JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

AMERICAN WATER RESOURCES ASSOCIATION

ESTIMATING POTENTIAL *E. COLI* SOURCES IN A WATERSHED USING SPATIALLY EXPLICIT MODELING TECHNIQUES[1]

K.J. Riebschleager, R. Karthikeyan, R. Srinivasan, and K. McKee[2]

ABSTRACT: The Spatially Explicit Load Enrichment Calculation Tool (SELECT) was automated to characterize waste and the associated pathogens from various sources within a mixed land use watershed. Potential *Escherichia coli* loads in Lake Granbury watershed were estimated using spatially variable governing factors, such as land use, soil condition, and distance to streams. A new approach for characterizing *E. coli* loads resulting from malfunctioning on-site wastewater treatment systems (OWTSs) was incorporated into SELECT along with the Pollutant Connectivity Factor (PCF) module. The PCF component was applied to identify areas contributing *E. coli* loads during runoff events by incorporating the influence of potential *E. coli* loading, runoff potential, and travel distance to waterbodies. Simulation results indicated livestock and wildlife are potential *E. coli* contributing sources in the watershed. The areas in which these sources are potentially contributing are not currently monitored for *E. coli*. The bacterial water quality violations seen around Lake Granbury are most likely the result of malfunctioning OWTSs and pet wastes. SELECT results demonstrate the need to evaluate each contributing source separately to effectively allocate site specific best management practices (BMPs) utilizing stakeholder inputs. It also serves as a powerful screening tool for determining areas where detailed investigation is merited.

(KEY TERMS: nonpoint source pollution; pathogens; point source pollution; total maximum daily loading; water quality.)

Riebschleager, K.J., R. Karthikeyan, R. Srinivasan, and K. McKee, 2012. Estimating Potential *E. coli* Sources in a Watershed Using Spatially Explicit Modeling Techniques. *Journal of the American Water Resources Association* (JAWRA) 1-17. DOI: 10.1111/j.1752-1688.2012.00649.x

INTRODUCTION

Bacterial pathogens (fecal *Coliform* and *Escherichia coli* [*E. coli*]) are the leading cause of water quality impairments in the United States (USEPA, 2008). The total maximum daily load (TMDL) program, mandated by the Clean Water Act (CWA) Section 303, is a process to develop pollutant specific management plans integrating water quality assessment for protection of impaired watersheds. The goal of the CWA is to restore and maintain the chemical, physical, and biological integrity of the nation's waters. To meet the criteria of these mandates, models are often developed to study the current status of water quality and the impacts of various management plans (Bora and Bera, 2004). The Soil and Water Assessment Tool (SWAT) and hydrologic simulation program — FOR-

[1]Paper No. JAWRA-10-0079-P of the *Journal of the American Water Resources Association* (JAWRA). Received May 12, 2010; accepted January 24, 2012. **Discussions are open until six months from print publication.**

[2]Respectively, Former Graduate Research Assistant (Riebschleager), Associate Professor (Karthikeyan), and Research Associate (McKee), Department of Biological and Agricultural Engineering, Texas A&M University, 2117 TAMU, College Station, TX 77843-2117; Professor and Director (Srinivasan), Spatial Sciences Laboratory, Texas Agricultural Experiment Station, College Station, TX 77845 (E-Mail/Karthikeyan: karthi@tamu.edu).

TRAN (HSPF) are both watershed hydrologic simulation models used for evaluating best management practices (BMPs) and characterizing pollutant sources. Unfortunately, due to the complexity of modeling living organisms more research is needed to determine the fate of *E. coli* in aquatic environments as highlighted by Ferguson *et al.* (2003).

Different models have been developed to add decision makers through the process of a TMDL. Chen *et al.* (1999) have developed a decision support system for calculating TMDLs that employs stakeholder involvement along with watershed models. The decision support system, includes its own watershed simulation model, database, consensus building module, and a TMDL module with a calculation worksheet. The system generates various combinations of waste load and nonpoint load allocations to meet the water quality criteria.

The Center for TMDL and watershed studies at Virginia Tech has developed a software tool in Microsoft Excel, the Bacteria Source Load Calculator (BSLC), to support the bacterial source characterization process of the TMDL and automate the creation of input files for water quality modeling (Zeckoski *et al.*, 2005). The BSLC uses a systematic process that includes inventorying bacterial sources, estimating loads from these sources, and distributing estimated loads across the landscape as a function of land use and source type, and generating bacterial load input parameters for watershed-scale simulation models for source characterization. This loosely coupled model will become spatially referenced only if tied to a GIS-based model. In addition, the data for source populations are often available by county, not by subwatersheds. Consequently, the user has to redistribute the data on a subwatershed basis. The HSPF is used with the BSLC tool to simulate accumulation and die off of *E. coli* (Moyer and Hyer, 2003; Zeckoski *et al.*, 2005). This model does not provide maps, charts, or any other visual aid for decision making in the TMDL process. Thus, improved user-friendly tools are needed for conducting TMDL studies.

A representative watershed-scale water quality model is needed to address microbial pollution (primarily fecal *Coliform* and *E. coli*) issues. A comprehensive model can aid decision makers evaluate multifaceted problems and determine the appropriate course of action (Benham *et al.*, 2006; Deepti *et al.*, 2009; Jamieson *et al.*, 2004; Paul *et al.*, 2006; Santhi *et al.*, 2001). Geographic information systems (GISs) can aid in the difficult task of characterizing nonpoint source pollution in a watershed. A spatial semi-qualitative approach can aid the initial stages of TMDL development by concentrating efforts in the appropriate locations within the watershed as well as addressing the appropriate sources. The Spatially Explicit Load Enrichment Calculation Tool (SELECT) methodology was developed to assist in the source characterization component of the TMDL development process and watershed protection plans (WPPs) where bacterial contamination is a concern (Teague *et al.*, 2009). The SELECT is a pathogen load assessment tool, which can be combined with a watershed-scale water quality model using spatially variable governing factors, such as land use, soil condition, and distance to streams to support TMDLs and WPPs. This tool can be used to determine the actual contaminant loads resulting in streams when used in conjunction with a fate and transport watershed model. SELECT can simulate potential pathogen loading in a watershed for various management scenarios based on user defined inputs. Other more complex models, such as SWAT and WATFLOOD include bacteria fate and transport routines, but often these models are difficult to parameterize. Application of SELECT will help stakeholders identify the areas potentially contributing to pathogen contamination of waterbodies without using complex hydrologic models. A new addition to the SELECT is the Pollutant Connectivity Factor (PCF) component developed based on three indicative factors for contamination: (1) potential pollutant loading, (2) runoff potential, and (3) travel distance to streams and other waterbodies. The PCF component of SELECT offers stakeholders a less expensive, less time-consuming, and easier approach for evaluating BMPs by linking watershed loads to capability to contribute.

Previous application of the SELECT approach was performed through a series of manual operations in ArcGIS. The SELECT is now automated and an example of its applicability is provided in this research. A graphical user interface (GUI) was developed in visual basic for applications (VBA) within ArcGIS 9.X (ESRI, Redlands, CA, USA), where project parameters can be adjusted for various pollutant loading scenarios. From the visual output of the program a decision maker or stakeholder can identify areas of greatest concern for contamination contribution and incorporate that information, while developing the WPP or the TMDL. Details of the automated model development and results of applying SELECT to the Lake Granbury watershed in Texas to estimate daily potential *E. coli* loads are presented in this article.

METHODOLOGY

Spatially explicit modeling technique developed by Teague *et al.* (2009) to characterize *E. coli* sources in a watershed was automated to extend its application to other mixed land use watersheds and expanded to include on-site sewage facilities and the PCF component.

Spatially Explicit Approach

First, spatial factors that have the greatest influence on bacterial impairment of waterbodies were identified to develop a spatially distributed approach for estimating potential *E. coli* sources in the Lake Granbury watershed. This identification was carried out by consulting with agricultural and wildlife experts as well as stakeholders (primarily property owners, public service providers, and businesses) in the watershed during WPP meetings organized by the Brazos River Authority. Land use was identified as the factor that has the greatest effect on potential *E. coli* loading because the type of land use/land cover dictates whether an area contributes to bacterial contamination or not.

To characterize the production and distribution of waste and associated pathogens, contributing contaminant sources were determined. This was achieved by looking at the agricultural census information provided by National Agriculture Statistics Service (NASS), talking to the local extension agents and wildlife experts, obtaining permitted wastewater treatment plants' (WWTPs) discharges from the EPA Envirofacts data warehouse, and researching previous pathogen TMDLs. The fecal production rates for the various sources were calculated using the EPA protocol for developing pathogen TMDLs (USEPA, 2001), which includes a summary of source-specific pathogen and fecal indicator concentrations. Alternatively, local studies can and should be used when better information is available.

Finally, to integrate SELECT into a hydrologic simulation model, the potential loading on a daily time scale was needed. This was achieved by estimating the source populations, distributing the sources uniformly across suitable habitats, applying fecal production rates, and then aggregating to the level of interest (here, the subwatersheds) for analysis.

Watershed Description

Lake Granbury is a man-made lake within the Middle Brazos-Palo Pinto watershed. The Lake Granbury watershed was delineated into 34 subwatersheds (Figure 1) using ArcSWAT (SWAT, 2005). The city of Granbury is located in north-central Texas approximately 32 km southwest of Fort Worth, Texas. This is a diverse watershed characterized by multiple land use classifications (Figure 2). This lake is used for recreation and is a water source for municipalities, industries, and agriculture. This popular area is rapidly growing with a large number of people populating the areas around the lake.

