# Quantitative Microbial Risk Assessment Tutorial
## SDMProjectBuilder: Import Local Data Files to Identify and Modify Contamination Sources and Input Parameters
## Updated 2017

**Gene Whelan**
**Rajbir Parmar**
**Kurt Wolfe**
**Michael Galvin**
U.S. Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Ecosystems Research Division
Athens, GA 30605


**Paul Duda**
**Mark Gray**
AQUA TERRA Consultants, a Division of RESPEC, INC.
Decatur, GA 30030

6/27/17

# Summary

Twelve example local data support files are automatically downloaded when the SDMProjectBuilder is installed on a computer. They allow the user to modify values to parameters that impact the release, migration, fate, and transport of microbes within a watershed, and control delineation of a watershed into subwatersheds. They represent templates for site-specific information, so data included in these files, when initially downloaded, provide only an example format for the data which must be modified to reflect conditions associated with the site being assessed. To specifically represent the user's watershed, separate files must be created using the formatted templates to replace existing files; the files are in a comma-separated variable (csv) format and, when modified, should be saved in a csv format. The files providing location information as Latitude-Longitude pairs can also be individually edited within the SDMProjectBuilder. Here, we review formats associated with the templates and explain information contained in each file.

# SDMProjectBuilder:
# Import Local Data Files to Identify and Modify
# Contamination Sources and Input Parameters

**PURPOSE**

Consolidate information for gaining access to software used in support of Quantitative Microbial Risk Assessment (QMRA) modeling.

**OBJECTIVE**

Describe 12 default local data files and their parameters which accompany the SDMProjectBuilder; provide guidance on modifying values of these parameters which impact the release, migration, fate and transport of microbes within a watershed; and describe steps to control delineation of a watershed into subwatersheds.

**GUIDANCE**

This tutorial provides instructions and information on

- Identifying locations of point and non-point sources of contamination;
- Providing guidance for modifying values of parameters that impact the release, migration, fate and transport of microbes within a watershed; and
- Controlling delineation of a watershed into subwatersheds by identifying key locations that represent subwatershed boundaries.

**SOFTWARE ACCESS, RETRIEVAL, AND DOWNLOAD**

The material in this document refers to the SDMProjectBuilder, and gaining access to the files that contain the data requires the user to access, retrieve, and download the software. If interested, Whelan et al. (2017a) provides instructions for access, retrieval, and download of the SDMProjectBuilder, HSPF, and BASINS software products.

**TABLE OF CONTENTS**

**INTRODUCTION**

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. Key elements include the Microbial Source Module (MSM), which has been seamlessly linked with a user interface, based on Data for Environmental Modeling (D4EM) and Site Data Manager Project Builder (SDMProjectBuilder or SDMPB).

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 subsequent 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, and 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 these are addressed through the user interface that externally supports MSM and only represent place holders that will eventually be replaced with actual time series.

D4EM manages, accesses, retrieves, analyzes, and caches web-based environmental data (EPA, 2013). 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.

The MSM and SDMProjectBuilder consume microbial-related data contained in 12 input Comma Separated Values (CSV) files. The data contained in these files capture microbial related properties of microorganisms, as well as data related to the release of microorganisms associated with fecal material. Although a user-friendly interface to manage them does not yet exist, these files allow users to modify input parameters that directly impact microbial loadings to a watershed. They contain microbial source locations, animal numbers and types, correlated with land-use type within subwatersheds. Although there is no user-friendly interface to manage them, these files allow users to identify specific locations of domestic animals, sampling/monitoring, septic systems, and point sources [e.g., Publicly Owned Treatment Works (POTWs) and Wastewater Treatment Plants (WWTPs)].

**DESCRIPTION OF THE 12 DEFAULT LOCAL DATA CSV FILES**

Twelve default local data CSV files allow users to modify values of parameters that impact the release, migration, fate, and transport of microbes within a watershed, and control delineation of a watershed into subwatersheds. The files are included in the SDMProjectBuilder software download and stored in a subfolder, "LocalData." The "LocalData" folder exists in two locations:

1. The "LocalData" subfolder is created within the "SDMPB" folder, when the SDMProjectBuilder is installed, as illustrated in Figure 1.
2. When the user creates a working folder during execution of the SDMProjectBuilder, a "LocalData" subfolder is concurrently and automatically created within the working folder. "TESTA" is an example working folder illustrated in Figure 2, and the subfolder "LocalData" was automatically created to house the 12 files.

The 12 user-defined CSV files and their microbial source-term input data requirements for a microbial assessment are listed in Table 1. 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. A number of parameters are functions of domestic animal, wildlife, land-use type, and/or urbanized built-up areas, and are defined as follows:

There are seven domestic animal designations:
- DairyCow: Dairy Cow
- BeefCattle: Beef Cattle (sometimes noted as BeefCow)
- Swine
- Poultry
- Horse
- Sheep
- OtherAgAnimal: Other Agricultural Animal

There are six wildlife designations:
- Duck
- Goose
- Deer
- Beaver
- Racoon
- OtherWildlife: Other Wildlife

There are four land-use type designations:
- Forest
- Cropland
- Pasture
- Urbanized (a.k.a. Builtup)

There are four Urbanized designations:
- CommercialAndServices: Commercial and Services
- Residential
- MixedUrban: Mixed Urban

- TransportationCommunicationUtilities: Transportation, Communication, Utilities

Table 1 should be used to help correlate parameter names and their units with the information contained within the 12 CSV files.

**Figure 1.** Example of the 12 support files located in the "LocalData" subfolder within the "SDMPB" download folder
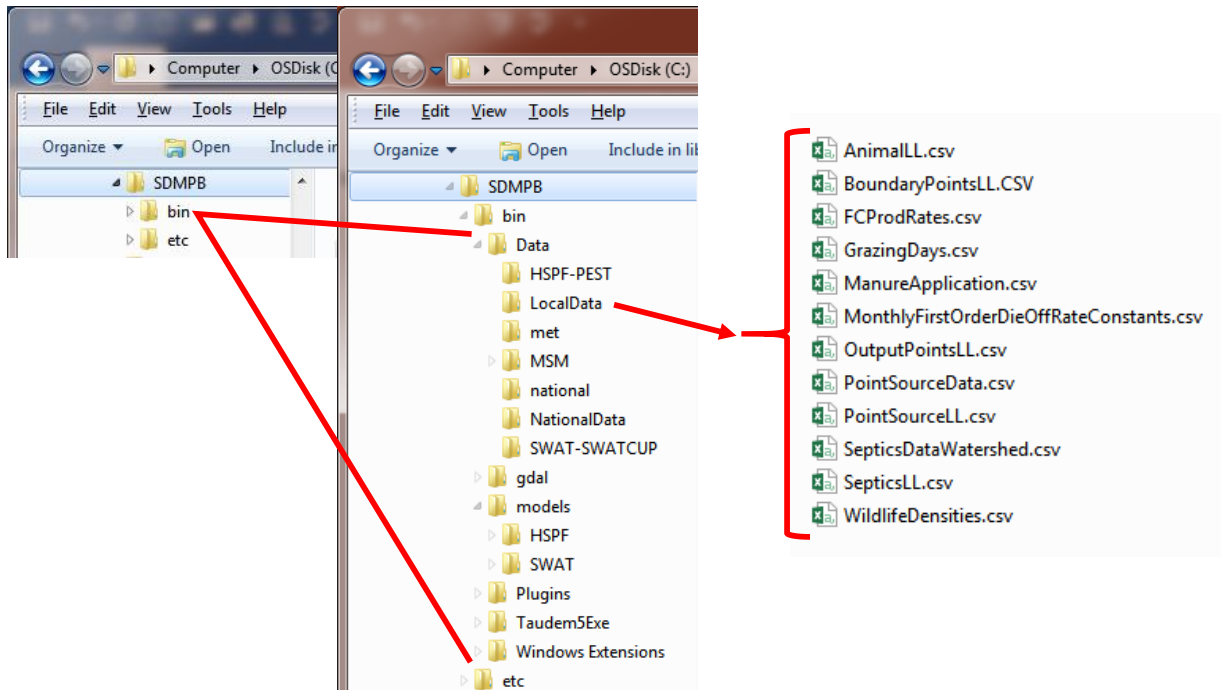


