to the increased effect of having too much overlap due to the wider plumes generated by the Domenico solution in this case.

The only general conclusions that can be stated regarding the performance of the model under heterogeneous conditions are that: a), the more the system deviates from the underlying assumptions of the Domenico solution, the greater the error of the model and b), the error cannot be quantified a priori due to the infinite number of possible ways a system can be made heterogeneous and the dependence of the model response on the specific configuration of the heterogeneities.

Based on the results from Section 3.4.4, when applying the model to a real-world scenario, it is difficult to perform a direct comparison with real world data without first understanding the behavior of the real system. Additionally, without performing calibration with actual site-specific parameters, a direct comparison may or may not yield results comparable to the field data. Another factor that complicates the comparison for the tested case is the fact that the data represents a snapshot in time where transient effects may be significant. These transient effects may be significant because the model considers only steady-state conditions. Since this model is a screening model, the results should instead be taken as a bulk estimate.

# 4

## Nitrate Load Estimation Module

The Nitrate Load Estimation module deals with the calculation of nitrate load to a surface water body (or bodies). The calculations required by this module are spread over the Transport and Nitrate Load Estimation Modules . The reason being that denitrification is modeled by the Transport module and it therefore becomes convenient, from an implementation point of view, to carry out some of the calculations required for load estimation in the Transport module. The parts of the nitrate load estimation that are carried out by the Transport module are the estimation of the input mass rate and the rate of mass removal due to denitrification.

After a discussion of the Nitrate Load Estimation module, a brief verification is carried out using MT3DMS as a reference.

## 4.1 Load Calculations

The approach taken in this model to calculate the nitrate load, is a mass balance approach. The general mass balance expression is given by

$$\frac{dm_{in}}{dt} = \frac{dm_{out}}{dt} + \frac{dm_{aq}}{dt} + \frac{dm_{dn}}{dt}$$
$$M_{in} = M_{out} + M_{aq} + M_{dn}$$

(4.1)

where $M_{in}$ in the mass inflow rate, $M_{out}$ is the mass outflow rate at a water body, $M_{dn}$ is the mass rate of removal due to denitrification, and $M_{aq}$ is the rate of change of mass storage in the aquifer. All units are [M/T]. The quantity of interest in our model is the load to the target water body, i.e. $M_{out}$. When the system reaches the steady-state, the total mass inflow rate equals the total mass outflow rate and $M_{aq}$ vanishes. Then Eq. (4.1) becomes

$$M_{in} = M_{out} + M_{dn}$$

(4.2)

solving for $M_{out}$,

$$M_{out} = M_{in} - M_{dn} \tag{4.3}$$

from which $M_{out}$ can be easily calculated if $M_{in}$ and $M_{dn}$ are known. In Eq. (4.2), the reason $M_{aq}$ becomes zero is analogous to filling a bathtub with the drain unplugged. After a while, the amount of water flowing into the bathtub and the amount flowing out are the same, the effect being that the water level in the bathtub remains the same.

Mass balance based approaches are common in simplified screening models (McCray et al., 2005). Although the transport module is capable of transient state simulations, their use is not supported by the mass balance approach to calculating load described here. In order to calculate loads for a transient state case, an approach similar to the one used by Andres (1992) must be employed. In the approach by Andres (1992), the contaminant flux to the water body is calculated on a cell-by-cell basis for each cell in contact with the boundary of the water body. The downside to this approach is the increased complexity and subsequently longer computation times.

## 4.2   Conceptual Model

The load calculation model is based on the principle of mass balance in the steady-state. In order to calculate the nitrate load from Eq. (4.3), it is necessary to calculate the $M_{in}$ and $M_{dn}$ terms. The calculation of the $M_{dn}$ term, i.e., the mass removal rate due to denitrification, is straightforward. Because denitrification is modeled as a first-order decay process, we can use the definition of first order decay to calculate the rate of denitrification for every cell, in the plume. The rate is then multiplied by the volume of solute to obtain the mass rate.

$$M_{dn} = \sum_i k C_i V_i \theta \tag{4.4}$$

where $k$ [$\text{T}^{-1}$] is the rate constant for the plume, $C_i$ [M/L$^3$] is the concentration in cell $i$, $V_i$ [L$^3$] is the volume of cell $i$ and $\theta$ [-] is the soil porosity.

The calculation of the mass input rate, $M_{in}$, is more complicated. The input load calculation accounts not only for mass input due to advection but also through dispersion. Fig. 4.1 shows a schematic of the mass input from the Domenico source plane. $M_{in}$ consists of two terms: the mass input due to the advective flow of groundwater and the component due to dispersive flux across the boundary resulting from the concentration gradient across the boundary.

$$M_{in} = M_{adv} + M_{dsp} \tag{4.5}$$

The advective term is calculated using the volume of water that flows across the interface in unit time, multiplied by the solute concentration.
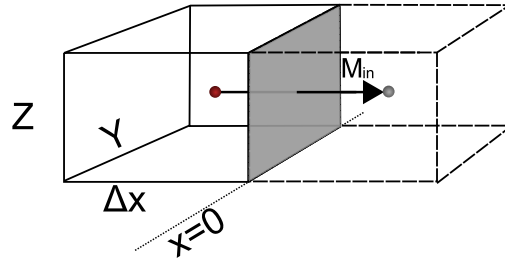
$$M_{adv} = C_0 Y Z \theta v \tag{4.6}$$

**Figure 4.1:** Mass flux from the Domenico source plane (shaded area) into the system.

$C_0$ [M/L$^3$] is the concentration of the source plane, $Y$ and $Z$ [L] are the dimensions of the source plane, $v$ [L/T] is the seepage velocity, and $\theta$ [-] is the porosity. The dispersion term is calculated by assuming dispersion is governed by Fick's Law (Freeze and Cherry, 1979). The dispersion through the source plane in Fig. 4.1, is written as

$$M_{dsp} = -D_{xx}YZ\theta\frac{\partial C}{\partial x} \tag{4.7}$$

where $D_{xx}$ [L$^2$/T] is the component of the dispersion tensor along the x-direction; the remaining parameters are the same ones from Eq. (4.6). There are two things to note regarding Eq. (4.7). First, the dispersion parameter is actually a second order tensor in three dimensions, represented by a $3 \times 3$ matrix. Because the direction of principal flow has been aligned with the coordinate system, and the flow is assumed to be only in the x-direction (in accordance with the Domenico solution), all z-terms and x-y cross terms vanish, leaving only the $D_{xx}$ term. If we were to consider dispersive mass transport from the remaining 5 faces of the cube in Fig. 4.1 (as does MT3DMS), then the $D_{yy}$ component would need to be considered. The second point to note regarding Eq. (4.7), is the presence of a negative sign. The negative sign serves a similar purpose as the one in Darcy's Law (Eq. (1.4)). Because the mass flows from high concentration to low concentration, the negative sign ensures that movement in this direction is positive. Disregarding molecular diffusion, the dispersion coefficient is calculated as (Freeze and Cherry, 1979)

$$D_{xx} = \alpha_x v \tag{4.8}$$

where $\alpha_x$ [L] is the medium's dispersivity in the x-direction and $v$ [L/T] is seepage velocity as before. Putting everything together

$$M_{in} = C_0 YZ\theta v - \alpha_x v YZ\theta\frac{\partial C}{\partial x} \tag{4.9}$$

Using Eq. (3.6) to calculate the partial derivative in $C$, $M_{in}$ can be expressed analytically as

$$M_{in} = C_0 YZ\theta v\left(\frac{1 + \sqrt{1 + \frac{4k\alpha_x}{v}}}{2}\right) \tag{4.10}$$

The derivation for Eq. (4.10) is provided in the following section.