Lake Granbury citizens are currently concerned about rising levels of bacteria within the coves of the lake. According to a recent water quality study (Espey Consultants Inc., 2007), there are four coves nearing bacteria impairments and one already impaired. In addition, four coves do not meet the dissolved oxygen standard, eight exceed the chloride standard, and one is approaching the nitrogen screening level. Currently, the main body of the lake is not impaired due to bacteria, but if conditions continue to worsen in the coves it is possible the lake would potentially be contaminated. There are two centralized sewage systems and new residential areas have on-site wastewater treatment systems (OWTSs) near the coves of the lake. Unfortunately, much of the soil around the lake is not suitable for traditional septic tank and gravity trench soil treatment areas in addition to small lot sizes. SELECT was applied to characterize and estimate potential *E. coli* loads in the Lake Granbury watershed. The authors would like to note that at the time this article was developed the WPP was still under development and results do not necessarily reflect the final inputs or conclusions of the Lake Granbury WPP. The focus of this article was to describe the new automated SELECT approach and the flexibility and applicability of the tool.

Geographic Information Systems Modeling Framework

The development of the automated tool started with using the model builder application within GIS to conceptualize the file processing and determining appropriate input parameters for each type of source assessment (livestock, wildlife, OWTSs, pets, and WWTPs). A GUI was developed in VBA to create a tightly-coupled model within ArcGIS 9.X. The GUI was used to create the watershed project setup, add layers to the map, and input parameters, such as appropriate habitats, source populations, and fecal production rates. The next step was to process the spatial files using the inputs from the GUI (Table 1). The map processing code was written using ArcObjects relationship classes and divided into several modules.

A central module processes information from the GUI and then initializes the appropriate subroutines within the various modules in an ordered sequence of events. The remaining modules contain subroutines

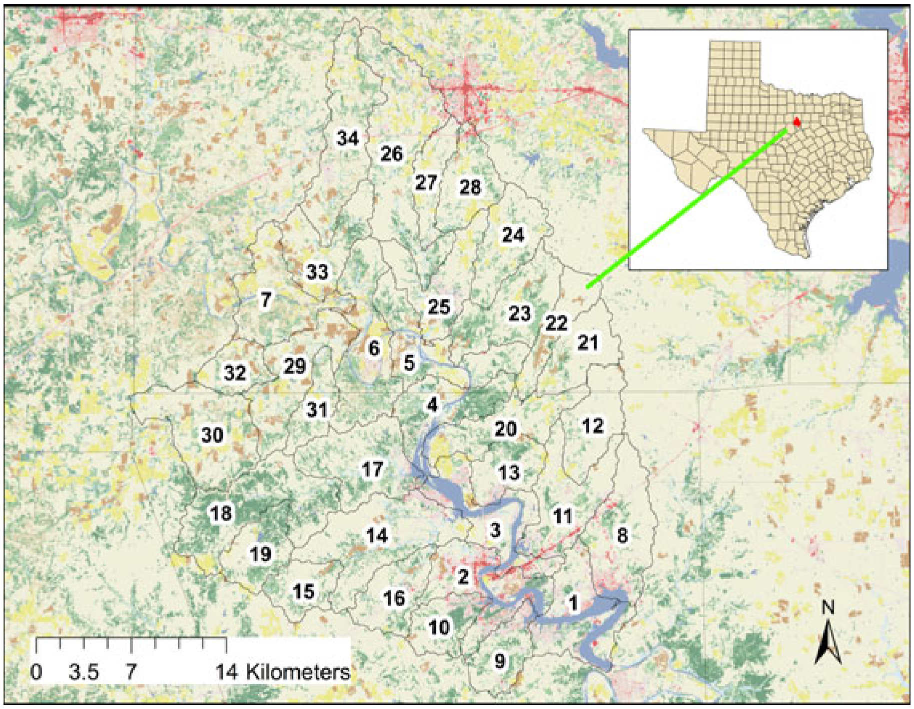

FIGURE 1. Location of Lake Granbury with Subwatersheds Delineated Using SWAT (Parker County — northern portion of the watershed; Hood County — southern portion of the watershed).

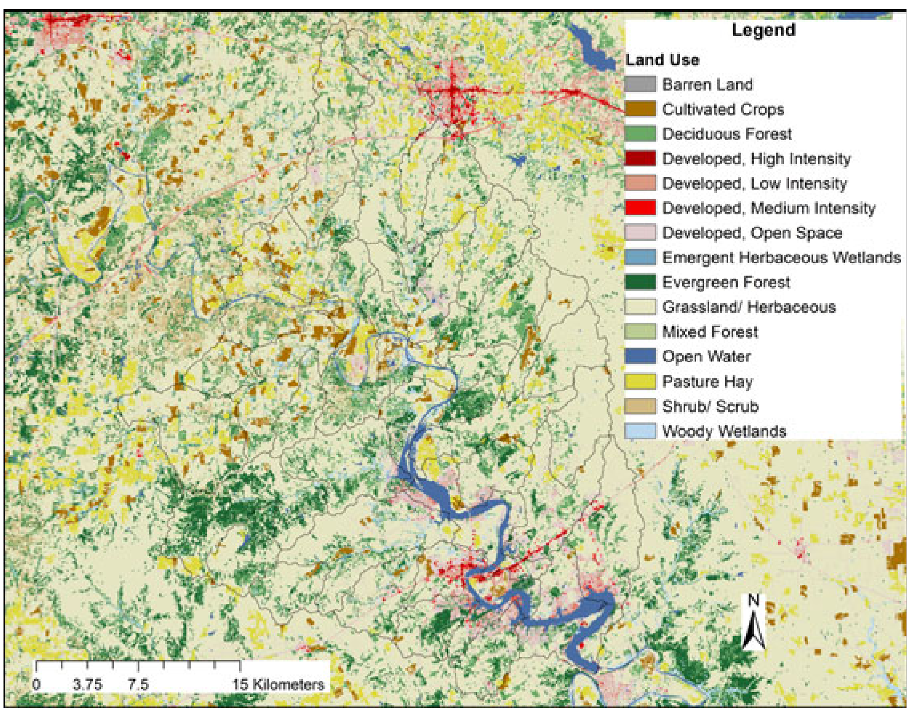

FIGURE 2. Land Use Classification of Lake Granbury Watershed (NLCD, 2001).

for determining the potential loading from both point (WWTPs) and nonpoint (livestock, wildlife, and domestic) sources. The livestock module has separate subroutines for cattle, dairy, sheep/goats, horses, and swine. The wildlife module calculates potential loading for deer, feral hogs, and two generic (other1 and

TABLE 1. Data Sources and Format Used in SELECT to Predict Potential *E. coli* Load in Lake Granbury Watershed.

Pollutant Source	File	Format	Data Source	Comments
Livestock	Counties	Shapefile	NASS	Include only needed counties in file
	Ag inventory	Tabular		Program does not read from file
Wildlife	Suitable habitat	Shapefile	Local wildlife census	Needed for Method 1
	Urban areas	Shapefile	TIGER census	Method 2: (optional)
	Streams	Shapefile	NHD plus	Method 2: feral hogs
OWTSs	Subdivisions	Shapefile	Appraisal district	Method 1: need age and no. of permit records fields
	Census blocks	Shapefile	TIGER census	Method 2: merged for all counties
	Demographics	Tabular	TIGER census	Method 2: state demo. table
	Soils	Shapefile	SSURGO	Separate for each county
	Soil properties	Tabular	SSURGO	
Pets	Census blocks	Shapefile	TIGER census	Separate for each county
	Demographics	Tabular	TIGER census	State census block demo graphics table
WWTP	Outfall locations	Shapefile	State regulatory agency	Remove nonpathogenic outfalls and inactive permits
	Permitted discharge	Field in shapefile	EPA Envirofacts warehouse	Create field in outfall locations file

2) sources. Subroutines for OWTSs and pets are part of the domestic module. The urban module has a subroutine for calculating *E. coli* contributions from WWTPs. Finally, the pollutant connectivity module is a set of subroutines for weighting the driving forces of pollutant contributions reaching waterbodies to create the PCF.

Model Simulation

The pathogen sources selected for the Lake Granbury example were beef cattle, OWTSs malfunction, deer, and WWTPs. The default fecal production rates used in the simulation were the highest from the range of values provided in the EPA protocol for developing pathogen TMDLs (USEPA, 2001) for all *E. coli* sources in the Lake Granbury watershed (Table 2). Spatial analysis was conducted at $30 \text{ m} \times 30 \text{ m}$ resolution and the results were aggregated at subwatershed level (Figure 1).

Potential *E. coli* Sources in Lake Granbury Watershed. SELECT simulated potential *E. coli* loads resulting from cattle, deer, pets, malfunctioning OWTSs, and WWTPs.

