FLORIDA DEPARTMENT OF HEALTH

Florida Onsite Sewage Nitrogen Reduction Strategies Study

Task D

Quality Assurance Project Plan

Final Report

August 2010

4423-001

HAZEN AND SAWYER
Environmental Engineers & Scientists

In association with

SCHOOL OF MINES 1874 COLORADO

AET
Applied Environmental Technology

OTIS ENVIRONMENTAL CONSULTANTS, LLC

Florida Onsite Sewage Nitrogen Reduction Strategies Study

TASK D
FINAL REPORT

Quality Assurance Project Plan

Prepared for:

Florida Department of Health
Division of Environmental Health
Bureau of Onsite Sewage Programs
4042 Bald Cypress Way Bin #A-08
Tallahassee, FL 32399-1713

FDOH Contract CORCL

August 2010

Prepared by:

HAZEN AND SAWYER
Environmental Engineers & Scientists

In Association With:

SCHOOL OF MINES
1874
COLORADO

Table of Contents

O:\44237-001R011\Wpdocs\Report\Draft

List of Figures

List of Tables

Section 1.0
Introduction

1.1 Background

Nitrogen transport in the subsurface is a complex process, especially when considering the nitrogen inputs from onsite sewage treatment and disposal systems (OSTDS). Figure 1-1 summarizes the conceptual understanding of the inputs of nitrogen and the transformative, advective, and dispersive processes that lead to measurable nitrogen concentrations in the groundwater. The dominant transformation processes in groundwater include advection, dispersion (due to heterogeneities) and denitrification (conversion of nitrate to nitrogen gas).

Additional discussion regarding modeling the fate and transport of nitrogen and its movement and distribution in groundwater related to OSTDS was presented in the Task D Literature Review (submitted previously).

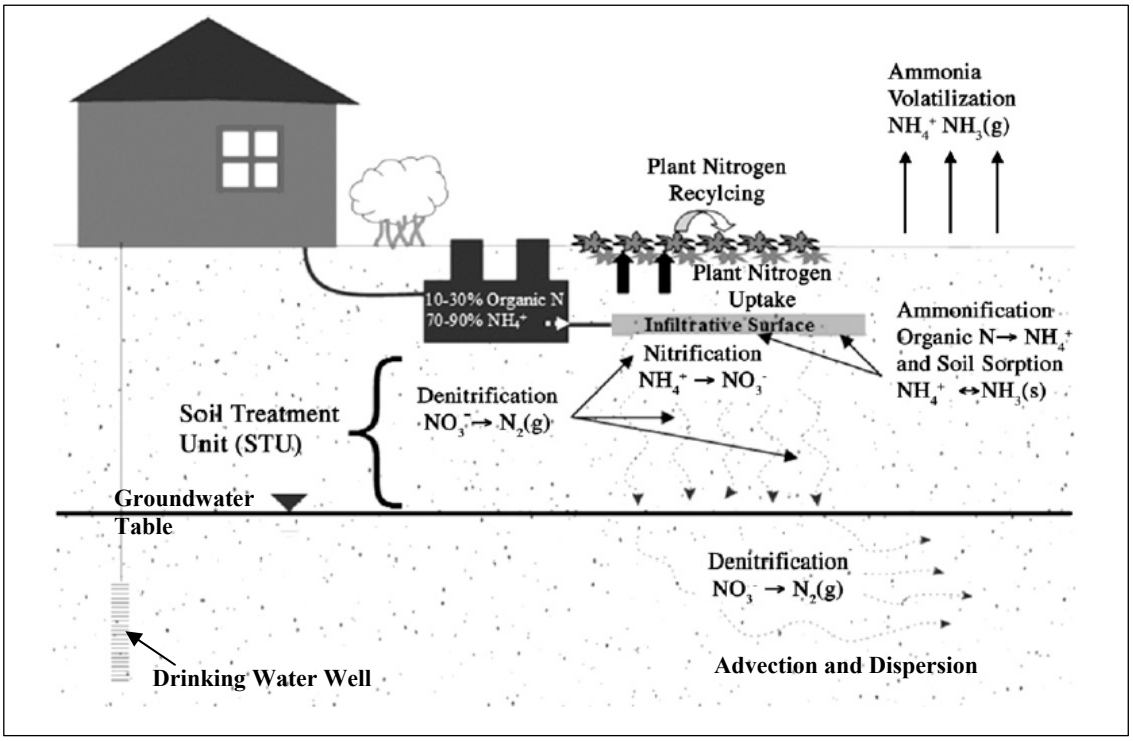

Figure 1-1: Nitrogen Processes Occurring in a Typical OSTDS (after Heatwole and McCray, 2007)

1.2 Project Scope and Purpose

For Task D of the Florida Onsite Sewage Nitrogen Reduction Strategies Study (FOSNRS), Colorado School of Mines (CSM) will develop a simple modeling tool to evaluate the fate and transport of nitrogen in groundwater related to the use of OSTDS. The model development will include the model conceptualization, design, and model-performance evaluation.

The goal of Task D is to develop a user-friendly modeling-tool that can be used to simulate nitrogen transport and transformation in groundwater, and to predict spatial and temporal nitrogen concentrations and fluxes, for a robust set of conditions relevant to OSTDS. Specifically, model output conditions that are important in Florida include the effects of seasonal loading from OSTDS, seasonal precipitation patterns, a spatial distribution of OSTDS, soil treatment, groundwater transformation and transport, plume concentrations, and mass flux at a downstream boundary. In addition, the load per OSTDS may be important for land planning purposes and/or for use in large-scale models to evaluate the specific inputs to groundwater or to back calculate the maximum load per system allowed without exceeding water quality levels.

The project organization is described in Section 1.3, and the technical approach is described in detail in Section 2.0.

1.3 Project Organization

The work described in this Quality Assurance Project Plan (QAPP) encompasses the entire scope of the project (see Section 2 for more detailed descriptions of activities). Task D is comprised of five interrelated activities as listed below (see also Figure 1-2). Items 1, 2, 3a, 3b, and 4a and the associated reporting will be completed in the first phase of this study. The remaining items will be completed in the subsequent phases of the project, building on the findings from field and modeling studies (Tasks C and D) using the observational method. Task D activities include:

1) Literature review,

2) Plan development,

3) Model development,

    a. Simple tool to calculate or estimate nitrogen removal in different soil types in Florida for input into the groundwater modeling tool,

    b. Analytical modeling tool to be used to predict temporal and spatial concentrations and fluxes of nitrate in groundwater,

    c. Integration of complex soil treatment module with the groundwater analytical modeling framework,

    d. Incorporate spatially variable OSTDS inputs (i.e., development scale model),

4) Performance evaluation of the model,

    a. Selection of existing site data (including from Task C) for model-performance evaluation,

    b. Calibrate models using existing site data (including from Task C),

    c. Validate models,

    d. Conduct uncertainty analysis of model input parameters,

5) Decision support framework,

   a. Guidance for determining model input parameters, and

   b. Risk-based approach for model selection.

The literature review has been previously submitted to the Florida Department of Health (FDOH) and the Research Review and Advisory Committee (RRAC) for review (Task D.1). This QAPP describes the approach for the proposed modeling-tool development, performance evaluation, and decision support development each building off of the existing knowledge of OSTDS performance and modeling techniques.

