MODSIM versus RiverWare:

A comparative analysis of two river reservoir modeling tools

Final Report 2014.3669

U.S. Department of the Interior
Bureau of Reclamation
Pacific Northwest Region
Boise, Idaho

November 2014

Mission of the U.S. Department of the Interior

Protecting America's Great Outdoors and Powering Our Future

The U.S. Department of the Interior protects America's natural resources and heritage, honors our cultures and tribal communities, and supplies the energy to power our future.

Mission of the Bureau of Reclamation

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Cover Photographs: Left – Middle Deschutes River below North Dam. Right – Aerial view of A.R. Bowman Dam (Prineville Reservoir). *Photographs by Dave Walsh, Reclamation.*

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*
T1. REPORT DATE SEP 2014	**T2. REPORT TYPE** Research		**T3. DATES COVERED**
T4. TITLE AND SUBTITLE MODSIM versus RiverWare: A comparative analysis of two surface water modeling tools			**5a. CONTRACT NUMBER** 2014.3669
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER 1541 (S&T)
6. AUTHOR(S) Jennifer Johnson jmjohnson@usbr.gov 208-378-5225			**5d. PROJECT NUMBER** 3669
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER PN-6216
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) PN Regional Office U.S. Department of the Interior, Bureau of Reclamation 1150 N. Curtis Rd, Boise, ID 83706			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Research and Development Office U.S. Department of the Interior, Bureau of Reclamation PO Box 25007, Denver CO 80225-0007			**10. SPONSOR/MONITOR'S ACRONYM(S)** R&D: Research and Development Office BOR/USBR: Bureau of Reclamation DOI: Department of the Interior
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 3669
12. DISTRIBUTION / AVAILABILITY STATEMENT Final report can be downloaded from Reclamation's website: https://www.usbr.gov/research/			
13. SUPPLEMENTARY NOTES			

14. ABSTRACT *(Maximum 200 words):* Reclamation hydrologists and engineers use river-reservoir models to plan for future water conditions and to understand the impact of potential changes to the system that may result from physical or operational changes. Choosing an appropriate river-reservoir modeling framework can be challenging because the criteria used to make the selection are not always quantifiable.

This project developed a method for comparing two river-reservoir modeling frameworks using Decision Matrix Analysis with Analytical Hierarchical Process. The method allows for the comparison of both qualitative and quantitative metrics. It is also flexible so that it can be used in many different situations and to compare many different types of models.

The project resulted in the submission of a journal article to the Journal of American Water Resources in October 2014.

15. SUBJECT TERMS
Model comparison, decision analysis matrix, MODSIM, RiverWare, Deschutes Basin

16. SECURITY CLASSIFICATION OF: U			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Jennifer Johnson
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	U	48	**19b. TELEPHONE NUMBER** 303-445-2125

S Standard Form 298 (Rev. 8/98)
P Prescribed by ANSI Std. 239-18

PEER REVIEW DOCUMENTION

Project and Document Information

Project Name MODSIM versus RiverWare: A comparative analysis of two surface water modeling tools.
WOID X3669
Document MODSIM versus RiverWare: A comparative analysis of two river-reservoir modeling tools
Document Author(s) Jennifer Johnson Document date: November 2014
Peer Reviewer: Bob Lounsbury

Review Certification

Peer Reviewer: I have reviewed the assigned items/sections(s) noted for the above document and believe them to be in accordance with the project requirements, standards of the profession, and Reclamation policy.

Reviewer Bob Lounsbury Date reviewed: 10/15/14
(Signature)

Executive Summary

Reclamation hydrologists and engineers use river-reservoir models to plan for future water conditions and to understand the impact of potential changes to the system that may result from physical or operational changes. Choosing an appropriate river-reservoir modeling framework can be challenging because the criteria used to make the selection are not always quantifiable.

This project developed a method for comparing two river-reservoir modeling frameworks using Decision Matrix Analysis with Analytical Hierarchical Process. The method allows for the comparison of both qualitative and quantitative metrics. It is also flexible so that it can be used in many different situations and to compare many different types of models.

The project resulted in the submission of a journal article to the Journal of American Water Resources in November 2014.

Table of Contents

This page intentionally left blank.

Project Summary

Reclamation hydrologists and engineers use river-reservoir models to plan for future water supply conditions and to understand the impact of potential physical or operational changes to distribution and storage systems. Many river-reservoir modeling frameworks are available and choosing appropriate software for a study can be challenging. This is because the criteria that are used to select a modeling framework can be qualitative in nature and difficult to both quantify and justify.

The need for this type of analysis originated in the Pacific Northwest region where two tools are typically used to simulate river-reservoir systems, MODSIM and RiverWare, and in some cases have been used to develop models for a single basin. MODSIM has been used historically because for many years, it was the only tool that could distribute water based on priority water rights. In recent years, RiverWare developed similar capability. Both tools have strengths and weaknesses and could be used for many of the water management studies that Reclamation addresses

Project Goal

The goal of this project was to develop a method for comparing two river-reservoir modeling frameworks. The method should be flexible enough so that it could be used in the case that (1) models were already developed for a particular basin or (2) before models were developed. It should be able to both quantify and help to justify the decision by removing the subjectivity that often enters the decision making process.

Partnerships

This project was conducted in collaboration with a researcher at the Pacific Northwest National Laboratory, Sara Niehus. The RiverWare model that was used in this analysis was developed for the Basin Scale Assessment of the Deschutes basin funded by the Department of Energy. The MODSIM model that was used in this study had been developed over many years with many funding sources including Reclamation, Oregon Department of Water Resources, and the Confederated Tribes of Warm Springs Reservation.

Results and Conclusions

A method was developed to compare two modeling frameworks using Decision Matrix Analysis and Analytical Hierarchy Process. This method allows for the comparison of two frameworks using both qualitative and quantitative metrics, since the decision to use one framework over another is not always a purely quantitative decision.

The case study was designed to test the method by comparing a MODSIM model and a RiverWare model that simulated similar networks in the Deschutes basin. Three modelers experienced in both tools were surveyed to determine the selection metrics that should be used to make the selection. The metrics were independently weighted by each modeler using Analytical Hierarchy Process which assigns a relative weight to each metric based on relative importance to the modeler. The metrics were then scored for two possible modeling scenarios, adding a new hypothetical reservoir and adding a new hypothetical instream flow requirement. The weighted scores for the three modelers were averaged into one score for each model. For both scenarios, the modeler's scores were higher for MODSIM.

This case study is merely an example of how to use the Decision Matrix Criteria for model selection. Although the modeler's scores showed that MODSIM was the preferred tool for the hypothetical scenarios, different modelers may score the tools differently or the scores may be different for different scenarios. The resulting method is flexible and can be applied to different model types and different model needs.

Products

A draft journal article documents the development of the method and the results of the case study. It was submitted to the Journal of American Water Resources in November 2014 (see Appendix A).

