A Workflow to Model Microbial Loadings in Watersheds

Kurt Wolfe
U.S. Environmental Protection Agency, wolfe.kurt@epa.gov

Rajbir Parmar
U.S. Environmental Protection Agency, parmar.rajbir@epa.gov

Gene Whelan
U.S. Environmental Protection Agency, Whelan.Gene@epa.gov

Gerard F. Laniak
U.S. Environmental Protection Agency, laniak.gerry@epa.gov

Mike Galvin
U.S. Environmental Protection, galvin.mike@epa.gov

See next page for additional authors

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Part of the Civil Engineering Commons, Data Storage Systems Commons, Environmental Engineering Commons, Hydraulic Engineering Commons, and the Other Civil and Environmental Engineering Commons

Wolfe, Kurt; Parmar, Rajbir; Whelan, Gene; Laniak, Gerard F.; Galvin, Mike; Kim, Keewook; Molina, Marirosa; Zepp, Richard; Duda, Paul; and Keiser, Dennis D., "A Workflow to Model Microbial Loadings in Watersheds" (2016). International Congress on Environmental Modelling and Software. 4.
https://scholarsarchive.byu.edu/iemssconference/2016/Stream-A/4
A Workflow to Model Microbial Loadings in Watersheds


*U.S. Environmental Protection Agency, Athens, GA USA
(Wolfe.Kurt@epa.gov, Parmar.Rajbir@epa.gov, Whelan.Gene@epa.gov, Laniak.Gerry@epa.gov, Galvin.Mike@epa.gov, Molina.Marirosa@epa.gov, Zepp.Richard@epa.gov)
bIdaho Falls Center for Higher Education, University of Idaho, Idaho Falls, ID USA
(kkeewook@uidaho.edu, dennisk@uidaho.edu)
cAQUA TERRA Consultants, a Division of RESPEC, INC, Decatur, GA USA
(Paul.Duda@respec.com)

Abstract: Many watershed models simulate overland and instream microbial fate and transport, but few actually provide loading rates on land surfaces and point sources to the water body network. This paper describes the underlying general equations for microbial loading rates associated with 1) land-applied manure on undeveloped areas from domestic animals; 2) direct shedding 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 linked within a workflow containing eight models and a set of databases that form a loosely configured modeling infrastructure which supports watershed-scale microbial source-to-receptor modeling by focusing on animal-impacted catchments. A hypothetical example application – accessing, retrieving, and using real-world data – demonstrates the ability of the infrastructure to automate many of the manual steps associated with a standard watershed assessment, culminating with calibrated flow and microbial densities at the point of a watershed.

Keywords: Integrated Modeling; BASINS; Pathogens; Watershed modeling; QMRA; Source Term

1 INTRODUCTION

There are many sources of pathogenic loadings to streams, such as agronomic practices that utilize animal manures resulting in runoff, direct animal shedding to streams, septic systems, and Publicly Owned Treatment Works/Wastewater Treatment Plants (POTWs/WWTPs). Several watershed models (e.g., Neitsch et al., 2011; Bicknell et al., 1997) simulate the release, fate, and transport of microbes from land sources through water body networks to sensitive receptors of concern. A module, which automatically estimates loading rates on land surfaces and as selected point sources (e.g., seeps, direct animal shedding) to water body networks, would fill an important gap in mixed-use watershed assessments. A Quantitative Microbial Risk Assessment (QMRA) is a modeling approach that integrates disparate data including those linked to fate/transport, exposure, and human health effect relationships to characterize potential health impacts/risks from exposure to pathogenic microorganisms within a watershed (Soller et al., 2010; Whelan et al., 2014a; Haas et al., 1999). Whelan et al. (2014b) described an automated process-based QMRA workflow (Figure 1), containing eight models and a set of databases that form a loosely configured modeling infrastructure which supports watershed-scale microbial source-to-receptor modeling by focusing on animal-impacted catchments. Table 1 describes each model and Figure 1 illustrates their interactions. Because source loadings directly impact spatial and temporal distribution of downstream densities, source-term modules are of utmost importance but tend to be lacking when considering microbial loading. This paper provides general equations to estimate point and non-point source microbial loadings in a mixed-use watershed, and an example application that demonstrates how a QMRA infrastructure can automate many manual steps in a standard watershed assessment.
Figure 1. An automated process-based QMRA workflow (after Whelan et al., 2014b)