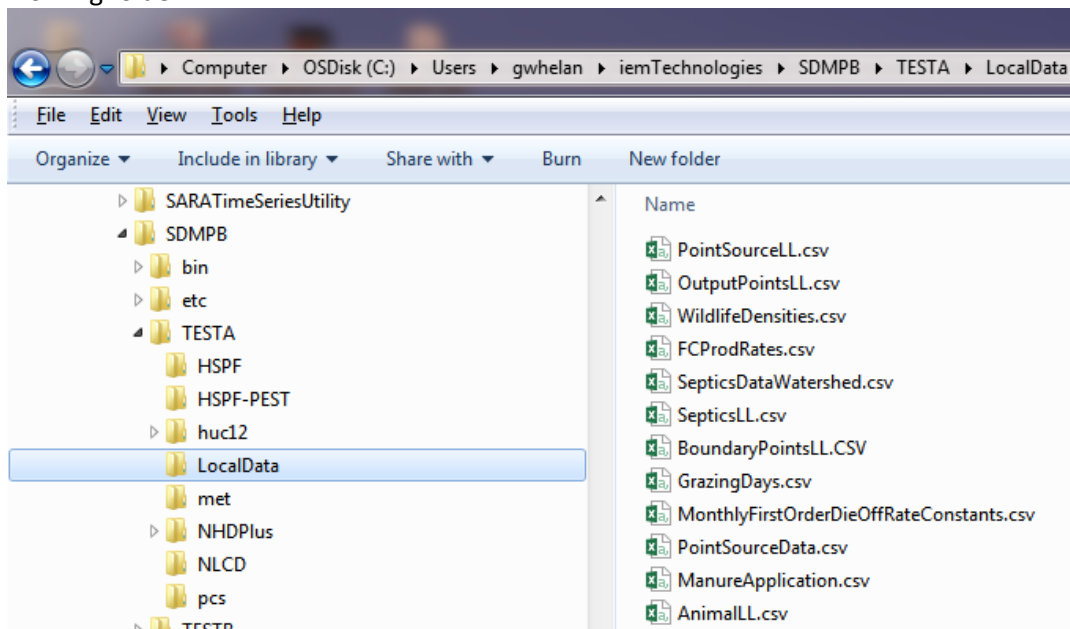**Figure 2.** Example of the 12 support files located in the "LocalData" subfolder within the "TESTA" working folder

**Table 1.** Local Data CSV files providing data consumed by SDMProjectBuilder (SDMPB) or Microbial Source Module (MSM) (after Whelan et al., 2017b; Wolfe et al., 2016)

| CSV File Name and Model Consuming Data | Data and Definition, as contained in the CSV File | Units in CSV File | Parameter Consumed as Input by MSM (unless noted) |
|---|---|---|---|
| | | | Area and AreaFraction* |
| **Domestic Animals and Wildlife** | | | |
| AnimalLL.csv SDMPB | Domestic animal location by latitude and longitude | Degree (by fraction) | Subwatershed* |
| | Domestic animal numbers by type by latitude and longitude location | Number | NumberOfAnimals |
| FCProdRates.csv MSM | Production or shedding rate of microbes from domestic animal | Cells/d/animal | MicrobeAnimalProductionRates |
| | Microbial production or shedding rate per wildlife per area | Cells/d/wildlife | MicrobeWldlifeProductionRates |
| | Microbial loading rate by sub-urbanized category | Cells/d/ac | SubUrbanizedBuiltupRate |
| GrazingDays.csv MSM | Number of grazing days per domestic animal per month | Number | GrazingDays |
| | Fraction of the number of grazing days that Beef Cattle spend in stream per month | Fraction | TimeSpentInStreams |
| ManureApplication.csv MSM | Fraction of manure applied to soil each month per domestic animal | Fraction | Application |
| | Fraction of amount of manure shed by the domestic animal incorporated into soil | Fraction | ManureIncorporatedIntoSoil |
| MonthlyFirstOrderDieOffRateConstants.csv MSM | First-order microbial inactivation/die-off rate on the land surface per month | 1/d | Die-off |
| WildlifeDensities.csv MSM | Number of wildlife per unit area by land-use type | 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* (not used by MSM) |

| | | | |
|---|---|---|---|
| | Annual-average discharge (Load) for each point source ID (PtSrcId) and facility name (FacName). | ft$^3$/s | PointFlow (not used by MSM) |
| PointSourceData.csv SDMPB | Annual-average microbe ($k$) 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* SepticNumber |
| | Number of people per septic unit | Number/septic | SepticNumberPeople |
| SepticsDataWatershed.csv MSM | 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 point locations by latitude and longitude | Degree (by fraction) | Subwatershed* (not used by MSM) |
| OutputPointsLL.csv SDMPB | Output point locations by latitude and longitude | Degree (by fraction) | Subwatershed* (not used by MSM) |

*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.

Within the Urbanized designation, there are five Sub-urbanized designations:
- Commercial
- SingleFamilyLowDensity: Single Family Low Density
- SingleFamilyHighDensity: Single Family High Density
- MultiFamilyResidential: Multi-family Residential
- Road

Single-family low-density is a single-detached dwelling, single-family residence, or separate house which 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 (e.g., 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).

These CSV files **only** represent **templates** for capturing site-specific information, so their contents must be updated by the user to represent the user's specific site. Five of the 12 files provide location information, as Latitude-Longitude pairs, and of these five, three allow the user to identify and manage sources of microbial loadings:

1. land-applied manure (AnimalLL.csv);
2. point sources (PointSourceLL.csv) such as WWTP, POTW, etc.; and
3. septic systems (SepticsLL.csv).

The other two location-based files (BoundaryPoints.csv and OutputPoints.csv) identify locations where the user may want the watershed subdivided, such as sampling/monitoring locations. The remaining seven files describe parameters that impact the actual loadings within a watershed. Whelan et al. (2017b) describe how these parameters are used in the Microbial Source Module, which determines microbial source loadings within a watershed. Use of these files is exemplified by Whelan et al. (2017c 2017d, 2017e, 2017f). Descriptions of each file are presented in the following sections.


**Domestic Animals, Wildlife, and Urban Built-up Area Loading Rates, and Microbial Die-off**

*Domestic Animals*

If domestic animals are located on the watershed, their locations should be reported using the AnimalLL.csv template which captures the location of each animal source and number of animals at that location. Farms are typically associated with these locations. An example of the AnimalLL.csv template is presented in Table 2. These data are usually collected at the state level. At least one Latitude-Longitude pair must be identified. If no sites are within the watershed, define one location and enter zeros for the Latitude and Longitude. Example locations of dairy farms are shown in Figure 3. These were captured with Latitude-Longitude pairs, along with the number of animals at each location, within the AnimalLL.csv template. If domestic animal sources are identified through the AnimalLL.csv template (Table 2), the user must update the following related templates:

**Table 2.** AnimalLL.csv template and example data, including locations of domestic animals and their respective numbers*

| Latitude | Longitude | BeefCow | Swine | DairyCow | Poultry | Horse | Sheep | OtherAg |
|---|---|---|---|---|---|---|---|---|
| 44.22582 | -88.0305 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.2114 | -88.0873 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.09547 | -88.0263 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.23976 | -88.0939 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.19051 | -88.1283 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.22693 | -88.1872 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.19139 | -88.1327 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.18323 | -88.146 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.22671 | -88.1462 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.12401 | -88.0751 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.2271 | -88.2093 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.19682 | -88.0229 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.1845 | -88.1445 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.21851 | -88.0042 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.12406 | -88.0718 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.23874 | -88.1834 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.17717 | -88.0314 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.21836 | -88.1342 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.20785 | -88.0638 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.22697 | -88.1897 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.21111 | -88.0748 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.20104 | -88.0435 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.19683 | -88.0259 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.09548 | -88.0354 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 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 |

