Environmental Modelling & Software 99 (2018) 126–146

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Capturing microbial sources distributed in a mixed-use watershed within an integrated environmental modeling workflow

Gene Whelan [a, *], Keewook Kim [b, 1], Rajbir Parmar [a], Gerard F. Laniak [a], Kurt Wolfe [a], Michael Galvin [a], Marirosa Molina [a], Yakov A. Pachepsky [c], Paul Duda [d], Richard Zepp [a], Lourdes Prieto [a], Julie L. Kinzelman [e], Gregory T. Kleinheinz [f], Mark A. Borchardt [g]

[a] *U.S. Environmental Protection Agency, Office of Research and Development, Athens, GA, USA*
[b] *Idaho Falls Center for Higher Education, University of Idaho, Idaho Falls, ID, USA*
[c] *U.S. Department of Agriculture, Agricultural Research Service, Beltsville, MD, USA*
[d] *AQUA TERRA Consultants, A Division of RESPEC, INC, Decatur, GA, USA*
[e] *Public Health Department, Racine, WI, USA*
[f] *Department of Engineering Technology, University of Wisconsin Oshkosh, Oshkosh, WI, USA*
[g] *U.S. Department of Agriculture, Agricultural Research Service, Marshfield, WI, USA*

ARTICLE INFO

Article history:
Received 28 November 2016
Received in revised form
17 June 2017
Accepted 8 August 2017

Keywords:
Integrated environmental modeling
QMRA
Risk assessment
Pathogens
Manure
Watershed modeling

ABSTRACT

Many watershed models simulate overland and instream microbial fate and transport, but few provide loading rates on land surfaces and point sources to the waterbody network. This paper describes the underlying equations for microbial loading rates associated with 1) land-applied manure on undeveloped areas from domestic animals; 2) direct shedding (excretion) on undeveloped lands by domestic animals and wildlife; 3) urban or engineered areas; and 4) point sources that directly discharge to streams from septic systems and shedding by domestic animals. A microbial source module, which houses these formulations, is part of a workflow containing multiple models and databases that form a loosely configured modeling infrastructure which supports watershed-scale microbial source-to-receptor modeling by focusing on animal- and human-impacted catchments. A hypothetical application – accessing, retrieving, and using real-world data – demonstrates how the infrastructure can automate many of the manual steps associated with a standard watershed assessment, culminating in calibrated flow and microbial densities at the watershed's pour point.

Published by Elsevier Ltd.

1. Introduction

The United States Environmental Protection Agency (EPA) is interested in characterizing, managing, and minimizing the risks of human exposure to pathogens in water resources impacted by effluents and runoff from both agricultural activities and built infrastructure. EPA (2016a) indicates that 52.8% of the assessed river and stream miles are impaired, with pathogens being the main cause followed by sediment contamination and nutrients. The designation "pathogen" is used in the broadest sense based upon detection of fecal indicator bacteria (FIB), *Escherichia coli* (*E. coli*), and fecal coliforms. Monitoring for the presence of pathogens in manure- and sewage-contaminated waters is extremely challenging, as pathogen concentrations in water samples are often low. Such low concentrations make detection unfeasible, unless large volumes of water are analyzed. Most monitoring approaches and microbial water quality regulations are based on indicator bacteria, since they are easier to sample and quantify (EPA, 2012, 2015), although good correlations between indicators and pathogens may be suspect. For example, Haack and Duris (2013) note that "… there is a widely acknowledged variable relationship between FIB and pathogen concentrations (Field and Samadpour, 2007; Savichtcheva et al., 2007)." Therefore, states might avail themselves of water quality criteria, if they can demonstrate an equivalent level of public health protection with higher indicator concentrations.

Agriculture is one of the most likely causes of pollution, affecting almost 13% of the total river miles assessed, since applying manure

---

* Corresponding author. U.S. Environmental Protection Agency, Office of Research and Development, Athens, GA 30605, USA.
E-mail address: Whelan.Gene@epa.gov (G. Whelan).
[1] Current address: Busan Development Institute, Busan, South Korea.

http://dx.doi.org/10.1016/j.envsoft.2017.08.002
1364-8152/Published by Elsevier Ltd.

for crop nutrition and production and animal shedding due to grazing are common practices. Manure applications may carry environmental contaminants such as pathogens, organic chemical residues and heavy metals (Edwards and Daniel, 1992). These contaminants adversely affect water quality mainly due to runoff-producing rainfall events. Among the various animal fecal sources, poultry are responsible for 44% of the total feces production in the United States, followed by cattle (31%) and swine (24%) (Kellog et al., 2000). In comparison, humans contribute only a small fraction (0.7%) on an equal weight basis; however, human sewage/wastewater is generally thought to constitute a much higher risk to public health due to the likelihood of viral pathogen presence (Soller et al., 2010; Schoen et al., 2011; Dufour, 1984).

Models can play a role in assessing the distribution of microbes in a mixed-use watershed and the potential risks associated with both measured and predicted indicator concentrations (i.e., degree to which concentrations indicate threats to public health under varying circumstances). Assessment of potential risks is critical in determining the appropriateness of waivers to criteria and concentration standards based on site-specific environmental settings and source conditions. Site surveys, coupled with modeling tools, are a basic way to identify sources, characterizing them, associate a level of infectivity with the source, and assess its level of impact at the point of exposure.

A Quantitative Microbial Risk Assessment (QMRA) is a source-to-receptor modeling approach that integrates disparate data – such as fate/transport, exposure, and human health effect relationships – to characterize the distribution of indicator and pathogenic microbes within a watershed, and the potential health impacts/risks from exposure to pathogenic microorganisms (Soller et al., 2010; Whelan et al., 2014a, 2014b; Haas et al., 1999; Hunter et al., 2003). As Whelan et al. (2014b) note, a QMRA's conceptual design fits well within an integrated, multi-disciplinary modeling perspective which describes the problem statement; data access retrieval and processing [e.g., D4EM (EPA, 2013a; Wolfe et al., 2007)]; software frameworks for integrating models and databases [e.g., FRAMES (Whelan et al., 2014b; Johnston et al., 2011)]; infrastructures for performing sensitivity, variability, and uncertainty analyses [e.g., SuperMUSE (Babendreier and Castleton, 2005)]; and risk quantification. Coupling modeling results with epidemiology studies allows policy-related issues (EPA, 2010; EPA and USDA, 2012; for example) to be explored. An important aspect of the integrated environmental modeling (IEM) (Laniak et al., 2013) microbial workflow is its ability to define spatial and temporal microbial loadings from human and animal sources within a mixed-use watershed. Multiple software tools have been developed to estimate microbial source loadings to a watershed, such as MWASTE, COLI, SEDMOD, modifications to SWAT, SELECT, BIT, and BSLC.

Moore et al. (1989) developed MWASTE to simulate waste generation and calculate bacterial concentrations in runoff from the land-applied waste of various animals and management techniques. MWASTE only considers animal-borne bacteria and allows only one animal per execution, so multiple runs are required for the consideration of different animal species.

Walker et al. (1990) developed the COLI model to predict bacteria concentration in runoff resulting from a single storm occurring immediately after land application of manure. It uses a Monte Carlo simulation to combine a deterministic relationship with rainfall and temperature variations and calculates maximum and minimum bacteria concentration in runoff.

Fraser et al. (1996) developed a GIS-based Spatially Explicit Delivery Model (SEDMOD) that estimates spatially-distributed delivery ratios for eroded soil and associated nonpoint source pollutants. The model predicts fecal coliform loading in rivers and calculates pollutant loadings in streams by multiplying livestock fecal coliform output and a delivery ratio, estimated for each watershed cell, to predict the proportion of eroded sediment (or other non-point source pollutant) transported from the cell to the stream channel.

Parajuli (2007) manually estimated fecal bacterial loading – considering different sources such as livestock (manure application, grazing), human (septic), and wildlife – for the SWAT bacteria sub-model. Guber et al. (2016) followed this up with a limited effort that integrated infection and recovery of white-tailed deer and cattle into the watershed model SWAT. It predicted pathogen transmission between livestock and deer by considering seasonal changes in deer population, habitat, and foliage consumption; ingestion of pathogens with water, foliage, and grooming soiled hide by deer and grazing cattle; infection and recovery of deer and co-grazing cattle; pathogen shedding by infected animals; survival of pathogens in manure; and kinetic release of pathogens from applied manure and fecal material.

Teague et al. (2009) developed the Spatially Explicit Load Enrichment Calculation Tool (SELECT) to identify potential *E. coli* sources in Plum Creek Watershed in Texas; SELECT is a grid-based load assessment tool that considers multiple point and non-point sources (wastewater treatment plant, livestock, pets, wildlife, septic, urban). Riebschleager et al. (2012) automated SELECT within ArcGIS and added the Pollutant Connectivity Factor component which is based on potential pollutant loading, runoff potential, and travel distance. SELECT has been used to identify *E. coli* (Teague et al., 2009; McKee et al., 2011; Riebschleager et al., 2012; McFarland and Adams, 2014; Borel et al., 2015) and enterococci (Borel et al., 2015) sources in multiple watersheds in Texas.

The Bacterial Indicator Tool (BIT) estimates microbial loading from domestic animals, wildlife, and human activities to a mixed-use watershed (EPA, 2000). It accounts for land-application of manure and direct shedding from certain domestic animals to pasture and cropland, and from wildlife to cropland, pasture, and forest. It also estimates point source loadings from septic system failures and direct shedding to the stream from certain domestic animals. Finally, it accounts for loading in urban (built-up) areas such as residential, commercial, transportation, etc. BIT uses Microsoft Excel for calculations and considers only 10 subwatersheds when distributing loads. Land-applied loading rates are adjusted for die-off. All loadings vary monthly, except for those from wildlife, in urban areas, and from septic systems which use constant loading rates to the stream based on the fraction of septic systems that fail. Urbanized areas include categories such as commercial, mixed-urban or built-up, residential, and roadways. Loading rates to urbanized areas are supplied by the user, although default values are suggested. Stormwater runoff through drainage pipes and combined and non-combined sewer systems are not accounted for.

In a similar manner to BIT, the Bacterial Source Load Calculator (BSLC) was designed to organize and process bacterial inputs for a Total Maximum Daily Load (TMDL) bacterial impairment analysis (Zeckoski et al., 2005). BSLC calculates bacterial loads based on animal numbers and default values for manure and bacterial production rates, accounting for die-off and the fraction of domestic animal confinement. It uses externally-generated, user-supplied inputs of watershed delineations, and land-use distribution, as well as domestic animal, wildlife, and human population estimates to suggest monthly land-based and hourly stream-based bacterial loadings. Neither BIT nor BSLC offer software that supports data collection to meet model input requirements, although their documentation suggests some default values.

Prior to allocating microbial sources within a watershed, the watershed must first be delineated into subwatersheds which are the smallest modeling units. To do so, many models require users to manually and externally delineate a watershed, then manually assign environmental characteristics, animal numbers and types, farming practices, and human activities to each subwatershed. This can be a daunting task, especially if the user re-delineates the watershed. Because the delineation pattern determines size and location of subwatersheds, it has a significant impact on distribution and magnitude of microbial loading rates within them; hence, it is desirable to have an automated process to delineate a watershed; populate its subwatersheds with environmental characteristics [land-use types, waterbody network, slope, soil type, meteorological (MET) data, etc.]; and overlay sources of microbial contamination so appropriate loading rates can be easily computed on land and in stream.

The work reported here describes the expansion and modification of BIT, developing a new Microbial Source Module (MSM) (Wolfe et al., 2016; Whelan et al., 2015a). Additionally, its use and implementation was demonstrated as a component of an IEM workflow. The workflow automates the manual processes that perform QMRAs on mixed-use watersheds anywhere in the United States by determining microbial sources and estimates of microbial loadings to land and streams. The mathematical formulations of MSM and its context within an IEM workflow are described here.

2. Materials and methods

The MSM organizes, analyzes, and supplies data that calculates microbial loading rates within subwatersheds, the smallest spatial units for data that it consumes and produces. MSM correlates sources to cropland, pasture, forest, and urbanized/mixed-use land-use types for each subwatershed. Microbial sources include numbers and locations of domestic agricultural animals (dairy and beef cattle, swine, poultry, etc.) and wildlife (deer, duck, raccoon, etc.), with estimated shedding rates; manure application rates where manure is directly incorporated into cropland's and pasture's soil; and loading rates due to urbanized/mixed-use activities (commercial, transportation, etc.). Manure contains microbes and the monthly maximum microbial storage and accumulation rates on the land surface, adjusted for die-off, are computed over a season to represent the source for overland fate and transport to instream locations. Monthly point source microbial loadings to instream locations are also determined for septic systems and instream shedding by cattle. The type of septic system (e.g., gravity, pressure distribution, sand filter, or mound) is not differentiated in the model. Flow, microbial, and chemical loadings also originate from point sources such as Publicly Owned Treatment Works/Wastewater Treatment Plants (POTWs/WWTPs); although they are not simulated, their discharge time series can be accounted for as direct input to the watershed model. The user interface that externally supports MSM automatically formats the watershed input file, so when the user registers the POTWs/WWTPs time series, it is seamlessly incorporated into the input file. The MSM has been seamlessly linked with a user interface, based on Data for Environmental Modeling (D4EM) and Site Data Manager Project Builder (SDMProjectBuilder or SDMPB). These three components are part of an IEM workflow also containing the Hydrologic Simulation Program – FORTRAN (HSPF) watershed model (Bicknell et al., 1997), Better Assessment Science Integrating point and Nonpoint Sources (BASINS) modeling infrastructure (EPA, 2001a), and Parameter ESTimation and Uncertainty Analysis (PEST) inverse model (Doherty, 2005).

2.1. Assumptions and constraints

Assumptions and constraints associated with the MSM are presented in Table 1, and correlations between manure application, land-use type, domestic animal, and wildlife are summarized in Table 2. An index glossary that correlates subscripts used in the mathematical formulas is provided in Table 3.

2.2. Land application of domestic animal manure

As indicated in Table 2, cropland and pasture are land-use types that receive land-applied domestic animal manure, and whose mathematical formulations are described in the following sections.

2.2.1. Domestic animal waste available for land application

The fraction of annual manure application available for runoff each month by domestic animals, based on the monthly fraction applied and incorporated into the soil, is computed as follows (EPA, 2013b, 2013c):

$$\begin{aligned}&FractionManureAvailableRunoff_{m,q}\\&\quad = (Application_{m,q})(AnimalFractionAvailable_m)\end{aligned} \tag{1}$$

in which

$$\begin{aligned}AnimalFractionAvailable_m &= 1 - \frac{(ManureIncorporatedIntoSoil_m)}{2} \quad \mathrm{m} = 1,\ 2,\ 3,\ \text{or}\ 5\\ &= 1 - \frac{(ManureIncorporatedIntoSoil_m)}{3} \quad \mathrm{m} = 4\end{aligned} \tag{2}$$

where $FractionManureAvailableRunoff_{m,q}$ is the fraction of annual manure application available for runoff by month (q) by domestic animal (m) [equivalent to the ratio of microbial cells available for runoff each month to cells available for runoff per year] (Ratio); $Application_{m,q}$ is the fraction of annual manure applied each month (q) by domestic animal (m) [equivalent to the ratio of cells applied each month to cells applied per year] (Ratio); $ManureIncorporatedIntoSoil_m$ is the fraction of applied manure incorporated into the soil by domestic animal (m) (Ratio); and $AnimalFractionAvailable_m$ is the fraction of domestic animal (m) manure available for runoff (Ratio).

