Sensitivity Analysis and Uncertainty Assessment for ArcNLET-Estimated Nitrate Load from Septic Systems to Surface Water Bodies

Manuscript Completed: March 2012

Prepared by

Liying Wang[1], Ming Ye[1], J. Fernando Rios[2], and Paul Z. Lee[3]

[1]Department of Scientific Computing, Florida State University, Tallahassee, FL 32306

[2]Department of Geography, State University of New York at Buffalo, Buffalo, NY 14260

[3]Groundwater and Springs Protection Section, Florida Department of Environmental Protection, Tallahassee, FL 32399

Richard W. Hicks, Contract Manger, Florida Department of Environmental Protection, Tallahassee, FL 32399

Contents

Figures

Tables

1. INTRODUCTION

Nitrate (NO_3^-), as a commonly identified groundwater and surface-water pollutant, is associated with a number of adverse health and environmental impacts. Nitrate concentration higher than 10 mg/l (measured as nitrogen, EPA drinking water primary standard) in drinking water may cause methemoglobinemia, also known as blue baby syndrome. Discharge of nitrate-rich groundwater to surface water bodies can lead to fish kills, algal growth, hypoxia, eutrophication, and outbreaks of toxic bacteria. One important source of nitrate in environment is due to wastewater treatment using Onsite Sewage Treatment and Disposal Systems (OSTDS) (a.k.a., septic systems) (U.S. Environmental Protection Agency (EPA) 1993, 2002; McCray et al. 2005). The nitrate contribution from septic systems to surface water bodies may be significant in areas where septic systems are located in close proximity to surface water bodies. Groundwater in areas with a shallow water table is also vulnerable to nitrate contamination, because of direct discharge of effluent from septic systems into soil. This may pose a threat to public health if drinking water supply depends on shallow domestic wells (Hitt and Nolan 2005). In the U.S., approximately 25% of the population and 30% of all new development utilize septic systems (Hazen and Sawyer 2009). In the state of Florida, nearly a third of households use septic systems and 92% of drinking water supply is from groundwater (Ursin and Roeder 2008; Hazen and Sawyer 2009). Therefore, for protection of the environment and public health, it is important to simulate nitrate transport in groundwater due to septic systems and to estimate corresponding nitrate load to surface water bodies.

For this purpose and with support of the Florida Department of Environmental Protection, an ArcGIS-based Nitrate Load Estimation Toolkit (ArcNLET) was developed by the Florida State University (Rios, 2010; Rios et al., 2011a-c, 2012). It is based on a simplified model that simulates groundwater flow and solute transport under advection, dispersion, and denitrification. ArcNLET has been applied to estimate nitrate load at two neighborhoods in the City of Jacksonville, Florida (Wang et al., 2011a,b,c). It provides nitrate load from thousands of septic systems in the neighborhoods to water bodies within and surrounding the neighborhoods including Julington Creek and St. Johns River. The nitrate load estimation can be used for water resources and environmental management such as the implementation of the Total Maximum Daily Load (TMDL) program.

However, the ArcNLET-based nitrate load estimates are inherently uncertain, and the uncertainty needs to be quantified before the estimates (and/or their statistics) are used for environmental management and planning (Haan et al., 1998; Hession et al., 1996; Reckhow, 1994; Wang et al., 2005). The uncertainty may be caused by the following reasons:

(1) Model parameters vary in space and time, but the variability cannot be fully characterized by available data that is sparse in real-world modeling. This is always s referred to as parametric uncertainty.

(2) Model structure used for numerical simulation is inadequate to describe field observations of groundwater flow and nitrate transport, and there may be multiple conceptual-mathematical models that are plausible with given data and information. This kind of uncertainty is always referred to as model uncertainty.

(3) Model input data are subject to spatial and temporal variability that cannot be fully characterized. For example, future recharge to groundwater is unpredictable in nature and septic use is subject to epistemic error. There is no consensus on the name of this kind of uncertainty, and it is here referred to as scenario uncertainty.

(4) Measurements of model parameters and observations of state variables (e.g., hydraulic head and nitrate concentration) are subject to measurement errors, spatial and temporal variability, scarcity, and/or scale discrepancy. These may give rise to data uncertainty.

There are other kinds of categories of uncertainty sources. For example, for general transport-transformation models (not limited to groundwater models), uncertainty sources are categorized by Isukapalli (1999) as natural, model, and data uncertainties. While it is still an open question how to quantify model and scenario uncertainties, quantification of parametric uncertainty has been studied for decades and mature techniques are available.

In this study, we focus on quantification of parametric uncertainty associated with the ArcNLET model in the following procedure:

(1) Identify uncertainty parameters whose values are not known with certainty. Which parameters are uncertain is problem specific and depends on available data. For example, an uncertain parameter at one site may become more certain at another site where sufficient measurements are available.

(2) Characterize the parametric uncertainty using statistical methods. The uncertainty may be described by probabilistic distribution of the parameters that include parameter ranges and probability of the parameter values.

(3) Conduct sensitivity analysis to better understand the relations between model parameters and the load estimate. The sensitivity analysis can also be used to select the parameters to which the load estimates are sensitive. This can reduce the number of parameters needed for quantification of parametric uncertainty.

(4) Quantify the parametric uncertainty using statistical methods such as Monte Carlo methods.

Instead of providing a single estimate, the uncertainty assessment provides multiple realizations of possible load values and gives probability of the values (e.g., which values are more likely than other values and to what extent). The uncertainty is summarized in the probability distributions of load estimates in the form of histograms. The distributions can be used to assess risk, i.e., the probability that the load exceeds a specific threshold or performance measure target value, which will provide more valuable information for decision/policy makers and help them making or evaluating decisions (Loucks et al., 2005).

ArcNLET has a total of six hydraulic and transport parameters: hydraulic conductivity, porosity, longitudinal dispersivity, horizontal transverse dispersivity, first-order denitrification coefficient, and source nitrate concentration. ArcNLET uses an additional parameter, the smoothing factor, to process Digital Elevation Model (DEM) for obtaining the shape of water table. Among these parameters, hydraulic conductivity and porosity are modeled as heterogeneous parameters at the neighborhood scale. There is usually more than one soil zone, and different zones have different values of hydraulic conductivity and soil porosity that can be obtained by processing the Soil Survey Geographic Database (SSURGO) (Wang et al., 2011b). While hydraulic conductivity and porosity are treated as constants within each zone, they actually vary in space. This deviation gives rise to parametric uncertainty, which is taken account into the SSURGO database by assigning representative, minimum, and maximum hydraulic conductivity for each soil zone. Since spatial variability of porosity is significantly smaller than that of hydraulic conductivity, porosity is always treated as a deterministic variable and there is only one porosity value for each soil zone in the SSURGO database. It is worth mentioning that this way of handling parametric uncertainty in hydraulic conductivity and porosity is restricted to the way of processing the SSURGO database in Wang et al. (2011b), in which only the parameter values of dominant component at the bottom horizon is used. If the SSURGO database is processed in different ways (e.g., Tesfa et al., 2009), the values of hydraulic conductivity and porosity may vary.

In the current ArcNLET application, longitudinal dispersivity, horizontal transverse dispersivity, first-order denitrification coefficient, source plane concentration, and smoothing factor are treated as homogeneous parameters, because there is no field measurement of these parameters. Their effective values are estimated during model calibration in which the model is fitted to field observations of hydraulic head and nitrate concentration. When there is no field data to characterize uncertainty of these parameters, the uncertainty is characterized using literature data. For example, the cumulative frequency distributions of denitrification rates and ranges of source nitrate concentration given by McCray et al. (2005) are used in this study. When field data is available that the parameter can be estimated, the estimated parameters can be used to better quantify the parametric uncertainty. Theoretically speaking, the model parameters should have spatial and/or statistical correlation. However, due to lack of data to characterize the correlation, the correlation is not considered in this study, except that longitudinal dispersivity is assumed to be 10 times as large as horizontal transverse dispersivity; this ratio is commonly used in groundwater transport modeling. In a summary, the random parameters explicitly considered in this study are hydraulic conductivity at various soil zones, longitudinal dispersivity, first-order denitrification coefficient, source nitrate concentration, and the smoothing factor.

Sensitivity of nitrate load estimate to the model parameters is conducted in this study to better understand their relations, in particular, the major source of prediction uncertainty, i.e., which parameters cause relatively large uncertainty in the load estimate. Among various methods of sensitivity analysis (e.g., Wang et al., 2005), statistical methods are the most comprehensive, because they estimate the probability distribution of model outputs based on the distributions of model inputs (Andersson et al., 2000). Among the statistical methods such as analysis of variance (ANOVA) and response surface methods (RSM), and Fourier amplitude sensitivity test (FAST), the variance based methods are versatile and effective (Saltelli et al., 2010). Highly effective variance based methods, such as Sobol' (Sobol', 1990) and FAST (Fourier Amplitude Sensitivity Test), have been used in the field of hydrology (Perera et al., 2006) for investigating the propagation of parameter uncertainty through groundwater flow and solute transport. These methods however are computationally expensive. It is suggested to use screening methods first to exclude less influential parameters before the variance-based sensitivity analysis is conducted so that the number of model evaluations can be reduced without affecting results of sensitivity analysis (Zhao et al., 2011).

In this study, a qualitative analysis is first conducted to understand the relations between nitrate load estimate and model input parameters. Subsequently, the Morris method (Morris, 1991) is conducted as a mathematical method to obtain an estimation of parameters ranking at global scale. At last, the Sobol' method is used to provide quantitative measures of parameter sensitivity, i.e., to what extent prediction uncertainty is caused by uncertainty in which parameters.

Uncertainty in the nitrate load estimation is quantified using the the Monte Carlo (MC) methods in the following procedure (Kalos and Whitlock, 1986; Fishman, 1996):

(1) Generate multiple samples of model parameters based on their distributions characterized based on available data and information.

(2) Execute ArcNLET for the parameter samples. A command-line version of ArcNLET is used in this step so that the multiple execution of ArcNLET can be automated.

(3) Evaluate probabilistic distribution of the load estimate for uncertainty quantification.

The Latin Hypercube Sampling (LHS) method (Iman and Conover,1980; Loh, 1996) is selected for the random number generation since it ensures that generated random samples span the full coverage of the random variables. This property of LHS reduces the computational cost of Monte Carlo simulations, since in comparison with the random sampling method LHS requires a smaller number of parameter realizations to yield representative parameter distribution functions (Helton et al., 2003). The LHS method is one of the widely used methods, and it has been used in various fields (Barrett, 1996; Camilo, 1992; Hession et al., 1996; Ogle et al., 2003) including groundwater contaminant transport modeling (Haan et al., 2003a, 2003b; Baalousha, 2006; Rong et al., 1998; Gwo et al., 1996; Liou et al., 1997; Osnes, 1998).

In the remaining part of this report, the methods of sensitivity analysis (centered parameter method, Morris method, and Sobol' method) and uncertainty assessment (LHS-based MC method) are described in Section 2, followed by descriptions of three sites to which the methods are applied in Section 3. Results of the applications are discussed in Section 4, and the conclusions from the applications are given in Section 5.

2. METHODOLOGY

In this section, we first briefly describe a qualitative sensitivity study and then explain more quantitative Morris method and Sobol' method of sensitivity analysis. The LHS-based MC method for uncertainty quantification is discussed at the end of this section.

2.1. Qualitative Sensitivity Study

The basic sensitivity study is to examine how model outputs vary with model inputs. When a model has multiple parameters, in general, the model is executed when only one parameter varies and other parameters are fixed at their nominal values. After model execution, the model outputs are plotted against corresponding parameter values; the slope of the curve is an indicator of local sensitivity. This method however does not consider interaction between different parameters.

The nominal parameter values are always those of interest to the modeler such as average or certain reference parameter values. In this study, the nominal values listed in Table 2-1 are taken from literature that are also listed in Table 2-1; the rationale of selecting these values can be found in Rios (2010) and Rios et al. (2011a). Although the nominal values are called default values in ArcNLET, they are simply used for convenience of the users to learn operation of the software. In particular, the values are not specific to any sites in Jacksonville, and the load estimate based on the parameter values should not be viewed as the results for Jacksonville. The users of ArcNLET should use the parameter values appropriate to their sites of interest.

Table 2-1 also lists the parameter ranges that are used to generate multiple parameter values for the qualitative sensitivity study. Although hydraulic conductivity and porosity are heterogeneous within a neighborhood, it is assumed that they are homogeneous in the qualitative sensitivity analysis. Heterogeneous hydraulic conductivity is used in the more advanced Morris method and variance based method of sensitivity analysis.

Table 2-1. Nominal parameter values embedded in ArcNLET and parameter ranges used in centered parameter study.