Appendix A

Journal Submission to American Water Resources Association, November 2014

Decision Matrix Analysis with an Analytical Hierarchy Process: A Methodology for Comparatively Evaluating River-Reservoir Modeling Platforms

Jennifer Johnson[1] and Sara Niehus[2]

ABSTRACT

With changing climate and weather patterns, increased occurrence and duration of drought and implementation of environmental objectives, it has become increasingly difficult to manage the many water needs of river basins in the Western United States. Water managers often rely on river-reservoir modeling tools to understand the impacts and plan for possible changes to systems. Several of these modeling tools have been selected based on historical use, staff and stakeholder experience, and licensing costs; among other reasons that can be difficult to quantify.

A decision matrix analysis framework using an analytical hierarchy process for comparing river-reservoir modeling tools was developed to aid water managers and modelers in making decisions based on equally weighted qualitative and quantitative decision metrics. To demonstrate the functionality of the method, two case studies are evaluated using existing river-reservoir models developed using two modeling tools, RiverWare and MODSIM, for the Deschutes River Basin in Central Oregon. This decision framework is able to quantitatively analyze specific desired goals and modeling outcomes, which may be qualitative or quantitative, so that an appropriate modeling tool is selected. The framework can be used by both modelers and policy makers to make collaborative decisions to meet the needs of the application. The framework is flexible and

[1] Hydrologic Engineer, Bureau of Reclamation, Pacific Northwest Region, Boise, ID 83706
2 Engineer, Pacific Northwest National Laboratory, Richland, WA 99352

its potential application could be expanded beyond river-reservoir model evaluations.

Keywords: Decision Matrix Analysis, Analytic Hierarchy Process, MODSIM, RiverWare, River-reservoir modeling,

INTRODUCTION

Water resource managers are increasingly being asked to determine potential impacts to systems that involve the interests of many different stakeholder groups including local, state, and Federal governments, Tribal Nations, environmental groups, and private interests. They use river-reservoir management modeling tools to understand the impact of potential changes in water distribution and storage systems so that they may appropriately plan for future conditions by simulating possible changes to water deliveries, reservoir storage, in-stream flows, water quality, and power production.

Choosing an appropriate river-reservoir management modeling tool can be challenging because modeling platforms have differing capabilities and strengths. For example, one tool may be better suited for evaluating flood control options while another one may be better suited for evaluating water delivery options. However, often the decision to use a particular tool may often be due to administrative constraints such as licensing costs, prior knowledge and expertise in a particular tool, or organizational comfort or familiarity with a particular tool.

This paper seeks to develop a decision framework for comparing river-reservoir management modeling tools that can be used for both qualitative and quantitative decision metrics. Two case studies were evaluated using existing models developed with two river-reservoir management modeling tools, RiverWare (RiverWare, 2014; http://www.riverware.org) and MODSIM-DSS (MODSIM, 2014; http://modsim.engr.colostate.edu) (hereafter referred to as MODSIM), for the

Deschutes River Basin in Central Oregon. The resulting framework may be used in determining the appropriate tool for long range planning studies or day-to-day operations.

BACKGROUND

River-reservoir management modeling tools are designed to simulate the distribution of water within a regulated river system. Many river-reservoir management tools have been developed to address differing objectives within a geographic region that including MODSIM, RiverWare, MIKE BASIN (DHI Water & Environment, 2006), CALSIM (Draper, et al. 2004), IQQM (New Wales Department of Infrastructure, Planning and Natural Resources; Hameed and O'Neill, 2005), RIBASIM (delft Hydraulics, 2006), and WEAP (Stockholm Environmental Institute-Boston; Yates et. al. 2005). Some of these tools are designed with only particular objectives in mind and are therefore not suitable for every application.

Water management groups have conducted comparisons of river-reservoir modeling tools to try to select or justify an appropriate selection (Sulis and Sechi, 2012; Wurbs, 2012). These comparisons typically consist of a general discussion of each tool's capabilities, the pros and cons of each tool, and possibly a comparison of output for a given scenario. Although these analyses can be useful in choosing an appropriate tool, there are no standard metrics used to select an appropriate one to achieve maximum desired objectives while minimizing subjectivity. Developing a set of standard metrics can be difficult as many of the metrics used to select a modeling tool are considered qualitative and may change depending on the study objectives.

River-Reservoir Modeling Platforms

In the Western United States, the common modeling objectives include flood control, distribution of water based on priority, environmental flows, and hydropower production.

Although many tools are available, two (MODSIM and RiverWare) have often been used by Federal and state agencies to simulate management options, as they both have the capability to simulate the distribution of water based on water right.

MODSIM is a river basin Decision Support System (DDS) and network flow tool that allows for the integrated analysis of water sector elements and optimization of resource management by allocating limited water resources (Berhe et al. 2013). The tool was developed at Colorado State University in 1978(Shafer and Labadie 1978), making it the longest continuously maintained river basin management software available. For many years, MODSIM was one of the few river-reservoir modeling tools that could simulate water allocation based on priority water rights, and therefore most water allocation models have been developed using MODSIM, especially in the Western United States in collaboration with the Bureau of Reclamation (Reclamation). MODSIM uses a minimum cost optimization solver which routes water based on network costs. Network costs are assumed to be properly configured, so modelers are required to have a thorough understanding of network cost structures and their implementation in MODSIM to get the desired result. Without this knowledge, modelers run the risk of incorrectly simulating the system. Debugging problems can be difficult because no graphic user interface or debugging tool is provided to view the optimization of the network flow. Users are required to follow within the software's executed syntax and interpret it to identify bugs. On this basis alone, MODSIM requires significant user investment to learn and due diligence from the user.

RiverWare is a river-reservoir modeling platform that is capable of simulating reservoir and system operations, responsive forecasting, operational policy evaluation, system optimization, water accounting, and water rights administration (Zagona et al. 2001). It was developed in 1986 by the Center for Advanced Decision Support for Water and Environmental Systems

(CADSWES) at the University of Colorado Boulder. The RiverWare model relies on logic within rules that drive the simulation. These rules are complex algorithms that express multifaceted operational policy that can be prioritized. Rules are executed in a user specified order at each time step and on simulation objects within the model (Zagona et al. 2001). Water accounting in RiverWare models are a network of "paper water" accounts that are separate from the simulation of "physical water" and the concepts of accrual, carryover, transfers and exchanges are represented in the model (Zagona et al. 2010).

RiverWare's ability to simulate the distribution of water based on water right is relatively new, so MODSIM was used for many years in basins that required that capability. However, MODSIM's "black-box" nature can make it difficult to use and transfer to other users. Since RiverWare's rule structure can be seen as more transparent, once RiverWare had similar capabilities to MODSIM, models were developed in basins where MODSIM models already existed. However, both models continue to be used because there is no clear way to evaluate the benefits of using one tool over the other.