● ●
● ●
● ●

| Latitude | Longitude | BeefCow | Swine | DairyCow | Poultry | Horse | Sheep | OtherAg |
|---|---|---|---|---|---|---|---|---|
| 44.05756 | -88.1512 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.08862 | -88.1774 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 43.97918 | -87.8456 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.30336 | -88.1727 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 43.91199 | -88.3076 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.11344 | -87.8003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 44.12612 | -88.2568 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*BeefCow = BeefCattle

**Figure 3.** Locations of dairy farms and assumed septic systems in the Manitowoc River basin near Manitowoc, WI. [Gray area designates the extent of the Manitowoc River basin, and farm and septic locations (i.e., same location) are designated by red dots.]
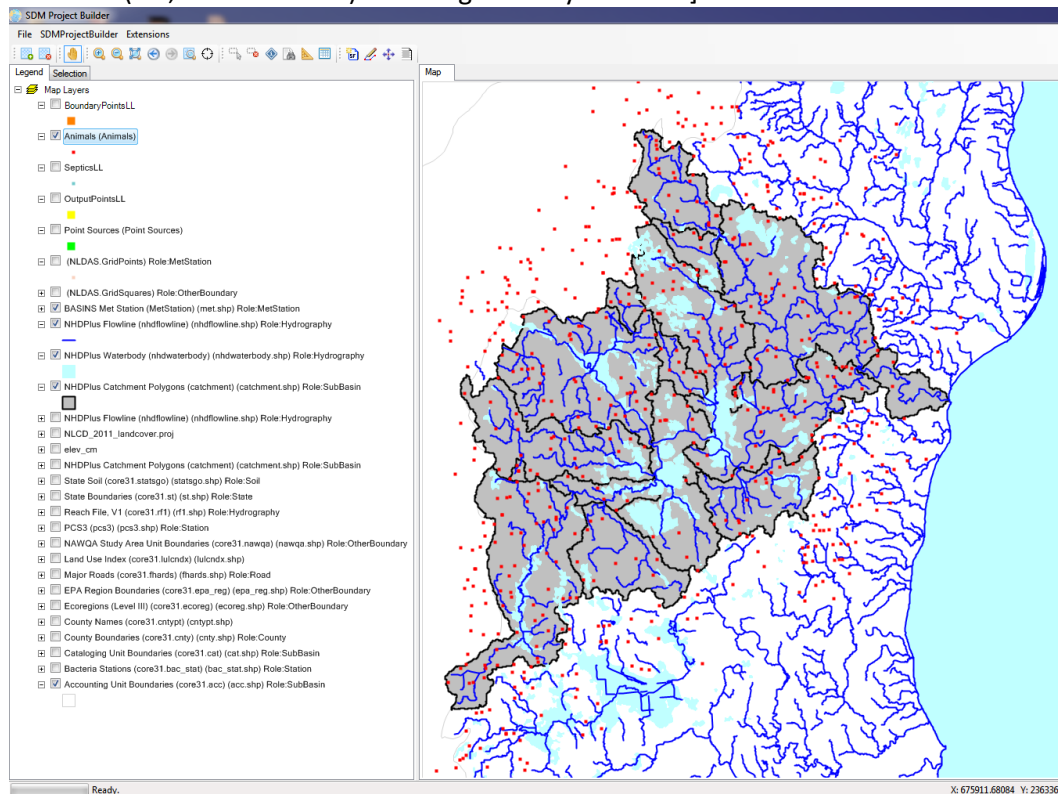


AnimalLL.csv (Table 2), FCProdRates.csv (Table 3), GrazingDays.csv (Table 4), and ManureApplication.csv (Table 5). Microbial production rates, domestic animal grazing days, and fraction of the manure applied to the land surface are documented in the FCProdRates.csv, GrazingDays.csv, and ManureApplication.csv templates, respectively. If no animals are present, it is not necessary to modify FCProdRates.csv (Table 3), GrazingDays.csv (Table 4), and ManureApplication.csv (Table 5) templates, since animals are required to produce microbial loadings. The user should not change column headings in any tables or row labels in Tables 3 through 5.

*Wildlife*

If wildlife are deemed to be a source of microbial contamination, they are assumed to be uniformly distributed in cropland, pasture, and forested areas, so only density is required. If shedding from wildlife must be considered, the FCProdRates.csv (Table 3) and WildlifeDensities.csv (Table 6) templates should be updated. The user should not change column headings or names of the animal. If wildlife is not an important source of microbes, all values in WildlifeDensities.csv template should be zeroed out.

*Urban Built-up Areas*

If Urban built-up areas are considered significant sources of microbial contamination, the FCProdRates.csv template (Table 3) must be updated for roads (Road), commercial property (Commercial), single-family low-density residence (SingleFamilyLowDensity), single-family high-density

residence (SingleFamilyHighDensity), and multi-family residential area (MultifamilyResidential). If they are not, loading rates associated with these parameters in the FCProdRates.csv template should be zeroed out (Table 3). The user should not change column headings or row labels.

**Table 3.** FCProdRates.csv template and example production rate data* [Green signifies loadings from domestic animals and wildlife. Blue signifies loadings from urbanized 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 |

*BeefCow = BeefCattle

*Microbial Die-off*

Die-off (i.e., inactivation) rates are allowed to vary by month for the microbe of interest. Because only one microbe is assessed at a time, die-off rates may need to be updated each time the microbe of interest changes. An example of monthly microbial first-order die-off rate constants is presented in Table 7 (MonthlyFirstOrderDieOffRates.csv). If there are no microbial loadings, having values in the file will not affect the results. The user should not change column headings or row labels.

**Point Sources**

If point sources are located within the watershed, then PointSourceLL.csv and PointSourceData.csv templates should be used, as illustrated in Tables 8 and 9, respectively. Publicly Owned Treatment Works (POTWs) and Waste Water Treatment Plants (WWTPs) are typically associated with these

**Table 4**. 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 5.** ManureApplication.csv template and example data

| ManureType | JanFractionApplied | FebFractionApplied | MarFractionApplied | AprFractionApplied | MayFractionApplied | JunFractionApplied | JulFractionApplied | AugFractionApplied | SepFractionApplied | OctFractionApplied | NovFractionApplied | DecFractionapplied | 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 6**. WildlifeDensities.csv template and example data

| Animal | DensityPerSqMile_Cropland | DensityPerSqMile_Pasture | DensityPerSqMile_Forest | DensityPerSqMile_Builtup |
|---|---|---|---|---|
| Duck | 14.13 | 14.13 | 14.13 | 14.13 |
| Goose | 2.22 | 2.22 | 2.22 | 2.22 |
| Deer | 20.51 | 20.51 | 20.51 | 20.51 |
| Beaver | 0 | 0 | 0 | 0 |
| Raccoon | 0 | 0 | 0 | 0 |
| OtherWildlife | 0 | 0 | 0 | 0 |

**Table 7.** MonthlyFirstOrderDieOffRateConstants.csv template and example data

| Month | DieOffRateContant |
|---|---|
| January | 0.027 |
| February | 0.035 |
| March | 0.042 |
| April | 0.05 |
| May | 0.058 |
| June | 0.065 |
| July | 0.073 |
| August | 0.065 |
| September | 0.058 |
| October | 0.05 |
| November | 0.042 |
| December | 0.035 |

**Table 8.** PointSourceLL.csv template for two point sources

| Latitude | Longitude | PtSrcId |
|---|---|---|
| 44.112 | -88.256 | PT001 |
| 44.06 | -88.191 | PT002 |

locations. Point sources are assumed to be directly discharged to the river/stream. Their locations should be reported using the PointSourceLL.csv template which captures the Latitude-Longitude location of each point source, along with a point-source identifier. If point source IDs are reported in the

PointSourceLL.csv, the user must have the same point source IDs in the PointSourceData.csv file. Users should only change the values, not column headings. Single, annual-average flows, and microbial and chemical loading rates are associated with each location and reported in the PointSourceData.csv template (Table 9) and only represent a place holder within the WDM file within BASINS. Actual point source data must replace these values for a site-specific assessment.