2.2.2. Land application of manures from domestic animals

The monthly microbial loading rate from land application of domestic animal manure associated with each subwatershed by land-use type is equal to:

Table 1
Assumptions and constraints associated with the Microbial Source Module (after Whelan et al., 2015a; Wolfe et al., 2016).

1. The MSM considers only one microbe at a time and must be individually executed.
2. Overland microbial loading rates, accounting for die-off, are computed for each subwatershed by land-use type on a monthly basis.
3. The MSM considers microbial loadings from sources correlated to four land-use types for each subwatershed, where it is the smallest area associated with watershed modeling: 1) Cropland: Land application of some domestic animal waste (Beef Cattle, Dairy Cow, Swine, and/or Poultry) and Wildlife shedding; 2) Pasture: Some domestic animal grazing with shedding (Beef Cattle, Horse, Sheep, and/or Other), Land application of some domestic animal waste (Beef Cattle, Dairy Cow, and/or Horse), and Wildlife shedding; 3) Forest: Wildlife shedding; and 4) Built: Urban-related releases: Commercial and Services, Residential, Mixed Urban, Transportation, and Communication, Utilities.
4. The MSM considers instream beef cattle shedding, where loading rates are identified with each subwatershed.
5. The MSM currently assumes that manure loadings from land application and shedding are computed monthly and represent a typical year.
6. The land-use types associated with the National Land Cover Database (NLCD) are consolidated into Cropland, Pastureland, Forest, and Urbanized, providing a more manageable modeling set when land use is the index, since supporting data for finer granularity are not available.
7. Urbanized land is subdivided into Commercial and Services; Mixed Urban or Built-Up; Residential; and Transportation, Communications, and Utilities. A single, weighted Urbanized loading rate is quantified for each subwatershed (all months) based on all individual Urbanized land uses present. Each Urbanized category considers a weighted combination of the following five attributes: Commercial, Single-family low density, Single-family high density, Multi-family Residential, and Road.
   a. Commercial and Services: Commercial
   b. Mixed Urban or Built-up: Average microbial accumulation rates for Road, Commercial, Single-family low density, Single-family high density, and Multi-family residential
   c. Residential: Average microbial accumulation rates for Single-family low density, Single-family high density, and Multi-family residential
   d. Transportation, Communications, and Utilities
8. Fecal shedding from animals is used for microbial loading estimates to all land-use types except Urbanized.
9. Manures from Swine and Poultry are assumed to be collected and applied to Cropland.
10. Beef Cattle/Dairy Cow manure is assumed to be applied only to Cropland and Pastureland by the same method.
11. Dairy Cows are only kept in feedlots; therefore, all of their waste is used for manure application, divided equally between Cropland and Pastureland.
12. Beef Cattle are kept in feedlots or allowed to graze. During grazing, a specified percentage of cattle also have direct access to streams; therefore, Beef Cattle waste is either applied as manure to Cropland and Pastureland, or contributes directly to Pasture (shedding) or Streams (shedding). Direct contribution of microbes from Beef Cattle to a stream through shedding is thus represented as a monthly point source. Dairy Cows are not allowed to graze and, therefore, do not have access to streams.
13. Horse manure not deposited in Pastureland during grazing is assumed to be collected and applied to Pastureland.
14. Manures from Beef Cattle, Horses, Sheep, and Other domestic animals are assumed to contribute to Pastureland in proportion to time spent grazing. Sheep and Other domestic animal manures not deposited to Pastureland during grazing are assumed to be collected and treated or transported out of the watershed.
15. Domestic animal designations are designed as placeholders to differentiate grazing and non-grazing animals by land-use type and manure application (land-applied versus direct shedding). For example, if Dairy Cows graze and/or shed directly to the stream, then they can be designated as Beef Cattle.
16. Wildlife densities are provided for all land uses except Built-up and assumed to be the same in all subwatersheds. The wildlife population is the only microbial contributor considered to Forest.
17. Fraction of annual domestic animal manure application available for runoff each month (EPA, 2013b, 2013c)
    = [Fraction of manure applied] * {1 - [Fraction of manure incorporated]/3} for poultry
    = [Fraction of manure applied] * {1 - [Fraction of manure incorporated]/2} for other domestic animals (dairy cow, beef cattle, swine, and horse)
18. One input time series for direct input to streams is allowed per subwatershed; multiple septics and instream shedding are each aggregated separately, then combined to provide monthly loadings.

Table 2
Correlation of manure application with land-use type by domestic animal and wildlife (after Wolfe et al., 2016).

Manure Application Correlated to Land Use	Domestic Animals and Wildlife							
	Dairy Cow	Beef Cattle	Swine	Poultry	Horse	Sheep	Other	Wildlife
SHEDDING								
Cropland Grazing/Shedding								x
Pasture Grazing/Shedding		x			x	x	x	x
Forest Shedding								x
In Stream Shedding		x						
LAND APPLIED								
Cropland Application	x	x	x	x				
Pasture Application	x	x			x			

Table 3
Index glossary used in mathematical formulations (after Wolfe et al., 2016).

Index	Description
i	Subwatershed ID
k	Land-use type (1 = Cropland, 2 = Pasture, 3 = Forest, 4 = Urbanized)
m	Domestic Animal [1 = Dairy Cow (DairyCow), 2 = Beef Cattle (BeefCattle), 3 = Swine, 4 = Poultry, 5 = Horse, 6 = Sheep, 7 = Other Agricultural Animal (OtherAgAnimal)]
n	Wildlife (1 = Duck, 2 = Goose, 3 = Deer, 4 = Beaver, 5 = Racoon, 6 = Other Wildlife)
q	Month of the year (January to December)
r	Urbanized category (1 = Commercial and Services; 2 = Mixed Urban or Built-up; 3 = Residential; and 4 = Transportation, Communications and Utilities)
u	Urbanized sub-category (1 = Commercial, 2 = Single-family Low Density, 3 = Single-family High Density, 4 = Multi-family Residential, 5 = Road)

$$\begin{aligned} \mathit{MicrobeRateApply}_{i,k,m,q} &= 0 \quad \text{for } m = 1 \text{ or } 2 \text{ and } k = 3 \text{ or } 4; m = 3 \text{ or } 4 \text{ and } k = 2,\ 3,\ \text{or } 4; \\ &\qquad m = 5 \text{ and } k = 1,\ 3,\ \text{or } 4; m = 6 \text{ or } 7 \text{ and } k = 1,\ 2,\ 3,\ \text{or } 4 \\ &= (\mathit{NumberOfAnimals}_{i,m})(\mathit{MicrobeAnimalProductionRates}_{m}) \\ &\times \left(\mathit{FractionManureAvailableRunoff}_{m,q}\right)(\mathit{ApplyMonths}_{q}) \\ &/(\mathit{ApplyArea}_{i,k}) \end{aligned} \tag{3}$$

in which

$$\begin{aligned} \mathit{ApplyMonths}_{q} &= (365/\mathit{DayInMonth}_{q}) && \text{for } m = 1 \text{ and } k = 1 \text{ or } 2; m = 3 \text{ or } 4 \text{ and } k = 1 \\ &= (365 - \mathit{TotalGrazeDays}_{m})/(\mathit{DayInMonth}_{q}) && \text{for } m = 2 \text{ and } k = 1 \text{ or } 2; m = 5 \text{ and } k = 2 \end{aligned} \tag{4}$$

$$\begin{aligned} \mathit{ApplyArea}_{i,k} &= \mathit{Area}_{T,i} && \text{for } m = 1 \text{ or } 2 \text{ and } k = 1 \text{ or } 2 \\ &= \mathit{Area}_{i,k} && \text{for } m = 3 \text{ or } 4 \text{ and } k = 1; m = 5 \text{ and } k = 2 \end{aligned} \tag{5}$$

$$\mathit{TotalGrazeDays}_{m} = \sum_{q} \mathit{GrazingDays}_{m,q} \tag{6}$$

$$\mathit{Area}_{T,i} = \sum\nolimits_{k=1}^{2} \mathit{Area}_{i,k} \tag{7}$$

where $\mathit{MicrobeRateApply}_{i,k,m,q}$ is the microbial loading rate per area to land-use type (k) from land application of domestic animal (m) manure by month (q) by subwatershed (i) (Cells/Time/Area); $\mathit{NumberOfAnimals}_{i,m}$ is the number of domestic animals (m) by subwatershed (i) (Number of domestic animal); $\mathit{MicrobeAnimalProductionRates}_{m}$ is the microbial production rate shed per domestic animal (m) (Cells/d/domestic animal) [equals the multiple of domestic animal shedding rate of waste in mass of wet weight (ww) per time (Mass/d/domestic animal), and microbial density (concentration) based on mass of waste shed by domestic animal (Cells/Mass)]; $\mathit{ApplyMonths}_{q}$ is the conversion for the number of months per year, weighted by the actual number of days in month (q), when land application of manure might be possible (mo/yr); $\mathit{ApplyArea}_{i,k}$ is the area associated with the land application of manure by subwatershed (i) by land-use type (k); "365" is the conversion constant for days in a year (d/yr); $\mathit{DayInMonth}_{q}$ is the conversion constant for days per month (q) [January = 31, February = 28, …, December = 31, in which the months are indexed by "q" as 1 = January, …, 12 = December] (d/mo); $\mathit{GrazingDays}_{m,q}$ is the number of grazing days by domestic animal (m) per month (q) (d/mo); $\mathit{TotalGrazeDays}_{m}$ is the total number of grazing days per year for domestic animal (m) (d/yr); $\mathit{Area}_{i,k}$ is the subwatershed (i) area by land-use type (k) (Area); and $\mathit{Area}_{T,i}$ is the total summed area for cropland ($k = 1$) and pasture ($k = 2$) by subwatershed (i) (Area).

2.3. Domestic animal and wildlife shedding rates to land surfaces

Land-use types that receive domestic animal and wildlife shedding are captured in Table 2. The monthly microbial loading rates to different land-use types due to shedding from domestic animals while grazing, by subwatershed, is equal to:

$$\begin{aligned} \mathit{MicrobeRateShed}_{i,k,m,q} &= 0 \quad \text{for } m = 1,\ 3,\ \text{or } 4 \text{ and } k = 1,\ 2,\ 3,\ \text{or } 4; \\ &\qquad m = 2,\ 5,\ 6 \text{ or } 7 \text{ and } k = 1,\ 3,\ \text{or } 4 \\ &= (\mathit{NumberOfAnimals}_{i,m})(\mathit{ShedMonths}_{m,q}) \\ &\times(\mathit{MicrobeAnimalProductionRates}_{m})/(\mathit{Area}_{i,k}) \end{aligned} \tag{8}$$

in which

$$\begin{aligned} \mathit{ShedMonths}_{m,q} &= (\mathit{GrazingDays}_{m,q})(1 - \mathit{TimeSpentInStreams}_{m,q})/(\mathit{DayInMonth}_{q}) && \text{for } m = 2 \text{ and } k = 2 \\ &= (\mathit{GrazingDays}_{m,q})/(\mathit{DayInMonth}_{q}) && \text{for } m = 5,\ 6,\ \text{or } 7 \text{ and } k = 2 \end{aligned} \tag{9}$$

where $\mathit{MicrobeRateShed}_{i,k,m,q}$ is the microbial loading rate to land-use type (k) due to grazing of domestic animal (m) by month (q) by subwatershed (i) (Cells/Time/Area); and $\mathit{ShedMonths}_{m,q}$ is the fraction of the month (q) that domestic animal (m) spends grazing/shedding (Ratio); and $\mathit{TimeSpentInStreams}_{m,q}$ is the fraction of the number of grazing days that domestic animal (m) spends in a stream each month (q) (Ratio).

The monthly microbial loading rates to different land-use types due to shedding from wildlife equals

$$\begin{aligned} \mathit{WildLifeMicrobeRateShed}_{k,n} &= \left(\mathit{Density}_{k,n}\right)(\mathit{MicrobeWildlifeProductionRates}_{n}) && \text{for } k = 1,\ 2,\ \text{or } 3 \\ &= 0 && \text{for } k = 4 \end{aligned} \tag{10}$$

where $WildLifeMicrobeRateShed_{k,n}$ is the microbial shedding rate per area by wildlife (n) by land-use type (k) (Cells/Time/Area), $Density_{k,n}$ is the number of wildlife (n) per area by land-use type (k) (Number of wildlife/Area), and $MicrobeWildlifeProductionRates_n$ is the microbial shedding rate per wildlife (n) (Cells/Time/Number of wildlife). The total microbial shedding rate per land-use type per area summed across all wildlife is:

$$WildLifeMicrobeRateShedSum_k = \sum_{n=1}^{6} WildLifeMicrobeRateShed_{k,n} \tag{11}$$

where $WildLifeMicrobeRateShedSum_k$ is the microbial shedding rate per area by land-use type (k) summed across all wildlife (Cells/Time/Area).

2.4. Accumulated microbial loading rates on urbanized areas

Urbanized land-use is divided into four Urbanized categories ($r = 1$ for Commercial and Services; $r = 2$ for Mixed Urban or Built-up; $r = 3$ for Residential; and $r = 4$ for Transportation, Communications and Utilities) which are further divided into Urbanized subcategories ($u = 1$ for Commercial, $u = 2$ for Single-family Low Density, $u = 3$ for Single-family High Density, $u = 4$ for Multi-family Residential, and $u = 5$ for Road). Single-family low density is a single-detached dwelling, single-family residence, or separate house that is a free-standing residential building (Wikipedia, 2015a). Single-family high density is a suite of smaller-scale single-family dwellings, representing a more compact single-family residential development (13–40 units/ac) (Garnett, 2012). Multi-family residential is a unit with multiple separate housing units for residential inhabitants contained within one building, or several buildings within one complex such as an apartment or condominium (Wikipedia, 2015b). Accumulation rates in median microbial cells per Urbanized land-use type (r) per area per time, indexed by the Urbanized subcategories (u), are computed as follows:

$$\begin{aligned} BuiltUpRate_{k,r} &= SubUrbanizedBuiltUpRate_u && \text{for } k = 4,\ r = 1,\ u = 1 \\ &= \left[\sum_{u=1}^{5} SubUrbanizedBuiltUpRate_u\right] \Big/ 5 && \text{for } k = 4,\ r = 2 \\ &= \left[\sum_{u=2}^{4} SubUrbanizedBuiltUpRate_u\right] \Big/ 3 && \text{for } k = 4,\ r = 3 \\ &= SubUrbanizedBuiltUpRate_u && \text{for } k = 4,\ r = 4,\ u = 5 \end{aligned} \tag{12}$$

where $BuiltUpRate_{k,r}$ is the accumulation rate in median microbial cells per Urbanized land-use type ($k = 4$) per area per time, indexed by the Urbanized category (r) by Urbanized sub-category (u) (Cells/Time/Area); and $SubUrbanizedBuiltUpRate_u$ is the general microbial loading rate by Urbanized sub-category (u) (Cells/Time/Area). Accumulated microbial loading rate associated with the Urbanized land-use type per subwatershed, weighted by the areas associated with the four Urbanized categories for all months (i.e., applicable throughout the year), is computed as follows:

$$\begin{aligned} AccumBuiltupRate_{i,k} &= 0 && \text{for } k = 1,\ 2,\ \text{or } 3 \\ &= \sum_{r=1}^{4} \left[\left(AreaFraction_{i,k,r}\right) \left(BuiltUpRate_{k,r}\right) \right] && \text{for } k = 4 \end{aligned} \tag{13}$$

where $AccumBuiltupRate_{i,k}$ is the accumulated microbial loading rate associated by land-use type (k) by subwatershed (i), weighted by areas associated with four Urbanized categories (r) for all months (i.e., throughout the year) (Cells/Time/Area); and $AreaFraction_{i,k,r}$ is the fraction of land-use type (k), indexed to the four subcategories of Urbanized (r) by Subwatershed (i).

