An approach to creating lumped-parameter rainfall–runoff models for drainage basins experiencing environmental change

EARL BARDSELY
Department of Earth Sciences, University of Waikato, Private Bag 3105, Hamilton, New Zealand
e.bardsley@waikato.ac.nz

SUXIA LIU
Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Datun Road, Building 917, An Wai, Beijing 100101, China

Abstract Many drainage basins experience hydrological change due to land-use modification and the consequential recalibration of conceptual rainfall–runoff models may be cumbersome. We describe a model-constructing approach which leads to straightforward calibration/recalibration using a hybrid of manual and automated methods. Hydrographs are simulated using weighted finite-mixture distributions where the “distributions” are just lists of hydrograph-like constants for general application. This creates many-parameter models which enable the use of standard, stable, least-squares algorithms subject to a non-negativity constraint. An illustrative model application is given for drainage basins in New Zealand and China which have experienced land-use change. The example model employs manual calibration of hydrograph lag and length, with automated calibration of 85 hidden parameters. In addition to land-use change applications, the straightforward nature of the automated calibration process should make models of this type amenable to the auto-updating requirements of flood forecasting models.

Key words China; estimation; model; New Zealand; parameter; runoff

INTRODUCTION

Conceptual rainfall–runoff models are usually formulated without any particular consideration of subsequent calibration to data (Houghton-Carr, 1999). However, it has been recognized for some time that constructing the models is a more difficult task than the process of model calibration (Gupta & Sorooshian, 1985). The non-trivial nature of obtaining best calibration parameter estimates from data arises from the complex topology of the parameter response surface, which is influenced by interaction of the objective function and model structure (Duan et al., 1992; Xiong & O’Connor, 2000). The calibration problem is likely to be exacerbated if frequent parameter updating is required in response to drainage basin land-use change, or if the model is to be used in real-time flood forecasting. Even in the absence of land-use change there may be different optimal parameter values in wet and dry years (Sorooshian et al., 1983).

The purpose of this paper is to outline an alternative approach to formulating and calibrating conceptual rainfall–runoff models. Rather than first creating models and...
then seeking a means to calibrate them, we advocate model creation within recognized mathematical frameworks for which there are already established parameter estimation algorithms designed to exploit the model structure.

PARAMETERS AND PARSIMONY

An increased number of parameters is likely to be required to maintain a hydrological flavour within a simpler mathematical framework. The issue of “over-parameterization” is therefore discussed briefly here, prior to outlining specifics of model construction and calibration.

Conceptual rainfall–runoff models are conceptualizations of the processes assumed to give rise to observed hydrographs. Concepts do not always equate to reality and an over-parameterized model will have mathematical flexibility to give good data calibration, but without capturing hydrological information. At the same time, however, it is questionable whether models with smaller numbers of parameters are any better for revealing information about the hydrological processes they supposedly represent. This concern arises from the observation that conceptual models often exhibit equifinality, whereby plausible but very different conceptual models fit the same data equally well (Beven, 2001).

Rainfall–runoff models might therefore be thought of as providing a mimic of the overall hydrograph-generation process rather than producing individual parameter values which yield details of specific processes (Perrin et al., 2001). If few or many parameter values contain equally little information, then it can be argued that it is of no scientific concern if conceptual rainfall–runoff models have few or many parameters.

This goes against the current thinking of conceptual model development, which looks toward less complex models where “complexity” is equated to the number of unknown parameters (Jakeman & Hornberger 1993; Perrin et al., 2001; Ye et al., 1997). Occam’s principle would certainly suggest that models should be no more complex than required to describe the data. However, model structure would seem at least as important as parameter numbers in any measure of complexity. That is, a conceptual model with a large number of parameters could, by some definition, be “simpler” than a more complexly-structured model with fewer parameters. Conceptual rainfall–runoff models are therefore as much at risk of being “overstructured” as “overparameterized”.