If the user only wants to assess microbes, not chemicals, the file is modified by removing the rows for chemicals in a manner similar to the example in Table 10. For chemicals only, the table is modified similarly by eliminating the rows for microbes. Likewise, if there is only one point source, rows related to the other point sources are removed. Also, multiple point sources can be identified, as illustrated in Tables 11 and 12.

If the user only wants to assess microbes, not chemicals, the file is modified by removing the rows for chemicals in a manner similar to the example in Table 10. For chemicals only, the table is modified similarly by eliminating the rows for microbes. Likewise, if there is only one point source, rows related to the other point sources are removed. Also, multiple point sources can be identified, as illustrated in Tables 11 and 12.

If the default data, as represented by the Latitude-Longitude pairs (Table 8), are outside the watershed boundary, these data will not be assigned to any location within the watershed; hence, no modifications are necessary to the PointSourceLL.csv or PointSourceData.csv default templates. If the default Latitude-Longitude pairs (Table 8) are inside the watershed boundary (and the user would not like them to be), the Latitude-Longitude pairs can be changed to locations outside the watershed boundary (e.g., zero out the Latitude-Longitude pair).

**Table 9.** Example PointSourceData.csv template for two point sources with flow, microbes, and chemicals

| 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 10.** Example PointSourceData.csv template for two point sources with flow and microbes

| PtSrcId | FacName | Load | Parm |
|---------|--------------|------|----------|
| PT001 | PointSource1 | 1 | FLOW |
| PT001 | PointSource1 | 1000 | Microbes |
| PT002 | PointSource2 | 2 | FLOW |
| PT002 | PointSource2 | 2000 | Microbes |

**Table 11.** Example PointSourceLL.csv template for a single point source

| Latitude | Longitude | PtSrcId |
|----------|-----------|---------|
| 44.112   | -88.256   | PT001   |

**Table 12**. Example PointSourceData.csv template for a single point source with flow and microbes

| PtSrcId | FacName      | Load | Parm     |
|---------|--------------|------|----------|
| PT001   | PointSource1 | 1    | FLOW     |
| PT001   | PointSource1 | 1000 | Microbes |
| PT001   | PointSource1 | 5    | Diazinon |

**Septic Systems**

If septic systems are located on the watershed, the SepticsLL.csv and SepticDataWatershed.csv templates should be used, as illustrated in Tables 13 and 14, respectively. Their locations should be reported using the SepticsLL.csv template which captures the Latitude-Longitude location of each; locations of septics are illustrated by Figure 3. Information supporting septic releases is reported in the SepticDataWatershed.csv template (Table 14); these data 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 septics within each subwatershed are combined to represent a single loading to the respective watershed stream segment. At least one Latitude-Longitude pair must be identified, but if not relevant, enter zeros in for the Latitude and Longitude in the SepticsLL.csv template (Table 13). If zeros are identified for the Latitude-Longitude pair, it is not necessary to zero out the values in SepticsDataWatershed.csv Template (Table 14) since these data will not be assigned to any location in the watershed. Users should not change column headings.

**Intermediate Watershed Locations**

Sampling and monitoring traditionally occur at locations with easy access, not locations that are most advantageous for modeling. Software-based automated delineation programs used for watershed modeling typically divide a watershed into subwatersheds without regard to sampling/monitoring locations. Hence, locations where simulation results are available do not necessarily correspond to the sampling/monitoring locations.  For example, two inter-watershed locations of interest (red dots) within a delineated watershed where the subwatersheds are outlined are presented in Figure 4; however, the automated delineation (thin black lines) does not coincide with the sampling/monitoring locations (see Figure 4a), and these data cannot be used in model calibration. In this case, the user must manually delineate the watershed to ensure subwatershed boundaries coincide with the sampling/monitoring locations.

The SDMProjectBuilder automatically delineates the watershed so that the subwatershed boundaries coincide with intermediate watershed locations of interest. Sampling/monitoring locations were identified prior to delineating the watershed into subwatersheds, and they were factored into the delineation process (Figure 4b). Subwatershed boundaries, therefore, coincide with the sampling/monitoring locations, and no manual delineation is required. Two additional files associated

with intermediate points of interest (output locations and boundary conditions) request Latitude and Longitude locations within the SDMProjectBuilder. Each is discussed below.

**Table 13.** SepticsLL.csv template

| Latitude | Longitude |
|---|---|
| 44.22582 | -88.0305 |
| 44.2114 | -88.0873 |
| 44.09547 | -88.0263 |
| 44.23976 | -88.0939 |
| 44.19051 | -88.1283 |
| 44.22693 | -88.1872 |
| 44.19139 | -88.1327 |
| 44.18323 | -88.146 |
| 44.22671 | -88.1462 |
| 44.12401 | -88.0751 |
| 44.2271 | -88.2093 |
| 44.19682 | -88.0229 |
| 44.1845 | -88.1445 |
| 44.21851 | -88.0042 |
| 44.12406 | -88.0718 |
| 44.23874 | -88.1834 |
| 44.17717 | -88.0314 |
| 44.21836 | -88.1342 |

● ●

● ●

● ●

| | |
|---|---|
| 44.05756 | -88.1512 |
| 44.08862 | -88.1774 |
| 43.97918 | -87.8456 |
| 44.30336 | -88.1727 |
| 43.91199 | -88.3076 |
| 44.11344 | -87.8003 |
| 44.12612 | -88.2568 |

*Output Locations*

Output locations are intermediate points, where the user would like the watershed model to produce simulation results, and only exist at subwatershed boundaries (see Figure 4b). As noted earlier, by ensuring that simulation results coincide with sampling/monitoring locations, they can be directly compared to sampling/monitoring data and can help calibrate the model for subwatersheds upstream from this location. The OutputPoints.csv template (Table 15) captures output Latitude-Longitude locations. When modifying the csv file, users may identify more than one location, but should not change column headings. At least one Latitude-Longitude pair must be identified, and if not relevant, enter zeros in for each Latitude and Longitude in the OutputPoints.csv template (Table 15).

**Table 14.** SepticsDataWatershed.csv template

| NumberOfPeoplePerSepticUnit | SepticFailureRate_Fraction | SepticOverchargeFlowRate_gallonsPerDayPerPerson | FCConcentrationReachingStreamFromSepticOvercharge_CountsPerLiter |
|---|---|---|---|
| 3.75 | 0.12 | 70 | 10000000 |

**Table 15.** OutputPoints.csv Template

| Latitude | Longitude |
|---|---|
| 44.11171 | -87.78116 |
| 44.13477 | -87.76533 |

*Boundary Conditions*

After a model has been calibrated at one output location or if data exist that provide necessary boundary conditions, these data could be directly consumed by the watershed model without re-running the simulation upstream of the boundary condition location. In essence, those subwatersheds would be removed from the assessment and replaced by a file containing previously collected data or calibration results; for example, a user-specified location (red dot) where known flow data exists is illustrated by Figure 5a. As opposed to simulating that portion of the watershed upstream from this point (aqua region in Figure 5b), flow information at that location could be provided as the upstream boundary condition in a file, eliminating the need to model the upstream portion (aqua region in Figure 5b).