Description of the case study area

Named "River of the Falls" ("Riviere des Chutes" in French), the Deschutes River originates in the Cascade Mountains of Central Oregon and runs 405.5 kilometers (km) to join the Columbia River near The Dalles, Oregon. The Deschutes River Basin covers approximately 27,701 square kilometers (km^2) and is the second largest river basin in Oregon (Aylward and Newton 2006). Major tributaries to the Deschutes River include the Little Deschutes River, Crooked River, Whychus Creek, and Tumalo Creek. The Deschutes River discharge is considered one of the most uniform and stable streams within the United States, not only month to month, but also year to year (Russell, 1905; Henshaw et al., 1914; USDA 1996; O'Connor et al. 2003). This stability is due to the strong influence from groundwater - in part by natural occurrence of high permeable Cenozoic volcanic rocks (O’Connor et al. 2003), and artificially augmented by inefficient irrigation practices.

The Deschutes River Basin can be divided into three sub-basins (Figure 1). This case study focuses on two of these sub-basins: the Upper Deschutes, which extends from the river’s headwaters downstream to Lake Billy Chinook reservoir formed by the Portland General Electric Pelton-Round Butte Hydroelectric Project; and the Crooked River, which extends from the river’s headwaters to its confluence with the Deschutes River near Madras, Oregon. Overall the study area has a drainage area of 11,700 km^2 with an average annual runoff of 5.2×10^9 m^3.

Irrigation development began in the Deschutes River Basin in 1899 by the Deschutes Reclamation and Irrigation Company (DRIC) now known as the Swalley Irrigation District. Currently there are seven irrigation districts (Arnold, Central Oregon, North Unit, Ochoco, Swalley, Three Sisters, and Tumalo) that store and divert water from the upper Deschutes and Crooked River and its tributaries (Figure 1). In addition to irrigation, there are several small hydropower facilities that exist within the Deschutes River and irrigation canals.

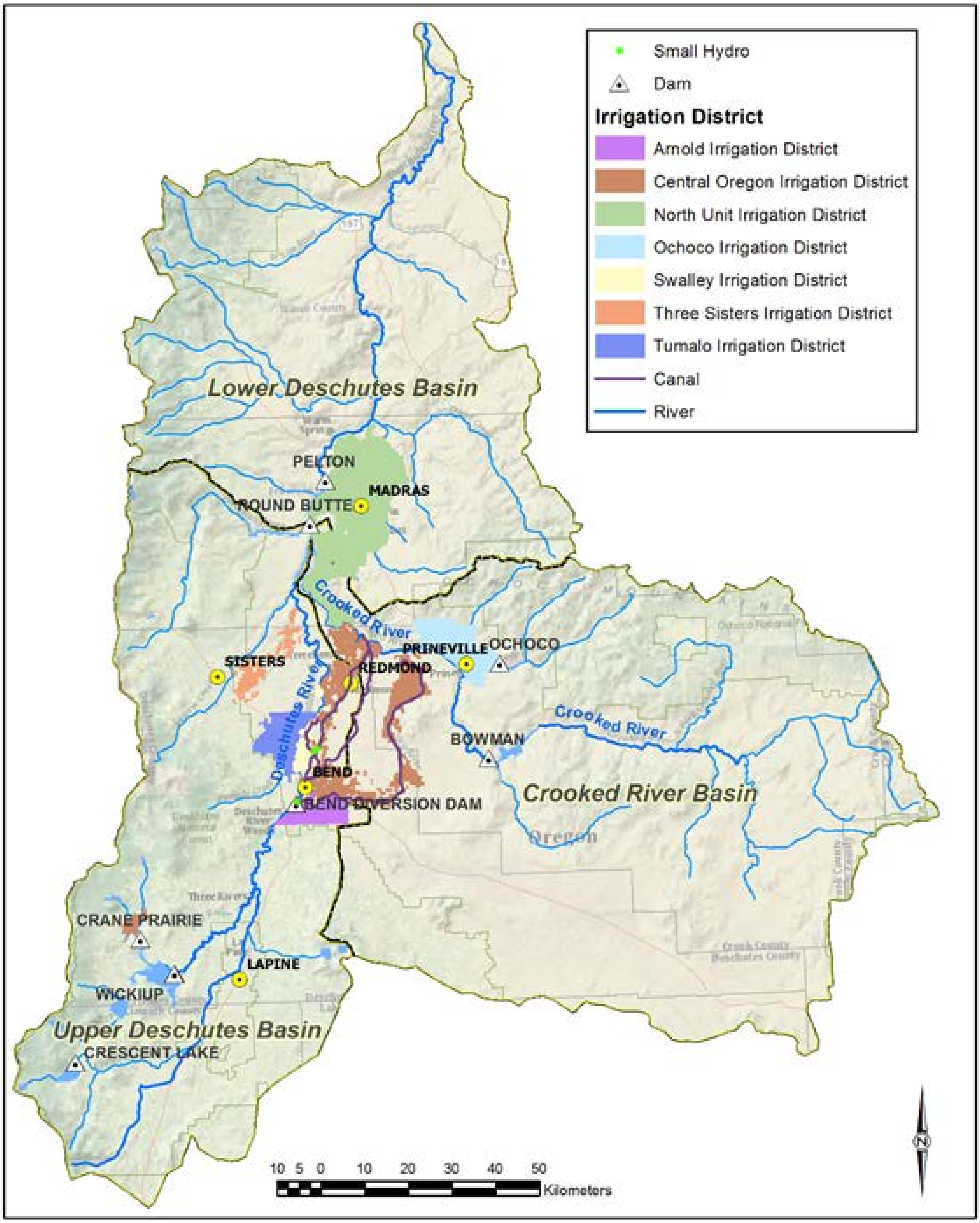

Figure 1: Deschutes River basin showing three major subbasins, existing dams, hydropower facilities, and irrigation districts and canals.

Description of the case study models

Existing MODSIM (Johnson, J. and J. LaMarche, 2013) and RiverWare (Larson et al. 2014) models of the Deschutes Basin were used for the river-reservoir model evaluation. The models were built using a similar network design and with similar operating logic. The models simulate the Upper Deschutes and its reservoirs, Crane Prairie, Wickiup, and Crescent Lake, and the Crooked River and its reservoirs, Prineville and Ochoco. Both models represent natural flow and stored water rights, diversions, groundwater return flows that result from irrigation inefficiencies, and minimum flow requirements.

The MODSIM model of the Deschutes basin was developed over many years with contributions from Oregon Water Resources Department, Reclamation, and Natural Resource Consulting Engineers (NRCE) (funded by the Confederated Tribes of Warm Springs Reservation). It has been used for multiple studies including biological assessments (Reclamation, 2005), in-stream flow and storage assessments, and climate change evaluations (Reclamation, 2011), and has earned stakeholder trust. The model operates at a monthly timestep and simulates the period 1928 through 2005. Although the model has proved to be useful for many studies, the monthly timestep limits the model’s usefulness for some application like minimum flow assessments, that require a finer temporal resolution as monthly timestep could mask a flow violation that occurred in only one or two days of a month.