2.5. Accumulated overland microbial loading rates to land surfaces, and maximum microbial storage adjusted for removal

Land-use types that receive domestic animal and wildlife shedding are captured in Table 2 and described in the following sections.

2.5.1. Accumulated overland microbial loading rates to land surfaces

The overland microbial loading rates to land surfaces, accumulated for shedding and land application, are computed by subwatershed, by month, across all domestic animals, and wildlife are computed for the different land-use types as follows:

$$\begin{aligned} AccumulationRateMonth_{i,k,q} &= WildLifeMicrobeRateShedSum_k + \sum_{m=1}^{4} MicrobeRateApply_{i,k,m,q} && \text{for } k=1 \\ &= WildLifeMicrobeRateShedSum_k + \sum_{m=1,2,5} MicrobeRateApply_{i,k,m,q} + \sum_{m=2,5,6,7} MicrobeRateShed_{i,k,m,q} && \text{for } k=2 \\ &= WildLifeMicrobeRateShedSum_k && \text{for } k=3 \\ &= AccumBuiltUpRate_{i,k} && \text{for } k=4 \end{aligned} \tag{14}$$

where $AccumulationRateMonth_{i,k,q}$ is the microbial loading rate by subwatershed (i) by month (q), across all domestic animals (m) and wildlife (n) for each land-use type (k) (Cells/Time/Area).

2.5.2. Maximum microbial storage adjusted for removal

The maximum microbial storage accumulation on the land surface is based on a formulation associated with HSPF. Removal from overland surfaces is simulated as a function of the input accumulation rate and maximum storage of microbes which represents accumulation without removal. The unit removal rate of stored microbes (e.g., Cells removed per day) represents processes such as die-off and wind erosion (Bicknell et al., 1997) and is computed as the microbial accumulation rate divided by the maximum microbial storage accumulation (storage limit):

$$\begin{aligned} StorageLimitMonth_{i,k,q} &= \left(AccumulationRateMonth_{i,k,q}\right) \int_{0}^{DayInMonth_q} 10^{-DieOff_q \cdot t} dt \\ &= \left[\left(AccumulationRateMonth_{i,k,q}\right) / \left(2.303 DieOff_q\right)\right] \times \left[1 - 10^{-\left(DayInMonth_q \cdot DieOff_q\right)}\right] \\ &\approx \left(AccumulationRateMonth_{i,k,q}\right) / \left(2.303 DieOff_q\right) \quad for\ 1 \gg 10^{-\left(DayInMonth_q \cdot DieOff_q\right)} \\ &= \left(AccumulationRateMonth_{i,k,q}\right)\left(DayInMonth_q\right) \quad as\ DieOff_q \rightarrow 0 \end{aligned} \tag{15}$$

where $StorageLimitMonth_{i,k,q}$ is the maximum microbial storage by subwatershed (i) by month (q) by land-use type (k), across all domestic animals and wildlife, adjusted for die-off (removal) (Cells/Area); and $DieOff_q$ is the first-order microbial removal/inactivation/die-off rate on the land surface by month (q) (1/Time).

2.6. Microbial point source loading rates

Monthly microbial point source loadings to a stream include shedding of beef cattle while wading in the stream and leakage from an average septic system, which are used by MSM.

2.6.1. Shedding rates of beef cattle in streams

The monthly microbial loading rate of beef cattle shedding to a stream by subwatershed is as follows:

$$\begin{aligned} BeefCattleShedRateStream_{i,q} &= \left(NumberOfAnimals_{i,m}\right)\left(MicrobeAnimalProductionRates_m\right) \\ &\quad \times \left[\left(GrazingDays_{m,q}\right)/\left(DayInMonth_q\right)\right] \\ &\quad \times \left(TimeSpentInStreams_{m,q}\right) \quad \text{for } m=2 \end{aligned} \tag{16}$$

where $BeefCattleShedRateStream_{i,q}$ is the microbial loading rate of beef cattle $(m=2)$ shedding into a stream by subwatershed (i) by month (q) (Cells/Time).

2.6.2. Microbial loadings due to septics

The average septic flow rate to the stream by subwatershed is as follows:

$$\begin{aligned} SepticStreamFlowRate_i \; &= (SepticNumber_i)(SepticNumberPeople)(SepticOvercharge) \\ &\quad \times (SepticFailureRate) \end{aligned} \tag{17}$$

where $SepticStreamFlowRate_i$ is the average septic flow rate to the stream by subwatershed (i) (Volume/Time), $SepticNumber_i$ is the number of septic systems associated with subwatershed (i) (Number of septics), $SepticNumberPeople$ is the average number of people per septic system (Number of people/septic), $SepticOvercharge$ is the typical septic overcharge flow rate (Volume/Time/Person), and $SepticFailureRate$ is the typical fraction of septic systems that fail (Ratio). The microbial loading rate associated with septic systems by subwatershed is as follows:

$$\begin{aligned} SepticStreamLoadingRate_i &= (SepticStreamFlowRate_i) \\ &\quad \times (SepticConc) \end{aligned} \tag{18}$$

where $SepticStreamLoadingRate_i$ is the microbial loading rate to the stream from leaking septic systems by subwatershed (i) (Cells/Time), and $SepticConc$ is the typical microbial density (concentration) in septic system waste (Cells/Volume).

2.7. Workflow components

A software infrastructure is developed to automate the manual process of characterizing transport of pathogens and indicator microorganisms, from sources of release to points of exposure, by loosely configuring a set of modules and process-based models. A design schematic of the workflow, which tracks data from sources to downstream locations within a watershed and visualizes simulation results, is presented in Fig. 1. Models, in addition to MSM, include D4EM, SDMPB, HSPF, PEST, and BASINS.

D4EM manages, accesses, retrieves, analyzes, and caches web-based environmental data (EPA, 2013a). It is an open source automated data access and processing library that accesses a variety of data types including water quality, land use, hydrology, soils, meteorological (MET), stream flow, groundwater levels, and crop data; uses DotSpatial geo-processing libraries to perform cartographic re-projections, intersection, clipping, overlaying, joining and merging of geographic features, and areas-of-interest delineation; performs statistical processing (extraction, interpolation, and averaging) of time series data; incorporates automatic data access functionality; and consists of a collection of .Net dynamic link libraries that can be linked to a modeling utility such as a batch processor or script to access data for multiple sites, or used with a custom-built user interface. The SDMPB leverages D4EM; provides geographical information system (GIS) capabilities using DotSpatial technology; converts DotSpatial-based project files to MapWindow-based project files (MapWindow, 2013; Watry and Ames, 2008); and pre-populates input files of fate and transport models automatically.

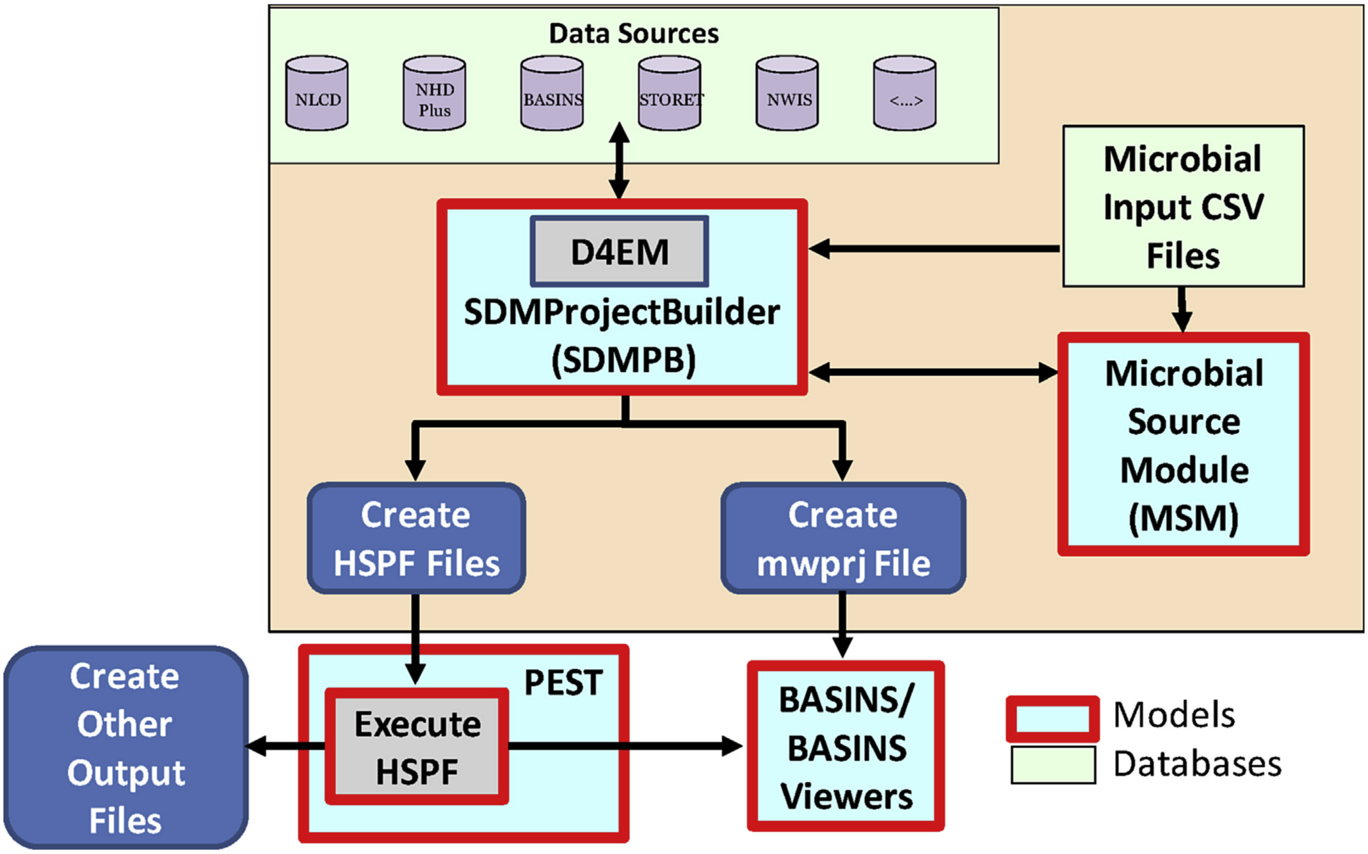

Fig. 1. An automated process-based QMRA workflow (after Wolfe et al., 2016).

HSPF (Bicknell et al., 1997) is a comprehensive package for simulating watershed hydrology and water quality for conventional (e.g., sediment or nutrients) and nonconventional pollutants (e.g., toxic organics) and microbes. It uses basin-scale analysis for integrated simulation of land and soil contaminant runoff processes with instream hydraulic and microbial interactions on user-defined time scales (hour, day, month, or year); and provides a history of runoff flow rates and microbial concentrations at any point in a watershed. Interflow and groundwater supplying the streams are also simulated, but contaminant levels are accounted for by a user-specified constant concentration level. HSPF executes as a stand-alone or within BASINS.

PEST is a model-independent nonlinear parameter estimation package that can estimate parameter values for almost any existing computer model, whether a user has access to the model's source code (Doherty, 2005) or not. PEST is designed to interface with an existing model, modify designated input, run the model as often as needed and adjust its parameters until differences between simulated and monitored output results are minimized, in a weighted least squares sense. PEST communicates with a model through the model's own input and output files. PEST implements a variant of the Gauss-Marquardt-Levenberg method of nonlinear parameter estimation, and also allows fine-tuning of parameter estimation via control variables.

BASINS (EPA, 2001a) provides graphical and tabular viewers of input data, and flow and concentration output. It is a multipurpose environmental analysis infrastructure that performs watershed- and water quality-based analyses by integrating environmental data, analysis tools, and watershed and water quality models. A MapWindow-based GIS organizes spatial information that displays maps, tables, or graphics; analyzes landscape information; and integrates and displays relationships among data at a user-chosen scale.

3. Results

3.1. Input requirements

The MSM has been seamlessly linked with a suite of Comma Separated Values (CSV) files that supply user-defined microbial data. Additional data supplied by SDMPB/D4EM on watershed characteristics are also consumed by MSM. Microbial loadings data produced by MSM represent input to SDMPB which passes the information to downstream models. An example application to a real-world watershed, using the QMRA workflow, illustrates flow and microbial loadings at the pour point.

3.1.1. Microbial input Comma Separated Values files

Twelve user-defined CSV files and their microbial source-term input data requirements for a microbial assessment using MSM (Whelan et al., 2015b) are listed in Table 4. Column 1 identifies the CSV file name and corresponding model (SDMPB or MSM) that consumes data. Columns 2 and 3 define each parameter and its corresponding units, respectively. SDMPB uses some of these data in calculations to produce output (Column 4), which is the input to MSM. For example, input location points defined by latitude and longitude for SDMPB (Column 2) are spatially overlaid automatically onto the watershed to identify corresponding subwatersheds, which are required input to MSM. Example CSV files and input requirements are presented in Appendix A, Tables A1 through A10.

3.1.2. Watershed characteristics

The SDMPB manages and acquires mixed-use watershed information from standard national databases (Table 5) using D4EM, which is based on the BASINS watershed modeling system download tool which accesses, retrieves, analyzes, and caches web-based data. A suite of GIS map layers includes gaging station locations, NHDPlus flowlines, waterbody network, subwatersheds, elevation (e.g., slope), soil types, land-use types, MET stations, and multiple legal boundaries [state, county, roads, eco regions, National Water-Quality Assessment (NAWQA) regions, etc.].