Table 1. Model descriptions in QMRA infrastructure

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>D4EM</td>
<td>Manages, accesses, retrieves, analyzes, and caches web-based environmental data</td>
<td>EPA (2013a); Wolfe et al. (2007)</td>
</tr>
<tr>
<td>SDMProjectBuilder</td>
<td>Leverages D4EM to retrieve environmental data; provides geographical information system (GIS) capabilities; converts DotSpatial-based project files to MapWindow-based project files; and automatically pre-populates input files of fate and transport models</td>
<td>Watry and Ames (2008)</td>
</tr>
<tr>
<td>MSM</td>
<td>Organizes, analyzes, and supplies the necessary data to determine microbial loading rates within a watershed to support HSPF input data requirements</td>
<td>Whelan et al. (2015a)</td>
</tr>
<tr>
<td>HSPF</td>
<td>Simulates watershed hydrology and water quality for conventional and nonconventional pollutants and microbes</td>
<td>Bicknell et al. (1997)</td>
</tr>
<tr>
<td>BASINS</td>
<td>Provides graphical and tabular viewers of flow and concentration output</td>
<td>EPA (2001)</td>
</tr>
<tr>
<td>PEST</td>
<td>Provides a nonlinear parameter estimation package for calibration</td>
<td>Doherty (2005)</td>
</tr>
<tr>
<td>TimeseriesUtility</td>
<td>Supports analysis and management of time-varying environmental data</td>
<td>Whelan et al. (2015b)</td>
</tr>
<tr>
<td>MRA-IT</td>
<td>Characterizes human-health risk from ingestion</td>
<td>Soller et al. (2008), Soller and Eisenberg (2008)</td>
</tr>
</tbody>
</table>

2 METHODS AND MATERIALS

This section describes underlying general equations for microbial loading rates associated with 1) land-applied manure on undeveloped areas from domestic animals (e.g., swine, poultry, dairy cows, cattle, etc.); 2) direct shedding on undeveloped lands by domestic animals and wildlife (e.g., duck, goose, deer, beaver, raccoon, etc.); 3) built (i.e., urban or engineered) areas; and 4) point sources that directly discharge to streams from septic systems and shedding by domestic animals proposed by EPA (2013b, 2013c). Measures (e.g., volume, time, etc.) are used to define units. Parameter units are defined using measures. Direct discharges due to Publicly Owned Treatment Works (POTWs) and Waste Water Treatment Plants (WWTPs) are also accounted for but as boundary conditions for point source loadings within the watershed model. For manure deposited on overland surfaces, an algorithm adjusts microbial levels for die-off. All land-applied and shedding loading rates are computed on a monthly basis to account for seasonal effects. The underlying equations can be a function of indices, represented by
time, and different microbes, animals (domestic and wildlife), and/or land use types (cropland, pasture, forest, and urbanized) and their respective areas.

2.1 Land Application

Concentrated Animal Feeding Operations (CAFOs) confine domestic animals for more than 45 days during a growing season, in an area that does not produce vegetation and meets a production process that concentrates large numbers of animals in relatively small, confined places (MacDonald and McBride, 2009). The manure from CAFOs are typically collected in holding ponds and applied to agricultural lands (i.e., cropland, pasture) prior to the growing season; therefore, the microbial land-applied loading rate is estimated as follows:

\[
\text{ARA}_{\text{DA}} = N_{\text{DA}} \times SR_{\text{DA}} \times f_{\text{RO}} \times (365 - N_{\text{GD}}) / N_{\text{Month}} \times \text{Area}
\]  

(1)

\[
f_{\text{RO}} = f_{\text{Apply}} \times (1 - f_{\text{incorp}} / 2)
\]

(2)

where \(\text{ARA}_{\text{DA}}\) is the microbial land-applied loading rate by domestic animal (Cells/Time/Area), \(N_{\text{DA}}\) is the number of domestic animals by location (Number), \(SR_{\text{DA}}\) is the microbial shedding rate by domestic animal (Cells/Time), \(N_{\text{GD}}\) is the total number of grazing days per year by domestic animal (Number), \(N_{\text{Month}}\) is the conversion constant by month for days per month (i.e., January = 31, February = 28, ..., December = 31), Area is that associated with a particular land use type (e.g., cropland, pasture, forest, builtup) (Area), \(f_{\text{RO}}\) is the fraction of annual manure application available for runoff each month by domestic animal (Ratio), \(f_{\text{Apply}}\) is the fraction of annual manure applied each month per domestic animal (Ratio), \(f_{\text{incorp}}\) is the fraction of applied manure incorporated into the soil by domestic animal (Ratio), and 365 is the conversion constant for days in a year. Equation (2) was proposed by EPA (2013b, 2013c).  

