Modeling Integration Package

Linking SWMM to Goldsim

Chris York, Steven Switzer, Travis Christensen, Jeff Huber, and Stephen Daras

Dr. Steven Burian- CVEEN 7970

Department of Civil & Environmental Engineering

University of Utah

Salt Lake City, USA

Abstract— **A need to link SWMM (Storm Water Management Model) and GoldSim together was identified in the Urban Water Research Group at the University of Utah. This is an effort to link model results together to answer a greater number of research questions and enhance modelling results by using multiple software programs. The most efficient way to link SWMM and GoldSim was determined to be through a central database. Once model results were stored in the database, the values could then be shared with both programs. Additional analysis was performed using R and model results were shared via HydroServer.**

I. INTRODUCTION

The Problem

An opportunity was identified within the Department of Civil and Environmental Engineering at the University of Utah, by students in the Urban Water Research Group. The goal of the project was to determine the most efficient way to connect two software programs to utilize the benefits and capabilities of two separate models. The objective was then to make this process easily repeatable and available to apply to any research question. As such, this report will provide guidance as to how the process was optimized and instructions so that a researcher could use this report and apply it to their own research questions.

The two programs that will be highlighted in this report are SWMM and GoldSim. The report provides basic steps so that the process could be repeated using different software other than SWMM, if this program does not fit research needs.

There are more detailed steps that provide in-depth detail as to how the linkage was achieved, so that the process could be replicated by other researchers with a common goal.

Background

The Jordan River is a large watershed that encompasses the Salt Lake Valley. With the future water supply and demand in the Salt Lake Valley in question because of population and climate changes, a need for an integrated model has been identified. SWMM is a dynamic rainfall-runoff simulation model that can model quality and quantity of stormwater. Linking the modelling results from SWMM to a modelling platform such as GoldSim will allow researchers to link results that will ultimately aid in more accurate modelling, as well as provide increased analyses for various research projects.

The process that the group optimized was centralizing a database in Microsoft SQL (MS SQL) that stores data from the Jordan River collected by United States Geological Survey (USGS) stream gauges along the river. This data was used to create a simple GoldSim model that models the flow in the Jordan River and downstream wetlands. This model is a downscaled version of an existing model created by a researcher in the Urban Water Research Group.

A SWMM model of Salt Lake County was also obtained from a researcher in the Urban Water Research Group. The model tracks a rainfall event over a 24-hour period and displays results as

volume of water that discharges to the Jordan River as a result of the storm.

Overview

The goal was to integrate two modelling tools with a database. The two software programs that the group wanted to integrate were SWMM and GoldSim. A third link to Microsoft SQL was established as an intermediate step to store historical data and model results. Further analysis was then performed using R Statistical Software that GoldSim was not capable of, such as statistical analysis and exceedance probability. Finally, a HydroServer was established so that model results could be shared publicly. A schematic detailing this process can be seen in Figure 1 below:
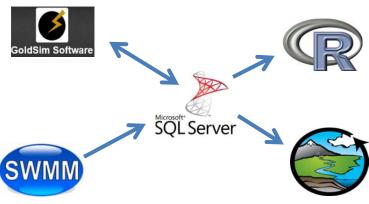

Figure 1: Schematic of the Modeling Integration Package

Results from SWMM were output into a text file and converted to a CSV (Comma Separated Value) file. These results were then imported into MS SQL. MS SQL was then linked to GoldSim where the stormwater runoff results could be brought into the dynamic model and linked to the Jordan River Hydraulic Model. These results could be viewed using GoldSim, however, the tools available for further analysis were limited. R was employed to perform further statistical analysis and to create exceedance plots to understand differences in historical and model results. Finally, HydroServer was used to make the historical data and SWMM results available so that this process can easily be reproduced.

II.   SWMM AND GOLDSIM

Description of SWMM and GoldSim

Using results from SWMM in a GoldSim model can be a useful method to incorporate modeled results into another model. In order to do this, previously developed SWMM and GoldSim models were used. The SWMM model contains 27 subcatchments that flow to the Jordan River. The GoldSim model uses four data elements representing the Lower Jordan River Basin, the Surplus Canal Basin, the Sewage Canal Basin, and an element representing the modeled SWMM results.

Connecting GoldSim to MS SQL

GoldSim allows for the use of a database connection as a data source. The type of database selected was a MS SQL database on a SQL server. Although it is possible to use a Microsoft Excel file as a data source, the SQL server was selected due to the simplicity with which it is able to be used by third-party software such as R and HydroDesktop. Generally, servers are more robust and accessible by multiple users than stand-alone files like Excel spreadsheets.

