
An investigation of site-similarity approaches to generalisation of a rainfall–runoff model

A.L. Kay, D.A. Jones, S.M. Crooks, T.R. Kjeldsen and C.F. Fung

Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

Email for corresponding author: alkay@ceh.ac.uk

Abstract

This paper investigates a new approach to spatial generalisation of rainfall–runoff model parameters — site-similarity with pooling groups — for use in flood frequency estimation at ungauged sites using continuous simulation. The method is developed for the generalisation of a simple conceptual model, the Probability Distributed Model, with four parameters which require specific estimation. The study is based on a relatively large sample of catchments in Great Britain. Various options are investigated within the approach. In the final version, the pooling group comprises the 10 calibrated catchments closest, in catchment property space, to the target site, where the catchment properties used to define the space differ for each parameter of the model. An analysis that, explicitly, takes account of calibration uncertainty as a source of error enables the uncertainty associated with generalised parameter values to be reduced, justifiably. The approach uses calibration uncertainty estimated through jack-knifing and employs a weighting scheme within pooling groups that uses weights which vary both with distance in the catchment property space and with the calibration uncertainty. Models using generalised values from this approach perform relatively well compared with direct calibration. Although performance appears to be better in some areas of the country than others, there are no obvious relationships between catchment properties and performance.

Keywords: flood frequency, continuous simulation, uncertainty, ungauged catchments, generalisation, regionalisation

Introduction

This paper looks at spatial generalisation of the parameters of a continuous simulation rainfall–runoff model for use in flood frequency estimation. Continuous flow simulation capitalises on advances in hydrological modelling, together with computing technology and the increasing availability of good quality fine-time-resolution data. It provides catchment modelling of the whole time series (including peaks, durations and hydrograph shapes), including effects of antecedent wetness conditions, river junctions and, if required, it can incorporate changes in climate drivers. However, even in a developed country such as the UK, many catchments have insufficient data for the calibration of such a model. The transfer of information from data-rich to data-poor catchments via relationships derived between model parameters and catchment properties is, therefore, an important area of research.

Regression analysis has traditionally been profitable in developing tools for estimation of model parameters at

ungauged sites. Many examples of the development of spatial generalisation of rainfall–runoff model parameters through regression analysis are available, in locations such as, in the UK (Sefton and Howarth, 1998; Calver *et al.*, 2001), in mainland Europe (Seibert, 1999; Xu, 1999; Parajka *et al.*, 2005), in the USA (Abdulla and Lettenmaier, 1997), in Japan (Yokoo *et al.*, 2001), in Australia (Post and Jakeman, 1999) and in Africa (Servat and Dezetter, 1993). Such studies have used a variety of models and their results have been variable, with only weak relationships between catchment properties and some parameters of some models. This might mean that such relationships do not exist but the explanation is more likely to lie in the model itself (e.g. parameter interdependence, Kokkonen *et al.*, 2003), in the lack of appropriate independent variables (catchment properties), or in the restrictive assumptions of regression.

This paper assesses an alternative method of spatial generalisation for parameter estimation, namely site-similarity, as a potential way of overcoming the limitations

of regression. Often the term site-similarity is used, simply, to mean the direct transfer of parameters from the gauged site most similar, in some sense, to the target, ungauged site. In a study involving 320 catchments in Austria, Parajka *et al.* (2005) found transferring parameters from a donor catchment (defined by mean catchment elevation, stream network density, lake index and areal proportion of porous aquifers, land-use, soils and geology) generally outperformed regression. Similarly, in Austria, Merz and Blöschl (2004) found that averaging parameters from the nearest catchments up- and down-stream performed better than global or local regression; this is more likely to be due to site-similarity rather than spatial proximity. Kokkonen *et al.* (2003) suggested that the direct transfer of parameters from a donor catchment, rather than the use of regression equations, may be preferable “when there is reason to believe that, in the sense of hydrological behaviour, a gauged catchment resembles the ungauged catchment”.

Site-similarity can be used to define a pooling group — a set of gauged sites most-similar to the target site. In the UK, this form of site-similarity approach is an important part of the statistical method for flood frequency estimation in the Flood Estimation Handbook (FEH) (Institute of Hydrology, 1999, Volume 3) and in Low Flows 2000 (Holmes *et al.*, 2002), where it is used to estimate statistics for the ungauged site. The idea behind this form of site-similarity evolved from regional flood frequency analysis, presented as the region-of-influence approach by Burn (1990), based on work by Acreman and Wiltshire (1987, 1989). Burn and Boorman (1992) tested a number of site-similarity-type approaches for estimating model parameters in the flood-event models developed in the Flood Studies Report (FSR) (NERC, 1975); the best of these methods outperformed the linear regression approach originally suggested in the FSR.

It is the latter form of site-similarity, for pooling group formation, which will be investigated here for estimation of continuous simulation rainfall–runoff model parameters. To the authors’ knowledge, such a site-similarity/pooling group approach has not previously been applied in this context. McIntyre *et al.* (2004) used a site-similarity approach with a continuous simulation model but did not use the pooling group to estimate parameters for the target site. Instead, they applied pooling group parameter sets at the target site directly and used Bayesian averaging on the resulting set of flow series to construct a single time-series for the site. The resulting series generally performed better, in terms of fit to observed flows, than flows simulated with parameters estimated through regression (for six catchments in the UK). In this paper, an attempt is made to define and test the best version of site-similarity for estimation of

rainfall–runoff model parameters, specifically in the context of flood frequency estimation at ungauged sites.

Methods

DATA

This study is based on data from 119 catchments across England, Wales and Scotland; of these, 46 have hourly rainfall and flow data, while the remaining 73 catchments have daily data. The map in Fig. 1 shows the catchment boundaries and locations of the catchment outlets, labelled with the catchment number from the UK National River Flow Archive. Information on the UK flow gauging network can be obtained from the series *Hydrological Data UK: Hydrometric Register and Statistics*, published annually by the Centre for Ecology and Hydrology and the British Geological Survey.

The catchments range in size from about 1 km² to 1200 km², with a mean of around 250 km²; the smallest daily catchment had an area of over 50 km² and the largest hourly catchment was about 450 km² in area. Daily data were incorporated alongside the hourly sites to improve the coverage of the UK (both spatially and in terms of catchment properties), as well as providing longer data sets.

The catchment average rainfall data required by the runoff model were produced from daily rainfall totals using the triangle method (Jones, 1983). For hourly sites, the available hourly raingauge data were quality-checked against daily data; then, any good quality hourly data were used to distribute catchment average daily rainfall through the day. For days where hourly data were missing or of bad quality, average profiles (dependent on location and daily rainfall total) were used to spread the daily catchment rainfall (Pilgrim *et al.*, 1969). Catchment potential evaporation (PE) data were calculated from monthly MORECS data, available on a 40 km × 40 km grid across the UK (Thompson *et al.*, 1982), with monthly values divided equally down to the time-step required for input to the runoff model. For model calibration, hourly/daily flow data were acquired. Catchments with hourly flow and rainfall data generally covered the period 1985–2001; for daily catchments, the data length was generally longer, with the earliest records from 1961.

To provide the means of spatial generalisation, sets of 24 catchment properties were derived for each catchment, covering aspects of topography, soil and geology, lakes and reservoirs, land cover, rainfall and drainage networks. To improve the potential to gain information from catchment properties, some of them were transformed before use in generalisation, with the transformation chosen to make the distribution of the quantity less skewed.

