Quantitative Microbial Risk Assessment Tutorial Publishing a Microbial Density Time Series as a TXT File

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

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

10/10/15

Summary

A SARA Timeseries Utility supports analysis and management of time-varying environmental data including listing, graphing, computing statistics, computing meteorological data and saving in a WDM or text file. File formats supported include WDM, HSPF Binary (.hbn), USGS RDB, and Text (txt) files (using the same import scripts as WDMUtil and BASINS). WDMUtil users will find that most of its functionality is included in this utility. In particular, the SARA Timeseries Utility is being used to publish a time series of microbial densities that can be viewed or consumed by other components, such as the Microbial Risk Assessment-Interface Tool (MRA-IT), which is a MathCad-based, integrated software tool based on the pathogen of interest, exposure, intake, and dose. MRA-IT consumes txt files using a Unicode Transformation Format using 8-bit code units (UTF-8), which is a character encoding capable of encoding all possible characters in Unicode using 8-bit code units. The SARA Timeseries Utility tool produces UTF-16, which is a character encoding capable of encoding all possible characters in Unicode using 16-bit code units. The objective of this effort was to modify the SARA Timeseries Utility to produce txt files in a UTF-8 format, which contributes to a seamless communication with components such as MRA-IT. This tutorial reviews some of the screens, icons, and basic functions of the SARA Timeseries Utility tool, which allows users to publish Hydrological Simulation Program – FORTRAN (HSPF) data as a UTF-8 txt file.

Publishing a Microbial Density Time Series as a TXT File

PURPOSE

Publish a microbial density time series as a Unicode Transformation Format using 8-bit code units (UTF-8) file.

OBJECTIVE

The objective is to publish a time series of microbial densities that can be viewed or consumed by other components, such as the Microbial Risk Assessment-Interface Tool (MRA-IT), which is a MathCad-based, integrated software tool based on the pathogen of interest, exposure, intake, and dose (Whelan et al., 2014; Soller et al., 2008, 2004). A C++ version of MRA-IT also exists, called the Irrigation and Recreation Risk Assessment (IRRA) module (Pelton, 2015).

DEMONSTRATION

The SARA Timeseries Utility supports analysis and management of time-varying environmental data including listing, graphing, computing statistics, computing meteorological data, and saving in a WDM or text file (Aqua Terra, 2011). File formats supported include: WDM, HSPF Binary (.hbn), USGS RDB, and Text files (using the same import scripts as WDMUtil and BASINS). WDMUtil users will find that most of its functionality is included in this new utility. This tutorial reviews some of the screens, icons, and basic functions of the SARA Timeseries Utility tool (Aqua Terra, 2011), which allows users to publish HSPF data as a txt file.

SOFTWARE ACCESS, RETRIEVAL, AND DOWNLOAD

Instructions for access, retrieval, and download of the SARA Timeseries Utility software product are provided under the heading, titled SARA Timeseries Utility, in Whelan et al. (2015), Installation of Software for Watershed Modeling in Support of QMRA. When installed, a shortcut icon should now be on your computer screen for SARA Timeseries Utility:



TUTORIAL: USING THE SARA TIMESERIES TOOL TO SAVE DATA FROM AN HSPF MODEL AS A TEXT FILE

The SARA Timeseries Utility tool is an independent software tool that can read watershed WDM files created by HSPF and convert WDM binary information to provide multiple types of plots including txt files. Typically, a watershed WDM file is created when the user executes HSPF, and the file exists within the working folder (e.g., <working folder>\HSPF\). For example, if the user's working folder is titled "TestA", as in the example below, then the watershed WDM (e.g., 040301010406.WDM) would be located in the HSPF folder:

Select WDM Time Series file to open		Section of the local division of the local d	and the second second	Street and	×
	s ▶ S	DMProjectBuilder_100214 TestA HSPF	-	Search HSPF	٩
Organize 🔻 New folder					• 🔳 🔞
SDMProjectBuilder_100214	*	Name	Date modified	Туре	Size
bin 🕒 data		🖪 040301010406.wdm	10/6/2014 6:08 PM	WDM File	1,400 KB
I TestA		🖾 met.wdm	10/6/2014 6:04 PM	WDM File	6,920 KB
HSPF					
huc12	-				
🕌 met	=				
📔 pcs					
Jan TestB					
🍐 TestC	-				
File <u>n</u> ame: 040301010406.wdm			•	WDM Files (*.wdm) 🔻
				<u>O</u> pen	Cancel

C:\Users\gwhelan\iemTechnologies\SDMProjectBuilder_100214\TestA\HSPF\

Note that there may be multiple WDM files, and the correct one has a watershed ID designation (e.g., 040301010406) as illustrated above.