2.2 Direct Shedding

Shedding herein refers to defecation by domestic animals or wildlife; animals that are free to roam can shed while grazing on agricultural lands (domestic animals) or in forested areas (wildlife). The microbial shedding rate per wildlife by microbe is computed as follows:

\[
\text{SRA}_{\text{w}} = \rho_{\text{w}} \times SR_{\text{w}}
\]

(3)

where \(\text{SRA}_{\text{w}}\) is the microbial shedding rate by wildlife by microbe by area (Cells/Time/Area), \(\rho_{\text{w}}\) is the number of wildlife per unit area (i.e., density) (Number/Area), and \(SR_{\text{w}}\) is the microbial shedding rate by wildlife by microbe (Cells/Time). The domestic animal microbial loading rate due to grazing is equal to:

\[
\text{SRA}_{\text{DA}} = N_{\text{DA}} \times SR_{\text{DA}} \times N_{\text{GD}} \times (1 - f_{\text{in}}) / \text{Area}
\]

(4)

Where \(\text{SRA}_{\text{DA}}\) is the microbial shedding rate during grazing by domestic animal by microbe (Cells/Time/Area), and \(f_{\text{in}}\) is the fraction of grazing days in the stream for the domestic animal (Ratio).  

2.3 Built Areas

Microbial loading rates by microbe, associated with an urbanized land type (i.e., commercial and service; residential; mixed urban; transportation, communication, and utilities), applicable throughout the year, is computed as follows:

\[
\text{BRA} = \sum_u (f_u \times BR_u)
\]

(5)

where \(\text{BRA}\) is the accumulated microbial loading rate across all urbanized categories “u” (Cells/Time/Area), \(f_u\) is the fraction of the land associated with the urbanized category “u” (Ratio), \(BR_u\) is the microbial accumulation rates by microbe per urbanized land areas (Cells/Time/Area); and “u” is the index on urbanized category for built-up areas: commercial and services; mixed urban or built-up; residential; or transportation, communications, and utilities.
2.4 Point Sources

The microbial loading rate from septic systems directly to the stream is computed as follows:

\[ QC_{\text{septic}} = N_s N_{sp} O_s f_{\text{Fail}} C_{\text{septic}} \]  

(6)

where \( QC_{\text{septic}} \) is microbial loading rate to the stream from leaking septic systems by microbe (Cells/Time), \( C_{\text{septic}} \) is the microbial density in septic system waste by microbe (Cells/Volume), \( Q_{\text{septic}} \) is the average septic flow rate to the stream (Volume/Time), \( N_s \) is the number of septic systems associated with the area (Number), \( N_{sp} \) is the average number of people per septic system (Number), \( O_s \) is the typical septic overcharge flow rate (Volume/Time/Person), and \( f_{\text{Fail}} \) is the typical fraction of septic systems that fail (Ratio). Grazing animals also use stream water for cooling and drinking; therefore, they can shed directly to the stream. \( SR_{DA} \) is used to define the shedding rate directly to the stream.

2.5 Die-off Adjustment

Die-off on overland surfaces is simulated as a function of the rate of microbial accumulation with die-off and total accumulation (maximum storage) without die-off. The unit removal rate represents processes such as die-off and wind erosion (Bicknell et al., 1997). The unit removal rate of microbes stored in the soil (number removed per day) is computed as the microbial accumulation rate (Cells/Area/Time), divided by the maximum microbial storage accumulation (storage limit) (Cells/Area). So, the removal rate = (accumulation rate) / (storage limit). The maximum microbial storage accumulation on the land surface, computed as the sum of storages for each day of the month, is as follows:

\[ N_t = N_o \int_{t=0}^{t=N_{\text{Month}}} \frac{10^{-6} dt}{[k \ln(10)]} = N_o / (2.303 k) \]  

(7)

where \( N_t \) is the maximum microbial storage accumulation on the land surface (Cells/Area), \( N_o \) is the initial uniform loading to the overland surface (Cells/Area), k is the first-order microbial die-off rate (1/Time), and 2.303 is a conversion constant for \( \ln(10) \). \( N_o \) is represented by the sum of loadings by microbe by land use type (e.g., BRA for urbanized, and sum of ARA, SRA_{DA}, and SRA_{w} for all animals for non-urban).