### 4.2.1   Derivation of the Analytical Mass Input Rate Expression

The expression for $M_{in}$ can be derived using Eq. (3.6) as follows. For reference purposes, the two-dimensional, steady-state Domenico solution (Eq. (3.6)) is reproduced here:

$$C(x, y) = \frac{C_0}{2} F_1(x) F_2(y, x) \tag{4.11}$$

where

$$F_1 = exp\left[ \frac{x}{2\alpha_x}\left( 1 - \sqrt{1 + \frac{4k\alpha_x}{v}} \right) \right] \tag{4.12a}$$

$$F_2 = erf\left( \frac{y + Y/2}{2\sqrt{\alpha_y x}} \right) - erf\left( \frac{y - Y/2}{2\sqrt{\alpha_y x}} \right) \tag{4.12b}$$

The mass input rate from the source plane is given in Eq. (4.13) below:

$$M_{in} = C_0 Y Z \theta v - \alpha_x Y Z \theta v \frac{\partial C}{\partial x}\bigg|_{x=0} \tag{4.13}$$

Where $Y$ and $Z$ are the dimensions of the source plane. The mass input rate includes the contributions from both advective and dispersive flow. The $\partial C/\partial x$ term can be calculated from Eq. (4.11) as shown below. The remaining terms in Eq. (4.9) are specified parameters or are otherwise known.

Differentiating Eq. (4.11) (using the chain rule) and evaluating it at the location of the source plane,

$$\frac{\partial C}{\partial x}\bigg|_{x=0} = \frac{C_0}{2} F_2 \frac{\partial F_1}{\partial x}\bigg|_{x=0} + \frac{C_0}{2} F_1 \frac{\partial F_2}{\partial x}\bigg|_{x=0} \tag{4.14}$$

Calculating $\partial F_1/\partial x$ is straightforward:

$$\frac{\partial F_1}{\partial x}\bigg|_{x=0} = \frac{1 - \sqrt{1 + \frac{4k\alpha_x}{v}}}{2\alpha_x} \tag{4.15}$$

The calculation of $\partial F_2/\partial x$ requires multiple steps. For illustration, the derivative of the first $erf$ term in $F_2$ (call it $F_{2(1)}$) is calculated (the second $erf$ term is identical with only a change in sign of the $Y/2$ term). Let

$$g(x) = \frac{y + Y/2}{2\sqrt{\alpha_y x}} \tag{4.16}$$

The error function, defined in terms of g(x) is

$$F_{2(1)}(g(x)) = \frac{2}{\sqrt{\pi}} \int_0^{g(x)} e^{-t^2} dt \qquad (4.17)$$

Then, using Leibniz's integral rule to differentiate Eq. (4.17) and substituting $g(x)$,

$$\begin{aligned}
\frac{\partial F_{2(1)}(g(x))}{\partial x} &= \frac{2}{\sqrt{\pi}} e^{-g(x)^2} \frac{d}{dx} g(x) \\
&= \frac{2}{\sqrt{\pi}} exp\left[-\left(\frac{y+Y/2}{2\sqrt{\alpha_y x}}\right)^2\right]\left[-\frac{y+Y/2}{2\sqrt{\alpha_y}} \frac{x^{-3/2}}{2}\right] \\
&= -\frac{1}{\sqrt{\pi}} exp\left[-\frac{(y+Y/2)^2}{4\alpha_y x}\right]\left[-\frac{y+Y/2}{2\sqrt{\alpha_y}} \frac{1}{x^{3/2}}\right]
\end{aligned} \qquad (4.18)$$

A similar expression is obtained for $\partial F_{2(2)}/\partial x$ (the derivative of the second $erf$ term in $F_2$):

$$\frac{\partial F_{2(2)}(g(x))}{\partial x} = -\frac{1}{\sqrt{\pi}} exp\left[-\frac{(y-Y/2)^2}{4\alpha_y x}\right]\left[-\frac{y-Y/2}{2\sqrt{\alpha_y}} \frac{1}{x^{3/2}}\right] \qquad (4.19)$$

In the limit as $x$ goes to zero, both Eqs. (4.18) and (4.19) go to zero therefore,

$$\left.\frac{\partial F_2}{\partial x}\right|_{x=0} = 0 \qquad (4.20)$$

Additionally, in the limit as $x$ goes to zero,

$$\left.F_1\right|_{x=0} = 1 \qquad (4.21)$$

and

$$\left.F_2\right|_{x=0} = 2 \qquad (4.22)$$

Then, substituting Eqs. (4.15) and (4.20) to (4.22) into Eq. (4.14),

$$\left.\frac{\partial C}{\partial x}\right|_{x=0} = C_0 \frac{1 - \sqrt{1 + \frac{4k\alpha_x}{v}}}{2\alpha_x} \qquad (4.23)$$
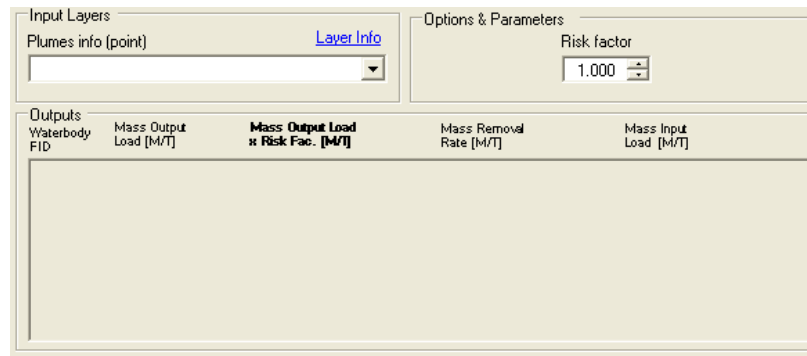
**Figure 4.2:** User interface of the Nitrate Load Estimation module.

Finally, substituting Eq. (4.23) into the mass input rate equation (Eq. (4.9)) and simplifying,

$$
\begin{aligned}
M_{in} &= C_0 YZ\theta v - \alpha_x YZ\theta v \left( C_0 \frac{1 - \sqrt{1 + \frac{4k\alpha_x}{v}}}{2\alpha_x} \right) \\
&= C_0 YZ\theta v \left( 1 - \frac{1 - \sqrt{1 + \frac{4k\alpha_x}{v}}}{2} \right) \\
&= C_0 YZ\theta v \left( \frac{1 + \sqrt{1 + \frac{4k\alpha_x}{v}}}{2} \right)
\end{aligned}
\tag{4.24}
$$

## 4.3   Algorithm and Implementation

The output of the Nitrate Load Estimation module is the load to the target water bodies, taking into account the mechanism of denitrification. Unlike the Groundwater Flow and Transport modules, the Nitrate Load Estimation module does not output a shapefile or raster. The output is a single number corresponding to the load to each target water body. The interface to the module is shown in Fig. 4.2.

### 4.3.1   Input and Output

The inputs to the model (using the nomenclature from Fig. 4.2) are:

1. **Plumes Info** Specifies the associated plumes information shapefile generated by the Transport module.