For simplicity, a WDM file was created *a priori* (i.e., 040301010406.WDM), and it accompanied the software download described in Whelan et al. (2015): "040301010406.wdm". For the sake of this tutorial, it is assumed that the WDM is located in

C:\Users\gwhelan\iemTechnologies\SDMProjectBuilder_100214\TestA\HSPF\

although in actuality, this particular file is located in the folder designated by the user during software installation.

1. Run SARA Timeseries Utility, using the icon from the desktop, or double click the executable, titled TimeseriesUtility_2015-08-13.exe, located in the folder where the user saved the file, when the software was originally downloaded. See the example below, which is consistent with the information contained in the software download tutorial (Whelan et al., 2015).

● ビン・クロール 単分 かかく			
G → Lomputer → OSDisk (C:) → Users → gwh	nelan	Utility	
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>T</u> ools <u>H</u> elp			
Organize 🔻 Include in library 👻 Share with 👻 E	Burn New folder		
Links MaxStat Lite MicrobialPropertyDB MicrobialPropertyDB	Name ■ 040301010406.wdm ☆ SARA_Timeseries_Utility_2015	Date modified 10/6/2014 6:08 PM 8/13/2015 11:31 AM	Type Size WDM File 1,400 KB Application 2,768 KB
My Music My Pictures My Videos	👯 😨 😑 💿 🗢 WinZip - SARA_Times_Seri	es_Utility.zip	
 SametimeTranscripts Saved Games Saveder 	Folder name: SARA_Times_Series_Utility		Opgrade Opgrade Opgrade Opgrade Opgrade Opgrade
June Scale Has Time Series Utility	Tg- Location: C:\Users\gwhelan\TimeSerie Unzip	Utility Unzip Dropbox Google S Drive Unzip to Clou	kyprive Search O Selected Files
SARA_Times_Series_Utility	Name SARA_Timeseries_Utility_2015-08-13.exe A	type Modified Application 8/13/2015 11:31 AM NDM File 10/6/2014 6:08 PM	5/28 2,834,229 1,433,600
	Selected 0 files, 0 bytes	Total 2 files, 4,168KB	

The following screen appears:

😤 SARA Timeserie	es Utility		
Open File	Manage Files No	files are open	
Select Timeseries	No Timeseries are selecte	d	
View	Save	Compute	KIVER AUTHORITY
List	Save As Text	Meteorologic	
Graph	Save As Data	Math	
Tree	Import Text to WDM		About Help

2. Choose the 'OpenFile'

¥	🗧 SARA Timeserie	es Utility		
6	Open File	Manage Files No	files are open	
	Select Timeseries	No Timeseries are selecte	d	SAN ANTONIO
	View	Save	Compute	RIVER AUTHORITY
	List	Save As Text	Meteorologic	
	Graph	Save As Data	Math	
	Tree	Import Text to WDM		About Help

3. Select the 'WDM Time Series'; choose 'Ok'.



4. In the HSPF folder within your working folder, locate the WDM file referenced to the watershed ID. In this example, it is the WDM that accompanied the install (i.e., 040301010406.wdm). For simplicity, it is assumed that the WDM file is located within the working folder "TestA":

C:\Users\gwhelan\iemTechnologies\SDMProjectBuilder_100214\TestA\HSPF\

and the WDM file is 040301010406.WDM, as noted in the figure below:

📩 Select WDM Time Series file to open					
😋 🔵 🗢 🕌 « Users 🕨 gwhelan 🕨 iemTechnologi	es ▶ S	DMProjectBuilder_100214 TestA HSPF	•	Search HSPF	٩
Organize 🔻 New folder				8==	• 🔳 🔞
SDMProjectBuilder_100214	*	Name	Date modified	Туре	Size
bin		B 040301010406.wdm	10/6/2014 6:08 PM	WDM File	1,400 KB
lestA		🖻 met.wdm	10/6/2014 6:04 PM	WDM File	6,920 KB
HSPF					
huc12	=				
NHDPlus					
NLCD					
\mu pcs					
JestB					
Ja Testo	~				
File <u>n</u> ame: 040301010406.wdm				WDM Files (*.wdm	n) 🔻
				<u>O</u> pen	Cancel

5. Click 'Open', and the name of the file becomes registered:

🚖 SARA Timeseri	es Utility		
Open File	Manage Files C:	\Users\gwhelan\iemTechnologies\SDMProjectBuilder_1	00214\TestA\HSPF\040301010406.wdm
Select Timeserie	s No timeseries are selecte	d our of 51	
List	Save As Text	Meteorologic	
Graph	Save As Data	Math	
Iree			About Help