3.2. Output requirements

Five pieces of information produced by MSM:

1. Microbial loading rate by subwatershed (i) by month (q), summed across all domestic animals (m) and wildlife (n) for each land-use type (k) without die-off (a.k.a. MON-ACCUM in HSPF)
2. Maximum microbial storage per land-use type (k) area per subwatershed (i) by month (q), summed across all domestic animals (m) and wildlife (n), adjusted for die-off (a.k.a. MON-SQOLIM in HSPF)
3. Microbial loading rate of domestic animal beef cattle ($m = 2$) shedding to streams by subwatershed (i) by month (q)
4. Average septic flow rate to the stream by subwatershed (i)
5. Microbial loading rate to the stream from leaking septic systems by subwatershed (i).

3.3. Modeling workflow

An example application of a QMRA workflow using MSM is presented in Fig. 1; although hypothetical, it accesses, retrieves, and uses real-world data. Coupled with microbial properties data contained in the CSV data files and D4EM-retrieved data, SDMPB produces a delineated watershed (Fig. 2) of 1358 km^2 (524 mi^2) containing a suite of GIS map layers. These layers include farms with domestic-animal types, numbers, and locations; septic-system locations; NHDPlus flowlines; subwatersheds; waterbody network; elevation (e.g., slope); soil types; land-use types; MET stations; gaging stations; and multiple legal boundaries (Table 5). Fig. 3 presents flow calibration results at the pour point of the watershed, including the initial uncalibrated simulation, monitored gage data, and the initial calibration with the inverse model PEST [correlation coefficient (r) of 0.86]. Fig. 4 presents enterococci calibration results at the pour point of the watershed, including uncalibrated simulation results, 41 monitored sample densities, and the initial calibration with PEST [correlation coefficient (r) of 0.45]. Initial microbial loadings and instream die-off rates within the watershed were used in the calibration. Microbial densities in interflow and groundwater were set to zero because local well data indicated an absence of enterococci.

4. Discussion

QMRA organizes, captures, and executes microbial data to address impacts to mixed-use watersheds within a modeling workflow which involves watershed characterizations, microbial source mapping, and instantiation of the workflow in an assessment. Source-term data are critical to development of a QMRA and, thus, emphasized in this manuscript.

4.1. Automating watershed delineation and microbial source mapping

The workflow that contains the MSM allows for automated watershed delineation and collation of microbial sources within each subwatershed. This allows users to easily change the number and size of subwatersheds, and microbial sources are automatically

Table 4
Files providing data consumed by SDMPB or MSM (after Whelan et al., 2015a; Wolfe et al., 2016).

CSV File Name and Model Consuming Data	Data and Definition, as contained in the CSV File[a]	Units in CSV File	Parameter Consumed as Input by MSM (unless noted)
			Area and AreaFraction[b]
Domestic Animals and Wildlife			
AnimalLL.csv SDMPB	Domestic animal (m) location by latitude and longitude	Degree (by fraction)	Subwatershed[b]
	Domestic animal (m) numbers by latitude and longitude location	Number	NumberOfAnimals
FCProdRates.csv MSM	Production or shedding rate of microbes from domestic animal (m)	Cells/d/animal	MicrobeAnimalProductionRates
	Microbial production or shedding rate per wildlife (n) per area	Cells/d/wildlife	MicrobeWildlifeProductionRates
	Microbial loading rate by sub-urbanized category (u)	Cells/d/ac	SubUrbanizedBuiltupRate
GrazingDays.csv MSM	Number of grazing days per domestic animal ($m = 2, 5, 6,$ and 7) per month (q)	Number	GrazingDays
	Fraction of the number of grazing days that Beef Cattle ($m = 2$) spend in stream per month (q)	Fraction	TimeSpentInStreams
ManureApplication.csv MSM	Fraction of manure applied to soil each month (q) per domestic animal ($m = 1 \rightarrow 5$)	Fraction	Application
	Fraction of amount of manure shed by the domestic animal ($m = 1 \rightarrow 5$) incorporated into soil	Fraction	ManureIncorporatedIntoSoil
MonthlyFirstOrderDieOffRateConstants.csv MSM	First-order microbial inactivation/die-off rate on the land surface per month (q)	1/d	Die-off
WildlifeDensities.csv MSM	Number of wildlife (n) per unit area by land use type (k)	Number/mi^2	Density
Point Sources			
PointSourceLL.csv SDMPB	Point source locations by point source ID (PtSrcId) and latitude and longitude	Degree (by fraction)	Subwatershed[b] (not used by MSM)
PointSourceData.csv SDMPB	Annual-average discharge (Load) for each point source ID (PtSrcId) and facility name (FacName).	ft^3/s	PointFlow (not used by MSM)
	Annual-average microbe loading rate (Load) for each point source ID (PtSrcId) and facility name (FacName).	Cells/yr	PointMicrobeRate (not used by MSM)
	Annual-average chemical loading rate (Load) for each point source ID (PtSrcId) and facility name (FacName).	Lbs/yr	PointChemRate (not used by MSM)
Septic Systems			
SepticsLL.csv SDMPB	Septic system locations by latitude and longitude	Degree (by fraction)	Subwatershed[b] SepticNumber
SepticsDataWatershed.csv MSM	Number of people per septic unit	Number/septic	SepticNumberPeople
	Average fraction of septic systems that fail	Fraction	SepticFailureRate
	Average septic overcharge rate per person	gal/d/person	SepticOvercharge
	Microbial density of septic overcharge reaching the stream	Cells/L	SepticConc
Intermediate Points			
BoundaryPointsLL.csv SDMPB	Boundary points are locations by latitude and longitude where upstream areas have been evaluated *a priori* and represent flow and concentration boundary conditions for downstream evaluation	Degree (by fraction)	Subwatershed[b] (not used by MSM)
OutputPointsLL.csv SDMPB	Output points are intermediate locations by latitude and longitude within the watershed where simulation results are produced	Degree (by fraction)	Subwatershed[b] (not used by MSM)

[a] Indices are defined in Table 3: Subwatershed (i), LandUse (k), Agricultural (m), Wildlife (n), MonthID (q), Urbanized (r), SubUrbanized (u).
[b] Produced by SDMPB, based on NHDPlus data and user-supplied delineation guidelines (i.e., minimum stream length and minimum subwatershed size). The SDMPB overlays and maps latitude-longitude locations to subwatersheds and supplies the corresponding subwatershed location to MSM, when appropriate.

placed within the correct subwatershed and collated accordingly; users, therefore, do not have to manually assign sources (domestic animals, humans, engineered point sources or septics) to subwatersheds. The SDMPB/D4EM automates watershed delineation and microbial source mapping as it

- links to a GIS system (MapWindow) to visualize map layers of data.
- accesses and retrieves web-based data from sources outlined in Table 5 to automatically create input files for MSM, including automatic delineation of watersheds into subwatersheds, areas for and land-use types in each subwatershed, etc.
- allows users to specify intermediate locations (e.g., gaging/monitoring) within a watershed to ensure that the automated delineation process has subwatershed boundaries going through those points.
- provides user control for watershed delineation as it relates to number of subwatersheds, minimum subwatershed size, and minimum stream length. The latter two prevent watershed modeling of areas and streams that are too small, although the smallest areas are those defined by the minimum NHD delineations.
- accesses and retrieves user-defined local data (Appendix A, Tables A1 through A10) which compute microbial loading rates distributed spatially and temporally by subwatersheds.
- allows manual manipulation of input data for more refined, site-specific assessments.

Table 5
Databases automatically accessed and used by D4EM and SDMPB.

- NASA NLDAS (North America Land Data Assimilation System)
- USGS NLCD (National Land Cover Data)
- USGS NWIS (National Water Information System)
- USGS NAWQA (National Water-Quality Assessment program)
- USDA NASS (National Agricultural Statistics Service)
- USDA Soils
  - SSURGO (Soil Survey Geographic database)
  - STATSGO (State Soil Geographic dataset)
- NOAA NCDC (National Climatic Data Center)
- NOAA NDBC (National Buoy Data Center)
- EPA STORET (STOrage and RETrieval)
- EPA Waters Web Services
- EPA BASINS (Better Assessment Science Integrating Point and Nonpoint Sources)
  - Land use/land cover
  - Urbanized areas
  - Populated place locations
  - Reach File version 1 (RF1)
  - Elevation [DEM (Digital Elevation Model)]
  - National Elevation Dataset (NED)
  - Major roads
  - USGS HUC (Hydrologic Unit Code) boundaries
    - Accounting unit
    - Cataloging unit
  - Dam sites
  - EPA regional boundaries
  - State boundaries
  - County boundaries
  - Federal and Indian lands
  - Ecoregions
  - Legacy STORET
- NHDPlus
  - NHD (National Hydrography Dataset)
  - NED (National Elevation Dataset)
  - WBD (Watershed Boundary Dataset)
- NatureServe

EPA = U.S. Environmental Protection Agency.
NASA = National Aeronautics and Space Administration.
NOAA = National Oceanic and Atmospheric Administration.
USDA = U.S. Department of Agriculture.
USGS = U.S. Geological Survey.

- facilitates linkage between microbial sources and loadings through MSM, with fate and transport modeling within a mixed-use watershed.
- assigns North American Land Data Assimilation System (NLDAS) radar or National Climatic Data Center (NCDC) land-based MET data to individual subwatersheds (Kim et al., 2014).
- creates the MSM input file.
- allows users to designate snow accumulation/melt, microbial fate and transport, and simulation time increments (e.g., hourly, daily, monthly, or annually) (Whelan et al., 2015c).
- creates map layers to visualize locations of subwatersheds, land-use types, farms, domestic animals, septics, engineered point sources, monitoring and gaging stations, and MET stations. An example watershed with delineated subwatersheds, water body network, gaging stations, and farms with domestic-animal and septic-system locations is presented in Fig. 2.

4.2. Microbial source characterization

Microbial source characterization identifies types and locations of sources and information that capture microbial loadings and influence fate and transport in the watershed, including locations and types of microbial sources, shedding and production rates, and die-off. Microbial sources include domestic animals, wildlife, septic systems, point sources (WWTPs and POTWs), and urban loadings. Microbial source characterization data supplied by the user are illustrated in Appendix A, Tables A1 through A10.

4.2.1. Microbial source locations and points of interest

Example file formats that document locations of farms containing domestic animals; point sources that discharge directly to the stream; and septic system, output, and boundary points as a function of latitude and longitude are captured (see Tables A1, A2, and A.3, respectively). Output points are intermediate locations within the watershed where simulation results are produced, and boundary points are locations where upstream areas have been evaluated *a priori* and represent flow and concentration boundary conditions for downstream evaluation (see Table 4). Latitude and longitude are used because the data-gathering process on microbial sources is typically determined prior to watershed delineation and can dictate how delineation proceeds. For example, output and boundary locations are of particular importance and must be identified in advance.

SDMPB automatically delineates a watershed into subwatersheds, accounting for the output and boundary locations; overlays latitude-longitude locations; and assigns these locations to subwatersheds. If locations are outside the watershed boundary, these data will be ignored. On the other hand, non-GIS-based models such as MWASTE (Moore et al., 1989), COLI (Walker et al., 1990) and BSLC (Zeckoski et al., 2005) assume that the watershed was delineated prior to manually assigning sources to subwatersheds and, therefore, do not consider points of interest such as monitoring locations or boundary conditions. If the watershed is re-delineated with these models, users must manually repeat the process. In MSM, delineation and overlaying of source locations are automated, so mapping microbial sources and output and boundary points align exactly to the correct subwatersheds without user intervention.

4.2.2. Domestic animals

The number and type of domestic animals associated with each farm location are considered (Table A1), and these numbers are used in Eqs. (3), (8) and (16). County-wide agricultural census data can be retrieved from USDA (2016), although many states track the numbers and types of domestic animals by farm – especially if the numbers exceed a threshold, as with designated concentrated animal feeding operations (CAFOs). In MWASTE, COLI, SEDMOD, BSLC, and SELECT, domestic animals are considered a major source of microbial loadings in watersheds, with some differences from MSM (i.e., number of animal species, consideration of manure application, grazing, in-stream shedding, etc.). MSM's approach is more consistent with BSLC, but BSLC also includes additional domestic animals such as goats and multiple types of chickens and turkeys; MSM includes a catch-all category for other animals (“OtherAg” in Table A1). When only county-level data are available, Zeckoski et al. (2005) suggests estimating and distributing the number of animals among subwatersheds, based on pasture area (e.g., heads per area of pasture).

A single shedding rate is captured (Table A4) and associated with each domestic animal, as defined in Table 2; these numbers are used in Eqs. (3), (8) and (16). A single shedding rate is consistent with other models (e.g., MWASTE, COLI, SEDMOD, BSLC, SELECT). If different age groups or types of domestic animals (calf, heifer, cow, bull, steer, etc.) are of concern for grazing/shedding, these may also be captured in MSM (using categories such as “Other” in Table 2, “OtherAg” in Table A1, and “OtherAgAnimal” in Table A4). MSM assumes that wastes generated from and associated with a location are released within the assigned subwatershed, which is consistent with models such as MWASTE, COLI, SEDMOD, and SELECT.

The schedule of land-application of manure for each domestic animal, as the fraction of manure applied to soil [used in Eq. (1)], is captured monthly (Table A5). In addition, the fraction of manure shed by each domestic animal eventually incorporated into soil

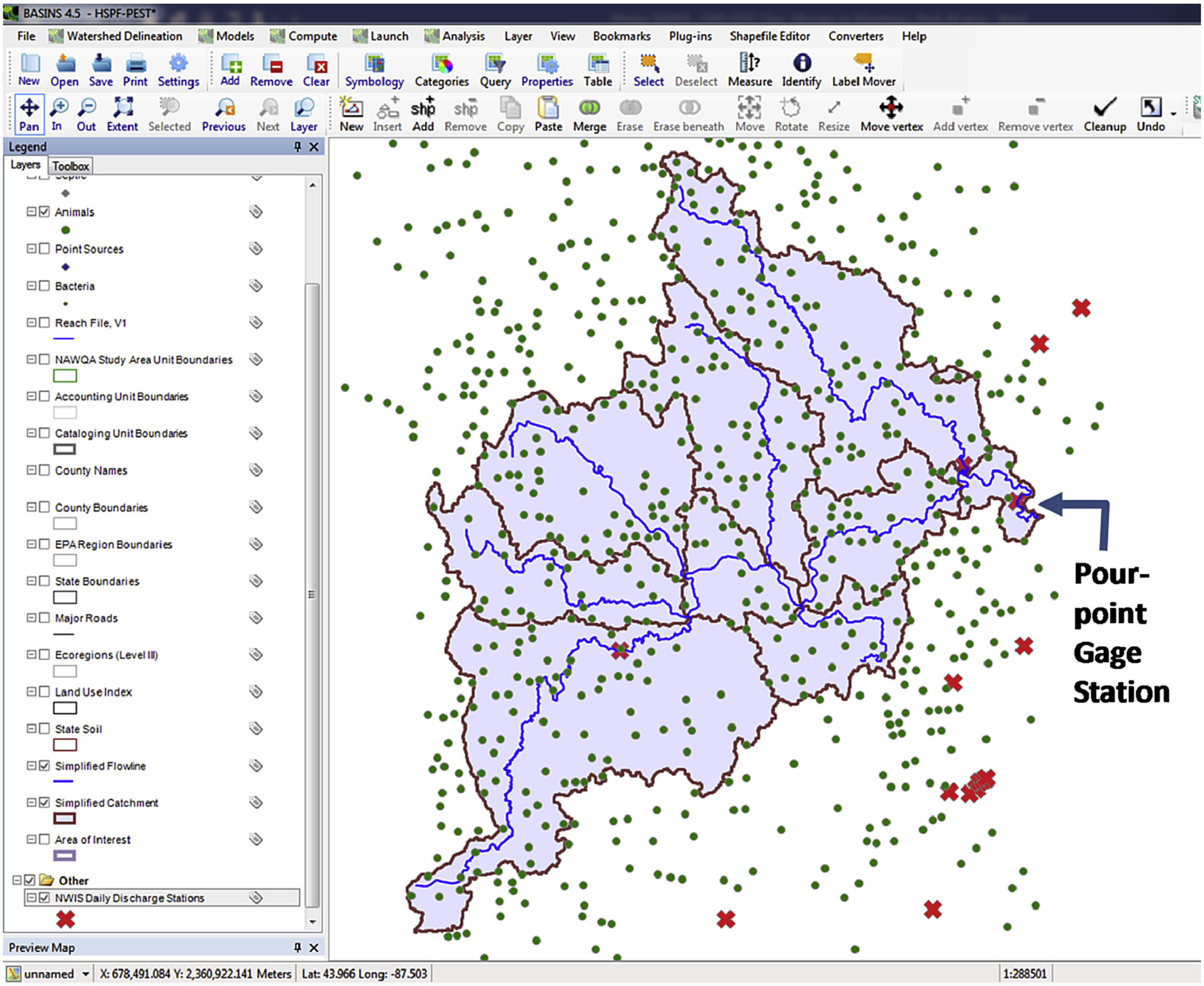

Fig. 2. Example watershed with subwatersheds (brown outline), water body network (blue lines), gaging stations (✖), and farms with domestic animal and septic system locations (●) (after Wolfe et al., 2016). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[used in Eq. (2)] is also captured, a concept consistent with COLI. Monthly schedules capture seasonal trends and are consistent with models like COLI and BSLC.