Parameter	Value	Range	Source
Hydraulic conductivity , *K*, (m/d)	2.113	0.001-20	Rios (2011b)
Soil porosity , Φ, (-)	0.25	0.25-0.50	Rios (2011b)
Smoothing factor , *SmthF*, (-)	50	20-80	Rios (2011b)
Longitudinal dispersivity, α_L (m)	2.113	0.21-21.34	Davis (2000) Gelhar et al. (1992)
Horizontal transverse dispersivity, α_T (m)	0.21	0.021-2.134	Davis (2000) Gelhar et al. (1992)
First-order decay coefficient, *k*, (1/d)	0.008	0.004-1.08	McCray et al. (2005)
Source plane concentration,C_0 , (mg/L)	40	25-80	McCray et al. (2005)

2.2. Morris Analysis

The Morris method implemented in the DAKOTA software is used to quantitatively investigate sensitivity of nitrate load estimation to the model parameters. The Morris method determines the factors (e.g., model parameters) that have significant effect on model outputs (e.g., nitrate load estimate). Morris method is based on elementary effect of each model parameter. The elementary effect, d_i, of input parameter x_i is defined as

$$d_i(\mathbf{x}) = \frac{y(x_1,...,x_{i-1},x_i+\Delta,x_{i+1},...,x_k) - y(\mathbf{x})}{\Delta} \tag{1}$$

where $x = (x_1, x_2, \ldots, x_k)$ is a parameter point within the defined parameter space, Δ is determined as $\Delta = p/(2p\text{-}2)$ (p being the partition value specified by the user to separate the parameter range into p segments), and y is a model output. By randomly sampling different x values from the parameter space and running multiple simulations with the parameter samples, a finite distribution (F_i) of elementary effects for the input parameter x_i is obtained. A high mean of the distribution F_i reveals that the input parameter x_i has an important overall influence on the output, whereas a high standard deviation shows that the parameter x_i is either interacting with other parameters or has a non-linear effect (Morris, 1991; Saltelli et al., 2004). The Morris method is a method of global sensitivity analysis that is capable of ranking the input factors in order of importance and is computationally more efficient than other methods of global sensitivity analysis. However, this method cannot give a quantitative measure about the percentage of total output uncertainty caused by uncertainty of each parameter, which can be achieved by the Sobol' method below.

2.3. Sobol' Method

The Sobol' method implemented in DAKOTA is used in this study to obtain quantitative measures on how the uncertainty in model outputs can be apportioned to uncertainty in individual input variables. The total unconditional variance, $V(Y)$ (Y being model output), can be decomposed as (Saltelli et al., 2002)

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \cdots + V_{12\ldots k}. \tag{2}$$

where V_i is the first order term of factor X_i that contributes to the variance of Y, V_{ij} is the second order terms that captures the contribution to output variance from X_i and X_j that is not described by the first order terms, and $V_{12\ldots k}$ is the term of order k. The decomposition of equation (2) has $2k - 1$ terms, k being the number of model parameters. The first k terms are the variance Vi of the first order, the next $k(k-1)/2$ terms are variance V_{ij} of the second order, and so on, till the last term of order k. Dividing every item of the variance decomposition by $V(Y)$ gives the global sensitivity indices (Saltelli et al., 2002)

$$S_i = \frac{V_i}{V(Y)} \tag{3}$$

$$S_{ij} = \frac{V_{ij}}{V(Y)} \tag{4}$$

$$TS_i = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i} \sum_{m>j} S_{ijm} + \ldots + S_{12\ldots k} \tag{5}$$

where S_i is the first-order sensitivity index of X_i (also called main effect), S_{ij} is the second-order sensitivity index which measures the interaction effect between X_i and X_j, and TS_i is the total sensitivity index of X_i (also called total effect) that includes all sensitivity indices pertaining to X_i. In the Sobol' method (Sobol', 1990), V_i and V_{ij} are calculated via

$$V_i = V\left(f_i(X_i)\right) = V_{X_i}\left[E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right] \tag{6}$$

$$\begin{aligned} V_{ij} &= V\left(f_{ij}(X_i, X_j)\right) \\ &= V_{X_i X_j}\left(E_{\mathbf{X}_{\sim ij}}(Y \mid X_i, X_j)\right) - V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) \\ &\quad - V_{X_j}\left(E_{\mathbf{X}_{\sim j}}(Y \mid X_j)\right) \end{aligned} \tag{7}$$

where V_{Xi} (*) and E_{Xi} (*) is variance and mean of argument (*) with respect to X_i ; $V_{X\sim i}$ (*) and $E_{X\sim i}$ (*) is variance and mean of argument (*) taken over all factors but X_i . Different formulas have been proposed to calculate these factors, and more details are referred to Saltelli et al., (2002). The sensitivity indices reveal immediately the percentage of output variance is contributed by each parameter and its interaction with other parameters.

2.4. Latin Hypercube Sampling Method

As discussed in Introduction, the Latin Hypercube Sampling (LHS) method is used to generate random parameter samples based on their probability distributions. When sampling *M* samples for a single parameter, the sampling procedure is as follows (Swiler and Wyss, 2004):

(1) The cumulative distribution function (CDF) is divided into *M* segments of equal width 1/*M*, considering that the range of CDF is between 0 and 1.

(2) For each segment, a probability value is randomly sampled using a uniform distribution.

(3) Map the probability value to the parameter space according to the probability distribution function of the parameter.

If there are *N* random parameters, repeating this procedure for each parameter leads to *N* vectors of parameters, each of which has *M* numbers. If the parameters are correlated, the order of the *M* numbers in each vector is adjusted until the correlation is achieved. Note that LHS uses the Spearman rank correlation, instead of the Pearson linear correlation, because the former is applicable for nonlinear correlation. If the parameters are not correlated (as assumed in this study), the ordering is also performed to ensure that there is no correlation between the parameters. In the LHS sampling, the random samples are generated from the entire parameter ranges, even with a relatively small sample number.

3. SITE DESCRIPTION AND CHARACTERIZATION OF PARAMETRIC UNCERTAINTY

The ArcNLET-based sensitivity analysis and uncertainty assessment is conducted for three neighborhoods in Jacksonville, Florida, where nitrate loads from OSTDS to surface water bodies are of interest to environmental management. The three neighborhoods are Lakeshore, Julington Creek, and Eggleston Heights, and their locations are shown in Figure 3-1. The sensitivity analysis of three methods is conducted only for the Lakeshore neighborhood, because there is only one surface water body there so that the results of sensitivity analysis can be fully understood. At Julington Creek and Eggleston Heights neighborhoods, multiple surface water bodies exist, which makes it difficult to understand the relations between model parameters and nitrate load estimate. For the latter two neighborhoods, only uncertainty analysis is conducted. The observations of hydraulic head and nitrate concentration are incorporated into the uncertainty assessment. Meanwhile, there is no observation of hydraulic head and nitrate concentration at Lakeshore.

This section gives brief description of the three sites and characterization of parametric uncertainty. Results of sensitivity analysis and uncertainty assessment are given in the next section.

3.1. Lakeshore Neighborhood

The Lakeshore neighborhood is located west to the St. Johns River (Figure 3-1). It 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; to the north, it is partially bounded by Big Fishweir Creek. As shown in Figure 3-2, there are a total of 265 septic systems located in this site and one target water body (confluence of the Cedar and Ortega rivers in the south). The SSURGO database indicates that there are nine soil zones in this area, which are delineated by the black polygons shown in Figure 3-2. The FID of each soil zone is labeled in green. Porosity values of each zone obtained from the SSURGO database are labeled in red.

Table 3-1 lists the random parameters, their distributions, and statistics that define the distributions. They characterize the parametric uncertainty and are used in the sensitivity analysis and Monte Carlo simulations. The reasons of selecting the distributions and their defining statistics are discussed below:

(1) **Hydraulic conductivity**. Hydraulic conductivity of the nine soil zones are treated as random variables. They are denoted as hy_conX in Table 3-1, X being the FID of each soil zone adopted from the SSURGO database. The database provides the high, low, and representative values for each zone; these are listed as maximum, minimum and mode values in Table 3-1. Although hydraulic conductivity follows lognormal distribution, the statistics in the SURROGO database do not support the lognormal distribution. For example, the minimum and maximum values are more or less symmetric with the mode values. Since the SSURGO database is site specific, the distribution of hydraulic conductivity is assumed to be triangular, whose two ends are the minimum and maximum values and the peak is the mode value.

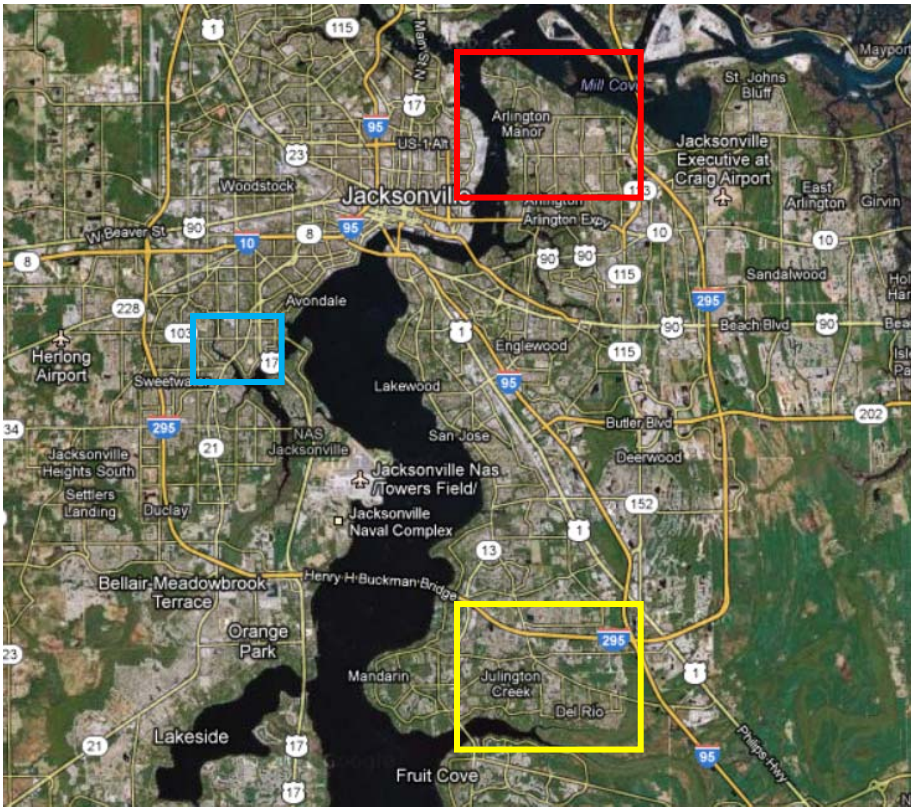

Figure 3-1. Locations of the Lakeshore (blue square), Julington Creek (yellow square) and Eggleston Heights (read square) neighborhoods. (map source: Google maps)

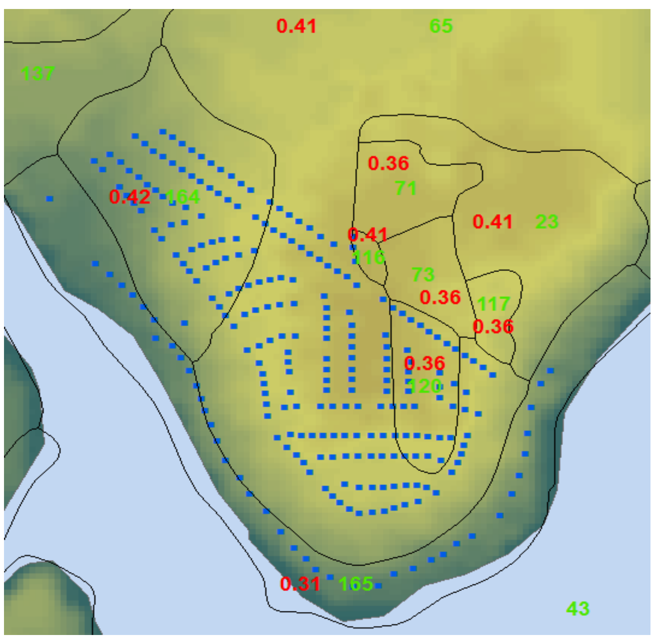

Figure 3-2. Locations of septic systems (blue square) and soil zones (black polygon) in Lakeshore neighborhood. FIDs of soil zones are labeled in green and soil porosity in red. The background is digital elevation model (DEM) of this area.

(2) **Longitudinal dispersivity** (α_L) and horizontal transverse dispersivity. The longitudinal dispersivity is assumed to follow normal distribution, and Table 3-1 lists its 1% and 99% percentiles as minimum and maximum, respectively (the mean value can be calculated as the average of the two percentiles). These values are determined based on literature data of Davis (2000) and (Gelhar et al. 1992), as field specific data are not available. Davis (2000) used a longitudinal dispersivity value of about 2.134m. Since longitudinal dispersivity may vary two orders of magnitude (Gelhar et al., 1992), the range of longitudinal dispersivity is determined as 0.21 m – 21.34 m. Following the tradition of contaminant transport modeling, the horizontal transverse dispersivity is assumed to be 10% of the longitudinal dispersivity. Based on this relation, when a random longitudinal dispersivity is generated, the horizontal transverse dispersivity is calculated.

(3) **Source nitrate concentration** (C_0). It is assumed that the source nitrate concentration follows normal distribution, and its range (again the 1% and 99% percentile) is based on the review article of McCray et al. (2005), as no site specific data is available.