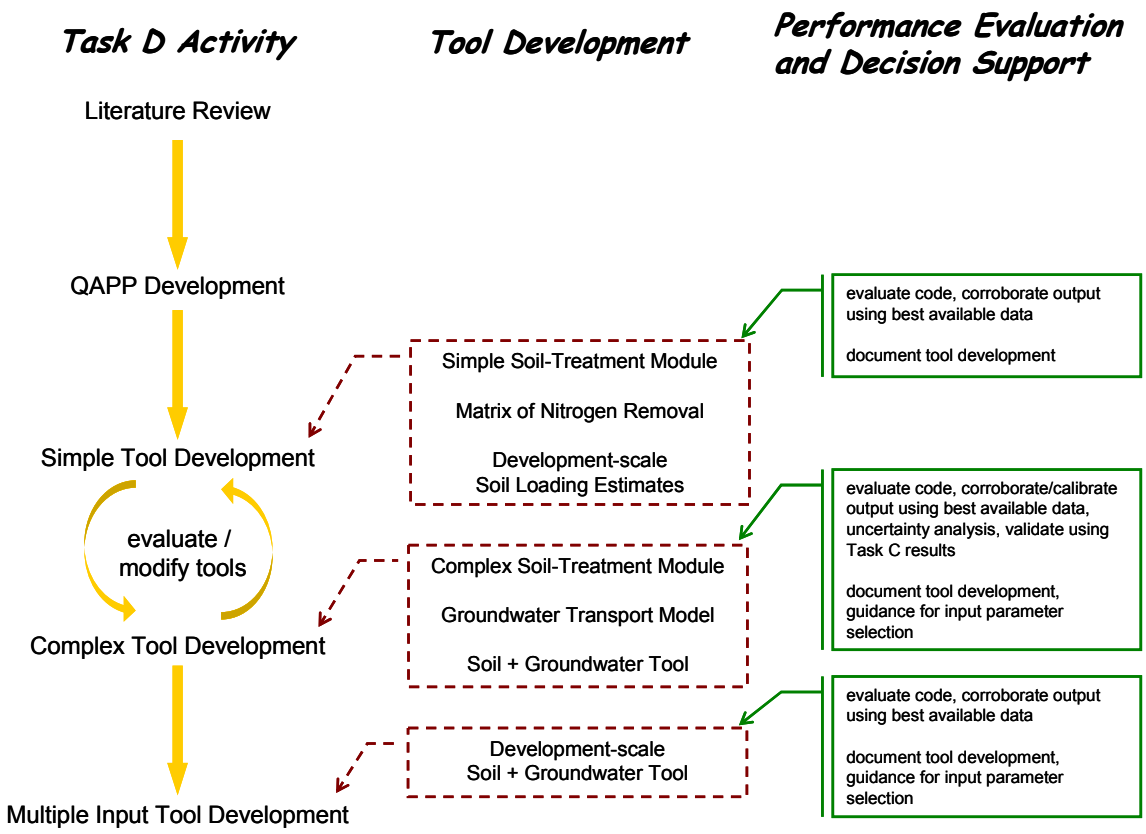

Figure 1-2: Task D Approach

Hazen and Sawyer will provide top-level management, task oversight, and direct reporting to the Florida Department of Health (FDOH). Dr. John McCray (Professor, Colorado School of Mines) is the principal technical manager for Task D. Ms. Kathryn Lowe (CSM) serves as the liaison to Hazen and Sawyer as well as coordinating Task C efforts with Task D needs. Dr. Mengistu Geza (CSM) is the lead technical expert providing model development and model-performance evaluation. Additionally, modeling projects are often dynamic by nature and frequently experience changes in conceptual and technical design as the development or model-performance evaluation progresses. Thus, the project team will be responsible for update reports that indicate needed changes in the project schedule, objectives, and path forward. Input to these changes will be consistent with the FOSNRS contract which includes timely FDOH reviews (30 days after submittal), project briefings at RRAC meetings, and solicitation of stakeholder input.

Section 2.0
Task D Description

The first phase of Task D will include development of a user-friendly analytical modeling tool that can be used to simulate nitrogen transport and transformation in groundwater, and to predict spatial and temporal nitrogen concentrations and fluxes, for a relevant set of conditions relevant to OSTDS (Section 1.3). The sophistication of the modeling tool is to be directed at the expertise level of an OSTDS technical practitioner (e.g., a soil-scientist, hydrologist, civil or environmental engineer, chemist, etc) who is not an expert in mathematical modeling. Performance evaluation of the model will be conducted in subsequent phases.

2.1 Description of Activities

The work scope described in this section is consistent with the scope of work and deliverables in the FOSNRS contract. The following description of activities provides detail related to development of the tools, performance evaluation of the developed module/models, and preparation of a complementary decision support framework. The general approach for tasks completed during future phases is described as additional detail depends on project outcomes and funding (precluding definitive detailed description at this time). To address this issue, the "observational approach" will be used which includes basing future tasks on information learned from previous tasks and obtaining appropriate consensus on changes to approach. FDOH, RRAC, and stakeholder input will be consistent with the FOSNRS contract.

The observational approach will include three general steps: 1) development of simple tools with generalized assumptions, 2) modification of simple tools to incorporate more complex mechanisms, and 3) up-scaling of the tools to evaluate multiple inputs (e.g., development- or subdivision-scale). The first step is to develop simple tools with generalized assumptions such as a simplified algorithm for soil treatment. Performance evaluation of these simple tools includes corroboration of the model output to existing data sets and qualitative checks for output reasonableness. As additional data is available from Task C, the second step will include modification of the simple models (from step 1) to incorporate more complex mechanisms based on robust field data sets. Performance evaluation of these more complex, but simple-to-use tools includes corroboration/calibration to rigorous data sets, sensitivity analysis, validation, and quantitative checks for output reasonableness. Finally, the third step will be to modify the existing

tools to incorporate multiple source inputs (i.e., multiple OSTDS) to evaluate subdivision or larger-scale impacts on groundwater.

2.1.1 Model Development

Two simple-to-use models will be developed and integrated: one for transport and treatment in the soil and vadose zone, and one for transport, dilution, and treatment in the aquifer. Development includes the conceptual model, coding and code evaluation. Existing models will first be reviewed to determine the appropriate starting point for model development. However, it is already apparent that no existing model other than complex numerical models (e.g., DRAINMOD, HYDRUS) can simulate all the desirable components of nitrogen fate and transport in Florida soils and aquifers as no existing simple modeling tools consider all of the following relevant OSTDS factors: time variable infiltration due to seasonal rain or prediction of increasing loads due to future development; variable source concentrations for dilution of OSTDS input by seasonal precipitation or a decrease in nitrogen mass-loading due to implementation of nitrogen reduction at the source; or chain degradation reactions such as from ammonium to nitrate to nitrogen-based gas. Even numerical models have not considered all these factors simultaneously. Thus, alterations to existing models, or development of new methods, will be employed to develop user-friendly yet robust tools to estimate nitrogen transport in soils and aquifers.

A vadose-zone modeling tool will be used to estimate transient concentrations of nitrogen reaching the water table (i.e., groundwater surface) from the infiltrative surface. Then a groundwater flow and transport model will be used to estimate spatially and temporally variable concentrations within a groundwater plume resulting from the vadose-zone source. The output from these two modules will be: 1) nitrogen concentrations at the water table, and 2) nitrogen concentrations in the aquifer downstream of the source. The models will be designed and linked together such that they can be implemented in a spreadsheet or using an alternate simple-to-use tool (e.g., a Java application).

2.1.1.1 Simple Tool ~ Spatially Averaged Nitrogen Removal in Florida Soils

The first modeling task is to develop a simple tool that can enable users to simulate groundwater plumes of nitrogen. Thus, we move directly toward this goal by first identifying a simple tool that can calculate or estimate the spatially averaged nitrogen removal in different soil types in Florida for input into the groundwater nitrogen-modeling tool. Figure 2-1 illustrates the difference between simplified spatially average assumptions and the more complex spatially variant conditions. Currently, most nitrogen groundwater modeling tools assume no treatment in the unsaturated soil. This approach is highly conservative and is usually not realistic or appropriate because significant treatment of nitrogen in most unsaturated soils has been rigorously documented (e.g., McCray et al.,

2010). To overcome this conservative approach, a module that provides realistic estimates of this treatment specific to Florida soils will be developed. The output from this simple soil treatment module will be based on a simplified approach such as described by Otis (2007) (e.g., under specific defined conditions x% removal can be achieved).