SQL Express and SQL Server Management Studio 2012

By installing Microsoft SQL Server 2012 Express (a free version of Microsoft SQL Server that contains all the necessary functionality), a table can be created that will hold the SWMM results and be accessed by GoldSim through a database connection. Microsoft SQL Server Management Studio 2012 can also be installed to facilitate importing the formatted Excel file of SWMM results. Immediately after running Microsoft SQL Server Management Studio, the user is prompted to create a connection to a SQL server. With SQL Server Express installed, the connection can be made by entering the following information.

Server name: [*COMPUTERNAME*]\SQLEXPRESS
Authentication: Windows Authentication

Although the SQL server is setup, GoldSim cannot connect directly to it; accessing the server must be done through an ODBC connection (Open Database Connectivity)[1].

The ODBC connection

In a general sense, ODBC is middleware that allows software to connect to a database management system such as MS SQL. It eliminates the need for software to know how to "talk" with individual types of servers and acts as a translator between the software and server. GoldSim therefore receives the SWMM results that are in the SQL server by accessing the ODBC instead of the SQL server [1]. It is worth noting that although ODBC data sources can be set up through Windows' Control Panel, the ODBC will not be recognized by GoldSim unless it is set up by navigating to C:\Windows\System32 and running the executable file "odbcad32.exe" (file path may vary). Running odbcad32.exe guides the user through the Microsoft SQL Server DSN Configuration. Input data for each setup window are found in Appendix I.

Getting Data from MS SQL to GoldSim

As mentioned, incorporating the SWMM results to GoldSim is accomplished by accessing the previously created ODBC [1]. To do this, a new element must be created in the GoldSim model. Figure 2 shows the full model with the new data element (Data1) linked to an expression element (Expression1) which links to the Jordan River inflows element.
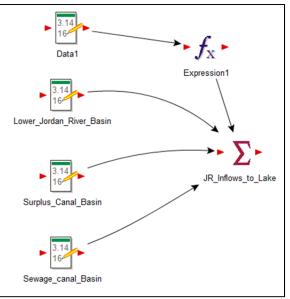

Figure 2: SWMM results connected to the GoldSim model through the Data1 element

The data source connection is accomplished by right-clicking the Data1 element and selecting the Definition tab. The Display Units must be the same units as the SWMM runoff results [2]. Select the Type button then set the Type to Value, and Order to Matrix 2D-Array. Next create new row labels equal to the number of rows of data in the SWMM results table, and create new column labels with names equal to the column headers in the SQL table that contains the SWMM results. Set the Data Source Type to Generic Database. In the Database tab find the database name in the drop-down menu. After doing so the Table field in the Generic Database Definition will be populated with the table names in the database. Set incoming units to cubic feet per second. Click the Define Fields button and add the Days and Runoff options. In the Where text field, enter Days = '1' (or any option the user desires) to define how many days from the SWMM results that the GoldSim model will evaluate. To download and use the data in the model select Download Now.

Creating an expressions element allows the downloaded data in the Data1 element to be used in the Jordan River inflow element. The expressions element is also necessary because it can process large amounts of data by evaluating the Data1 element with an elapsed time expression. After the

data was successfully downloaded, the GoldSim model was run to incorporate the SWMM results.

Results from the GoldSim Model

Running the GoldSim model produces a summary of contributions from each data element in KAF/yr (thousands of acre-feet per year). Evaluating one day of the SWMM results in the following table of data:

- Lower_Jordan_River_Basin: 17 KAF/yr
- Surplus_Canal_Basin: 20 KAF/yr
- Sewage_Canal_Basin: 9 KAF/yr
- Expression1: 23216.5 KAF/yr

Expression1 overwhelms the rest of the results because it is taking runoff from a significant storm for the entire Salt Lake County and loading it into one location along the Jordan River. The other values are simply averages for inputs that comprise the Lower Jordan River. In reality, a significant amount of water would infiltrate or meet a different fate that was not modeled. Model results can be optimized based on changes in the model specific to the users needs.

III. R CODING

R was employed to perform further statistical analysis and to create exceedance probability plots to understand differences in historical and model results.