(4) **First-order denitrification coefficient** (k_{den}). This parameter, as shown in the next section, is the most critical parameter to the load estimation. Based on a literature review, McCray et al. (2005) gave a range of 0.004 - 2.27 1/d for this parameter (Figure 3-3). Heatwole and McCray (2007) plotted the cumulative frequency distribution of the literature data (Figure 3-4) and showed that the parameter follows lognormal distribution. Figure 3-4 shows that although the possible maximum value is as large as 2.27/d, more than 80% of the data is below 0.5/d and 90% below 1.0/d, which indicates that the value is not likely to be larger than 1 in most occasions. Heinen (2006) summarized the first-order denitrification values used in several models (Figure 3-5). Among the models, since models DAISY, DRAINMOD-N and SWATNIT have the same unit (mass per volume) of nitrate-nitrogen used in ArcNLET, k_{den} values used in their models are of special value to this study. The range used in DRAINMOD-N is 0.004-1.08/d, which supports the observation above that the values of k_{den} are unlikely to be larger than 1.0. In addition, the values used in model DAISY and SWATNIT (0.1 and 0.01/d, respectively) fall in the range of 0.004-1.08. Therefore, the range of the first-order denitrification coefficient in this study is determined to be 0.004-1.08/d.

(5) **Smoothing factor** (smthF). This parameter is specific to ArcNLET and it is used to smooth the topography to generate the shape of the water table. Larger smoothing factor results in smoother water table shape and thus smaller hydraulic gradient. The parameter is site specific and strongly correlated to site topography. A fine resolution of topography requires large value of smoothing factor. Rios (2010) tested several smoothing factors for a groundwater model developed for the U.S Naval Air Station (NAS), Jacksonville. The model was calibrated by Davis et al. (1996) and the calibrated water table was used as the reference to evaluate the best smoothing factor. It was found that a value of 50 yielded a good approximation to the water table. In this study, the smoothing factor is assumed to follow uniform distribution and the range is set as 20 ~ 80 empirically.

Table 3-1. Probability distributions and their defining statistics of random parameters used in sensitivity and uncertainty analysis of Lakeshore neighborhood.

Parameter (unit)	Distribution	Minimum	Mode	Maximum
hy_con23 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con65 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con71 (m/d)	TRIANGULAR	0.122	0.6705	1.2198
hy_con73 (m/d)	TRIANGULAR	0.122	0.6705	1.2198
hy_con116 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con117 (m/d)	TRIANGULAR	0.122	0.6705	1.2198
hy_con120 (m/d)	TRIANGULAR	1.2198	6.696	12.1976
hy_con164 (m/d)	TRIANGULAR	0.122	0.6912	1.2198
hy_con165 (m/d)	TRIANGULAR	12.1824	21.3408	30.4992
C_0 (mg/L)	NORMAL	25		80
α_L (m)	NORMAL	0.21		21.34
k_{den} (/d)	LOGNORMAL	0.004		1.08
smthF (-)	UNIFORM	20		80

Process/Reaction Order	Median	Range	Number of Data
Nitrification—zero order, μ' (mg/L/d)	264	156 to 1464	7
Nitrification—first order, k_1 (per day)	2.9	0.0768 to 211.2	19
Denitrification—first order, k_2 (per day)	0.025	0.004 to 2.27	53

[1]Geng et al. (1996); Cho (1971); Ling and El-Kadi (1998); Yamaguchi et al. (1996); Starr et al. (1974); Starr and Gillham (1993); Misra et al. (1974); Ardakani et al. (1974a, 1974b); Lind (1983); Slater and Capone (1987); Anderson (1998); Trudell et al. (1986); Smith and Duff (1988); Bengtsson and Annadotter (1989); Francis et al. (1989); Obenhuber and Lowrance (1991); Smith et al. (1991); Ekpete and Cornfield (1965); Christensen et al. (1989); Bradley et al. (1992); Tesoriero et al. (2000); Smith et al. (1996); Lawrence and Foster (1986); Korom (1992); Hiscock (1991).

Figure 3-3. Range of first-order denitrification coefficient based on literature review. (source: Table 2 of McCray et al., 2005)

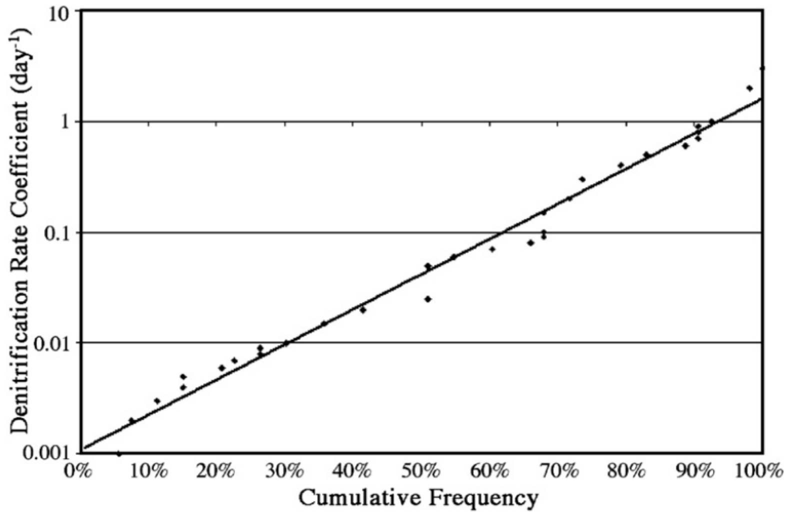

Figure 3-4. Cumulative frequency distribution for first-order denitrification coefficient. (source: Figure 4 of Heatwole and McCray, 2007)

k_d	Units	Units of N	Model
0.1	d^{-1}	$kg\ N\ m^{-3}$	DAISY
0.004–1.08	d^{-1}	$g\ N\ m^{-3}$	DRAINMOD-N
0.1	d^{-1}	$kg\ N\ ha^{-1}$	IMPACT
0.001–0.005	d^{-1}	$g\ N\ m^{-2}$	LEACHM
10	yr^{-1}	$mol\ N\ ha^{-1}$	NUCSAM
0.01	d^{-1}	$mg\ N\ L^{-1}$	SWATNIT

Figure 3-5. Values of first-order denitrification coefficient (k_d) used in several models (source: Table 3 of Heinen, 2006)

3.2. Julington Creek

The location of Julington Creek neighborhood is shown in Figure 3-1 **Error! Reference source not found.**. Julington Creek is bounded to the east and south by the Julington Creek. To the west, the neighborhood is bounded by the Oldfield Creek. As shown in Figure 3-6, there are a total of 587 septic systems located in this neighborhood. The SSURGO dataset indicates that there are eight soil zones in this area, whose boundaries are delineated by the black polygons shown in Figure 3-6. The FID and soil porosity of each zone are labeled in green and red respectively. In this neighborhood, 13 monitoring wells (shown as purple circles in Figure 3-6) were installed by the St. Johns River Water Management District (SJRWMD), and a total of 451 observations of water level depth and 484 observations of nitrate concentration have been collected from these wells 2003 – 2010. These data were used to calibrate the ArcNLET model by Wang et al. (2011b,c). In this study, these observations are used in the uncertainty analysis to exclude realizations in which the observations cannot be reasonably simulated. Details of this filtering process are discussed in section 4.

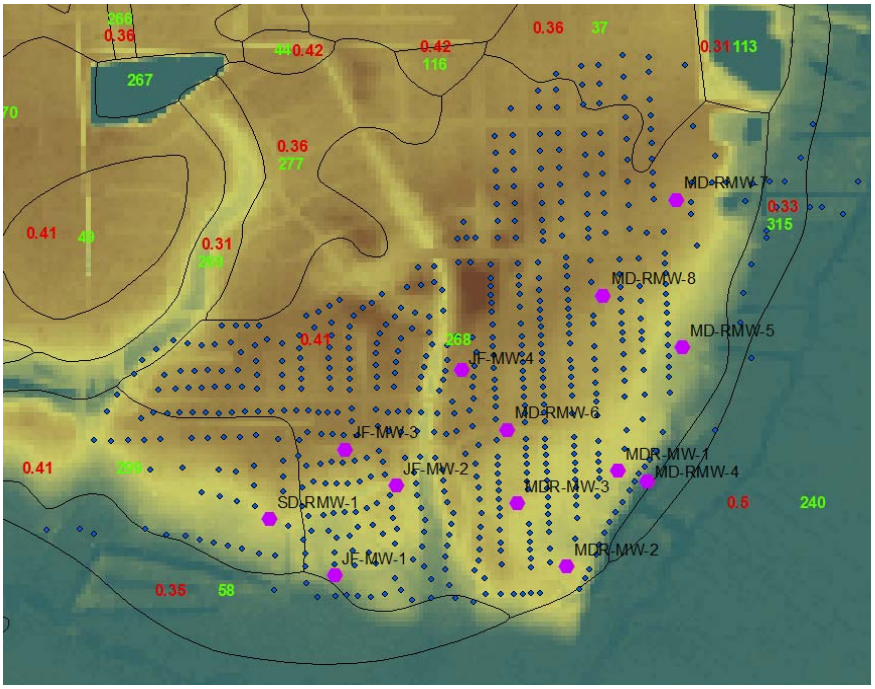

Figure 3-6. Locations of septic systems (blue squares), soil units (black polygons), and monitoring wells (purple circles) in Julington Creek neighborhood. FID and soil porosity of each soil zone are labeled in green and red, respectively. Names of monitoring wells are labeled in black. The background is a DEM of this area.

Table 3-2 lists the parameter distributions and the statistics that defines the distributions. While the parameter distributions are the same as those listed in Table 3-1, the defining statistics are not solely determined by literature data. Because model calibration was conducted for this neighborhood, the calibrated parameter values are incorporated in the process of determining the defining statistics. This is only done for source plane concentration and first-order denitrification coefficient, the most influential parameter to the nitrate load estimate, and doing so for other parameters is straightforward. The parameter distributions and their defining statistics are described below:

(1) **Hydraulic conductivity**. The parameter distributions and their defining statistics are determined from the SSUGRO database in the same manner of that of Lakeshore neighborhood.

(2) **Source plane concentration**. While this parameter is still assumed to follow normal distribution, its range for this neighborhood is 40 - 160 mg/L, larger than that of 25-80 mg/L used for Lakeshore neighborhood. The reason is that isotope data collected from this area September – October, 2010, suggests a high possibility of mixed nitrate sources from septic systems and fertilizer uses (Wang et al., 2011c). In order to simulate the observed nitrate concentration, the source plane concentration was increased to 100mg/L during the model calibration to empirically include the nitrate source from fertilizer use (Wang et al., 2011c). Since the literature-based range of 25-80mg/L is only for septic systems, it is considered to be necessary to enlarge the range to incorporate the fertilizer source

so that observed nitrate concentration can be simulated. The new range of 40 - 160 mg/L is minus/plus 60 mg/L from the calibrated value of 100 mg/L, assuming that 60 mg/L is sufficient to reflect variability of the source plane concentration.

(3) **Longitudinal dispersivity.** The probability of longitudinal dispersivity and its defining statistics are the same as those of Lakeshore neighborhood.

(4) **First-order denitrification coefficient**. The range of first-order denitrification coefficient used for Julington Creek is 0.004-0.036/d, which is smaller than that of 0.004-1.08/d used for Lakeshore neighborhood. The reason is that the calibrated value for Julington Creek neighborhood is 0.012/d (Wang et al., 2011b,c). By assuming that this is the mode of the lognormal distribution and taking 0.004/d as the minimum (i.e., 1% percentile), the maximum (i.e., 99% percentile) is calculated as 0.036/d. The uncertainty analysis (Section 4) confirms that this narrow range is reasonable, and details are given in section 4.2.

(5) **Smoothing factor.** Because observations of hydraulic head exist at the neighborhood, the smoothing factor can be estimated with more certainty. According to Wang et al. (2011c), the calibrated smooth factor is 100. This parameter is considered to be a deterministic variable and is fixed to the calibrated value of 100.

Table 3-2. Probability distributions and their defining statistics of random parameters used in sensitivity and uncertainty analysis of Julington Creek neighborhood.

Parameter (unit)	Distribution	Minimum	Mode	Maximum
hy_con37 (m/d)	TRIANGULAR	1.2198	6.696	12.1976
hy_con58 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con113 (m/d)	TRIANGULAR	12.1824	21.3408	30.4992
hy_con240 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con268 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con269 (m/d)	TRIANGULAR	0.122	0.6912	1.2198
hy_con299 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con315 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
C_0 (mg/L)	NORMAL	40		160
α_L (m)	NORMAL	0.21		21.34
k_{den} (/d)	LOGNORMAL	0.004	0.012	0.036
smthF (-)			100	

3.3. Eggleston Heights

The location of Eggleston Heights neighborhood is shown in Figure 3-1. Eggleston Heights is bounded to the west and north by the St. Johns River. For the domain of interest to this study, the targeted water body is the Red Bay Branch that is in dark blue of Figure 3-7.