Livestock. All livestock populations (beef cattle, dairy cattle, sheep/goats, swine, and horses) were obtained from the 2002 NASS inventory on a per county basis. In this watershed, the livestock included only range cattle. Once appropriate land use classification (indicated within the SELECT GUI) was chosen, the automated program clipped the land use file to create a land use grid for each county. A raster from the indicated land use for each county was reclassified into suitable (value of 1) and nonsuitable (0). Next, the population density grid was created by multiplying the suitable habitat grid times the population and divided by the number of cells. The population density grids for each county were combined using the mosaic operation into one population density grid. Finally, the population density grid was multiplied by the fecal *Coliform* production rate indicated in the user form and converted into an *E. coli* production rate using a conversion factor of 0.5 (Doyle and Erikson, 2006). The conversion factor is a needed adjustable parameter in the project setup as the transition from fecal *Coliform* to *E. coli* as the indicator organism for pathogens is a recent development and the ratio tends to be specific to the geographic area. It is recommended that overlapping fecal *Coliform* and *E. coli* data from the same observation location and time within the watershed are compared to estimate this ratio. Finally, a zonal sum was performed to aggregate the resultant load for each subwatershed.

The cattle populations for Hood and Parker counties were 30,059 and 71,601 cattle, respectively (USDA-NASS, 2002). The cattle population was distributed uniformly on grasslands (NLCD Classification 71) and pasture/hay (NLCD Classification 81), as cattle graze mainly on these land uses. There are no concentrated animal feeding operations (CAFOs) in the watershed.

Wildlife. Using SELECT a user can account for wildlife contributions by distributing population

TABLE 2. Calculation of *E. coli* Loads from Source Populations.

Source	Calculation
Cattle	$EC = \#\text{ Cattle} \times 10 \times 10^{10}\text{ CFU/day} \times 0.5$
Deer	$EC = \#\text{ Deer} \times 3.5 \times 10^{8}\text{ CFU/day} \times 0.5$
Dogs	$EC = \#\text{ Households} \times \frac{0.8\text{ dogs}}{\text{Household}} \times 5 \times 10^{9}\text{ CFU/day} \times 0.5$
Malfunctioning OWTSs	$EC = \#\text{ OWTSs} \times \text{Malfunction Rate} \times \frac{1 \times 10^{6}\text{ CFU}}{100\text{ ml}} \times \frac{60\text{ gal}}{\text{person/day}} \times \frac{\text{Ave }\#}{\text{Household}} \times \frac{3758.2\text{ ml}}{\text{gal}} \times 0.5$
WWTP	$EC = \text{Permitted mgd} \times \frac{126\text{ CFU}}{100\text{ ml}} \times \frac{10^{6}\text{ gal}}{\text{mgd}} \times \frac{3758.2\text{ ml}}{\text{gal}}$

estimates across suitable habitats as determined by consultation with wildlife experts. The first step in calculating wildlife pollutant loading is to identify the types of wildlife most likely contributing the most significant amounts of pollution and consider the sources that only minimally contribute. This was achieved by consulting wildlife experts such as the Texas Parks and Wildlife Department (TPWD), thorough literature review, and gathering stakeholder input. A stakeholder group typically consisted of farmers, ranchers, common public, administrators, and extension personnel living in the watershed. Stakeholder input was gathered at public meetings discussing the watershed protection and water quality. It is also important to identify the land uses wildlife prefer/need for survival, along with population estimates. Many agencies such as the TPWD have published studies that address these issues. Currently, SELECT provides the option to evaluate pollutant loading of *E. coli* from deer, feral hogs, and two other generic sources. The program allows for two methods for estimating wildlife loadings. In the first method, the user inputs a suitable habitat shapefile and then the program assumes the wildlife will graze only in these areas. In the second method, the user indicates appropriate land use and whether or not to include urban areas and allows the model to distribute the populations on the suitable habitat based on built-in assumptions. The final suitable habitat for population distribution is determined based on the selected land use and other assumptions (for deer at least 20 acres of contiguous terrain should be available). Once the suitable habitat is created, fecal production rates are multiplied by the population density and then the total loading for the source to each zone of interest (here, subwatershed) is aggregated.

The population density of 13 deer per 1, 000 acres for the Lake Granbury watershed was estimated (Lockwood, 2005) based on the Resource Management Unit (RMU) within Lake Granbury watershed. It was assumed that deer roam in forested areas (land use codes: 41, 42, and 43) and shrubland (52). Urban areas were removed from the suitable habitat for this study.

On-site Wastewater Treatment Systems. Another need for bacteria load assessment is an improved understanding of when OWTSs malfunction, how much these systems contribute to contamination, and how to reasonably predict such occurrences. For evaluating the potential *E. coli* loading from malfunctioning OWTSs, a new approach different from Teague *et al.* (2009) was developed. Clark *et al.* (2001) indicated that the age of OWTSs, soil condition, and vicinity to water bodies have the greatest influence on contamination due to OWTSs. Methods for developing a sewage pollution risk assessment have been developed and were used as a guideline (Kenway and Irvine, 2001). Combining this methodology for OWTSs risk assessment with soil landscape mapping can assess the individual system contribution to the cumulative risk of sewage pollution (Chapman *et al.*, 2004). Two methods for OWTSs malfunction prediction have been created for the SELECT. The first method can be used when detailed OWTSs permit information is available. The second method relies only on readily available public data sources.

Method 1: This method was developed based on the age of subdivisions and the septic absorption field limitation ratings (slight, moderate, and severe) provided with National Resource Conservation Service (NRCS) Soil Survey Geographic (SSURGO) soils data (USDA-NRCS, 2004). The user inputs the appropriate OWTSs shapefile and indicates the "fields" within the attribute table containing the number of permits and the average estimated age of the subdivision/OWTSs in each polygon. This information can be gathered from health department permit records where available, parcel data with the year the home was built, or the years the subdivision was under development can be gathered from the homeowner associations. The number of systems contributing to the potential *E. coli* load is determined from the number of homes on OWTSs multiplied by the expected percent malfunction. The percent malfunction is a reclassification of the OWTSs suitability rating for a given area. The suitability rating is calculated as:

$$\text{Suitability Rating} = 0.7 \times \text{Soil Rate} + 0.3 \times \text{Age Rate} \tag{1}$$

The program creates an age rating for the OWTSs shapefile (Table 3), and a soil rating based on the SSURGO soil limitation ratings of severely limited

TABLE 3. Age Rating for Subdivisions in Lake Granbury Watershed to Calculate OWTSs Index.

Age (Years)	Age Rate
0-15	1
16-30	2
>30	3
No data	−99

TABLE 4. Interpretative Soil Properties and Limitation Classes for Septic Tank Soil Absorption Suitability (Source: SCS, 1986).

	Limitation Class		
Interpretive Soil Property	**Slight**	**Moderate**	**Severe**
Total subsidence (cm)	-	-	>60
Flooding	None	Rare	Common
Bedrock depth (m)	>1.8	1-1.8	<1
Cemented pan depth (m)	>1.8	1-1.8	<1
Free water occurrence (m)	>1.8	1-1.8	<1
Saturated hydraulic conductivity (μm/s)			
Minimum 0.6-1.5 m[1]	10-40	4-10	<4
Maximum 0.6-1 m[1]			>40
Slope (Pct)	<8	8-15	>15
Fragments >75 mm[2]	<25	25-50	>50
Downslope movement			[3]
Ice melt pitting			[3]
Permafrost			[4]

[1]0.6 to 1.5 m pertains to percolation rate; 0.6 to 1 m pertains to filtration capacity
[2]Weighted average to 1 m.
[3]Rate severe if occurs.
[4]Rate severe if occurs above a variable critical depth (see discussion of the interpretive soil property).

(3), somewhat limited (2), and slightly limited (1). The soil and age rating was estimated on a lot by lot basis. The NRCS limitation ratings are based on geophysical factors, such as soil classification, depth to bedrock, and slope (Table 4). The soil file with the suitability rating is intersected with the age rate and then weighted with 70% to soil rate and 30% to the age rating to create a new OWTSs malfunction index. This weighting scheme is based on the assumption that soil treatment capability has the greatest role in contribution, followed by malfunction due to limited maintenance (related to age of system) (Bruce Lesikar, Texas A&M University, December 7, 2007, personal communication). Areas missing soil or age information are assigned index ratings of −99. In this case, the higher the suitability rating, the less effluent the system can treat. A malfunction index based on the suitability rating is converted to a raster file and then reclassified into percent malfunctioning (contributing to load potential) (Table 5). After determining the number of homes potentially contributing per subdivision (polygon), a flow rate (gal/person × day), effluent rate (CFU/100 ml), the average population per home, and necessary conversion

TABLE 5. OWTSs Index Reclassification to Percent Malfunction Used in Determining OWTSs Malfunction Rates in Lake Granbury Watershed.

Index	% Malfunction
<0	8
0-1.5	5
1.5-2.5	10
2.5-3	15

factors are applied to estimate the potential *E. coli* loading in CFU/day.

Method 2: The second method is conceptually similar to Method 1; however, using only publicly available information. To determine the number of OWTSs without detailed permit information, the number of homes is estimated using the U.S. Census Bureau census block shapefile with demographics and then creating a raster grid (USCB, 2000). Areas using municipal sewage are removed, determined from the Texas Commission on Environmental Quality (TCEQ) shapefile (TCEQ, 2008a) with Certificates of Convenience and Necessity (CCN) sewer service areas, by creating a "not sewered" grid and then multiplying by the number of homes grid. The potential loading is then determined in the same manner as in Method 1, except the suitability rating is simply the SSURGO soil rate when age data were not available.