Statistical analysis that can be performed in R includes determining the mean, median, minimum, and maximum. The variance, standard deviation, and the normal distribution are other useful functions to know for a statistical analysis. The mean or the sample mean is many times interchangeable with the average and is important in many analyses because it generally is a good representation of the population mean. The minimum and maximum also will give an idea what the range of the data set is and the standard deviation will tell if the data points are close or spread out from the mean. The normal distribution can display how the values are distributed and express what the probability that a number will be greater or lower than a set value. This would be useful if there was limit or some historical design value that was of interest. If major storm data produced large runoff amounts in the tributaries flowing to the Jordan River in the model had great probability the downstream implications of the runoff could be assessed.

The R script used can be seen in Appendix II. There are many methods to load data into R, but a copy and paste method was used in this case. The user must have the Excel sheet open, copy the data of interest and then use the command *read.delim*. This gives a data frame and typing *copdat* in R will show the vector. Displayed in Appendix II are two ways to find the minimum, maximum, mean and medium. One is to use the summary function and the other is to solve each individual. When solving for the normal distribution function, R requires a value such as a limit being analyzed then the mean and standard deviation. The function will use the lower tail value as a default if not specified.

The probability exceedance program was created to process flow data which had been uploaded into the MS SQL database from the GoldSim Model. The probability exceedance program determines the likelihood in terms of a percentage that an event will occur and plots the results graphically. Exceedance probability plots are commonly used in representing runoff and stream flow data. The produced plot also allows the user to see differences between modeled and historic results. To calculate the exceedance probability the equation below was used.

$$Exceedence\ Probability = \frac{s}{(n+1)}$$

s = rank of the value within the total number of values
n = total number of values

Within the program the total number of values in the dataset was found using the *count* command in R. Ranking the values to determine the *s* variable was performed using the *vect* and *rank* commands in R. The probability exceedance program creates a flow exceedance plot displaying the exceedance probability versus the stream flow. An example of this plot can be seen as Figure 3. Additionally the script written for this program can be found in Appendix II.
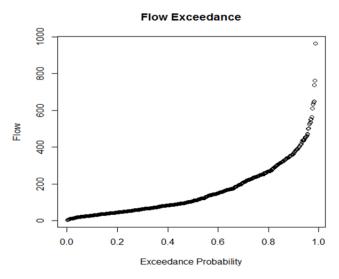

Figure 3: Example of the Output Exceedance Plot from the Exceedance Probability Program

Description of HydroServer HydroServer Lite is a hydrological data server. It is a lightweight version of the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) HydroServer written in PHP [3]. It can easily be placed on a website which then allows a user to geo-reference on a map at which the data was taken as well as quickly developing tables and graphs for visualizing the data or model results. See Figure 4 below for an example of a graph generated by the HydroServer Lite website. A HydroServer can then be registered with the CUAHSI Hydrologic Information Systems (HIS) so that the data can be discoverable to anyone using HydroDesktop.
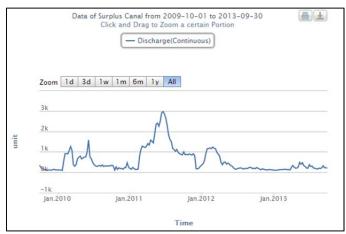

Figure 4: Discharge at the Surplus Canal

Use of HydroServer Lite

For one aspect of this project, a HydroServer has been developed and employed on a free webhosting service. The HydroServer was used as a means for making the data in addition to generated results, readily available to the public and easily located and visualized. Data can easily be imported into HydroServer as a CSV file and organized according to the user's need. Instructions for setting up a HydroServer on a public website are found at: http://hydroserverlite.codeplex.com/documentation and requires very little knowledge of computer programming or website building. For demonstrative purposes, raw time series data that was used in a model was downloaded from HydroDesktop. HydroDesktop is a free and open source GIS enabled desktop application [4]. This

data has been uploaded into the HydroServer as spatial data at six locations, as can be seen in Figure 5.

These sites all contribute to flows in the Jordan River. The results from the SWMM/GoldSim model were uploaded into the HydroServer and given a spatial reference to the point of diversion from the Jordan River that flows into the Great Salt Lake. With this HydroServer functional, a professor, scientist, engineer, student or research team using this system would also be able to search, use or upload their own data into the HydroServer. This would contribute to the growth of the public data as well as provide a tool that would store their data on a spatial database and easily develop tables or graphs of their own model results. Furthermore, if a user fills their server with relevant and useful data, it can be registered on the broader CUAHSI HIS system so others can discover the data. Instructions for registering a HydroServer can be downloaded from http://hydroserverlite.codeplex.com/documentation. Once registered and properly set up the data can be searched in HydroDesktop where it can be compared with similar data nearby. This is beneficial not only from a sharing standpoint, but from a data analysis standpoint as well because HydroDesktop offers a larger spectrum of data analysis tools.
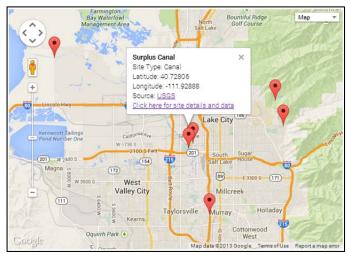