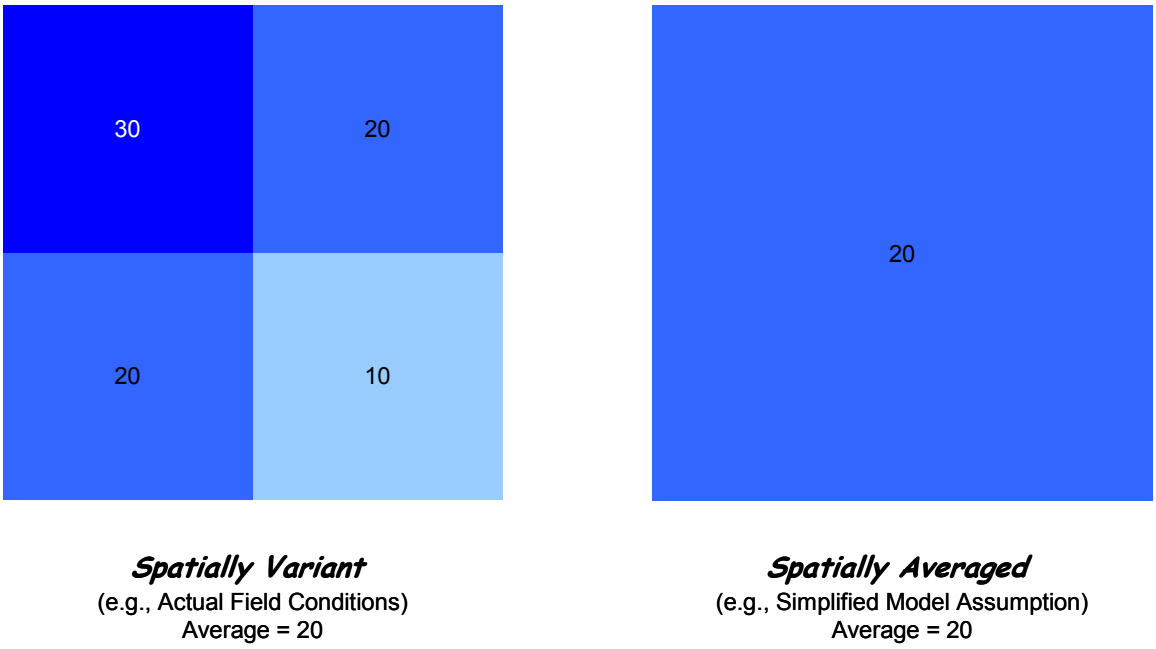

Figure 2-1: Illustration of Spatial Model Conditions

We intend to incorporate information from existing soil-treatment approaches that calculate or estimate treatment based on specific soil types that exist in Florida. We will evaluate the approach described by Otis (2007) specific to Florida soils for estimating soil reduction of nitrogen in the vadose zone to determine nitrogen loading to the aquifer. We will also evaluate the approach used by McCray et al. 2010 (i.e., STUMOD), which enables a prediction of nitrogen removal and vadose zone pore-water concentrations for each soil type among the 12 from the USDA soil triangle. Combination of these two approaches will enable an estimate of percent nitrogen reduction between the OSTDS and an aquifer water table. This concentration or loading will then be used as a source term for the groundwater nitrogen-modeling tool.

As data becomes available from Task C or from other Florida sites, a more robust soil-treatment module will be developed that can account for spatially variant OSTDS (not averaged input concentration), evapotransporation, and/or the effect of high/seasonal

variable water tables on nitrogen removal in the soil. However, initially a more rigorous soil-treatment module would delay completion of the groundwater modeling tool (which requires input from OSTDS). Thus, a simple soil-treatment module is used initially if groundwater modeling moves forward prior to complex soil model development. The output from the simple soil-treatment module will be a series of look-up tables providing estimated nitrogen removal based on common OSTDS operating conditions. Table 2-1 provides an example of the simple soil treatment output.

Table 2.1
Example of Look-up Table to be Developed for Simple Soil Treatment Evaluation

Soil Texture	Design HLR (gpd/ft^2)	Effluent Quality	Separation to Seasonal High Groundwater	Estimated Nitrogen Removal
Fine sand	0.8	60 mg-N/L as NH_4^+	12 inches	10%
Fine sand	0.8		24 inches	45%
Fine sand	0.8		36 inches	50%
Fine sand	0.8	30 mg-N/L as NO_3^-	12 inches	10%
Fine sand	0.8		24 inches	30%
Fine sand	0.8		36 inches	30%

Note: Table values are arbitrary and intended to illustrate the type of information and format of look-up values only rather than expected performance or actual modeled conditions.

The performance of this module will initially be evaluated using previous Florida OSTDS studies (such as USF Lysimeter Facility research) or other existing soils data provided by FDOH, Hazen and Sawyer, or Colorado School of Mines. Development and performance testing of the more robust soil-treatment module will be based on data collected from Task C when it becomes available. Additional description of model-performance evaluation is described in Section 2.1.2.

2.1.1.2 Analytical Modeling Tool ~ Temporal and Spatial Concentrations and Fluxes of Nitrate in Groundwater

The model must link soil and vadose zone fate described above and transport with aquifer fate and transport. Results of the literature review (Task D.1) suggested the use of the Horizontal Plane Source (HPS) model as the basis for the groundwater nitrogen-plume modeling tool. The model solution assumes a horizontal-plane contaminant source zone (Figure 2-2). Figure 2-2 shows a single OSTDS but the plane source could

also represent an averaged input for a development including fertilizer loads and groundwater recharge. The HPS model is based on the analytical solution developed by Carslaw and Jaeger (1959) for the transport of heat in solids. Galya (1987) adapted this analytical model for contaminant transport in groundwater. Heatwole and McCray (2006) used this model for OSTDS applications assuming spatially and temporally averaged inputs to the aquifer with no soil treatment.

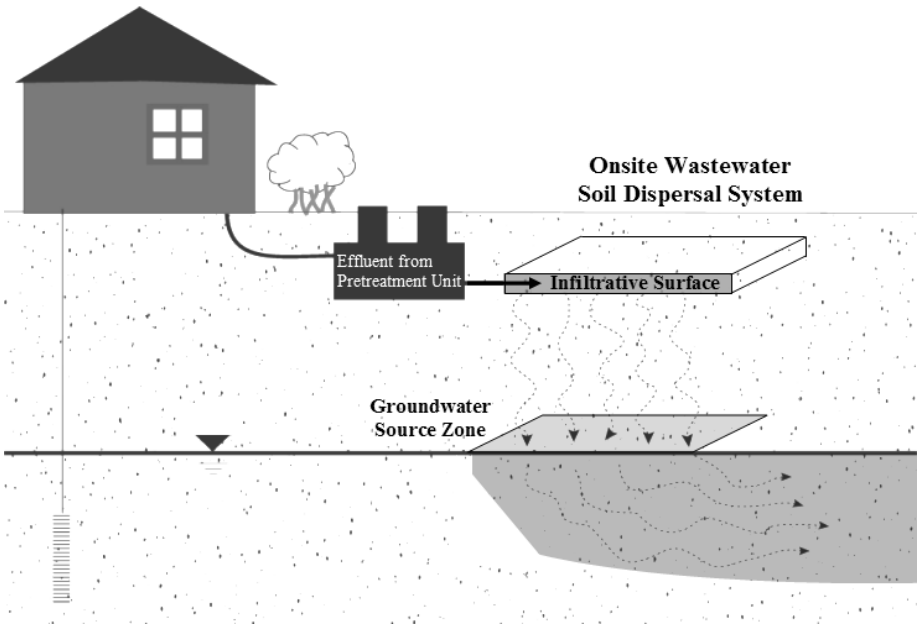

Figure 2-2: Conceptual Model for OSTDS Contamination of Groundwater Including the Contaminant Horizontal Plane Source (HPS) to the Aquifer

The HPS model is a transient, three-dimensional analytical model capable of simulating advective-dispersive transport and first order degradation (e.g., from denitrification) in a homogeneous aquifer with uniform horizontal flow. Assumptions of homogeneous media are required to develop the user-friendly model that is the goal of this research. Detailed heterogeneity can only be accounted for with the use of numerical models and a considerable amount of site data. However, successful models can be developed for heterogeneous media using our proposed modeling tool by accounting for aquifer heterogeneity through a macro-dispersion model-input parameter. To illustrate this concept, Figures 2-3 depicts a plume calculated by the HPS model, while Figures 2-4 and 2-5 illustrate measured plumes below OSTDS at Florida sites. Note that the HPS model captures the dominant features of the nitrogen plume: a relatively shallow plume below the water table, with considerable longitudinal spreading. The model could also produce a

"rounder" lateral plume similar to that shown in Figure 2-4 by decreasing the ratio of longitudinal to transverse dispersivity. The figures depict 2-D plumes, but the HPS model can calculate concentrations in three dimensions in an aquifer.

