

## ***Interactive comment on “Hydrological appraisal of operational weather radar rainfall estimates in the context of different modelling structures” by D. Zhu et al.***

**H. McMillan (Referee)**

h.mcmillan@niwa.co.nz

Received and published: 12 November 2013

This paper provides an interesting comparison of the relative importance of three factors – rainfall estimation method, storm type and model structure – on the quality of flow forecasts. The analysis is based on a chalk catchment in south-east England. I believe the paper is a worthwhile addition to the hydrological literature for its contribution to enable practitioners to untangle different sources of error in flow forecasts and their relative magnitudes. I have several comments that the authors should address to improve the description and discussion of the modelling results.

C6230

1. The PRTF event-based model is fundamentally unsuited for a chalk catchment with a strong baseflow influence. This is clear in the performance (Figures 2, 3; Table 1). The authors need at least to justify the inclusion of this model (is it used operationally?) and explain why it performs poorly so the reader is clear on this.

2. Also related to the nature of the chalk catchment, the model warm-up, calibration and validation periods do not appear to be sufficiently long. I would expect a multi-year calibration period to be required to ensure the models are trained to include seasonal baseflow cycle and long-term water balance. This is particularly the case for the MIKESHE and PDM models which the authors say were started from a dry state. I also question whether the lack of sufficient wet-up period is responsible for the poor performance of these two models in the convective storm event (Figure 11). This would explain the large difference in the flow volumes produced by these models vs the PRTF model.

3. The authors make sweeping generalisations in the conclusions section which are not justified based on the analysis in this paper (Section 6, last bullet; Section 7, bullets 2-4). This section requires a more thoughtful and nuanced discussion of the findings in this study, how they are affected by the choice of catchment and model structures, and to what extent they may be generalisable to other situations.

4. The quality of the English needs to be improved, especially in Sections 6 and 7.

Minor points

p 10500 line 7. Units of  $\text{min km}^{-1}$  are unusual

p 10502 line 25. Expand on what kind of ‘trial-and-error’ method was used to calibrate MIKESHE. Also, why were different methods used to calibrate the different models, and may this affect the conclusions?

p 10503 why use both the RMSE and Nash Sutcliffe when they are linear transforms of each other?

C6231

p 10504 line 3. Zhu (2009) not in reference list.

Section 4.2. This section is long-winded; I suggest it could be shortened.

Section 5. Could the authors comment on whether the consistent differences between raingauge and radar forecast distributions implies that hydrological models intended to be driven with radar data should be calibrated on radar data?

Figure 5 and 6. These figures seem to be showing identical data. I suggest only one of them is chosen for the paper.

Figure 7. It would be clearer here to use  $1 - FAR$  so that the perfect score is 1 for all three indicators.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 10495, 2013.