MSM considers the number of days per month that a domestic animal grazes [used in Eqs. (6), (9) and (16)] and fraction of the number of grazing days per month that beef cattle spend in stream [used in Eqs. (9) and (16)] (Table A.6). The BSLC, on the other hand, gives users the option to define the fraction of time livestock is confined and also considers that both dairy cows and wildlife have access to streams. The MSM assumes that the category “Dairy Cow” is confined and that “Beef Cattle” are allowed to graze and enter the stream; these terms only differentiate between non-grazing and grazing animals, respectively. For example, if dairy cows graze and/or shed directly to the stream, users can designate them as beef cattle. If beef cattle are restricted from entering the stream, then time spent in stream can be set to zero. Schedules (see Table A.6) apply across the watershed, while BSLC allows schedules to vary by subwatershed. The number of grazing days per month cannot be defined in other models such as MWASTE, COLI, SEDMOD, and SELECT.

4.2.3. *Septic systems and point sources*

Instream loadings from septic systems and point sources are considered (see Tables A.7 and A.8, respectively). Septics data accounts for the average number of people per septic unit, fraction of systems that fail, and overcharge flow rate [all used in Eq. (17)] as well as microbial density associated with the overcharge [used in Eq. (18)]. Information supporting septic releases represent the watershed as a whole – that is, the same average usage rate, failure and overcharge rates, and microbial densities are applied to each septic location. All septic systems within each subwatershed are combined to represent a single loading to the respective subwatershed stream segment. The same parameters and methodology are used in SELECT to estimate loadings from septic systems. BSLC is similar, since it assumes an average number of people per septic unit, a human shedding rate, and fraction of failures that cause septic material to rise to the land surface and be carried away by overland runoff. BSLC also considers septic system age (oldest, mid-age, and newest) and homes that discharge directly to the stream.

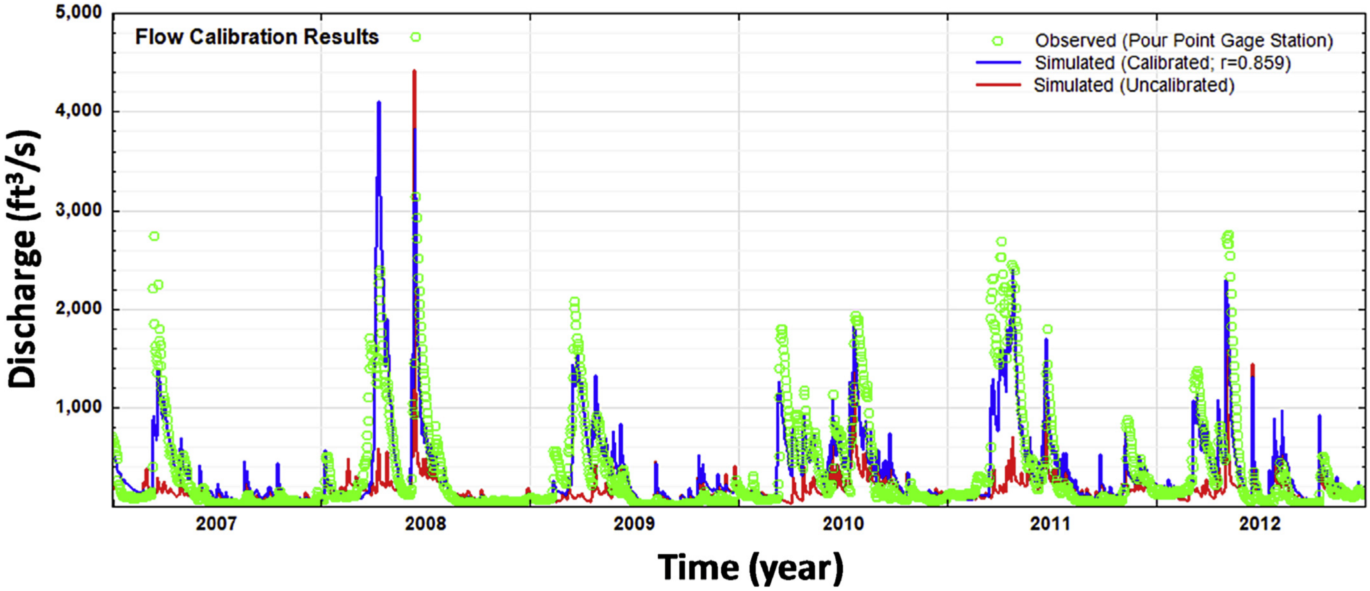

Fig. 3. Flow calibration results at the pour point of the watershed.

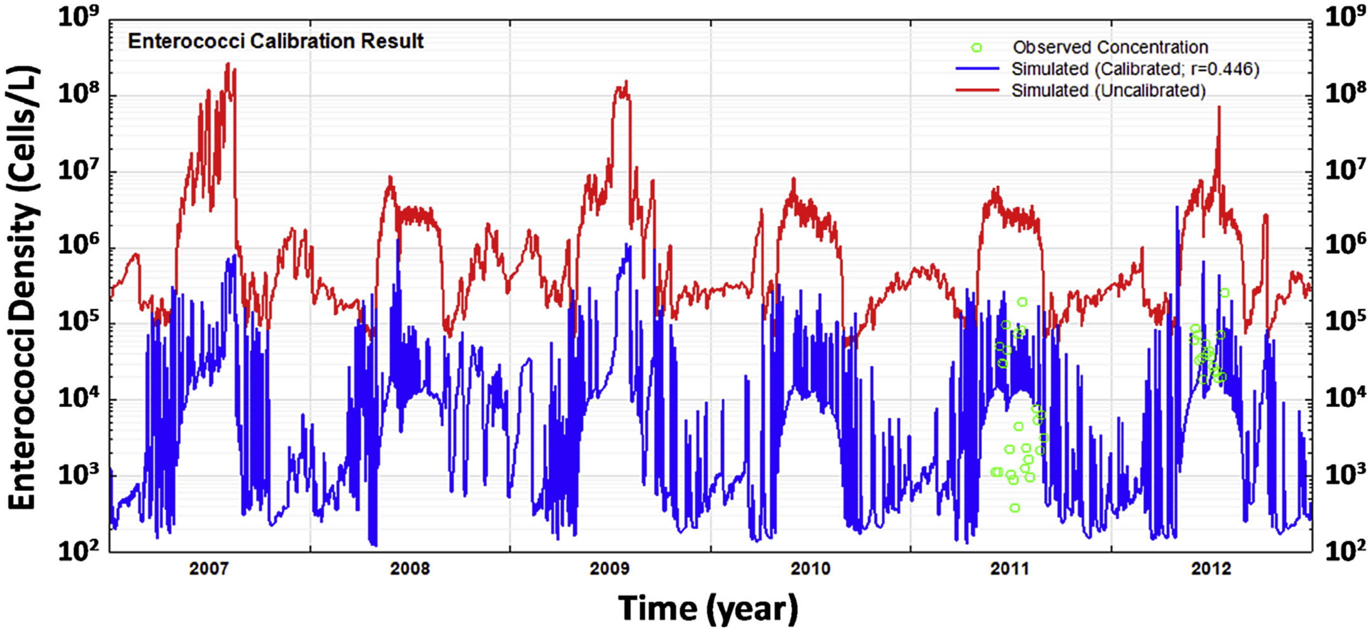

Fig. 4. Enterococci calibration results at the pour point of the watershed.

Point source discharges also include direct input to streams from engineered sources such as WWTPs and POTWs. The SDMPB consumes a single annual average discharge and microbial and chemical loading rates at specified point locations. If the user wants to assess only microbes, the file is modified by removing rows for chemicals. If there is only one point source, rows related to other point sources are removed. Additional point sources can be identified and added. SDMPB consumes these data as a function of point source ID, name, and latitude and longitude (see Table A2). It spatially overlays latitude and longitude locations onto the watershed, mapping to corresponding subwatershed locations, and produces the following data for watershed model consumption: annual-average discharge by subwatershed (*i*) (ft^3/s), annual-average microbe loading rate by subwatershed (*i*) (Cells/yr), and annual-average chemical loading rate by subwatershed (*i*) (Lbs/yr). The annual average point-source data are essentially placeholders for when the user replaces these constant values with actual point source time series (e.g., daily values); this is accomplished by editing the HSPF WDM file through the BASINS inteface and activating the point-source location in the HSPF interface. Because MSM consumes only microbial data, chemical data are of no importance for a microbial assessment in this case and, thus, are not discussed. Automatically including a place holder in the watershed input file allows it to be more easily updated with the actual time series through watershed modeling user interfaces (HSPF and BASINS). Other models such as BSLC and SELECT do not directly address point sources, although the user can manually account for them separately when using a watershed model. Other models (MWASTE, COLI, and SEDMOD) do not consider loading from point sources.

4.2.4. *Wildlife*

MSM considers six wildlife categories (Table 3). Wildlife shedding rates (Table A4) and microbial densities (Table A.9) vary by land-use type and appear in Eq. (10). Wildlife is assumed to shed on forest, cropland, and pasture, but not in urban areas or streams, although certain wildlife (e.g., geese) may shed in large quantities in urban areas and streams. Numbers for selected wildlife such as deer are typically available for each state. BSLC considers seven default wildlife categories; only migratory waterfowl loadings are distributed between the land surface and stream on a monthly basis, although users can add additional wildlife as needed. BSLC assumes that wildlife sheds on forest, cropland, and three pasture designations. SELECT also considers wildlife by distributing the population across suitable habitats (Riebschleager et al., 2012). Other models (MWASTE, COLI, and SEDMOD) do not consider loadings from wildlife.

4.2.5. *Microbial die-off*

Monthly first-order microbial die-off rates due to manure on surface soils, captured by Eq. (15) (Table A10), allow users to account for variations in die-off by season. Die-off rates apply to both domestic animals and wildlife loadings to land surfaces. BSLC uses a constant die-off rate (no monthly variations), while a temperature correction factor (in COLI and MWASTE) and soil pH factor (in MWASTE) are considered. SEDMOD and SELECT do not consider die-off. BSLC uses a similar approach, although it provides an approximate of maximum storage limit [StorageLimitMonth (a.k.a. MON-SQOLIM in HSPF) in Eq. (15)], based on the asymptotic limit over the month. MSM, however, integrates the first-order microbial die-off equation over the month to obtain the exact, closed-form solution [see Eq. (15)]. For example, for a die-off rate of 0.36 d^{-1}, BSLC has a surface accumulation multiplier of 1.77 $[(1-10^{-0.36})^{-1}]$ versus the exact solution used in MSM of 1.21 $\{= [0.36 \cdot \mathrm{Ln}(10)]^{-1}\}$.

4.2.6. *Urban sources*

User-supplied loading rates (i.e., wash off) directly to the stream from urban areas, according to the following categories and subcategories [used in Eq. (12)], are accounted for (see Table A4):

- Commercial and Services: Commercial
- Mixed Urban or Built-up: Road, Commercial, Single-family low density, Single-family high density, and Multi-family residential
- Residential: Single-family low density, Single-family high density, and Multi-family residential
- Transportation, Communications, and Utilities

Urbanized built-up areas include roads (Road), commercial property (Commercial), single-family-low-density residence (SingleFamilyLowDensity), single-family high density residence (SingleFamilyHighDensity), and multi-family residential (MultifamilyResidential). A single, weighted urbanized loading rate is quantified for each subwatershed (all months) based on individual urbanized land uses [Eqs. (13) and (14)].

BSLC and SELECT do not consider different urbanized land-use types, although they capture "residential" areas through consideration of septic systems. "Residential" areas do not include those on a sewer network (Zeckoski et al., 2005) or within a city limit (Teague et al., 2009) and are mostly found in urban areas. Thus, "residential" areas in these models are more applicable to rural, not urban, settings. MWASTE, COLI, and SEDMOD do not consider microbial loadings to urban areas.

4.2.7. *Supporting literature information*

Normal microbial composition of animal feces is different from human feces and can change dramatically over time and space (Boehm et al., 2002; Dorner et al., 2007), so animal and human sources of pathogens and indicators can be treated differently, depending on characterization of fecal material and availability of technology that can accurately and reliably differentiate between sources (EPA, 2009). Wide variability remains within and between sites (Fraser et al., 1998) and in relevant literature. For example, Kim et al. (2016) performed detailed monitoring of microbial release from manure and subsequent overland runoff on 36 identically prepared, side-by-side plots in the same field which resulted in 144 plot-scale, rainfall-runoff events. The range in microbial densities was more than eight orders of magnitude. Wolfe et al. (2016) demonstrated that microbial loading rates to a mixed-use watershed – based on numbers and types of domestic animals and wildlife; microbial densities; shedding and production rates by domestic animal, wildlife, and septics; and microbial die-off rates – represent only estimates and require calibration using observed densities downstream. A summary of densities and production rates is not tabularized herein, although suggested values are found in the published literature.