Model options are:
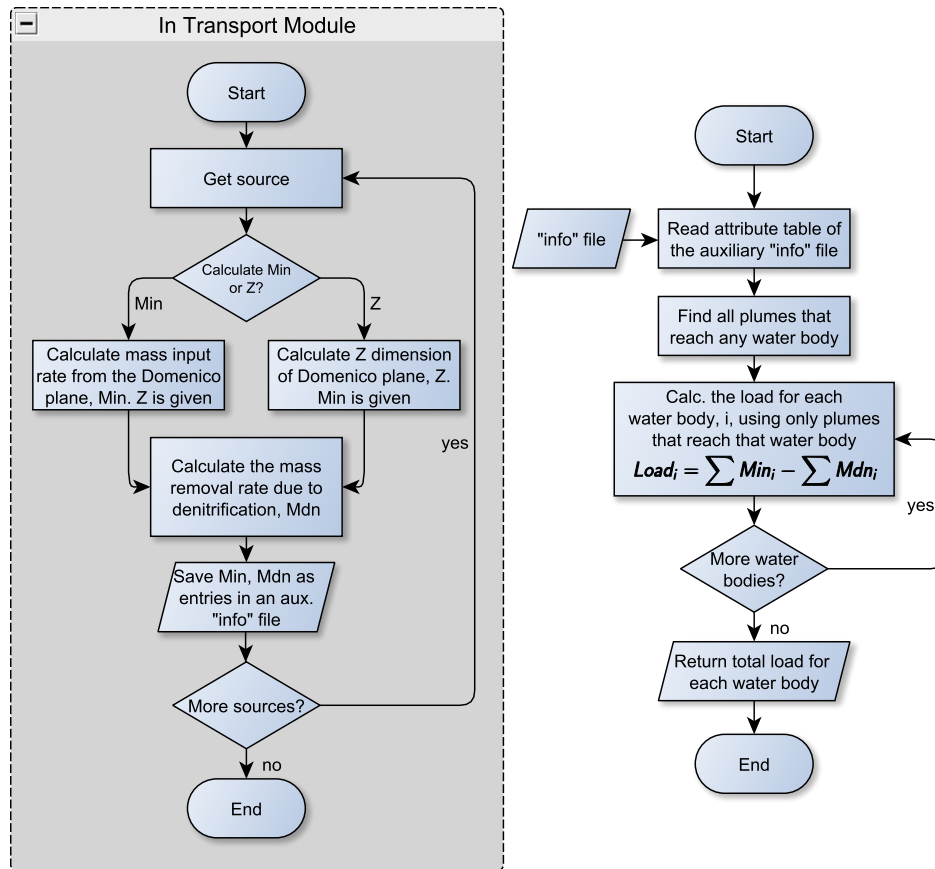
**Figure 4.3:** Logic flow of the Nitrate Load Estimation module. $Load_i$ is the nitrate load to the $i^{th}$ water body. $Min_i$ is the mass input rate due to sources whose flow path intersects water body $i$. $Mdn_i$ is the mass removal rate due to denitrification due to sources whose flow path intersects water body $i$.

1. **Risk Factor** A multiplicative factor used to multiply the calculated load to each water body. Used as a safety margin to compensate for model uncertainty. This factor is determined by the user.

### 4.3.2 Process

Fig. 4.3 shows the logic flow of the algorithm implementing the conceptual model of Section 4.2. The first part of the process involves the calculation of the input load, $M_{in}$ or $Z$ (depending on which one was specified) and the nitrate mass removal rate $M_{dn}$. These calculations are carried out by the Transport module (left side of Fig. 4.3) and proceed as follows. For each nitrate source, the input load or the value of $Z$ is calculated depending on which one was specified by the user (see

the Domenico Bdy. parameter in Section 3.3). If $Z$ is given, the input load is calculated by

$$M_{in} = C_0 Y Z \theta v \beta \left( \frac{1 + \sqrt{1 + \frac{4k\alpha_x}{v}}}{2} \right) \qquad (4.25)$$

Note that Eq. (4.25) is identical to Eq. (4.10) except an extra parameter, $\beta$ has been added. The parameter $\beta$ corresponds to a factor that converts units of volume obtained from the units of $Y$, $Z$, and $v$ to the volume units of $C_0$. For example, if the length units of $Y$, $Z$ and $v$ are meters and the units of $C_0$ are mg/l, then $\beta$ should be set to 1000 since 1 cubic meter equals 1000 liters. If $M_{in}$ is specified instead, $Z$ is obtained by solving for it in Eq. (4.25). The next part is to calculate the mass removal rate due to denitrification, $M_{dn}$. The calculation of $M_{dn}$ involves several steps. First, the volume of solute in cell, $i$, of each plume is calculated via

$$V_{aq}^i = V_i \beta \theta \qquad (4.26)$$

where $V_i$ [L$^3$] is the cell volume (each cell having dimensions $\Delta\ell \cdot \Delta\ell \cdot Z$ where $\Delta\ell$ is the Plume Cell Size parameter of the Transport module and $Z$ is the height of the source plane), and $\theta$ [-] is the soil porosity. The $\beta$ factor is the same as the volume conversion factor mentioned previously (it has the same value as well), however its interpretation is slightly different. In ArcGIS, the linear units of the raster file may be different from raster to raster, thus the need for a factor to convert between cell volume units and concentration volume units. For example, if the concentration, $C_0$, is provided in units of mg/l and the raster has linear units of meters, then $\beta$ should be set to 1000 since one cubic meter equals 1000 liters. $V_{aq}^i$ will then have units of liters. After the calculation of $V_{aq}^i$, the denitrification rate is calculated, again on a cell-by-cell basis

$$R_{dn}^i = k C_i \qquad (4.27)$$

where $C_i$ [M/L$^3$] is the concentration in cell $i$ and $k$ [T$^{-1}$] is the decay constant (obtained from the Plumes Info file). Multiplying the previous two equations and summing over all cells, $M_{dn}$ from Eq. (4.4) is obtained. Cells corresponding to the source plane are not counted in the $M_{dn}$ calculation, similar to the behavior of MT3DMS which does not include constant concentration cells when calculating mass balances. Note that although the flow model supports a fully heterogeneous porosity, the Nitrate Load Estimation module supports only a constant porosity for each plume (the porosity can vary from plume to plume). Similar to the treatment of spatially heterogeneous in the Domenico solution, the porosity is averaged along the flow line, using an arithmetic mean instead of a harmonic mean used for averaging the velocity. Because of this averaging, the calculated porosity for a given plume can be considered as a representative value for that plume. The calculation of $M_{in}$ and $M_{dn}$ is repeated for each plume and the results saved in the auxiliary "info" file.

The Nitrate Load Estimation Module reads the "info" file and calculates the load for each water body as shown on the right side of Fig. 4.3. For any particular water body, the total input load from septic tanks whose flow paths intersect that water body is calculated by summing the $M_{in}$
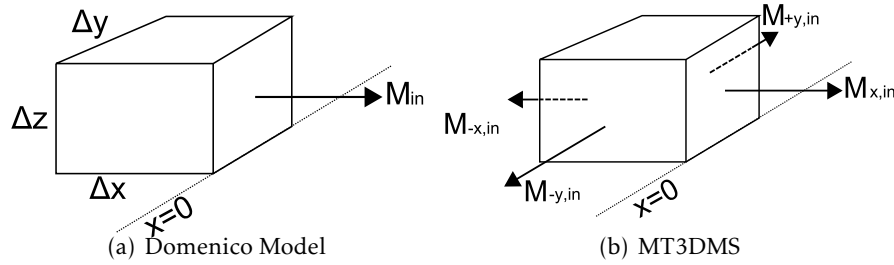
**Figure 4.4:** Source fluxes considered by our Domenico-based model compared to the ones considered by MT3DMS. The box represents a single simulation cell.

values for each corresponding source. If a flow path reaches the water body but the plume does not, it is assumed that 100% denitrification has taken place and the $M_{dn}$ value for that plume is set to $M_{in}$. If the plume does reach the water body, the $M_{dn}$ value for that plume is taken from the "info" table. After summing the $M_{in}$ and $M_{dn}$ values for each source whose path terminates at a particular water body, a simple subtraction is carried out to calculate $M_{out}$, which corresponds to the load to the target water body, in units of mass per time; the mass units correspond to the same mass units used for $C_0$ and the time units correspond to the same time units used for hydraulic conductivity and the decay constant.

## 4.4 Analysis of Model Performance

The performance of the load calculations is evaluated using the same synthetic models as the ones used for testing the transport module in Section 3.4 so as to keep results comparable. In this case, although it is the load values that are of interest, direct comparison with the load calculated from MT3DMS is complicated by the numerical approach taken by MT3DMS as well as the slightly different source conditions. As a result, MT3DMS load calculations are not *exactly* comparable to Domenico-based load calculations. This was not a important issue for the comparison carried out in Section 3.4 because far from the source, the differences in the source geometry between the Domenico model and the numerical model become less important. The source conditions become important here because the calculation of the input load involves concentrations in the immediate vicinity of the source cells. The main difference in the source conditions between the Domenico solution and the MT3DMS solution is the fact that MT3DMS considers dispersive flux from the faces of the source cells in the -x, +y and -y directions, shown in Fig. 4.4, while this Domenico based model cannot. In order to make the MT3DMS solution more comparable, flux in the -x direction was eliminated with the method described in Section 3.4.1. Due to the nature of the finite-difference evaluations of MT3DMS, the +y and -y dispersive mass transfer can be reduced but cannot be eliminated completely, without affecting the mass flux from the +x face. This may be remedied by setting the MT3DMS cell to inactive, which renders the model not equivalent to the one from Section 3.4.1. The other issue hindering direct comparison is the way in which MT3DMS

calculates mass balances. In MT3DMS, mass balances are calculated directly from fluxes into and out of each cell, at every time step, while in this algorithm, output load is calculated using Eq. (4.3), which may result in differences in the results.

