

## ***Interactive comment on “Household water use and conservation models using Monte Carlo techniques” by R. Cahill et al.***

### **Anonymous Referee #5**

Received and published: 12 June 2013

#### General Comments:

There is an underdeveloped argument made in this paper that the authors' “mechanistic” model, outfitted with parameters derived from a statistical sampling process (Monte Carlo method), are better than... regression analyses, other models? If it is better, there needs to be some measure by which the authors can argue it is better.

There is an equally underdeveloped argument for the importance or usefulness of the results, who/what models their results agree or disagree with, and what is novel about the findings. There is an argument made that the method itself is novel to household water use and conservation studies, but the methods themselves are not new.

In both the existing conditions and least-cost conservation models, the specific Monte  
C2436

Carlo methods used in each model are not clearly described. For example, in the existing conditions model, I read that parameter distributions are sampled, and each sample/iteration represents a modeled household for which water use is calculated. These samples, in aggregate, characterize a neighborhood's water use. What the authors mean by convergence in reference to this method remains unclear. The authors' application of MC methods in the least-cost model is also unclear.

The introduction is difficult to follow. Please provide introductory sentences to each paragraph stating the purpose of mentioning the literature cited, and how those references relate to the authors' study.

It does not become clear until section: 2.1.1 that what the authors have done in – at least with the existing conditions model (and I believe also with the least cost conservation model) - is modify/improve upon an existing model, and then applied and calibrated the model to new data. This should be stated in the Introduction.

The “basic modeling process” descriptions for both models are very helpful. These should be summarized (or simply just repeated in sentence form) sooner – in the introduction.

Was any sensitivity analysis performed on the model? Was it ever demonstrated that the number of parameters estimated for each model is necessary?

The authors list two DeOreo et al., 2011 references that are not distinguished when cited in the text (e.g. Page 4872; line 11, 16).

I'm not sure why the title references MC methods - there are a variety of methods used in their modeling process.

#### Specific Comments:

1. Page 4871; line 12-13: “effective” and “adequacy” are unclear terms; what is the measure of adequacy and effectiveness that the authors refer to?

2. Page 4871; line 20: I don't understand what the authors mean by "data management programs" – this is too vague a description.
3. Page 4871; line 23,26: What "measured data" is this referring to – household water use data?
4. Page 4871; line 24: a model does not "depict" a good fit; a model has a good fit (or not) according to a measure like  $R^2$ , AIC, BIC; model predictions/estimates demonstrate a good fit by minimization of a loss function like the RMSE. The language in this section makes it sound like the authors are not familiar with model validation.
5. Page 4872; line 15: "The strength of the regression analysis..." Why is this information provided? To what effect? Is this related to the "reasonable" performance mentioned in the next sentence? I can't tell.
6. Page 4872; line 23-25:
  - a. Delete this first sentence; the second and third say the same thing. Then, just mention further down in this paragraph that the authors' methods make use of more, newly available measurement data.
  - b. If this sentence is not deleted – it sounds like the authors are saying that "mechanistic" and "deductive" models are complimentary (or are synonyms); this does not make sense.
7. Page 4873; line 2:
  - a. Specify what Monte Carlo approach is taken (since this is the first mention made in the paper after the abstract); there are many applications of MC methods.
  - b. Please consider if use of the phrase "to include variability" is what the authors meant.
8. Page 4873, line 6-8: This sentence is confusing:
  - a. Does the "novel way" that differs from the "more deductive (statistically-based) ap-

C2438

proaches" mean that the author's model IS deductive, but less so than other models? MC sampling methods ARE a statistically-based approach, but it sounds like the authors are saying that their approach is "novel" because it is more "mechanistic". Please clarify.

9. Page 4873, line 18: what does it mean that the virtual households are "given" by an MC iteration? Still don't know what the "iteration" is.
10. Page 4875, line 12: explain what you mean by convergence; what are the convergence criteria?
11. Page 4875, line 16: "are used to capture uncertainty" – this is a weird way to phrase this, consider: "probability distributions capture variability in water use parameters"
12. Page 4875, line 19:
  - a. "end use studies" from which parameter distributions are formed: mention that these are the studies previously cited in the introduction, or cite them here if they are different.
  - b. "other literature" needs citation, or reference to where they are already cited.
  - c. Are the send use studies and other literature all specific to the study area used here? If not, need to justify how they still represent a relevant distribution for this paper's study area.
  - d. Re cite to Cahill, 2011 (here and elsewhere): The citations from the thesis should be provided here instead of the thesis.
13. Page 4875, line 23 – Page 4876, line 2: This makes sense, but would be unclear to someone not familiar with MC sampling of model parameter inputs. My understanding is this: this method captures study area water use by way of sampling from the full (empirically-based) distributional range of each parameter; individual realizations from parameter distributions, when plugged into a formula for water use, describe one realization of that water use. Many samples will therefore describe many realizations

C2439

of individual household water use that when combined, describe a community of water users = a modeled neighborhood based on real neighborhood data.

14. Page 4876, line 2-3: This point needs serious clarification and justification:

a. Clarify what correlation/covariance exists, between what parameters, and to what potential effect (in the model). It is known that correlation between model input parameters can bias results, and either increase or decrease variability in model output. This needs to be acknowledged, and the disregard of potential or known correlation between input parameters needs to be explicitly stated and justified. When correlation does exist between parameters, there are data transformation and sampling methods that can be applied.

b. The phrasing of this sentence: 'including covariance' and 'such relations exist' are vague, and makes it sound like the authors do not understand the problem of correlated model input parameters, or how it may be affecting their results.

15. Page 4876, line 13-15: Need to say why wet/dry seasons are relevant for urban household water use; I'm assuming it's because of the outdoor/landscaping water component, but need to say this if that's the case.