As shown in Figure 3-7, there are a total of 393 septic systems located in this area. The SSURGO database indicates that there are four soil zones in the neighborhood; their boundaries are delineated by the black polygons in Figure 3-7, in which FID and soil porosity of each zone are labeled in green and red, respectively. SJRWMD installed four monitoring wells (represented as purple circles in Figure 3-7), and a total of 136 observations of hydraulic head and 143 observations of nitrate concentration have been collected from these wells during the period of 2003 – 2010. The observations were used to calibrate the ArcNLET model in Wang et al. (2011b,c). In this study, these observations are used in the uncertainty analysis to exclude realizations in which the observations cannot be reasonably simulated. Details of this filtering process are discussed in section 4.

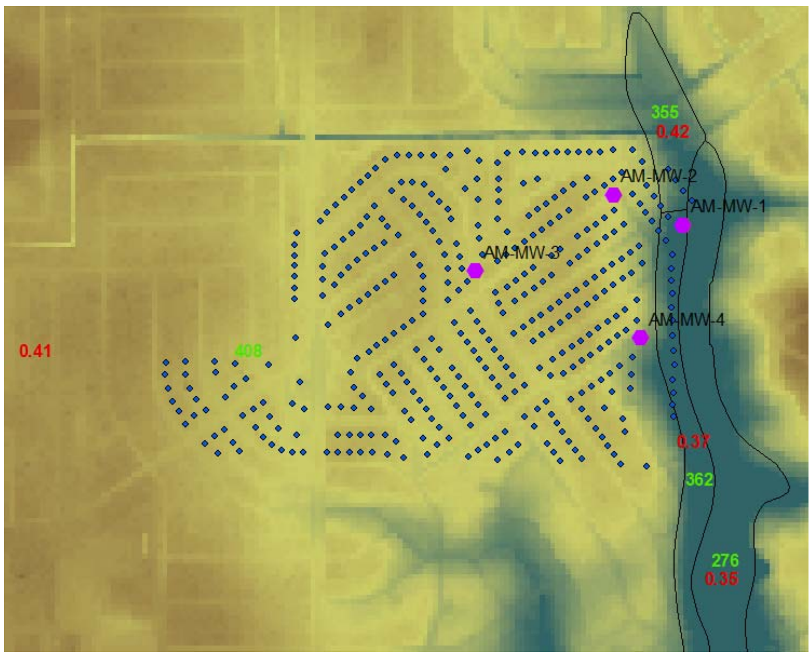

Figure 3-7. Locations of septic systems (blue square), soil units (black polygon with FID labeled as green number and soil porosity as red number), and monitoring wells (purple circles with well name labeled in black text) in Eggleston Heights neighborhood. The background is DEM of this area.

Table 3-3 lists the parameter distributions and the statistics that defines the distributions. While the parameter distributions are the same as those listed in Table 3-1, the defining statistics are not determined solely by literature data. Because model calibration was conducted for this neighborhood, the calibrated parameter values are incorporated in the process of determining the defining statistics. Similar to the procedure for Julington Creek neighborhood, only the calibrated first-order denitrification coefficient is used, which is discussed below:

(1) **Hydraulic conductivity**. The parameter distributions and their defining statistics are determined from the SSUGRO database in the same manner of that of Lakeshore neighborhood.

(2) **Source plane concentration**. The parameter distribution and its defining statistics are the same as those of Lakeshore neighborhood, because the isotope data collected from this area September – October, 2010, does not suggest other nitrate sources such as fertilizer use.

(3) **Longitudinal dispersivity.** The probability of longitudinal dispersivity and its defining statistics are the same as those of Lakeshore neighborhood.

(4) **First-order denitrification coefficient**. The range of first-order denitrification coefficient used for this neighborhood is 0.004-0.036/d, the same as that used for Julington Creek neighborhood. However, the reason of determining this range for this neighborhood is different from that for Julington Creek neighborhood. Given that the calibrated value of this parameter is 0.005/d (Wang et al., 2011), if the parameter range were determined in the same manner as that for Julington Creek neighborhood, the range would be 0.004-0.0062/d. This range is considered to be too small to cover feasible parameter values, which is confirmed in the uncertainty analysis presented in Section 4.4 in that the posterior range is 0.004-0.0082 after the filtering process. Therefore, the parameter range is determined to be the same as that of Julington Creek neighborhood.

(5) **Smoothing factor.** For the reason explained above for Julington Creek neighborhood, this parameter is considered to be a deterministic variable and its value is fixed at 60 obtained from the model calibration of Wang et al. (2011c).

Table 3-3. Probability distributions and their defining statistics of random parameters used in sensitivity and uncertainty analysis of Eggleston Heights neighborhood.

Parameter (unit)	**Distribution**	**Minimum**	**Mode**	**Maximum**
hy_con276 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
hy_con355 (m/d)	TRIANGULAR	0.122	0.6912	12.198
hy_con362 (m/d)	TRIANGULAR	0.0367	0.0864	1.2198
hy_con408 (m/d)	TRIANGULAR	3.6593	7.9488	12.1976
C_0 (mg/L)	NORMAL	25		80
α_L (m)	NORMAL	0.21		21.34
k_{den} (/d)	LOGNORMAL	0.004	0.012	0.036
smthF (-)			60	

4. RESULTS

This section presents results of sensitivity analysis and uncertainty quantification for Lakeshore neighborhood. The conclusions of sensitivity analysis for Lakeshore neighborhood are expected to be applicable to Julington Creek and Eggleston Heights neighborhoods. Therefore, for the latter two neighborhoods, only uncertainty assessment is performed to save computational cost.

4.1. Results of Sensitivity Analysis for Lakeshore Neighborhood

For Lakeshore neighborhood, local sensitivity analysis is performed using the qualitative method, followed by the global sensitivity analysis first using Morris method and then the more quantitative Sobol' method.

4.1.1. Qualitative sensitivity analysis

Figure 4-1 plots nitrate load estimates with different values of model parameters: first-order denitrification coefficient (k_{den}), hydraulic conductivity (K), smoothing factor (smthF), source plane concentration (C_0), longitudinal dispersivity (α_L), and porosity (Φ). For an individual parameter, its values used in the sensitivity analysis are generated manually to cover its range listed in Table 2-1. The following are observed from Figure 4-1:

(1) The estimated nitrate load increases with hydraulic conductivity, source plane concentration, and longitudinal dispersivity. Because it is assumed that horizontal transverse dispersivity is 1/10 of longitudinal dispersivity, the dispersivities are changed simultaneously.

(2) The estimated nitrate load decreases when smoothing factor and soil porosity increase.

(3) The estimated nitrate load has linear relations with source plane concentration and longitudinal dispersivity.

(4) The relation between the estimate nitrate load and the first-order denitrification coefficient is more complicated. When k_{den} increase from 0.004 to 0.01, the estimated load decreases from 40 to 5. At the parameter range of 0.01 ~ 0.06, the decrease is slow and some increase because of nonlinear relations between the parameter and nitrate concentration as well as the mass of nitrate influx and denitrification. When the value is larger than 0.06, the estimated load is 0, indicating that there is no nitrate plume reaching the surface water bodies.

While Figure 4-1 is useful to understanding of qualitative relations between the estimated nitrate load and the model parameters, it is insufficient to draw definitive conclusions on parameter sensitivity, because the figure does not provide a quantitative measure of parameter sensitivity. One can only conclude that porosity is least influential to the load estimate. While quantitative measures of local sensitivity can be calculated based on the figure, we prefer not to do so because of the limitations of local sensitivity. Instead, we conduct the quantitative global sensitivity analysis below.

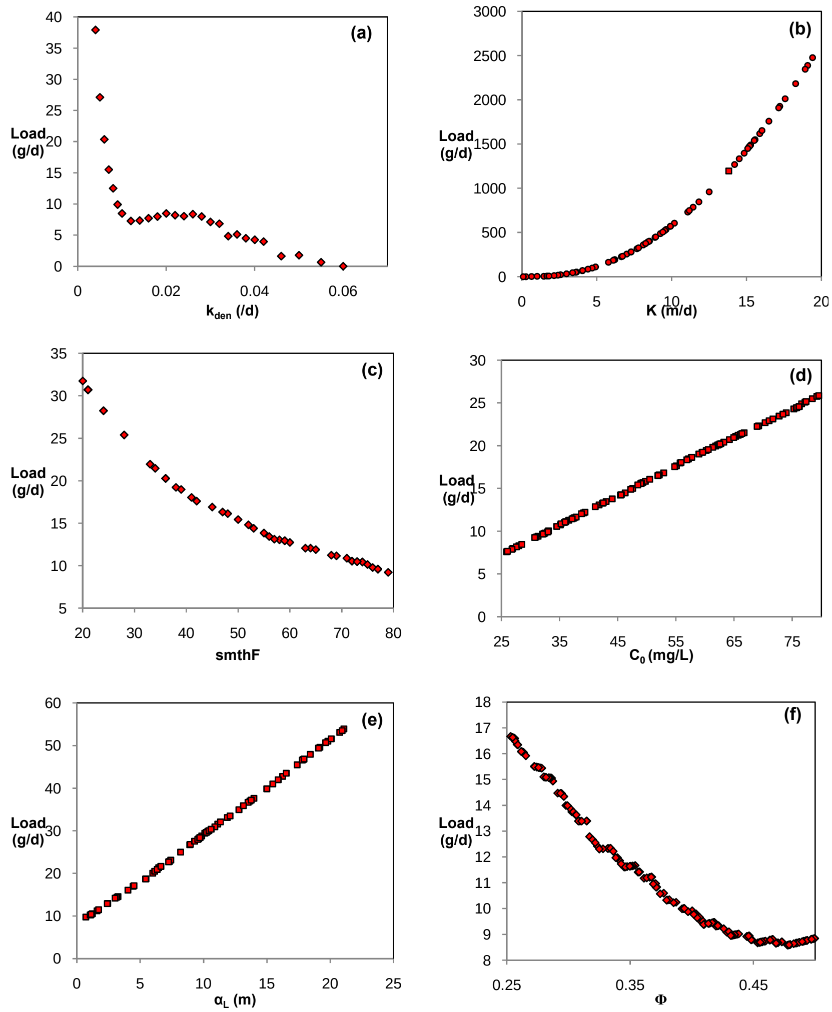

Figure 4-1. Variation of nitrate loads estimate with (a) first-order denitrification coefficient, (b) hydraulic conductivity, (c) smoothing factor, (d) source plane concentration, (e) dispersivity, and (f) soil porosity.

4.1.2. Results of Morris analysis

In the Morris sensitivity analysis, porosity is excluded because of its least influence on the load estimate, as discussed above. In addition, instead of using homogeneous hydraulic conductivity values in the last section, hydraulic conductivities of all the nine soil zones are considered in the Morris sensitivity analysis.

Table 4-1 lists the calculated mean (μ) and standard deviation (σ) of the elementary effects for all the parameters, and the values are plotted in Figure 4-2. As discussed in Section 2.2, parameters with large mean values have overall important influence on the load estimation. Parameters have large standard deviation values have nonlinear effect or interaction with other parameters, which cannot be revealed from the previous local sensitivity analysis. Based on Table 4-1 and Figure 4-2, the parameter sensitivity can be categorized into four groups as follows:

(1) The first-order denitrification coefficient is the most influential parameter to the load estimate, because the parameter has the largest mean and standard deviation

of the elementary effect. The large mean elementary effect is reasonable, because the parameter determines the loss of nitrate due to denitrification. The large standard deviation is caused by both nonlinear effect shown in Figure 4-1(a) and interactions between this and other parameters. Because of the interaction, the denitrification coefficient needs to be adjusted together with other parameters during model calibration.

(2) The group of second most important parameters includes longitudinal dispersivity, source plane concentration, and hydraulic conductivity of the soil zone with FID 165. For these three parameters, our conclusions are as follows:

1) For longitudinal dispersivity and source plane concentration, since Figure 4-1 of local sensitivity analysis suggests linear relations between the two parameters and the load estimate, the large standard deviation of two parameters indicates interaction with other parameters that influence the load estimate. Therefore, the two parameters need to be adjusted with other parameters during model calibration.

2) Figure 4-2 shows that the source plane concentration has smaller mean but larger standard deviation than the longitudinal dispersivity. It indicates that, in comparison with the longitudinal dispersivity, the source plane concentration has less overall impact on load estimation but stronger interaction with other parameters.

3) Hydraulic conductivity is not identified as the most influential parameter. Although this is inconsistent with Figure 4-1(b), the Morris results are more reasonable, considering that it is assumed that hydraulic conductivity is homogeneous over the entire neighborhood in the local sensitivity analysis.

4) Among all hydraulic conductivities of all soil zones, that of soil zone 165 is the most important parameter. This is physically reasonable, because as shown in Figure 3-2. Locations of septic systems (blue square) and soil zones (black polygon) in Lakeshore neighborhood. FIDs of soil zones are labeled in green and soil porosity in red. The background is digital elevation model (DEM) of this area.Figure 3-2, this soil zone is along the rivers and all the flow paths pass through it before reaching the rivers. In other words, it controls the nitrate travel time from all septic systems. In addition, hydraulic conductivity of this soil zone has the largest variability as indicated by the largest range listed in Table 3-1. This result suggests that, in order to have reliable nitrate load estimation and reduce predictive uncertainty, it is critical to have parameter measurements and field observations of heads and nitrate concentrations within or near this soil zone.

