Review comments – "Understanding and seasonal forecasting of hydrological drought in the Anthropocene" by Yuan et al

This paper analyzes precipitation and streamflow in the Yellow River basin using the standardized precipitation and streamflow indices (basically just Z-values, where the percentiles come from a fitted gamma distribution). They analyze both naturalized flows (it's not entirely clear where they came from, but probably an attempt by someone to back out the effects of irrigation and reservoir storage) and observed flows. As is well known, irrigation development in the basin has increased rapidly in recent decades, and during the irrigation season flows have been greatly diminished in some reaches. This shows up in their Figure 4, where for basins like Lijin in particular, the observed flows have SSI values that go as low as -9, which would be an absurdly small value for the naturalized flows (for which something in excess of 99% of the values should be between +/- 3).

The problem I have with this paper is the motivation and general approach. Basically, they are analyzing the managed flows as if the management effects were some kind of natural phenomenon. But in fact, the river flows post-management are the result of a set of decisions – either planned, or ad hoc. Those management decisions should be predictable – the irrigated area can be estimated, as can the crop water demands, hence the diversions and return flows. There's a whole branch of water resources systems analysis that does just that. So why treat those decisions as a black box, and do a statistical analysis at all? What they have found essentially is that if you look at the season in which the irrigation diversions are made, the river flows go down. Some of that water gets back into the river, later, and perhaps at a different place, and that effect may increase the flows relative to natural (say, somewhere downstream). But all of that should be predictable at some level.

Also, another concern I have – perhaps not so much with the Yellow River, but with the basins referenced in the Zhang et al, (2014) and Wen et al. (2011) studies that they cite as motivation. The Zhang et al. publication is somewhat obscure, and I could only get the abstract. I did read Wen et al., which is a study of a basin in Australia. In Table 1 of Wen, they give the various water management perturbations to the basin, which include construction of what appears to be a couple of km^3 of reservoir storage. There must be an operating policy for those reservoirs, and it must be based on an objective function, presumably having something to do with meeting the irrigation demand. Whether or not the policy deals with instream flows at all isn't clear in the paper. My point is that if you look at the statistics of the reliability of the reservoir system in meeting the irrigation demands, presumably it's higher than without the reservoirs – after all, that's the reason for building reservoirs. But if the operating policy doesn't consider the instream flows, of course eventually enough irrigation will be added to dry up the river. But we don't need a statistical study to tell us that. My concern is that in all of these papers that look at instream drought statistics (including the authors'), that's completely ignored. The situation is slightly different in the Yellow, as I think (I could be wrong) that there isn't currently a lot of

storage, so the diversions for irrigation are mostly run of the river. But as I implied above, that could be modeled as well.

A final concern I have about the paper is that the ensemble prediction doesn't seem to fit. What was the purpose of including it? Is it to show that if some change in operation was made based on the forecasts, the hydrologic drought statistics would improve? I don't see any argument to that effect. So to me, that part of the paper seems not to fit. I do question the results they show in Figure 7, take for instance for Lijin, which is the sub-basin most affected by diversions for irrigation per their Figure 4. That basin has all kinds of SSI values in the -3 to -9 range, but none of their ensemble members are anywhere close to those values – their smallest forecasts are in the -3 range. The reason of course is that they're using VIC, which (I assume in the version they used) doesn't deal with water management. So they must be forecasting naturalized flows. But who would care about a forecast of naturalized flows? What a management agency needs is a forecast of how much water will be in the river. So we're back to the same thing – to make this paper meaningful, they need a water management model.

I think the authors need to go back to the drawing board with the entire concept, and take a physically based, rather than statistical approach. As it is currently written, I don't find that the paper provides the reader with many insights into causality, which they could do.