### 4.4.1   Comparison Using Domenico-Generated Plumes

In order to evaluate the performance of the method of calculating loads described in Sections 4.2 and 4.3, the algorithm was applied to Domenico plumes and the results compared with similar plumes generated with MT3DMS. The parameters used to generate the plumes were the same as in Section 3.4.1. Again, to approximate the steady-state condition, the simulation time was set to 200 years, as explained in Section 3.4.1. The only thing that was varied in the simulations was the decay rate $k$. Several decay rates were selected, so as to a), produce both a plume that does not reach the water body and one that does and b), be representable with the same cell size used in Section 3.4.1 (4 m). Note: in the simulation with $k = 8 \times 10^{-3}$, the plume does not reach the water body (i.e, the mass flux at the boundary according to MT3DMS is insignificant) while the other plumes do. This fact is reflected in the MT3DMS-calculated values of the output load.

**Single Plume**

First, the case of a single isolated plume is analyzed. The comparison of the model to the results obtained fro MT3DMS are shown in Table 4.1. From the table, it is apparent that there are some differences in the results obtained.

**Table 4.1:** Load values calculated directly from MT3DMS and using the algorithm from Section 4.3 using Domenico plumes. Load values in [kg/day]. Cell size for all cases is 4 m.

|           | k    | Input Mass Load | Denitrification Mass | Output Load |
|-----------|------|-----------------|----------------------|-------------|
| MT3DMS    | 0    | 5.2735E-3       | 0                    | 5.2735E-3   |
|           | 1E-5 | 5.2778E-3       | 9.5842E-4            | 4.3194E-3   |
|           | 8E-3 | 6.6247E-3       | 6.6247E-3            | 1.0480E-10  |
|           |      |                 |                      |             |
| Algorithm | 0    | 5.2267E-3       | 0                    | 5.2267E-3   |
|           | 1E-5 | 5.1027E-3       | 9.2953E-4            | 4.1731E-3   |
|           | 8E-3 | 7.7620E-3       | 4.5451E-3            | 3.2169E-3   |

In regards to the differences in the input loads in Table 4.1, the cases of $k = 0$ and $k = 1E-5$ are comparable to the values generated by MT3DMS. The value of the input load for $k = 8E-3$ is notably higher for our model. These differences are due to the use of the analytical solution to find the concentration gradient (Eq. (4.10)). In the $k = 0$ and $k = 1E-5$ cases, the advection

term dominates over the dispersion term in Eq. (4.10) due to the relatively small value of $k$. In the MT3DMS results for those two cases, the dispersion term likely has a higher contribution to the input load than our result due to the numerical approximation of the solution by MT3DMS (i.e., discretization error). For the $k = 8E-3$ result, the increased $k$ value causes the dispersion term in Eq. (4.10) to have a larger contribution to the input load. In principle if the MT3DMS cell size is decreased, the input load value should tend towards the value given by Eq. (4.10), due to the fact that Eq. (4.10) is an analytical solution to the boundary value problem approximated by MT3DMS; this is demonstrated in Section 4.4.2.

Concerning the differences in the denitrification mass, our model generally predicts a lower value than the MT3DMS results. The differences are most likely caused by the underestimation of plume concentrations by the Domenico solution (the under-estimation of concentration by the Domenico solution was established in Section 3.4.1). A second contributing factor to the underestimation is the cell size used (discussed in Section 4.4.2). The largest differences for the input mass load and denitrification mass occur for the case of $k = 8 \times 10^{-3}$ and are approximately 17% (MT3DMS underestimates input load) and 30% (our model underestimates denitrification mass) respectively.

The plume with $k = 8 \times 10^{-3}$ does not reach the simulated water body, however the results from the Domenico solution for that case indicates there is a load when in fact there should not be according to MT3DMS. This is problematic, although this conservative estimation is preferable to underestimating the load. The software handles this situation by checking if the plume has reached the water body. If it has not, it sets the denitrification mass equal the input load, regardless of the actual calculated vale of the denitrification mass.

**Overlapping Plumes**

The analysis is now repeated except using the case of overlapping plumes of Section 3.4.3. The case of $k = 0$ is omitted, since the behavior is the same as the single plume case. From Table 4.2,

**Table 4.2:** Load values calculated directly from MT3DMS and using the algorithm from Section 4.3 using overlapping Domenico plumes. Load values in [kg/day]. Cell size for all cases is 4 m.

|  | k | Input Mass Load | Denitrification Mass | Output Load |
|---|---|---|---|---|
| MT3DMS | 1E-5 | 10.4264E-3 | 1.8932E-3 | 8.5332E-3 |
|  | 8E-3 | 12.4968E-3 | 12.4968E-3 | 1.0487E-10 |
| Algorithm | 1E-5 | 10.2054E-3 | 1.8591E-3 | 8.3464E-3 |
|  | 8E-3 | 15.5245E-3 | 9.0902E-3 | 6.4343E-3 |

it is apparent that the MT3DMS values for the loads are slightly less than double the values

of the single plume case, due to the interaction between plumes. As in the single plume case, our model results are comparable in the $k = 1E - 5$ case however in the $k = 8E - 3$ case, the input load remains noticeably higher (due to the use of the analytical solution and the increased contribution of the dispersion term in Eq. (4.10) due to a higher $k$ value). Similar to the single plume case, the denitrification mass remains underestimated due to the general underestimation of the concentration distribution of the Domenico solution, in combination with the numerical error introduced by discretizing the domain. This underestimation of the denitrification mass results in a conservative estimate of the output nitrate load. The underestimation of the denitrification mass by our model reaches 27% for this case.

### 4.4.2   Effect of Grid Resolution

As mentioned previously, the input load and denitrification mass values calculated by MT3DMS are dependent on the grid resolution. In principle, as the cell size is decreased, those values should more closely represent the "true" values. For our model, the input load does not depend on the grid size since the analytical solution presented in Eq. (4.10) is used. However, the denitrification mass in our model does depend on the grid resolution since the domain is discretized and the Domenico solution evaluated over the discretized domain. As in MT3DMS, as the cell size is decreased, the denitrification mass approaches the "true" value determined by the Domenico solution. Note that since the Domenico solution is an approximate analytical solution, the values of the input load and denitrification mass differ from the actual values.

Table 4.3 shows the effect of decreasing the grid resolution for the $k = 8E - 3$ case for both the MT3DMS results and our model. For the 1 m grid size case, the MT3DMS-calculated input

**Table 4.3:** Load values calculated directly from MT3DMS and using the algorithm from Section 4.3 using Domenico plumes, higher grid resolution. Load values in [kg/day]. Grid size 4 m except where indicated.

|           | k      | Input Mass Load | Denitrification Mass | Output Load |
|-----------|--------|-----------------|----------------------|-------------|
| MT3DMS    | 8E-3   | 6.6247E-3       | 6.6247E-3            | 1.0480E-10  |
|           | 8E-3*  | 7.4718E-3       | 7.4718E-3            | 4.6082E-11  |
| Algorithm | 8E-3   | 7.7620E-3       | 4.5451E-3            | 3.2169E-3   |
|           | 8E-3*  | 7.7620E-3       | 6.8415E-3            | 9.2050E-4   |

*1 m grid size

load has now moved closer to the analytical value calculated by our model, as expected. For the 1 m grid size case, the denitrification mass value calculated by our model has increased. This is the expected result since a higher grid resolution results in a more accurate evaluation of the Domenico solution. However, our model still underestimates the denitrification mass resulting in

an overestimation of the output load (recall that the plume in the $k = 8E-3$ case does not reach the water body, therefore the output load should theoretically be zero). When using a 1 m resolution, the difference between the MT3DMS input mass and denitrification mass compared to our model are 4% (MT3DMS underestimates input load) and 8% (our model underestimates denitrification mass) respectively

### 4.4.3   Conclusions

Tests indicate that for low denitrification rates, the model approximates the results from an MT3DMS simulation well. For higher denitrification rates, the underestimation of the concentration by the Domenico solution becomes evident. However, the estimation of the input load is generally better. Additionally, due to the Domenico underestimation of concentration, in combination with the discretized domain over which the Domenico solution is evaluated, a nitrate load to a test target water body is indicated for the case of $k = 8 \times 10^{-3}$, even though there should be no such load since the plume does not actually reach the water boundary according to MT3DMS. Although this underestimation of denitrification mass can be reduced by increasing the grid size, the underestimation remains. In cases where it is known that the plume does not reach the water body the software assumes that 100% denitrification has taken place and the input load equals the mass removal rate due to denitrification.