(3) Smoothing factor and hydraulic conductivity of soil zone 65 also have influence on the load estimation. This is physically reasonable, because the smoothing factor determines the shape of water table and soil zone 65 covers the most area of the modeling domain (Figure 3-2).

(4) The group of least influential parameters includes hydraulic conductivity of soil zones other than zones 165 and 65. This again suggests importance of considering heterogeneity of hydraulic conductivity. On the other hand, due to the small sensitivity to these zones, collecting more data from these zones may not be valuable.

In a summary, the Morris sensitivity analysis provides important insights regarding influence of model parameters on the load estimate that can be used for guidelines of model calibration and data collection.

However, there is still an unanswered question, i.e., to what extent is one parameter more influential than another parameter? For example, the first-order denitrification coefficient is identified as the most influential parameter. Is it overwhelmingly or slightly more influential than other parameters? To answer this kind of question requires conducting more quantitative sensitivity analysis such as the Sobol' method below.

Table 4-1. Calculated values of mean and standard deviation of the elementary effects in Morris analysis.

Parameter	**Mean (μ)**	**Standard deviation (σ)**
hy_con23	1.04E+001	9.29E+001
hy_con65	4.92E+004	1.61E+005
hy_con71	0.00E+000	0.00E+000
hy_con73	0.00E+000	0.00E+000
hy_con116	0.00E+000	0.00E+000
hy_con117	7.84E+000	7.55E+001
hy_con120	2.35E+003	1.43E+004
hy_con164	1.72E+002	1.02E+003
hy_con165	3.77E+005	4.79E+005
C_0	3.97E+005	4.66E+005
α_x	4.61E+005	3.47E+005
k_{den}	5.11E+005	8.77E+005
smthF	2.18E+005	2.81E+005

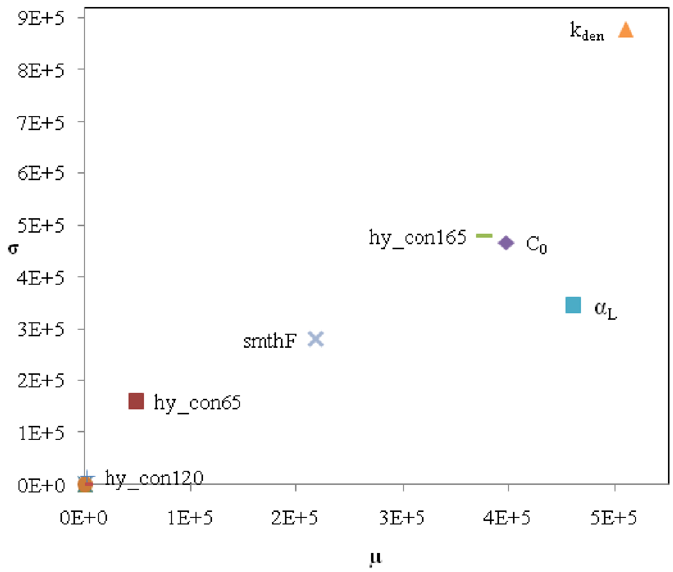

Figure 4-2. Mean and standard deviation of the elementary effects.

4.1.3. Results of Sobol' method

The Sobol' method is computationally expensive and requires thousands of executions of the ArcNLET model. To save computational cost, based on the results of Morris sensitivity analysis, hydraulic conductivities of the soil zones are fixed at their SSURGO representative values (listed as mode in Table 3-1), except that of soil zone 165 because of its large influence on the load estimate. Using Morris method before Sobol' method to exclude least influential parameters is typical in sensitivity analysis (e.g., Zhao et al., 2011).

Table 4-2 lists the first-order (or main effect) and total (or total effect) sensitivity indices calculated in Sobol' analysis. As discussed in Section 2.3, the main effect measures sensitivity to an individual parameter and the total effect considers interactions between an individual parameter and other parameters. The table shows that:

(1) The first-order denitrification coefficient is the overwhelmingly most critical parameter for load estimation. Uncertainty of this parameter contributes 66.3% to the uncertainty of the load estimation. In other words, 66.3% of the variance of the load estimate is due to variability of this parameter. This parameter also interacts with other parameters, and the interaction contributes 5.7% (72.0% - 66.3%) to the variance of the load estimation.

(2) Among all the parameters, the interaction between hydraulic conductivity of soil zone 165 and other parameters is the strongest. While the contribution from this parameter is 4.5%, the interaction contributes 7.5% to the predictive uncertainty.

(3) Between source plane concentration and longitudinal dispersivity, the former has stronger interaction with other parameters, because the difference between of total effect and main effect is 6.7% (9.8%-3.1%) for source plane concentration but only 1.8% (3.9%-2.1%) for longitudinal dispersivity. This is consistent with the Morris results shown in Figure 4-2.

(4) The results of Sobol' sensitivity analysis indicate that the smoothing factor is the second most influential parameters (in terms of both main effect and total effect), which is different from the order given by the Morris results. While the reason is not exactly known, we believe that the results of Sobol' method are more reliable, because Morris method is designed as a screening method and less quantitative than the Sobol' method. Future research is warranted to fully understand the difference with regard to the smoothing factor.

Table 4-2. Main and total effect sensitivity indices.

Parameters	Main effect	Total effect
First-order denitrification coefficient (k_{den})	66.3%	72.0%
Smoothing factor (smthF)	6.6%	13.3%
Hydraulic conductivity of zone 165 (hy_con165)	4.5%	12.0%
Source plane concentration (C_0)	3.1%	9.8%
Longitudinal dispersivity (α_L)	2.1%	3.9%

4.2. Results of Uncertainty Analysis for Lakeshore Neighborhood

Predictive uncertainty of the nitrate load estimate is conducted using the MC methods. A total of 2,000 random samples of model parameters are first generated using the LHS method, and the ArcNLET model is executed 2,000 times for the parameter samples. Although insensitive parameters identified in the sensitivity analyses above can be excluded from the uncertainty analysis, all the parameters are included because the 2,000 MC simulations are computationally affordable in this study.

Figure 4-3 plots the histograms of the LHS-generated random samples of four model parameters. This figure is to demonstrate that the generated samples follow the parameter distributions specified in Table 3-1. For example, Figure 3-1(c) shows that the histogram follows the triangle distribution with the minimum of 12.44, mode of 21.8, and maximum of 30.24, which are close to the theoretical values of 12.18, 21.34, and 30.49, respectively. It indicates that the LHS method is capable of generating satisfactory random samples. This examination is also conducted for Julington Creek and Eggleston Heights neighborhoods, but results are not shown in the report.

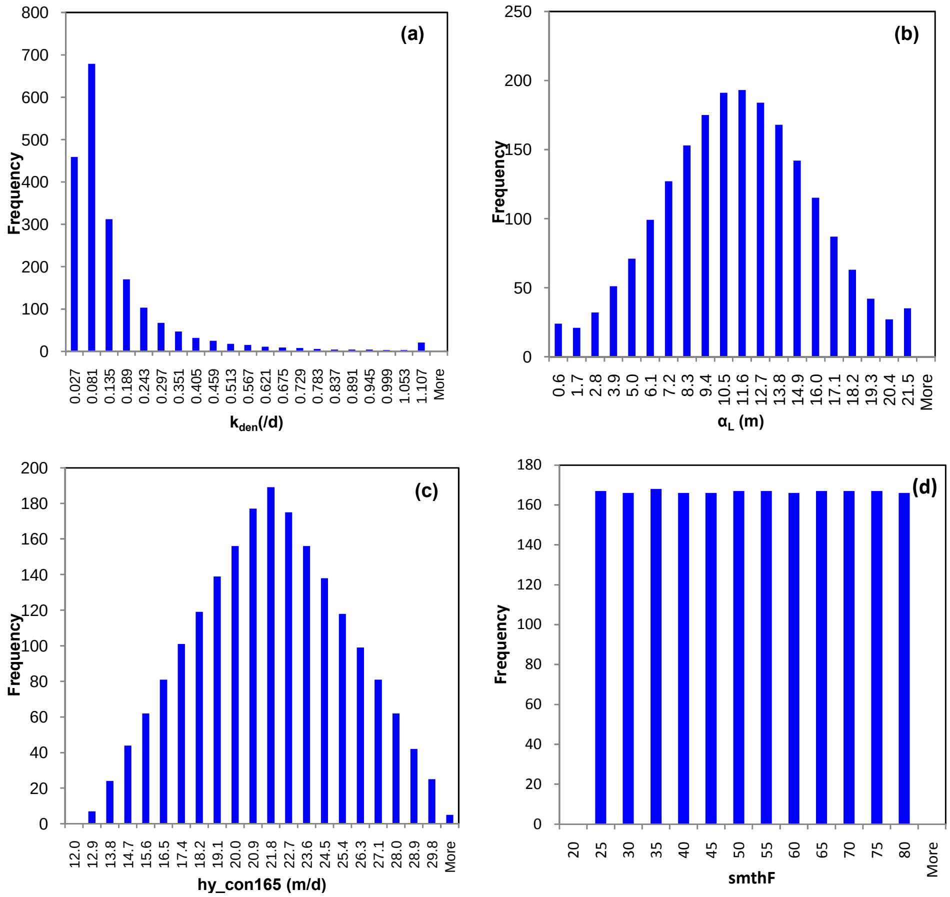

Figure 4-3. Histograms of random samples generated for (a) first-order denitrification coefficient (lognormal distribution), (b) longitudinal dispersivity (normal distribution), (c) hydraulic conductivity of soil zone 165 (triangular distribution) and (d) smoothing factor (uniform distribution).

Figure 4-4 plots the histogram of the nitrate load estimates of the 2,000 MC realizations. The histogram shows that the load estimate is highly uncertain. The histogram appears to be lognormal distribution, which may be attributed to the lognormal distribution of the first-order denitrification coefficient whose uncertainty contributes 72% to the variance of the load estimate (Table 4-2).

Table 4-3 lists the mean, median, standard deviation, minimum, and maximum of the 2,000 realizations. It also lists the nitrate load estimate obtained using the nominal parameter values listed in Table 2-1. The estimate using the nominal values is extremely small, only about 2% of the mean value. The major reason is that the nominal values of hydraulic conductivity and longitudinal dispersivity are significantly smaller than their mode values (Table 3-1) used for the random sampling and Monte Carlo simulation. As shown in Figure 4-1, smaller values of the two parameters correspond to smaller load estimates. Another reason is, when the nominal value is used, that hydraulic conductivity the entire domain is assumed to be homogeneous and it takes the small value of 2.113 m/d from literature. However, in the Monte Carlo simulation, site-specific information of hydraulic conductivity is used. The domain is divided into nine soil zones, each of which has its own hydraulic conductivity value. In particular, hydraulic conductivity of the most influential soil zone (FID 165) is large, in the range of 12.18 ~ 30.50. The increased hydraulic conductivity results in larger load estimate. This result indicates importance of site-specific data to accuracy of load estimation.

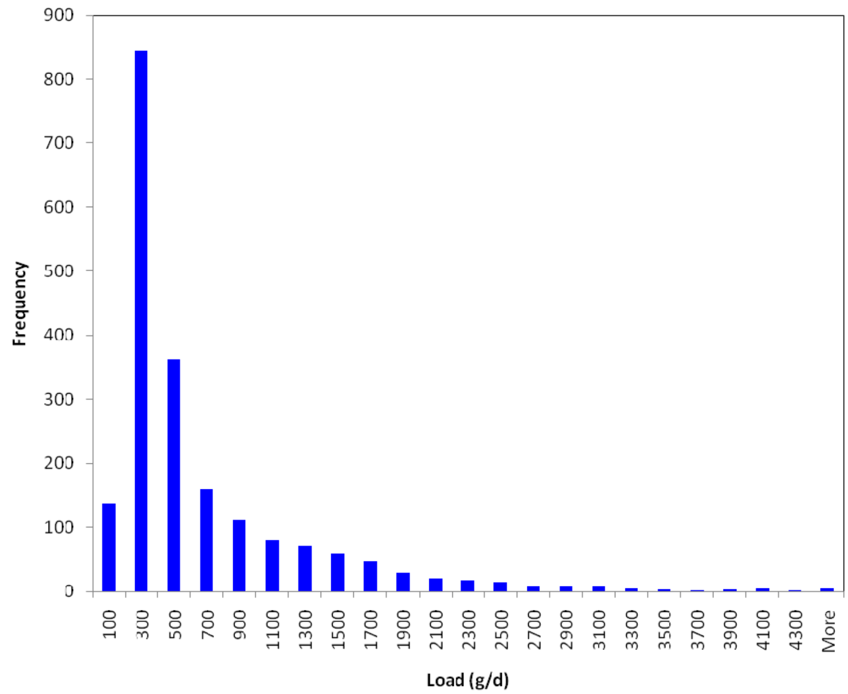

Figure 4-4. Histogram of nitrate load estimates of 2,000 MC realizations for Lakeshore neighborhood.