The BoundaryPoints.csv template (Table 16) captures Latitude-Longitude boundary locations. In the SDMProjectBuilder, these are associated with an instream location, not an overland segment. When modifying the csv file, users may identify more than one location but should not change column headings. Currently, the SDMProjectBuilder only allows the user to identify the boundary condition location – consuming boundary condition data has NOT been activated. At least one Latitude-Longitude pair must be identified and if not relevant, enter zeros for each Latitude and Longitude in the BoundaryPoints.csv template (Table 16).
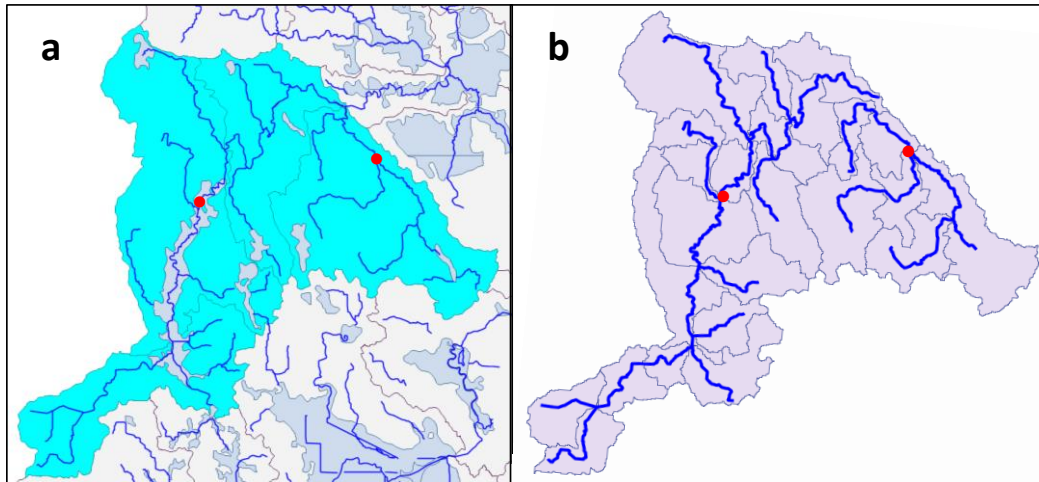


**Figure 4.** Delineated watershed (a) without considering intermediate points of interest to the user within the watershed, and (b) considering intermediate points of interest
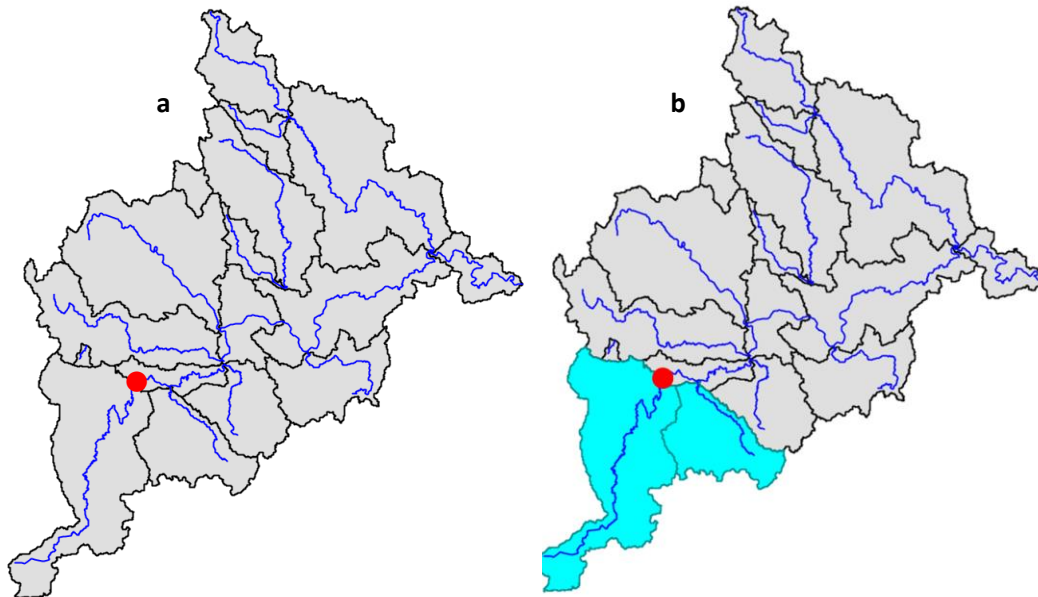


**Figure 5.** Example of using a boundary condition to identify a location where data are known: (a) red dot symbolizes the boundary condition location on a watershed and (b) aqua region symbolizes the region where known data would replace and modeling simulations.

**Table 16.** BoundaryPoints.csv template

| Latitude | Longitude |
|----------|-----------|
| 44.02731 | -88.1643  |

**NAVIGATING THE SDMProjectBuilder TO IDENTIFY A HUC-8**

Whelan et al. (2017c) describe navigation of the SDMProjectBuilder and how to identify a watershed of interest by providing step-by-step instructions to

- Identify a New project ("New SDM Project"):
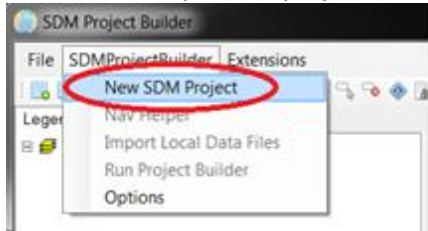


- Navigate through the GIS with the Navigation Helper ("Nav Helper") to identify the relevant 8-digit HUC (HUC-8) of interest:
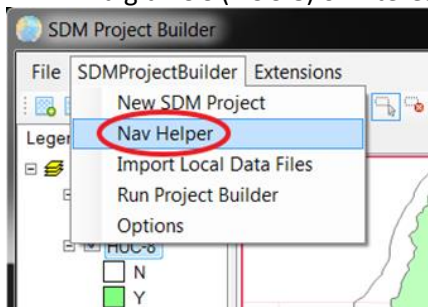
## IMPORTING LOCAL DATA FILES AS MAP LAYERS

Of the 12 default local data files, five report locations that can be imported and displayed as map layers (i.e., AnimalLL.csv, BoundaryPointsLL.csv, OutputPointsLL.csv, PointSourceLL.csv, and SepticsLL.csv), using the following procedure.

It should be noted that after replacing or updating file in the "LocalData" folder within the working folder, data will register automatically with the SDMProjecBuilder (SDMPB). It is unnecessary to choose "Import Local Data Files under the "SDMProjectBuilder" tab, as noted by the "X" on the screen below, unless the user would like to view the results as a map layer.
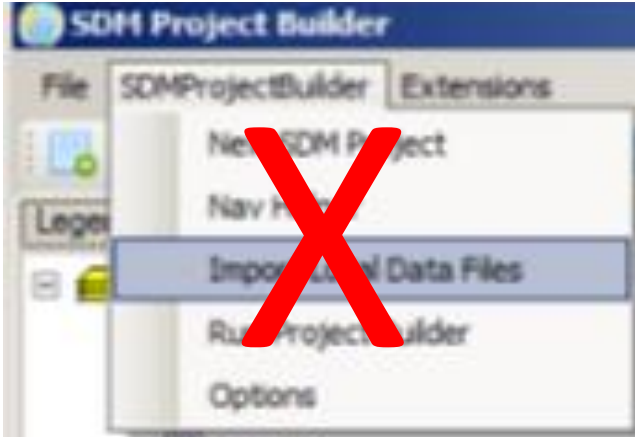


This section explains how to view the local data files as a map layer.