Soller et al. (2015) and EPA (2010) provided studies related to occurrence and abundance (shedding densities in cells/g manure) of pathogens (e.g., *E. coli* O157:H7, Campylobacter, *Salmonella* spp., *Cryptosporidium* spp., *Giardia* spp.) in manures from domestic animals (beef cattle, dairy cows) and disinfected secondary effluent. EPA (2009) provided representative fecal indicator bacteria and zoonotic (i.e., passage from animals to humans) pathogen densities in human and animal feces and sewage. Soller et al. (2010) documented ranges used to characterize densities of indicators (*E. coli* and enterococci) and reference pathogens (*E. coli* O157:H7, *Cryptosporidium* spp., *Salmonella* spp., *Giardia* spp., Norovirus) in the fecal sources (primary sewage, secondary chlorinated effluent, gulls, cattle, pigs, chickens) (Schoen and Ashbolt, 2010). EPA (2010) and Butler et al. (2008) provided example shedding rates for cows (Whelan et al., 2014b). Geldreich (1978) and ASAE (2005) provided manure production rates and fecal coliform shedding rates associated with various domestic animals and wildlife (Zeckoski et al., 2005). Overcash et al. (1983) provided fecal coliform densities from domestic animal manures (Moore et al., 1989). Walker et al. (1990) used Geldreich's (1978) values for fecal coliform densities in manure; they also provided suggested values for the fraction of manure incorporated by month. EPA (2000) provided example values for fecal coliform shed from domestic animals and wildlife, fecal coliform production rates associated with urban areas (road; commercial; single-family low and high density; and multi-family residential), and supporting information on septics.

Users of the SELECT model (e.g., McFarland and Adams, 2014; Riebschleager et al., 2012; McKee et al., 2011; Teague et al., 2009) provided example shedding rates for *E. coli*, based on fecal coliform production rates for domestic animals and wildlife (EPA, 2001b), assuming a fecal coliform-to-*E. coli* conversion factor of 0.5 recommended by Doyle and Erickson (2006). Riebschleager et al. (2012) and Teague et al. (2009) provided production rates for septic (on-site wastewater treatment) systems, and Riebschleager et al. (2012) also considered urban development and built areas including low-, medium- and high-density land use consisting of single- and multi-family housing, commercial service, industrial and utilities/transportation.

Kim et al. (2016) provided a summary of *E. coli* die-off rates from published literature (Crane and Moore, 1986; Wang et al., 2004; Meals and Braun, 2006; Gu et al., 2012; Blaustein et al., 2013; Martinez et al., 2013; Oladeinde et al., 2014). Based on Moore et al. (1988), Walker et al. (1990) and Moore et al. (1989) tried to account for bacterial die-off in stored manure.

4.3. Workflow instantiation

Instantiation of the QMRA workflow (Fig. 1) begins with the user initiating SDMPB/D4EM, then navigating the United States by state, county, and 8-digit Hydrologic Unit Code (HUC-8), which typically represents an area of 1800 km^2 (~700 mi^2). From here, the user may stay with the HUC-8 selection, or select a different pour point or HUC-12 [100 km^2 (~40 mi^2)] of interest (Whelan et al., 2015c). With user-defined simulation output intervals (hourly, daily, etc.), simulation start and end times, selected data sources and a pour-point selection, SDMPB automatically identifies the upstream basin boundary and registers the 12 user-defined CSV data files (Table 4 and Appendix A) (Whelan et al., 2015b). The number of subwatersheds can also be controlled with user-defined minimum subwatershed areas and stream lengths. The SDMPB manages data acquisition from standard national databases with D4EM and caches web-based data (Table 5). Coupled with boundary and output points (e.g., Table A3) and user-defined minimum stream lengths and subwatershed areas, the SDMPB produces a delineated watershed (Fig. 2) of 1358 km^2 (524 mi^2) containing a suite of GIS map layers that include gaging stations, farms with domestic-animal and septic-system locations, waterbody network, elevation (e.g., slope), soil types, land-use types, and MET stations. Number and type of domestic animals, as well as wildlife density, were collected *a priori*; although these data exist, they are not always routinely known due to privacy/security. The MSM develops microbial loadings (e.g., Cells/Area/Time), adjusted for die-off, to the overland subwatershed areas by land use and to instream (e.g., Cells/Time) locations within a watershed.

The SDMPB automatically pre-populates input needs of the fate and transport watershed model HSPF by automatically creating its Users Control Input (*.uci) file, a collection of geospatial data files, a DotSpatial-based project file, and a MapWindow-based project file (*.wmprj) currently used by BASINS. Using HSPF/BASINS Windows interfaces, non-spatially related data may be modified without re-delineation. For example, if a point source exists within the watershed, its time series loadings can be registered within BASINS prior to HSPF execution, replacing the annually averaged default values provided in the CSV file (e.g., Table A8). HSPF is then executed, creating flows and microbial concentrations that are spatially and temporally distributed throughout the watershed (e.g., Whelan et al., 2015c).

BASINS (e.g., EPA, 2013b, 2013c; Whelan et al., 2015c) provides a user interface and visualization tool for HSPF, and accesses gage data for subsequent inverse modeling. PEST uses HSPF flow and microbial density simulations with monitored flow and microbial density data at the pour point for an initial calibration that will require a final manual calibration. HSPF flow calibration has been discussed by Duda et al. (2012). Key calibration parameters produced by MSM and consumed by HSPF included loadings by microbe and by land-use type, maximum microbial storage accumulation on the land surface, and point source loading rates to the stream from septic systems and direct shedding. Key HSPF microbial calibration parameters include rate of surface runoff (which removes 90% of stored microbes per hour), instream first-order die-off rate, and temperature correction for first-order die-off. Although microbial densities in interflow and groundwater outflow can be considered, they were set to zero because local well data indicated an absence of enterococci.

Using BASINS as a viewer, Figs. 3 and 4 overlay time series of the initial uncalibrated results, those calibrated with PEST, and the monitored gage data at the watershed pour point used in the calibration. Fig. 3 plots flow data, while Fig. 4 plots microbial densities. The improvement is significant when gage data are factored into the assessment. The correlation coefficient (r) for flow was 0.86, while it was 0.45 for microbial density based on only 41 samples over two summers (24 from 2011 to 17 from 2012). The microbial calibration shows poor correlation at very low densities, as illustrated by the 2011 observations in Fig. 4. PEST minimizes the squared-error by not weighting the smaller values as heavily; hence, the better correlation at higher, more important, densities. Wolfe et al. (2016) demonstrated this, when only 17 higher-valued observations were initially employed in their calibration ($r = 0.70$). Their densities during winter months were not much different than those during the summer months, however. By incorporating additional data – even though they contained more variability (i.e., 2011 data in Fig. 4) – base flow due to groundwater is now more appropriately captured during non-summer months, resulting in densities one to two orders of magnitude lower than summer values. Additional sampling would confirm these findings and add to the final calibration. A typical problem with microbial watershed assessments is lack of data which limits the ability to fully understand and capture nuances of cause-and-effect of microbial levels within the watershed (Wolfe et al., 2016). The initial uncalibrated simulation illustrates the need to anchor modeling to sampling/monitoring; in fact, even with limited samples, an initial calibration can significantly improve the uncalibrated results, as illustrated by Figs. 3 and 4. Because results are for an indicator, not a pathogen, there is no risk of infection. Although the framework is designed to simulate risk of infections to pathogens, risk is not computed in the example because results are for an indicator. Whelan et al. (2014b) presents an example where risk is computed for pathogens.

5. Limitations and future efforts

5.1. Limitations

- Although MSM and HSPF are not specific to the continental United States, the QMRA system is currently designed to access databases that specifically support assessments within the United States.
- MSM accounts for variations in monthly agricultural practices in a typical one-year cycle (i.e., January–December) but does not account for variations from year to year.
- The only direct input to streams from shedding is from cattle, neglecting other domestic animals, wildlife, and birds.
- Source-term data requirements documented in Appendix A are typically available in various forms at the state level (Whelan et al., 2017) but currently in a form that requires the user to manually search existing documentation. Capturing these data in a searchable state-by-state database would be a significant improvement; data access and retrieval would be automated,

which would result in more consistency, fewer errors associated with data transfer, and faster assessments.

- The user is currently required to re-run SDMPB to change the spatial delineation of the watershed (e.g., number of subwatersheds), once SDMPB has been executed. Even though the current process is automated, executing re-delineations on the fly would speed the selection of the most appropriate user-required resolution.
- Once the spatial delineation of the watershed is determined by executing SDMPB, the user can change parameter values but not the spatial orientation or delineation of the watershed through HSPF and BASINS user interfaces.
- Yakirevich et al. (2014) note that bottom sediments have been shown to be an important reservoir of *E. coli*, since they can survive and grow in streambed sediments (Cho et al., 2016a, 2016b; Pachepsky and Shelton, 2011). Existing frameworks for modeling microbial transport in steams, including our QMRA framework, disregard the effect of transient storage (dead-end zones represented by stagnant pools, eddies, etc.; Bencala and Walters, 1983, Gooseff et al., 2008), which does not allow one to simulate long tails observed in the time-series of microbial densities (Yakirevich et al., 2014).
- Cho et al. (2016b) introduced temperature as a critical parameter in simulating bacterial growth/die-off and reproducing seasonal variability of bacteria. MSM allows the user to account for seasonal variations in inactivation/die-off with monthly values (see Table A10). HSPF has limitations in simulating seasonal variability of microbes, as it excludes simulations of alternating bacterial growth and die-off by temperature.

5.2. Future efforts

- A storm drain or sewer is designed to drain excess rain and groundwater from impervious surfaces such as paved streets, parking lots, footpaths, sidewalks, and roofs to and through a water distribution system such as pipes. Flow from impervious areas in the current QMRA framework is simulated in HSPF by assuming runoff directly enters the stream; therefore, stormwater runoff through a channel network is unaccounted for. Recent research and modifications have been made to HSPF, using SWMM (EPA, 2016b), to more appropriately account for urbanization by incorporating stormwater drainage into HSPF (Mohamoud and Flaishans, 2013). If deemed an improvement in urban areas, the current version of HSPF will be replaced by this modified version.
- The QMRA system is being exercised on watersheds across the United States (e.g., Manitowoc, WI; Tillamook, OR; Biloxi, MS; Teton River, ID, Salt River, ID) to determine its efficacy in supporting microbial source allocation assessments. A number of these studies investigate 1) which sources and their locations and under what environmental conditions have the greatest impact on downstream microbial density levels within the basin and 2) which strategies mitigate adverse impacts on downstream locations.
- There is currently a two-step process for registering point-source data (direct discharge from POTWs/WWTPs) within the QMRA process: 1) SDMPB formats the HSPF *.uci file to receive and account for point-source data, recorded as annual-average loadings; then 2) HSPF and BASINS user interfaces replace these data with the actual time series. Automating the process of seamlessly registering point-source data would result in more consistency, a reduction in errors associated with data transfer, and faster assessments.
- The current version of the QMRA framework is loosely configured. D4EM, SDMPB, and MSM are seamlessly linked and executed within a single user interface. The user currently executes HSPF, BASINS, PEST, and MRA-IT separately; input files for execution are automatically constructed, although the user can edit them. Even though these assessments can be completed in hours or days, versus months and years, expanding the single-user-interface concept to seamlessly execute all components would improve usability and repeatability.

6. Summary

A Quantitative Microbial Risk Assessment (QMRA) integrates databases and interdisciplinary, multiple media, exposure and effects models. Assessment of predicted indicator concentrations from modeling could be used in determining the appropriateness of waivers to criteria and standards concentration numbers on the basis of site-specific environmental settings and source conditions. Although QMRA does not preclude using source-term data and source and watershed models, it starts and is applied most commonly at the exposure point, skipping release and transport. The work described here documents a new Microbial Source Module (MSM) and describes its mathematical formulations and integration in an environmental modeling workflow to support QMRAs. The MSM estimates microbial loadings within a mixed-use watershed due to land application and shedding of manure from domestic animals and wildlife to pervious areas (e.g., cropland, pasture and forest), from cattle shedding directly to streams, and from failing septic systems discharging to streams. MSM loadings are monthly.

The MSM has been seamlessly linked with a user interface that consumes a suite of Comma Separated Values (CSV) files that supply user-defined microbial data. The interface is based on the Data for Environmental Modeling (D4EM) and Site Data Manager Project Builder (SDMProjectBuilder or SDMPB)]. The two components are linked to include SDMProjectBuilder, D4EM, MSM, accessing and reading the microbial input CSV files, and constructing and populating the Hydrological Simulation Program-FORTRAN (HSPF) input file. The process has been automated from source to receptor within a loosely coupled workflow which allows many of the components to operate individually or within the workflow. MSM works with the SDMPB to read a pre-populated file with microbial source locations, animal numbers and types [agricultural animals (cattle, swine, poultry, etc.)], wildlife (ducks, deer, etc.), manure application schedules, grazing patterns, and point-source releases (septic systems), where applicable. Source locations are defined by latitude-longitude and correlated with land-use type (built-up/impervious, pastureland, cropland, or forest) and subwatersheds, delineated by the SDMPB. The MSM develops monthly microbial loadings (adjusted for die-off) to the soil surface, by land-use type, and to the stream. Engineered point source releases are captured by SDMPB as annual average point-source data that represent placeholders, so the user can replace these constant values with actual point source time series (e.g., daily values) using HSPF and the Better Assessment Science Integrating point & Non-point Sources (BASINS) system. The SDMPB automatically pre-populates the HSPF input file with microbial loading rates and other data needs. Automating data collection and input needs significantly reduces data transfer errors, decreases time for assessment set-up and execution, and allows the user to focus on interrogating results. HSPF routes microbes to the watershed pour point, and BASINS provides a user interface and visualization tool for HSPF and accesses gage data for subsequent inverse modeling, using Parameter ESTimation and Uncertainty Analysis (PEST). PEST uses HSPF flow and microbial density simulations with monitored flow and microbial density data at the pour point to provide an initial calibration of microbial loadings and densities. An example is presented, and, although hypothetical, it accesses, retrieves, and uses real-world data. It also illustrates the importance of using microbial observations in an assessment.

Acknowledgements

The United States Environmental Protection Agency (EPA) through its Office of Research and Development partially funded and collaborated in the research described here under DW-089-92399101 to Idaho National Laboratory, DW-012-92348101 to the U.S. Department of Agriculture Agricultural Research Service, and EP-C-12-021 WA 3-52 to Eastern Research Group, Inc. Thanks are extended to Ms. Fran Rauschenberg of EPA for editing the document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. It has been subjected to Agency review and approved for publication.

Appendix A. User-supplied Microbial Source-term Data

Examples of 12 Comma Separated Values (CSV) files that capture data needs for a microbial assessment (three of which have the same format) are documented in the following 10 tables. Table 4 can be used to help correlate parameter names and their units with the information contained within Tables A1 through A10. Locations of farms that contain domestic animals; point sources that discharge directly into the stream; and septic system, output, and boundary points are documented in Tables A1, A2, and A3, respectively. Output points are intermediate locations within the watershed where simulation results are produced. Boundary points are locations where upstream areas have been evaluated *a priori* and represent flow and concentration boundary conditions for downstream evaluation. Information associated with domestic animals is captured in Tables A1 and A4 through A6. Instream loadings from septic systems and point sources are shown in Tables A3 and A7 and Tables A2 and A8, respectively. Wildlife shedding rates and wildlife densities by land-use type are captured in Tables A4 and A9, respectively; Table A4 also captures loading rates associated with four different urbanized (built-up) areas, with each sub-divided as appropriate, into Urbanized sub-categories. Monthly first-order microbial die-off rates in surface soils appear in Table A10.