6. Choose 'Select Timeseries',

😤 SARA Timeserie	s Utility		
Open File	Manage Files No	files are open	
Select Timeseries	No Timeseries are selecte	d	
View	Save	Compute	KIVER AUTHORITY
List	Save As Text	Meteorologic	
Graph	Save As Data	Math	
Tree	Import Text to WDM		About Help

7. A screen similar to the following will appear:

😤 Select Timeseries		
File Attributes Select Hel	р	
Select Attribute Values to Filter Availab	le Data	
Scenario 💌	Location -	Constituent
04030101	IMP102	DQAL
	IMP202	FLOW
	PER101	PERO
	PER102	POQC
	PER103	RO
	DED 104	5000
Matching Data (51 of 51)		
04030101	RCH2	FLOW
04030101	PER101	PERO
04030101	PER101	POQC
04030101	PER102	PERO
04030101	PER102	POQC
04030101	PER103	PERO
04030101	PER103	POQC
04030101	PER104	PERO +
Selected Data (1 of 51)		
04030101	RCH2	DQAL
Dates to Include		
All Common		
Start 2000/01/01 2000/01/01	2000/01/01	
End 2000/12/31 2000/12/31	2000/12/31	
Apply month/day range to each ye	ar	
Change Time Step To: 1	Day	Ok Cancel

8. To add an attribute, such as the Time Unit (e.g., hour, day, month, etc.), choose 'Attributes', then 'Add'. Go to the down arrow, and choose 'Time Unit'. This is not a necessity, but it allows to user to ensure that they are choosing the data with the correct units. In this case, hourly output is needed.

Select Attributer values I	o Filter Available Data			
Scenario	 Location 	 Constituent 	•	Data Source
04030101	IMP102	 DQAL 	<u>~</u>	Data Source
	IMP202	FLOW	-	Description
	PER101	PERO	=	ID
	PER102	POQC		Interval
	PER103	RO		Point
	000104	T 0000	•	STANAM
Matching Data (51 of 51)			TGROUP
04030101	RCH2	FLOW		TSBYR
04030101	PER101	PERO		TSFORM
04030101	PER101	POQC		TSTYPE
04030101	PER102	PERO		Time Step
04030101	PER102	POQC		Colordated Attributers
04030101	PER103	PERO		Coefficient of Variation
04030101	PER103	POQC		Count
04030101	PER104	PERO		Count Missing Count Positive
				Count Zero
Selected Data (1 of 51)				Geometric Mean
04030101	RCH2	DQAL		Last
				MaxDate
				Mean
	MinDate			
Dates to Include	Serial Correlation Coefficient			
All	Skew Standard Deviation			
Start 2000/01/01	2000/01/01 200	0/01/01		Standard Error of Skew
End 2000/12/31	2000/12/31 200	0/12/31		Start Date
Apply month/day ra	nge to each year			SumAnnual

9. Within the 'Select Attribute Values to Filter Available Data' section, under the 'Constituent' heading, choose 'DQAL', as this is the constituent that provides the microbial density time series. Choose the location of interest under the 'Matching Data' section (i.e., click on 'RCH2'). In this case, we choose 'RCH2', this location appears in the 'Selected Data' section.

🚖 Select Timeseries			
File Attributes Sele	ct Help		
Select Attribute Values to Filte	er Available Data		
Scenario -	Location -	Constituent -	Time Unit 🔹
04030101	IMP102	DQAL	Hour
	IMP202	FLOW	Day
	PER101	PERO	
	PER102	POQC	
	PER103	RO	

Matching Data (8 of 51)			
04030101	RCH1	DQAL	Hour
04030101	RCH6	DQAL	Hour
04030101	RCH7	DQAL	Hour
04030101	RCH5	DQAL	Hour
04030101	RCH4	DQAL	Hour
04030101	RCH3	DQAL	Hour
04030101	RCH8	DQAL	Hour
04030101	RCH2	DQAL	Hour
Selected Data (1 of 51)			
04030101	RCH2	DQAL	Hour
Detecto la chude			
Start 2000/01/01 2000	/01/01 2000/01/01	_	
E-1 2000/01/01 2000	(12/21 2000/01/01	_	
Ena 2000/12/31 2000/	2000/12/31		
Apply month/day range to	o each year		
Change Time Step To:	1 Day	rage/Same 🔻	Ok Cancel

If the user wants to specify a subset of the time series, they can modify the 'Start' and 'End' dates located in the lower most section, titled 'Dates to Include'. In the example provided, the time series extends over one year (i.e., 2000). If the user is interested in a particular day or weekend (e.g., when a beach is being used), this is the place where the user chooses which days to assess. In this particular example, we would like to assess a three day exposure period from March 1, 2000 through March 3, 2000.

