Review of “The Analogue Method for Precipitation Prediction: Finding Better Analogue Situations at a Sub-Daily Time Step.” (HESS-2016-246) by Horton et al. 2016.  
  
This is my first time reviewing this manuscript, though it is my understanding that it was previously reviewed by others, with major revisions requested. I have not studied the comments of other reviewers, with the desire to provide guidance as independent as possible.  
  
My recommendation, with regret, is rejection. The article is needs significant reorganization and rethought, above what is required for a major revision. Some of my most significant issues include: (a) it provided detail of a minor improvement to a rather antiquated post-processing methodology; (b) it uses data sets for which there are better alternatives; (c) it doesn’t describe all of the procedures clearly; (d) it doesn’t consider other alternatives as controls against which to evaluate the methodology. In sum, I think the authors need to re-evaluate their research from top to bottom.  
  
Here are some more details on my most significant concerns.  
  
Poor organization. Journal articles are generally minor variations on a standard organization, with an introduction, data, methods, results, conclusion. Here are some of the issues with such sections:  
  
Introduction. There is now a rather rich body of literature on the statistical post-processing to produce probabilistic precipitation forecasts. These include Gamma-distribution fitting methods, Bayesian Model Averaging, Extended Logistic Regression, and more. With the use of an older analog approach, one wonders why such an approach is considered, in the first instance, and how one should place in context the results to follow. Without a thorough review of other possible alternatives and some explanation of why your approach is being considered despite these others, the reader is left wondering why they should bother with continuing to read the rest of the manuscript.  
  
Data. What information was being used for the forecast was not made clear, and this is crucial information to find out right away. See (2a) below for more.  
  
Methods. Section 2.2 and 2.6 seem to be describing two different methods. There should be one, single, clear, unambiguous description of the methodology to be used. Also in this section, while verification using CRPSS is relatively standard, so are other verification diagnostics like reliability diagrams, which are not presented.  
  
Choice of data sets.   
  
Forecast data. What is used as the forecast, and why, is not clear. Is another analysis from the NCEP-NCAR reanalysis time series used as a surrogate for a numerical forecast? Is ECMWF, or other model forecast data used? I could not tell. If NCEP-NCAR reanalysis data is used, this then begs the question: why? This would not be a practical forecast methodology, where one needs information in advance of the event. If a numerical weather prediction forecast is used, then there are potential issues of forecast bias; the forecast model and the analysis may be different in character, leading to the issue of whether perfect-prog (your analog) type approaches are suitable or whether more model-output statistics approaches are needed.  
  
Reanalysis data. There are more modern, more accurate reanalysis data sets available now such as ERA-20C, available at higher temporal resolution (3 hourly), which seems to be crucial for an article examining the usefulness of temporal shifts of the data. You dismiss this in your section 4.6 with an older reference, but I think given the focus in this article on temporal shifts, you need to reconsider higher temporal resolution reanalysis data.  
  
Observation data. While the geographic details of the observation locations are definitely different from many other locations in Europe, there are still other locations in the mountains. Why not consider approaches that supplement the training data with other locations’ observations, potentially allowing you to get more without needing a very lengthy reanalysis? You may have objections to this, but at the least it would be worth explaining your choices.  
  
Methodology. Finding some other reference methodology, e.g., Bayesian Model Averaging, would certainly be desirable. Another logical control would be unadjusted ensemble forecasts from the ECMWF ensemble prediction system. This would allow you to have a point of comparison against which to judge the analog methodology. But even considering the analog methodology in isolation, one wonders why you chose the particular approach to selecting analogs. In particular, I found myself wondering why, for example, you didn’t use canonical correlation analysis approaches to determine what information in the reanalysis data set was most directly related to precipitation variability. Analog approaches, in my experience, are not very “efficient” with their training data; either a previous day’s data is selected as an analog date, or not. This, then, drops on the floor all the other data, which may yet have some useful information. To use an analog approach, then, it seems especially incumbent to have demonstrated that you have chosen the most important predictors that will be used in selecting past dates.  
  
In summary, I regret recommending rejection. I hope the authors will constructively use this feedback in the spirit intended.