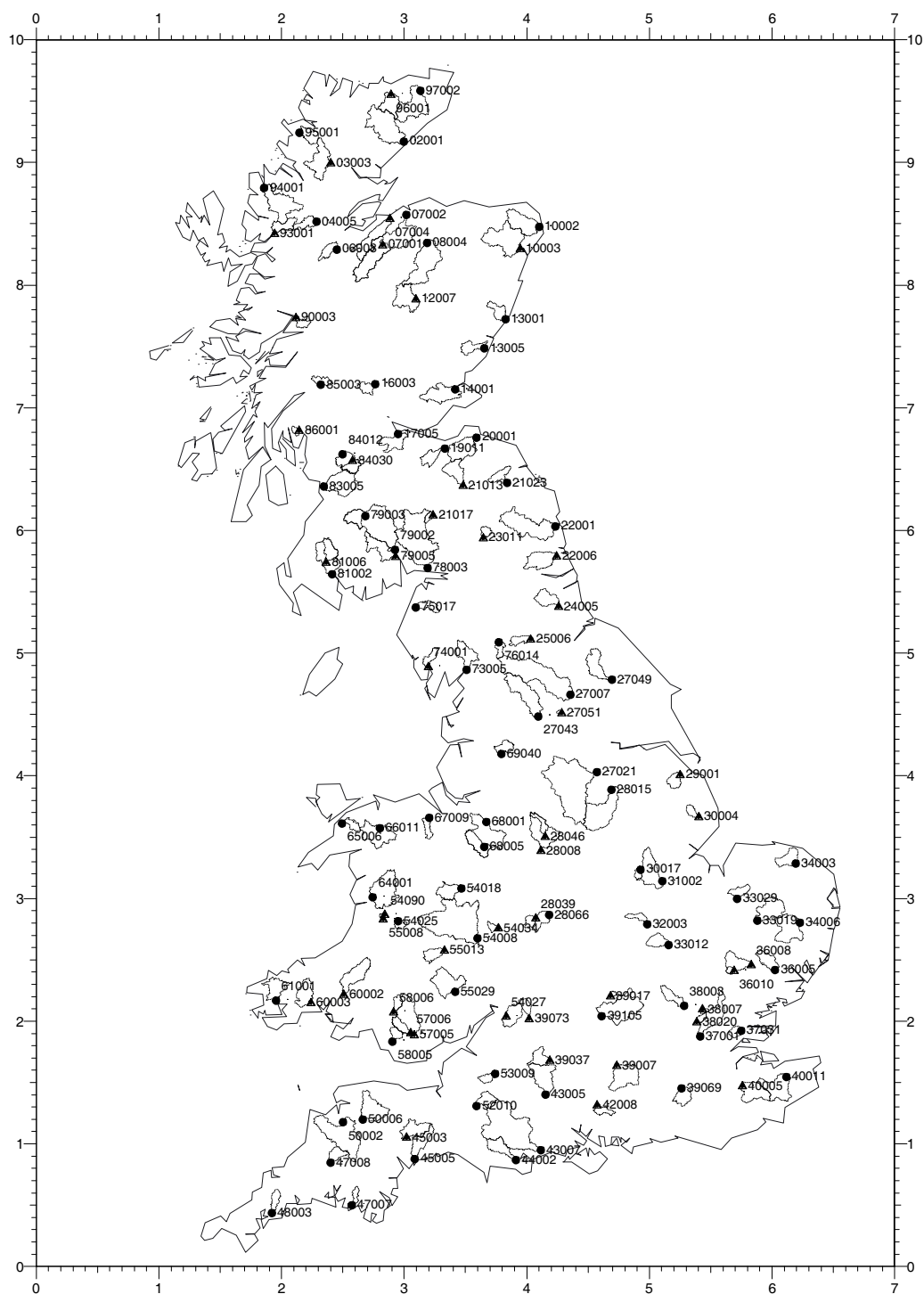


Fig. 1. Locations of catchment outlets (hourly – triangles, daily – circles), with catchment boundaries and ID numbers.

THE RAINFALL–RUNOFF MODEL

The model used here is a simplified version of the Probability Distributed Model (PDM) (Moore, 1985, 1999; 2007). It is based on conceptual water stores and represents non-linearity in the transformation of rainfall to runoff by a

probability distribution of soil moisture storage. This determines the time-varying proportion of the catchment which contributes to runoff, through either ‘fast’ or ‘slow’ pathways. The form used here (Fig. 2) has five parameters: reducing the number of parameters improves the

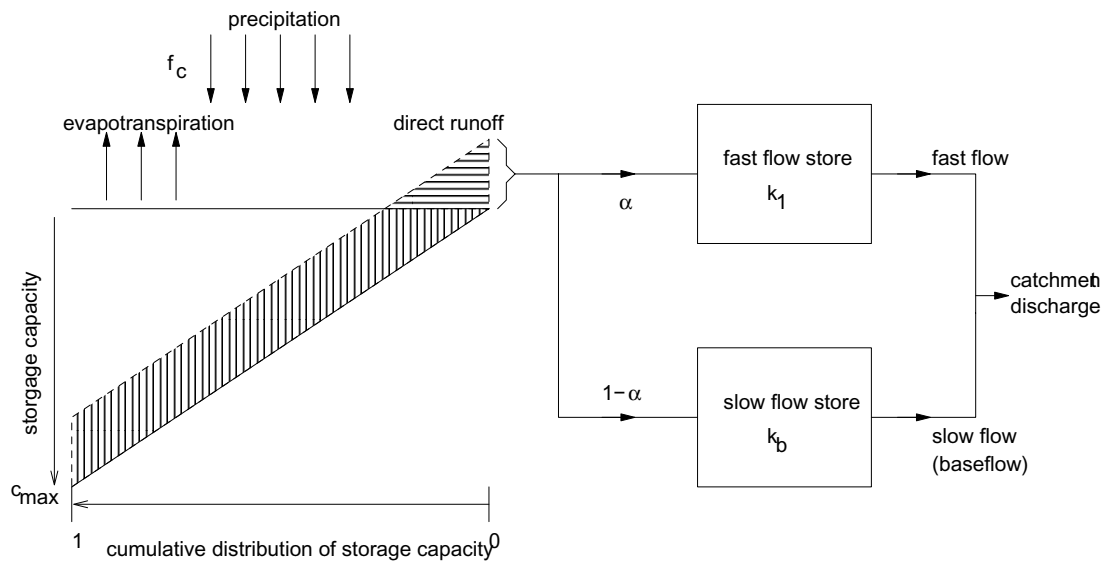


Fig. 2. Structure of the 5-parameter version of the PDM rainfall-runoff model

performance of spatial generalisation (Lamb *et al.*, 2000). Lee *et al.* (2005) tested various combinations of runoff and routing schemes, although not this particular one, and conclude that a form very similar to this is one of the best candidates for regionalisation.

Rainfall inputs to the soil store are first multiplied by a rainfall correction factor f_c , which can also, if required, compensate for loss or gain of water via lateral, sub-surface routes. The soil moisture store can be depleted by evaporation, with content of the store determining the proportion of the potential evaporation which actually occurs. The distribution of the soil moisture storage capacity with store depth is assumed to be uniform; the minimum store depth is set to zero and the maximum is given by the parameter c_{max} . The soil store, then, generates direct runoff from a varying proportion of the catchment area, depending on how full it is. It is assumed that a proportion α of the direct runoff goes to the (linear) fast flow store, whilst $1-\alpha$ goes to the (cubic) slow flow store. The time constants of the fast and slow flow stores are k_1 and k_b respectively. The catchment discharge, then, results from a combination of fast and slow flow.