1. Click on "Import Local Data Files" under the "SDMProjectBuilder" tab.
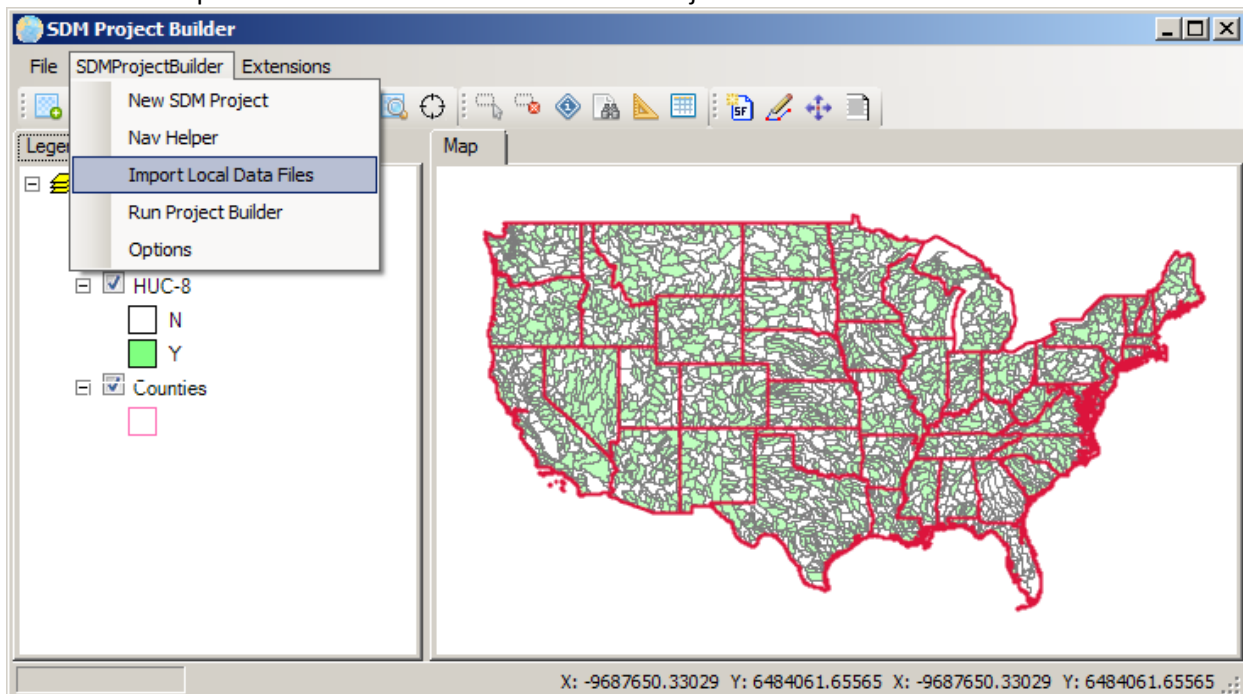
2. A screen will appear that allows the user to import data associated with 12 "Local Data Files" of which the five files listed in the following screen identify specific Latitude-Longitude locations for domestic animals (e.g., farms), intermediate points (i.e., boundary conditions, output points), point sources (e.g., POTWs), and septic systems.
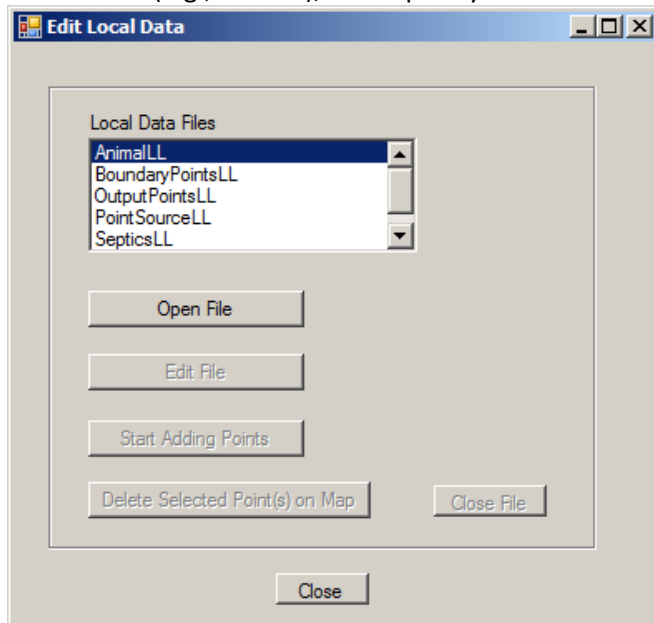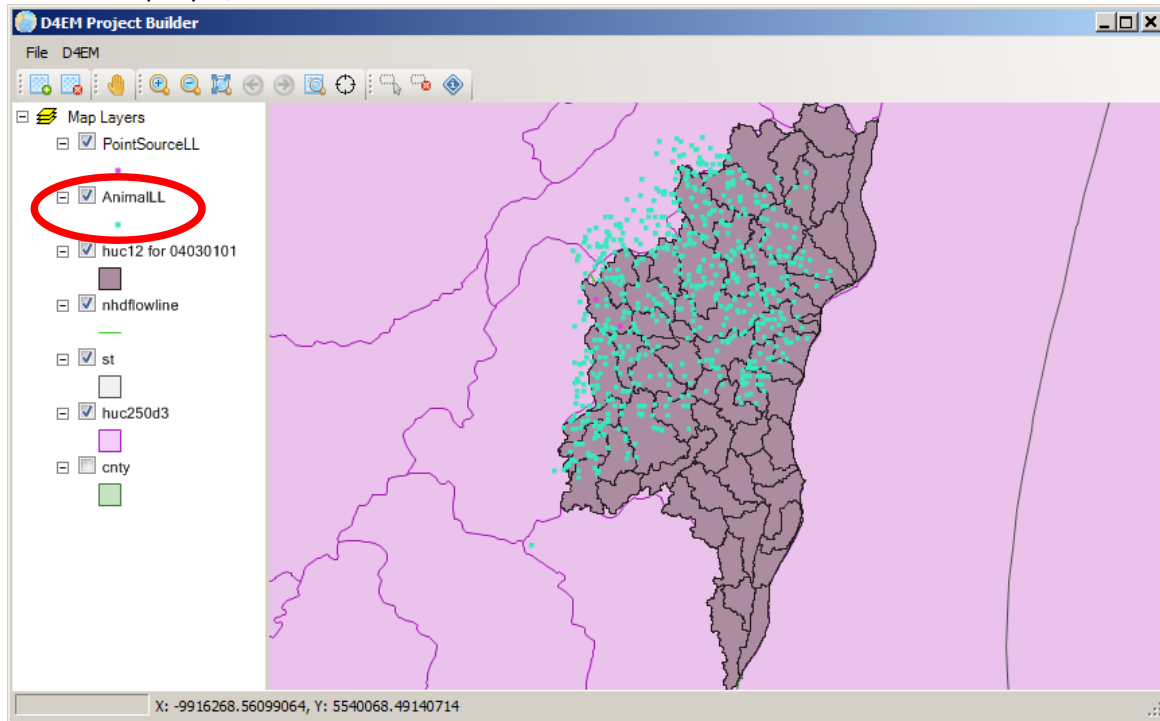


To register the relevant local data files containing points of interest,
- choose "Open File"
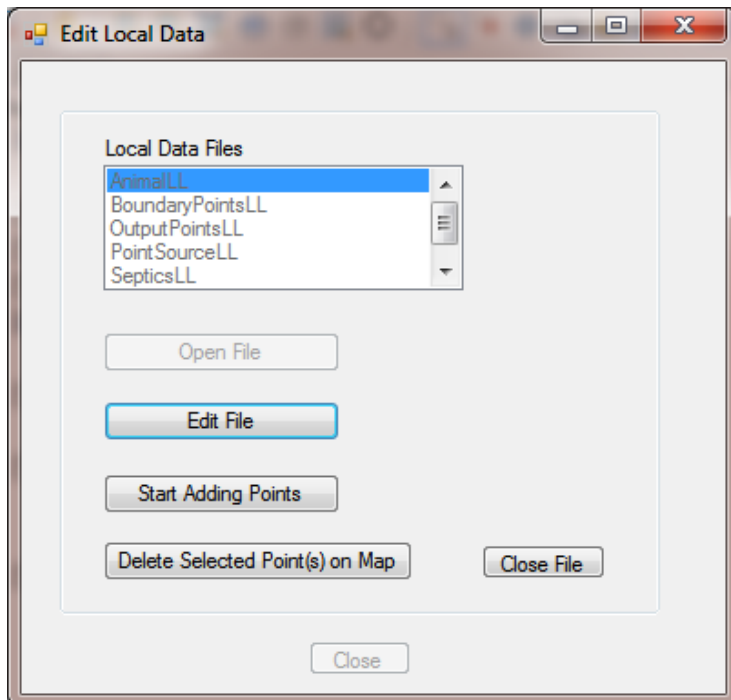- wait until the screen changes,
- then choose "Close File."