*5*

<div style="background:#d9d9d9">

## Closing Remarks

</div>

## 5.1 Conclusions

A simplified GIS-based model to estimate nitrate fate and transport from OWTS in the surficial aquifer is developed and implemented within the GIS system ArcGIS, using the VB.NET programming language. The model is implemented as three submodules: the flow, transport, and denitrification modules.

### 5.1.1 Flow Module

The Groundwater Flow module implements the portion of the model dealing with groundwater flow. The amount of input data to the flow model is reduced by making several simplifying approximations. These approximations are:

- The shape of the water table is a subdued replica of the topography.

- Flow conditions are in the steady-state.

- Flow occurs only in the surficial aquifer.

- The surficial aquifer consists solely of an unfractured, porous medium such that Darcy's Law is applicable.

- The Dupuit approximation is valid. In other words, vertical flow can be ignored, and the slope of the water table is equal to the hydraulic gradient.

- Recharge to the aquifer due to rainfall and OWTS effluent is not considered

By making these approximations, a DEM, after being appropriately processed, can be used to obtain an approximation to the hydraulic gradient which is then used in conjunction with Darcy's Law to obtain an estimate of the seepage velocity magnitude and direction. The DEM is processed by using a kernel-based smoothing filter. The amount of smoothing is a user-adjustable parameter. Additionally, the model is able to handle issues with using a DEM as input, namely, the treatment of sinks and flat areas. Furthermore, the flow model accepts spatial heterogeneity in the hydraulic conductivity and porosity.

It is found that for areas in the vicinity of the U.S. Naval Air Station Jacksonville, Florida (NAS), the water table is likely a subdued replica of the topography, to varying degrees. Based on a comparison with a calibrated MODFLOW model of the area, the correlation between the simulated water table elevation and the ground surface elevation is 0.87 (Pearson's correlation coefficient). Applying a rule-of-thumb to the NAS indicates that the water table in the vicinity of the NAS is weakly topography controlled. This result is extended to the nearby Lakeshore neighborhood (which is a neighborhood of interest). For the Lakeshore neighborhood, applying the rule-of-thumb indicates that the water table is not topography controlled however a topography-controlled water table may not be completely ruled out due to a) the proximity of the NAS to Lakeshore and b), a semi-synthetic model of the area that suggests a relationship between the water table and the topography.

Using the calibrated MODFLOW model as a reference, it is found that the smoothing factor of the simplified model is location and scale dependent, when predicting groundwater flow paths and travel times. For the two sub-domains of the NAS, the simplified model can predict the travel times for the combined sub-domains to within ±20 years of the MODFLOW model, 75% of the time. It is found that the agreement between this model and the MODFLOW model decreased with distance. There is good agreement up to a distance of approximately 500 meters. This distance is acceptable since OWTS of interest are usually close to the water body of interest.

### 5.1.2   Transport Module

A simplified model that simulates the fate and transport of nitrates from OWTS in the saturated zone within the surficial aquifer was developed and implemented in the Transport module. The Transport module implements the model by using an analytical solution to the advection-dispersion equation (ADE) with first-order decay, based on the one developed by Domenico and Robbins (1985), Domenico (1987) and Martin-Hayden and Robbins (1997), a.k.a the Domenico solution. Due to the use of the Domenico solution, the model inherits its assumptions resulting in several limitations, some of which are addressed by the model implementation. The assumptions in the Domenico solution are:

- The domain is infinite in y and z and semi-infinite in the +x direction.

- The contaminant source is represented by a finite vertical plane of constant concentration centered at the origin, and remaining constant in time and space.

- The dispersivities are homogeneous.

- The first-order decay parameter is homogeneous.

- The velocity flow field is steady, homogeneous and exclusively in the +x direction.

- There is only a single contaminant source in the domain.

A consequence of the above assumptions is that Domenico plumes are always straight, extending in the +x direction. In order to speed up processing and reduce memory requirements, the transport model is a 2.5D model, meaning that a two dimensional contaminant plume is calculated, which is then extended vertically downwards to approximate a 3D plume. This approximation is valid when the vertical dispersivity is low, as is normally the case. The assumption that physical aquifer parameters, such as dispersivities and decay rate parameters, are homogeneous is justifiable within a study area of small scale.

The transport module addresses the heterogeneity of flow velocity, and the inability of the Domenico solution to directly consider multiple contaminant sources simultaneously. In regards to the flow field, there are two types of heterogeneity in the flow velocity: magnitude and direction. In order to adapt the Domenico solution to a steady flow field with a flow direction that is not exclusively in the +x direction, the Domenico solution is mapped to the curved flow path generated by the flow module, achieved by mapping the plume centerline to the flow path and using a warping algorithm ($1^{st}$ or $2^{nd}$ order polynomial, thin-plate spline) to transform the entire plume to this curved flow path. The variation in velocity along the plume centerline is considered by averaging the velocity along the flow path and then using the average value in the Domenico solution. Variation in the flow field (direction and magnitude) are not considered for areas of the plume away from the centerline. Additionally, the use of GIS to manage input data allows for the possibility to consider a certain degree of heterogeneity in the physical aquifer parameters (dispersivity, decay rate). This heterogeneity can be introduced when considering multiple plumes by making use of feature class attributes to store a potentially distinct set of parameters for each contaminant source. Essentially, parameters can vary over all plumes although each set of parameters will remain constant for any particular plume.

Under ideal conditions (i.e., all of the Domenico solution assumptions are satisfied), a comparison with a numerical model (constructed using MT3DMS) shows that the Domenico solution generally underestimates concentrations, in agreement with Srinivasan et al. (2007) and West et al. (2007). The degree of underestimation depends on the scenario and for the tested scenarios, the concentration was underestimated by up to 60%. For the test cases, as the decay rate increases, the concentration is gradually overestimated at locations moving away from the plume centerline, resulting in plumes that are wider than the corresponding MT3DMS plume. This underestimation/overestimation is present in all other test cases as well and is an unavoidable consequence of using the Domenico solution, the consequence of which is to produce shorter/wider or shorter/narrower plumes than the numerical solution, depending on the specific scenario.

Under non-ideal conditions, a general conclusion regarding the performance of the model cannot be reached, given that the results will depend on the nature and extent of the deviations of the system from the model assumptions. However a conclusion *can* be reached regarding the specific scenarios tested. The test cases were:

- A flow in the +x direction with a linearly increasing magnitude. This tests the effect of averaging the flow velocity along the flow path.

- A flow with a constant magnitude but a flow path in the shape of an arc. This tests the effect of warping the plume to the flow path.

In the case of flow in the +x direction with a linearly varying magnitude, the model underestimates the width of the plume in the region of the flow field that is slower than the average velocity and overestimates the width when the flow field velocity is faster. Near the plume advection front, the algorithm produces a shorter plume than the numerical solution in the transient case, due to the fact that the Domenico-based model cannot take into account the increased velocity in this region (in addition to the general concentration underestimation). Overall, the model results were comparable to the numerical model (except in the above mentioned region), for this configuration of the flow field. When the flow path is not straight, the plume is warped to conform to this curved path. The test case shows that warping the plume to the flow path provides a reasonable approximation to the behavior predicted by the numerical model. The only major difference occurs near the advection front where the plume generated by MT3DMS is slightly asymmetrical, due to the variation in the flow field velocity magnitude along the plume transverse. A small but potentially significant effect is the one introduced by the warping of the plume. By warping the plume, the concentration values of the warped plume compared to the unwarped plume changed by 4% on average for a transverse plume cross section at 500 m from the source. This change may impact the load calculations carried out by the denitrification module.