16. Page 4878, line 25: Please also state what is sampled in the MC process; presumably this model is "solved" with linear programming methods; what in the model list above (page 4877) is sampled iteratively?

17. Page 4879, line 20 - Page 4880, line 2: Are these methods the authors' own design, or do they stem from another model or method cited previously? If they stem from, or build upon, another model or method, please mention and cite (as was done in the 'existing conditions' model)

18. Page 4887, line 5-8: Again, did the authors considering conducting a sensitivity analysis of their model to its parameters to see if it's actually necessary to estimate/sample all parameters?

C2440

19. Page 4887, line 9: Nice that the authors mention this, it's a really important point.

20. Page 4887, line 20: clarify what you mean by robust; robust to what?

21. Page 4887, line 7-8: "more" detailed and "more" mechanistic than what? Again, why is this better, or for what purposes/applications is more detailed and mechanistic better? Need to be explicit about that.

#### Technical Corrections

1. Page 4870; line 4-6: this sentence is awkward, re-phrase; a suggested re-phrasing:

a. "This study simulates water use in a single-family residential neighborhood using end-water-use parameter probability distributions generated from Monte Carlo sampling."

b. Monte Carlo methods are "iterative" by nature; remove word "iterative"

2. Page 4870; line 7: "existing conditions" of/pertaining to what? Define this.

3. Page 4870; line 22:

a. What kind of "devices" does this refer to? Specify.

b. Remove "quite" – unnecessary.

4. Page 4871; line 3: Add comma after "homes"

5. Page 4871; line 5: Remove "etc."

6. Page 4871; line 14-17: "Sauri (2003) presents a qualitative approach..." to what?

7. Page 4871; line 21-25: Does the "end-use model" discussed refer to Gumbo et al. (2003) or Blokker et al. (2010)? The two sentences following the citation (Gumbo et al., 2003) on line 21 are unclear as to their reference.

8. Page 4871; line 26:

C2441

- a. does “it” refer to aggregate demand?
  - b. “close to”: what does this mean?
9. Page 4872; line 1: In reference to “water end use data”: at the beginning of the introduction, define “water end use(s)”, and list a couple specific examples. Then throughout the rest of the text, refer to this term simply as “end use”.
10. Page 4872; line 1: models don’t “attempt” things. . . . Replace “some models have attempted to estimate” with “models estimate”
11. Page 4872; line 6-9:
- a. replace “has use” with “is useful for”;
  - b. “prediction” of what (the conservation potential)?; “estimation” of what (again, conservation potential)? If not the same thing, specify; if the same thing, just say “long-term, large-scale prediction of X”
12. Page 4872; line 6-9:
- a. do the authors really mean “much” or “any”?
  - b. Does “which can cause varying effectiveness. . .” mean “these models therefore result in varying effectiveness [again, what measure of effectiveness]. . .”. Make the “which. . .” statement a new sentence; this sentence is over-long.
13. Page 4872; line 10: add comma after “Still”
14. Page 4872; line 17-19: This sentence is awkwardly phrased and ungrammatical; some suggestions:
- a. Replace “have a reasonable performance” with “perform reasonably” or “show reasonable performance”
  - b. Replace “are useful to estimate” with “useful in estimating”

C2442

- c. Add “of” after “effectiveness” in line 18.
15. Page 4872; line 20: “usually absent. . .” from these models, or from what?
16. Page 4872; line 20: Use of “However” in this line (after the previous line starting with “However” also) is unclear.
17. Page 4873, line 24:
- a. Replace “builds” with “that builds” or “built”
  - b. Replace “finding” with “that finds”
18. Page 4873, line 26: should “of conservation savings” be “on conservation savings”?
19. Page 4874, line 3-4: this sentence is awkward, passive voice, and vague with respect to “insights” – say what insights are, or provide an example, in this sentence.
20. Page 4874, line 6: “Both models extend of a models” should be “Both models are extensions of models”
21. Page 4874, line 10: Still don’t know what the authors mean by “reasonable”
22. Page 4874, line 9: I believe these are the same “other” models mentioned before in the introduction; just call them something in the introduction so they can be specifically referred to here and anywhere else in the document.
23. Page 4874, line 11: Add comma after “yet”, and just make this a new sentence (starting with “Yet”).
24. Page 4874, line 16: Add comma after “CA”
25. Page 4874, line 23:
- a. Replace “to the” with “to this”
  - b. It is unclear whether or not the “study area” is just this neighborhood, or some area containing this neighborhood (that is larger than the neighborhood).

C2443

26. Page 4875, line 4-5: “are presented in their assumed market penetration rates”: this is unclear; what does “presented” mean in terms of the model?
27. Page 4875, line 6-7: re-phrase sentence:
- In “it allows”, is “it” the model?
  - Change to “. . . specific alternatives’ affect on total water use”
28. Page 4877, line 25: re-phrase “define and solve the optimization equations mixed linear programming”. I don’t understand what this means; did the authors mean “using mixed linear programming methods”?
29. Page 4878, line 22-23: please cite the “EBMUD water shortage contingency plant”
30. Page 4879, line 21:
- Insert “a” between “cutoff” and “proportion”, or re-phrase otherwise
31. Page 4880, line 9-10:
- Replace “to” with “that” in line 9
  - I find the first half of this paragraph (starting with line 9) unclear. The second half is clear.
32. Page 4885, line 3: define lphd and gphd.
33. Page 4885, line 16: add “of” between “rates” and “most”
34. Page 4885, line 28: re-phrase as “These preliminary model results also. . .”
35. Page 4886, line 4-5: re-phrase; awkward and ungrammatical.
36. Page 4886, line 23-25: to what affect?
37. Page 4887, line 6: add “it” after the comma following “studies”.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 4869, 2013.

C2444