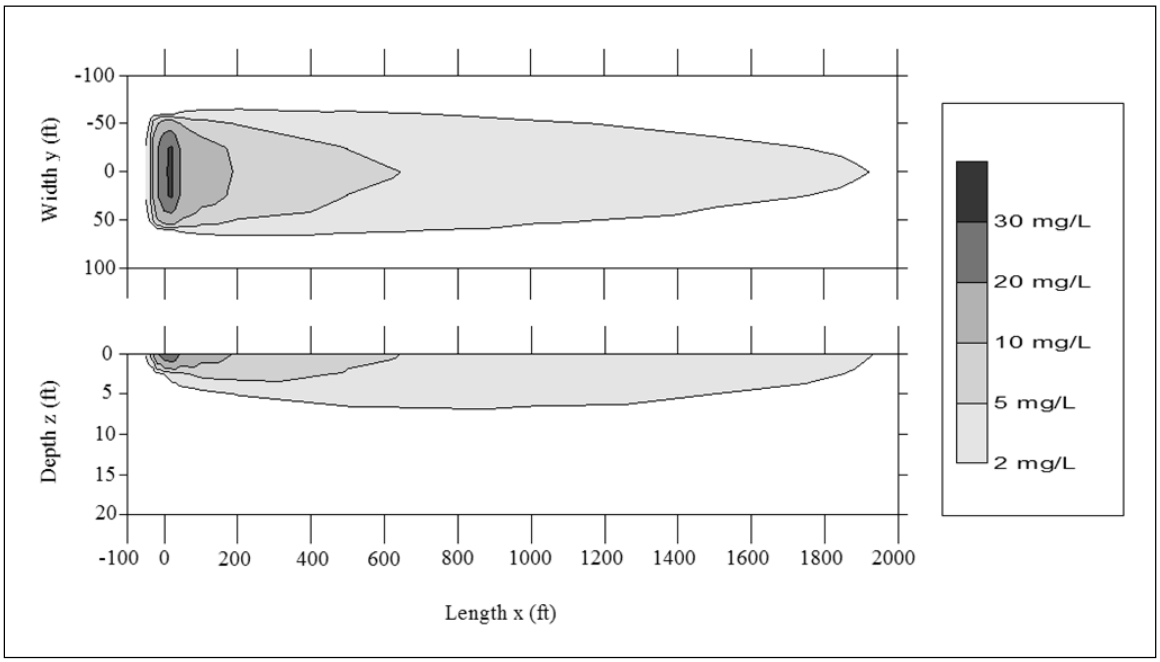

Figure 2-3: Plume Illustration Showing Aerial (Top) and Cross-sectional (Bottom) Views of a Nitrate Plume Generated by the HPS Model

Figure 2-4: Plume Illustration Showing a Plot of Measured Concentrations for Aerial Views of a Nitrate Plume in Lake County, Florida (Ellis & Associates, 2007)

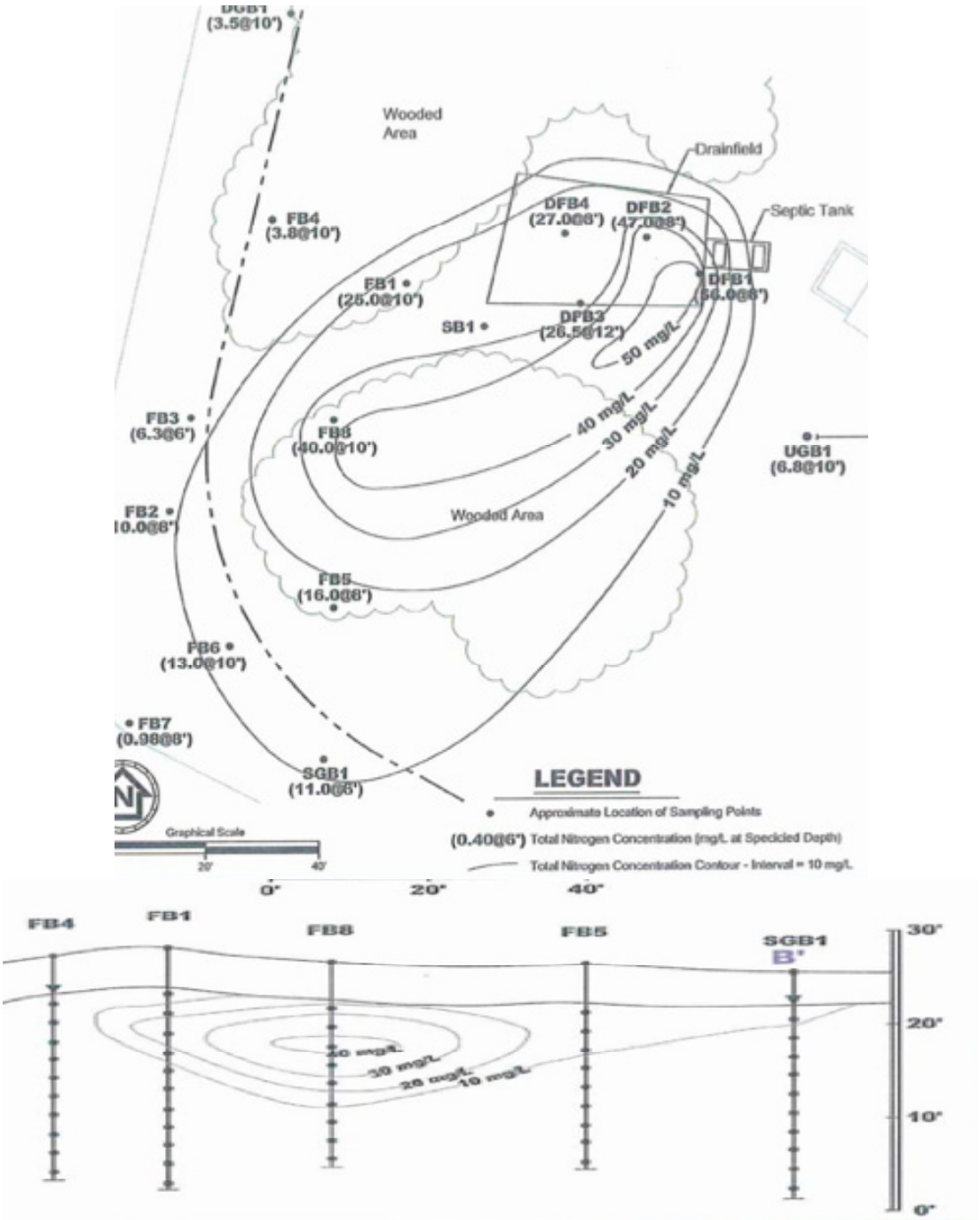

Figure 2-5: Plume Illustration Showing a Plot of Measured Concentrations for Aerial (Top) and Cross-sectional (Bottom) Views of a TN Plume in Seminole County, Florida (Ellis & Associates, 2007)

The HPS approach is capable of simulating aquifer impacts from single or multiple OSTDS or other sources through the law of superposition, but the modeling procedure becomes more complex and computationally demanding. Thus, the simplest approach for using this model is when the source zone is the same area as the footprint of OSTDS source. Then, the input to the vadose zone must consider the average loading per unit area.