The RiverWare application for the Deschutes was developed by Pacific Northwest National Laboratory (http://basin.pnnl.gov). It has been used for the Basin Scale Opportunity Assessment (BSOA) (Larson et al. 2014) and is still in the process of earning stakeholder trust due to its relative infancy. The BSOA initiative is a US Department of Energy (DOE) national effort to develop an approach to basin-scale hydropower and environmental assessments that emphasize sustainable, low-impact or small hydropower and related renewable energies within the context of environmental protection/restoration. RiverWare was selected for its capability to simulate reservoir storage accounting and inline canal hydropower on a daily timestep. The model network was configured using the MODSIM network layout and, where functionally possible, implemented identical inputs. The model operated at a daily timestep and simulated the period 1983 through 2010.

TECHNICAL APPROACH

River-Reservoir Model Evaluation

The first step in any model comparison is to understand what is the question that these models are looking to answer and does any of these have the ability to provide those answers. If only one model has these capabilities, then there is a clear path forward. It is when both have the ability but have various strengths and weaknesses is where a method of evaluation can be a challenge.

Decision matrix analysis, first made popular by Pugh (1996), is a method commonly used to make decisions when the selection metrics (model strengths and weaknesses) are difficult to quantify; however it has not been applied to determining an appropriate river-reservoir modeling tool. Five general step have been developed for this evaluation method that include: (1) determine selection metrics; (2) determine the relative importance of the selection metrics and assign weights; (3) score the selection metrics for the given scenario; (4) multiply the scores by the weights resulting in the relative score; (5) add the relative scores and choose the scenario with the highest relative score.

The decision matrix analysis approach that was used for this study attempts to evaluate both quantitative and qualitative metrics in a quantitative fashion using the five steps described above.

The methods were designed to address the primary questions that modelers consider when trying to determine the appropriate river-reservoir modeling tool to use.

Step One: Determine Selection Metrics

Engineers and hydrologists who have experience with multiple river-reservoir modeling tools brainstormed selection metrics. The metrics were selected to best understand the pros and cons of choosing a particular river-reservoir modeling tool, including factors like time to develop a new system model or modify an existing system model, cost of the software, and transferability between modelers. The metrics were separated into quantitative and qualitative categories. If a metric could be quantified using the case study, it was considered to be quantitative.

Step Two: Assigning Weights to Model Metrics

Weights were assigned to all of the selection metrics using Analytic Hierarchy Process (AHP), where the metrics were compared against each other, one-by-one, to determine a relative importance value. AHP, developed by Saaty (1977 and 1994), can be described as establishing a rank of desirable functions for a modeling platform by making a series of judgments based on pairwise comparisons of these described modeling functions. For example, when comparing river-reservoir modeling platforms, the model developer might say they prefer flexible software integration over software price and software price over model run time. Table 1 shows an example of the pairwise comparison matrix using four selection metrics, A, B, C, and D.

Table 1: Example of pairwise selection matrix using four metrics, A, B, C, and D.

	A	B	C	D
A	-	A or B	A or C	A or D
B	-	-	B or C	B or D
C	-	-	-	C or D
D	-	-	-	

The metrics are set up in the matrix so that they are column and row names, and then each metric is compared against the others individually. The number of times a metric is selected is called a selection, s. The sum of possible selections, S, is defined by Equation 1:

$$S = 0.5N(N-1) \tag{1}$$

where N is equal to the number of metrics. For the example matrix, N is 4, so S calculates to be 6, which is the number of decision points shown in the matrix in Table 1. The weights for each metric are then calculated by dividing the number of individual selections, s, by the sum of the possible scores, S. So that none of the weights calculate to zero, even if a metric is not selected in the matrix, an adjustment factor is added to the weight equation of 1/N to the numerator and 1 to the denominator. The adjusted weight, W, equation 2 for each metric, i, is then:

$$W_i = \frac{s + 1/N}{s + 1} \tag{2}$$

For the purposes of this paper, the AHP was completed by three separate modelers, two with experience in both MODSIM and RiverWare, and one with experience in only RiverWare. They completed the process separately and their adjusted weights were then averaged together.

Step Three: Scoring the Selection Metrics for the Given Scenario

The decision matrix analysis involves a qualitative and quantitative evaluation of both MODSIM and RiverWare with a Mixed Method Evaluation (Driscol et al. 2007). Strict quantitative studies sometimes fail to capture nuances within the groups or communities studied, and the analysis can often lack the depth and detail of qualitative methods (Bamberger 2012). Qualitative studies can be powerful however data replication can be difficult and it can be difficult to isolate specific elements that are driving results (Bamberger 2012). By using combinations of qualitative and quantitative data the evaluation can be improved by ensuring that the limitations of one type of data are balanced by the strengths of each other (Driscol et al. 2007). Typically in these Mixed Method Evaluations, one method is the dominating evaluation. For this study the quantitative method will dominate in that the qualitative analysis is transferred to a quantitative scoring to aid modelers in model suitability and selection.

Qualitative Metric Evaluation

The same three modelers were asked to score the model evaluation metrics established in Step One with the idea that the models would be used to at least simulate two scenarios: (1) new reservoir storage by adding a new reservoir and (2) new minimum flow conditions at control points. The results of the three modelers were recorded separately and then averaged. The scores ranged from one to five with one indicating that the model is least successful for that metric and five indicating that it is most successful.

The initial quantitative metrics were evaluated using existing MODSIM and RiverWare models of the Deschutes Basin case study and focused on three metrics that were identified in Step One: (1) the time it takes to develop a new scenario in an existing model, (2) model run time, and (3) the quality of the model calibration. Model calibration consisted of the comparison of simulation results to the observed historical records of discharge and storage at several reservoirs and gage locations.

Quantitative Metrics Evaluation: Scenario Development Time

The metric evaluated was the time needed to develop new scenarios within an existing model. Two new scenarios were created: (1) a new storage reservoir and (2) a new minimum instream flow requirement. The new reservoir storage scenario evaluated a fictional 0.049 cubic kilometers (km^3) (40,000 acre-foot) reservoir on the Little Deschutes River where the stored water was used to supplement river flows downstream during the months of April through September. The new minimum instream flow requirement evaluated a 1.42 cubic meters per second (m^3/s) (50 cubic foot per second (cfs)) on the Crooked River above the confluence with the Upper Deschutes. The minimum instream flow requirement was simulated using a flow-through demand node in MODSIM and a control point with an instream flow account in RiverWare. The time required to set up each scenario and resolve any issues was recorded. The scenarios were designed and implemented by the same modeler that was equally experienced in using the MODSIM and RiverWare models, to minimize the potential for one model to benefit over another because of experience or knowledge.