At a more pragmatic level, large numbers of parameters may create a risk of over-fitting whereby a good calibration masks the reality that the model has not captured sufficient hydrological information to have useful predictive value. However, such over-fitting is readily detected in practice by the dual effect of good fits to calibration data and poor fits to validation data (Perrin et al., 2001).

It might also be argued that large numbers of parameters make model output comparisons difficult. However, a reasonable representation of the hydrological information from any calibrated rainfall–runoff model can be achieved by the graphical device of plotting model hydrographs generated from a number of specified rainfall sequences. These predicted hydrographs might then be compared, for example, with the corresponding estimated hydrographs produced from the same rainfall...
An approach to creating lumped-parameter rainfall–runoff models

sequences when the model is calibrated to a different land use. The estimation errors of the individual model parameters are still important, but only to the extent of their total contribution to the error of the estimated hydrograph. It is the magnitude of this final estimation error which is of relevance, rather than the number of parameters contributing to the error.

Finally, concern might be raised that large numbers of parameters may result in difficult calibrations. The alternative argument here is that the calibration process is actually likely to be more reliable if the model has been constructed within a specific mathematical structure allowing use of calibration algorithms optimized to that structure.

THE CONCEPTUAL MODEL

A largely scale-independent runoff generation process is the spatial convergence of drainage basin water from a precipitation event to an observation point where the hydrograph is recorded. A simple hydrograph with limited multimodality is anticipated from a single event, with perhaps more subdued and delayed hydrograph forms associated with larger drainage basins. Of course, very different dominant water transport processes will operate at different scales (Blöschl, 2001), but this will only have a modifying effect on hydrograph forms.

A conceptual rainfall–runoff model for application over a wide range of scales is therefore proposed on the basis of simulating a simple hydrograph response from individual rainfall events, with the total hydrograph estimate being baseflow plus the addition of all individual event hydrographs. The event hydrographs are taken to have a common base length, with the form and magnitude of each event hydrograph being influenced by the rainfall amount and antecedent drainage basin wetness, as quantified by the sequence of previous rainfall events. This approach is in the category of unit-hydrograph models incorporating slow flows.

The novel feature of the approach advocated here is that the event hydrographs are specified as rescaled weighted finite-mixture probability distributions where the component distributions are selected so as to create “hydrograph-like” forms from any combination of weights. A single peak and recession curve is possible, for example, but two separated sharp peaks are not. This is somewhat analogous to a wavelet hydrograph model (Parada et al., 2002) although arguably more “hydrological” in approach. The finite mixture component distributions are simply lists of constants rather than mathematical expressions. This allows the event hydrographs to be specified as linear functions with a large number of parameters, all subject to a non-negativity constraint. These parameters in effect replace a much smaller number of parameters in an equivalent nonlinear model. The advantage with the many-parameter approach is that standard algorithms can be utilized for automated calibration.

A hybrid manual–automated calibration procedure is utilized in a manner somewhat similar to that of Boyle et al. (2000). In a given iteration, hydrograph lag and length are first specified by the user. The many “hidden” parameters that define the forms of the event hydrographs are estimated automatically by a standard algorithm.
A full description of the model structure is given by Bardsley (2003). No explicit allowance is made for seasonal evaporation effects in the model structure although the model could be optimized to a particular season by applying a seasonal weighting to the objective function.

EXAMPLE APPLICATIONS

An example model was formulated using weighted mixtures of seven component hydrographs, some of which are illustrated in Fig. 1. The discrete nature of these hydrographs is masked by the graph scale. In all, a total of 85 hidden parameters are estimated during automated calibrations. A MATLAB graphical interface provided a convenient framework for both automated fits and manual adjustment of hydrograph lag and length. The automated component used the linearly-constrained least squares procedure “lsqlin” in the MATLAB Optimization Toolbox, utilizing a specific large-scale optimization algorithm. This was found to give consistently stable convergence in keeping with the design of the model structure toward that end. There is still a possibility of convergence to local optima but the hope here is that the simple structure of positive linear parameters gives rise to fitted parameters not too far removed from their global values.

