UtahState
UNIVERSITY

David Rosenberg
Utah Water Research Laboratory
8200 Old Main Hill
Logan, UT 84322-8200
Telephone: (435) 797-8689
FAX: (435) 797-3663
Email: david.rosenberg@usu.edu
http://www.engr.usu.edu/cee/faculty/derosenberg/

February 24, 2015

SUBJECT: Re-Submission of Manuscript #2013WR014667RR to *Water Resources Research.*

Ximing Cai, Editor in Chief
Water Resources Research
American Geophysical Union

Dear Dr. Cai

Please find attached an electronic copy of the revised manuscript #2013WR014667RR, titled "Blended near-optimal alternative generation, visualization, and interaction for water resources decision making", by myself as the sole author with I am resubmitting for publication in *Water Resources Research*.

Below, please also find a listing of the revisions in response to the most recent round of comments from Reviewer #5. Should you need to contact me, please do so via email at david.rosenberg@usu.edu.

Sincerely,

David E. Rosenberg (the corresponding author)
Associate professor

Attachments

- Revised manuscript
- List of revisions and response to reviewer #5′s comments

Response #3 to Reviewers' Comments

February 2015

General Comments

Again, thank you for this opportunity to further revise and improve the manuscript. This version of manuscript adds improved phrasing, sentences, and text to clarify or expand upon issues raised by the reviewer's comments. Below, I elaborate on these changes. I list reviewer comments in *blue italic*, and indent my further responses in plain black text.

Reviewer #5 (Formal Review for Authors):

The manuscript presents a mathematical programming approach conceived to identify near-optimal solutions (alternatives) in water resources decision-making problems. In particular, the tool adopts a Monte-Carlo Markov Chain sampling to identify a large number of solutions that are spread through the objective and decision space of the near-optimal region. The approach is complemented by a visual analytic tool (built on parallel coordinate visualization) that 1) shows the relation between the value of the objective function(s) and decision variables, and 2) supports the analyst/decision-maker in the selection of sensible solutions, exploration of the near-optimal region and re-formulation of the optimization problem. I found this topic quite interesting and of potential relevance to the water resources management community, mostly because un-modelled objectives/constraints can arise during the formulation of decision-making problems and hence influence the final results.

> Thank you for the succinct summary of the manuscript and its contribution. This stated contribution is exactly the intended contribution.

This said, I believe that the contribution and overall quality of the manuscript does not meet the requirements of this journal. The methodological contribution is rather marginal, since the proposed approach leverages existing and well-established techniques (e.g., linear programming and Monte-Carlo Markov Chain sampling) to explore the near-optimal region.

> Yes, the manuscript uses existing optimization and applies Monte-Carlo Markov Chain sampling methods. However, the application is novel and important as it is a first application of Monte-Carlo Markov Chain methods to generate alternatives that span the full extent of the near-optimal region. And it is this application that overcomes a shortcoming of prior near-optimal methods which is that they can only generate a limited number of alternatives in reasonable time that give an incomplete picture of the region. Further, the parallel coordinate visualization component allows the user to make sense of all the Monte-Carlo Markov chain sampled alternatives to see the full near-optimal region. Together, the generation and visualization tools show the full extent of the near-optimal region and a much broader and more flexible set of near-optimal alternatives than prior near-optimal methods.

To better emphasize this contribution, I have reordered, condensed, and streamlined several parts of the introduction. The review of near-optimal methods (lines 94-134) is now followed by:

- Line 136: "To generate a more comprehensive set of alternatives in reasonable time…",
- A review Monte-Carlo Markov chain methods (lines 137-149), and
- Line 140 which explains that Monte-Carlo Markov chain methods "have yet to applied to near-optimal problems."

The applicability of the approach is indeed limited to single- and multi-objective linear (and mixed-integer) problems with a convex and bounded feasible region, so it cannot be combined with any of the stochastic optimization methods commonly adopted to handle uncertainty or derivative-free problems.

Yes, the manuscript only applies the methods to single- and multi-objective linear and mixed-integer problems with convex bounded feasible regions. These examples are provided as demonstrations of the method and manuscript notes in multiple places that further work remains to apply the near-optimal stratified random sampling alternative generation techniques to other types of problems. To further emphasize this point, I have added the adjectives "example" and "linear-programming" in front of "water quality management problem" at several points throughout the manuscript including in the key points (Line 30), abstract (Line 68), introduction (Line 183), application (Line 442), and conclusions (Line 764). I also substituted the word "three" for different" on line 304 so the text now reads "… for three classes of optimization problems."

The water supply/demand planning problem for Amman, Jordan that is briefly discussed is actually a mixed-integer, two-stage stochastic program with recourse so the reviewer's statement that the near-optimal methods cannot be combined with any stochastic methods is incorrect. I have added text on lines 321-324 and 736 that more fully describe the Amman, Jordan program.