Quantitative Metrics Evaluation: Model Run Time & Calibration Error Evaluation

The model run time and calibration error metrics were evaluated using the historical/calibration model runs in MODSIM and RiverWare of the Deschutes Basin. Model runtime was divided by the number of timesteps to allow for a comparable value between models. The quality of calibration was determined by the mean absolute error (MAE) and the root mean squared error (RMSE) and were calculated at key locations within each model at all simulated timesteps. The RMSE and MAE are measurements that express average model-prediction error and can be used to measure average difference between model calibration quality. MAE was calculated as:

$$\mathrm{MAE} = [\mathrm{n}^{-1} \textstyle\sum_{\mathrm{i}=1}^{n} |\mathrm{Observed_i} - \mathrm{Modeled_i}|] \quad \textbf{(3)}$$

The MAE is the sum of the magnitude (absolute value) of the error between the observed and modeled value. The RMSE was calculated as:

$$RMSE = n^{-1}[\textstyle\sum_{i=1}^{n} |Observed_i - Modeled_i|^2]^{1/2} \quad \textbf{(4)}$$

Both storage and discharge values were used in the calculation and since the storage values were a larger, by a two orders of magnitude, than the flow values, the MAE and RMSE of the storage values were divided by 100. This method was replicated from what was developed for and used in PEST, a parameter estimation software package, when calculating objective functions for quantities of differing units (Doherty, 2005). By implementing this methodology, both the MAE and the RMSE become unitless. A larger MAE or RMSE value indicates a lower quality calibration. When calibrating a river-reservoir model, it is common to find that the model cannot always match historical data to the degree that can be expected in a more detailed physically based model. This is because river-reservoir operation decisions are made by human operators that use many factors, including past experience, to determine how the system is operated day to day, which is not always repeatable by computer logic.

Once the qualitative and qualitative methods have been evaluated they are combined using the Mixed Method Evaluation described above. As the three quantitative results are not in the same comparative scale as the qualitative (one to five) they must be resolved by being scaled to fit into the one to five range and added to the table of qualitative scores. All of the quantitative metrics were such that a lower score indicated a favorable outcome, so the values were scaled using the following Equation 5:

$$\frac{5-V_1}{5(V_1-V_2)} \tag{5}$$

Where V_1 is the first quantified value and V_2 is the second. The results of the equation were rounded to the nearest whole number.

Step Four and Five: Developing the MODSIM and RiverWare Relative Matrix Score

The relative matrix score was developed by multiplying the developed weights in Step Two by the assigned scores for each metric in Step Three and adding up the relative scores for each modeling tool.

RESULTS

Step One: Determine Selection Metrics

The results of the decision matrix analysis are presented in the order of the steps. The list of qualitative metrics resulting from Step One is shown in Table 2 along with a short description of each metric, and the quantitative metric are shown in Table 3.

Table 2: Qualitative selection metrics along with brief descriptions.

Selection Metrics	Description
A. New model development time	How long does it take to develop a new model including network design and setup, data population, simulation logic, and output?
B. Has known deficiencies/bugs	Are there limitations that prevent the model from performing a particular function?
C. Institutional knowledge	How experienced is the modeler or work environment with the tool? How familiar is management and the stakeholders with the tool?
D. User support	Does the model have user support and how effective is the user support; i.e. are problems addressed in a reasonable time frame?
E. Transparency	Can the model be easily transferred to another modeler? How easy is the transfer to another modeler? Can the logic be easily understood by a modeler that has limited experience with the tool? Can the logic and constraints be easily explained to non-modelers?
F. Cost	Are there license fees and if so, how much? Is it a one time purchase or recurring? What is the cost of user support?
G. GUI features	What are the tools available within the GUI that improve the modeling experience, i.e. graphs, tables, calculators, input/output tools, debugging capabilities?
H. Integration with other tools	Can the tool be easily integrated with other modeling tools?

Table 3: Quantitative selection metrics along with brief descriptions.

Selection Metrics	Description
I. Model scenario development time	How long does it take to set up a new scenario in an existing model including the time to adjust the network, rules, and debug any logic?
J. Model run time	How long does it take for the model to simulate a given time period, considering time step size?
K. Calibration quality	How well is the model calibrated?

Step Two: Assigning weights to Model Metrics

Table 4 shows a pairwise comparison matrix that was constructed and used to evaluate the metrics listed in Table 2 and Table 3. Table 4 also shows the criteria compared one-to-one and shows the results from one of the sampled modelers. When comparing metric A, new model development time, to metric B, has known deficiencies/bugs, the sampled modeler chose metric B as having more importance. The remaining metrics were scored similarly.

Table 4: A resulting pairwise comparison of model metrics scoring from one of the sampled modelers.

		A	B	C	D	E	F	G	H	I	J	K
New model development time	A	-	B	C	D	A	A	A	A	I	J	A
Has known deficiencies/bugs	B	-	-	B	D	B	B	B	B	B	B	B
Institutional knowledge	C	-	-	-	D	E	C	C	C	I	J	K
User support	D	-	-	-	-	E	D	D	D	D	D	D
Transparency	E	-	-	-	-	-	E	E	E	I	J	K
Cost	F	-	-	-	-	-	-	F	F	I	J	K
GUI Features	G	-	-	-	-	-	-	-	G	I	J	K
Integration with other tools	H	-	-	-	-	-	-	-	-	I	J	K
Model scenario development time	I	-	-	-	-	-	-	-	-	-	I	I
Model run time	J	-	-	-	-	-	-	-	-	-	-	J
Calibration quality	K	-	-	-	-	-	-	-	-	-	-	-

Figure 2 shows results of the AHP analysis. The spread of the weights assigned by the modelers and the average weights are shown for each metric. This plot shows that there are some metrics where the modelers had similar opinions of the relative importance, like new model development time, and some where the modelers disagreed about the relative importance, like user support and cost.

