Hydrol. Earth Syst. Sci. Discuss., 7, C4985–C4990, 2011
www.hydrol-earth-syst-sci-discuss.net/7/C4985/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Recent developments in predictive uncertainty assessment based on the model conditional processor approach" *by* G. Coccia and E. Todini

Anonymous Referee #1

Received and published: 4 February 2011

1 General Summery

This paper is generally well written and could be an informative contribution to the hydrological literature; it addresses an operationally important problem. The title and abstract reflect the content of the paper. While the techniques used are generally well outlined (though some of the mathematical notation could be made clearer) some aspects of the paper could be substantially improved. Comments on these areas are in the following section.

2 Main Comments

1. The authors are correct to point out in the introduction (Pg 9221 Line 27) that is the exceedance of the actual not predicted water level that results in damage. However in all the techniques used they appear to substitute the observed water level (or worse discharge) for the actual value. Much as the model value isn't the actual value neither is the observed value. This need to be noted, comparison could be drawn to other Bayesian analysis which (at least theoretically if not in practise) allow the actual value to be predicted (;). Prediction of the observed value is still a useful exercise but may require a different interpretation in a risk context.

2. The probabilistic threshold paradigm is an natural consequence of producing probabilistic forecasts. Figures 1 & 2 are informative of the challenges in interpreting probabilistic forecasts when issuing warnings and these are worth keeping to help illustrate the latter results. However, since no attempt is made to define a method for giving warnings based on probabilistic forecasts, I believe this section could be substantially shortened. Section 1.2 is also poorly referenced, even mentioning the two texts on probabilistic decision making referenced on Page 9221 would be a start!

3. Section 2 reviews a number of Bayesian statistical methodologies for expressing prediction uncertainty. All of the methods reviewed appear to work on the principle of building a probabilistic model of residuals of a given model. In this context the work of Kennedy and O'Hagan () may be of interest. Also this is just one way of assessing predictive uncertainty, alternative methods e.g. (; ;) which are more often used in calibration could be applied to generate Bayesian predictive uncertainty bounds. This should be mentioned and comparisons drawn.

4. Unlike the outlined formulation of the HUP (Section 2.1) the model conditional

HESSD

7, C4985–C4990, 2011

Interactive Comment

processor (Section 2.3) makes no use of auto-correlation in the residuals. It would appear that at times these are significant (Fig 3 & 4) and could be of use in forecasting (depending upon the lead time). The authors should explain why this is not included. Note that the formulation used in the multimodel case (Section 2.4) would appear to allow for inclusion of lagged residuals as alternative 'models' and the later work (Section 2.4.3) for altering properties of the temporal pattern of residual dependency with output.

5. Much of Section 2.4 outlines the new work in this paper, the use of truncated normal distributions. As such putting it a section titled 'existing approaches' is confusing. I suggest that work should be in a separate section.

6. At the start of Section 2.4 the authors highlight some of the difficulties encountered due to transforming the data before describing the joint distribution. In the processors outlined earlier in the work the transformation had clear advantages, i.e. the marginal distributions were Gaussian and a multivariate Gaussian distribution was used. What is the purpose of maintaining the transformation with the truncated normal joint distribution?

7. Section 2.4.3 should be written to improve the clarity of the exposition. In particular

   - It is often unclear which of the two truncated normal distributions was being discussed;
   - The estimation of the parameters of the truncated normal distributions is reported in the appendix, yet this is entirely unmentioned in the text;
   - Some notation is reused for different variables resulting in repetition e.g. Eq 25 and 26;
   - Eq 21 and the purpose of $\hat{\eta}^{*}$ is opaque. If $\tilde{\eta}^{*}$ is a realisation of the random variable $\hat{\eta}$ do you mean $f\left(\eta \,|\hat{\eta} = \hat{\eta}^{*}, \hat{\eta}^{*} > a\right)$?