🛣 Select Timeseries	🛧 Select Timeseries 📃 🗕 📼					
File Attributes Sele Select Attribute Values to Fil	ect Help ter Available Data					
Scenario 👻	Location -	Constituent -	Time Unit 🔹			
04030101	IMP102	DQAL 🔺	Hour			
	IMP202	FLOW	Day			
	PER101	PERO				
	PER102	POQC				
	PER103	RO				
		500C				
Matching Data (8 of 51)						
04030101	RCH1	DQAL	Hour			
04030101	RCH6	DQAL	Hour			
04030101	RCH7	DQAL	Hour			
04030101	RCH5	DQAL	Hour			
04030101	RCH4	DQAL	Hour			
04030101	RCH3	DQAL	Hour			
04030101	RCH8	DQAL	Hour			
04030101	RCH2	DQAL	Hour			
Selected Data (1 of 51)						
04030101	RCH2	DQAL	Hour			
Dates to Include						
All Common						
Start 2000/01/01 2000/01/01 2000/03/01						
End 2000/12/31 2000/12/31 2000/03/03						
Apply month/day range to each year						
Change Time Step To:	1 Day - Av	erage/Same 🔻	Ok Cancel			

10. Select 'OK'. The selected time series is registered next to the label 'Select Timeseries':

T SARA Timeserie	es Utility		-
Open File Select Timeserie	Mapage Files C 04030101 RCH2 DQAL	Users \gwhelan \em Technologies \SDMProjectBuilder_100214\TestA\HSPF\040301010406.wdm	
View List Graph	Save Save As Text Save As Data	Compute Meteorologic Math	
Tree	Import Text to WDM	About Help	

11. Select 'List'.

gwhelan\iemTechnologies\SDMProjectBuilder_100214\TestA\HSPF\040301010406.wdm
ipute
Meteorologic
Math
About Help

and a 'Timeseries List' appears:

🚖 Time Se	ries List 💷 💷 💻	×
File Edit	View Analysis Help	
History 1	from 040301010406.wdm	
Constituent	DQAL	
ld	151	
Min	9,203.1	
Max	21,029	=
Mean	15,126	
	10173	
	10275	
	10000	
	9377.3	
	9315	
	9203.1	
	9303.4	
	9489.6	
	9734.9	
	10014	
	10320	
	10656	
	11024	
	11427	
	11853	
	12250	
	12532	
	12712	
	12879	-
	13030	-
	13164	-
	13280	-
	13378	

Note that the information between the pull-down menus and the data is denoted as "Attributes" (e.g., History, Constituent, Id, Min, etc.). The view may also appear as follows, because the last settings from your machine are saved, so one screen capture in this tutorial will not necessarily capture everyone's view.

🚖 Timeseries List — 🗖 🗙						
File Edit View Analysis Help						
History 1	from 040301010406.wdm	from 040301010406.wdm				
Constituent	RO	DQAL				
ld	138	139				
Min	0.0061348	103.36				
Max	40.172	329,280,000				
Mean	2.7202	12,202,000				
1990/01/01 01:00	14.893	103.36				
1990/01/01 02:00	11.199	109.61				
1990/01/01 03:00	8.4653	118.88				
1990/01/01 04:00	6.44	131.4				
1990/01/01 05:00	4.9386	147.29				
10000010101.00.00	2.0242	100 C				

12. The output that will be saved to file can be customized. Because we will use this example to cater to the input needs of the MRA-IT model, it requires only a listing of the microbial densities without any other identifying information, such as headers or the times associated with each density. To produce such a txt file, choose 'File', then 'Select Attributes':

File Edit View
Analysis Help
10173
10275
10000
9377.3
9315
9203.1
9303.4
9489.6
9734.9
10014
10320
10656
11024
11427
11853
12250
12532
12712
12879
13030
13164
13280
13378

13. Deselect all Attributes by clicking on the highlighted ones, producing:

🚰 Select		x
Coefficient of Variation		A
Constituent		
Count		
Count Missing		
Count Positive		
Count Zero		
Data Source		
Description		
End Date		=
Geometric Mean		
History 1		
ld		
ID I		
Interval		
Last		
Location		
Max		
MaxDate		
Mean		
Min		
MinDate		
Point		
Scenario		
Serial Correlation Coefficient		
Skew		*
All None Ok	Car	ncel

14. Choose 'OK'. Resulting in all Attributes being removed. It just so happens that this screen capture produces the same results as the previous one.