Fig. 1 Examples of shapes of the seven component distributions in the finite-mixture model: (a) and (b) show end-point distributions 1 and 7 respectively; (c) is an equal weighting of distributions 1 and 7, corresponding to an extreme separation of quickflow and slowflow components; (d) gives weighted mixtures incorporating other distributions with (i) having greater weight for distribution 1, and (ii) greater weight for distribution 7.
The model was applied to 30-day discharge averages of the Luo River in Eastern China, which drains to the Yellow River and occupies a drainage basin of 4423 km². The effect of extensive forest planting appears to have modified the flow regime of the river with a tendency to produce lower peak flows since about 1988. The model was calibrated independently to both the pre- and post-1988 data, with the hydrograph lag and base length being adjusted to zero and 180 days respectively, for both data sets. The observed and model 30-day hydrographs are shown in Fig. 2 for pre-1988 calibration with predictions of the later flow data. The coarseness of the averaging process allows for only an approximate calibration and no verification exercise was carried out. However, the post-1988 forecast hydrograph does give some impression of peak discharges being under-predicted by the pre-1988 calibration. Model comparisons by way of specified rainfall sequences are illustrated in Fig. 3, where the difference between the predicted flows of the pre- and post 1988 calibrations is clearly evident.

The second application of the model was to daily discharge data from the Tarawera River, which occupies a 984 km² drainage basin in the North Island of New Zealand. Similar to the Luo River, the Tarawera River has experienced discharge modification since about 1981 as a consequence of forest planting—in this case pine plantations associated with production forestry.

The model was calibrated for the 12-year period 1966–1968, 1972–1980. The 1970 discharges were used as the validation data set. The years 1969 and 1971 were not included for calibration to ensure unbiased validation. The manually-calibrated hydrograph lag and length were adjusted to zero and 150 days respectively. The overall calibration fit to the daily discharges was not particularly good, with a Nash-Sutcliffe efficiency of 54%. The hydrograph base length was a rather insensitive parameter and there was little difference in fit with an alternative base length of 100 days (Nash-Sutcliffe efficiency of 53%). Interestingly, the validation data shown in Fig. 4 gave a higher degree of fit than the calibration set with an efficiency value of
Fig. 3 Predicted discharges for the Luo River, based on pre- and post-1988 calibrations, for (a) greater and (b) lesser 30-day rainfall totals.

Fig. 4 Observed and predicted discharges for the 1970 validation data of the Tarawera River at Awakaponga, New Zealand. Discharges are daily means.

72%. However, the validation peak discharges were not very well matched. The degree of fit nonetheless gives encouragement that many-parameter models of this type are able to capture hydrological information. Figure 5 indicates that the model has detected the forestry impact on flow reduction because the model calibrated to the pre-forestry state predicts considerably higher river discharges for 1988. The sharp form of the hydrograph rises were not fully captured by the model, suggesting the possibility of adding an eighth component distribution to allow for more rapid rises.
An approach to creating lumped-parameter rainfall–runoff models

CONCLUSION

The seven-hydrograph model may find some further application. However, the finite-mixture approach is still a method rather than a final model. Further investigation is needed on the optimal number and form of the component hydrographs to achieve best application over a range of drainage basin scales. Extensions of the model applicability might include flood forecasting where frequent recalibrations are required to reflect the current drainage basin state. This might be achieved by fixing the nonlinear parameters just once, by manual calibration, and then having automated real-time updates of the hidden parameters.

Acknowledgements Thanks go to Professor Weimin Zhao and the Yellow River Commission of Water Resources, Zhengzhou, China, for providing the Luo River data. This work was supported by the Chinese National Natural Science foundation (Project nos 49771019, 49890330), and a grant by the Asia 2000 Foundation of New Zealand.

REFERENCES