Although the model as described here has five remaining parameters, only four of these require calibration or generalisation for a specific catchment. The parameter determining the split between the fast and slow flow stores, α , is set as $SPRHOST/100$, where $SPRHOST$ is standard percentage runoff inferred from soil information (a readily-available catchment property). This was deemed an appropriate simplification, because of the directly comparable meanings of α and $SPRHOST$.

MODEL CALIBRATION

The large number of catchments meant that automatic calibration was necessary, so a sequential procedure was developed which calibrated each of the model parameters in two passes. For the first pass, each parameter is calibrated in turn, following Monte-Carlo sampling of the parameter space of the so-far uncalibrated parameters. A different objective function is chosen for fitting each parameter, according to hydrological judgement and plots of objective functions versus parameter values. Each objective function considers the whole flow time-series but some give more weight to certain aspects of it. For example, when calibrating the parameter determining the overall water availability (f_c), only 30-day average flows are considered, whereas when calibrating the parameter of the fast flow store (k_1) more weight is given to the higher flows. For the second pass, each parameter is calibrated in turn, following Monte-Carlo sampling of its own value, with the values of the other parameters held at their previous calibrated values. This second pass allows a re-adjustment of parameter values, once other values have been estimated, and uses the same objective functions as the first pass. As this method is sequential, alternative ordering of parameters in the calibration is possible, and no testing has been done to assess the effect of this. The order used for each pass was f_c, k_b, c_{max}, k_1 , with the choice of order based on hydrological judgement and on the behaviour of the response surfaces in the objective function dot plots. Lastly, it was found helpful to allow a final re-calibration of the first parameter, this time using an objective function that concentrates on fit of the flood frequency curve.

A key consideration was how to merge information from catchments with data at different time steps. In theory, the PDM is structured so that its parameters have the same intrinsic meaning (in particular, the same units) regardless of the length of the internal time-step. However, tests on a small number of the hourly catchments showed that the parameters would be calibrated differently if they were treated at a daily time-step. Thus, using a daily internal time-step for the daily sites was not appropriate. Instead, a common internal time-step of one hour was adopted for all of the catchments. Tests showed that spreading the daily rainfall uniformly over the 24 hours of the day to produce the hourly inputs for the daily sites, and then averaging the simulated hourly flows to compare with observed daily flows, led to calibrated parameter values matching much more closely those using full hourly data rather than the daily time-step. Hence, this method was adopted.

An advantage of automatic calibration is that it enables the estimation of calibration uncertainty. In this case, a variation of the standard statistical method of jack-knifing (Shao and Tu, 1995) was used. This entails initial calibration of the model using all T years of data, then performing T recalibrations, in each of which one year of flow data is treated as missing (all rainfall data are retained, to maintain the water balance). The spread of these values gives an indication of the calibration uncertainty for each catchment due to the finite amount of data available. For a given catchment, the variance, σ^2 , of the estimation error of a given model parameter, α , is calculated from the T values for that parameter using the formula

$$\sigma^2 = \frac{T-1}{T} \sum_{i=1}^T \left(\alpha_{(i)} - \frac{1}{T} \sum_{j=1}^T \alpha_{(j)} \right)^2, \tag{1}$$

where $\alpha_{(i)}$ is the parameter estimate treating the i th year of flow data as missing.

SITE-SIMILARITY GENERALISATION

The idea behind site-similarity approaches to generalisation is to find a set of *calibrated* sites most similar to the target site — a pooling group — and then to form the parameter estimates for the target site by combining the calibrated parameter values for the sites in the pooling group.

Similarity is defined by Euclidean distance in a space of catchment properties that must be determined for a given model parameter:

$$dist_{a,b} = \sqrt{\sum_{j=1}^J \left(\frac{X_{a,j} - X_{b,j}}{\sigma_{X,j}} \right)^2}, \tag{2}$$

where j indicates one of a total of J catchment properties,

$X_{a,j}$ is the value of that catchment property at the a th site and $\sigma_{X,j}$ is the standard deviation of the property across all the N sites. The M closest neighbours (minimum distance) are selected to create a pooling group for the target site. The catchment property space, and so the pooling group for a target site, can differ for each model parameter. Note that, just as some of the catchment properties are transformed before use in generalisation, so too are some of the model parameters: log-transformations are used for k_1 and k_b .

Having formed the pooling group, the estimate α_a^{PG} of the model parameter at a target site a is calculated as a weighted average of the corresponding parameters from the sites in the pooling group. However, it is more convenient to write the expression for this as a weighted average of the estimated parameter values, α_m , over all catchments:

$$\alpha_a^{PG} = \frac{\sum_{m=1}^N h_{am} \alpha_m}{\sum_{m=1}^N h_{am}}. \tag{3}$$

Catchments not in the pooling group are given a weight h_{am} equal to zero but those in the pooling group are assigned weights to reflect their importance where this can reasonably be based on the distance measure $dist_{a,m}$ defined in Eqn. (2). In their simplest form, the weights for pooling group members can be expressed as

$$h_{am} = 1 - s_{am}, \tag{4}$$

where

$$s_{am} = \begin{cases} 0 & \text{for equal weights} \\ dist_{a,m}/dist_{a,max} & \text{for linearly decreasing weights} \\ (dist_{a,m}/dist_{a,max})^2 & \text{for quadratically decreasing weights} \end{cases}$$

(with $dist_{a,max}$ set to be 10% larger than the maximum distance of a pooling group member from the target site a).

In addition to subjective weighting schemes such as the distance weighting schemes above, schemes can be constructed to take some account of the differing uncertainties inherent in the estimates α_m for each catchment; an uncertainty-weighting scheme of the form

$$h_{am} = \frac{1}{1 + \sigma_{m,\epsilon}^2 / \sigma_\eta^2} \tag{5}$$

can be derived as providing an optimal set of weights. Here $\sigma_{m,\epsilon}^2$ describe the uncertainty (variance) in the calibrated model parameters for catchment m , and σ_η^2 the variation between the parameter values for catchments in the pooling group. In this scheme, less weight is given to those catchments for which the calibration variance $\sigma_{m,\epsilon}^2$ is high.

The different types of weighting considered here are

expressed in a combined form which includes both distance-weighting and uncertainty-weighting:

$$h_{am} = \frac{1 - s_{am}}{1 + k\sigma_{m,\varepsilon}^2} \quad (6)$$

where k is either estimated iteratively as the reciprocal of the generalisation variance σ_η^2 , or is set to zero so that the weighting is simply a distance weighting. The latter option is equivalent to assuming that the calibration variances $\sigma_{m,\varepsilon}^2$ are all zero: the calibration uncertainty is ignored in forming the weights.

MODEL FOR UNCERTAINTY

The initial development of an uncertainty-weighted scheme such as outlined above can be based on a simple statistical model that can also form the basis of measures for assessing how well various weighting schemes perform. The model is a local one, specific to each target catchment and the catchments included in the pooling group for that catchment. For convenience, the subscript m is used for catchments within the pooling group, while the special subscript $*$ is reserved for the target catchment. Firstly, it is assumed that the estimated parameters α_m are related to notional true values of the parameters μ_m via an error model:

$$\alpha_m = \mu_m + \varepsilon_m \quad (7)$$

Here, the true values μ_m are defined to be the values to which the parameter estimates would converge as the record length available for model calibration increases. The error terms ε_m correspond to calibration errors. There is an assumption that, within the pooling group, the true values μ_m are scattered about an unknown central value, β , and the true parameter value μ_* for the target catchment is an unknown value from this same population:

$$\mu_m = \beta + \eta_m \quad (8)$$

Here η_m represents the generalisation error, the extent to which catchment parameters are predictable within a pooling-group method. The term β represents a notional central value for sites in the pooling group, which will differ across pooling groups.

