Hydrology and Earth System Sciences, 3(2), *305–307* (1999) © EGS

Hydrology & Earth System Sciences

RESEARCH NOTE

Hydrograph Prediction—How much skill?

Max Beran

OB Research Services, The Croft, East Hagbourne, Didcot, Oxon OX11 9LS
e-mail for corresponding author: max@oldboot.demon.co.uk

Abstract

The matching of estimated to observed hydrograph shape is central to much hydrological analysis. This research note quantifies built-in biases that tend to inflate goodness of fit indicies, biases that arise from the similarity of geometry between observed and estimated hydrographs.

Introduction

How much genuine skill is displayed by hydrograph prediction techniques? This paper shows that the answer is, 'a lot less than is commonly supposed', a conclusion which is based on substituting a more realistic 'null hypothesis' for the ones used currently in testing the significance of hydrograph predictions.

A typical instance occurs with flood analysis by means of the unit hydrograph. The accuracy of the temporal aspects of the procedure is judged by applying the unit hydrograph to observed flood events, essentially by convoluting it with an observed effective precipitation hyetograph and comparing the predicted with the recorded immediate response hydrograph. Objective functions have been proposed for hydrograph simulation models based on explained variance and on likelihood (for example Nash, J.E. and Sutcliffe, 1970; Binley *et al.*, 1991; Jakeman, A.J. and Hornberger, G.M., 1993; Gupta, H.V. *et al.*, 1998). However, this past work has been directed mainly to parameter optimisation and model identification. The question of how much skill is represented by a given value of the objective function is not addressed directly.

To answer the question in the title, it is necessary to set a baseline or null-hypothesis with which a test statistic can be compared. An inappropriate null-hypothesis has led in the past to a tendency to exaggerate the significance of the hydrograph fit and hence to exaggerate the real level of modelling skill. Though expressed here in the context of unit hydrograph analysis, this research note's conclusions apply to other hydrological input:output models in which a transfer function is evaluated by measuring how closely it is able to replicate the shape (as opposed to the volume) of a recorded hydrograph.

Evaluating goodness of fit

Two measures of goodness of fit are considered—the Nash and Sutcliffe (1970) Coefficient of Efficiency, E and the product moment correlation coefficient, R. Both E and R are measures of explained variance and may be regarded as quite severe tests inasmuch as they operate on corresponding hydrograph ordinates and so make no allowance for small temporal shifts. Values of E or R in the region of 0.8 or 0.9 are often regarded as significant, meaning that the transfer function model is safe to apply.

$$E = 1 - \left\{ \Sigma(Q_{rec} - Q_{est})^2 \;/\; \Sigma(Q_{rec} - Q_{mean})^2 \right\} \qquad (1)$$

where Q_{rec}, Q_{est} and Q_{mean} are the recorded, estimated and mean values of discharge and the summation is taken over all ordinates within the hydrograph range.

While at first sight such values do suggest a high level of skill, in the case of the coefficient E, this measure is not normalized so there is no simple relationship with R or R^2, the conventional statistical measure of explained variance. Thus $E = 0$ can occur simultaneously with high positive correlation between Q_{rec} and Q_{est}, see Fig. 1. Also, as can be seen by inspection of Eqn. 1, large negative values can build up. This may occur even with moderate positive correlation between the hydrograph ordinates when the estimated values depart from the recorded value more than the latter do from their own mean value. There is also an element of arbitrariness in the calculation as both measures can be inflated by the inclusion of lengths of lead-in and hydrograph tail within the summation range.

These difficulties are ones of implementation. The more important matter, with which this paper is primarily concerned, is of a conceptual nature, and lies with the