RECLAMATION

Managing Water in the West

Investigation of Software Tools for Visualization of Results from Water Resources Planning Models

Research and Development Office
Science and Technology Program
(Scoping Report) ST-2015-3265-1

U.S. Department of the Interior
Bureau of Reclamation
Research and Development Office

February 2016

Mission Statements

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.

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.

REPORT DOCUMENTATION PAGE		*Form Approved OMB No. 0704-0188*
T1. REPORT DATE February 2016	**T2. REPORT TYPE** Research (Scoping)	**T3. DATES COVERED FY '15**
T4. TITLE AND SUBTITLE Investigation of Software Tools for Visualization of Results from Water Resources Planning Models		**5a. CONTRACT NUMBER:** WOID: Z3265 Fund: 15XR0680A1 WBS: RY1541DS201513265
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER 1541 (S&T)
6. AUTHOR(S) David O'Connor (916) 978-5188 doconnor@usbr.gov		**5d. PROJECT NUMBER** 3265
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER MP-740
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Thomas Fitzhugh (initial PI, no longer with USBR, formerly MP-740)** **Nancy Parker (TSC), Cameron Bracken (TSC), Travis Yonts (MP-740),** **David O'Connor (second PI, MP-740), Karl Tarbet (PN)**		**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) ST-2015-3265-1
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)*

The focus of this scoping study was preliminary testing of combinations of open source programming languages with open source database formats using Mid-Pacific Region river temperature model output as example model output, to estimate the time and resources needed for more in depth open source data visualization package testing and water resource tool development. The native model output format used, HEC-DSS, (U.S. Army Corps of Engineers' Hydrologic Engineering Center Data Storage System) is typical to water resource models. The first project step involved development of tools to efficiently convert data in HEC-DSS format into open source database formats (HEC-DSS source code is proprietary). Popular open source database formats NetCDF and SQLite were chosen based on project research and team experience, and we accessed and processed the data for data visualization testing via the popular open source programming languages R and Python, each of which contain open source, robust packages for data visualization.

Tools to convert HEC-DSS format data to NetCDF and SQLite formats were developed and are provided and documented here: https://github.com/usbr/convertdss. In addition, USBR's Pisces data visualization application was updated to read data in SQLite format (http://www.usbr.gov/pn/hydromet/pisces/). The second project step involved data analysis and visualization using R and Python with data in NetCDF and SQLite format. All combinations yielded promising results, for both database and data visualization performance, especially considering the availability of open source interactive interfaces ("data dashboards") that could likely be easily paired with R and Python scripting. Follow up project (USBR S&T: 4878) was not awarded funding in 2016.

15. SUBJECT TERMS Database, Data Visualization, Open data

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON David B. O'Connor**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	**U**	11	**19b. TELEPHONE NUMBER 303-445-2125**

PEER REVIEW DOCUMENTATION

Project and Document Information

Project Name: *Investigation of Software Tools for Visualization of Results from Water Resources Planning Models*

Document Name: *Investigation of Software Tools for Visualization of Results from Water Resources Planning Models*

WOID: *Z3265*

Document Author: *David B. O'Connor*

Document date: *4 March 2016*

Peer Reviewer: *Jun Wang, PhD, PE*

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 (Signature) Jun Wang Date reviewed 3/4/2016

Report

Water resource planning studies typically require multiple runs of complex models that have large numbers of inputs and outputs. Efficiently analyzing results of these models can be challenging because of the data volume and complexity of model behavior involved. For effective model use, innovative approaches are typically necessary for quickly summarizing and visualizing model output and determining primary factors influencing model behavior. Existing standard USBR software (eg HEC-DSS) is generally limited in terms of one or more of the following factors: (1) the volume of data that can be efficiently handled, (2) availability of visualization options relevant to water resource analysis, and (3) availability of a scripting language for customization and control diagnostics. High end database and visualization tools are commonly cost prohibitive and proprietary, which limits their usefulness in situations where model results need to be distributed.

Hence we realized a need for open source, generic database and visualization packages with customization capabilities that could be used to analyze results from a variety of water resources planning models. While outputs from water resources planning models vary depending on model type (e.g., operational, groundwater, hydrodynamic) it is believed that a common visualization framework can be developed that is both user-friendly and applicable to results from most water resource models, allowing not only visualization of results from individual models but also integration of results from different models. A scoping level study was conducted, so analysis is incomplete. However, important conclusions were drawn and some effective tools were developed.

The focus of this scoping study was preliminary testing of combinations of open source programming languages with open source database formats using Mid-Pacific Region river temperature model output as example model output, to estimate the time and resources needed for more in depth open source data visualization package testing and water resource tool development. The native model output format used, HEC-DSS, (U.S. Army Corps of Engineers' Hydrologic Engineering Center Data Storage System) is typical to water resource models. The first project step involved development of tools to efficiently convert data in HEC-DSS format into open source database formats (HEC-DSS source code is proprietary). Popular open source database formats NetCDF and SQLite were chosen based on project research and team experience, and we accessed and processed the data for data visualization testing via the popular open source programming languages R and Python, each of which contain open source, robust packages for data visualization.

Tools to convert DSS format data to NetCDF or SQLite formats were developed and are provided and documented here: https://github.com/usbr/convertdss. In addition, USBR's Pisces data visualization application was updated to read data in SQLite schema, and is documented here: www.usbr.gov/pn/hydromet/pisces/ .

No weaknesses in any combination of programming language and database format were identified. It is possible that NetCDF handles gridded data more efficiently than SQLite, but full testing was not conducted. Full development of flexible tools with modern open source data dashboards (interactive interfaces such as Shiny, iPython, Gadfly) using these platforms appeared very promising. Both R and Python contain continuously growing and well-tested libraries of statistical, graphical, and geoprocessing tools (e.g., RSQLite, PyTables) that very efficiently process data of virtually any volume typical to modern water resource studies. In addition, near the project's completion, the project team learned of the quickly growing popularity of NASA's open source database format HDF5, but time constraints did not allow for testing. The follow-up proposal, Project 4878, titled Model Agnostic Tools for Visualizing Water Resource Model Data, was not awarded funding in 2016.