File Edit View	٦
Analysis Help	
10173	^
10275	
10000	
9377.3	_
9315	
9203.1	
9303.4	
9489.6	
9734.9	
10014	
10320	
10656	
11024	
11427	
11853	
12250	
12532	
12712	
12879	
13030	
13164	
13280	-
13378	

15. Choose 'View', then 'Date and Value Formats' to select/deselect formatting information. One may or may not receive a screen that looks like the following:

☆ Date and Va	lue Formats		×
Date Format Year Month Day Month Day Year Day Month Year Days starting 1900 Years Two Digit Years Months Month Names Days Hours Minutes Seconds Date Separator Image Separator Time Separator Help Cancel Apply Apply	Value Format Standard Format Exponent Format Significant Digits Maximum Length If Value Cannot Fit	##0.######## 5 10 #	

16. Remove the comma in the 'Standard Format', if it exists. If the time formats (e.g., Years, Months, etc.) have been checked, deselect them by clicking on them. Finally, remove the title.

☆ <u>- □ × </u>
File Edit View
Analysis Help
10173
10275
10000
9377.3
9315
9203.1
9303.4
9489.6
9734.9
10014
10320
10656
11024
11427
11853
12250
12532
12712
12879
13030
13164
13280
13378

17. Click 'Apply', then 'Ok', and the commas will be removed as well as Attributes:

- 18. When satisfied with the selected data, attributes, and format, Choose 'File', then 'Save Changes'.
- 19. Close by clicking on the X in the upper right hand corner.
- 20. Choose 'Save as Text'

SARA Timeseries Utility	
Open File Manage Files No files are open	$\overline{\mathbf{A}}$
View Save Compute	S A N A N TO N LO RIVER AUTHORITY
List Save As Text Meteorologic	
Graph Save As Data Math	
Tree Import Text to WDM	About Help

21	Create a	name for the	file and	Save it to	your working	folder
Z I.	Create a	name for the	: me, anu	Javentio	your working	ioiuei.

Save as		Contraction of the local division of the loc	×
🔾 🖓 🖓 🐨 Users 🕨 gwhelan 🔸 iemTechnologies 🕨 SDMProjectBuilder_10021	14 • TestA • HSPF 🔹	€ Search HSPF	Q
Organize 🔻 New folder		:= -	0
SARATimeSeriesUtility	Name	Date modified	Туре
SDMPB SDMProjectBuilder_100214 bin data	TestAOutput.txt	10/12/2014 12:17 8/11/2015 10:44 AM	Text [Text [
TestA HSPF huc12			4
File name: TestAOutput_Rev2			-
Save as type: Text Files (*.bxt)		Save	• •

After carefully choosing the format the first time using the 'Timeseries List' form, the 'Save As Text' button on the main window can be used to save other data using the same format without having to open the 'Timeseries List' form again.

22. Close by clicking on the X in the upper right hand corner. This file is ready to be consumed by other modules. In this case, by the MRA-IT risk assessment module.

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication.

REFERENCES

Aqua Terra. 2011. SARA Timeseries Utility <<u>http://www.aquaterra.com/resources/downloads/saratsutility.php</u>)> (last accessed 12.08.15).

Pelton, M.A. 2015. IRRA Requirements, Design and Specifications Documentation. Battelle, Pacific Northwest Division, Richland, WA (Draft)

Soller, J.A., Olivieri, A.W., Eisenberg, J.N.S., Sakaji, R., Danielson, R., 2004. Evaluation of Microbial Risk Assessment Techniques and Applications, 00-PUM-3. Water Environmental Research Foundation, Alexandria, VA.

Soller, J.A., Seto, E., Olivieri, A.W., 2008. Microbial Risk Assessment Interface Tool: User Documentation. Water Environmental Research Foundation, Alexandria, VA.

Whelan, G., K. Kim, K. Wolfe, R. Parmar, M. Galvin. 2015. Quantitative Microbial Risk Assessment Tutorial: Installation of Software for Watershed Modeling in Support of QMRA. EPA/600/B-15/276. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.

Whelan, G., K. Kim, M.A. Pelton, J.A. Soller, K.J. Castleton, M. Molina, Y. Pachepsky, J. Ravenscroft, R. Zepp. 2014. An Integrated Environmental Modeling Framework for Performing Quantitative Microbial Risk Assessments. *Environ Modell Softw*, 55:77-91.