3 RESULTS AND DISCUSSION

3.1 Assumptions and Constraints

Relationships between Equations (1) through (7) and their indices for this assessment are illustrated in Table 2 and covered in more detail in Parmar et al. (2016). Although the equations are written generically, certain constraints ensure proper relevancy; for example, one would not expect wildlife to shed significant amounts in urban areas; likewise, land-applied manure is not in forests or urban areas. The following assumptions and constraints are applied to ensure relevancy (Whelan et al., 2015a).

Table 2. Correlation between equations and indices

<table>
<thead>
<tr>
<th>Indices</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microbe</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Varies by Month</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Domestic Animal</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wildlife</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Land Use Type and Area</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. The 22 (or more) land use types associated with the National Land Cover Database (NLCD) are consolidated into cropland, pasture, forest, and urbanized, which provides a more manageable modeling set when land use is used as an index, since supporting data for finer granularity are not available.
2. Time varies monthly.
3. Fecal shedding from animals is used for loading estimates of all land uses except urbanized.
4. Manures from swine and poultry are assumed to be collected and applied to cropland.
5. Beef cattle and dairy cow manures are assumed to be applied only to cropland and pasture by the same method.
6. Dairy cows are kept in feedlots, so all of their waste is used for manure applications to cropland and pasture.
7. Beef cattle are kept in feedlots or allowed to graze by month; during grazing, a specified percentage of cattle also have direct access to streams. Beef cattle waste is, therefore, applied as manure to cropland and pasture, or contributed directly to pasture (shedding) or streams (shedding).
8. Horse manure not deposited in pasture during grazing is assumed to be collected and applied to pasture.
9. Manures from beef cattle, horses, sheep, and other domestic animals are assumed to contribute to pasture in proportion to time spent grazing. Sheep and other domestic animal manures not deposited to pasture during grazing are assumed to be collected and treated or transported out of the watershed.
10. Wildlife densities are provided for all land uses except urbanized (built-up) and assumed to be the same in all subwatersheds. The wildlife population is considered to be the only microbial contributor to forest.
11. Loading rates are calculated for four urbanized land-use categories:
   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 Multifamily residential
   c. Residential: Average microbial accumulation rates for Single family low density, Single family high density, and Multifamily residential
   d. Transportation, Communications, and Utilities
12. A single, weighted urbanized loading rate is quantified for each subwatershed (all months) based on individual urbanized land uses present.
13. Because beef cattle are allowed to graze, they are assumed to have access to streams; direct contribution of microbes to a stream from beef cattle through shedding is represented as a point source. Since dairy cows are not allowed to graze, they do not have access to streams.
14. Direct contributions of microbes from septics and point sources to a stream are represented as point sources.
15. Point sources are aggregated in subwatersheds.

Based on these assumptions and constraints, Table 3 summarizes relationships between manure applications with land use types by domestic animal and wildlife. Table 4 lists user-supplied input data.

<table>
<thead>
<tr>
<th>Manure Application Correlated to Land Use</th>
<th>Domestic Animals and Wildlife</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beef Cow</td>
</tr>
<tr>
<td>Cropland Grazing/Shedding</td>
<td></td>
</tr>
<tr>
<td>Pasture Grazing/Shedding</td>
<td>x</td>
</tr>
<tr>
<td>Forest Shedding</td>
<td>x</td>
</tr>
<tr>
<td>In Stream Shedding</td>
<td>x</td>
</tr>
<tr>
<td>Cropland Application</td>
<td>x</td>
</tr>
<tr>
<td>Pasture Application</td>
<td>x</td>
</tr>
</tbody>
</table>

**Table 3.** Correlation of manure application with land use type by domestic animal and wildlife (after Whelan et al., 2015a)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic animal locations by Latitude and Longitude</td>
<td>Degree</td>
</tr>
<tr>
<td>Domestic animal numbers by type and location</td>
<td>Number</td>
</tr>
<tr>
<td>Shedding rate of microbes from domestic animal</td>
<td>Cells/Time/animal</td>
</tr>
<tr>
<td>Shedding rate per wildlife per area</td>
<td>Cells/Time/Area</td>
</tr>
<tr>
<td>Number of grazing days per domestic animal per month</td>
<td>Number</td>
</tr>
<tr>
<td>Fraction of grazing days that Beef Cattle spend in a stream per month</td>
<td>fraction</td>
</tr>
<tr>
<td>Fraction of manure applied to soil each month per domestic animal</td>
<td>fraction</td>
</tr>
</tbody>
</table>