The HPS model can consider a time-varying source rate and generates output for a transient, three-dimensional aquifer solute concentration in relation to the source zone. The time variable mass input is given by:

$$\dot{m} = i_r L W c_0 \qquad \text{(eqn. 1)}$$

The variable, *m*, is the temporally variable nitrogen mass loading rate, *L* and *W* are the dimensions of the source zone horizontal plane source "footprint" (length and width), *i* is the time variable infiltration rate and c_o is the time variable concentration in the total infiltrating water reaching the water table. The infiltration can include that from climate inputs as well as from OSTDS, and c_o may include the effects of rainfall dilution. The mass of the nitrogen input does not change with dilution, but the time-dependent temporal strength of the horizontal plane source would vary, as would the aquifer dilution relative to groundwater flow.

The analytical solution calculated groundwater concentrations in space and time, $C_p(x,y,z,t)$, given by (Galya, 1987):

$$c_p(x,y,z,t) = \int_0^t \dot{m} c_s(x,y,z,t) \qquad \text{(eqn. 2)}$$

and

$$c_s(x,y,z,t) = \frac{e^{-kt}}{4nRLWb}\left(erf\left(\left(\frac{x - v't + L/2}{\sqrt{4D_{xx}'t}}\right)\right) - erf\left(\frac{x - v't - L/2}{\sqrt{4D_{xx}'t}}\right)\right)$$
$$\times\left(erf\left(\frac{y + W/2}{\sqrt{4D_{yy}'t}}\right) - erf\left(\frac{y - W/2}{\sqrt{4D_{yy}'t}}\right)\right) \times \left(1 + 2\sum_{j=1}^{\infty} \exp\left(-\frac{j^2\pi^2 D_{zz}'t}{b^2}\right)\cos\left(\frac{j\pi z}{b}\right)\right)$$

(eqn. 3)

Where $c_s(x, y, z, t)$ are calculated groundwater concentrations, v' is the contaminant velocity [L/T], D' are the dispersion coefficients in the x, y, and z directions including any effects of contaminant retardation [L^2/T], R is the retardation factor, b is the aquifer thickness [L], n is porosity, and k is a first-order degradation rate constant (e.g., for denitrification) [T^{-1}]. For equation 3, j is a numerical counter for the number of depth interval (from 0 to infinity).

The solution to the above equations requires integration, but can be solved through numerical approximation using Simpson's rule. The analytical solution can thus be implemented in a Java application and compiled to create an executable file capable of calculating multiple space-time inputs, and can likely be achieved with a spreadsheet program.

A sensitivity analysis was conducted to evaluate the importance of input parameters for the HPS model (Heatwole and McCray, 2006). These are input parameters that have the most impact on model output, and must be carefully considered when developing model input as well as parameterizing the model. The most sensitive parameters were vertical dispersivity, velocity, porosity, the ratio of source zone length to width, solute infiltration rate, OSTDS nitrate concentration, and denitrification rates. For example, aquifers with faster groundwater velocities will result in higher dispersion and be much better equipped to attenuate nitrate inputs from typical OSTDS through dilution. Alternatively, slower groundwater velocities will enhance denitrification through longer reaction times for a given distance downgradient.

The performance of the groundwater model will be evaluated using data collected from Task C when it becomes available as well as any existing available data sufficient to validate the model. Additional description of model-performance evaluation is described in Section 2.1.2. In addition, guidance will be provided for parameter selection in the decision support framework (Section 2.1.3). Specifically, during model development and performance evaluation, easily measured field parameters will be identified that can be used as simple surrogates for model parameterization by practitioners.

2.1.1.3 Additional Tool Development

Depending on the level of funding available, adaptation of, or development and calibration of, a model that describes multiple inputs for transport of nitrogen to either deeper aquifer zones or to surface water will be conducted. Numerous simple tools already exist to predict nitrogen infiltration to soils versus runoff to streams from multiple sources using GIS-based watershed scale models (Lasserre et al., 1999; Stark et al., 1999; Kellog et al., 1997). Development of this model will enable simulation of nitrogen concentrations and mass flux in space and time from several OSTDS in a development-scale

area. This aquifer model must be calibrated using existing data from a development-scale plume, based on metrics such as average concentration in the plume and/or mass flux crossing a boundary. Finally, a multi-development scale model may be developed incorporating spatially averaged OSTDS inputs over a sub-basin area.

2.1.2 Model-Performance Evaluation

The performance of the model will be evaluated based on whether or not it adequately simulates site data, and whether model calibration is required to produce an acceptable modeled representation of site data. To accomplish this task, the module/model will be calibrated to the best available data (e.g., Task C GCREC data) and validated using existing data (e.g., Task C home sites). The first test, evaluation of whether or not the model adequately simulates site data, is qualitative based on whether or not the model can simulate the most important attributes of the plume (e.g., plume shape, length, maximum concentrations, etc). In addition, uncertainty analysis allows the user to evaluate if the probably outcome is compatible with specific treatment objectives. The second test, whether model calibration is required to produce an acceptable modeled representation of site data, is quantitative and includes calibration and validation. The performance of each module developed will be evaluated based on the complexity of the model. For example, the simple soil module will be corroborated with existing site data, while the complex soil module will be calibrated with data from Task C. Data from Task C homesites as well as other available data sources will be used to validate the models.

2.1.2.1 Selection of Site Data for Model-Performance Evaluation

Actual data from OSTDS sites will be used to evaluate the performance of the groundwater nitrogen-modeling tool (Section 2.1.1.2). At a minimum the site data should include spatial nitrogen concentrations in groundwater, fundamental hydrogeology parameters (especially the ones the model is most sensitive to, as described in the previous section), and information on the OSTDS loading. The primary source of this data will be the Task C controlled pilot-scale testing conducted at the Gulf Coast Research and Education Center (GCREC) to characterize nitrogen fate and transport under a variety of typical operating conditions (described in the Task C QAPP). Other sites were also identified as described in the Task D.3 report submitted previously, and will be evaluated for use in model-performance evaluation:

- Primary Candidate Studies for Task D Model-Performance Evaluation
  1) Multiple Nitrogen Loading Assessments from Onsite Waste Treatment and Disposal Systems Within the Wekiva River Basin, by Ellis & Associates, 2007 and summarized in Wekiva Nitrogen Source Study, by Briggs et al., (2007) and Roeder (2008).

2) St. George Island Study, described by Corbett and Iverson (1999) and Corbett et al., (2002).
3) Contaminant Transport Investigation from an Onsite Wastewater Treatment System (OWTS) in Fine Sand: Phase 3 Report, by Ayres Associates, 1996 and summarized by Nielsen et. al., 2002, and McAvoy et. al., 2002.,
4) Septic Tank Treatment System (SepTTS) Model for Cleaning Product Chemicals, User's Manual, Version 1.0, by The Soap and Detergent Association, 1997 and applied to St. Johns County, Florida site in McAvoy et. al., 2002 (using data from Ayres Associates, 1996 above)
5) An Investigation of Surface Water Contamination Potential from Onsite Sewage Disposal Systems in the Turkey Creek Sub-basin of the Indian River Lagoon, by Ayres Associates, 1993 and summarized by McNeillie et. al.1994, and Anderson, 1998.

- Alternate Candidate Study for Task D Model-Performance Evaluation
  1) Lake Okeechobee, described by ESE (1993)

2.1.2.2 Model Corroboration/Calibration

Corroboration/calibration will be used to better understand the quality and quantity of data required to enable a rigorous calibration using data from Task C. Model performance is quantitatively evaluated by comparing simulated parameter values to the corresponding measured values. These parameters are called calibration targets. Calibration targets for this work will include nitrogen concentrations (weighted equally in space), the mass of contaminant in the plume, and plume dimensions. Because concentration calibration targets are spatially and temporally variable, the goodness of calibration (model performance), is assessed by some representative measure (average) of performance that describes the "match" of measured data to simulated data for all locations in the hydrogeologic domain. The corroboration/calibration procedure will be an iterative process and may suggest revisions in the data collection plan (Task C) or in the model itself.