In order to consider multiple plumes simultaneously, the model combines multiple Domenico solutions using additive superposition. While the behavior of this procedure is complex, it can generally be said that the effect of superimposing Domenico solutions in this way is that in the overlapping region, the concentration will generally be slightly overestimated compared to the numerical solution. The overestimation originates from two sources: the inability of the Domenico solution to consider the presence of other plumes and the fact that depending on the parameters (e.g., decay rate, dispersivities, etc.), Domenico plumes can be wider than plumes calculated via MT3DMS. As plumes become wider compared to the MT3DMS, the latter effect will dominate. This overestimation is counter-balanced by the general effect of underestimation of the Domenico solution. This means that in overlapping areas, the concentration may still be under-estimated depending on how much the individual plumes overlap.

The only general conclusion that can be reached regarding the performance of the transport model under heterogeneous conditions is that the model can be expected to produce a reasonable estimate of the contaminant plume as long as the system under study does not deviate considerably

from the assumptions underlying the Domenico solution. However when comparing concentration values at specific locations, the issue of the Domenico solution producing shorter/wider plumes remains.

Application of the model using parameter values from the literature to a study site (Waterside Dr.) and comparing with real-world measurements shows that the model has the potential to predict the general behavior of nitrate plumes. In order to realize this potential in the study site, model calibration with site specific parameter values is required.

### 5.1.3   Nitrate Load Estimation Module

Load calculations are carried out by the Nitrate Load Estimation module, which uses mass balance considerations to determine the nitrate load to a target water body by calculating the input mass load and mass removed due to denitrification. The nitrate load is calculated by considering the system in the steady-state. The load is calculated by subtracting the mass removed due to denitrification from the input mass due to advection and dispersion.

The results for an ideal case were compared to MT3DMS results for the same case. Results indicate that for low denitrification rates, the load calculation algorithm approximates the results from the numerical simulation well. For higher denitrification rates, the underestimation of the concentration by the Domenico solution, in combination with the discretization error, gives an underestimation of denitrification mass of about 9% in the best case (single plume, small grid size) and 27% in the worst case (overlapping plumes). Due to the use of an analytical solution by our model for the calculation of the input load, it is shown that the input load is underestimated by MT3DMS. This is demonstrated by the fact that as the grid size becomes smaller, the MT3DMS input load tends to the value calculated by our model. It is also shown that decreasing the grid size yields a denitrification mass estimation by our model that is closer to the MT3DMS results.

### 5.1.4   General Conclusions

The model, taken as a whole, has the potential to be a useful tool in providing quick estimates of nitrate loads to surface water bodies due to its ease of use and low data requirements. Initial feedback from the FDEP and others involved in subsurface contamination in Florida has been positive, with several commenting on the usefulness of having a modeling tool that can directly use GIS data layers as input to the modeling process. As a result, further development of this tool is a worthwhile endeavor.

It is important to note that due to the simplified nature of this model, exact agreement with field data is an unrealistic expectation. Additionally, due to the approximate nature of the model, results should be taken as a bulk estimate and not an exact prediction.

## 5.2 Future Work

There are many avenues for future work. For clarity, they have been separated into future work related to the conceptual model, the algorithmic implementation, and model analysis.

### 5.2.1 Conceptual Model

Due to the approximate nature of the Domenico solution, it may be preferable to use a different analytical solution. Because all of the alternate solutions are more computationally intensive than the Domenico solution, further investigation is required to determine the impact they will have on computation speed.

An improvement to the flow model could be achieved through the use of an adaptive smoothing algorithm. In this algorithm, water table measurements could be used by applying less smoothing where the water table is known to be shallow (an therefore more likely to resemble the topography) and more smoothing would be applied to areas where the water table was measured to be deeper. Another approach to would be to directly incorporate field measurements of the water table using a method similar to the one presented by Desbarats et al. (2002). In the method of Desbarats et al. (2002), head measurements were used in conjunction with DEM information in a kriging model to obtain an estimate of the water table elevation.

According to the literature, the concept of biodegradation and the instantaneous reaction approximation may be a better way of considering natural attenuation compared to first-order decay. However, little research has been done regarding its application to nitrate attenuation via the mechanism of denitrification. Additionally, investigation is required to determine whether the instantaneous reaction approximation is suitable for nitrate load calculations.

Currently, the transport model is a steady-state, 2D model. Moving to a 3D transient state model would be beneficial since it would allow for a more accurate representation of nitrate plumes as well as allowing for the analysis of the plumes in early times. The feasibility of this improvement needs to be further evaluated in terms of the computational impact on the model as well as the impact on development time and budgetary constraints due to increased complexity

### 5.2.2 Implementation

Due to the approximate nature of the Domenico solution, it may be preferable to modify the load calculations to a flux-based approach. This would involve calculating the advective and dispersive mass flux at the boundary of the water body. By calculating the flux in this way, it would remove the requirement that the system be in the steady state, thereby enabling the analysis of transient state cases. Additionally, this method of calculating load is transferable, without modification, to other solution methods such as the instantaneous reaction model.

### 5.2.3   Analysis

Although each module was tested independently, a full evaluation of model performance using a known solution of a more realistic system, would provide a better indicator of how the model performs in real-world cases, given the various approximations in each submodule.

# Bibliography

ADAMUS, C. L. AND BERGMAN, M. J. 1995. Estimating non-point source pollution loads with a GIS screening model. *Water Resources Bulletin 31*, 4, 647–655.

ANDERSON JR., W. P., EVANS, D. G., AND SNYDER, S. W. 2000. The effects of Holocene barrier-island evolution on water-table elevations, Hatteras Island, North Carolina, USA. *Hydrogeology Journal 8*, 390–404.

ANDRES, A. S. 1992. *Estimate of nitrate flux to Rehoboth and Indian river bays, Delaware, through direct discharge of ground water – Open File Report No. 35.* Delaware Geological Survey.

ART, H. W. 1993. *A Dictionary of Ecology and Environmental Science*, 1st ed. Henry Holt and Company.

ARTHUR, J. A., BAKER, A. E., CICHON, J. R., WOOD, A. R., AND RUDIN, A. 2005. *Florida Aquifer Vulnerability Assessment (FAVA): contamination potential of florida's principal aquifer systems.* Florida Geological Survey.

ARYA, A. 1986. Dispersion and reservoir heterogeneity. Ph.D. thesis, University of Texas, Austin.

AZIZ, C. E., NEWELL, C. J., GONZALES, J. R., HAAS, P., CLEMENT, T. P., AND SUN, Y. 2000. *BIOCHLOR Natural Attenuation Decision Support System User's Manual Version 1.0.* U.S. Environmental Protection Agency. Available online at `http://www.epa.gov/ada/csmos/models/biochlor.html`.

BANDILLA, K., SURIBHATLA, R., AND JANKOVIĆ, I. 2006. *SPLIT - Win32 computer program for analytic-based modeling of single-layer groundwater flow in heterogeneous aquifers with particle tracking, capture-zone delineation, and parameter estimation.* Department of Civil, Structural and Environmental Engineering, University at Buffalo, Buffalo, NY.

BATU, V. 2008. Discussion of "On the Use and Error of Approximation in the Domenico (1987) Solution" by Michael R.West, Bernard H. Kueper, and Michael J. Ungs, March-April 2007, v. 45, no. 2: 126-135 and "Domenico Solution–Is It Valid?" by V. Srinivasan, T.P. Clement, and K.K. Lee, March-April 2007 issue, v. 45, no. 2: 136-146. *Ground Water 46*, 5, 664–666.

BEAR, J. 1972. *Dynamics of Fluids in Porous Media.* Elsevier N.Y.

BEAR, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill, New York.

BEAUJOUAN, V., DURAND, P., RUIZ, L., AUROUSSEAU, P., AND COTTERET, G. 2002. A hydrological model dedicated to topography-based simulation of nitrogen transfer and transformation: rationale and application to the geomorphology-denitrification relationship. *Hydrological Processes 16*, 2, 493–507.