Method 1 for predicting OWTSs *E. coli* contributions was applied to the Lake Granbury watershed. OWTSs information was obtained from county permit records (Hood County Appraisal District). The population density, 1.94 people per home, was estimated from the year 2000 Hood County Census (U.S. Census Bureau). SSURGO soil shapefiles for each county and the associated soil properties tables were obtained from the NRCS Soil Datamart. Detailed OWTSs information was not available for Parker County. Method 2 was not utilized based on stakeholder request to focus study on areas close to the lake and were not interested in potential loading from OWTSs in the upper watershed for this WPP.

Pets. Generally, dogs are the primary pet allowed to defecate outside the home and most often the defecated waste is not cleaned up. Cats and other pets are primarily kept in homes and waste disposed of directly to solid waste management therefore these contributions were neglected. The assumption of a constant one dog per home for Texas (AVMA, 2002) was an adjustable model parameter included in SELECT. The program creates a raster that represents the number of homes from the census block demographics table joined to the census block shapefile. Again the program applies the fecal production rate and then aggregates the potential load to zones of interest, here subwatersheds. Census block shapefiles are needed for each county to determine the spatial distribution of homes.

Wastewater Treatment Plants. Contribution of potential *E. coli* from point sources such as WWTPs in the watershed was estimated by providing spatial information and permitted discharges of WWTPs *E. coli* loading was calculated by simply multiplying the effluent *E. coli* standards by the discharge and applying conversion factors to determine the loading in CFU per day. For this study, seven wastewater outfall locations in the watershed (covering 60% of the total population) were obtained from TCEQ as GIS shapefiles (TCEQ, 2008b). The permitted flows were obtained from the EPA Envirofacts data warehouse (USEPA, 2006). There are no CAFOs in the watershed hence WWTPs are the only point source included in this study.

Once all individual source inputs were selected summation of potential *E. coli* loads from all sources was carried out. Thus, potential loading from the most prevalent sources in Lake Granbury were spatially distributed and summarized at the subwatershed level of interest.

Pollutant Connectivity Module. The Italian Environmental Protection Agency has developed the Potential Nonpoint Pollution Index (PNPI), a GIS-based watershed scale tool (Munafo *et al.*, 2005). PNPI is a simple method designed to inform decision makers about the potential environmental impacts of different land management scenarios. This tool helps the user detect and display areas that are likely to produce pollution due to their land use, geo-morphology, and location with respect to the stream network. This approach uses expert knowledge to generalize the relationship between the land cover indicator (LCI), run-off indicator (ROI), and the distance indicator (DI) to study the driving forces of pollution instead of impacts (Munafo *et al.*, 2005). A similar approach was taken here to weigh the influence of the driving forces of *E. coli* contamination with the total *E. coli* load present in the watershed by PCF.

The PCF indicates areas within the watershed vulnerable to contributing bacteria to waterbodies. Using this module, the user can screen the relative impact of loads from the contributing watershed to the nearest waterbodies by combining the SELECT potential loading with the curve number, which directly relates to runoff potential, and the distance to streams, which directly relates to fate and transport. The total PCF was calculated using a weighted combination of the potential loading (normalized on a scale of 0-100), curve number grid, and the inverse of the flow length to streams (normalized on a scale of 0-100) (Figure 3). The average flow length for each subwatershed was derived from a digital elevation model (DEM) using ArcHydro Tools within ArcGIS. The curve number grid was created from intersecting the SSURGO soils hydrologic soil grouping (HSG) and the NRCS 2001 land use classification and then using a user defined NRCS Curve Number lookup table. The NRCS Curve Number indicates the runoff potential of an area based on the hydrologic soil group, land use type, and antecedent moisture condition of the soil (Haan *et al.*,

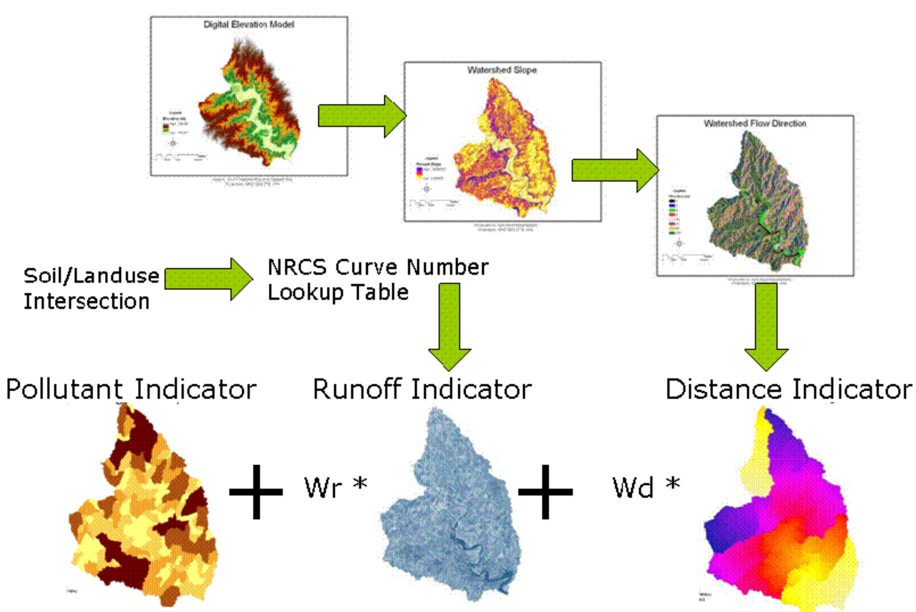

FIGURE 3. Spatial and Hydrologic Processes to Determine the Pollutant Connectivity Factor.

1994). The resulting PCF is a ranking of potential contribution from subwatersheds without considering any detailed fate and transport processes in the watershed. The following is the weighted overlay expression for determining the PCF:

$$\mathrm{PCF} = W_{\mathrm{P}} \times P_{\mathrm{I}} + W_{\mathrm{R}} \times R_{\mathrm{I}} + W_{\mathrm{D}} \times 1/D_{\mathrm{I}} \tag{2}$$

where PCF = Pollutant Connectivity Factor, W_{P} = weighting factor for the pollutant indicator, P_{I}, P_{I} = pollutant indicator, normalized pollutant load on scale from 0 to 100, W_{R} = weighting factor for the runoff indicator, R_{I}, R_{I} = runoff indicator, curve number, W_{D} = weighting factor for the distance indicator, D_{I}, and D_{I} = distance indicator, normalized flow length on scale from 0 to 100.

In this study, multiple trials of the PCF with an array of weighting factors were run and then averaged. An example weighting scheme is presented in Table 6. Alternatively, a weighing scheme developed based on stakeholder recommendations and expert knowledge for the most important factors was also used for comparison. If a particular subwatershed consistently is determined to be a “hot spot” for contributing to potential *E. coli* contamination, then it is likely that this subwatershed is of great concern and should be more readily addressed. On the other hand, if a particular watershed is consistently rated low (regardless of weighting factors), then this watershed should not be of concern when determining management practices. Consideration should be given to scale of watersheds when analyzing these results. If great disparities in the watershed size distribution are present, then it may be appropriate to area weight the potential load prior to using the PCF application.

TABLE 6. Example Weighting Schemes for Sensitivity Analyses of Pollutant (W_{p}), Runoff (W_{r}), and Distance Indicators (W_{d}) for Determining the PCF.

Trial Number	W_{p}	W_{r}	W_{d}
1	5	3	2
2	5	2	3
3	4	4	2
4	4	3	3
5	4	2	4
6	3	5	2
7	3	4	3
8	3	3	4
9	3	2	5
10	2	5	3
11	2	4	4
12	2	3	5
13	3.33	3.33	3.33

RESULTS AND DISCUSSION

Potential *E. coli* loadings from livestock, wildlife, and domestic sources in the Lake Granbury watershed were estimated by SELECT. The loadings from the individual sources were combined and aggregated on a subwatershed basis (Figure 4). By doing this aggregation, potential source contributions were spatially distributed across the watershed. However, this was only a daily estimate of the potential of *E. coli* load present in the watershed under the assumed scenario. The PCF provided helpful information to determine whether *E. coli* from various sources potentially contaminate the waterbodies or not by applying weighting to important fate and transport factors, such as runoff capabilities, and travel distance. This weighting scheme when based on watershed characteristics provides a screening tool to indicate the areas of highest concern for *E. coli* contamination (Figure 5a). For the Lake Granbury watershed, PCF analyses were based on applying multiple weighting schemes and then ranking the subwatersheds for potential water quality problems due to *E. coli* (Figure 5b). It should be noted that not all the subwatersheds had monitoring stations when this study was conducted. This limited the scope of validating the results obtained from this screening tool. However, the results from SELECT and the PCF rankings were compared with the available water quality data to help decision makers and stakeholders develop a spatially explicit WPP and develop a new water quality monitoring plan (Table 7).

Daily Potential E. coli *Loading in Lake Granbury Watershed*

The potential *E. coli* loading is divided into two classes for analyses: nonpoint (Figure 6) and point sources (Figure 7). For each of these classes it is important to consider how potential loads are compared with actual *E. coli* concentrations in waterbodies, as measured at water quality monitoring locations (Figure 8 and Table 7). This verified significant contributions from OWTSs as other major sources (livestock and wildlife) are not significant in these monitored subwatersheds (Figures 6, 7, and 8).