The approach can work with derivative free methods so long as the derivative free optimization method can generate the optimal solution. Then, the tools simply sample objective function values between the optimal objective function value and near-optimal tolerance and proceed according to the steps listed in Section 3 for alternative generations. Please see also the response to the next comment.

For example, I wonder how the approach could work in the presence of multiple local/global optima.

The tools work as expected both for multiple global optima and local optima. The tools work because global optima are part of the near-optimal region, local optima are as well (when their objective function values are in the range defined by the optimal objective function value and near-optimal tolerance), and the tools don't use objective function slope/derivative information to find solutions/alternatives as conventional optimization and optimization-based near-optimal approaches do. Instead, the new tools sample across the maximum extents of the objective function (from the optimal value to near-optimal tolerance). The text now better explains how results include multiple global/local optimal and how to use the tools to generate these optima:

- Adding the word "global" to the list of methods on line 193,
- Extending line 630 that discusses the Figure 7 results to explain the darkest green lines show the sub-set of alternatives that are either multiple global optima (have the same optimal removal cost but a different mix of phosphorus removal practices) or alternatives with removal costs that are very near the optimal removal costs.
- Expanding the caption for Figure 7 to further explain that the darkest green lines include multiple global optima.
- Adding lines 632 which reads: "To highlight only the problem's multiple global optima, set the tolerable deviation parameter value to 1.0 and resample (results not shown)."

The figure below compares the results of such sampling for the multiple global optima (purple lines) to alternatives generated from the broader near-optimal region (green lines).

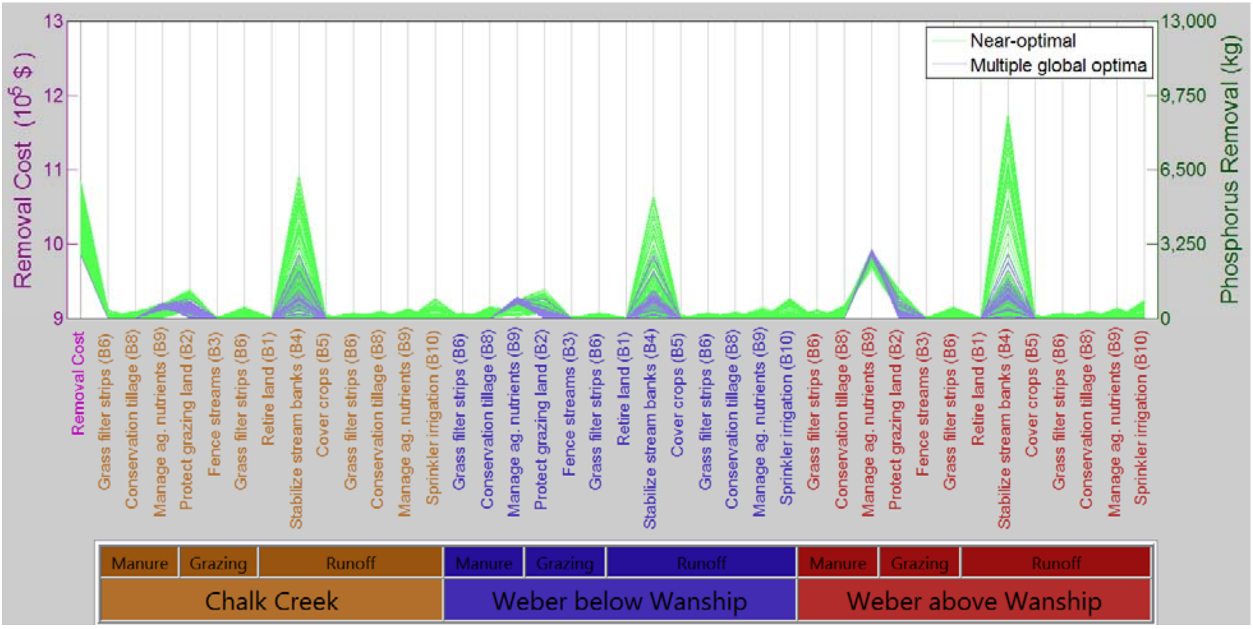

Figure 1 of the Letter of Response. Comparing multiple global optima (purple lines) and near-optimal alternatives generated for the Echo Reservoir phosphorus removal problem.

Lastly, please also note Figures 3-7 already show generated alternatives have objective function values (removal costs) that span the near-optimal range between 100% and 110% of the optimal (global) removal cost and will include local optimum should such local optima fall in that range. Obviously, a linear programming problem does not have local optima, but the point is that the new stratified sampling approach can still capture local optima (should they exist) even without making use of the derivative or curvature information about the objective function. Further, local optima are often of mathematical interest (i.e., the user wants to know if the algorithm ends up at the local rather than the global optima), these local optima are of little management interest. If a local optimum is of management interest, the user can also simply add constraints to model formulation that make the local optimum the global optimum (within the newly constrained feasible region), then use the tools to explore the near-optimal region in the neighborhood of this (now) new global optimum.