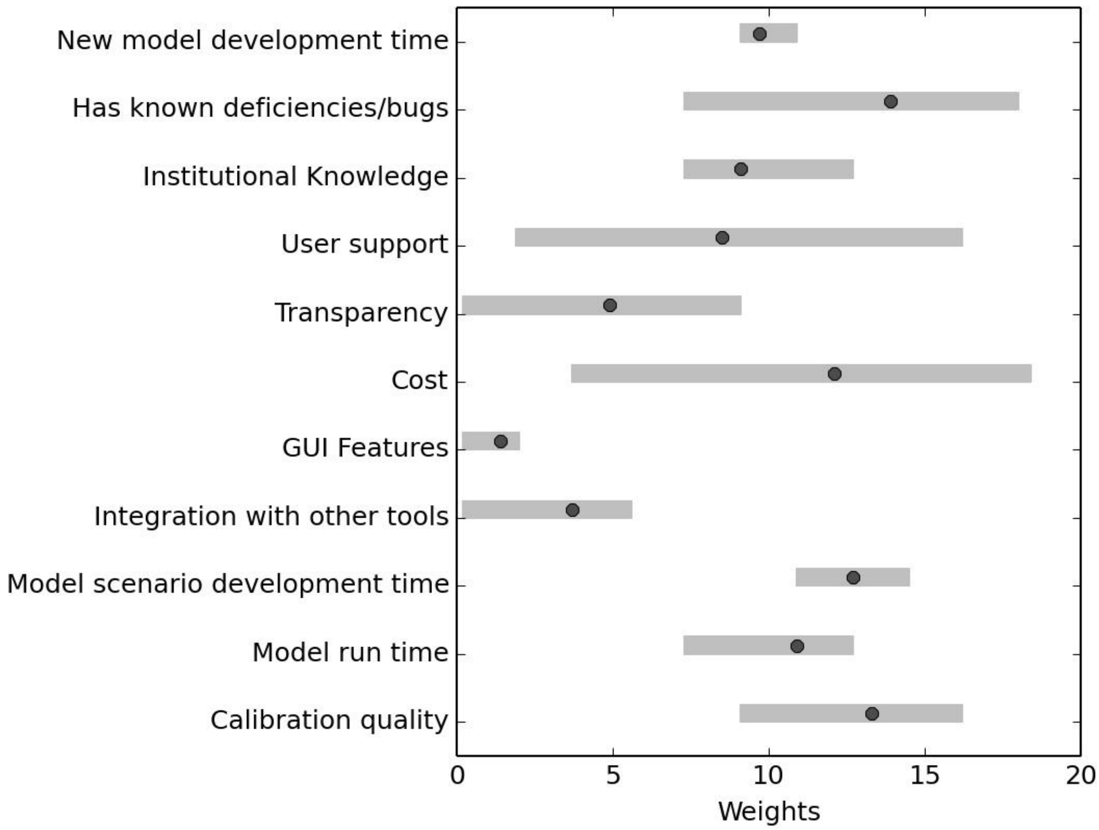

Figure 2: Results of AHP analysis for the case study. The light grey bars indicate the range of the responses from the three modelers and the dark grey dots indicate the average value that will be carried forward.

Step Three: Scoring the Selection Metrics for the Given Scenario

The scores that were assigned by the individual modelers for the qualitative metrics are shown in Figure 3.

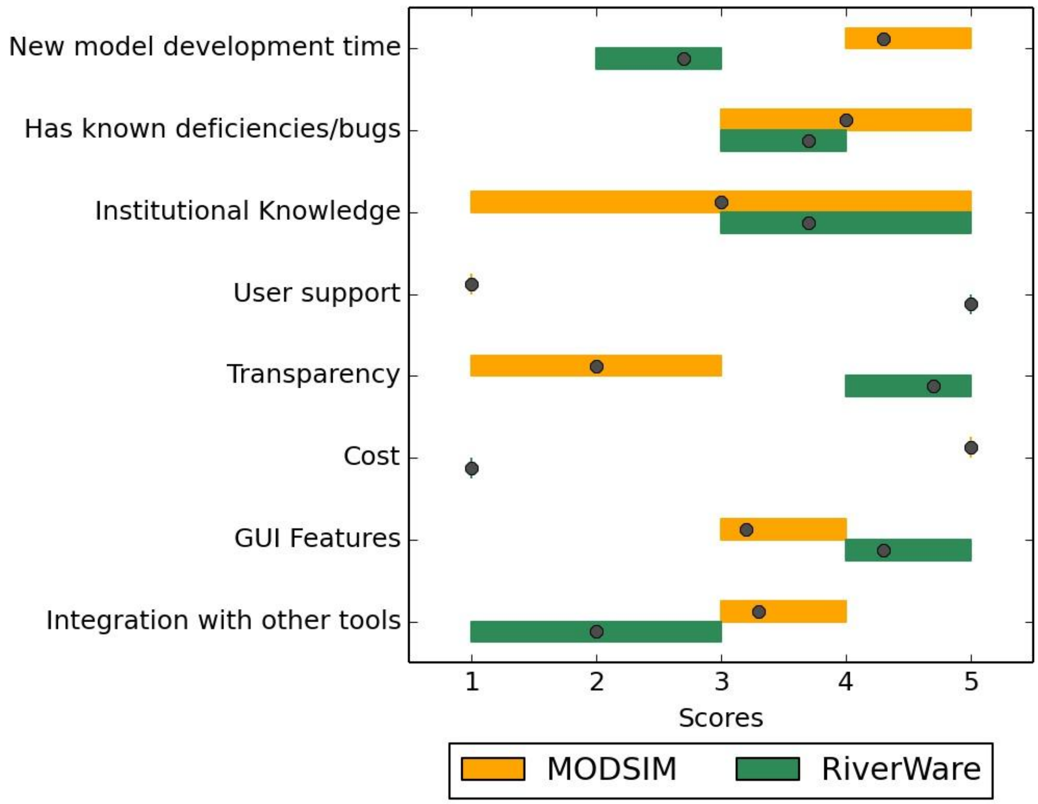

Figure 3: Scores for the qualitative analysis from the three modelers. The bars indicate the range of scores and the dots indicate the average value that will be carried forward in the analysis. The dots without bars indicate that all three modelers gave the same score for the metric.

The three results of the analysis for the quantifiable metrics are described here and shown in Table 5. The time required to set up the first scenario, adding a fictional new storage reservoir on the Little Deschutes River, was about two hours for both models. This included the time to set up the network structure, add the data, adjust any logic, and ensure that the model ran without any errors. The time required to set up the second scenario, adding a new minimum flow requirement on the Crooked River, was about three minutes for MODSIM and 17 minutes for RiverWare.

For the historical calibration models, the MODSIM model has a runtime per timestep equal to 0.05 seconds and the RiverWare model has a runtime per timestep equal to 0.2 seconds. For these two models, the RiverWare model takes longer to run for each timestep.

The RMSE and MAE were calculated for both models for the storage reservoirs (Crane Prairie, Wickiup, Crescent Lake, Prineville, and Ochoco) and flow locations (outflow from each reservoir, the gage at Benham Falls, and the gage below Bend). The MODSIM model RMSE was 20,736 for the storage reservoirs and 293 for the flow locations, with an adjusted total RMSE of 500. The RiverWare model RMSE was 36,210 for the storage reservoirs and 459 for the flow locations, with an adjusted total RMSE of 821. The MODSIM model MAE was 6,818 for the storage reservoirs and 90 for the flow locations, with an adjusted MAE of 158. The RiverWare model MAE was 17,937 for the storage reservoirs and 140 for the flow locations, with an adjusted MAE of 319. Both the RMSE and MAE are reported in Table 5.

For each metric, the scores were calculated using the actual values and converting them to a value between one and five using the relative values of the scores. For example, for the model runtime per timestep values were calculated using:

$$\frac{5 - 0.05}{5(0.05 + 2)}$$

The values are subtracted from five since a lower actual score should have a higher scaled score.

Table 5: Results of quantifiable metric analysis.

		Actual Values		Scaled Values (Scores)	
Selection Metrics		MODSIM	RiverWare	MODSIM	RiverWare
Model Scenario Development Time (hours)	Minimum instream flow requirement scenario	0.05	0.28	4	1
	New reservoir storage scenario	2	2	3	3
Model runtime per timestep (second)		0.05	0.2	4	1
Calibration Statistics (weighted RMSE)		500	821	3	2
Calibration Statistics (weighted MAE)		158	319	3	2

The results of Steps Four and Five are shown in two tables to account for the variation in scores between the two scenarios (Table 6 and Table 7). Table 6 shows the results of Scenario 1 for, minimum flow conditions at control points, and in Table 7 shows Scenario 2, for new reservoir storage either by expanding an existing reservoir or adding a new reservoir.