Nonpoint Sources. High potential *E. coli* load resulting from cattle (Figure 6a) occurs in the northern most subwatersheds 26 and 34 as well as in subwatersheds 14 and 30 (Figure 1). These subwatersheds have a landscape dominated by grasslands with a mixture of pasture/hay (Figure 2 and Table 8). The middle of the watershed has lower loads mainly due

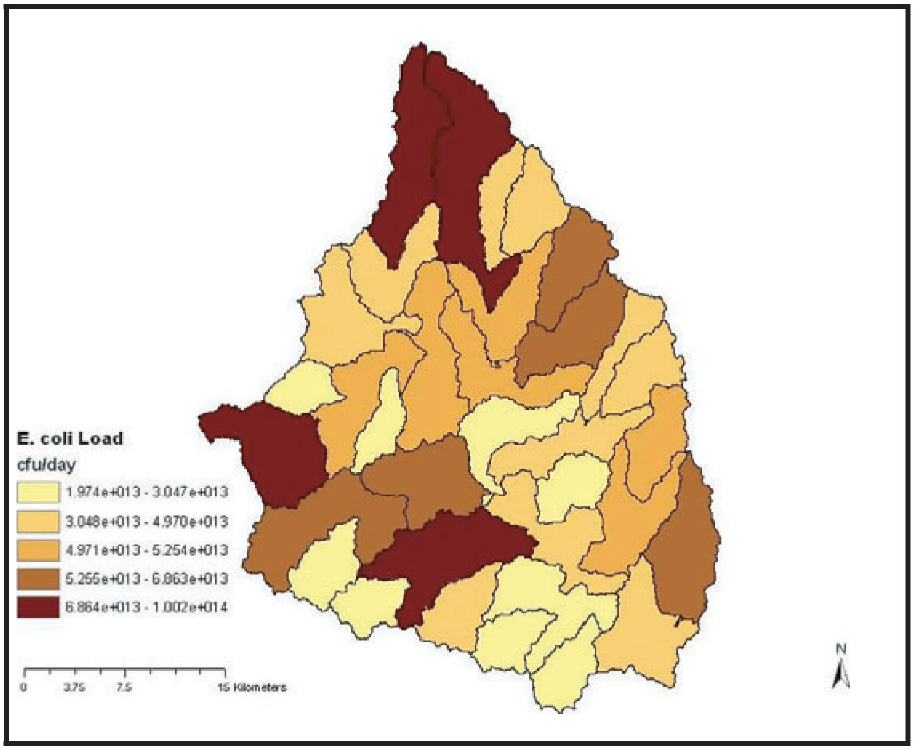

FIGURE 4. Total Potential *E. coli* Load from All Sources in Lake Granbury Watershed.

TABLE 7. Comparison of *E. coli* Monitoring Data with SELECT-PCF Analysis.

Subwatershed	No. Monitoring Stations	Total Number of Samples[1]	% Samples Exceeding 126 CFU/100 ml	SELECT Estimated Potential *E. coli* Load (CFU/day)	SELECT-PCF Ranking[2]
1	21	265	3-43	$(3.048\text{-}4.97) \times 10^{13}$	9
2	14	174	5-22	$(1.97\text{-}3.047) \times 10^{13}$	10
3	12	180	4-43	$(3.048\text{-}4.97) \times 10^{13}$	8
4	7	96	8-22	$(1.97\text{-}3.047) \times 10^{13}$	10
5	1	12	Not exceeding	$(4.98\text{-}5.25) \times 10^{13}$	6
7	1	12	Not exceeding	$(3.048\text{-}4.97) \times 10^{13}$	8
8	6	75	2-11	$(5.26\text{-}6.86) \times 10^{13}$	4
11	3	27	Not exceeding	$(4.98\text{-}5.25) \times 10^{13}$	5
13	5	60	3-53	$(1.97\text{-}3.047) \times 10^{13}$	10
17	3	45	6-11	$(5.26\text{-}6.86) \times 10^{13}$	5
20	1	15	Not exceeding	$(3.048\text{-}4.97) \times 10^{13}$	6

[1]*E. coli* samples were collected on a quarterly basis with a few additional samples in some cases. The sampling duration was three years.
[2]PCF ranging range was between 1 and 11 (Figure 5b). Watersheds ranked 1 by PCF analysis had the highest potential for *E. coli* contamination; and watersheds ranked 11 had the lowest *E. coli* contamination concern.

to higher human population. Subwatershed 14 is an area of potential concern due to its close proximity to the lake with highest *E. coli* potential load. Further analysis using the PCF was applied to verify this concern (Figure 5b). However, this could not be verified with actual monitored *E. coli* data because there was no monitoring station in this watershed (Figure 8 and Table 7). During a runoff event the highest ranked "hot spots" are the most likely to significantly contribute to contamination in the waterbodies. The same subwatersheds with high-potential loads were determined to be the three highest ranked, by PCF, areas likely to be contributing to contamination in the waterbodies. The highest average PCF ranking was subwatershed 34. Water quality data could be used to verify the PCF results; however, the subwatersheds with high-loading resulting from cattle are not monitored for *E. coli* concentrations (Figure 8 and Table 7).

The highest potential *E. coli* loading resulting from deer (Figure 6b) can be seen in the northern portions of the watershed where human population is less

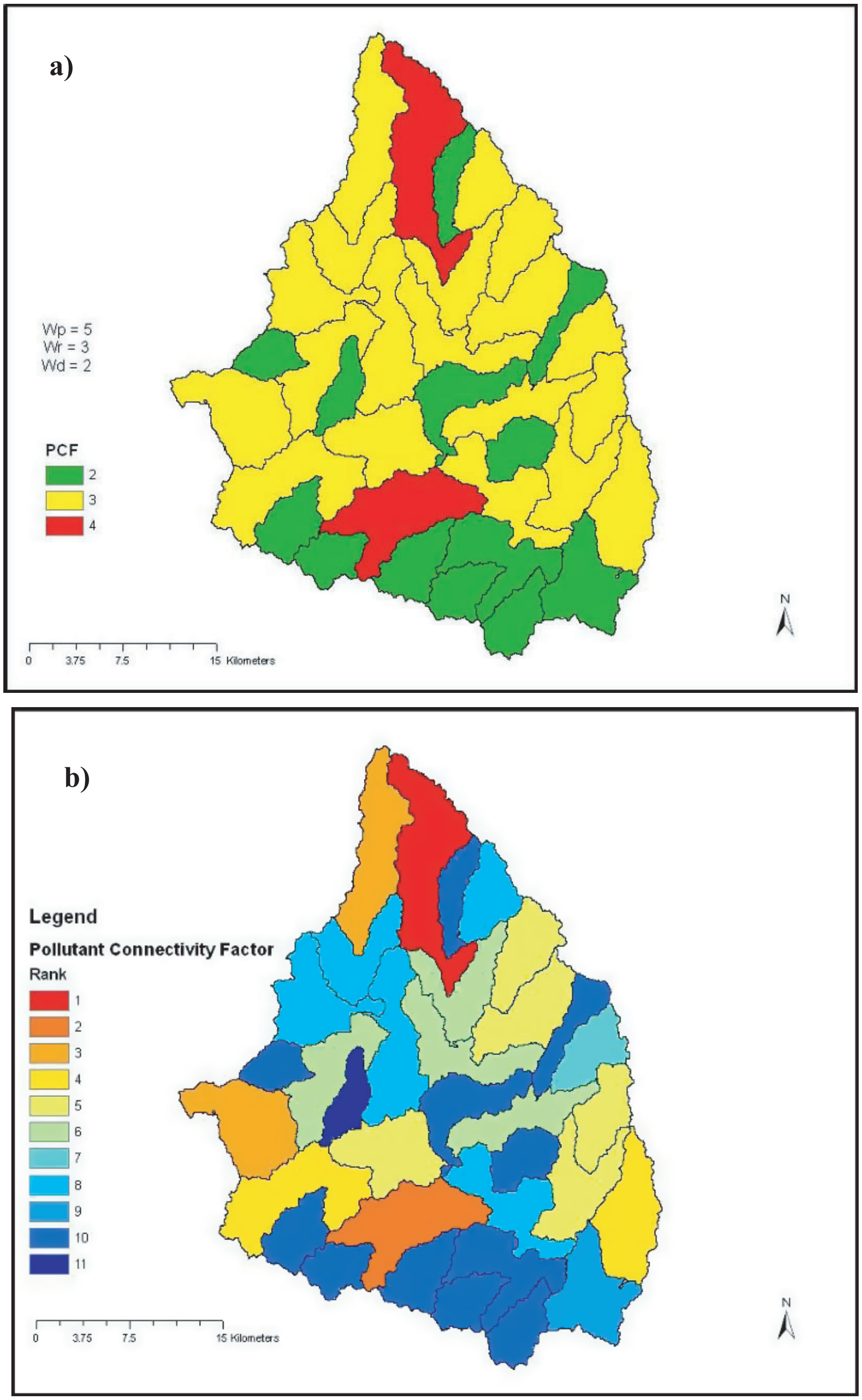

FIGURE 5. Pollutant Connectivity Factor for Total *E. coli* Potential Load Determined by (a) Expert Knowledge Weighting, and (b) Ranked Subwatersheds Averaged over Multiple Weighting Scenarios.