Table A.2
PointSourceLL.csv template and example data for two point sources.

Latitude	Longitude	PtSrcId
44.112	−88.256	PT001
44.06	−88.191	PT002

Table A.3
Template and example data for SepticsLL.csv, OutputPointsLL.csv, and BoundaryPointsLL.csv.

Latitude	Longitude
44.23752	−88.0046
44.19700	−88.0954
●	●
●	●
●	●
44.09854	−88.0243
44.08787	−87.9167

Table A.4
FCProdRates.csv template and example shedding rates by domestic animal and wildlife, and production rates by urban (built-up) areas.

Source	Value	Units
DairyCow	2.50E+10	CountPerAnimalPerDay
BeefCow	3.30E+10	CountPerAnimalPerDay
Swine	1.10E+10	CountPerAnimalPerDay
Sheep	1.20E+10	CountPerAnimalPerDay
Horse	4.20E+08	CountPerAnimalPerDay
Poultry	1.31E+08	CountPerAnimalPerDay
Duck	2.40E+09	CountPerAnimalPerDay
Goose	8.00E+08	CountPerAnimalPerDay
Deer	3.50E+08	CountPerAnimalPerDay
Beaver	2.50E+08	CountPerAnimalPerDay
Raccoon	1.25E+08	CountPerAnimalPerDay
OtherAgAnimal	0.00E+00	CountPerAnimalPerDay
OtherWildlife	0.00E+00	CountPerAnimalPerDay
Road	2.00E+05	CountPerAcrePerDay
Commercial	6.21E+06	CountPerAcrePerDay
SingleFamilyLowDensity	1.03E+07	CountPerAcrePerDay
SingleFamilyHighDensity	1.66E+07	CountPerAcrePerDay
MultifamilyResidential	2.33E+07	CountPerAcrePerDay

Table A.1
AnimalLL.csv template and example data.

Latitude	Longitude	BeefCow	Swine	DairyCow	Poultry	Horse	Sheep	OtherAg
44.23752	−88.0046	30	0	30	0	0	0	0
44.19700	−88.0954	0	0	0	0	0	0	0
44.16789	−88.0410	75	0	75	0	0	0	0
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
44.05979	−87.8475	0	0	0	93	0	0	0
44.09854	−88.0243	0	0	0	66	0	0	0
44.08787	−87.9167	0	0	0	25	0	0	0

Table A.5
ManureApplication.csv template and example data.

ManureType	JanFraction Applied	FebFraction Applied	MarFraction Applied	AprFraction Applied	MayFraction Applied	JunFraction Applied	JulFraction Applied	AugFraction Applied	SepFraction Applied	OctFraction Appied	NovFraction Applied	DecFraction applied	FractionIncorporatedIntoSoil
SwineManure	0	0	0	0.15	0.3	0.2	0.15	0.1	0.05	0.05	0	0	0.8
CowManure	0	0	0	0.15	0.3	0.2	0.15	0.1	0.05	0.05	0	0	0.75
CattleManure	0	0	0	0.15	0.3	0.2	0.15	0.1	0.05	0.05	0	0	0.75
HorseManue	0	0	0	0.1	0.1	0	0	0	0	0.4	0.4	0	0.75
PoultryLitter	0	0	0	0.15	0.3	0.2	0.15	0.1	0.05	0.05	0	0	0.96

Table A.6
GrazingDays.csv template and example data.

Month	BeefCattleGrazingDays	HorseGrazingDays	SheepGrazingDays	OtherAgAnimalGrazingDays	FractionOfTimeBeefCattleInStreams
January	0	3.1	31	0	0
February	0	2.8	28	0	0
March	0	3	31	0	0
April	30	27	30	0	0
May	31	27.9	31	0	0
June	30	27	30	0	0.1
July	31	27.9	31	0	0.1
August	31	27.9	31	0	0.1
September	30	27	30	0	0
October	31	27.9	31	0	0
November	0	27	30	0	0
December	0	3.1	31	0	0

Table A.7
SepticsDataWatershed.csv template and example data, correlated with SepticsLL.csv.

NumberOf PeoplePerSepticUnit	SepticFailure Rate_Fraction	SepticOvercharge FlowRate_gallonsPerDayPerPerson	FCConcentrationReachingStreamFrom SepticOvercharge_CountsPerLiter
3.75	0.12	70	1.0E+7

Table A.8
PointSourceData.csv template and example data for two point sources with flow, microbes, and chemicals, correlated with PointSourceLL.csv.

PtSrcId	FacName	Load	Parm
PT001	PointSource1	1	FLOW
PT001	PointSource1	1000	Microbes
PT001	PointSource1	5	Diazinon
PT002	PointSource2	2	FLOW
PT002	PointSource2	2000	Microbes
PT002	PointSource2	4	Diazinon

Table A.9
WildlifeDensities.csv template and example data.

Animal	Density PerSqMile_Cropland	Density PerSqMile_Pasture	Density PerSqMile_Forest
Duck	14.13	14.13	14.13
Goose	2.22	2.22	2.22
Deer	20.51	20.51	20.51
Beaver	0	0	0
Raccoon	0	0	0
OtherWildlife	0	0	0

Table A.10
Template and example data for MonthlyFirstOrderDieOffRateConstants.csv.

Month	DieOffRateContant
January	0.027
February	0.035
March	0.042
April	0.050
May	0.058
June	0.065
July	0.073
August	0.065
September	0.058
October	0.050
November	0.042
December	0.035

Appendix B. Software/Data availability

Categories of free and nonfree software include the following (GNU, 2013a, 2013b): Free (i.e., open source) software that permits anyone to use, copy, and/or distribute, verbatim or with modifications, gratis or for a fee. Most free software is copyrighted, and the copyright holders have legally given permission to use it with a free software license. Public domain software is not copyrighted. Proprietary (i.e., non-free) software is licensed under the exclusive legal rights of its owner.

MSM
Software Developer EPA
Address EPA, 960 College Station Road, Athens GA 30605 USA
Tel +1 706-355-8306
Fax +1 706-355-8302
E-mail Parmar.Rajbir@epa.gov
First available 2016
Hardware requirements 1.5 Ghz CPU 2 GB RAM
Software requirements Windows 7
Availability Open source (no license, source code available upon request)
Cost Free
Program language C#
Program size: 202 KB on disk after installation
Software Access https://github.com/USEPA/MicrobialSourceModule

SDMPB
Software Developer EPA, RESPEC
Address EPA, 960 College Station Road, Athens GA 30605 USA
Tel +1 706-355-8306
Fax +1 706-355-8302
E-mail Parmar.Rajbir@epa.gov
First available 2016
Hardware requirements 2 Ghz CPU 4 GB RAM
Software requirements Windows 7
Availability Open source (no license, source code available upon request)
Cost Free
Program language VB.net, C#
Program size: 523 MB MB on disk after installation
Software Access https://github.com/USEPA/D4EM/tree/master/SDMPB

D4EM
Software Developer EPA, RESPEC
Address EPA, 960 College Station Road, Athens GA 30605 USA
Tel +1 706-355-8306
Fax +1 706-355-8302
E-mail Parmar.Rajbir@epa.gov
First available 201X
Hardware requirements 2.0 Ghz CPU 4 GB RAM
Software requirements Windows 7
Availability Open source (no license, source code available upon request)
Cost Free
Program language VB.net, C#
Program size: 200 MB on disk after installation
Software Access https://github.com/USEPA/D4EM

HSPF
Software Developer EPA, RESPEC
Address EPA, 960 College Station Road, Athens GA 30605 USA
Tel +1 706-355-8306
Fax +1 706-355-8302
E-mail Parmar.Rajbir@epa.gov
First available 1970
Hardware requirements 1 Ghz CPU 4 GB RAM
Software requirements Windows 7

Availability Open source (no license, source code available upon request)
Cost Free
Program language FORTRAN
Program size: 6.4 MB on disk after installation
Software Access https://www.epa.gov/exposure-assessment-models/hspf

BASINS

Software Developer EPA, RESPEC
Address EPA, 960 College Station Road, Athens GA 30605 USA
Tel +1 706-355-8306
Fax +1 706-355-8302
E-mail Parmar.Rajbir@epa.gov
First available 2008
Hardware requirements 2 Ghz CPU 4 GB RAM
Software requirements Windows 7
Availability Open source (no license, source code available upon request)
Cost Free
Program language VB.net, C#
Program size: 2 GB on disk after installation
Software Access https://www.epa.gov/exposure-assessment-models/basins

PEST

Software Developer John Doherty
Address Flinders University, GPO Box 2100, Adelaide 5001, South Australia
Tel +1 301-718-8900 (S.S. Papadopulos & Associates)
E-mail john.doherty@flinders.edu.au
First available 1994
Hardware requirements 1 Ghz CPU 2 GB RAM
Software requirements Windows 7
Availability Open source (no license)
Cost Free
Program language FORTRAN or UNIX
Program size: 50 MB (3 MB on disk after installation)
Software Access http://www.pesthomepage.org/Downloads.php

References

ASAE (American Society of Agricultural Engineers) Standards, 2005. Manure Production and Characteristics. ASAE D384.2. St. Joseph, MI USA. http://extension.psu.edu/animals/dairy/nutrient-management/certified-dairy/tools/manure-prod-char-d384-2.pdf. (Accessed 7 October 2016).

Babendreier, J.E., Castleton, K.J., 2005. Investigating uncertainty and sensitivity in integrated, multimedia environmental models: tools for FRAMES-3MRA. Environ. Modell. Softw. 20 (8), 1043–1055.

Bencala, K.E., Walters, R.A., 1983. Simulation of solute transport in a mountain pool-and riffle stream: a transient storage model. Water Resour. Res. 19, 718–724.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Donigian Jr., A.S., Johanson, R.C., 1997. Hydrological Simulation Program – FORTRAN, User's Manual for Version 11. EPA/600/R-97/080. U.S. Environmental Protection Agency, Athens, GA, 755 pp.

Blaustein, R.A., Pachepsky, Y.A., Hill, R.L., Sheldon, D.R., Whelan, G., 2013. *Escherichia coli* survival in waters: temperature dependence. Water Res. 47 (2), 569–578.

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E., 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. Environ. Sci. Technol. 36 (18), 3885–3892.

Borel, K., Karthikeyan, R., Berthold, T.A., Wagner, K., 2015. Estimating *E.coli* and enterococcus loads in a coastal Texas watershed. Tex. Water J. 6 (1), 33–44.

Butler, D.M., Randells, N.N., Franklin, D.H., Poore, M.H., Green Jr., J.T., 2008. Runoff quality from manured riparian grasslands with contrasting drainage and simulated grazing pressure. Agri Ecosyst. Environ. 126, 250–260.

Cho, K.H., Pachepsky, Y.A., Oliver, D.M., Muirhead, R.W., Park, Y., Quilliam, R.S., Shelton, D.R., 2016a. Modeling fate and transport of fecally-derived microorganisms at the watershed scale: state of the science and future opportunities. Water Res. 100, 38–56.

Cho, K., Pachepsky, Y.A., Kim, M., Pyo, J., Park, M., Kim, J., Kim, J., 2016b. Modeling seasonal variability of fecal coliform in natural surface waters using the modified SWAT. J. Hydrol. 535, 377–385.

Crane, S.R., Moore, J.A., 1986. Modeling enteric bacterial die-off: a review. Water Air Soil Poll. 27 (3/4), 411–439.

Doherty, J., 2005. PEST: Model-independent Parameter Estimation User Manual, fifth ed. Watermark Numerical Computing.

Dorner, S.M., Anderson, W.B., Gaulin, T., Candon, H.L., Slawson, R.M., Payment, P., Huck, P.M., 2007. Pathogen and indicator variability in a heavily impacted watershed. J. Water Health 5 (2), 241–257.

Doyle, M.P., Erickson, M.C., 2006. Closing the door on the fecal coliform assay. Microbe 1 (4), 162–163.

Duda, P.B., Hummel, P.R., Donigian Jr., A.S., Imhoff, J.C., 2012. BASINS/HSPF: model use, calibration, and validation. T ASAE 55 (4), 1523–1547.

Dufour, A.P., 1984. Health Effects Criteria for Fresh Recreational Waters. U.S. Environmental Protection Agency, Washington, DC.

Edwards, D.R., Daniel, T.C., 1992. Environmental impacts of on farm poultry waste disposal - a review. Bioresour. Technol. 41, 9–33.

EPA (U.S. Environmental Protection Agency), 2000. Bacterial Indicator Tool: User's Guide. EPA-823-B-01–003. Office of Water, Washington, D.C.

EPA (U.S. Environmental Protection Agency), 2001a. Better Assessment Science Integrating Point and Nonpoint Sources, BASINS 3.0, User's Manual. EPA-823-B01-001, Washington, DC, 343 pp.

EPA (U.S. Environmental Protection Agency), 2001b. Protocol for Developing Pathogen TMDLs. EPA 841-R-00–002, first ed. Office of Water, Washington, DC USA. https://nepis.epa.gov/Exe/ZyPDF.cgi/20004QSZ.PDF?Dockey=20004QSZ.PDF. (Accessed 4 October 2016).

EPA (U.S. Environmental Protection Agency), 2009. Review of Published Studies to Characterize Relative Risks from Different Sources of Fecal Contamination in Recreational Water. EPA 822-R-09–001. Office of Water, Washington, DC USA.

EPA (U.S. Environmental Protection Agency), 2010. Quantitative Microbial Risk Assessment to Estimate Illness in Freshwater Impacted by Agricultural Animal Sources of Fecal Contamination. EPA 822-R-10–1005. Office of Water, Washington DC.

EPA (U.S. Environmental Protection Agency), 2012. Recreational Water Quality Criteria. 820-F-12–058. Office of Water. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf. (Accessed 18 October 2016).

EPA (U.S. Environmental Protection Agency), 2013a. Data for Environmental Modeling (D4EM). Office of Research and Development, Athens, GA. http://www.epa.gov/AthensR/research/d4em.html. (Accessed 7 October 2016).

EPA (U.S. Environmental Protection Agency), 2013b. BASINS/HSPF Training, Exercise 10 – Bacterial and Temperature Modeling. https://www.epa.gov/sites/production/files/2015-07/documents/exercise-10-bacteria-and-temperature.pdf. (Accessed 7 October 2016).

EPA (U.S. Environmental Protection Agency), 2013c. BASINS User Information and Guidance, BASINS Tutorials and Training. https://www.epa.gov/exposure-assessment-models/basins-user-information-and-guidance. (Accessed 7 October 2016).