Table 6: Results of scoring for Scenario 1, minimum flow conditions at control points.

	MODSIM			RiverWare		
Selection Metrics	Weight (%)	Score	Weighted Score	Weight (%)	Score	Weighted Score
New model development time	9.7	4.3	0.4	9.7	2.7	0.3
Has known deficiencies/bugs	13.9	4.0	0.6	13.9	3.7	0.5
Institutional knowledge	9.1	3.0	0.3	9.1	3.7	0.3
User support	8.5	1.0	0.1	8.5	5.0	0.4
Transparency	4.9	2.0	0.1	4.9	4.7	0.2
Cost	12.1	5.0	0.6	12.1	1.0	0.1
GUI Features	1.4	3.2	0.0	1.4	4.3	0.1
Integration with other tools	3.7	3.3	0.1	3.7	2.0	0.1
Model scenario development time	12.7	4.0	0.5	12.7	1.0	0.1
Model run time	10.9	4.0	0.4	10.9	1.0	0.1
Calibration quality	13.3	3.0	0.4	13.3	2.0	0.3
Sum			**3.5**			**2.5**

Table 7: Results of Scoring for Scenario 2, New Reservoir Storage.

	MODSIM			RiverWare		
Selection Metrics	Weight (%)	Score	Weighted Score	Weight (%)	Score	Weighted Score
New model development time	9.7	4.3	0.4	9.7	2.7	0.3
Has known deficiencies/bugs	13.9	4.0	0.6	13.9	3.7	0.5
Institutional knowledge	9.1	3.0	0.3	9.1	3.7	0.3
User support	8.5	1.0	0.1	8.5	5.0	0.4
Transparency	4.9	2.0	0.1	4.9	4.7	0.2
Cost	12.1	5.0	0.6	12.1	1.0	0.1
GUI Features	1.4	3.2	0.0	1.4	4.3	0.1
Integration with other tools	3.7	3.3	0.1	3.7	2.0	0.1
Model scenario development time	12.7	2.5	0.3	12.7	2.5	0.3
Model run time	10.9	4.0	0.4	10.9	1.0	0.1
Calibration quality	13.3	3.0	0.4	13.3	2.0	0.3
Sum			**3.3**			**2.7**

The results of both scenarios show that MODSIM would be the preferred tool for the modelers that were surveyed for this study. The results of this survey may have been different had different modelers been surveyed or if the models were to be used for different scenarios.

DISCUSSION

The results of the decision matrix analysis presented in this paper were specific to the case study and were presented to simply illustrate the usefulness of decision matrix analysis technique when trying to select a river-reservoir modeling platform. The power of this method is revealed by focusing the opinions and values of multiple people down to a single comparable number for each modeling platform.

The decision metrics that were selected for this analysis were chosen based on the experience and knowledge of the modelers that were surveyed for this case study. They included metrics that were easily quantifiable using the existing models, like model run time, and metrics that were not easily quantifiable, like user support.

Assigning weights to model metrics provided a mechanism for assigning weights to the metrics using AHP. This method attempts to remove the subjectivity from the assignment of weights by simply asking the modeler to determine which metrics are more highly valued than others.

For the case study, the AHP process was completed by three separate modelers and the results were averaged together for the final selection. The AHP process revealed that there were several evaluation metrics that the modelers agreed upon the relative importance when it came to selecting a river-reservoir modeling platform, and therefore they had a smaller difference between the maximum and minimum weights assigned by each modeler. These included: new model development time, GUI features, model scenario development time, and model run time.

The relative importance of other metrics was not agreed upon by the modelers including: has known bugs or deficiencies, user support, and cost. This is an interesting result of this analysis because it shows that the values of modelers with similar backgrounds can vary. This method can help to provide a unified answer that can be used to select a modeling platform, even when opinions vary widely about the importance of each metric. A similar outcome can be seen in step three where the modelers were asked to score each model for the two modeling scenarios.

This decision matrix analysis was able to provide flexibility to evaluate two specific modeling scenarios; minimum flow conditions and new reservoir storage. This process allows modelers to hone in on model selection based on specific modeling tasks or simulation goals and get meaningful quantitative feedback. This is important because some river-reservoir models do not provide capabilities that need to be simulated or could be better simulated in another modeling platform.

The two above mentioned scenarios that were evaluated for the case study scored similarly for all except one of the metrics and MODSIM received more favorable scores for both scenarios. This is because the highest weighted metrics, has known deficiencies/bugs, cost, model scenario development time, and calibration quality, also were scored in favor of MODSIM by the surveyed modelers. The results of this analysis reflect the values of the modelers that were surveyed and might be different if different modelers or non-modelers were surveyed.

Two of the metrics were quantified using the Deschutes models, model scenario development time and calibration quality. Since a single modeler experienced in both RiverWare and MODSIM developed the model scenarios, a large amount of subjectivity was removed from this analysis. However, it could be argued that another modeler may develop the scenarios faster or slower in either modeling tool. It could also be argued that the calibration quality could be improved if more time was spent on the RiverWare model rules. However, this is a good illustration of the situation that is likely to exist when a decision is being made about an appropriate river-reservoir modeling platform.

It is important to recognize that when selecting a modeling tool for any application, there are certain criteria that may rule out all other options. For example, if the application of the tool requires that it be able to simulate daily flow values and stream temperature and there is only one tool that meets that criteria. If that is the case, this pass-fail type of criteria may preclude any other analysis of other qualitative or quantitative metrics. The method presented in this paper is designed for the case when models have similar capabilities and both could perform the required analysis for study in question.

It is also important to recognize that the analysis presented in this paper was conducted with three modelers that were not chosen at random but were chosen based on their expertise and availability. The results presented in this paper were based on their individual preferences and it is impossible to say that if others were surveyed, they may have chosen criteria, weights, and scores that would result in an opposite model selection. In addition, if the same modelers had been asked to select a tool for other scenarios or applications, they may have chosen different criteria, weights, and scores that would have resulted in a different outcome. The purpose of this paper was not to select the best or most useful river-reservoir model, but rather to show how the decision matrix analysis could be used to select a modeling tool.

CONCLUSIONS

The decision matrix analysis with AHP was designed to provide a generic quantitative framework to effectively compare two river-reservoir modeling platforms. The approach allows for the assignment of numerical values to modeler specific metrics that previously could only be discussed in a qualitative fashion. The framework provides a more defensible mechanism for justifying the use of one modeling platform over another.

The decision matrix analysis was applied to select between MODSIM and RiverWare models that had been developed for the Deschutes River Basin. The approach was applied by three modelers in order to determine which modeling platform would be more appropriate for two hypothetical scenarios, a new storage reservoir and a new minimum instream flow requirement. The surveyed modelers had varied levels of experience in the Deschutes Basin and with MODSIM and RiverWare. For both scenarios, MODSIM scored higher in the decision matrix analysis. It is possible that if different modelers used the same process to analyze the river-reservoir modeling tools, RiverWare may score higher.