dense. The subwatersheds with the highest potential loading (6, 18, 23, 26, and 34, [Figure 1]) have large amounts of forest land use. The second highest group of potential loading tends to have significant amounts of forests, but these areas are more scattered and broken up by streams and intermixed with open range and grasslands. The southern half of the watershed generally has lower potential loads resulting from deer mainly due to the influence of higher human populations. When these loads are compared

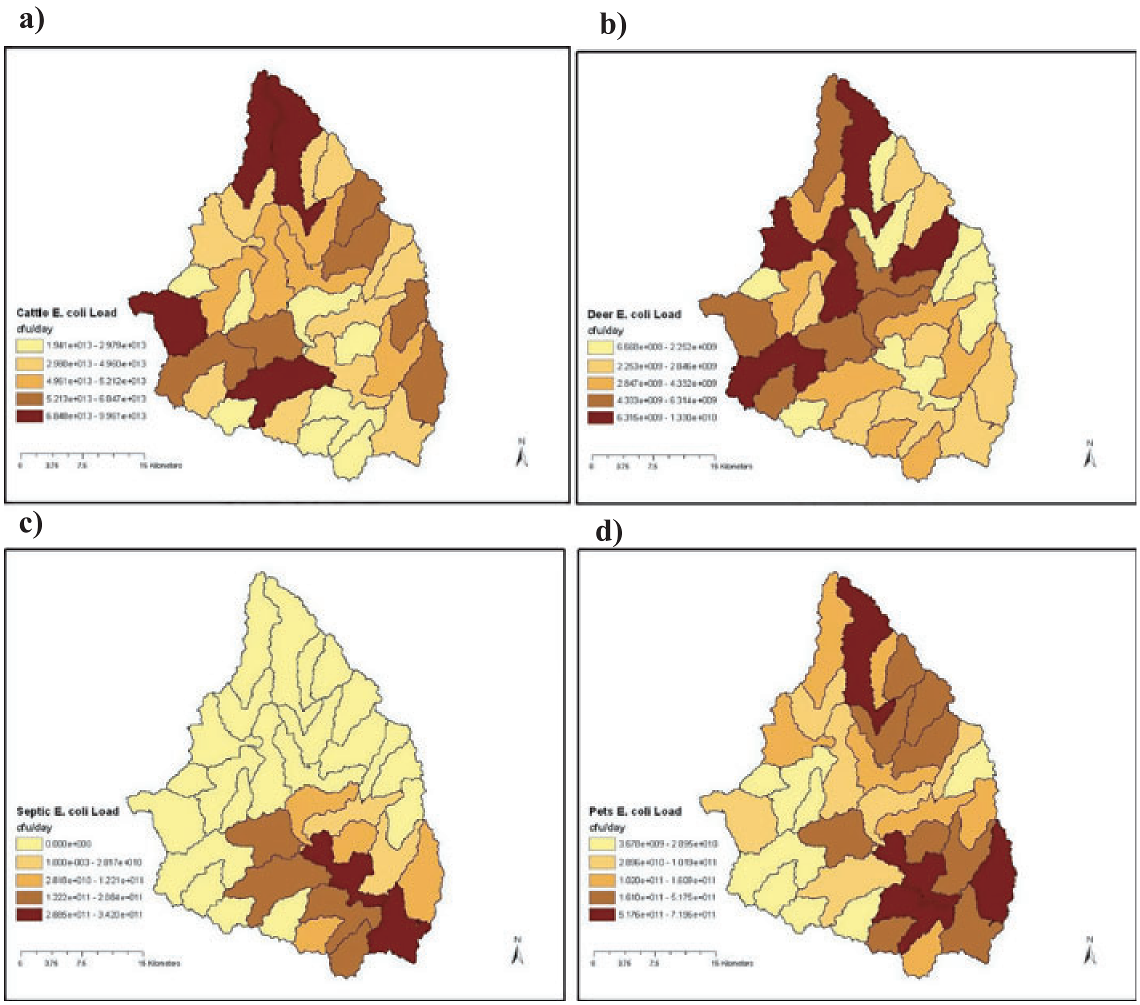

FIGURE 6. Potential *E. coli* Load in Lake Granbury Watershed Resulting from Various Nonpoint Sources: (a) Cattle, (b) Deer, (c) On-site Wastewater Treatment Systems (OWTSs), and (d) Pets.

with the PCF ranking, again subwatersheds 26 and 34 are among the areas of high concern. Subwatersheds 6, 18, and 23 are in the middle range of PCF ranking (fourth through eighth). Unfortunately, all of the subwatersheds with high-loading resulting from deer are not monitored for *E. coli* concentrations as well (Figure 8 and Table 7).

Potential *E. coli* loading resulting from malfunctioning OWTSs (Figure 6c) was calculated for Hood County only where descriptive permit data was gathered to create a spatial subdivision OWTSs file by the Brazos River Authority from the Hood County Appraisal District. This information has not been gathered for Parker County (T. Morgan, Brazos River Authority, Waco, Texas, 2008, personal communication). This does not pose a significant problem as the northern portion of the watershed in Parker County is much further from the waterbodies of concern. In addition, the only areas with significant populations are on the northeastern edge of the watershed where the populations are quite dense and most likely on combined sewer networks. Method 2 for OWTSs malfunction potential loading without detailed permit information could be run to verify this assumption. At the request of the WPP coordinator and stakeholders, Method 2 was not run since future modeling efforts would focus primarily on the potential sources within a 2-mile buffer of the lake to help focus source identification and implementation efforts under limited time and budget constraints. Subwatersheds 1 and 3 are located along the middle of Lake Granbury and had the highest potential *E. coli* loads resulting from malfunctioning OWTSs. Subwatershed 1 is characterized by developed low-intensity land use, mostly with single-family housing units. Subwatershed 3 has developed medium- and high-intensity land use, which includes single-family housing units with higher percent impervious land cover. The second highest potential loading group is located west of the lake and characterized by residential development

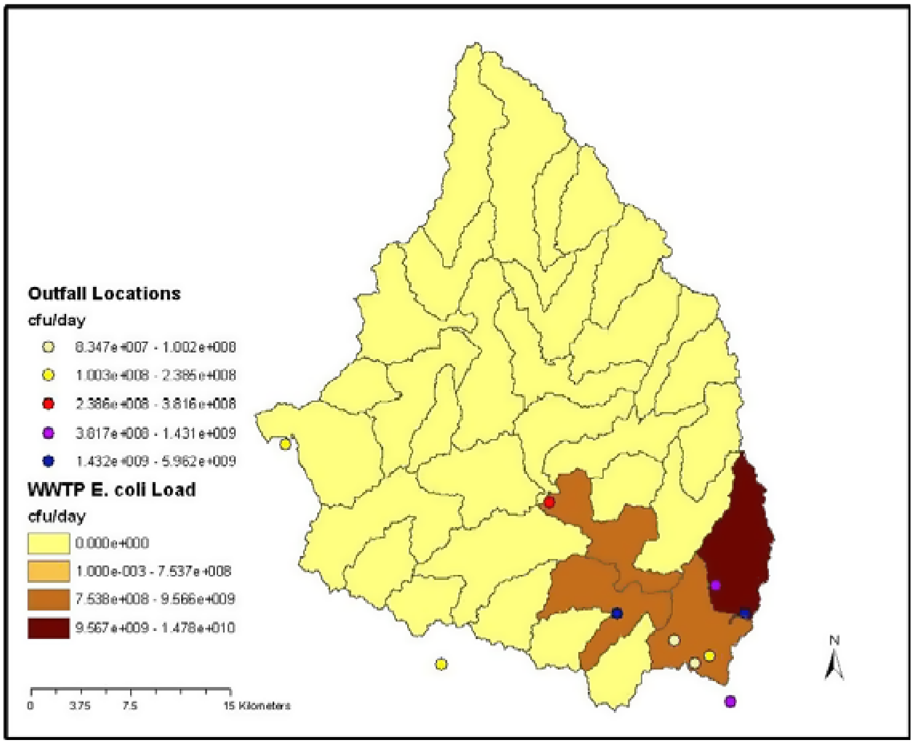

FIGURE 7. Potential *E. coli* Loading from Point Sources (wastewater treatment plants).