2.1.2.3 Model Validation

Validation will be used to compare the corroborated/calibrated model to actual field data. Model validation ensures that the module/model meets the intended requirements and identifies the range of appropriate conditions (e.g, capabilities and limitations). Data from Task C homesites as well as other available data sources will be used to validate the modules/models.

2.1.2.4 Model Uncertainty Analysis

Uncertainty testing provides a methodology to use models for decision making even if sufficient data does not exist to calibrate the model. Probability-based ranges for model input parameters are used to generate probable model outcomes.

Uncertainty testing assesses the range of model outcomes for nitrogen concentration and the theoretical probability that any particular outcome will occur. Figure 2-6 is a nomograph illustrating uncertainty in the model outcome based on the range of model inputs. In this example, the 50 percentile value (median case or most likely model prediction) of all model predictions is read from the y axis as approximately 7.5 mg/L, which is below the MCL for drinking water, but above background. Alternatively, the 70th percentile concentration is 10 mg/L (the MCL). One way to interpret this result is to conclude that there is a 30% probability that the MCL will be reached or exceeded, based on the model simulations.

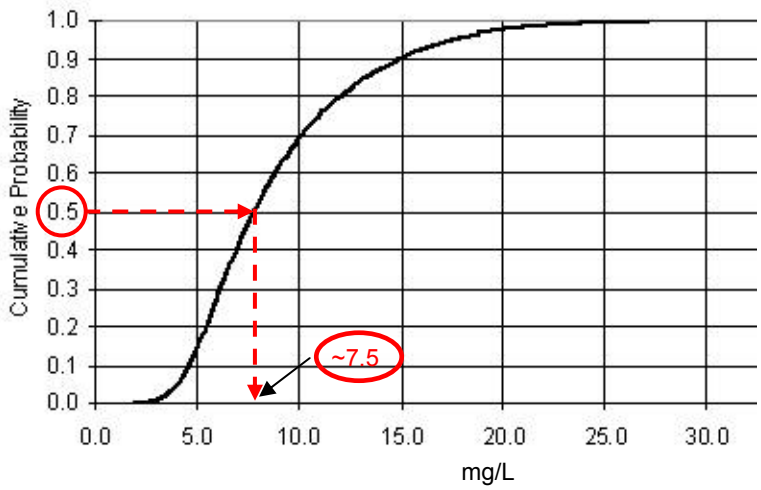

Figure 2-6: Example Nomograph Illustrating Model Output Uncertainty

This information can lead to informed, risk-based decisions. While this approach appears to make the decision-making process more complicated, it has considerable advantages. For example, this type of information can result in a stakeholder agreement to proceed with a decision that assumes an MCL will not be exceeded, provided that water-quality monitoring is conducted, and that contingency plans are made ready for rapid implementation. As the target goal (e.g., 10 mg/L MCL) reaches the upper percentile values (e.g., 95th percentile), the risk of violating the water quality standard is reduced and a majority of stakeholders may be more willing to accept the uncertainty in OSTDS performance.

To the extent possible (without precluding performance evaluation of the aquifer model in Phase 1), model uncertainty and sensitivity analyses will be conducted. Uncertainty testing results will be provided in the form of nomographs for selected conditions. Because an infinite number of nomographs cannot be prepared, the simple tool will also be developed to incorporate the risk-based software (which must be purchased by the user), enabling the user to conduct uncertainty testing for specific conditions of interest.

2.1.3 Decision Support Framework

The decision support framework will be in the form of a guidance manual describing model development, input parameter selection, and uncertainty assessment. Probability-based ranges for model input parameters will be used to generate probable model outcomes, providing planners with the option of using the most-probable model outcome in the decision making process, or the model outcome that would lead to a more conservative or liberal decision as the specific case warrants.

2.2 Task D Performance Assessment

The performance assessment of Task D will be evaluated by the acquisition of sufficient data to calibrate and validate the simple models developed in Task D. If a model is to be used, then some level of understanding of how model fit is evaluated is a necessary criterion for model quality assurance. While the general user will most likely assess performance by comparing model output to field observations, a more rigorous approach is required for technical users. For this case, the model-performance assessment will be conducted by using model-evaluation statistics (i.e., acceptance criteria) to determine whether the model can appropriately simulate the observed data. Measures of model performance are classified into different measures and correlation measures. Root Mean Square Error (RMSE) is perhaps the most common measure. Other difference measures include the mean bias error (MBE), the index of agreement (d), and the Nash-Sutcliffe efficiency (NSE). Correlation measures and graphical techniques are also useful for evaluating model performance.

Moriasi et al. (2007) reviewed several model-evaluation techniques, including statistical measures and graphical techniques. They reported ranges of values and corresponding performance ratings for each recommended statistic and gave recommendations for acceptable criteria for each statistic. Based on this analysis, they recommend use of three quantitative statistics and a graphical technique. The statistical measures were the NSE, normalized mean bias (NMB), and RMSE-observations standard deviation ratio (RSR). Thus, we will use multiple methods for evaluating the model performance. By using multiple methods, the model quality assurance evaluation is not unduly hindered

by the specific limitations of a single calibration statistic (each test has strengths and disadvantages). However, by using each of these methods, if poor performance is suggested by one of the statistical outcomes, further evaluation of the model may be warranted. The following describes these statistical performance assessment tests.

Root Mean Square Error-observations standard deviation ratio (ROSR, also called RSR). Singh et al. (2004) developed the RSR which can account for the bias due to variability in the data set. RSR standardizes RMSE using the observations' standard deviation. RSR is calculated as the ratio of the RMSE and SD of measured data:

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_{i}^{n}(o_i - p_i)^2}}{\sqrt{\sum_{i}^{n}(o_i - o_{mean})^2}} \qquad \text{(eqn. 4)}$$

where p_i is simulated parameter value, o_i is observed (or measured) value, n is the number of observations, and o_{mean}, is the mean of the observed values. The smaller the RSR value for a given hydrogeologic model, the better the calibration. A RMSE value closer to zero indicates a better fit to observed values. The denominator in the RSR serves to minimize the influence of a few observations that have very large or small values relative to the observations as a whole. RSR varies from the optimal value of 0, which indicates zero RMSE or residual variation and therefore perfect model simulation, to a large positive value. The lower RSR, the lower the RMSE, and the better the model simulation performance. Moriasi et. al. (2007) conducted a detailed study of model-calibration-evaluation measures and recommended the following criteria for RSR: 0 to 0.5 is considered to be very good, 0.50 to 0.60 is good, 0.60 to 0.70 is satisfactory and greater than 0.70 is unsatisfactory.

The index of agreement (d): The index of agreement (d) developed by Willmott (1981) is another measure of a standardized measure of the degree of model prediction error. It is calculated as:

$$d = 1.0 - \frac{\sum_{i=1}^{N}|o_i - p_i|}{\sum_{i=1}^{N}\left|p_i - \overline{o_i}\right| + \left|o_i - \overline{o_i}\right|} \qquad \text{(eqn. 5)}$$

where o_i = measured value, p_i = simulated value and $\overline{o_i}$ is mean of measured values. The d value varies between 0 and 1. A value of 1 indicates a perfect agreement between the measured and predicted values, and 0 indicates no agreement at all (Willmott,

1981). This method is similar in concept to the RSR. Additional literature review is required to develop standardized acceptance criteria for this statistic.

Nash-Sutcliffe Efficiency (NSE): The Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) determines the model efficiency as a fraction of the measured value variance that is reproduced by the model. The underlying concepts justifying this statistic are similar to the RSR. NSE is given is calculated as:

$$NSE = 1 - \frac{\sum (o_i - p_i)^2}{\sum (o_i - \overline{o_i})^2} \qquad \text{(eqn. 6)}$$

where: $\overline{o_i}$ is mean of measured values. The closer the NSE value to 1.0 the better is the model estimation. NSE $\geq$ 0.75 is considered to be an excellent estimate, and NSE between 0.75 and 0.36, is regarded to be satisfactory (Motovilov et al., 1999).