Table 4-3. Statistics of the loads estimates (g/d) of 2000 MC realizations for Lakeshore neighborhood and load estimates using nominal values (Table 2-1) and an empirical method developed by CDM-Smith consulting company.

Mean	Median	Standard deviation	Minimum	Maximum	Nominal	CDM
566	307	654	0	4653	12	5015

Table 4-3 also lists a load estimate using an empirical method developed by the CDM-Smith consulting company. According to a technical report (City of Jacksonville, 2008), the method first determines the number of septic systems that contribute nitrate into surface water bodies and then multiplies the number by nitrate load from each septic system. In the CDM method, the nitrate load from each septic system is 22.49 g/day, estimated by the formula 11.2 g/d/person × 2.51 persons/household × 80% (assuming that nitrate reduction is 20% for failing septic systems). The number of contributing septic systems is determined in the following three steps:

(1) Locate the neighborhoods of septic system failure areas based on the City of Jackson Master Stormwater Management Plan (MSMP).

(2) Create a 300-meter buffer zone around the conveyance system based on the National Hydrograph Database (NHD).

(3) Count the number of septic systems in the 300-meter buffer zone within the neighborhood of Step 1.

By implementing this empirical method for Lakeshore neighborhood, Figure 4-5 shows that 223 out of 265 septic systems are considered to contribute nitrate to the rivers. This leads to the nitrate load estimate of 5,015 g/d.

Figure 4-5. According to the empirical method developed by CDM-Smith, 223 out of 265 septic systems (blue squares) fall in the 300-meter buffer zone (green zone) and contribute nitrate to the rivers.

The main reason that the CDM-estimated load is larger than those of the 2,000 MC simulations is the consideration of denitrification in ArcNLET. Figure 4-6 plots the ArcNLET simulated plumes corresponding to the realization of the maximum load estimate of 4,653 g/d. In this realization, the first-order denitrification coefficient is 0.004/d, the lowest one (Table 3-1); hydraulic conductivity and dispersivity values are relatively high. Out of 265 septic systems, 205 septic systems contribute nitrate load to the rivers. Comparing Figure 4-5 and Figure 4-6 shows that the spatial distributions of contributing septic systems are similar. However, due to denitrification, the ArcNLET-estimated load is 7% smaller than that of CDM-estimated load.

Due to the large sensitivity of load estimate to the first-order denitrification coefficient, uncertainty in the parameter significantly affects the number of contributing septic systems. Figure 4-7 shows that when the denitrification coefficient increases from 0.004/d to 0.087/d, the number of contributing septic systems decreases from 205 to 69, and the 136 septic systems located inside the neighborhood do not contribute nitrate load to the rivers. Even for the 69 contributing septic systems, their contribution is smaller due to denitrification. In an extreme realization in which the first-order denitrification coefficient is 0.3954/d, none of the 265 septic systems contribute nitrate load to the rivers.

While the first-order denitrification coefficient is critical, site-specific measurements of this parameter are not available. The literature data is too uncertain to be useful to make reliable estimation of nitrate load. This problem may be partly resolved if observations of head and nitrate concentration are available, because the observations can be used to help better constraint this and other parameters. While there is no field observation at Lakeshore neighborhood, a relatively large amount of observations are available at Julington Creek and Eggleston Heights neighborhood. Use of them in the load estimation is discussed below.

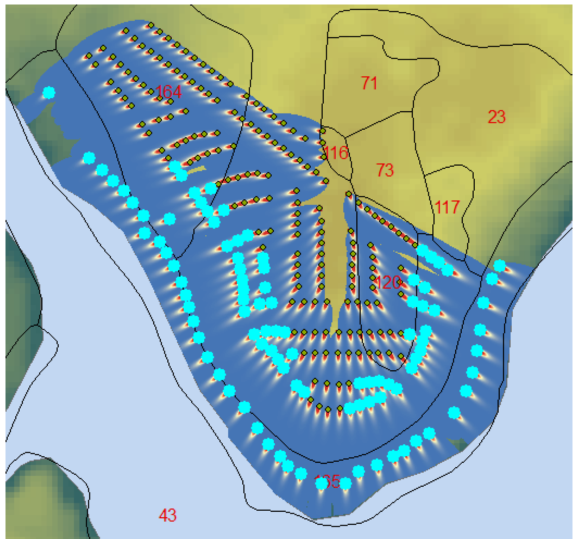

Figure 4-6. Arc-NLET simulated nitrate plumes with the following model parameters: 23.9 m/d for hydraulic conductivity of soil zone 165, 9.351 m for longitudinal dispersivity, 54.76 mg/L for source plane concentration, 0.004/d for first-order denitrification coefficient, and 24 for smoothing factor. The estimated nitrate load is 4,653 g/d from 205 (out of 265) septic systems marked by blue circles.

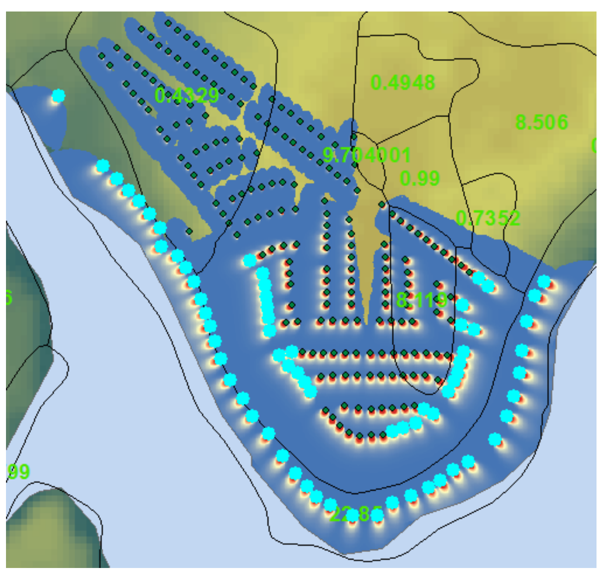

Figure 4-7. Arc-NLET simulated nitrate plumes with the following model parameters: 22.85 m/d for hydraulic conductivity of soil zone 165, 15.13 m for longitudinal dispersivity, 57.89 mg/L for source plane concentration, 0.087/d for first-order denitrification coefficient, and 30 for smoothing factor. The estimated nitrate load is 567 g/d from 69 (out of 265) septic systems marked by blue circles.

4.3. Results of Uncertainty Analysis for Julington Creek Neighborhood

As shown in Figure 3-6, 13 monitoring wells are available in this neighborhood, and a total of 451 observations of water level depth and 484 observations of nitrate concentration have been collected. The ArcNLET model was calibrated by adjusting model parameters to match these observations. The procedure and results of calibration can be found in Wang et al. (2011b,c). The calibrated value of the first-order denitrification coefficient has been used to help characterize the parametric uncertainty by narrowing the parameter range from 0.004 ~ 1.08 (3.2Table 3-1) to 0.004 ~ 0.036 (Table 3-2). In addition, the smoothing factor is fixed at the calibrated value (Table 3-2).

The field observations of concentration are further used in the uncertainty assessment to exclude MC realizations that cannot reasonably simulate the observations. This is achieved by a filtering process designed as follows: for an individual MC realization, if its simulated nitrate concentrations fall within the ranges of minimum and maximum observations at less than twelve wells, the realization is not used for uncertainty assessment. Repeat the filtering process for all the MC realizations, and the remaining realizations are used for uncertainty assessment. In other words, the realizations retained for uncertainty assessment can match observed nitrate concentrations at twelve or thirteen wells. The reason of using twelve wells (instead of all the thirteen wells) in the filtering process is that ArcNLET simulations may not match observed concentrations at all wells because ArcNLET is based on the simplified model and operated with limited data.

Figure 4-8 plots the histogram of nitrate load estimates of the 2,000 Monte Carlo realizations before the filtering process is applied. Similar to Table 4-3, Table 4-4 lists descriptive statistics of the realizations and the estimates using the nominal parameter values and the CDM method. The load estimate resulted from the manual calibration of Wang et al. (2011b,c) is also listed. The estimate is highly uncertain, varying from 178 to 5656 g/d. The load estimate corresponding to the nominal parameter values is again extremely small. The mean and median load values are close to the load estimate of model calibration, largely because the calibrated denitrification rate is used in the characterization of parametric uncertainty.

The maximum load estimate is still smaller than the CDM-estimated one, and the reason is again attributed mainly to consideration of denitrification in ArcNLET. For this neighborhood, ArcNLET and the CDM method have similar number of contributing septic systems. The CDM number is 369; the ArcNLET number is 354 for the calibrated model and 560 for the realization that gives the maximum load estimate.

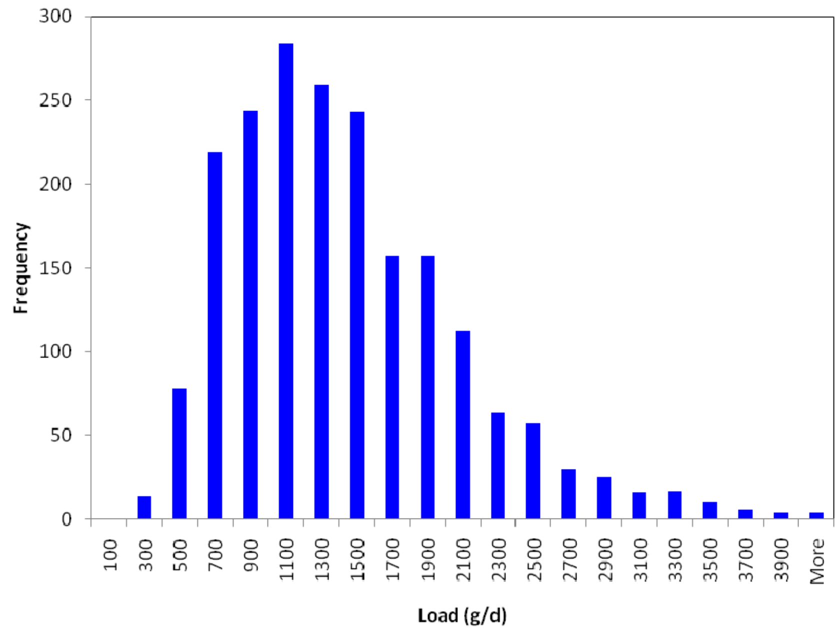

Figure 4-8. Histogram of estimated nitrate load from 2,000 MC realizations for Julington Creek neighborhood.

Table 4-4. Statistics of the loads estimates (g/d) before and after the filtering process for Julington Creek neighborhood and load estimates using nominal parameter values (Table 2-1), calibrated parameter values, and an empirical method developed by CDM-Smith consulting company.

	Mean	Median	St. Dev.	Min	Max	Nominal	Calibration	CDM
Before filtering	1334	1225	652	178	5656	91	1032	8292
After filtering	1504	1466	257	1049	2078	91	1032	8292

After applying the filtering process using field observations of nitrate concentration, only 19 realizations are retained. Figure 4-9 plots the ranges of observed nitrate concentrations and simulated concentrations of the 19 realizations at the wells. The figure shows that nitrate concentration at Well JF-MW-4 is underestimated in the 19 realizations. However, the estimation is not typical, because at several wells, ArcNLET-simulated nitrate concentrations are toward high observation values.

According to the histogram (Figure 4-10) and the statistics (Table 4-4) of the 19 realizations, uncertainty in the load estimate is significantly reduced by incorporating field observations. For example, the range of estimated loads is reduced from 178 – 5656 g/d to 1049-2078 g/d, and the standard deviation is reduced about 60%, from 652 to 257. The uncertainty reduction is also observed in the ranges of model parameters.

Table 4-5 lists the minimum and maximum parameter values before and after the filtering. The table shows that all the parameter ranges become smaller after the filtering. In particular, the range of the firs-order denitrification coefficient reduces from 0.004-0.036/d to 0.005-0.015/d. This confirms that it is reasonable to adjust the parameter range in Section 3.2 by changing the maximum value from 1.08 to 0.036.

Table 4-5 indicates that even the value of 0.036 is large to reasonably simulate the field observations. It is expected that the uncertainty in load estimate and parameters can be further reduced if more data are collected. However, the ranges may vary, depending on data collected in future. For example, if data of high nitrate concentration are collected, the load estimate may increase to reflect updated knowledge of the site.