This decision matrix analysis was set up generically so that it could be applied, not only to two river-reservoir modeling platforms, but to any two models that are being compared for use in a particular situation. In the test case, both models were already developed for the Deschutes Basin, but the method could just as easily be applied to determine the appropriate modeling platform before models are developed. The method is also flexible enough that modelers or non-modelers can contribute to the list of metrics and the scoring thus providing cohesive collaboration and decision making.

ACKNOWLEDGMENTS

Thank you to Bob Lounsbury, Marshall Richmond, Cindy Rakowski, Simon Geerlofs, and Nathalie Vision for their guidance and technical input.

REFERENCES

Aylward, B. and D. Newton. 2006. Long-Range Water Resources Management in Central Oregon: Balancing Supply and Demand in the Deschutes Basin. Final Report, Deschutes Water Alliance, Bend, Oregon. Unpublished manuscript.

Berhe, F. T., A.M. Melesse, D. Hailu, and Y. Sileshi. 2013 Modsim-based water allocation modeling of awash river basin, ethiopia. CATENA, 109(0), 118-128. Retrieved from http://dx.doi.org/10.1016/j.catena.2013.04.007

Bamberger, M., Rugh, J. and Mabry, L. 2012 . RealWorld Evaluation Second Edition. Sage Publications.

Bureau of Reclamation (Reclamation). 2005. Ongoing Operation and maintenance of the Deschutes River Basin Projects. Pacific Northwest Region. February 2005.

Bureau of Reclamation (Reclamation). 2011. Climate and Hydrology Datasets for Use in the RMJOC Agencies' Longer-Term Planning Studies: Part 2 - Reservoir Operations Assessment for Reclamation Tributary Basins. Pacific Northwest Region. January 2011.

Doherty, J. 2005. PEST: Model-Independent Parameter Estimation User Manual 5th Edition. Watermark Numerical Computing.

Draper, A. J., A. Munévar, S. Arora, E. Reyes, N. Parker, F. Chung, and L. Peterson, 2004. CalSim: A generalized model for reservoir system analysis.” Journal of Water Resources Planning and Management, ASCE, Vol. 130, No. 6, pp. 480-489.

DHI, 2006. MIKE BASIN Users Manual. DHI Water and Environment, copenhagen, Denmark

Driscoll, D. Appiah-Yeboah, A., P. and Rupert, D. 2007. Merging Qualitative and Quantitative Data in Mixed Methods Research: How To and Why Not. Ecological and Environmental Anthropology (University of Georgia).Paper 18.http://digitalcommons.unl.edu/icwdmeea/18

Hameed, T. and R. O’Neill. 2007. River management decision modelling in IQQM, Proceedings of International Congress on Modelling and Simulation, A. Zerger and R. Argent (eds.), Modelling and Simulation Society of Australia and New Zealand, 170-176.

Gannett, M.W., Lite, K.E., Jr., Morgan, D.S., and Collins, C.A., 2001, Ground-water hydrology of the upper Deschutes Basin, Oregon: U.S. Geological Survey Water Resources Investigation Report 00-4162, 77 p.

Henshaw, F. F., J. H. Lewis, and E. J. McCaustland. 1914. Deschutes River, Oregon, and its utilization, U.S. Geol. Surv. Water Supply Pap. 344, 200 pp.

Inter-district Contract, Data January 4, 1938, between Central Oregon Irrigation District, Jefferson Water Conservation District, Arnold Irrigation District and Crook County Improvement District No 1.

Johnson, J. and LaMarche. 2013. Water Management Modeling in the Deschutes Basin using MODSIM. Paper presented at the American Water Resources 2013 Annual Water Resources Conference, Nov 5, 2013. Portland, OR.

LaMarche, J. 2001. Upper and Middle Deschutes Basin surface water distribution model. Surface Water Open File Report #SW02-001. Published by Oregon Water Resources Department. http://www1.wrd.state.or.us/pdfs/reports/SW02-001.pd

Larson KB, SE Niehus, JD Tagestad, KD Ham, SH Geerlofs, and MC Richmond. 2014. The Integrated Basin-Scale Opportunity Assessment Initiative: Pilot Assessment for the Deschutes River Basin . PNNL-23197, Pacific Northwest National Laboratory, Richland, WA.

MODSIM. 2014. http://modsim.engr.colostate.edu/

O'Connor, J.E., Grant, G.E. and T.L. Haluska, 2003a. A peculiar river: Geology, geomorphology, and hydrology of the Deschutes River, Oregon: American Geophysical Union Water Science and Application Series, no. 7-29 p.

Pugh, S., D. Clausing, R. Andrade (1996). Creating Innovative Products Using Total Design. Addison Wesley Longman; 1st edition (April 24, 1996). ISBN-13: 978-0201634853.

Russell, I. C. 1905. Preliminary report on the geology and water resources of central Oregon, U.S. Geol.Surv.Bull. 252, 138 pp.

RiverWare. 2014. http://www.riverware.org/

Saaty, T.L. 1977. A Scaling Method for Priorities in Hierarchical Structures, Journal of Mathematical Psychology,15: 57-68.

Saaty, T.L. 1994.Fundamentals of Decision Making and Priority Theory with the AHP.RWS Publications, Pittsburgh, PA, U.S.A.

Shafer, J., and Labadie, J., 1978 Synthesis and Calibration of a River Basin Water Management Model, Completion Report No. 89. Colorado Water Resources Research Institute, Colorado State University, Ft. Collins, CO.

Sulis, A., and Sechi, G.M., 2012. Comparison of generic simulation models for water resource systems. Environmental Modelling and Software, 40(2013_ 214-225.

USDA. 1996a. Upper Deschutes Wild and Scenic River: Record of Decision and Final Environmental Impact Statement. Bend, OR: United States Department of Agriculture Forest Service. 316 pp.

Wurbs, R.A. 2012. Reservoir/River System Management Models. Texas Water Journal 3(1), 26-41.

Yates, D., D. Purkey, J. Sieber, A. Huber-Lee, and H. Galbraith, 2005a.WEAP21—A demand-, priority-, and preference-driven water planning model. part 2: Aiding freshwater ecosystem service evaluation, Water Int., 30, 501–512.

Zagona, E. T. Fulp, R. Shane, T. Magee, and H. M. Goranflo. 2001. Riverware: A generalized tool for complex reservoir systems modeling. Journal of the American Water Resources Association, 37(4), 913-929. doi: DOI: 10.1111/j.1752-1688.2001.tb05522.x

Zagona, E., E. Kandl, J. Carron and S. Bowser. 2010. Water Accounting and Allocation in RiverWare. In Proc. of the Fourth Federal Interagency Hydrologic Modeling Conference, Las Vegas, Jun 27 – Jul 1, 2010.