Journal of Hydrology 386 (2010) 198–209

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Calibration of the semi-distributed PDM rainfall–runoff model in the Upper Lee catchment, UK

I.G. Pechlivanidis *, N.R. McIntyre, H.S. Wheater

Department of Civil and Environmental Engineering, Imperial College London, UK

ARTICLE INFO

Article history:
Received 12 August 2009
Received in revised form 10 March 2010
Accepted 13 March 2010

This manuscript was handled by
K. Georgakakos, Editor-in-Chief

Keywords:
Calibration
Regionalisation
Semi-distributed
PDM
Sensitivity analysis

SUMMARY

Distributed hydrological models have the capability to incorporate spatially variable inputs, to represent spatial heterogeneity of catchments and to generate outputs at interior locations. However calibrating a distributed model is challenging so much so that a less powerful lumped model is often preferred. This paper describes a method of calibrating a semi-distributed model based on regionalisation of parameters, maintaining as far as possible their physical meaning, and applying spatial multipliers to optimise performance. A semi-distributed implementation of the Probability Distributed Moisture (PDM) model is employed using hourly data from the Upper Lee catchment, UK. Regression relationships between known catchment descriptors (land cover, soil type, climate and topography) and parameters of the PDM are developed using 10 gauged subcatchments and applied to give prior estimates of the parameters at smaller ungauged subcatchment scales. These prior estimates are adjusted using multipliers which are spatially uniform for each parameter. The analysis reveals that the semi-distributed PDM usually (but not always) performs better than a lumped version at gauged points in the catchment, without the introduction of additional degrees of freedom into the calibration. This result applies even when the rainfall input is assumed uniform over the catchment. However, predictions at ungauged points are poor, and the scope for using for additional sources of information, including using more physics based model components and the outputs of broader scale regionalisation studies, is discussed. Results also highlight the sensitivity of the model parameter estimates to the number of gauged catchments used in the regression analysis and to uncertainty associated with parameter equifinality.

Introduction

Modelling the flow response at ungauged interior points of a catchment is one of the principal challenges in hydrological modelling. Addressing this challenge through semi-distributed rainfall–runoff models has been facilitated by improved availability of spatial data. However, it is well known that the parameters of such models in general need to be calibrated in order to achieve a useful degree of reliability, and this calibration is rarely a straightforward task due to the large number of parameters that need to be estimated (Carpenter and Georgakakos, 2006). In particular, as the number of model parameters increases with the degree of spatial discretisation, distributed models can easily become over-parameterised and subsequently ill-posed with respect to the information content of available input–output data. Thus uncertainty in parameter estimates becomes a problem which should be addressed within any application (Madsen et al., 2002; Orellana et al., 2008).

Suitable calibration strategies for semi-distributed models are not as easily defined as strategies for lumped models. Spatial inconsistencies can arise, for example, optimised parameter values can vary spatially without any physical basis, due to uncertainties associated with parameter interactions or data errors (Koren et al., 2003). This raises doubt about the applicability of the model for making predictions, particularly if parameter values need to be adjusted, for example, to represent land cover change (Hundecha and Bardossy, 2004). Ideally, the variability across the catchment of calibrated parameter values would be consistent with the intended physical significance of the parameters (Young, 2006; Pagano et al., 2009). An important challenge therefore is to evaluate the potential use of spatial data to inform the calibration process.

The objective of this study is to test the extent to which regional relationships developed through lumped modelling, using observed data from a relatively small number of sites within the region of interest, can provide useful information for this spatial disaggregation problem. The practical motivation is to maintain spatial consistency of parameter estimates to inform the estimation of flows at ungauged interior points in a catchment and hence, for example, the effect of local-scale land cover changes.

* Corresponding author.
E-mail address: i.pechlivanidis05@imperial.ac.uk (I.G. Pechlivanidis).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.03.022

The paper is organised as follows. Section 2 gives a review of calibration strategies for distributed rainfall–runoff models. Section 3, the study area and data are introduced. Section 4 describes the rainfall–runoff model and the regionalisation or downscaling methodology. In Section 5, the results are presented, including sensitivity and uncertainty analyses. Finally Section 6, states the conclusions.

Calibration strategies for distributed hydrological models

Until recently, there have been few formalised parameter estimation schemes for distributed models, mainly because of the lack of distributed observations of rainfall and runoff and the time that the calibration process requires (Khakbaz et al., in press). The importance of considering multiple flow gauges during the calibration of a semi-distributed model was highlighted by various authors, for example, Andersen et al. (2001) and Wooldridge et al. (2001), who showed that using information from only the catchment outlet gauge exposed significant shortcomings in parameter estimation for some of the upstream tributaries. The general expectations of distributed models, however, are that they should provide information about responses within gauged catchments. For example, applications may require estimates of flow or soil moisture at interior ungauged locations, and/or may require spatially distinct parameters to be used for the purpose of exploring localised catchment management effects. Therefore, the modelling may require more spatial information about parameters to be included than can be identified using traditional calibration approaches. Some calibration approaches aim to extract additional spatial information from a single flow gauge by explicitly considering the origins of the streamflow signals (e.g., Wagener et al., 2009), by using chemical tracers (e.g., McGuire and McDonnell, 2006), or by using distributed data about state variables (Koren et al., 2008). Where such information exists, such approaches seem promising.