EPA (U.S. Environmental Protection Agency), 2015. Human Health Water Quality Criteria. https://www.epa.gov/wqc/human-health-water-quality-criteria. (Accessed 18 October 2016).

EPA (U.S. Environmental Protection Agency), 2016a. National Summary of State Information. https://ofmpub.epa.gov/waters10/attains_nation_cy.control#total_assessed_waters. (Accessed 18 October 2016).

EPA (U.S. Environmental Protection Agency), 2016b. Storm Water Management Model (SWMM): Version 5.1.011 with Low Impact Development (LID) Controls. https://www.epa.gov/water-research/storm-water-management-model-swmm. (Accessed 18 July 2016).

EPA (U.S. Environmental Protection Agency) and USDA (U.S. Department of Agriculture/Food Safety and Inspection Service), 2012. Microbial Risk Assessment Guideline: Pathogenic Microorganisms with Focus on Food and Water. EPA/100/J-12/001; USDA/FSIS/2012-001. Washington, DC.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. 41, 3517–3538.

Fraser, R.H., Barten, P.K., Tomlin, C.D., 1996. SEDMOD: a GIS-based method for estimating distributed sediment delivery ratios. 137–146. In: Hallem, C.A., et al. (Eds.), American Water Resources Symposium on GIS and Water Resources (AWRA TPS-96-3). Ft. Lauderdale, FL. 22-26 Sept 1996. AWRA, Herndon, VA.

Fraser, R.H., Barten, P.K., Pinney, D.A.K., 1998. Predicting stream pathogen loading from livestock using a geographical information system-based delivery model. J. Environ. Qual. 27, 935–945.

Garnett, L.W., 2012. 5 High Density Housing Concepts. http://www.probuilder.com/5-high-density-housing-concepts. (Accessed 7 October 2016).

Geldreich, E.E., 1978. Bacterial populations and indicator concepts in feces, sewage, stormwater and solid wastes. In: Berg, G. (Ed.), Indicators of Viruses in Water and Food, vol. 4. Ann Arbor Science Publishers, Inc., Ann Arbor, MI USA, pp. 51–97.

GNU, 2013a. What Is Free Software? http://www.gnu.org/philosophy/free-sw.en.html. (Accessed 7 October 2016).

GNU, 2013b. Categories of Free and Nonfree Software. http://www.gnu.org/philosophy/categories.html. (Accessed 7 October 2016).

Gooseff, M.N., Bencala, K.E., Wondzell, S.M., 2008. Solute transport along stream and river networks. Chapter 18. In: Rice, S.P., Roy, A.G., Rhoads, B.L. (Eds.), River Confluences, Tributaries and the Fluvial Network. John Wiley & Sons, Ltd.

Gu, Y., Molina, M., Bohrmann, T., Cyterski, M., Whelan, G., Purucker, T., 2012. Estimation of Decay Rates for Fecal Indicator Bacteria and Bacterial Pathogens in Agricultural Field-applied Manure. Annual Meeting of the Association of Southeastern Biologists, Athens, GA. April 04–06.

Guber, A.K., Williams, D.M., Dechen Quinn, A.C., Tamrakar, S.B., Porter, W.F., Rose, J.B., 2016. Model of pathogen transmission between livestock and white-tailed deer in fragmented agricultural and forest landscapes. Environ. Modell. Softw. 80, 185−200.

Haack, S.K., Duris, J.W., 2013. Dynamics of fecal indicator bacteria, bacterial pathogen genes, and organic wastewater contaminants in the Little Calumet River−Portage Burns Waterway, Indiana. J. Gt. Lakes. Res. 39, 317−326.

Haas, C.N., Rose, J.B., Gerba, C.P., 1999. Quantitative Microbial Risk Assessment. John Wiley & Sons, Inc, New York, 449 pp.

Hunter, P.R., Payment, P., Ashbolt, N., Bartram, J., 2003. Chapter 3. Assessment of risk. In: Ronchi, E., Bartram, J. (Eds.), Assessing Microbial Safety of Drinking Water: Improving Approaches and Methods. OECD/WHO Guidance Document. OECD/WHO, Paris, pp. 79−109.

Johnston, J.M., McGarvey, D.J., Barber, M.C., Laniak, G.F., Babendreier, J.E., Parmar, R., Wolfe, K., Kraemer, S.R., Cyterski, M., Knightes, C., Rashleigh, B., Suarez, L., Ambrose, R., 2011. An integrated modeling framework for performing environmental assessments: application to ecosystem services in the Albemarle-Pamlico basins (NC and VA, USA). Ecol. Model 222 (14), 2471−2484.

Kellog, R.L., Lander, C.H., Moffitt, D.C., Gollehon, N., 2000. Manure Nutrients Relative to the Capacity of Cropland and Pastureland to Assimilate Nutrients: Spatial and Temporal Trends for the United States. NRCS and ERS GSA Publ. No. NPS00−0579. U.S. Department of Agriculture, Washington, DC.

Kim, K., Price, K., Whelan, G., Galvin, M., Wolfe, K., Duda, P., Gray, M., Pachepsky, Y., 2014. Using remote sensing and radar meteorological data to support watershed assessments comprising integrated environmental modeling. In: Ames, D.P., Quinn, N. (Eds.), Proceedings of the 2014 International Congress on Environmental Modelling and Software. San Diego, CA USA.

Kim, K., Whelan, G., Molina, M., Purucker, S.T., Pachepsky, Y., Guber, A., Cyterski, M., Franklin, D., Blaustein, R.A., 2016. Rainfall-induced release of microbes from manure: model development, parameter estimation, and uncertainty evaluation on small plots. J. Water Health. http://dx.doi.org/10.2166/wh.2016.239.

Laniak, G.F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R., Hughes, A., 2013. Integrated environmental modeling: a vision and roadmap for the future. Environ. Modell. Softw. 39, 3−23.

MapWindow, 2013. MapWindow 6. http://mapwindow6.codeplex.com/. (Accessed 7 October 2016).

Martinez, G., Pachepsky, Y.A., Sheldon, D.R., Whelan, G., Zepp, R., Molina, M., Panhorst, K., 2013. Using the Q10 model to simulate *E. coli* survival in cowpats on grazing lands. Environ. Int. 54, 1−10.

McFarland, A., Adams, T., 2014. Characterizing Potential Bacteria Loads for the Leona River Watershed Using the Spatially Explicit Load Enrichment Calculation Tool (SELECT). Technical Report No. 1305. Texas Institute for Applied Environmental Research. Tarleton State University, Stephenville, Texas.

McKee, K., Karthikeyan, R., Gregory, L., 2011. Modeling Support for Buck Creek Watershed Protection Plan Development. Texas Water Resources Institute Technical Report. Texas A&M University System, College Station, Texas.

Meals, D.W., Braun, D.C., 2006. Demonstration of methods to reduce *E. coli* runoff from dairy manure application sites. J. Environ. Qual. 35, 1088−1100.

Mohamoud, Y., Flaishans, J., 2013. . A green infrastructure experimental site for developing and evaluating models. In: Presented at American Society of Civil Engineers (ASCE)-world Environmental & Water Resources Congress 2013 Meeting, Cincinnati, OH, May 19-23, 2013. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=246936&searchAll=hspf&searchAll=swmm&actType=&TIMSType=+&TIMSSubTypeID=&DEID=&epaNumber=&ntisID=&archiveStatus=Both&ombCat=Any&dateBeginCreated=&dateEndCreated=&dateBeginPublishedPresented=&dateEndPublishedPresented=&dateBeginUpdated=&dateEndUpdated=&dateBeginCompleted=&dateEndCompleted=&personID=&role=Any&journalID=&publisherID=&sortBy=title&count=25. (Accessed 31 May 2017).

Moore, J.A., Smith, J., Baker, S., Miner, J.R., 1988. Evaluating Coliform Concentrations in Runoff from Various Animal Waste Management Systems. Special Report 817. Agric. Exp. Station, Oregon State University, Corvallis USA. https://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/5558/SR%20no.%20817_ocr.pdf?sequence=1. (Accessed 7 October 2016).

Moore, J.A., Smyth, J.D., Baker, E.S., Miner, J.R., Moffitt, D.C., 1989. Modeling bacteria movement in livestock manure systems. T ASAE 32 (3), 1049−1053.

Oladeinde, A., Bohrmann, T., Wong, K., Purucker, S.T., Bradshaw, K., Brown, R., Snyder, B., Molina, M., 2014. Decay of fecal indicator bacterial populations and bovine-associated source tracking markers in freshly deposited cowpats. Appl. Environ. Microb. 80 (1), 110−118.

Overcash, M.J., Humenik, F.J., Miner, J.R., 1983. Livestock Waste Management, vol. 1. CRC Press, Inc.

Pachepsky, Y.A., Shelton, D.R., 2011. Escherichia coli and fecal coliforms in freshwater and estuarine sediments. Crit. Rev. Environ. Sci. Technol. 41 (12), 1067−1110.

Parajuli, P.B., 2007. SWAT Bacteria Sub-model Evaluation and Application. Doctoral Dissertation. Kansas State University.

Riebschleager, K.J., Karthikeyan, R., Srinivasan, R., McKee, K., 2012. Estimating potential *E.coli* sources in a watershed using spatially explicit modeling techniques. J. A Water Resour. As 48 (4), 745−761.

Savichtcheva, O., Okayama, N., Okabe, S., 2007. Relationships between Bacteroides 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. Water Res. 41, 3615−3628.

Schoen, M.E., Ashbolt, N.J., 2010. Assessing pathogen risk to swimmers at non-sewage impacted recreational beaches. Environ. Sci. Technol. 44, 2286−2291.

Schoen, M.E., Soller, J.A., Ashbolt, N.J., 2011. Evaluating the importance of faecal sources in human-impacted waters. Water Res. 45, 2670−2680.

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J., Ashbolt, N.J., 2010. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. 44 (16), 4674−4691.

Soller, J., Bartrand, T., Ravenscroft, J., Molina, M., Whelan, G., Schoen, M., Ashbolt, N., 2015. Estimated human health risks from recreational exposures to stormwater runoff containing animal fecal material. Environ. Modell. Softw. 72, 21−32.

Teague, A., Karthikeyan, R., Babbar-Sebens, M., Srinivasan, R., Persyn, R.A., 2009. Spatially explicit load enrichment calculation tool to identify potential *E.coli* sources in watersheds. T ASABE 52 (4), 1109−1120.

USDA (U.S. Department of Agriculture), 2016. Census of Agriculture: 2012 Ag Census Data. https://www.agcensus.usda.gov/. (Accessed 7 October 2016).

Walker, S.E., Mostaghimi, S., Dillaha, T.A., Woeste, F.E., 1990. Modeling animal waste management practices: impacts on bacteria levels in runoff from agricultural lands. T ASAE 33 (3), 807−817.

Wang, L., Mankin, K.R., Marchin, G.L., 2004. Survival of fecal bacteria in dairy cow manure. Trans. ASAE 47 (4), 1239−1246.

Watry, G., Ames, D.P., 2008. A Practical Look at MapWindow GIS, first ed. 316 pp.

Whelan, G., Kim, K., Parmar, R., Wolfe, K., Galvin, M., Duda, P., Gray, M., Molina, M., Zepp, R., Pachepsky, Y., Ravenscroft, J., Prieto, L., Kitchens, B., 2014a. Using IEM to automate a process-based QMRA. In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15−19, San Diego, California, USA, ISBN 978-88-9035-744-2.

Whelan, G., Kim, K., Pelton, M.A., Soller, J.A., Castleton, K.J., Molina, M., Pachepsky, Y., Ravenscroft, J., Zepp, R., 2014b. An integrated environmental modeling framework for performing quantitative microbial risk assessments. Environ. Modell. Softw. 55, 77−91.

Whelan, G., Parmar, R., Laniak, G.F., 2015a. Microbial Source Module (MSM): Documenting the Science and Software for Discovery, Evaluation, and Integration. EPA/600/B-15/315. U.S. Environmental Protection Agency, Athens, GA USA.

Whelan, G., Parmar, R., Wolfe, K., Galvin, M., Duda, P., Gray, M., 2015b. Quantitative Microbial Risk Assessment Tutorial − SDMProjectBuilder: Import Local Data Files to Identify and Modify Contamination Sources and Input Parameters. EPA/600/B-15/316. U.S. Environmental Protection Agency, Athens, GA USA.

Whelan, G., Kim, K., Wolfe, K., Parmar, R., Galvin, M., Molina, M., Zepp, R., 2015c. Navigate the SDMPB and Identify an 8-digit HUC of Interest. EPA/600/B-15/273. U.S. Environmental Protection Agency, Athens, GA USA.

Whelan, G., Parmar, R., Wolfe, K., Galvin, M., Duda, P., Gray, M., 2017. Quantitative Microbial Risk Assessment Tutorial − SDMProjectBuilder: Import Local Data Files to Identify and Modify Contamination Sources and Input Parameters − Updated 2017. EPA/600/B-15/316. U.S. Environmental Protection Agency, Athens, GA.

Wikipedia, 2015a. Single-family Detached Home. https://en.wikipedia.org/wiki/Single-family_detached_home. (Accessed 7 October 2016).

Wikipedia, 2015b. https://en.wikipedia.org/wiki/Multi-family_residential (last accessed 07.10.2016).

Wolfe, K.L., Parmar, R.S., Laniak, G.F., Parks, A.B., Wilson, L., Brandmeyer, J.E., Ames, D.P., Gray, M.H., 2007. Data for environmental modeling (D4EM): background and example applications of data automation. In: International Symposium on Environmental Software Systems, Prague, Czech Republic, May 22−25, 2007. http://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=166789&fed_org_id=770&SIType=PR&TIMSType=&showCriteria=0&address=nerl&view=citation&personID=18011&role=Author&sortBy=pubDateYear&count=100&dateBeginPublishedPresented=. (Accessed 7 October 2016).

Wolfe, K., Parmar, R., Whelan, G., Laniak, G.F., Galvin, M., Kim, K., Molina, M., Zepp, R., Duda, P., Keiser, D.D., 2016. A workflow to model microbial loadings in watersheds. In: Sauvage, S., Sánchez-Pérez, J.-M., Rizzoli, A. (Eds.), Proceedings of the International Environmental Modelling and Software Society (iEMSs), 8th International Congress on Environmental Modelling and Software, Toulouse, France.

Yakirevich, A., Pachepsky, Y.A., Guber, A., Kuznetsov, M., 2014. MaSTiS, Microorganism and Solute Transport in Streams, Model Documentation and User Manual. Software and User Manual Public Release. U.S. Department of Agriculture, Agricultural Research Service, pp. 1−23. https://www.ars.usda.gov/ARSUserFiles/80420525/EnvironmentalTransport/MaSTiSFiles/MaSTiS-manual.pdf. . (Accessed 31 May 2017).

Zeckoski, R.W., Benham, B.L., Shah, S.B., Wolfe, M.L., Brannan, K.M., Al-Smadi, M., Dillaha, T.A., Mostaghimi, S., Heatwole, C.D., 2005. BSLC: a tool for bacteria source characterization for watershed management. Appl. Eng. Agric. 21 (5), 879−889.