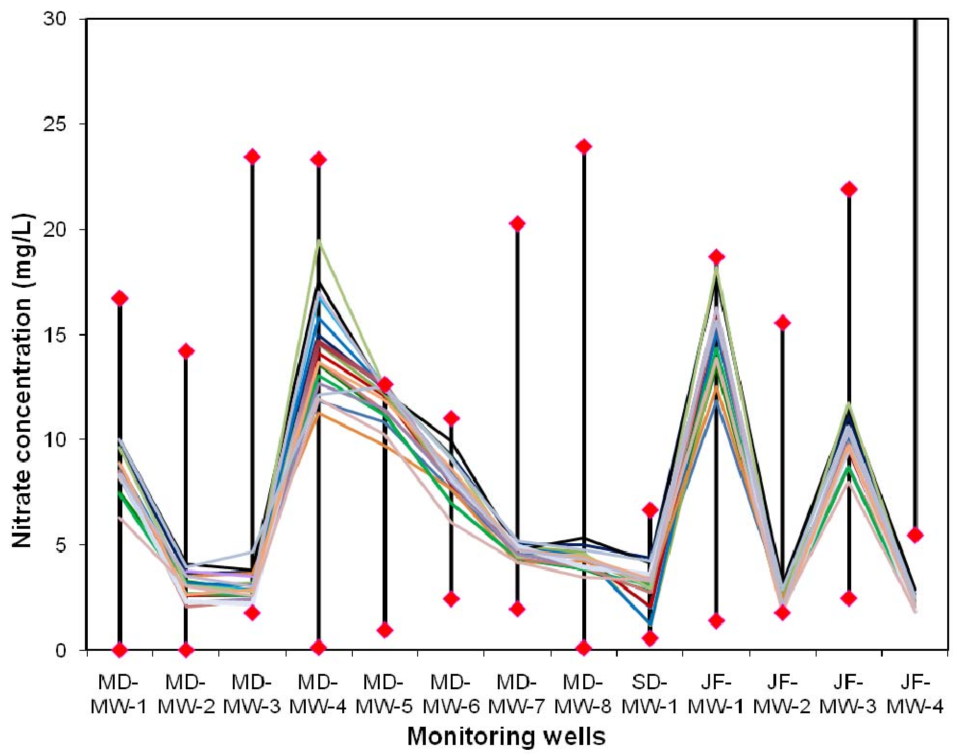

Figure 4-9. Ranges (between pink diamonds) of observed nitrate concentrations and simulated concentrations of the 19 realizations after the filtering process at the 13 monitoring wells.

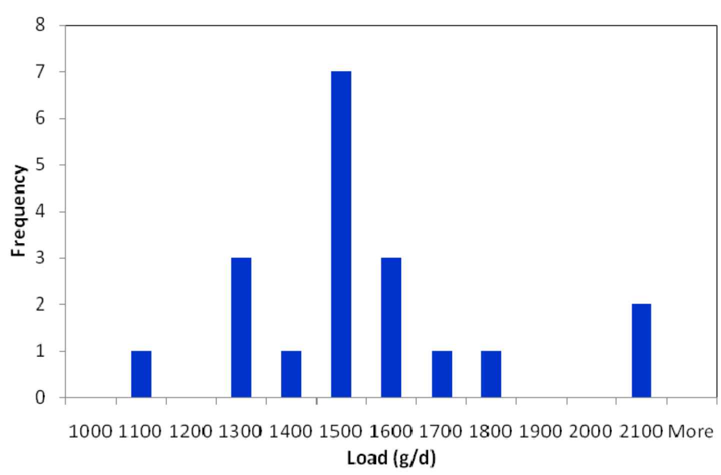

Figure 4-10. Histogram of nitrate load estimates of 19 MC realizations retained after the filtering process.

Table 4-5. Minimum and Maximum parameter values before and after the filtering process using filed observations.

Parameter	Before filtering		After filtering	
	Minimum	Maximum	Minimum	Maximum
hy_con37	1.220	12.198	2.505	10.880
hy_con58	3.659	12.198	6.132	11.300
hy_con113	12.182	30.499	12.600	29.210
hy_con240	3.659	12.198	4.218	10.420
hy_con268	3.659	12.198	4.839	11.380
hy_con269	0.122	1.220	0.359	1.075
hy_con299	3.659	12.198	5.748	11.080
hy_con315	3.659	12.198	3.772	11.470
C_0	40	160	40	121.5
α_L	0.21	21.34	12.25	21.34
k_{den}	0.004	0.036	0.005	0.015

4.4. Results of Uncertainty Analysis for Eggleston Heights Neighborhood

The uncertainty assessment for Eggleston Heights neighborhood is conducted in the same manner as that of Julington Creek neighborhood, except that 3,000 MC realizations are simulated because ArcNLET runs faster in Eggleston Heights neighborhood where the number of septic systems is about 2/3 of that of Julington Creek neighborhood. The histogram of nitrate load estimate of the 3,000 MC realizations is plotted in Figure 4-11, and the descriptive statistics are summarized in Table 4-6. The histogram has similar shape as those of Lakeshore and Julington Creek neighborhoods, but has smaller variability; the load estimate varies between 5 and 350 g/d. The load estimate corresponding to the nominal parameter values is again extremely small.

After applying the same filtering process to the 3,000 realizations using observations of nitrate concentration at the four monitoring wells shown in Figure 3-7, only 31 realizations are retained. Simulated nitrate concentrations in the 31 realizations are plotted in Figure 4-12 together with the ranges of the observations. The figure shows that the simulated concentrations are smaller than the observed. The reason of the underestimation is not entirely known, but it is attributes to the small hydraulic conductivity of soil zone 362 (Figure 3-7). Table 3-3 indicates that hydraulic conductivity of this zone is one order of magnitude smaller than that of other zones. The low hydraulic conductivity leads to low groundwater velocity, which in term gives low values of simulated nitrate concentration, as shown in Figure 3-5 of Wang et al. (2011b). To examine the ArcNLET simulated nitrate concentration and load, it is critical to investigate hydraulic conductivity of this soil zone in future field trips.

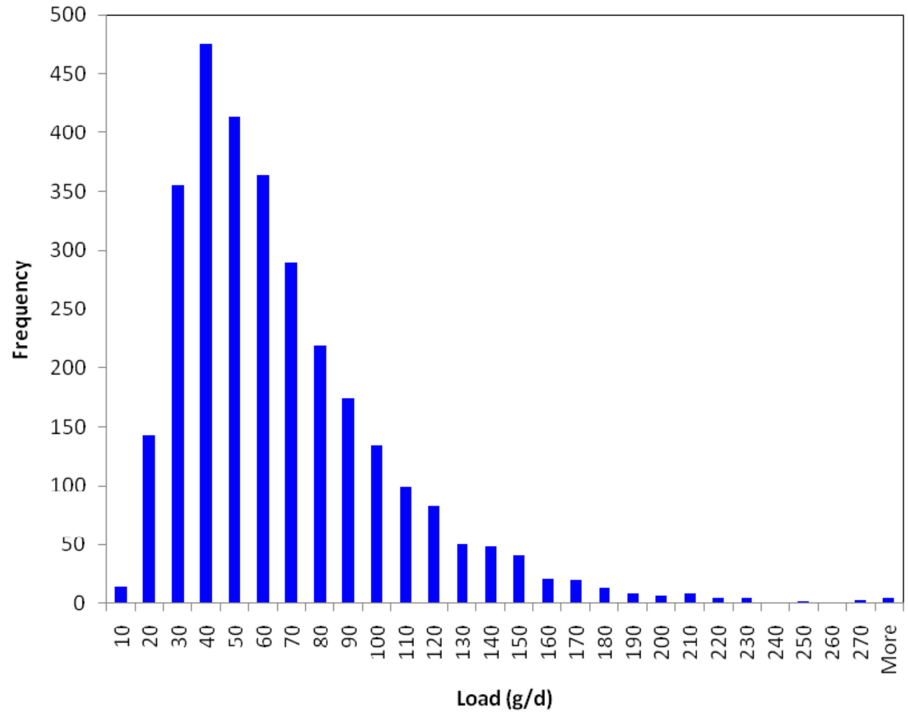

Figure 4-11. Histogram of estimated nitrate load from 3,000 MC realizations for Eggleston Heights neighborhood.

Table 4-6. Statistics of the loads estimates (g/d) before and after the filtering process for Eggleston Heights neighborhood and load estimates using nominal parameter values (Table 2-1), calibrated parameter values, and an empirical method developed by CDM-Smith consulting company.

	Mean	Median	St. Dev.	Min	Max	Nominal	Calibration	CDM
Before filtering	62	52	38	5	350	18	263	8838
After filtering	189	186	57	98	350	18	263	8838

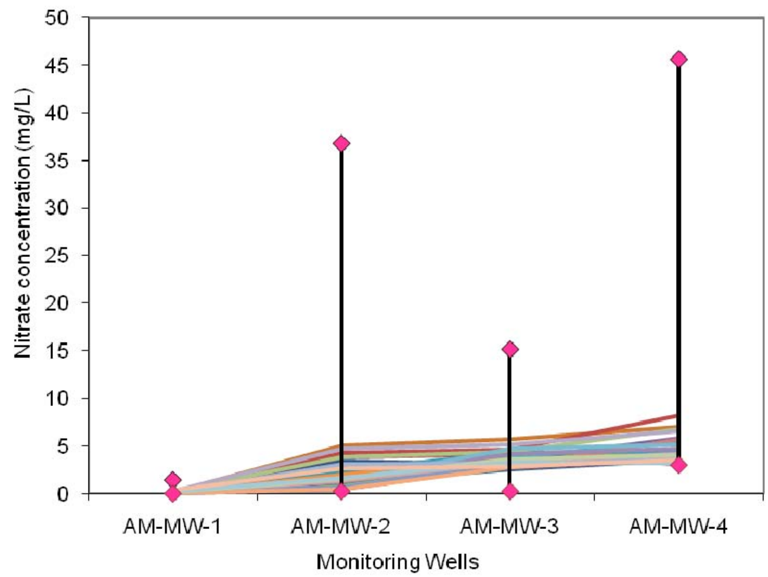

Figure 4-12. Ranges (between pink diamonds) of observed nitrate concentration and simulated concentrations of the 31 realizations after the filtering process at four monitoring wells.

The histogram of the retained 31 realizations is presented in Figure 4-13. As shown in Table 4-6, after the filtering, the range of estimated loads is reduced from 5-350 to 98-350 g/d. Comparing Figure 4-13 and Figure 4-11 shows that the majority of the 3,000 realizations are excluded; as a result, the median and mean values change significantly and the variance even increases after the filtering. It indicates that the parameter distributions and/or their defining statistics (listed in Table 3-3) are not reasonable, because most of the generated parameters cannot reasonably simulate the field observations. This however is not surprising, because the parameter distribution of first-order denitrification coefficient determined for the Julington Creek neighborhood (Table 3-2) is used for the MC simulations of Eggleston Heights neighborhood (Table 3-3). More discussion is given below. It is expected that this problem will disappear if site specific parameter distributions are used.

Table 4-7 lists the minimum and maximum parameter values before and after filtering. The table shows that high values of first-order denitrification coefficient are filtered out. For example, after the filtering, the maximum parameter value is 0.0082/d; it is even smaller than the mode of 0.012/d used for random number sampling. Recalling that the calibrated value of Julington Creek neighborhood is 0.012/d but the calibrated value of Eggleston Heights neighborhood is 0.005/d, it is not reasonable to use the parameter distribution of Julington Creek neighborhood. Instead, the parameter distribution specific to Eggleston Heights neighborhood should be determined and used for the uncertainty assessment. On the other hand, comparing

Table 4-5 and

Table 4-7 shows similarity in the ranges of longitudinal dispersivity. This suggests that it is possible to obtain representative distributions for certain parameters for different neighborhoods with similar hydrogeologic conditions. More research is warranted in future study.

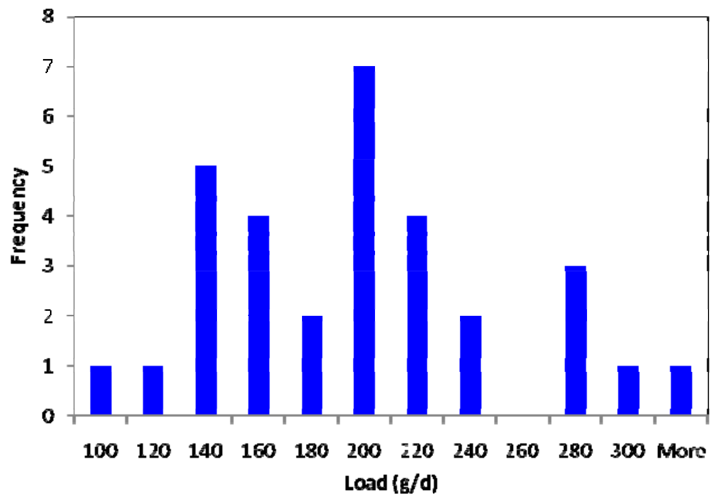

Figure 4-13. Histogram of nitrate load estimates of 31 MC realizations retained after the filtering process.

Table 4-7. Minimum and Maximum parameter values before and after the filtering process using filed observations.

