Thank you for taking the time to review this paper. Your review approaches the paper as if it is a collection of separate technical notes: one for a model, another for a method, and a third for a set of concepts and definitions. You then robustly criticize each note on the basis that its contents is not fit for purpose in general RR modelling (along the lines that the model does not have advantages relative to existing models, the method is poorer than existing methods, and the definitions and concepts are redundant).

The reality of the paper is quite different. It is about an experiment which is a scientific exploration of the very difficult problem of properly linking performance to hydrologic knowledge. Specifically, it is a report on the huge effort expended by a pair of researchers to develop numerous arguments and definitions, reframe some historical arguments, develop a full blown method and a completely new RR model, and thus make it possible to create a simple concrete example for the link between performance and hydrologic knowledge (that concrete example is a benchmark for the link and gives a basis for generating hypotheses). Somewhere, therefore, there must be a problem with communication or expectation. Note, for example, the paper does not even hint that the model is proposed for general use in the way you assume (and are very robust about) in your review. Perhaps that is a problem of expectation. If there are problems of communication, then surely they are quite easy to resolve.

We note that in the future there may be a more detailed account of your concerns. In the meantime, here are responses to three points you have made so far.

1. You commented that "The new model performs poorly against the benchmark". A benchmark is not a target; it is a reference point against which things can be measured. The performance benchmark used shows what can be achieved when employing the substantial advantages of allowing calibration and using evaporation data. KERR is simple, parameterless, and does not use evaporation data, so if its performance consistently matched or exceeded the benchmark the resulting shock would have marked the start of a revolution in RR modelling.

2. You commented that "In general, the model itself is not something that I can see any hydrologist using or being interested in for any practical or intellectual reason". Let us look
closely at KERR: (i) it is probably the best performing simple parameterless model available for temperate catchments; (ii) it gives information on similarity across time and place; (iii) it can use proxy catchment data which is not scaled or modified in any way; (iv) it is linked directly and comprehensively to a set of hydrologic knowledge; and (v) it preserves that knowledge and allows it to be quantified. It is easy to imagine a research hydrologist spending a decade or more working with KERR and its descendants.

3. Our response to your claim that the concept of the Modelled Hydrologist (MH) is redundant is as follows. A perceptual or conceptual model does not exist in a vacuum, it entails understanding and actions by human beings, and places restrictions on what that understanding and actions can or must be. There is, therefore, an entity (the MH) which is larger than the perceptual or conceptual model. This raises the question of what that larger entity is and how it relates to the search for hydrologic laws at the catchment scale. There is, though, another side to this. One of the points made in the blind validation work is that models and modellers must be seen as a package (Ewen and Parkin, 1996). Our experience is that hydrologists running an RR model sometimes forget the nature of the model. Sometimes it is treated as a statistical black box. The worst case is when the model is treated as if it is reality, and it is implicitly assumed that there are no constraints on what can be concluded from the resulting simulations. Putting all fine scientific arguments to one side, it is quite legitimate in an international research journal to encourage thinking, or a change in thinking, about the nature of RR modelling and RR models.

There are obvious dangers in creating a lot of things at one time. It gives multiple targets for criticism (fair and unfair criticism), and runs the risk that the baby will be thrown out with the bathwater.

John and Greg