Note that the objective in creating a weighted combination of the estimates α_m is not to estimate the parameters that might have been obtained had calibration data been available for the catchment, but rather to estimate the true value μ_* , where

$$\mu_* = \beta + \eta_* \quad (9)$$

If the generalised estimate is

$$\hat{\mu}_* = \sum g_{*m} \alpha_m, \quad (10)$$

with $\sum g_{*m} = 1$, it follows that the error in estimating μ_* is given by

$$\begin{aligned} \mu_* - \hat{\mu}_* &= \mu_* - \sum g_{*m} \alpha_m \\ &= \eta_* - \sum g_{*m} (\eta_m + \varepsilon_m). \end{aligned} \quad (11)$$

From the above formula, the variance of the estimation error can be calculated and this gives a criterion to be minimised in selecting the weights g_{*m} . This calculation assumes that the calibration and generalisation errors are all uncorrelated, that the generalisation errors η_m, η_* have constant variances,

$$\text{var}(\eta_m) = \text{var}(\eta_*) = \sigma_\eta^2, \quad (12)$$

and that the calibration errors have variances $\sigma_{m,\varepsilon}^2$. The weights giving minimum variance for the error in Eqn. (11) are found to be

$$g_{*m} = \frac{w_m}{\sum w_j}, \quad (13)$$

where, for catchments in the pooling group for the target site,

$$w_m = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_{m,\varepsilon}^2} = \frac{1}{1 + \sigma_{m,\varepsilon}^2 / \sigma_\eta^2}, \quad (14)$$

and $w_m = 0$ otherwise.

In general, the variance of the generalisation error, σ_η^2 , is estimated iteratively because the required weighting scheme may depend on this unknown quantity. Note that the weighting schemes considered are not limited just to the weights given in Eqn. (14), which are optimal given some special assumptions, but include the more general form given in Eqn. (6). As will be seen, the same estimation scheme can also be employed when uncertainty weighting is not used at all. Firstly, weights of the form

$$K_m = 1 + k\sigma_{m,\varepsilon}^2, \quad (15)$$

are constructed. Here k is either set to zero, or as

$$k = 1 / \sigma_\eta^2, \quad (16)$$

where σ_η^2 takes the value generated in a previous iteration. Note that the choice $k = 0$ is equivalent to imposing the assumption that all the calibration-error variances are zero,

meaning that all of the sample estimation variance is attributed to the generalisation error variance. In either case, the following relation is treated as true:

$$\sigma_{\eta}^2 + \sigma_{m,\varepsilon}^2 = K_m \sigma_{\eta}^2, \quad (17)$$

where the quantities on the left hand side are theoretical variances while the σ_{η}^2 on the right hand side is the value that will be calculated in the present iteration. Estimation of σ_{η}^2 is based on a weighted sum of squared residuals calculated by treating each catchment in the calibration set as a target for generalisation. The same procedure can be applied either including or excluding each target catchment from the pooling group being used for that catchment.

In the following, there is a need to account for the fact that the location parameter β , used in Eqn. (8), varies with the target catchment and that the same model cannot hold simultaneously for all target catchments. Use is, therefore, made of the implicit assumption that, if β_m and β_* are values corresponding to each catchment, these values vary slowly enough for the condition

$$\beta_* - \sum g_{*m} \beta_m = 0, \quad (18)$$

to be approximately valid. The values β_m represent the generalised values that would be produced by a best-possible generalisation procedure of the given type for all catchments having catchment properties identical to catchment m .

If the generalised estimate for target catchment i is

$$\hat{\mu}_i = \sum g_{ij} \alpha_j, \quad (19)$$

with $\sum g_j = 1$, the weighted sum of squared residuals is given by

$$S^2 = \sum_{i=1}^N K_i^{-1} (\alpha_i - \hat{\mu}_i)^2. \quad (20)$$

The expected value of S^2 can be found by noting that

$$\begin{aligned} \alpha_i - \hat{\mu}_i &= \alpha_i - \sum g_{ij} \alpha_j = \beta_i + \eta_j + \varepsilon_j - \sum g_{ij} (\beta_j + \eta_j + \varepsilon_j), \\ &= \eta_j + \varepsilon_j - \sum g_{ij} (\eta_j + \varepsilon_j), \end{aligned} \quad (21)$$

where use is made of Eqn. (18), and thus that

$$\alpha_i - \hat{\mu}_i = \sum u_{ij} (\eta_j + \varepsilon_j) \quad (22)$$

where

$$u_{ij} = \begin{cases} 1 - g_{ij}, & i = j, \\ -g_{ij}, & i \neq j. \end{cases} \quad (23)$$

Thus

$$\begin{aligned} E(S^2) &= \sum_{i=1}^N K_i^{-1} \left\{ \sum_{j=1}^N u_{ij}^2 \text{var}(\eta_j + \varepsilon_j) \right\}, \\ &= \sum_{i=1}^N K_i^{-1} \left\{ \sum_{j=1}^N u_{ij}^2 K_j \right\} \sigma_{\eta}^2. \end{aligned} \quad (24)$$

Hence, an unbiased estimate of σ_{η}^2 , given the assumption made in setting k , is given by

$$\hat{\sigma}_{\eta}^2 = S^2 \left[\sum_{i=1}^N K_i^{-1} \left\{ \sum_{j=1}^N u_{ij}^2 K_j \right\} \right]^{-1}. \quad (25)$$

While σ_{η}^2 is a useful measure of how well a generalisation procedure performs, it must be augmented to encapsulate all the uncertainty in the estimated value produced by the generalisation procedure. The generalisation variance is only a part of this: σ_{η}^2 relates to variations about an unknown local mean value for the pooling-group and the procedure needs to account for errors in estimating that local mean. The adjustment required varies from catchment to catchment. The total estimation error for an ungauged catchment is, as in Eqn. (11),

$$\mu_* - \hat{\mu}_* = \eta_* - \sum g_{*m} (\eta_m + \varepsilon_m). \quad (26)$$

The error variance is, therefore, given by

$$\begin{aligned} \text{var}(\mu_* - \hat{\mu}_*) &= \sigma_{\eta}^2 + \sum g_{*m}^2 (\sigma_{\eta}^2 + \sigma_{m,\varepsilon}^2), \\ &= \sigma_{\eta}^2 \left\{ 1 + \sum g_{*m}^2 K_m \right\}. \end{aligned} \quad (27)$$

where K_m is as given in Eqn. (15). Derivation of this formula has assumed that, for a 'new' target catchment *, η_* is not the same as any of the η_m for the calibrated catchments.

PERFORMANCE MEASURE

The requirement from the investigation is to define the site-similarity approach for each model parameter, in terms of the catchment properties used in the distance measure, the number of catchments in the pooling group, and the best weighting scheme when combining pooling group parameters to form the target estimate. To compare different approaches, a measure of overall performance is required, in terms of how well the estimated parameter values agree with the values calibrated. The estimation-error variance derived above (Eqn. (27)) is not immediately useful as a performance measure for the generalisation procedure since it varies for each target catchment. This difficulty can be overcome by arguing that the calibration catchments are representative catchments to which the procedure might be applied in future and, hence, they provide a representative collection of catchment properties. An overall measure of

generalisation performance can, therefore, be created by averaging the estimation error variances that would be obtained for a collection of ungauged catchments that happened to coincide in catchment properties with the calibration set. This gives the measure σ_{TA}^2 of the total accuracy of the generalisation procedure as the average estimation error variance

$$\sigma_{TA}^2 = \sigma_{\eta}^2 \left\{ 1 + N^{-1} \sum_{i=1}^N \sum_{m=1}^N g_{im}^2 K_m \right\}. \quad (28)$$

This measure aims to encapsulate all the uncertainty in the estimated value produced by the generalisation procedure, of which the generalisation variance σ_{η}^2 is only a part: σ_{η}^2 relates to variations about an unknown local mean value for the pooling-group and the procedure needs to account for errors in estimating that local mean. Note that lower values of σ_{TA}^2 indicate better performance but they cannot be used to compare performance across different model parameters, just for the same model parameter across different generalisation approaches.

Investigation of approaches

The number of options within the site-similarity method means that a thorough investigation of all possible combinations is impracticable. Instead, options are investigated subject to other choices being fixed, using a comprehensive search procedure in each case to find the best performing combinations of catchment properties

NUMBER OF SITES IN THE POOLING GROUP

The first option relates to the value of M , the number of closest neighbours chosen to form the pooling group for a target site. Investigations suggested that, although the choice of M was not as important as other factors, the use of a pooling group with around ten members was preferable to a much larger group, particularly when more catchment properties are used to define proximity for the pooling group (Fig. 3). Although a smaller pooling group seems to perform marginally better for some parameters when five or more

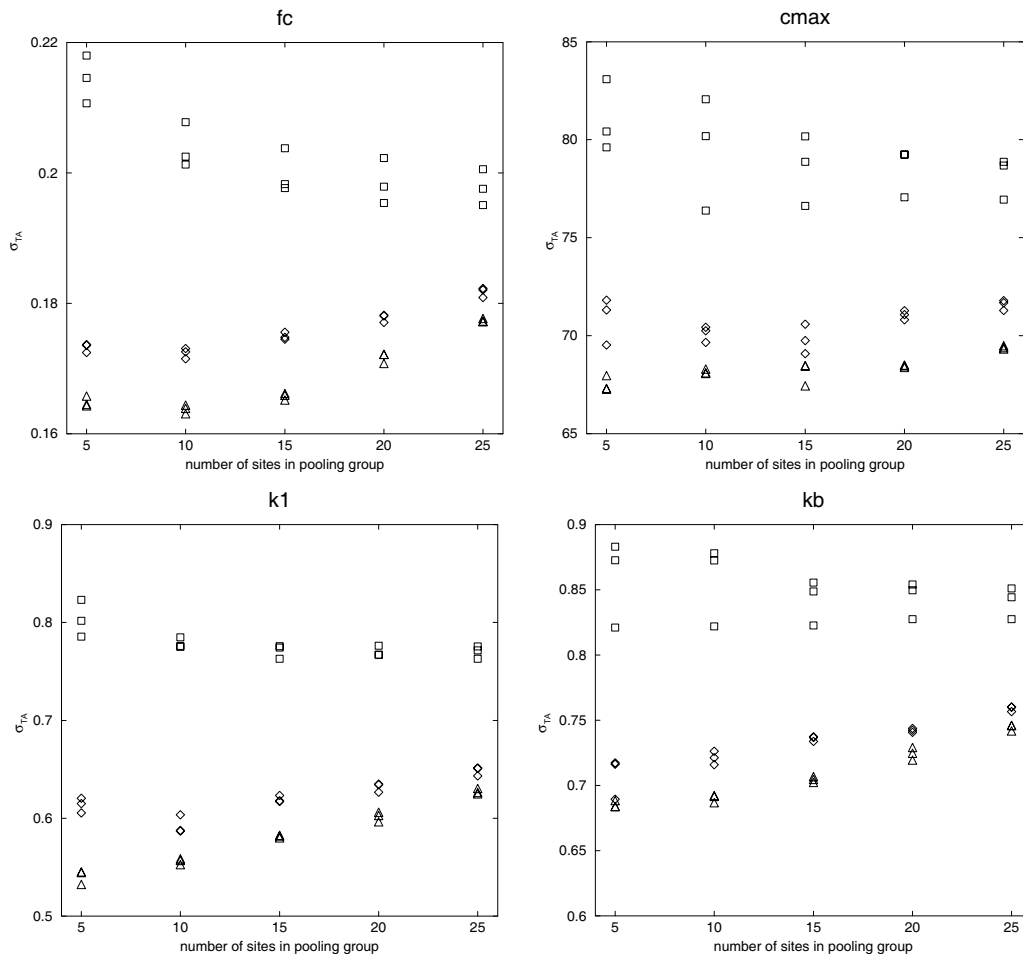


Fig. 3. The performance of site-similarity generalisation with size of pooling group, M . The performance measure σ_{TA} is plotted against the number of catchments used in the pooling group, for the top three combinations of 1, 3 and 5 catchment properties (respectively, squares, diamonds and triangles).

catchment properties are included, a pooling group size of ten was chosen for use in later investigations, due to concerns that, if pooling groups are too small, a single catchment might have an overly large impact on pooled parameter estimates.

DISTANCE WEIGHTING

The second option relates to the use of equal, linear or quadratic functions in the distance weighting. An exhaustive search of combinations of one up to six catchment properties was used to identify the top three combinations in each case. For each distance weighting, these are plotted against performance in Fig. 4. Although distance weighting does not make much difference to performance, linear distance is marginally preferable to either equal or quadratic for most parameters, particularly when more catchment properties are included. Linear distance weighting is, thus, used in later investigations.

ACCOUNTING FOR CALIBRATION UNCERTAINTY

The third option relates to the treatment of calibration uncertainty. Essentially similar analyses can be made which either account for the calibration errors as a specific source of variability, or ignore the existence of this source. In each case, the performance measure σ_{TA}^2 represents the typical uncertainty to be associated with the generalised parameter values. Results for the best three combinations of up to six catchment properties are shown in Fig. 5: the weighting scheme used in this example is linear distance weighting including weighting for calibration uncertainty but other cases reach the same conclusion. Clearly, the generalisation uncertainty is reduced substantially when the analysis takes the calibration uncertainty into account; then, the analysis can recognise that some of the variation in the empirical residuals $(\alpha_i - \hat{\mu}_i)$ can be ignored in assessing the proximity of the true value μ_* to the generalised value $\hat{\mu}_*$ for new target catchments. All the other results presented here take calibration uncertainty into account.

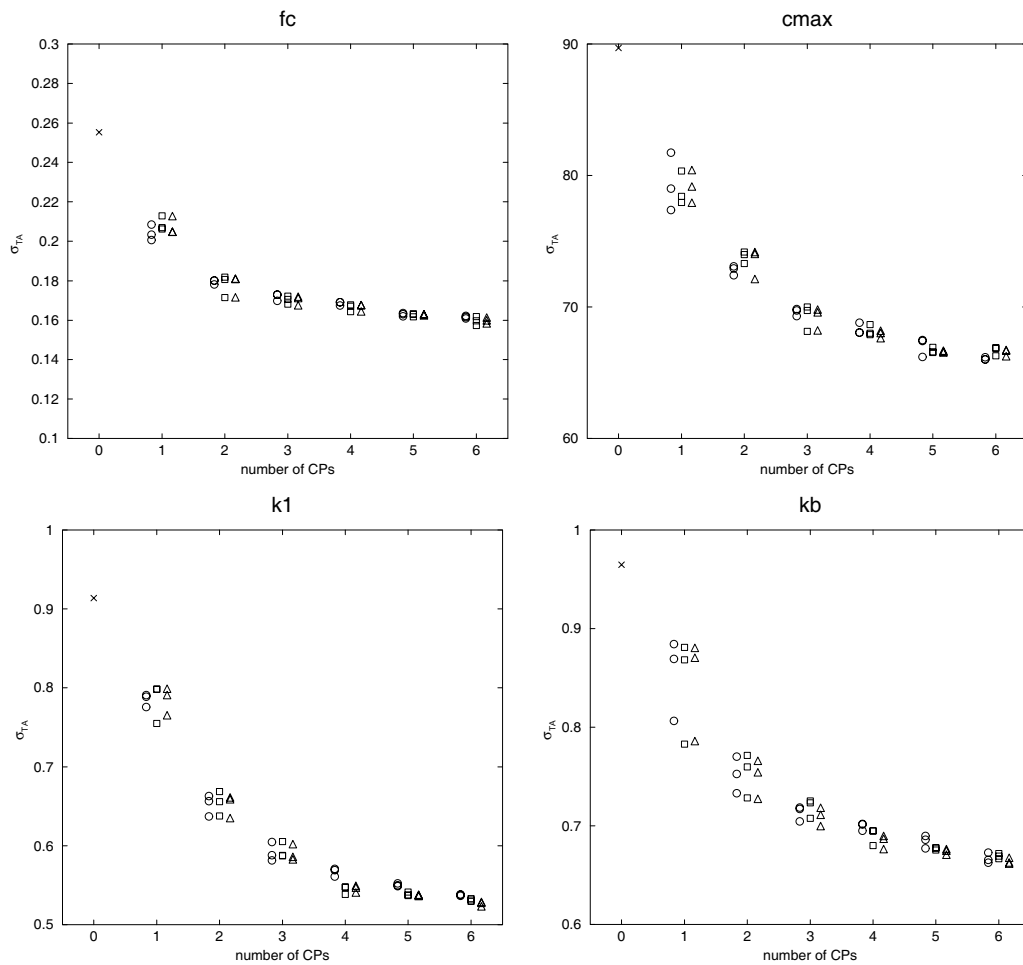


Fig. 4. Comparison of the performance of site-similarity generalisation with equal, linear and quadratic distance weighting (respectively, circles, squares and triangles). The performance measure σ_{TA} is plotted against the number of catchment properties used to define the pooling group, for the top three combinations of catchment properties (CPs) in each case. The cross, at 0 CPs on each figure, indicates the performance when the generalisation for any catchment is simply an average over all 119 catchments.

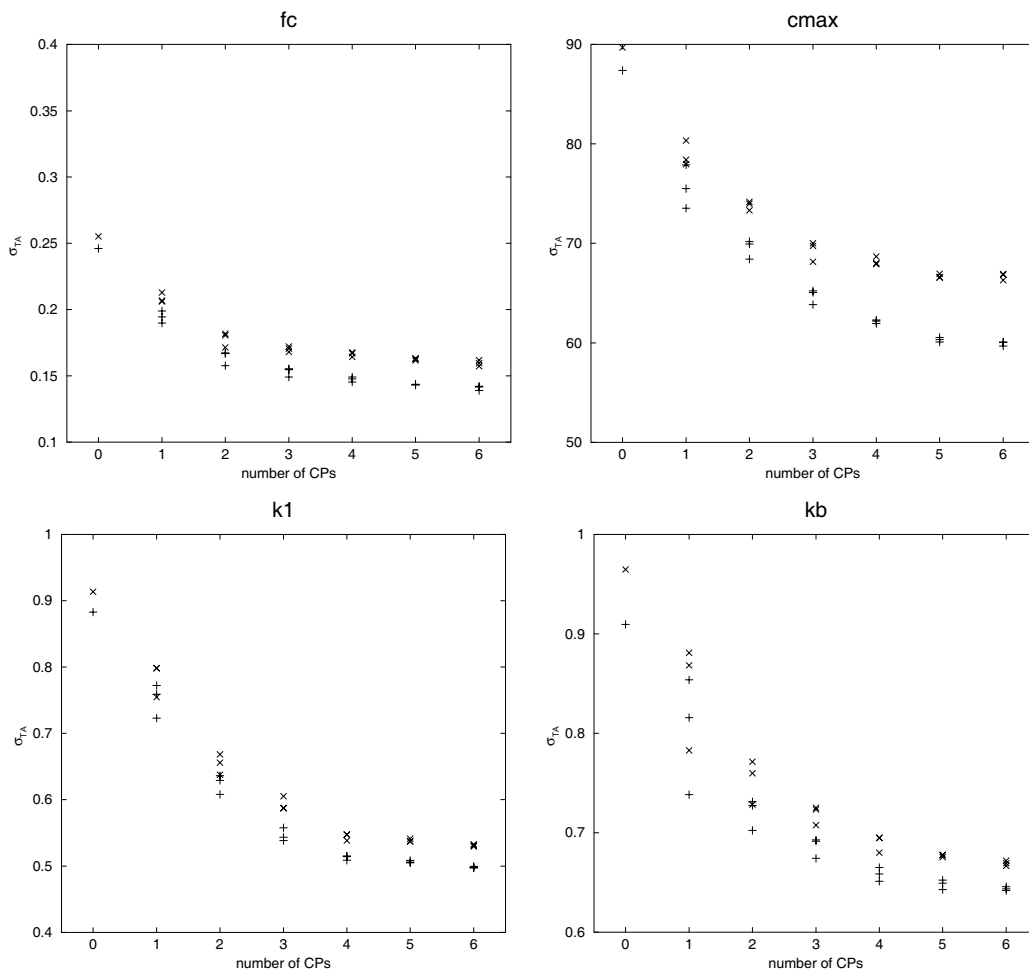


Fig. 5. Comparison of the performance of site-similarity generalisation using analyses which either do or do not take account of calibration uncertainty (respectively, plus signs and crosses). Both use linear distance weighting. The performance measure σ_{TA} is plotted against the number of catchment properties used to define the pooling group, for the top three combinations of catchment properties (CPs) in each case.

CHOICE OF CATCHMENT PROPERTIES

The final option, when forming the pooling group for a target site, is the choice of the catchment properties used to determine catchment similarity. This is not straightforward; many different combinations perform similarly in terms of overall parameter estimation. Hence, the hydrological relevance of particular properties to particular parameters must be considered.

The top ten combinations of one up to six catchment properties were listed and patterns of dominant properties (which appear in almost all combinations) and stable groupings (which pass down the lists as the number of catchment properties is increased) were identified. In addition, the inclusion of properties related to urban development (with the potential to increase peaks) and reservoirs and lakes (which can attenuate peaks) was important for parameters relating to the responsiveness of the catchment. No catchment property was ‘forced’ in but

combinations including certain properties were selected from the lists in preference to others. Plots of calibrated versus generalised parameter values were also inspected.

Accordingly, the set of catchments properties chosen for each model parameter is given in Table 1 (with catchment property definitions in Table 2), along with the R^2 values for the fit of calibrated and generalised parameters. Two R^2 values are given for each parameter: the first is where the estimate for a target catchment has included the site in its own pooling group while, for the second, the target site is excluded. The latter is more representative of potential performance for a truly ungauged site, whereas the former represents the potential to use the site-similarity method to augment the use of at-site data, for instance where only a short record is available. The large difference in performance is due to a target site potentially having a high weight when used in its own pooling group.

Table 1. The catchment properties (including their transformations) used for each model parameter. Definitions of the properties are given in Table 2.

Parameter	f_c	c_{max}	k_1	k_b
Catchment properties	$\sqrt{\text{AREA}}$	PROPWET	BFIHOST	BFIHOST
	BFIHOST	$\sqrt{\text{URBEXT}}$	DPLBAR	$\sqrt{\text{URBEXT}}$
	$\sqrt{\text{DPSBAR}}$	HOSTNG	$\sqrt{(1-\text{FARL})}$	HYDC
	$\sqrt{\text{SAAR}}$	$\sqrt{\text{HOSTP}}$	$\sqrt{\text{URBEXT}}$	$\sqrt{\text{LANDC}}$
	$\sqrt{\text{LANDB}}$	$\sqrt{\text{LANDA}}$	$\sqrt{\text{LANDB}}$	DRAIN2
	$\sqrt{\text{LANDC}}$			
R ² (inc. target)	0.78	0.70	0.82	0.74
R ² (exc. target)	0.55	0.41	0.66	0.50

Table 2. Definitions of catchment properties used in the generalisation

CP Name	Range, units	Source	Notes
AREA	[0,∞] km ²	FEH	DTM-derived
BFIHOST	[0,1] –	FEH	Base flow index, calculated from weighted average of HOST classes over the catchment
DPLBAR	[0,∞] km	FEH	Mean drainage path length
DPSBAR	[0,∞] m/km	FEH	Mean slope of DTM drainage paths to site
FARL	[0,1]	FEH	Index of flood attenuation due to reservoirs and lakes
PROPWET	[0,1] –	FEH	Proportion of time catchment wet (SMD<6mm)
SAAR	[0,∞] mm	FEH	Standard average annual rainfall, 1961-90
URBEXT	[0,1] –	FEH	Extent of urban/suburban land cover
HOSTNG	[0, 100] %	HOST	% of catchment area covered by HOST classes 16-29 (essentially ‘non-groundwater’)
HOSTP	[0,1]	HOST	Index of porosity as a weighted average of values inferred from HOST classes.
HYDC	[0,∞] cm/d	SEISMIC/ HOST	Saturated soil hydraulic conductivity, as weighted average of values inferred from HOST classes
LANDA	[0,1]	ITE	Proportion of catchment area covered by grassland based on ITE land cover data (classes 5-8,19,23)
LANDB	[0,1]	ITE	Proportion of catchment area covered by upland based on ITE land cover data (classes 9-13,17,24,25)
LANDC	[0,1]	ITE	Proportion of catchment area covered by trees based on ITE land cover data (classes 14-16)
DRAIN2	[0,∞] km/km ²	DTM	Drainage density (total length of river (km) divided by the catchment area (km ²))

Notes on sources:

FEH	Properties appearing on the FEH CD-ROM or based on FEH catchment properties (Institute of Hydrology 1999)
HOST	Properties derived from the HOST soil classification system (Boorman <i>et al.</i> 1995)
SEISMIC/HOST	Properties derived from the SEISMIC soils characteristics database for each HOST class
ITE	Properties derived from the ITE 1990 land cover classification (Fuller, 1993)
DTM	Properties derived from the CEH-Wallingford ‘Integrated Hydrological Digital Terrain Model’ (Morris and Flavin, 1990)

Generalised flood frequency results

In assessing the performance of the flood frequency curves estimated using the generalised parameters from the site-similarity approach defined earlier, curves for all 119 catchments cannot be presented because of space constraints.

However, examples are given in Fig. 6. Figure 7 summarises the results for each catchment in terms of percentage errors in estimated flood frequency (compared with flood frequency from observed flows for the same time period) at three return periods. Also shown, for comparison, are the

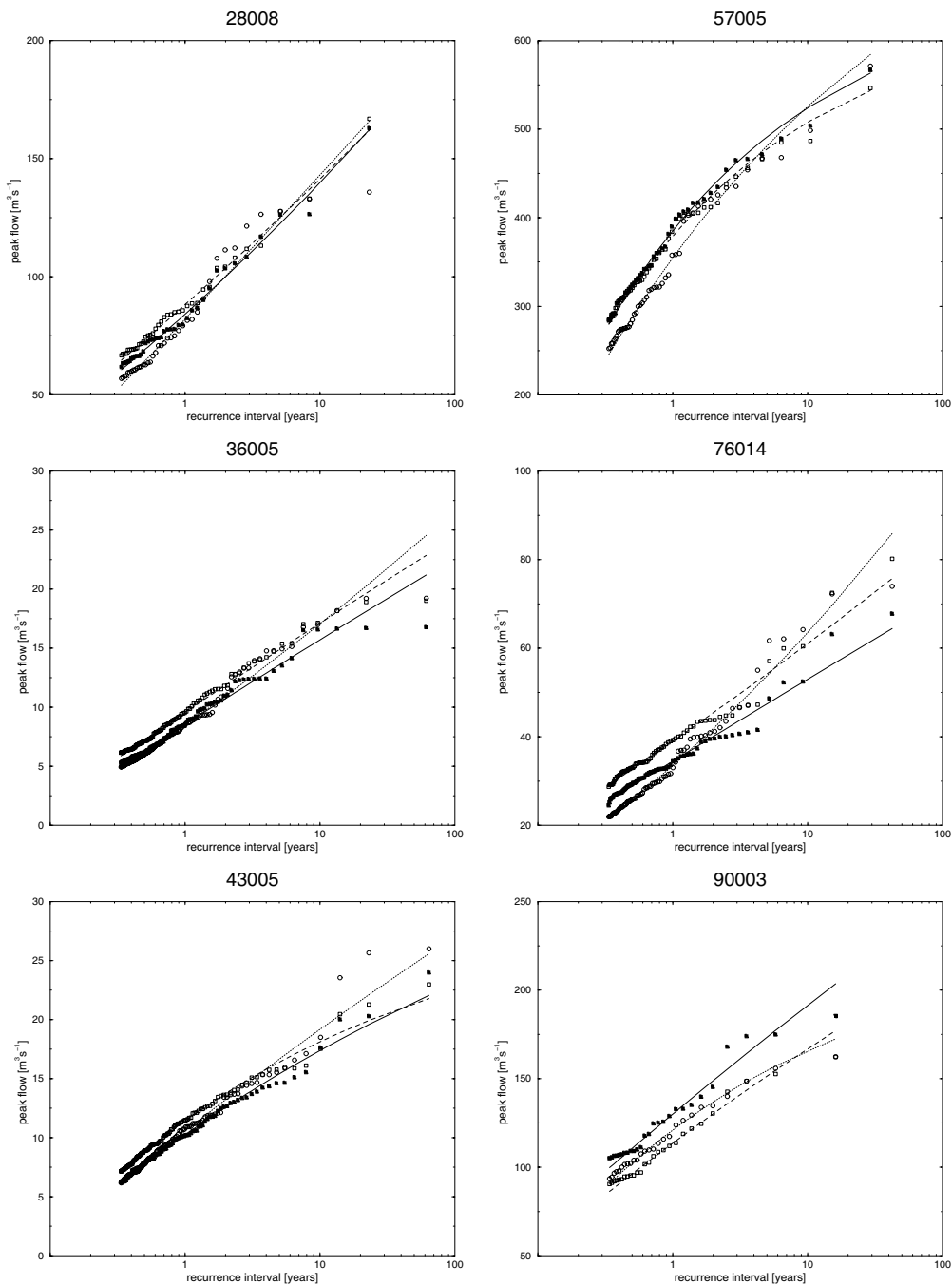


Fig. 6. Examples of generalised flood frequency curves (solid lines and filled squares) compared with those using calibrated parameters (dashed lines and open squares) and from observed flows (dotted lines and open circles).

errors when calibrated parameters are used for each catchment.