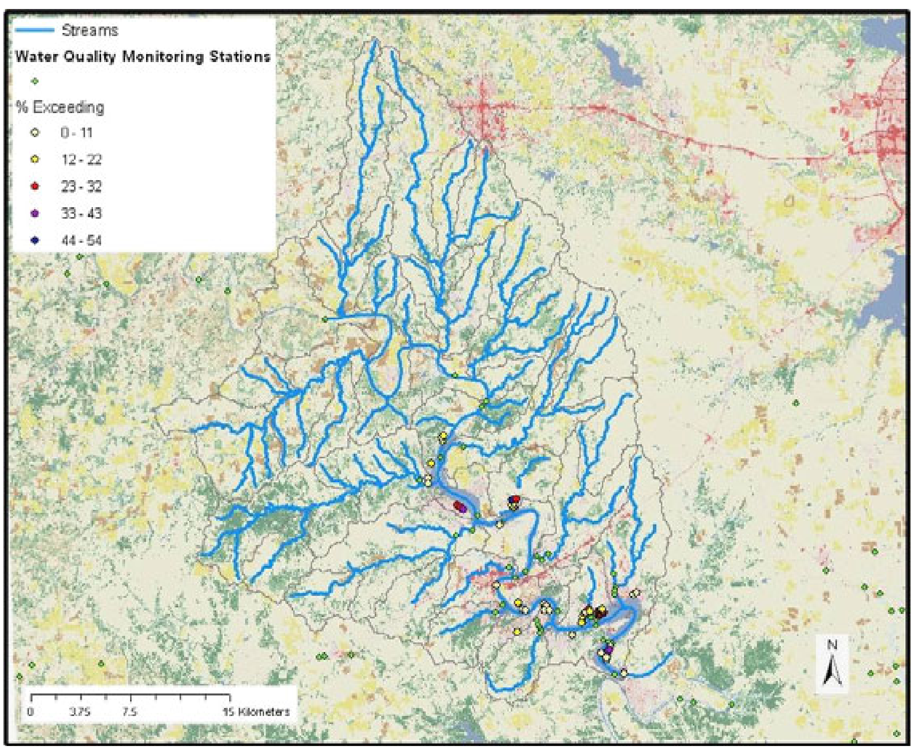

FIGURE 8. Water Quality Monitoring Stations Located within the Lake Granbury Watershed with Percent of Observations Exceeding *E. coli* Standard (126 CFU/10).

TABLE 8. Land Use Statistics for Lake Granbury (BRA, 2008).

Land Use	0-1 Mile (%)	1-2 Mile (%)	Total (%)
Multifamily residential	>1	>1	>1
Single-family residential	40	18	30
Commercial/services	4	2	3
Industrial	>1	>1	>1
Utilities/transportation	1	>1	>1
Recreational	3	>1	2
Cropland/pasture	26	31	28
Orchards	>1	2	>1
Other agriculture	>1	>1	>1
Rangeland	23	43	32
Quarries	>1	>1	>1

scattered among undeveloped grasslands, forests, and pastures. The areas potentially contributing significant *E. coli* loadings resulting from malfunctioning OWTSs range from a PCF ranking of 3 to 10. Water quality monitoring data for *E. coli* in subwatersheds 1 and 3 indicate several stations where from 23 to 43% of observations at these locations exceed the maximum concentration standard of 126 CFU/100 ml (Figure 8 and Table 7).

The potential *E. coli* loading resulting from pets (Figure 6d) is highest in subwatershed 26 in the northern portion of the watershed, subwatershed 8 along the southeastern edge, and in subwatersheds 2 and 3 around Lake Granbury (Figure 1). This is explained low- and medium-intensity housing developments within these subwatersheds. These are popular residential areas because of the lake in the southern portion of the watershed and the close proximity to the Fort Worth metropolitan area in the northeast. The PCF ranking incorporated driving forces of pollutant fate and transport. The subwatersheds with highest potential *E. coli* resulting from pets are ranked using the average PCF over several weighting schemes as 1st, 4th, 8th, and 10th. The next highest subwatersheds have a PCF ranking ranging from 4th to 10th. As noted earlier, subwatershed 26 (Figure 1) is not currently monitored for *E. coli* contamination (Figure 8 and Table 7). Several water quality monitoring stations are located in subwatershed 8, but the data do not indicate significant violations in water quality due to *E. coli* (Figure 8 and Table 7). Again subwatersheds 1 and 3 do indicate high *E. coli* concentrations from 23 to 43% out of all observations (Figure 8 and Table 7).

Point Sources. There are seven WWTP facilities operating within the watershed (Figure 7). The highest *E. coli* loading occurs in subwatershed 8 (Figure 1) on the southeastern edge of the watershed. These facilities contribute large amounts of treated effluents and could impact the environment if improper/inefficient treatment of wastewater were to occur. When localities are considering consolidating OWTSs into municipal sewage systems, the local officials should take into account the amount of pollutants, such as *E. coli* and nutrients, that would be discharged as a direct point source (with virtually zero travel time or attenuation). Currently, no water quality concerns from these facilities have arisen.

Combined Loading from all Sources. The final determination of the most likely sources is a consideration of the total potential load as determined with the SELECT program (one map output that aids in both comparison of sources and total subwatershed loads) and the PCF ranking map output, which considers watershed characteristics and relative subwatershed loads, but can only be used on a source by source basis and not meant for comparison between sources. The highest total potential *E. coli* loads (Figure 4) occur in subwatersheds 14, 26, 30, and 34 (Figure 1). Subwatersheds 30 and 34 have land uses appropriate for cattle and deer. Hence, it can be concluded that major *E. coli* contributors in these subwatersheds are cattle and deer. Subwatershed 14 is ranked as the third highest area of concern based on the PCF due to the combined effects of potentially higher loading from cattle and a potentially high load from deer and OWTSs. Subwatershed 26 has the greatest likelihood to contribute to bacterial contamination in waterbodies based on the PCF ranking. This particular subwatershed is characterized by grasslands, pastures, and forests in the majority of the region and with significant development on the northern edge. It can be concluded that the potential *E. coli* loading in this subwatershed with diverse land use is a result of combined contributions from cattle, deer, and pets.

The SELECT results including the PCF analysis indicate that across the entire watershed cattle is the largest contributor to *E. coli* loading to streams followed by deer, pets, OWTSs, and WWTPs (Figure 5b). Comparing the SELECT results with actual *E. coli* concentrations measured at water quality monitoring stations near the lake (Figure 8) indicates that malfunctioning OWTSs are potentially a major concern followed by pets. Currently, bacterial water quality is not monitored where SELECT predicts high-potential *E. coli* loads in the Lake Granbury watershed (Figures 4 and 8).

Versatility of Spatially Explicit Load Enrichment Calculation Tool

When potential *E. coli* loads simulated by SELECT are combined with the PCF module, decision makers can identify *E. coli* sources and areas of potential concern in a watershed. This will ultimately help decision makers choose cost effective BMPs to alleviate contamination issues in an impaired watershed. Once BMPs have been chosen, PCF analysis can be performed to determine the spatially explicit locations to implement source-specific BMPs. The PCF results can also be used to determine the locations for further water quality monitoring. Ideally, these locations should be in potential *E. coli* contributing areas and in areas where BMPs have been implemented to measure the success of the *E. coli* load reductions.

The current approach for many WPPs target load reductions from all sources applied uniformly across the watershed. It is evident from the geographical representation provided by SELECT that this is not practical and enforcement of pollutant reduction should only be in areas of greatest concern and should address each source separately. This will save both time and money by effectively developing BMPs that will preserve vital water resources.

It is very possible that the water quality data will indicate a different scenario than the simulated loads using SELECT as this is a potential load assessment. In this case, a more thorough investigation is imperative. It will be necessary to apply a more advanced hydrologic simulation model to route the pollutants through the watershed to more accurately predict pollutant loads reaching the waterbodies. The use of a transport model simulation could also be used to calibrate SELECT input parameters by comparing to water quality data. Unfortunately, unless species specific data is gathered, this calibration would be limited to scaling up and down total loading across the watershed. The better the input data available for assumptions when estimating loads, the more reliable the SELECT results will be. For example, if the WWTPs are not treating effluent properly or are discharging pollutants at higher than the permitted concentration, this actual amount should be determined through sampling and used in SELECT simulations.

Bacteria loading in a watershed can have seasonal variability due to migratory patterns of wildlife and grazing rotations for livestock. SELECT can easily simulate this temporal variability of *E. coli* with appropriate assumptions and modifying input data. The simulated potential *E. coli* loads can be fed into a comprehensive water quality model to predict *E. coli* loads at different spatial scales. This is important because some hydrologic simulation models use subwatersheds, whereas others such as SWAT use hydrologic response units (HRUs). If temporal information is available on *E. coli* sources, SELECT can generate spatial as well as temporal *E.coli* loads in a watershed based on chosen time scale. *E. coli* loads calculated using SELECT for each subwatershed or HRU can serve as input for SWAT to simulate *E. coli* concentrations occurring in a waterbody. SELECT could also be linked with SWAT to identify the areas of most concern, so that a SWAT user can focus on those areas instead of the entire watershed to simulate the effects of BMPs. This tool could be integrated into a wide range of simulation models. The SELECT approach can be modified to determine potential loads of other contaminants such as nutrients by using appropriate source inputs and loading rates.

The benefit of the automated SELECT is its ability to generate various scenarios to simulate potential contaminant loads with minimizing the errors inherent in manual approaches. The automated approach takes about 5 minutes to incorporate input files and parameters and 20 minutes to do the simulations for a watershed of 1, 100 km^2 evaluating five contaminant sources. Prior to the initial application some preprocessing of data is necessary, and then subsequent simulations are simple and fast.

CONCLUSIONS

The SELECT was developed and automated to characterize the production of pathogens from various pollutant sources across a watershed. SELECT was applied to the Lake Granbury watershed in Texas. On the basis of simulation results for Lake Granbury, BMPs are recommended to decrease *E. coli* loads from pets and OWTSs near the lake. Further investigation using watershed-scale water quality models, such as SWAT or HSPF is needed to determine the influence of various *E. coli* sources across the watershed. Travel time and decay rates from the subwatersheds with high-potential loading should be determined to characterize the amount of *E. coli* reaching the waterbodies after a rainfall event. In addition, incorporating source-related travel distances to waterbodies in the PCF module rather than subwatershed flow lengths would likely improve this tool. It is also recommended that water quality monitoring should be carried out in northern and western portions of the Lake Granbury watershed to monitor *E. coli* concentrations in the watershed. This will ultimately help in protecting Lake Granbury from contamination due to pathogenic bacteria.