Parameter	Before filtering		After filtering	
	Minimum	Maximum	Minimum	Maximum
hy_con276	3.659	12.198	4.222	11.67
hy_con355	0.122	12.198	1.023	9.733
hy_con362	0.0367	1.220	0.4506	1.196
hy_con408	3.659	12.198	4.793	11.65
C_0	25	80	36.5	80
α_L	0.21	21.34	7.109	21.34
k_{den}	0.004	0.036	0.004	0.0082

5. CONCLUSIONS

This report summarizes sensitivity analysis and uncertainty assessment for three neighborhoods at Jacksonville, FL. At the Lakeshore neighborhood, local sensitivity analysis is conducted using a qualitative method, and global sensitivity analysis using more the quantitative Morris method as well as a computationally demanding variance-based decomposition method that requires thousands of execution of ArcNLET. Except for the hydraulic conductivity and soil porosity from the SSURGO database, there is no site-specific data and information available at Lakeshore neighborhood. As a result, characterization of parametric uncertainty, sensitivity analysis, and uncertainty assessment are largely based on literature data. At Julington Creek and Eggleston Heights neighborhoods, field observations of hydraulic head and nitrate concentration are available, and they are used in characterization of parametric uncertainty and uncertainty assessment. The smoothing factor is determined solely by calibration against head observations. The calibrated values of first-order denitrification coefficient are used to determining defining statistics of lognormal distribution of the parameter. The observations of nitrate concentration are used to exclude MC realizations that cannot reasonably simulate the observations. Due to computational cost, sensitivity analysis is not performed for Julington Creek and Eggleston Heights neighborhoods.

The sensitivity analysis and uncertainty assessment lead to the following major conclusions:

(1) The first-order denitrification coefficient is the most critical parameter to ArcNLET-estimated nitrate load, and the relation between this parameter and load estimate is highly nonlinear. The results of variance-based decomposition method show that this parameter contributes more than 66% to predictive uncertainty in the load estimate. To obtain accurate load estimate and reduce predictive uncertainty, it is important to obtain direct measurements of this parameter and/or gather field observations of hydraulic head and nitrate concentration so that this parameter can be better estimated through model calibration.

(2) Interaction between model parameters is relatively strong. In particular, flow and transport parameters interact to affect simulated nitrate concentration and load estimate. This poses a challenge to model calibration, since multiple model parameters need to be adjusted simultaneously to improve goodness-of-fit between model simulations and field observations.

(3) Hydraulic conductivity of different soil zones has different degrees of influence on nitrate load estimate. The soil zones are important if a larger number of flow paths are through them. It is critical to identify the influential zones before conducting future field investigation so that data of maximum information content can be collected to improve accuracy of load estimation and reduction of predictive uncertainty.

(4) Predictive uncertainty of nitrate load estimate is significant for the three neighborhoods, especially when there is no field observation and uncertainty quantification is based solely on literature data. This uncertainty cannot be disregarded when using ArcNLET-calculated load estimate for science-informed decision-making.

(5) Field observations of hydraulic head and nitrate concentration are of great value to reduction of predictive uncertainty. It can help better characterize parametric uncertainty by using calibrated parameter values to determine defining statistics of parameter distributions. On the other hand, it can be used to exclude MC realizations that cannot reasonably simulate field observations. Incorporating field observations into uncertainty assessment can dramatically reduce uncertainty in load estimation.

(6) Consistency in parameter distribution longitudinal dispersivity is observed between Julington Creek and Eggleston Heights neighborhoods. This suggests that it is possible to identify representative parameter distributions for different neighborhoods with similar hydrogeologic conditions.

(7) Due to the large uncertainty for a single neighborhood, it is difficult to identify specific values (or default parameter values) representative for all neighborhoods. It has been found that the nominal parameter values give extremely small load estimate. On the other hand, due to large variability of predictive uncertainty between different neighborhoods, uncertainty assessment should be conducted for neighborhoods of similar hydrogeologic conditions. One cannot expect that uncertainty is similar for neighborhoods of different conditions. As a summary, it is inappropriate to use default parameter values for estimation of nitrate load of different neighborhoods and to use a uniform risk factor to quantify predictive uncertainty of different neighborhoods.

Given the conclusions above, we suggest collecting more measurements of model parameters (e.g., hydraulic conductivity and first-order denitrification coefficient) and observations of state variables (e.g., hydraulic head and nitrate concentration). The results of sensitivity analysis and uncertainty assessment provide insights into the future data collection, and ArcNLET can be used as a tool to facilitate model-guided field work plans. In addition, more advanced computational methods can be used to improve the uncertainty quantification. For example, we may use the Markov Chain Monte Carlo (MCMC) methods to estimate posterior parameter distributions. This method automatically incorporates field observations into the estimation and resolves the problem of unknown parameter ranges. However, field observations are still indispensible even if advanced computational methods are used.

6. REFERENCES

Adams, B.M., Dalbey, K.R., Eldred, M.S., Gay, D.M., Swiler, L.P., Bohnhoff, W.J., Eddy, J.P., Haskell, K., Hough, P.D. 2009. *DAKOTA, A Multilevel Parallel Objective-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis*, Version 5.0 User's Manual, Sandia National Laboratories, SAND2006-6337.

Andersson, F.O., Aberg, M., Jacobsson, S.P. 2000. Algorithmic approaches for studies of variable influence, contribution, and selection in neural networks. *Chemometrics and Intelligent Laboratory Systems 51(1)*: 61-72.

Barrett, T.S. 1996. *Probabilistic Design for Rotordynamic Simulations using Receptance-based Reanalysis*. PhD thesis, Texas A&M University.

Camilo, G.R. 1992. *Food Web Analysis of Macrobenthic Riffle Communities (South Llano River, Texas)*. PhD thesis, Texas Tech University.

City of Jacksonville. 2008. *Lower St. Johns River Basin Management Action Plan Project Specific Pollutant Load and Removal Evaluation.* Provided by the Florida Department of Environmental Protection.

Davis, J.H., Planert, M., 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.

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.

Fishman, G.S. 1996. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Verlag, New York.

Gelhar, L.W., Welty, C., Rehfeldt, K.R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research 28 (7)*:1955–1974.

Gwo, J.P., Toran, L.E., Morris, M.D., Wilson, G.V. 1996. Subsurface storm flow modeling with sensitivity analysis using a Latin-Hypercube Sampling technique. *Ground Water*. 4(5):811-818.

Haan, P.K., Skaggs, R.W. 2003a. Effect of parameter uncertainty on DRAINMOD predictions: I. Hydrology and yield. *Trans. ASAE 46(4)*: 1061-1067.

Haan, P.K., Skaggs, R.W. 2003b. Effect of parameter uncertainty on DRAINMOD predictions: II. Nitrogen loss. *Trans. ASAE* 46(4): 1069-1075.

Hazen and Sawyer, Environmental Engineers & Scientists. 2009. *Florida onsite sewage nitrogen reduction strategies study, task D.2, literature review of nitrogen fate and transport modeling*. Final Report.

Heatwole, K.K., McCray, J.E. 2007. Modeling potential vadose-zone transport of nitrogen from onsite wastewater systems at the development scale. *Journal of Contaminant 91*, 184-201.

Heinen, M. 2006. Simplified denitrification models: overview and properties. *Geoderma 133*:444–463.

Helton J.C., Davis F.J. 2003. Latin Hypercube Sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Safety, 2003;81:23–69.

Hession, W.C., Storm, D.E., Haan, C.T. 1996. Two-phase uncertainty analysis: An example using the Universal Soil Loss Equation. *Trans. ASAE* 39(4): 1309-1319.

Hitt, K.J., 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.

Husam Baalousha. 2006. Groundwater pollution risk using a modified Latin hypercube sampling. *Journal of Hydroinformatics* 08 (3):223-234.

Iman, R.L., Conover, W.J. 1980. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. *Communications in Statistics, Part A. Theory and Methods, 17*:1749-1842.

Isukapalli, S.S. 1999. Uncertainty Analysis of Transport-Transformation Models. PhD thesis, The State University of New Jersey.

Kalos, M.H., Whitlock, P.A. 1986. *Monte Carlo Methods : Basics*. John Wiley & Sons, New York.

Liou, T.S., Yeh, H.D. 1997. Conditional expectation for evaluation of risk groundwater flow and solute transport - one-dimensional analysis. *Journal of Hydrology.*, 199(3-4):378-402

Loh, W.L. 1996. On Latin Hypercube Sampling. *Annals of Statistics, 24(5)*:2058-2080.

Loucks, D.P., Beek, Eelco van. 2005. *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*. UNESCO Publishing, ISBN 92-3-103998-9.

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

Morris, M.D. 1991. Factorial Sampling Plans for Preliminary Computational Experiments, *Technometrics, 33(2),* 161-174.

Ogle, S.M., Breidt, F.J., Eve, M.D., Paustian, K. 2003.Uncertainty in estimating land use and management impacts on soil organic carbon storage for U.S. agricultural lands between 1982 and 1997. *Global Change Biology 9(11)*: 1521-1542.

Osnes, H. 1998. Stochastic analysis of velocity spatial variability in bounded rectangular heterogeneous aquifers. *Advances in Water Resources.* 21(3):203-215.

Perera, B.J.C., Hasofer, A.M., King, D.M. 2006. Sensitivity analysis of hydrological models: A historical survey [online]. In: *30th Hydrology and Water Resources Symposium: Past, Present and Future, Sandy Bay, Tasmania: Conference Design*, 241–246.

Reckhow, K.H. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecol. Modeling* 72(1-2): 1-20.

Rios, J.F., 2010. *A GIS-based model for estimating nitrate fate and transport from septic systems in surficial aquifers*. Master's thesis, Florida State University.

Rios, J.F., Ye, M., Wang, L., Lee, P.Z. 2011a. *Technical Manual of ArcNLET: An ArcGIS-Based Nitrate Load Estimation Toolkit*. http://people.sc.fsu.edu/~mye/ArcNLET/technical_manual.pdf.

Rios, J.F., Ye, M., Wang, L., Lee, P.Z. 2011b. *User's Manual of ArcNLET: An ArcGIS-Based Nitrate Load Estimation Toolkit*. http://people.sc.fsu.edu/~mye/ArcNLET/users_manual.pdf.

Rios, J.F., Ye, M., Wang, L., Lee, P.Z. 2011c. ArcNLET: Nitrate Load Estimation Toolkit, Esri Hydro Blog, available at http://blogs.esri.com/esri/arcgis/2012/03/06/arcnlet-nitrate-load-estimation-toolkit/.

Rios, J.F., Ye, M., Wang, L., Lee, P.Z., Davis, H., Hicks, R.W. 2012. ArcNLET: A GIS-based nitrate load estimation toolkit, Environmental Software and Modeling. Under Review.

Rong ,Y., Wang, R.F., Chou, R. 1998. Monte Carlo simulation for a groundwater mixing model in soil remediation of tetrachloroethylene. *Journal of Soil Contamination.* 7(1):87-102.

Saltelli, A. 2002. Making best use of model valuations to compute sensitivity indices. *Computer Physics Communications 145*: 280–297.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, John Wiley& Sons, Ltd.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S. 2010. Varance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications 181*: 259–270

Sobol', I.M. 1993. Sensitivity analysis for non-linear mathematical models.. *Mathematical Modelling and Computational Experiment 1:* 407–414; Translated from Russian: Sobol', I.M. 1990. Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie 2* : 112–118.

Swiler L.P., Wyss G.D. 2004. *A user's guide to Sandia's Latin Hypercube Sampling software: LHS UNIX library/standalone version.* Technical Report SAND2004-2439, Sandia National Laboratories, Albuquerque, NM.

Tesfa, T.K., Tarboton, D.G., Chandler, D.G., McNamara, J.P. 2009. Modeling soil depth from topographic and land cover attributes, *Water Resour. Res.*, 45, W10438, doi:10.1029/2008WR007474.

Ursin, E.L., 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. Environmental Protection Agency. 2002. *Onsite Wastewater Treatment System Manual, Office of Water*. Report No.: EPA/625/R-00/008.

U.S. Environmental Protection Agency. 1993. *Nitrogen Control.* Office of Research and Development, Cincinnati Ohio and Office of Water, Washington, DC.

Wang, L., Ye, M., Rios, J.F., Lee P.Z. 2011a. *Use of an ArcGIS based software for estimation of nitrate load from septic systems to St. Johns River in Jacksonville, FL.,* Proceeding of MODFLOW and More 2011Conference, June 5 - 8, 2011, Golden, CO.

Wang, L., Ye, M., Rios, J.F., Lee, P.Z. 2011b. *Application Manual of ArcNLET: An ArcGIS-Based Nitrate Load Estimation Toolkit*. http://people.sc.fsu.edu/~mye/ArcNLET/application _manual.pdf.

Wang, L., Ye, M., Rios, J.F., Fernandes, R., Lee, P.Z., Hicks, R.W. 2011c. Estimation of nitrate load from septic systems to surface water bodies using an ArcGIS-based software, *Environmental Earth Sciences*, Under Review.

Wang, X., He, X., Williams, J.R., Izaurralde, R.C., Atwood, J.D. 2005. Sensitivity and uncertainty analyses of crop yields and soil organic carbon simulated with EPIC. *American Society of Agricultural Engineers. 48(3)*: 1041−1054.

Zhao, J., Scheibe, T.D., Mahadevanl, R. 2011. Model-Based Analysis of the Role of Biological, Hydrological and Geochemical Factors Affecting Uranium Bioremediation. *Biotechnol. Bioeng. 108*: 1537–1548.