Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-691-RC2, 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Value of seasonal streamflow forecasts in emergency response reservoir management" *by* Sean W. D. Turner et al.

Anonymous Referee #2

Received and published: 9 February 2017

The paper analyses the relationship between seasonal forecast skill and value for reservoir operation. In the first computational experiment, the paper considers a set of synthetic forecasts with various skill scores (from perfect forecast to no-skill forecast) and decreasing skill with lead time (up to 12 months). In the second experiment, the paper considers forecasts produced by the "Forecast Guided Stochastic Scenarios" (FoGSS) forecasting system. In the computational experiments, forecast skill is computed with respect to climatology, while forecast value is computed by comparing the reservoir operation performances obtained with a forecast informed adaptive control scheme and the performances obtained with a conventional operating scheme based on Stochastic Dynamic Programming (SDP) and Deterministic Dynamic Programming (DDP). Numerical results are shown for 4 reservoirs in Australia, whose features (e.g., capacity, demand) are altered to analyze a broader range of stress situations induced

by droughts (for a total of 32 different configurations).

General comments

The paper concerns an interesting and emerging topic: quantifying the actual improvement of seasonal forecasts in water resources management and reservoir operation in particular. I think the experimental setting is valuable and provides interesting results, which are worth to be published, but I disagree about the way the results are presented, which may lead, in my opinion, to misleading interpretations.

The paper distinguishes between "continually adjusted operation", when decisions are adjusted at frequent intervals, and "emergency response operation", when key decisions are taken infrequently. As an example of "continually adjusted operation", the paper considers a reservoir operated to track a constant target reservoir storage. As an example of "emergency response operation", the papers consider a reservoir operated to meet a constant supply demand. The paper concludes that there is a clear relationship between forecast skill and value in the case of "continually adjusted operation" while this is not in the case for "emergency response operation". Another conclusion is that, in this second type of operation, it is of fundamental importance considering skillful forecasts at certain, well defined moments when critical decisions are taken. In my opinion, the results should not be commented in light of the type of decision process (in both cases decision are taken with the same frequency, i.e., every month), rather considering the type of reservoir operating objective and its dynamical response to forecasts. When the reservoir is operated for water supply, the release decision does not directly depend on the predicted inflow, because the reservoir storage partially buffers the variability of the inflows. As a consequence, poor inflow forecasts may not directly affect the operating performance (of course, this depends on reservoir capacity and storage-demand ratio). Prolonged poor forecasts (on long lead times, for example) may instead negatively affect the operating performance. When the reservoir is operated to track a certain target storage, the release decision must mimic the inflow much more closely (the reservoir storage does not provide any buffer in this case). For this reason, a skillful forecast implies high operating performance and viceversa, which turns into a clearer relationship between forecast skill and value. On top of this, we should consider that the storage tracking objective depends basically on the forecast of few lead times (if not on the first lead time only), while the supply objective is more dependent on forecasts on longer lead times (which allow for the hedging decisions mentioned in the paper). This means that, in general, the tracking storage operation benefits of more skillful forecasts than the supply objective, because both FoGSS forecasts and synthetic forecasts show a decreasing skill with lead time (Fig. 2 and 3). This may explain the fact that a certain forecast skill may be linked to a wider range of forecast values when considering the water supply objective (Fig. 5). I argue whether a different definition of forecast skill should be used in the two cases (for example, skill computed on different lead times, i.e., the ones relevant to the dynamic of each reservoir objective). I'll provide more comments on these points in the following part.

Summarizing, I think that the paper deserve publication after a major revision. It might bring valuable insights on the topic of seasonal forecast value for reservoir operation, but the results should be discussed highlighting the effect of the different dynamics of the operating objectives (fast response to forecasts for storage tracking and slow response for water supply) on the forecast skill-value relationship, rather than focusing on the duality between "continually adjusted operation" and "emergency response operation", as it is in the current version of the manuscript.

Specific comments

Page 3, Lines 4-8. How is the sampling of the "error injected" parameter performed? Is there one sampling for each forecast? How does the error increase with lead time? The interpretation of the results would benefit by a more comprehensive description.

Page 3, Lines 25-26. "In months where forecasts are not informative, FoGSS is designed to return a climatological forecast." How is this performed? If this is the case, why does it happen that in Fig. 3 there are some negative skills (this is particularly

evident for the Eppalock reservoir)?

Page 4, Lines 24-26. In the standard operating policy, how is the release modulated when the demand can not be fully met? I would include the description of this aspect also in Appendix 1.

Page 4, Lines 26 and followings. How does the maximum release change in the different reservoir configurations? Is it linked to the reservoir capacity? Does the maximum release affect the ability of meeting the operating objectives (for example, should it be designed so to be able to always meet the demand)? The authors mention something in Appendix 2, but I would add a sentence in the main text as well.

Table 2. I think the readability of the table would benefit from a description of the meaning of "draft ratio", "drift" etc. in the caption. If I'm not wrong, "storage ratio" is not defined in the text. What does it represent? What are the actual reservoir features? It would be nice to have the figures for better understanding the reservoir settings in the different experiments. Is the "critical period" computed assuming no inflow to the reservoir or climatology or what else? The numbers seem in some cases very high and I don't fully understand how to interpret these large numbers (see also the following comment).

Figure 4. "Emergency response" case: why do the Burrinjuck and Eppalock reservoir empty, if their critical periods are 84 and 102 months respectively? "Continually adjusted" case: why does it happen that the reservoir storage is above the target but the releases are equal to 0? (see for example, Burrinjick and Eppalock reservoirs at the beginning of the time series) This behavior seems to be sub-optimal, because a release greater than zero would contribute in reducing the objective cost.

Page 6, Lines 13-14. "We test two operating objectives: one that rewards a judicious response to an emergency (emergency response objective) response and one that rewards judicious continual adjustments (continual adjustment objective)". As already mentioned, I don't agree with this terminology because decisions are taken at a constant pace when considering both the objectives. The “emergency response objective” could easily become a “continual adjustment objective” if, for example, the demand was (relatively) large and the reservoir (relatively) small, because the release decision could change much more frequently. Viceversa, the “continual adjustment objective“ in Fig. 4 – Eppalock reservoir behave as an “emergency response objective”, because the release decision is most of the time equal to the maximum release and does not change frequently, just because the storage is for a long time higher than the target. I would suggest changing this terminology because it is misleading in my opinion.

Page 7, Lines 1-2. “For the continually adjusted operating setting we find that the release must be adjusted constantly through the operating horizon to keep storage close to the target level of 75%”. As mentioned before, this objective is directly influenced by the variability of the inflow and reservoir release should change frequently to keep a constant storage. An example can be found in Fig.4 – Eppalock reservoir, where the storage is at the target storage, but the release keeps changing because of the inflow to the reservoir. I would include this explanation in the text.

Page 7, Lines 24-29. I would appreciate a discussion on the reason why the two objectives behave differently. The interpretation I propose in the “General comment” may represent a possible interpretation.

Page 7, Lines 29-32. I am not sure that the two objectives can be compared just on the basis of the value of the “injected forecast error”. In fact, as commented in the “General comment”, the “continually adjusted objective” is mainly driven by the forecast at one or few lead times, while the “emergency response objective” is driven by forecasts on longer lead times. On these lead times, the forecast skill is poorer by construction (if I understood how the synthetic forecasts were produced). What is the authors’ opinion on this? Would it make sense to compute the forecast skill on different lead times, e.g. short lead times for the first objective and long lead times for the other objective? If, for example, it would turn out that the forecast skill in Fig. 5 a-d for injected error equal to 0.2 is comparable to the forecast skill of Fig.5 e-h for injected error equal to 0.6, the

spread of the forecast value would be similar in the two cases.

Page 7, Lines 33-34. Why do the Burrinjuck and Serpentine reservoirs behave differently from the other two reservoirs? What are the features that they have in common which may justify the observed behavior? I would be nice if this result could be generalized.

Page 7, Lines 38 and followings. “These results show that the measure of forecast error, quality, skill or goodness-of-fit - if based on the entire forecast period - cannot predict accurately whether that forecast will be valuable in an emergency-type operational setting”. What do you mean when you write: “if based on the entire forecast period”?

Page 8, Lines 2-4. “This may be because the emergency response objective is constructed to be sensitive to a few serious shortfalls in meeting demand, while the continual adjusted objective rewards consistent performance over all the months assessed”. Both the objectives have a squared term, which should penalize to the same extent big deviations from the target (It could be different if there was an absolute value in Eq. 5, instead of the squared term). I think that the explanation may rely in the buffer effect of the reservoir, as explained in the “General comment”.

Page 8, Lines 15-19. This sentence is not fully clear to me.

Page 8, Line 25. “our own prior experiments with this approach”. Please provide a reference, if any.

Page 8, Line 35. I would explicitly state that the second experiment focuses mainly on the “emergency response objective”.

Figure 6. I would include a marker on the left hand side figures. I was surprised to see that there are more critical periods in case of low draft ratios (see fig. a, c, e). I would expect that higher draft ratios drive more critical situations. Could the authors comment on this?

Page 9, Line 10-12. “since the reservoir is drawn down at the end of the simulation, meaning the implications of late, sacrificial decisions are unavailable to quantify overall performance”. Does this mean that the penalty on the final storage is not considered in the optimization? If so, why is it not considered?

Page 9, Line 28. “the hedge comes too late at the end of 2006 ”. This is currently not visible in Fig. 8. I wonder if including the trajectories obtained with the perfect forecast might clarify what is “too late”.

Page 9, Line 31-32. I don’t understand the comment. The reservoir does not fully recover in both the trajectories. This might be because the inflow is too small in comparison of the reservoir storage.

Page 9, Lines 41-42. “forecast skill must be consistently available”. “Consistently” means “on long lead times”?

Page 10, Lines 14 and following. I would revise the conclusion (as well as the abstract and title) to soften (or remove) the distinction between the “continually adjusted operation” and “emergency response operation”, as already commented several times in this review.

Page 10, Lines 19-20. It seems that the results of the second experiment are driven mostly by the Millennium Drought. This exceptional sequence of several dry years contributes in enhancing the value of forecasts. Would the results be much different if this extremely exceptional event was not considered? It would be interesting to have a look at the figures computed without the last years, to have an idea of the forecast value in case of less extreme droughts.

Technical corrections

Page 6, Equation 4-5. Should “T” be “H”, given the notation in Eq. 3? Define all the variables, not only “D”.

Page 8, Line 12. “24 reservoirs” should be “32 reservoirs”.

Page 9, Line 20. Fig. 8 seems to be cited before Fig. 7 in the text.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-691, 2017.