Figure 5: HydroServer Map Interface

V. CONCLUSION

To reemphasize, the overall goal was to demonstrate how to link two or more water resources modelling tools to answer questions that might not otherwise be possible and to enhance modelling accuracy. It was determined that creating a database as a central hub for data and linking the software, in this case SWMM and GoldSim to the database, was the most efficient process.

Model integration is an important research area that can be used to effectively answer countless questions related to water resources. This report outlined the process of linking SWMM, GoldSim, and Microsoft SQL to each other in an effort to be able to answer a greater number of research questions with the same amount of resources.

SWMM is an effective urban hydrology modelling software that was used to simulate runoff for the Salt Lake Valley. Precipitation data was then gathered for Salt Lake City and implemented into the model. A 24-hour time period in which a storm event occurred was used to simplify the model and limit the amount of data generated, as the focus was on the process. The runoff volumes from the SWMM model were then saved to a database created in Microsoft SQL.

A larger model created in GoldSim that encompasses the Jordan River and downstream wetlands was employed to create a link between the two models. A connection was made to the database that effectively allowed GoldSim access to the results that had been stored in the database from SWMM. GoldSim then used historical data and the existing model to show the effects that stormwater has on the flow in the Jordan River and downstream wetlands.

Finally, GoldSim results were then saved in a different table within Microsoft SQL. Recognizing that the data was saved to a local server, it was deemed useful to create a HydroServer that can be publicly accessed in case the user wants to mimic results or view data. The group also used the statistical software R to perform further analysis to demonstrate additional capabilities once the

connection to the software's and the database were made.

Based on findings over the course of three months, the methods described in this report are the most efficient in describing how to link SWMM and GoldSim. The group's intent is to have this document serve as an instruction guide for how to create this link or a link with similar software. The group has already identified a fellow researcher poised to use this document for such purposes upon completion.

ACKNOWLEDGMENT

REFERENCES

[1] Goldsim Technology Group, Goldsim Probabalistic Simulation Environment User's Manual, Vol 2, July 2013. Available at *www.**goldsim**.com/Web/Downloads/**UserManuals**/*.

[2] Lewis A. Rossman, Goldsim Storm Water Management Model User's Manual, Vol 5, July 2010. Available at http://www.epa.gov/nrmrl/wswrd/wq/models/swmm/.

[3] Conner, L., Ames, D.P., Gill, R.A., HydroServer Lite as an Open Source Solution for Archiving and Sharing Environmental Data for Independent University Labs, *Ecological Informatics*, Available online, September 5, 2013. Pending volume and issue assignment.http://www.sciencedirect.com/science/article/pii/S15749541 13000770

[4] Ames, D.P., Horsburgh, J.S., Cao, Y., Kadlec, J., Whiteaker, T., and Valentine, D., 2012. HydroDesktop: Web Services-Based Software for Hydrologic Data Discovery, Download, Visualization, and Analysis.*Environmental Modelling & Software*. Vol 37, pp 146-156.http://dx.doi.org/10.1016/j.envsoft.2012.03.013

Appendix I. ODBC Setup Instructions

Setup Window 1
- Name: [*example_name*]
- Description: [*example_description*]
- Server: [*COMPUTERNAME*]\SQLEXPRESS

Setup Window 2
- Select the "With Integrated Windows authentication" option.

Setup Window 3
- Check the "Change the default database to:" option and select the previously created database that holds the table which contains the SWMM results.
- Check the "Use ANSI quoted identifiers" option.
- Check the "Use ANSI nulls, paddings and warnings" option.
- Set the "Application intent:" value to READWRITE.

Setup Window 4
- Check the "Perform translation for character data" option and leave all others unchecked.
- Select "Finish" and click "Test Data Source…" to make sure that the connection is working.

When the setup is complete, the ODBC Data Source Administrator will look similar to the figure below. The ODBC connection that was created for this figure is *sql_generic_odbc*.