**Table 4.** User-supplied input data* (after Whelan et al. 2015c)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of manure shed by the domestic animal incorporated into soil</td>
<td>fraction</td>
</tr>
<tr>
<td>First-order microbial inactivation/die-off rate on the land surface per month</td>
<td>1/Time</td>
</tr>
<tr>
<td>Typical number of wildlife per unit area by land use type</td>
<td>Number/Area</td>
</tr>
<tr>
<td>Septic system locations by Latitude and Longitude</td>
<td>Degree</td>
</tr>
<tr>
<td>Number of people per septic unit</td>
<td>Number</td>
</tr>
<tr>
<td>Average fraction of septic systems that fail</td>
<td>fraction</td>
</tr>
<tr>
<td>Average septic overcharge rate per person</td>
<td>Volume/Time/person</td>
</tr>
<tr>
<td>Typical microbial density of septic overcharge reaching the stream</td>
<td>Cells/Volume</td>
</tr>
</tbody>
</table>

*Example values and formats for each parameter provided in software download.

### 3.2 Example Application

The SDMProjectBuilder/D4EM accesses, retrieves, analyzes, and caches web-based data; delineates the basin into subwatersheds; consumes source location data and overlays source locations as a map layer to identify subwatersheds which correspond to various source locations; consumes microbial properties data; and automatically pre-populates input needs of the fate and transport watershed model. A hypothetical example that implements the QMRA workflow is presented in Figure 1; although hypothetical, this example accesses, retrieves, and uses real-world data. Figure 2 presents the delineated watershed [1358 km² (524 mi²)] with water body network, gaging stations, and farms with domestic-animal and septic-system locations. The number and type of domestic animals, as well as wildlife density, have been collected and, therefore, are known; although these data exist, they are not necessarily routinely known (due to privacy/security). The MSM develops microbial loadings, adjusted for die-off, to the overland subwatershed areas by land use (e.g., Cells/Area/Time) and instream (e.g., Cells/Time) locations within a watershed. HSPF simulates hourly flow and microbial fate/transport within a watershed. BASINS (EPA, 2013b, 2013c) provides a user interface and visualization tool for HSPF, and accesses gage data for use in subsequent inverse modeling. PEST uses HSPF flow and microbial density simulations with monitored flow and microbial density data at the pour point to provide an initial calibration, one that will require a final manual calibration. HSPF flow calibration has been discussed by Duda et al. (2012). Key calibration parameters from the microbial source module to HSPF included loadings by microbe by land use type (e.g., BRA for urbanized, and sum of ARA, SRADA, and SRAw for all animals for non-urban), maximum microbial storage accumulation on the land surface, and point source loading rates to the stream from septics (QCSepic) and direct shedding (SRDA). Key HSPF microbial calibration parameters included the rate of surface runoff which will remove 90 percent of stored microbes per hour, microbial densities in interflow and active groundwater outflow, instream first-order die-off rate, and temperature correction for first-order die-off. Figure 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 0.95]. Figure 4 presents enterococci calibration results at the pour point of the watershed, including the initial uncalibrated.
simulation, results of 24 samples, and the initial calibration with PEST [correlation coefficient \((r)\) of 0.70]. Initial microbial loadings, densities in interflow and groundwater, and instream die-off rates within the watershed were used in the calibration. A typical problem with microbial watershed assessments is the lack of data, which limits the ability to fully understand and capture nuances of cause and effect of microbial levels within the watershed. Hence, the initial uncalibrated simulation illustrates the need to anchor the modeling to sampling; in fact, even with limited samples, an initial calibration can significantly improve the uncalibrated results, as illustrated in Figures 3 and 4. Because these data are indicators, not pathogens, there is no risk of infection; hence, the TimeseriesUtility and MRA-IT were not used.

![Flow Calibration Results](image)

**Figure 3.** Flow calibration results at the pour point of the watershed

![Enterococci Calibration Results](image)

**Figure 4.** Enterococci calibration results at the pour point of the watershed

4 ACKNOWLEDGMENTS

The United States Environmental Protection Agency through its Office of Research and Development funded and managed, and collaborated in the research described here under DW-89-92399101-1 and EP-C-12-021. It has been subjected to Agency review and approved for publication.

5 REFERENCES