3. Register locations of domestic animals; choose "AnimalLL," as illustrated below, then "Open File."
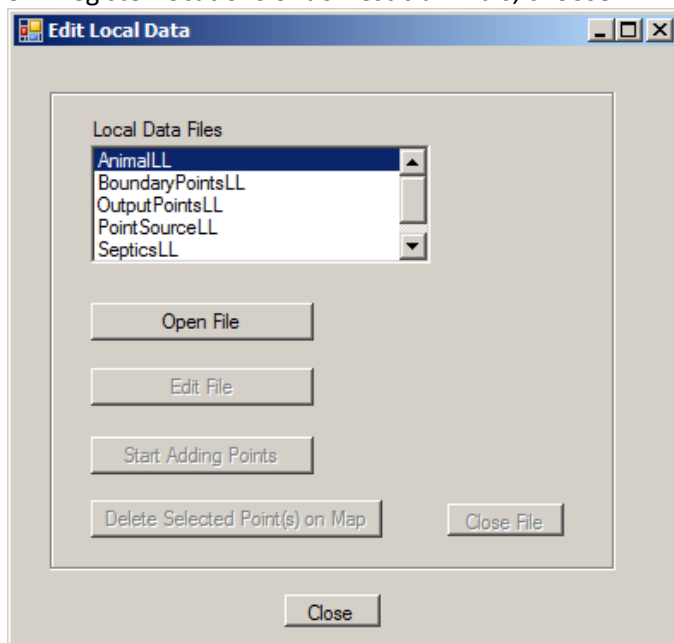
4. The locations contained in the AnimalLL.csv file, found in the "LocalData" folder, will be registered as a map layer, as illustrated below.
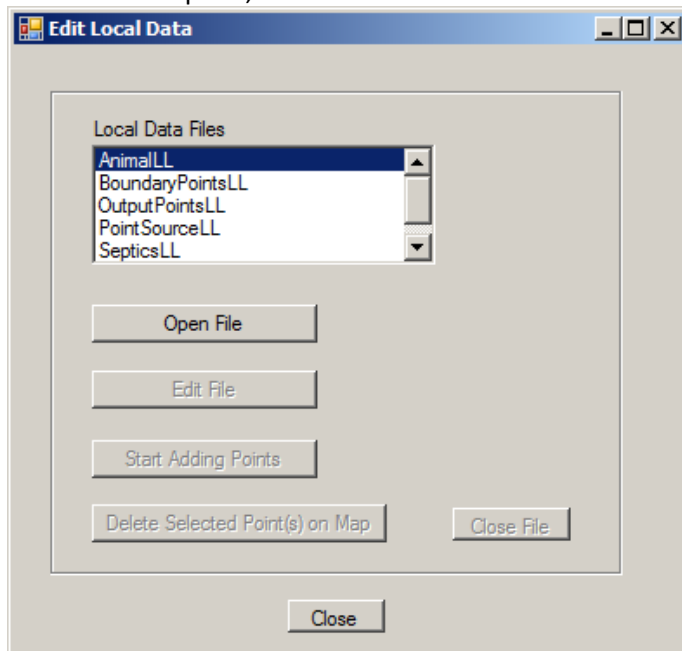


5. Choose the "Close File" button to exit the "AnimalLL" map layer registration.



6. **Repeat for the other relevant local data files: BoundaryPointsLL, OutputPointsLL, PointSourcesLL, and SepticsLL.**

7. When complete, click the "Close" button at the bottom of the screen to exit the screen.



Although the software will allow the user to "Edit File", "Start Adding Points", and "Delete Selected Point(s) on Map", these features have not been fully tested. It is recommended that the user directly modify the csv files using tools such as Excel, Notepad, or Textpad.

**SPECIAL NOTES ON IMPORTING POINT SOURCE TIME SERIES DATA**

This section discusses conditions for importing point source time series data through HSPF and BASINS, replacing the default data illustrated in Tables 8 and 9. It is assumed that the reader is familiar with SDMPB screens documented in various SDMPB tutorials (e.g., Whelan et al., 2017d). Whelan et al. (2017e) provides a tutorial for importing point source time series for flow and microbial loadings.

For HSPF to correctly model microbial and/or chemical point source loadings, the following conditions must be met:

1. For microbes, the "Microbes" text box in the SDMProjectBuilder user interface (Figure 6) **must** be checked. Because a microbial loadings and chemical point source loadings are different from land-applied chemical loadings, the check box for "Land-Applied Chemical" is irrelevant (Figure 6).
2. For point sources, the text box in the HSPF "Point Sources" window must be checked, and "YES" must be selected beside "MICROBES" in the "Details of POINT SOURCES (OBS)" section (Figure 7). To get to the HSPF "Point Sources" window, select "Functions" in the main HSPF menu bar, then "Point Source Editor" (Figure 8). If chemicals are being assessed, then the chemical would also have to have "Yes" by its name.
3. PointSourceLL.csv default template must be updated with the correct Latitude-Longitude.
4. The correct time series will have to be imported to the WDM file using BASINS.

If condition 1, 2, or 3 is NOT met, point sources will not be considered in the assessment.
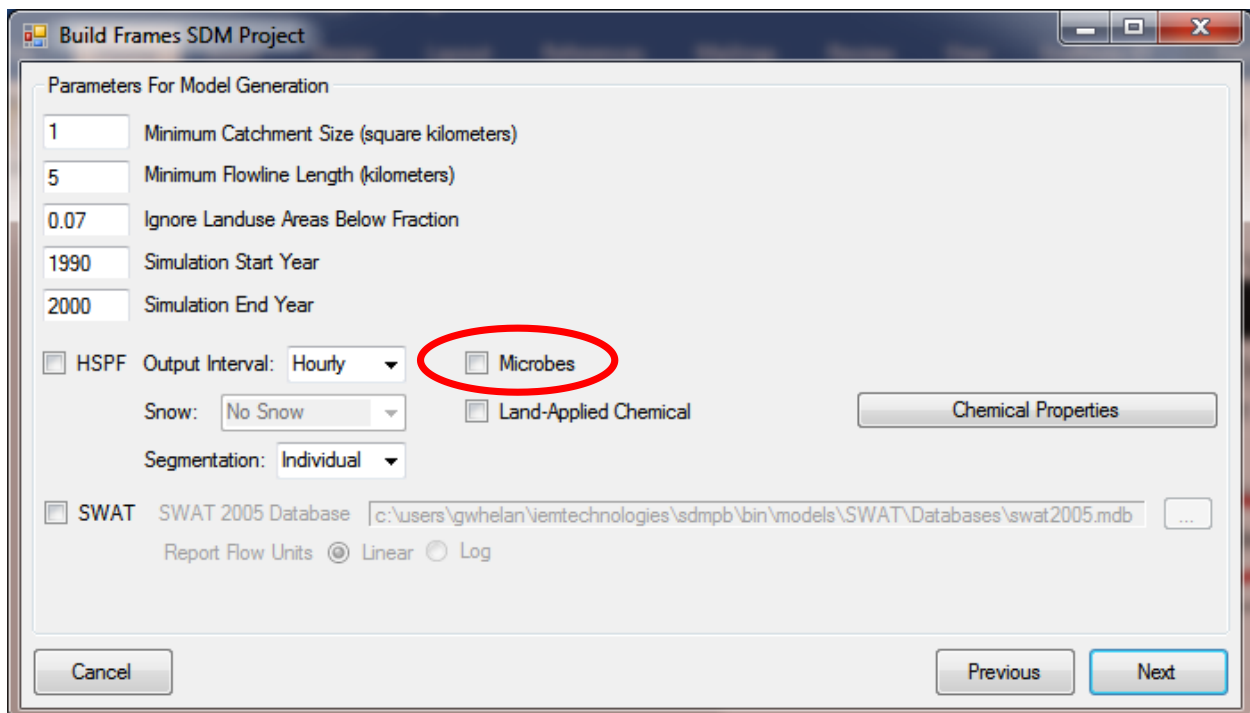


