Hydrol. Earth Syst. Sci. Discuss., 7, C3232–C3235, 2010
www.hydrol-earth-syst-sci-discuss.net/7/C3232/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Series distance – an intuitive metric for hydrograph comparison" *by* U. Ehret and E. Zehe

Anonymous Referee #1

Received and published: 28 October 2010

The submitted paper "Series Distance – An intuitive metric for hydrograph comparison" introduces a new approach for comparing hydrographs. It is well-written and relatively easy to follow. Although I am not a hydrologist, I think that the paper offers an interesting discussion on an important problem in hydrology. This is definitely the strongest point of the paper. Yet, I found a couple of serious shortcomings in the paper, namely the following ones:

1) The authors focus on the comparison of time series in a pure hydrological context. I missed in this paper an overall discussion of connections with related fields. A lot of distance measures for time series have been proposed in fields like statistics, dynamical systems, signal processing, physics, etc. Particularly, in time series forecasting

in general, one might find a lot of relevant papers. By a quick search on google for "similarity" or "distance" of time series, I found the following papers, which I consider as relevant:

a. Richard Moeckel and Brad Murray Measuring the distance between time series Physica D: Nonlinear Phenomena Volume 102, Issues 3-4, 1 April 1997, Pages 187-194 b. Chouakria-Douzal, Ahlame and Nagabhushan, Panduranga, Improved Fréchet Distance for Time Series, Data Science and Classification, 2006. c. Basic introduction to time warping distances for time series on Wikipedia: http://en.wikipedia.org/wiki/Dynamic_time_warping

So, in general, the paper definitely needs an overview of what already exists in related domains, and a discussion about the aspects that make the distance measure of the authors novel. I really cannot imagine that no similar approaches have been taken in related domains. Contacting an expert in one of these domains might definitely help in further improving the quality of this paper.

2) As a second remark, I missed in the introduction a clear definition of the exact goal for presenting a new distance measure. The authors talk about "comparing two hydrographs", which is to my opinion a very general goal. The most appropriate measure of distance between two objects depends on what type of behavior you want to quantify. From that perspective, I believe that such a thing like the "ideal distance measure for comparing hydrogrpahs" does not exist. Further on in the paper, the authors are focusing mainly on comparing hydrographs for forecasting reasons, by comparing the hydrograph of a fitted model with the hydrograph of true observations. To my opinion, it would be less confusing to state this goal immediately at the beginning of the text. Then, the discussion on presenting a new distance measure for hydrographs basically boils down to a discussion on choosing an appropriate performance measure for hydrological systems. I agree with the authors that better alternatives to RMSE should be investigated for evaluating hydrological models, but I am pretty sure that many alternatives exist already, if one takes a look at time series prediction in general, as indicated

in my first comment.

3) Concerning evaluating the output of predictive models, the authors seem to mix up a couple of things in the paper. For example, the authors are claiming in line 28-30 that RMSE should be avoided as performance measure because it consists of a weighted three-criteria objective function. This argument is used in a totally incorrect context. Statistically speaking, any error measure (loss function) of any predictive data-driven model can be decomposed into three parts. The expected prediction error of a model consists of: (a) the irreducible error (as a result of noise in the data); (b) the squared bias (as a result of choosing a too simplistic model); and (c) the variance (as a result of choosing a too complex model). The first component cannot be optimized because a model cannot predict random noise in data. The other two parts have to be optimized simultaneously, because the core of predictive modeling is nothing more than choosing the optimal trade-off between bias and variance, namely the optimal trade-off between simplicity and complexity. So, it is incorrect to state that one ends up with a weighted three-criteria optimization problem. More importantly, it is also incorrect to state that this is due to the nature of the RMSE. The bias-variance trade-off is a phenomenon that can be observed for any performance measure, thus also for the one presented by the authors. I refer to reference books like "T. Hastie, R. Tibshirani and J. Friedman, The elements of statistical learning, data mining, inference and prediction", chapter 2, for a more detailed discussion on this topic.

4) It is maybe due to the fact that my main background is not in hydrology, but I found it really strange to read a paper about distance measures or performance measures without seeing one mathematical formula. To my opinion, measures like RMSE, NSE or MPTE should be more formally defined, in terms of a simple mathematical formula. As a result of the clear writing style of the authors, I could understand the different measures without confusion, but a short comprehensive mathematical description will only improve the readability of the paper.

5) The presented procedure makes sense from a methodological perspective. It re-

minds of similar ideas that have been put forward in spectral data analysis. Of course, spectral data is in nature quite different from time series data, but similar requirements exist for good similarity measures in comparing spectra. More specifically, it also happens in spectral data that one has to account for shifts on the y-axis (amplitude) and the x-axis simultaneously. As such, one typically uses peak alignment algorithms to take shifts on the x-axis into account, by looking at local segments. Subsequently, distance measures are computed locally on the aligned peak segments and averaged over different segments. It might be interesting for the authors to take a look at such methods.

6) One shortcoming of the presentation of the procedure of the authors is that the description is quite informal. I think that it becomes almost impossible to reproduce the procedure of the authors, based on the description given. Would it be possible to give a precise formal description of the algorithms by means of pseudo-code and a few mathematical definitions?

I think that the topic of this paper is quite interesting, so I do not recommend to reject it. But it is clear that the article definitely needs a drastic revision before it can be published.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 8387, 2010.