Normalized mean bias (NMB): Normalized mean bias (NMB) measures the average tendency of the simulated data to be larger or smaller than their observed values (Gupta et al., 1999). This statistic normalizes the difference (model - observed) over the sum of observed values. NMB is defined as:

$$NMB = \frac{\sum_{1}^{n} (p_i - o_i)}{\sum_{1}^{n} (o_i)} * 100 \qquad \text{(eqn. 7)}$$

Positive values indicate that simulated values tend to be greater than observed values, while negative values indicate that simulated values tend to be smaller than observed values. A value of zero indicates no bias. Additional literature review is required to develop standardized acceptance criteria for this statistic.

Correlation measures: The relationship between measured and observed data such as covariation and correlation can be useful to evaluate model performance and “calibration goodness”. The correlation coefficient, R, or the coefficient of determination (R^2) is typically used. R^2 describes the degree of co-linearity between simulated and measured data. R^2 ranges from 0 to 1, with higher values indicating less error variance, and typically values greater than 0.5 are considered acceptable (Santhi, et al., 2001, Van Liew et al., 2003). Indeed, in hydrogeologic modeling, an R^2 value greater than 0.7 is considered excellent. However, R^2 is oversensitive to high extreme values (outliers) and insensitive to additive and proportional differences between model predictions and measured data (Legates and McCabe, 1999). That means it is possible to obtain a good R^2 value as long as simulation results capture the trend in observed values even when the

absolute differences are large. For example, it is possible to achieve a very high R^2 value even if a simulated hydraulic head-vs-time data series is visually offset (does not overlap) with the observed head-vs-time series, provided that the shape (or trend) of the two series was the same. Thus, while this value is a good measure of model goodness, it cannot be used alone.

Graphical techniques: Graphical techniques such as comparative time series plots, bar graphs comparing measured to simulated values, box plots showing the overall difference (including high and low ranges for the errors), can also be helpful. In particular, plots that can illustrate the spatial or temporal variation in measured-vs-simulated differences are helpful to understand how the model is performing in different geographic locations or through time. An example of a plot that can aid in spatial analysis is a "dot plot" where the size of the dot represents the difference in simulated vs. observed values. The dot plot can help place the model results in geographic context, or suggest areas where additional data collection or more careful data scrutiny is warranted. Graphical methods enable the modeler to insert professional judgment and "common sense" into the task of model-performance evaluations.

Model Uniqueness: During model calibration, it is often impossible to converge on a unique solution when estimating many parameters (Geza et al. 2009). That is, a similar model calibration can be achieved for different input-parameter values. This fundamentally suggests a non-physical model, or that the model is not likely to be effective for simulating conditions outside the calibration conditions. Consequently, one needs to pose a tractable calibration problem by limiting the number of parameters for which values will be estimated (i.e., simplifying the model to one that represents important aspects of the system (Hill, 1998)). This is accomplished by identifying the most sensitive, uncorrelated parameters. (Poeter et al., 2005; Saltelli et al., 2004), and evaluating whether input parameters are correlated. Accepted USGS guidance (Poeter et al. 2005) will be used for this purpose.

2.3 Contingency Measures

The observational method for technical decision making will be employed during Task D model development. This method is a continuous, integrated, process of design, monitoring, and review that enables modifications to be incorporated into the model design/development as appropriate. The observational method provides for initial model development based on the simple general conditions rather than complex poorly understood mechanisms. The gaps in the available information required for more complex model development are then filled by observations (e.g., Task C field monitoring) which aid in performance evaluation (e.g., calibration and validation) of the model. For Task D, three general steps will be included in the observational approach: 1) development of

simple tools with generalized assumptions, 2) modification of simple tools to incorporate more complex mechanisms, and 3) up-scaling of the tools to evaluate multiple inputs (e.g., subdivision). Each step will include model-performance evaluation and documentation of the model development. This approach enables inclusion of “learn as you go” understanding without stopping task progress unless data quality objectives (DQOs, Section 3.1) are not satisfied.

Section 3.0
Quality Assurance

3.1 Data Quality Objectives (DQOs)

The general quality assurance (QA) objectives for Task D are provided below:

1) Document the model theory.

2) Document the model development process.

3) Document model revisions.

4) Back-up the model software and associated electronic files.

5) Evaluate theory and mathematics used in the model to ensure they are accurately implemented.

6) Evaluate model performance using measurable acceptance criteria.

7) Provide guidance on how to use the model.

8) Identify and track QA documentation.

The process used to meet each DQO is described in more detail below.

3.2. Process to Meet Data Quality Objectives (DQOs)

3.2.1 Document the Model Theory

The model theory will be documented in detail as part of a written User's Manual (i.e., Decision Support Framework). A technical expert who was not involved with the model theory selection or development will review the document for appropriateness and correctness, and sign a written statement indicating that the document was reviewed and providing the date of the review.

3.2.2 Document the Model Development Process and Theory

The model development process (software used, methods used to develop the software, mathematics used to implement functions) will be documented in an electronic document and updated monthly. A signature sheet will be implemented that documents the developers acknowledgement that the electronic document has been updated.

3.2.3 Document the Model Revisions

Significant revisions to the model software will be documented in an electronic document and updated monthly. A signature sheet will be implemented that documents the developer's acknowledgement that these revisions have been documented.

3.2.4 Back-Up the Model Software and Associated Electronic Files

The most recent version of the software will be saved and backed up daily on electronic storage media located in a separate physical location at CSM from the computer used to implement the changes. An electronic version of the software will be saved at the end of each month and all these monthly versions will be kept until the end of the project. An electronic document that is updated monthly will document the name of the file and the significant changes to the document.

The file name will include a model identifier, developer initials, and date. For example, if the software is implemented in an XLS file:

Nmodel-MG-31Mar10.XLS

where "Nmodel" is the name of the model, MG is the developer initials (e.g., Mengistu Geza) and the date is March 31, 2010.

If software is used that requires separate input or output files, then designators of "in" and "out" will be used in the filename. An example of a Fortran input file is given below:

Nmodel-MG-31Mar10.in

3.2.5 Evaluate Theory and Mathematics in the Model to Ensure They are Accurately Implemented

Correct implementation of the theory and mathematics will be verified using two methods.

1) Hand calculations to evaluate selected model calculations. The code evaluation will be conducted by a technical expert who has not been directly involved with the model development.

2) Benchmarking the model against a tested software package where the initial, boundary, and run-time conditions are manipulated to be the same for both models.

These methods and outcomes will be documented in a short report to the FDOH.

3.2.6 Evaluate the Model Performance Using Measurable Acceptance Criteria

The model performance will be evaluated as described in Section 2.2. These evaluation methods and outcomes will be documented in a short report. Numerical statistical measures will be used to assess how well the model simulates measured field data. The statistics calculated for measured versus observed data will be compared to accepted values published in the peer-reviewed literature. While the "goodness" of a model is necessarily subjective, the use of numerical acceptance criteria provides a transparent means of documenting the model performance. Finally, the model's performance in simulating measured data will also be tested for non-uniqueness and input parameter correlation (recall Section 2.2), in accordance with the guidance provided in the USGS document written by Poeter et al. (2005).

3.2.7 Provide Guidance on How to Use the Model

A written guidance document (i.e., Decision Support Framework) will be provided that describes how to use the model. A technical expert not directly associated with model development will review the document. The reviewer will sign a written statement indicating that the document was reviewed and providing the date of the review.

3.2.8 Identify and Track QA Documentation

A written list and short description of all the documents associated with DQO's defined in Sections 3.2.1 through 3.2.8 will be maintained in a file titled "QA-Documentation.doc" and will be provided to the FDOH within the relevant report. This will enable the QA system to be fully auditable.