**Figure 6.** Screen capture of the SDMProjectBuilder interface, indicating a microbial assessment will be performed
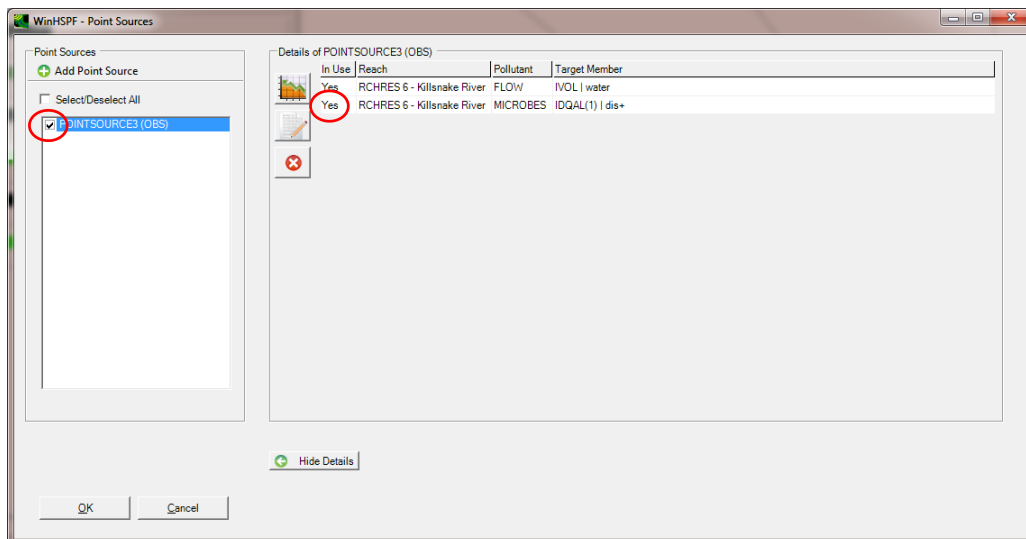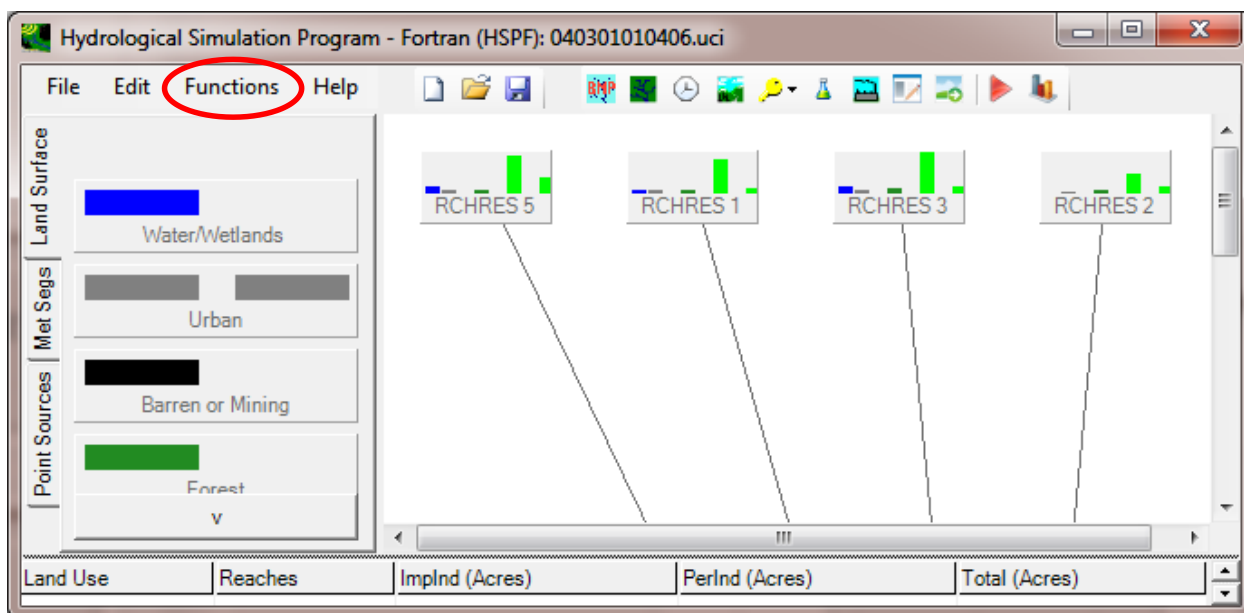
**Figure 7**. HSPF Point Sources window



**Figure 8**. Main HSPF screen highlighting "Functions" on the menu bar

Modifications to the PointSourceData.csv template update only a single annual-average value for flow, and microbial and chemical loading rates. Default values contained in the file are very small and typically will not impact any simulations, even if they are incorrectly included in the assessment. The file should not be directly modified to include site-specific, time-series data. The actual time series of flows, and microbial and chemical loading rates for each location can only be recorded through the HSPF user interface. A tutorial has been developed to show the user how to import point sources into an assessment; see Whelan et al. (2017e).

**DISCLAIMER**

The information in this document has been funded in part by the United States Environmental Protection Agency (EPA) under EP-C-12-021 to AQUA TERRA Consultants, a Division of RESPEC, INC. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document.

**ACKNOWLEDGMENTS**

**REFERENCES**

EPA (U.S. Environmental Protection Agency). 2013. Data for Environmental Modeling (D4EM). Office of Research and Development, Athens, GA. http://www.epa.gov/AthensR/research/d4em.html (last accessed 07.10.16).

Garnett, L.W. 2012. 5 High density housing concepts. <http://www.probuilder.com/5-high-density-housing-concepts> (Last accessed 11.9.2015).

MapWindow. 2013. MapWindow 6. http://mapwindow6.codeplex.com/ (last accessed 07.10.16).

Watry, G., Ames, D.P., 2008. A practical look at MapWindow GIS. First Edition, 316 p.

Whelan, G., K. Kim, K. Wolfe, R. Parmar, M. Galvin. 2017a. Quantitative Microbial Risk Assessment Tutorial: Installation of Software for Watershed Modeling in Support of QMRA – Updated 2017. EPA/600/B-15/276, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.

Whelan, G., R. Parmar, G.L. Laniak. 2017b. Microbial Source Module (MSM): Documenting the Science and Software for Discovery, Evaluation, and Integration; Updated – 4/17/17. EPA/600/B-15/315, U.S. Environmental Protection Agency, Office of Research and Development, Athens, GA.

Whelan, G., K. Kim, K. Wolfe, R. Parmar, M. Galvin, M. Molina, R. Zepp. 2017c. Navigate the SDMPB and Identify an 8-Digit HUC of Interest – Updated 2017. EPA/600/B-15/273, U.S. Environmental Protection Agency, Athens, GA.

Whelan, G., K. Kim, R. Parmar, K. Wolfe, M. Galvin, M. Gray, P. Duda, M. Molina, R. Zepp. 2017d. Quantitative Microbial Risk Assessment Tutorial: Land-applied Microbial Loadings within a 12-Digit HUC – Updated 2017. EPA/600/B-15/298, U.S. Environmental Protection Agency, Athens, GA.

Whelan, G., K. Wolfe, R. Parmar, M. Galvin, M. Molina, R. Zepp, P. Duda. 2017e. Quantitative Microbial Risk Assessment Tutorial: Point Source and Land-applied Microbial Loadings within a 12-Digit HUC. U.S. Environmental Protection Agency, Athens, GA.

Whelan, G., K. Kim, K. Wolfe, R. Parmar, M. Galvin, M. Molina, R. Zepp, P. Duda, M. Gray. 2017f. Quantitative Microbial Risk Assessment Tutorial: Pour Point Analysis of Land-applied Microbial Loadings and Comparison of Simulated and Gaging Station Results – Updated 2017. EPA/600/B-15/290. U.S. Environmental Protection Agency, Athens, GA.

Wikipedia. 2015a. Single-family detached home. <https://en.wikipedia.org/wiki/Single-family_detached_home> (Last accessed 11.9.2015).

Wikipedia. 2015b. <https://en.wikipedia.org/wiki/Multi-family_residential> (Last accessed 11.9.2015).

Wolfe, K., R. Parmar, G. Whelan, G.F. Laniak, M. Galvin, K. Kim, M. Molina, Z. Zepp, P. Duda, D.D Keiser. 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.