Furthermore, the scalability of the approach to multi-objective problems is poorly explored (only a two-objective problem is considered in Section 6).

Yes, correct, only a two-objective problem is demonstrated. This choice reflects a desire to use a simple example to demonstrate the key aspects of the multi-objective extensions which is now noted on lines 451-458. Also, I feel it is unrealistic to expect one journal article to singularly address all the different types of programming problems including, for example, many-objective problems. The manuscript already considers single objective, two-stage stochastic, mixed-integer, and multi-objective problems.

A similar comment applies to the parameterization of the Monte-Carlo Markov Chain sampling technique, which is not discussed in the manuscript.

Line 298 now fully describes the three principal parameters of the alternative generation method: the (i) total number of samples, (ii) stratified samples per decision variable or objective function group, and (iii) Monte-Carlo Markov chain samples per stratified sample (Monte Carlo Markov chain length). The example application in Section 6.2 (line 498) now explains that the total samples were split 15%/85% between the objective and decision variables groups and two Monte-Carlo Markov chain samples were drawn for each stratified sample within the decision variable groups.

The case study presented in Section 6 is a highly simplified linear program: I believe that it can be used to illustrate the main features of the tool, but it should be considered that it does not totally represent the complexity of water resources management problems.

Yes, correct and a good observation. As noted above in response to the prior comment on single and multi-objective programs, I have added the words "example" and "linear programming" in the key points (Line 30), abstract (Line 68), introduction (Line 183), application (Line 442), and conclusions (Line 764). Additionally, in section 6, I added line 455 which states that "Although the application does not represent the total complexity of the Echo Reservoir phosphorus removal problem, the results and discussion highlight several methodological contributions and management insights gleaned by use of the new blended near-optimal tools".

In addition, some aspects of the case study are questionable. For example, the decision-making problem is first formulated as a single-objective optimization problem (i.e., minimization of the removal cost with a constraint on the load reduction) and then as a two-objective problem aimed at minimizing the removal cost and maximizing the phosphorus removed. This second objective is considered as an un-modelled issue (Section 6.4), while I think that it is a common modelling *practice to substitute some constraints with objective functions in order to explore the decision/objective space in a more exhaustive way. In this case, for example, the problem should be formulated from the very beginning as a two-objective one to study the trade-off between removal cost and phosphorus removed.*

Yes, correct that if the second objective (phosphorus removed) was known a-priori, then the objective should be included from the beginning as a two-objective problem. However, here, the problem was first formulated as single-objective problem (cost removal) because cost was the sole criteria specified in the TMDL. The un-modelled issue (phosphorus removed) was not known a-priori and was only discovered through the process of using the near-optimal tools to generate, visualize, and interactively explore the near-optimal region. It is easy and natural to say in hindsight that the second objective should be included from the beginning. Further, the statement that the issue should be included from the outset additionally confirms the value of the near-optimal tools and results to identify un-modelled issues and use identified issues to update and improve the model formulation.

To clarify the motivation for the single-objective problem, I have added line 470 that reads: “A single objective function minimizes removal costs and reflects the TMDL criteria of cost to select practices to reduce phosphorus loads.”

Please additional note: Figures 5 and 6 already show the pareto front (black lines and triangles in Figures 5 and 6) that was generated by standard multi-objective methods (after the 2nd objective to increase the phosphorus removed was identified and added as a new modelled issue). If both objectives were known a priori, this front could be generated right away. But even after generating the front, the near-optimal alternatives in the figures further show numerous additional alternatives that lie on and near the front and represent additional flexibility managers have in locating phosphorus removal practices beyond the options offered by the pareto set. This flexibility relates to where to site practices, the mixture of practices, and implementation levels. The figures thus show the further benefit of the near-optimal tools to identify additional un-modeled issues beyond the initial modeled issues. Thus, even if the two objectives are known a priori, there are still *additional* un-modeled issues that persist and that near-optimal tools can help identify (see lines 580-614).

Finally, the interaction with the decision-maker is hypothesized but not demonstrated.

Yes, correct, this limitation is noted prominently in the description in Section 6 and discussion in Section 7. We look forward to explore these interactions further in future work.

The manuscript is well organized and appropriate references are given, while the description of the methodology (Section 3) and Interactive exploration (Section 5) should be improved. The former lacks of justifications aimed at motivating the adoption of these techniques, while the latter should be supported by a few graphical examples that illustrate and demonstrate the functionalities of the visual analytic tool.

The sentence at the beginning of section 3 (lines 271-275) was reordered and expanded to better justify the method. Six of the nine interactive tools are illustrated in Figures 3-7. In the first paragraph of Section 5 (line 386), I have added a sentence "Below, text describes nine interactive tools six of which are demonstrated later in Section 6." I have also edited text in Section 5 to more clearly describe the interactive tools and emphasize tools that are featured in the Figure 3-7 examples.