A more widely applied method of maximising the spatial information, to supplement that in the gauged rainfall–runoff data, is using prior knowledge about the distribution of catchment physical properties. There are two main associated methodological questions: how to turn this prior knowledge into estimates of model parameters; and how to then refine the estimates using whatever gauged data exist. The first question has received much attention from hydrological modellers, and it may be said that three types of method are used. Firstly, models may be parameterised such that some of the parameters have direct physical meaning, so that measurements (for example, of slopes and soil depths) may be used directly to estimate the associated parameter values. Secondly, the prior values of model parameters may be linked to catchment physical properties through experience of using the model, but without any well-defined objective basis. Thirdly, and of principal interest here, is parameter regionalisation, where parameter estimates for an ungauged catchment are objectively derived using information about streamflow responses at gauged locations in one or more physically comparable catchments.

The regionalisation method has particular practical appeal for conceptual rainfall–runoff modelling: although physical properties cannot be used directly to estimate the parameters of conceptual rainfall–runoff models, often there are strong associations between the two sets of variables which act as a basis for regionalisation (Merz and Bloschl, 2004; Young, 2006). There are alternative approaches to regionalisation, ranging from direct transfer of parameter sets between groups of similar catchments or hydrological response units (Perrin et al., 2008; Reichl et al., 2009) to the use of sophisticated relationships between parameters and physical catchment descriptors for example, based on neural networks (reviews are included in Vogel (2005) and Bardossy (2007)). The most common approach is the use of regression relationships to explicitly link model parameters to physical catchment descriptors (Kling and Gupta, 2009). This has the practical attraction of being simple, and many published relationships or their outputs are widely used (e.g., Boorman et al., 1995; FEH, 1999). There are various fundamental limitations, such as model structural and data errors (Wagener and Wheater, 2006), lack of power in objective functions (Gupta et al., 1998), spatial deficits (Kling and Gupta, 2009), lack of identifiability of calibrated parameters (Beven, 2001), potentially leading to large uncertainty in parameter estimates.

Regionalisation has commonly been used to estimate the prior spatial distribution of parameter values. For example, Bulygina et al. (2009) used a regression-based regional model to estimate base flow index as a function of soil type, and conditioned each $100 \times 100\ m^2$ element of a distributed rainfall–runoff model on this index; however they noted that additional information would be needed to achieve the sought accuracy. Koren et al. (2003) established regional equations by calibrating the Sacramento model parameters to soil properties from 18 catchments in the Ohio River Basin, finding that the regionalisation could maintain the spatial and physical consistency of parameter estimates but that the regionalised models could not perform as well as the calibrated lumped models. Moreda et al. (2006) used 67 subcatchments of the Susquehanna River Basin with areas ranging from 66 to 3848 km^2 to derive regression relationships between 11 physical catchment properties and 10 model parameters. This information was applied to estimate parameters for distributed $4 \times 4\ km^2$ grid scale models. Results were promising with the distributed model performing slightly better than the lumped model in most cases.

Given prior parameter estimates, if calibration data exist they should be used to further condition the model. For most distributed models the potentially large number of parameters means that attempting to optimise all parameters together, as traditionally done for a simple lumped model, is unlikely to be useful because of the large dimensionality of the optimisation space. Instead, some sequential method may be used, where a relatively small number of parameters are optimised while the others are fixed at prior estimates or previously optimised values. Ajami et al. (2004) followed this approach to calibrate the Sacramento model to flow at the outlet of the 1645 km^2 Illinois River basin. The results showed that although the method could in principle represent the catchment heterogeneity, the spatial consistency of the prior parameter values was lost due to the optimisation, and the improvement in model performance compared to using spatially uniform parameter values did not clearly justify the distributed model. To maintain the prior spatial distribution of model parameters, while improving performance, a common approach is to optimise spatial multipliers which move the absolute or relative value of each parameter up or down uniformly over the catchment (e.g., Bandaragoda et al., 2004; Yatheendradas et al., 2008; Pokhrel and Gupta, 2010). Irrespective of the number of spatial elements in the model, the number of variables to optimise is therefore equal to the number of parameters in an individual spatial element, making the calibration both numerically tractable and procedurally straightforward. However the accuracy of the estimated spatial distribution of prior parameters may become the primary control on model performance (e.g., Michaud and Sorooshian, 1994).

In the UK there has been significant research into regionalisation, in particular the estimation of parameters of conceptual models using regression models. National scale studies using spatially lumped models include Sefton and Howarth (1998), Lee et al. (2005), McIntyre et al. (2005), Young (2006), Wagener et al. (2004) and Lamb et al. (2000). However, there has been much less