Appendix I – Figure 1. Completed data source connection ready for use by GoldSim

Appendix II. R Script – Statistical Analysis

```
#Load the RODBC library
library(RODBC)

#Pastes the copied data from an open Excel spreadsheet, copdat is a data frame
copdat <- read.delim("clipboard")

#Gets the statistical summary of copdat$CFS
summary(copdat$CFS)

#Determines the mean, median, minimum, and maximum values. Also finds the variance and standard
deviation
mean(copdat$CFS)
median(copdat$CFS)
min(copdat$CFS)
max(copdat$CFS)
var(copdat$CFS)
sd(copdat$CFS)

#Counts the number of rows in the dataset
nrow(copdat)

#Probability that the peak runoff will be greater than 600 cfs or the upper tail in a normal distribution
pnorm(600, mean=265.72, sd=277.81, lower.tail=FALSE)

#The lower tail in a normal distribution
pnorm(600, 265.72, 277.81)
```

R Script – Exceedance Probability Plot

```r
hydro <- read.table("c:/Users/Travis/FlowData.csv",header=TRUE,sep=",")
plot(hydro$Date,hydro$DataValue,type="n", xlab="Time",ylab="Temperature",lty=3)
title("Temperature Plot")

#Count the number of rows in a dataset
n <- nrow(hydro)

#Order the data largest to smallest
attach(hydro)
order <- hydro[order(DataValue),]

#identify the largest value
max <- max(hydro$DataValue)

#rank the values
vect = c(hydro$DataValue)
rank <- cbind(vect,as.numeric(factor(vect)))

#Determine exceedance probability
exprob <- rank/(n+1)

#plot
plot(exprob[,c(2)],hydro$DataValue,xlab="Exceedance Probability",ylab="Flow")
title("Flow Exceedance")
```

Appendix III. Tables Used to Perform R Analysis

Table 1. Displays data from the sub-catchments for the watershed

Subcatchment	Total Precip in	Total Runon in	Total Evap in	Total Infil in	Total Runoff in	Total Runoff 10^6 gal	Peak Runoff CFS	Runoff Coeff
GSL	0.41	0	0	0.32	0.01	529.19	651.95	0.035
BN	0.41	0	0	0.34	0.02	21.28	591.77	0.06
DL	0.41	0	0	0.33	0.03	5.47	164.88	0.079
UCC	0.41	0	0	0.37	0.01	2.03	66.88	0.016
LJR	0.41	0	0	0.28	0.08	56.07	1038.87	0.194
URB	0.41	0	0	0.36	0.01	0.97	31.94	0.016
UPC	0.41	0	0	0.35	0.01	7.58	250.72	0.02
UEM	0.41	0	0	0.35	0.01	2.86	94.47	0.022
UMC	0.41	0	0	0.35	0.01	2.52	83.17	0.016
LBC	0.41	0.01	0	0.35	0.02	11.68	278.54	0.05
LLC	0.41	0.01	0	0.36	0.02	4.32	138.06	0.046
LDC	0.41	0	0	0.33	0.03	6.73	207.6	0.07
LWC	0.41	0.01	0	0.34	0.03	4.78	153.76	0.071
BG	0.41	0	0	0.33	0.03	18.51	553.78	0.072
CY	0.41	0	0	0.34	0.02	6.19	194	0.059
MBC	0.41	0	0	0.34	0.02	20.09	575.84	0.056
RC	0.41	0	0	0.34	0.02	10.3	301.63	0.052
LCC	0.41	0.02	0	0.36	0.02	2.46	81.18	0.046
LRB	0.41	0.02	0	0.36	0.02	1.1	36.15	0.057
LEM	0.41	0.03	0	0.37	0.02	1.84	59.93	0.041
LPC	0.41	0.07	0	0.4	0.03	2.81	90.98	0.053
LMC	0.41	0.01	0	0.34	0.03	8.6	220.24	0.078
UBC	0.41	0	0	0.36	0.01	6.02	199.21	0.017
ULC	0.41	0	0	0.35	0.01	3.28	108.31	0.017
UDC	0.41	0	0	0.36	0.01	0.7	22.93	0.016
UWC	0.41	0	0	0.35	0.01	0.94	31.22	0.019
UJR	0.41	0	0	0.28	0.08	37.4	946.41	0.194

Table 2. Statistical analysis results from R

Mean	265.72
Median	164.88
Min	22.93
Max	103.87
Variance	77178
Standard Deviation	277.81
Count	27
Probability of Peak Runoff > 600 cfs	11.40%