BECKER, M. W. AND JIANG, Z. 2007. Flux-based contaminant transport in a GIS environment. *Journal of Hydrology 343*, 3-4, 203–210.

BEVEN, K. J. AND KIRKBY, M. J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin 24*, 1, 43–69.

BODE-BÖGER, S. M. AND KOJDA, G. 2005. Organic nitrates in cardiovascular disease. *Celluar and Molecular Biology 51*, 307–320.

BOOKSTEIN, F. L. 1989. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions On Pattern Analysis And Machine Intelligence 11*, 6, 567–585.

BROWN, M. T., ANNABLE, M. D., DELFINO, J. J., JAWITZ, J. W., COHEN, M., HALL, E., HARDEN, H., CHANTON, J., BURNETT, W., ROSE, J. B., PAUL, J. H., GRIFFIN, D. W., LIPP, E., AND JOHN, D. 2001. *Determination of an Appropriate Onsite Sewage System Setback Distance to Seasonally Inundated Areas.* Florida Department of Health.

CANTER, L. W. 1996. *Nitrates in Groundwater*, 1st ed. CRC Press.

CRANK, J. 1975. *The Mathematics of Diffusion*, 2nd ed. Oxford University Press.

DAVIS, J. H. 2000. *Fate and Transport Modeling of Selected Chlorinated Organic Compounds at Operable Unit 3, U.S. Naval Air Station, Jacksonville, Florida.* U.S. Geological Survey, Tallahassee FL. Open-file report 00-255.

DAVIS, J. H., PLANERT, M., AND ANDREWS, W. J. 1996. *Simulation of Ground-Water Flow at the U.S. Naval Air Station, Jacksonville, Florida, with an Evaluation of Changes to Ground-Water Movement Caused by Proposed Remedial Designs at Operable Unit 1.* U.S. Geological Survey, Tallahassee FL. Open-file report 96-597.

DESBARATS, A. J., LOGAN, C. E., HINTON, M. J., AND SHARPE, D. R. 2002. On the kriging of water table elevations using collateral information from a digital elevation model. *Journal of Hydrology 255*, 1-4, 25 – 38.

DIERSCH, H. J. G. 2006. *FEFLOW Finite Element Subsurface Flow and Transport Simulation System*. Reference manual, WASY GmbH Institute for Water Resources Planning and Systems Research.

DOMENICO, P. A. 1987. An analytical model for multidimensional transport of a decaying contaminant species. *Journal of Hydrology 91*, 49–58.

DOMENICO, P. A. AND ROBBINS, G. A. 1985. A new method of contaminant plume analysis. *Ground Water 23*, 4, 476–485.

DOMENICO, P. A. AND SCHWARTZ, F. W. 1998. *Physical and Chemical Hydrology*, 2nd ed. John Wiley & Sons Inc., New York.

DONATO, G. AND BELONGIE, S. 2002. Approximate thin plate spline mappings. In *Computer Vision – ECCV 2002*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. Lecture Notes in Computer Science, vol. 2352. Springer Berlin / Heidelberg, 13–31.

ESRI. 2009. *ArcGIS Desktop Help 9.3 – Online User's Manual*. Environmental Systems Research Institute, Inc. Available at http://webhelp.esri.com/arcgisdesktop/9.3.

FAIRCHILD, R. W. 1972. *The Shallow-Aquifer system in Duval County, Florida*. Florida Bureau of Geology.

FREEZE, R. A. AND CHERRY, J. A. 1979. *Groundwater*. Prentice Hall, Inc.

GELHAR, L. W., WELTY, C., AND REHFELDT, K. R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research 28*, 7, 1955–1974.

GITTLEMAN, A. 2003. *Computing with C# and the .NET Framework*. Jones and Bartlett Publishers.

GOGU, R., CARABIN, G., HALLET, V., PETERS, V., AND DASSARGUES, A. 2001. GIS-based hydrogeological databases and groundwater modelling. *Hydrogeology Journal 9*, 6, 555–569.

HAITJEMA, H. M. AND MITCHELL-BRUKER, S. 2005. Are water tables a subdued replica of the topography? *Ground Water 43*, 6, 781–786.

HARBAUGH, A. W. 2005. *MODFLOW-2005, the U.S. Geological Survey modular ground-water model – the Ground-Water Flow Process*. USGS Techniques and Methods 6-A16. Reston, Virginia: USGS.

HARBAUGH, A. W., BANTA, E. R., HILL, M. C., AND McDONALD, M. G. 2000. *MODFLOW-2000, the U.S. Geological Survey modular ground-water model – User guide to modularization concepts and the Ground-Water Flow Process*. U.S. Geological Survey Open-File Report 00-92.

HEINEN, M. 2006. Simplified denitrification models: Overview and properties. *Geoderma 133*, 3-4, 444–463.

HITT, K. J. AND NOLAN, B. T. 2005. Nitrate in ground water–using a model to simulate the probability of nitrate contamination of shallow ground water in the conterminous United States: U.S. Geological Survey Scientific Investigations Map 2881.

Hoffman, J. L. and Canace, R. J. 2004. *A Recharge Based Nitrate Dilution Model for New Jersey, New Jersey Geological Survey Open File Report OFR 04-1*. New Jersey Department of Environmental Protection, Trenton, NJ.

Ibaraki, M. 2001. Source characteristics and contaminant plume evolution. *Transport in Porous Media 44*, 3, 577–589.

Jackson, W. A., Rainwater, K., Anderson, T., Lehman, T., Tock, R., Rajagopalan, S., and Ridley, M. 2004. *Distribution and Potential Sources of Perchlorate In the High Plains Region of Texas – Final Report*. Texas Commission on Environmental Quality, Austin TX.

Jenson, S. K. and Domingue, J. O. 1988. Extracting topographic structure from digital elevation data for geographical information system analysis. *Photogrammetric Engineering and Remote Sensing 54*, 11, 1593–1600.

Kalivarapu, V. and Winer, E. 2008. A multi-fidelity software framework for interactive modeling of advective and diffusive contaminant transport in groundwater. *Environmental Modelling and Software 23*, 12, 1370–1383.

Kellogg, D. Q., Joubert, L., and Gold, A. 1996. *Technical documentation, nutrient loading component of the MANAGE Geographic Information System-based risk assessment method*. University of Rhode Island Cooperative Extension.

King, F. H. 1899. Principles and conditions of the movements of groundwater. *U.S. Geological Survey Annual Report 19*, 2, 59–294.

Lee, S., McAvoy, D. C., Szydlik, J., and Schnoor, J. L. 1998. Modeling the fate and transport of household chemicals in septic systems. *Ground Water 36*, 1, 123–132.

Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. 2005. *Geographic Information Systems and Science*, 2nd ed. John Wiley & Sons Ltd.

Martin, P., LeBoeuf, E., Dobbins, J., Daniel, E., and Abkowitz, M. 2005. Interfacing GIS with water resource models: a state-of-the-art review. *Journal of the American Water Resources Association 41*, 6, 1471–1487.

Martin-Hayden, J. M. and Robbins, G. A. 1997. Plume distortion and apparent attenuation due to concentration averaging in monitoring wells. *Ground Water 35*, 2, 339–346.

Martz, L. W. and Garbrecht, J. 1998. The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models. *Hydrological Processes 12*, 6, 843–855.

McCray, J. E. 2009. *Modeling Onsite Wastewater Systems at the Watershed Scale: A User's Guide*. Water Environment Research Foundation, Alexandria, VA.

McCray, J. E., Kirkland, S. L., Siegrist, R. L., and Thyne, G. D. 2005. Model parameters for simulating fate and transport of on-site wastewater nutrients. *Ground Water 43*, 4, 628–639.

McDonald, M. G. and Harbaugh, A. W. 1988. *A modular three-dimensional finite-difference ground-water flow model: U.S. Geological Survey Techniques of Water-Resources Investigations, book 6, chapter A1,*.