Section 4.0
Health and Safety

Work associated with Task D is conducted in an office setting. Thus, only routine health and safety measures required (ground fault circuit interrupts, clutter around electrical connections not permitted, etc.).

Section 5.0
References

Anderson, D. L. (1998). Natural Denitrification in Groundwater Impacted by Onsite Wastewater Treatment Systems. In: On-site Wastewater Treatment. Ed. By D.M. Sievers. ASAE Publication 03-98, Proc. of the 8th National Symp. on Ind.and Small Comm. Sewage Syst., Orlando, FL. Amer. Soc. Of Agr. Engrs., St. Joseph, MI.

Ayres Associates (1993). An Investigation of the Surface Water Contamination Potential From Onsite Sewage Disposal Systems in the Turkey Creek Sub-basin of the Indian River Lagoon. St. Johns River Water Mgmt. District SWIM Project IR-1-110.1D Report under HRS Contracts No. LP114 and LP506. 2 volumes: Final Report and Appendices

Ayres Associates (1996). Contaminant Transport Investigation from an Onsite Wastewater Treatment System (OWTS) in Fine Sand: Phase 3 Report. Report to The Soap and Detergent Association, New York, NY.

Briggs, G.R., E. Roeder, and E. Ursin (2007). *Nitrogen impact of onsite sewage treatment and disposal systems in the Wekiva study area*. Prepared for Florida Department of Health, Tallahassee, FL, by Bureau of Onsite Sewage Programs, Division of Environmental Health.

Carslaw, H.W. and J.C. Jaeger (1959). Conduction of Heat in Solids. Oxford University Press, London, England.

Center for Climatic Research (2005). Department of Geography, University of Delaware. Vol. 30: 79–82.

Corbett, D.R., K. Dillon, W. Burnett, and G. Schaefer (2002). The spatial variability of nitrogen and phosphorus concentration in a sand aquifer influenced by onsite sewage treatment and disposal systems: A case study on St. George Island, Florida. *Environmental Pollution* 117 (2):337-345.

Corbett, D.R. and R. Iverson (1999). *Groundwater and nutrient dynamics on a strip barrier island served by on-site sewage treatment and disposal systems in the northern Gulf of Mexico*. Contract No C0005. Prepared for Florida Department of Health, Tallahassee, FL, by Florida State University, Department of Oceanography.

Ellis & Associates (2007). Multiple Nitrogen Loading Assessments from Onsite Waste Treatment and Disposal Systems Within the Wekiva River Basin. Report submitted to the Florida Dept. of Health, May, 2007

ESE (1993). *Onsite sewage disposal systemresearch on the northern periphery of Lake Okeechobee*. Contract No. LP555. Prepared for State of Florida, Department of Health

and Rehabilitative Services, Tallahassee, FL, by Environmental Science & Engineering, Inc, Gainsville, FL.

Galya, G.P. (1987). A Horizontal Plane Source Model for Ground-Water Transport. Ground Water. Vol. 25, No. 6, p. 733-739.

Geza, M., J.E. McCray, and E.P. Poeter (2009). Quantifying predictive uncertainty for a mountain-watershed model, J. Hydrology, 376, 170–181.

Gupta, H.V., S. Sorooshian, and P. Yapo (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibrati*on. J. Hydrologic Eng.*, 4(2): 135-143.

Heatwole, K. and J.E. McCray (2006). A simple model for predicting nitrate plumes, Proceedings of the NOWRA Annual Mtg, Denver CO, National Onsite Wastewater Recycling Association (NOWRA), Laurel, MD.

Hill, M.C. (1998). Methods and guidelines for effective model calibration. U.S. Geological Survey Water-Resources Investigation Report 98-4005. http://pubs.water.usgs.gov/wri984005/.

Kellogg, D.Q., L. Joubert, and A.J. Gold (1997). *MANAGE: A Method for Assessment, Nutrient-Loading, and Geographic Evaluation of Nonpoint Pollution*. University of Rhode Island Cooperative Extension, Department of Natural Resources Science, Kingston, RI.

Lasserre, F., M. Razack, and O. Banton (1999). "A GIS-linked Model for the Assessment of Nitrate Contamination in Groundwater." *Journal of Hydrology*, 224, 81-90.

Legates, D.R. and G.J McCabe (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*. 35(1): 233-241.

McAvoy, D. C.; A. J. DeCarvalho; A.M. Nielsen; and M. L. Cano (2002). Investigation of an Onsite Wastewater Treatment System in Sandy Soil: Modeling the Fate of Surfactants. Environmental Toxicology and Chemistry, Vol. 21, No. 12, pp 2623-2630.

McCray, J.E., M. Geza, K.S. Lowe, T.B. Boving, D. Radcliffe, M.B. Tucholke, A. Wunsch, S. Roberts, J. Amador, J. Atoyan, J. Drewes, D. Kalen, and G. Loomis (2010). Development of quantitative tools to determine the expected performance of unit processes in wastewater soil treatment units. Final Report for Water Environment Research Foundation Project No. DEC1R06, Colorado School of Mines, Golden, CO.

McNeillie, J. I.; D. L. Anderson; and T.V. Belanger (1994). Investigation of the Surface Water Contamination Potential from On-Site Wastewater Treatment Systems (OWTS) in the Indian River Lagoon Basin. In: On-Site Wastewater Treatment. Ed. By E. Collins. ASAE Publication 18-94. Proc. of the 7th Intern. Symp. on Ind. and Small Comm. Sewage Syst., Amer. Soc. of Agr. Engrs., St. Joseph, MI

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

Motovilov, Y.G., L. Gottschalk, K. Engeland, and A. Rodhe (1999). Validation of a distributed hydrological model against spatial observations. *Agr. Forest Meteorol.*, 98/99: 257-277.

Nash, J. E. and J. V. Sutcliffe (1970). River flow forecasting through conceptual models, part 1 - A discussion of principles, J. of Hydrology 10(3): 282-290.

Nielsen, A. M.; A.J. DeCarvalho; D.C. McAvoy; L Kravetz; M. L. Cano; and D.L. Anderson. (2002). Investigation of an Onsite Wastewater Treatment System in Sandy Soil: Site Characterization and Fate of Anionic and Nonionic Surfactants. Environmental Toxicology and Chemistry, Vol. 21, No. 12, pp 2606-2616.

Otis, R.J. (2007). Estimates of Nitrogen Loadings to Groundwater from Onsite Wastewater Systems in the Wekiva Study Area: Task 2 Report. Florida Department of Health Report.

Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, and S. Christensen (2005). UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation: U.S. Geological Survey Techniques and Methods 6-A11, 283 pp.

Roeder, E. (2008). *Revied estimates of nitrogen inputs and nitrogen loads in the Wekiva study area*. Prepared for Florida Department of Health, Tallahassee, FL, by Bureau of Onsite Sewage Programs, Division of Environmental Health.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto (2004). Sensitivity analysis in Practice: A Guide to Assessing Scientific Models, J. Wiley & Sons.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, R. Srinivasan, and L.M. Hauck (2001). Validation of the SWAT model on a large river basin with point and onpoint sources. *J. American Water Resources Assoc.*, 37: 1169-1188.

Septic Tank Treatment System (SepTTS) Model for Cleaning Product Chemicals: Users Manual, Version 1.0 (1997). The Soap and Detergen Association, New York, NY.

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

Stark, S.L., J.R. Nurkols, and J. Rada (1999). Using GIS to investigate septic system sites and nitrate pollution potential. *Journal of Environmental Health* April.

Van Liew, M.W., J.G. Arnold, and J.D. Garbrecht (2003). Hydrologic simulation on agricultural watersheds: Choosing between two models. *Trans. ASAE* 46(6): 1539-1551.

Willmott, C. J. (1981). On the validation of models. Physical Geography 2: 184-194.

Willmott, C. J. (1982). Some Comments on the evaluation of model performance. Bulletin American Metrological Society.

Willmott C.J. and K. Matsuura (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance.