Hydrol. Earth Syst. Sci. Discuss., 9, C338–C341, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C338/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Adaptive correction of deterministic models to produce accurate probabilistic forecasts" *by* P. J. Smith et al.

S. Regonda (Referee)

satish.regonda@noaa.gov

Received and published: 14 March 2012

The authors' work deals with producing probabilistic forecast from a real-time operational single-valued forecast – is a challenging problem for more than a few reasons. The real-time operational forecast corresponds to total uncertainty, which often times lacks a clear signal/pattern and makes it difficult to model. In addition, small sample size of forecast and corresponding verified observed flow imposes its own challenges. The proposed approach is new and interesting. The authors model observed flows as a function of deterministic model predictions plus noise, i.e., a simple linear regression model, and then model the regression coefficient, i.e., adaptive gain, using various time series techniques that allow the coefficient vary with time. The authors proposed two different methods to estimate parameters of the selected time series techniques. The

paper is concise and well organized, however, the content in sections 2.2 and 3 makes it tough to understand technical details and reproduce the technique. Improving these sections in terms of additional details and with a better presentation should improve the readability of the paper and assist if a person wants to reproduce the technique. My major concerns are (a) calculation of detailed verification metrics and (b) detailed discussion related to violation of assumptions of the methods. Here are my comments.

1. The authors produce the probabilistic forecast from the single-valued (deterministic) forecast. Therefore, it is suggested to calculate both deterministic and probabilistic verification metrics and verify that the probabilistic forecast is comparable with the deterministic forecast in the single-valued sense. In addition, it is suggested to evaluate the forecast in the context of probabilistic forecast aspects such as reliability, sharpness and event discrimination. Fraction of observations falling in a specific confidence interval does provide only partial information.

2. The authors mentioned about violation of assumptions of methods, i.e., prediction residuals not following a Gaussian distribution, etc. I agree that it is difficult to meet all assumptions of a method with real-time data, however, violation of assumptions warrants a detailed discussion. The forecast appears to perform well with respect to a few metrics, however, certain aspects of the forecast are affected with the violation of the assumptions – needs to be discussed and provide information what the user can expect in the context of flood forecasting.

3. Page 612, line 11-13: unless detailed evaluation is done, it is considered as an overstated statement.

4. It is suggested to develop plots for adaptive gains with respect to lead time, and draw interpretation of the values of adaptive gains with respect to the deterministic forecast skill. Are these adaptive gains follow a pattern and/or could be changed easily by the forecaster? (Page 610, line 28-29)

5. There is a clear difference in the performance of various time series techniques,

and is suggested to discuss in detail; only the poor performance of IRW and DLLT are mentioned. Why models with smoothing parameters have superior performance only at shorter lead times? And, it is not always consistent.

6. In general, using of one year for training and three years for validation considered as a stringent validation criterion, however, in this case, it is OK may be due to the year that used for training is a significant event. However, to test the robustness of the approach, it is suggested to present results from the leave-one-year out cross validation mode; drop a few years, calibrate the technique and then evaluate the method for the dropped period. Repeat the exercise for the entire period, thus all data will dropped and evaluated.

7. Page 609, line 24-25: It is suggested to exclude points that fall outside the Gaussian distribution, but, not mentioned how to deal with these types of forecasts. Additional guidance is needed particularly when these forecasts correspond to large flood events.

8. It is suggested to calculate conditional single-valued verification metrics such as mean error, mean absolute error and correlation for both ensemble mean and deterministic forecast; the conditioning can be done observed flow magnitude, forecast flow magnitude, rising limb, etc.

9. Page 611, lines 11-14: Gives impression that the rising limb is always correspond to long lead times. Is it always true?

10. Page 611, lines 21-23: Not completely true, did see bias in Figure 5 as well.

11. Page 609, lines 28: 'reasonable approximation' needs to be supported quantitatively

12. Cite literature related to time varying parameter techniques

13. In the abstract section, it is mentioned that the technique performs better in certain situations but unable to find related information in the manuscript

14. Page 598, line 29-31: is not clear

15. Page 599, line 16-17, what are these robust error models and how these models address the issue, requires more information

16. Page 600, line 12-17: suggested to provide more information related to "evolved stochastically according to local level or generalized random walk"

17. Page 600, line 15-17: Why generalized random walk is selected?

18. Page 601, line 3-7: define symbols, and how one estimates noise variance?

19. Page 602, line 23-24: Is it OK to assume that xt has similar properties as error? What is Pt?

20. Page 603, it is not clear how equations (17) through (19) are obtained? How the information from the new observed is utilized?

21. Page 604, how this initial variance is calculated?

22. Page 608, reference for ISIS hydrodynamic model

23. Page 609, line 26: are samples collected at every 0.25 hours, how 28 samples approximate 7hr?

24. Page 610: lines 1-6, Clearly SEFE methodology violated assumptions but it is mentioned that the forecasts are good. This highlights the need to discuss the forecast aspects for which forecasts are not good due to the violation of the assumption.

25. It is not clear why two calibration methodologies are selected.

26. Page 611, unless verification metrics conditioned on rising limb are calculated, it is hard to make comments from the figure. The figure is useful with respect to the qualitative interpretation.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 595, 2012.