Mueller, D. K. and Spahr, N. E. 2006. *Nutrients in streams and rivers across the Nation – 1992-2001: U.S. Geological Survey Scientific Investigations Report 2006-5107.* U.S. Geological Survey, Reston, Virginia.

Neuman, S. P. 1990. Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resources Research 26*, 8, 1749–1758.

Newell, C. J., McLeod, K. R., and Gonzales, J. R. 1996. *BIOSCREEN, Natural Attenuation Decision Support System, User's Manual Version 1.3. EPA/600/R-96-087.* U.S. Environmental Protection Agency. Available online at `http://www.epa.gov/ada/csmos/models/bioscrn.html`.

Ogata, A. and Banks, R. B. 1961. Solution of the differential equation of longitudinal dispersion in porous media. *U.S. Geological Survey Professional Paper 411-A*, 7.

Pan, F., Peters-Lidard, C. D., Sale, M. J., and King, A. W. 2004. A comparison of geographical information systems-based algorithms for computing the TOPMODEL topographic index. *Water Resources Research 40*, W06303.

Pennsylvania Department of Environmental Protection. 2002. *Pennsylvania's Land Recycling Program, Technical Guidance Manual.* Bureau of Land Recycling and Waste Management.

Peuquet, D. J. and Marble, D. F. 1990. *Introductory Readings In Geographic Information Systems*, 1st ed. CRC Press.

Pickens, J. F. and Grisak, G. E. 1981. Scale-dependent dispersion in a stratified aquifer. *Water Resources Research 17*, 4, 1191–1211.

Prakash, A. 2009. Comment on "On the Use and Error of Approximation in the Domenico (1987) Solution" by Michael R.West, Bernard H. Kueper, and Michael J. Ungs, March-April 2007, v. 45, no. 2: 126-135. *Ground Water 47*, 6, 758–759.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1996. *Numerical Recipes in Fortran 90*, 2nd ed. Cambridge University Press.

Savitzky, A. and Golay, M. J. E. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry 36*, 8, 1627–1639.

Schecher, W. 1997. *Septic Tank Treatment System (SepTTS) Model for Cleaning Product Chemicals, User's Manual v1.0.* The Soap and Detergent Association.

Schilling, K. and Wolter, C. 2007. A GIS-based groundwater travel time model to evaluate stream nitrate concentration reductions from land use change. *Environmental Geology 53*, 2, 433–443.

SCHOWENGERDT, R. A. 2006. *Remote Sensing: Models and Methods for Image Processing*, 3rd ed. Academic Press.

SEPÚLVEDA, N. 2002. *Simulation of Ground-Water Flow in the Intermediate and Floridian Aquifer Systems in Peninsular Florida*. U.S. Geological Survey, Tallahassee FL. Water-Resources Investigations Report 02-4009.

SHAHBAZI, M., ZAND, S., AND TODD, D. K. 1968. Effect of topography on ground water flow. *International Association of Scientific Hydrology 77*, 314–319.

SIEBERT, J., BISHOP, K. H., AND NYBERG, L. 1997. A test of TOPMODEL's ability to predict spatially distributed groundwater levels. *Hydrological Processes 11*, 9, 1131–1144.

SILAVISESRITH, W. AND MATOTT, L. S. 2005. *ArcAEM: GIS-based application for analytic element groundwater modeling documentation and user's guide*. University at Buffalo, Buffalo, NY. Available from: `http://www.groundwater.buffalo.edu`.

SINGER, J. 2003. JVM versus CLR: a comparative study. In *PPPJ '03: Proceedings of the 2nd international conference on principles and practice of programming in Java*. Computer Science Press, Inc., New York, NY, USA, 167–169.

SØRENSEN, R. AND SEIBERT, J. 2007. Effects of DEM resolution on the calculation of topographical indices: TWI and its components. *Journal of Hydrology 347*, 1-2, 79 – 89.

SRINIVASAN, V., CLEMENT, T. P., AND LEE, K. K. 2007. Domenico solution – is it valid? *Ground Water 45*, 2, 136–146.

STIEGLITZ, M., RIND, D., FAMIGLIETTI, J., AND ROSENZWEIG, C. 1997. An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. *Journal of Climate 10*, 118–137.

THOMPSON, J. A., BELL, J. C., AND BUTLER, C. A. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. *Geoderma 100*, 1-2, 67 – 89.

THOMPSON, J. C. AND MOORE, R. D. 1997. Relations between topography and water table depth in a shallow forest soil. *Hydrological Processes 10*, 11, 1513–1525.

TÓTH, J. 1963. A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research 68*, 4795–4812.

TROELSEN, A. 2008. *Pro VB 2008 and the .NET Framework 3.5*, 3rd ed. Apress.

UNGS, M. J., WEST, M. R., AND KUEPER, B. H. 2009. Reply to 'Comment on 'On the Use and Error of Approximation in the Domenico (1987) Solution' by Michael R.West, Bernard H. Kueper, and Michael J. Ungs, March-April 2007, v. 45, no. 2: 126-135'. *Ground Water 47*, 6, 759–760.

Ursin, E. L. and Roeder, E. 2008. An assessment of nitrogen contribution from onsite wastewater treatment systems (OWTS) in the Wekiva study area of central Florida. In *National Onsite Wastewater Recycling Association (NOWRA) Nitrogen Symposium*. Florida Department of Health.

U.S. Department of Agriculture. 2004. *Soil Survey Laboratory Methods Manual Version 4.0*. Natural Resources Conservation Service. Soil Survey Laboratory Investigations Report No. 42.

U.S. Environmental Protection Agency. 1993. *Nitrogen Control*. Office of Research and Development, Cincinnati Ohio and Office of Water, Washington, DC. Available online at `http://nepis.epa.gov/Adobe/PDF/30004MI0.PDF`.

U.S. Environmental Protection Agency. 2002. *A Homeowner's Guide to Septic Systems*. Office of Water, Washington, DC. Avaiable online at `http://www.epa.gov/owm/septic/pubs/homeowner_guide_long.pdf`.

Valiela, I., Collins, G., Kremer, J., Lajtha, K., Geist, M., Seely, B., Brawley, J., and Sham, C. H. 1997. Nitrogen loading from coastal watersheds to receiving estuaries: New method and application. *Ecological Applications 7*, 2, 358–380.

West, M. R., Kueper, B. H., and Ungs, M. J. 2007. On the use and error of approximation in the Domenico (1987) solution. *Ground Water 45*, 2, 126–135.

Wexler, E. J. 1992. Analytical solutions for one-, two-, and three-dimensional solute transport in ground-water systems with uniform flow. In *Techniques of Water Resources Investigations of the United States Geological Survey, Book 3: Applications of Hydraulics*. U.S. Geological Survey, Chapter B7, 53–55.

Winston, R. B. 2009. *ModelMuse - A graphical user interface for MODFLOW2005 and PHAST: U.S. Geological Survey Techniques and Methods 6A29*. U.S. Geological Survey. Available online at `http://pubs.usgs.gov/tm/tm6A29`.

Wolock, D. M. and Price, C. V. 1994. Effects of digital elevation model map scale and data resolution on a topography-based watershed model. *Water Resources Research 30*, 3041–3052.

Wörman, A., Packman, A. I., Marklund, L., Harvey, J. W., and Stone, S. H. 2006. Exact three-dimensional spectral solution to surface-groundwater interactions with arbitrary surface topography. *Geophysical Research Letters 33*, L07402.

Xu, M. and Eckstein, Y. 1995. Use of a weighted least-squares method in evaluation of the relationship between dispersivity and field scale. *Ground Water 33*, 6, 905–908.

Ye, Z., Maidment, D. R., and McKinney, D. C. 1996. *Map-based surface and subsurface flow simulation models: an object-oriented and GIS approach*. Center for Research in Water Resources, University of Texas at Austin. CRWR online report 1996-05. Available online at `http://hdl.handle.net/2152/6746`.

ZHENG, C. AND WANG, P. P. 1999. *MT3DMS: A modular three-dimensional multispecies model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; Documentation and User's Guide, Contract Report SERDP-99-1.* U.S. Army Engineer Research and Development Center, Vicksburg, MS.