HESSD

7, C4985–C4990, 2011

Interactive Comment

- No mention is made of how to select a, or how to select the number of truncated normal distributions. These selections are crucial to applying the technique;
- The NQT transform used results in η and $\hat{\eta}$ having Gaussian marginal distributions. Does the truncated multivariate normal representation maintain this property? If not what does this imply about the validity of using this transform? For example what is the predictive if there is no evaluation of think of forecasting without an evaluation $\hat{\eta}$? Also if one is not interested in maintaining this marginal property why not consider alternative representations (e.g. mixtures of normal distributions) and different transforms?

8. Section 2.4.4 should be altered to reflect the changes in Section 2.4.3.

9. The assumption around line 10 on pg 9238 in section 2.4.4 implies a single split and two truncated multivariate normal distributions. But if each model had a single a values there could be 2^M truncated normal distributions. The computational advantages of the assumption are clear, but the loss of accuracy in the representation of the predictive uncertainty compared to considering more complex splits need exploring.

10. The results of example application show that the the PU technique outlined does not appear to fully capture the timing errors on the rising limb (Fig 10 & 11). Please comment as to why this is the case.

11. Section 4 is very difficult to follow, though the explanation of the results in comparison to Figures 1 & 2 is worth reporting. The interpretation of Figure 14 is very unclear, the labelling of the y axis as ‘Observed occurrences’ makes little sense when related to the text and figure label.

12. The conclusions introduce a number of summary results that would be better reported and discussed in Sections 3 and 4. The results on line 6 Pg 9247

HESSD

7, C4985–C4990, 2011

Interactive Comment

indicate some misrepresentation of the predictive uncertainty that may require further discussion.

3 Other Remarks

- I would consider the errors given for the observation of water level (Pg 9223 Line 24) to be optimistic (especially at sites without gauging structures)
- The reference on Pg 9223 Line 23 is inadequate
- More evidence then one conference paper is required to suggest that a paradigm has been radically changed (Pg 9924 Line 11)
- Pg 9224 first paragraph. From personal experience I don't think I have met an operational flood forecaster who believes that a deterministic output from their model is correct. While they do use the model forecasts in issuing warnings to suggest an approach as naive as your text does is perhaps doing many a disservice.
- Pg 9224 Line 16: 'in' not 'into'. This statement need evidence
- Pg 9226 Line 15: To claim the HUP was the first 'correct' formulation of the prediction uncertainty needs further support. Even if excluding other inference methodologies other Bayesian formulations (with differing assumptions) had been given before (see references above)
- Pg 9226 Line 15: Why is the HUP 'hardly extendable' to multiple models?
- Pg 9230 Line 1: Please enlarge on the extensions required to accommodate the tails of the distribution or provide a reference.

HESSD
7, C4985–C4990, 2011

- Pg 9231: As the authors point out on page 9232 this page outlines computing the conditional distribution from a multivariate normal distribution. Since this is well known it could be shortened.
- Pg 9237 Line 9: Missing symbols?
- Section 3 onwards: The standard of the English in these later section is not a high as earlier in the text and requires further proof reading. For example the word 'pick' is repeatedly used instead of 'peak'.
- pg 9243 Line 25: What 'early stopping procedure'?
- Figures 3 & 4 appear quite blurred. It is not clear what Figure 5 adds to the presentation.

References

M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 63:425–450, 2001.

G. Kuczera. Improved Parameter Inference In Catchment Models.1. Evaluating Parameter Uncertainty. *Water Resources Research*, 19(5):1151–1162, 1983.

G. Kuczera, D. Kavetski, S. Franks, and M. Thyer. Towards a bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *Journal of Hydrology*, 331(1-2):161–177, 2006.

B. Renard, D. Kavetski, and G. Kuczera. Comment on "an integrated hydrologic bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction" by newsha k. ajami et al. *Water Resources Research*, 45, 2009.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 9219, 2010.