For the Lake Granbury watershed most of the high *E. coli* concentrations were observed on days of or immediately preceding significant precipitation on the day of measurement (BRA, 2008; NCDC, 2008). There are a few incidences where high *E. coli* concentrations were measured at water quality monitoring locations with no recent precipitation events (BRA, 2008; NCDC, 2008). This indicates that point source discharges either from WWTPs or illicit direct discharges were causing *E. coli* contamination on these days.

SELECT is a user-friendly tool to conduct spatial analysis under different land use scenarios. In addition to this, maps and tables resulting from SELECT can be used for technical and educational communication. This approach proves the need to evaluate each contaminant source separately to effectively allocate site specific BMPs and serves as a powerful screening tool for determining areas where detailed investigation is merited.

ACKNOWLEDGMENTS

We would like to thank Dr. Lesikar for his inputs and suggestions throughout the course of this research. We also want to extend our gratitude to the Texas Water Resources Institute and the United States Geological Survey for their gracious funding through the T. W. Mills Fellowship and the USGS Water Resources Grant. In addition, we would like to thank the Spatial Sciences Laboratory at Texas A&M University for providing data support. The authors would like to acknowledge the three anonymous reviewers and Associate Editor, Mr. R. Alexander, who provided an excellent review and feedback which significantly improved the final manuscript.

LITERATURE CITED

AVMA, 2002. U.S. Pet Ownership and Demographics Source Book. Center for Information Management, American Veterinary Association, Schaumburg, Illinois.

Benham, B.L., C. Baffaut, R.W. Zeckoski, K.R. Mankin, Y.A. Pachepsky, A.M. Sadeghi, K.M. Brannan, M.L. Soupir, and M.J. Habersack, 2006. Modeling Bacteria Fate and Transport in Watersheds to Support TMDLs. Transactions of American Society of Agricultural and Biological Engineers 49(4):987-1002.

Bora, D.K. and M. Bera, 2004. Watershed-Scale Hydrologic and Nonpoint-Source Pollution Models: Review of Applications. Transactions of the American Society of Agricultural Engineers 47(3):789-803.

BRA, 2008. Brazos River Watershed Information Center. Brazos River Authority, Waco, Texas. http://crpdata.brazos.org/, *accessed* April 2008.

Chapman, G., J. Gray, R. Irvine, and M. Barry, 2004. Using Soil Landscape Mapping for On-Site Sewage Risk Assessment. SuperSoil 2004: 3rd Australian New Zealand Soils Conference, University of Sydney, Australia. http://www.regional.org.au/au/asssi/, *accessed* April 2008.

Chen, C.W., J. Herr, L. Ziemelis, R.A. Goldstein, and L. Olmsted, 1999. Decision Support System for Total Maximum Daily Load. Journal of Environmental Engineering, American Society of Civil Engineers 125(7):653-659.

Clark, M.K., W.S. Heigis, B.F. Douglas, and J.B. Hoover, 2001. Decentralized Wastewater Management Needs Assessment: A Small Community's Approach to On-Site Treatment. Proceedings of the Ninth National Symposium on Individual and Small Community Sewage Systems. American Society of Agricultural Engineers, Warren, Vermont, pp. 427-434.

Deepti, P., R. Karthikeyan, and M. Babar-Sebens, 2009. Predicting the Fate and Transport of *E. coli* in Two Texas River Basins. Journal of the American Water Resources Association 45(4): 928-944.

Doyle, M. and M. Erikson, 2006. Closing the Door on the Fecal *Coliform* Assay. Microbe 1(4):162-163.

Espey Consultants Inc., 2007. Lake Granbury Water Quality Modeling Project PHASE 1 Draft Report — Data Trend Analysis, Modeling Overview and Recommendations. Prepared for Brazos River Authority. Espey Consultants, Austin, Texas.

Ferguson, C., A.M. Roda Husman, N. Altavilla, D. Deere, and N. Ashbolt, 2003. Fate and Transport of Surface Water Pathogens in Watersheds. Critical Reviews in Environmental Science and Technology 33(3):299-361.

Haan, C., B. Barfield, and J. Hayes, 1994. Design Hydrology and Sedimentology for Small Catchments. Academic Press, Inc., San Diego, California.

Jamieson, R., R. Gordon, D. Joy, and H. Lee, 2004. Assessing Microbial Pollution of Rural Surface Waters: A Review of Current Watershed Scale Modeling Approaches. Agricultural Water Management 70:1-17.

Kenway, S. and R. Irvine, 2001. Sewage Pollution Risk Assessment for Environmental Health. Conference Proceedings — Environmental Health Conference 2001, 11-12 September, Bathurst, Australia, pp. 1-8.

Lockwood, M., 2005. White-tailed Deer Population Trends. Federal Aid in Fish and Wildlife Restoration. Project W-127-R-14. Texas Parks and Wildlife Department, Austin, Texas.

Moyer, D. and K.E. Hyer, 2003. Use of the Hydrologic Simulation Program — FORTRAN and Bacterial Source Tracking for Development of the Fecal *Coliform* Total Maximum Daily Load (TMDL) for Accotink Creek, Fairfax County, Virginia. Water Resources Investigation Report 03-4160. United States Geological Survey and Virginia Department of Conservation and Recreation, Richmond, Virginia.

Munafo, M., G. Cecchi, F. Baiocco, and L. Mancini, 2005. River Pollution from Non-Point Sources: A New Simplified Method of Assessment. Journal of Environmental Management 77:93-98.

NCDC (National Climate Data Center), 2008. . U.S. Department of Commerce — NOAA, Washington, D.C. http://www.ncdc.noaa.gov/oa/mpp/freedata.html, *accessed* April 2008.

Paul, S., R. Srinivasan, J. Sanabria, P. Haan, S. Mukhtar, and K. Neimann, 2006. Groupwise Modeling Study of Bacterially Impaired Watersheds in Texas: Clustering Analysis. Journal of the American Water Resource Association 42(4):1017-1031.

Santhi, C., J. Arnold, J. Williams, L. Huack, and W. Dugas, 2001. Application of a Watershed Model to Evaluate the Management Effects on Point and Nonpoint Source Pollution. Transactions of American Society of Agricultural Engineers 44(6):1559-1570.

SCS, 1986. Urban Hydrology for Small Watersheds, Technical Release 55. U.S. Dept. of Agriculture, Soil Conservation Service, Washington, D.C., Chapter 2, pp. 2-5.

SWAT (Soil and Water Assessment Tool), 2005. Texas A&M University, College Station, Texas and USDA Agricultural Research Service: Grassland, Soil and Water Research Laboratory, Temple, Texas. http://www.brc.tamus.edu/swat/soft.html, *accessed* March 2008.

TCEQ (Texas Commission on Environmental Quality), 2008a. Regulatory/Administrative Boundaries. Texas Commission on Environmental Quality, Austin, Texas. http://www.tceq.state.tx.us/gis/boundary.html, *accessed* March 2008.

TCEQ (Texas Commission on Environmental Quality), 2008b. Site Layers. Texas Commission on Environmental Quality, Austin, Texas. http://www.tceq.state.tx.us/gis/sites.html, *accessed* March 2008.

Teague, A., R. Karthikeyan, M. Babar-Sebens, R. Srinivasan, and R. Persyn, 2009. Spatially Explicit Load Enrichment Calculation

Tool to Identify *E. coli* Sources in Watersheds. Transactions of ASABE 52(4):1109-1120.

USCB, 2000. Census 2000 TIGER/Line® Files. U.S. Census Bureau, Washington, D.C. http://www.census.gov/geo/www/tiger/index.html, *accessed* October 2007.

USDA-NRCS, 2004. Soil Survey Geographic (SSURGO) Database. United States Department of Agriculture National Resource Conservation Service, Washington, D.C. http://soils.usda.gov/survey/geography/ssurgo/, *accessed* December 19, 2006.

USDA-NASS, 2002. National Agricultural Statistics Service (NASS) Census of Agriculture: Texas. United States Department of Agriculture National Agricultural Statistics Service, Washington, D.C. http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/County_Profiles/Texas/index.asp, *accessed* December 2006.

USEPA, 2001. Protocol for Developing Pathogen TMDLs: Source Assessment (First Edition). EPA841-R-00-002. USEPA Office of Water, Washington, D.C, Chapter 5, pp. 1-18.

USEPA, 2006. Envirofacts Data Warehouse. U.S. Environmental Protection Agency, Washington, D.C. *accessed* February 2008.

USEPA, 2008. Total Maximum Daily Loads: National Section 303(d) List Fact Sheet: Top 100 Impairments. U.S. Environmental Protection Agency, Washington, D.C. http://oaspub.epa.gov/waters/national_rept.control#TOP_IMP, *accessed* February 2008.

Zeckoski, R.W., B.L. Benham, S.B. Shah, M.L. Wolfe, K.M. Brannan, M. Al-Smadi, T.A. Dillaha, S. Mostaghimi, and C.D. Heatwole, 2005. BSLC: A Tool for Bacteria Source Characterization for Watershed Management. Applied Engineering in Agriculture ASAE 21(5):879-889.