Figure 7 shows significant over-estimation in flood frequency at the 50-year return period for two catchments. The worst error is for 38003 (The Mimram at Panshanger Park, located just north of London, see Fig. 1.) This could

be due to significant groundwater abstraction from the catchment, whereas the generalised parameters could be representative of the catchment under more natural conditions. The other catchment for which there is significant overestimation is catchment 94001 (the Ewe at Poolewe) in north-west Scotland; this is dominated by a large loch and

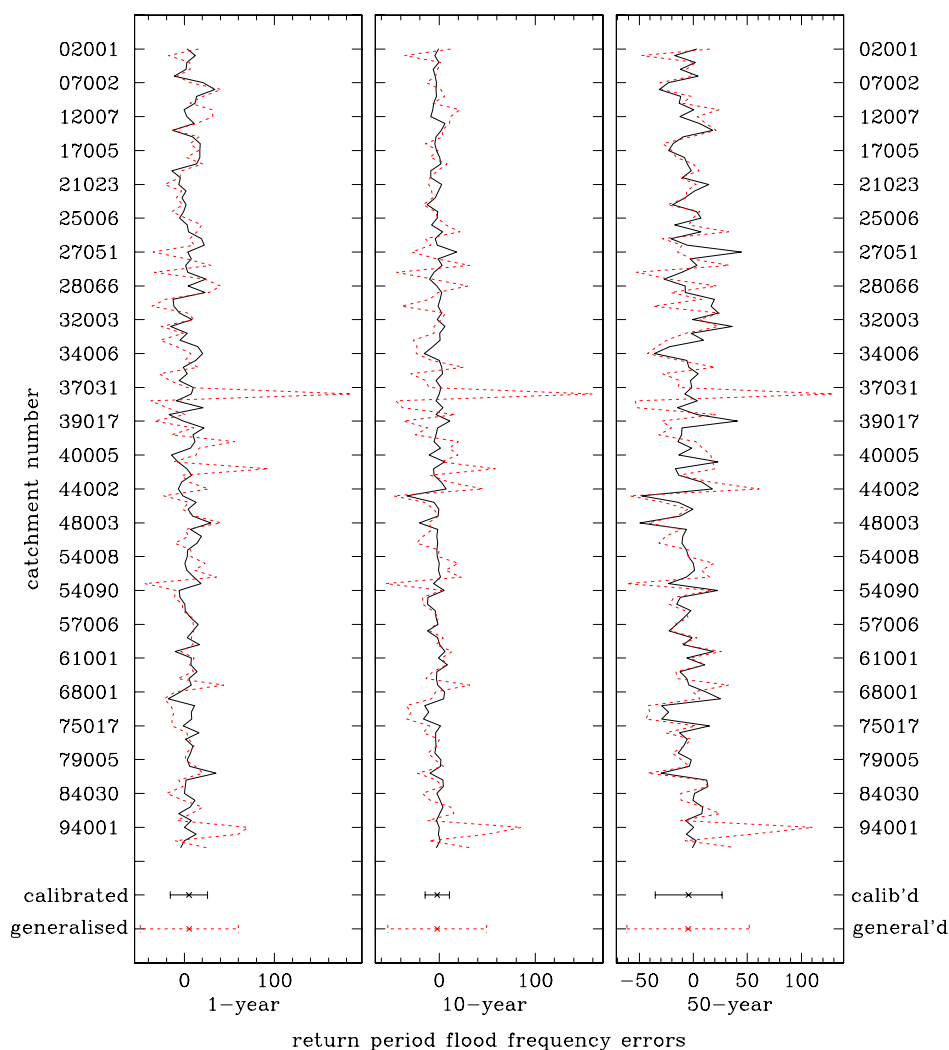


Fig. 7. Percentage errors in flood frequency at three different return periods, for parameters estimated using site-similarity (target included) (dotted line) compared with calibrated parameters (solid line) for each of 119 catchments. The vertical axis has every fifth catchment marked (in order of catchment number). The error bars at the bottom of the plots show the mean error \pm 2 standard deviations, for calibrated and generalised flood frequency curves.

so has a FARL value of 0.67, the lowest in the study set. The relative lack of similar catchments with which to form a pooling group for this catchment has thus affected the performance of the method there. A further catchment for which there is serious overestimation at higher return periods is 44002 (the Piddle at Baggs Mill, in southern England), perhaps reflecting difficulties of gauging higher flows from the catchment.

Figure 7 also suggests that, compared to *calibration*, generalisation tends to perform less well for catchments with numbers between 26000 and 45000, which lie roughly to the south/east of a line from Middlesbrough on the north east coast, to Bournemouth on the south coast. These catchments are more likely to have a significant groundwater component of flow, which could make consistent calibration

and generalisation more difficult, in the light of the simplified nature of the rainfall–runoff model. However, the errors for these catchments are not consistent in direction. For catchments outside this range, the trace for generalisation generally tracks that for calibration more closely. The same is true of the performance of generalisation when the target site is excluded from its pooling group but errors are generally marginally higher.

Table 3 summarises the generalisation performance in terms of the mean and standard deviation of the absolute percentage error at various return periods (simulated compared to observed flood frequency curve). The site-similarity generalisation including the target site gives average errors two to four times those for calibration; performance is slightly worse when the target site is

Table 4. Overall performance in terms of the mean and standard deviation (SD) across catchments of the absolute percentage errors in the estimated flood at a given return period. The performance using directly calibrated parameters is also given, for comparison.

	Mean and SD of absolute % errors	Return Period (years)					
		1.0	2.0	2.33	5.0	10.0	20.0
Calibration	Mean	9	6	6	4	5	7
	SD	7	5	4	4	5	7
Site-similarity(target included)	Mean	17	16	16	16	17	18
	SD	22	21	21	20	20	19
Site-similarity(target excluded)	Mean	23	23	23	23	23	24
	SD	36	36	35	34	34	33

excluded. These results compare well with those reported by Lamb *et al.* (2000), based on regression using a subset of the catchments and data used in this study. The mean errors when the target site is included are lower than those in the previous work, and those where the target site is excluded are similar or lower, particularly for higher return periods. The standard deviations reported here are larger than those in the previous work, particularly when the target site is excluded, but this is due to the very bad performance for catchment 38003 (discussed above), which is one of the additional catchments used in this study. When the results for this catchment are excluded from the analysis, the mean absolute errors are reduced by about 1% and their standard deviations to around 15% at each return period, when the target is included; this is an improvement over the performance in the previous work. When the target site is excluded, the corresponding figures are a reduction in mean absolute error of around 3% (consistently better than in the previous work), with standard deviations reduced to around 21% at each return period (still worse than in the previous work, particularly at lower return periods).

An alternative way of assessing overall performance is to look at the percentage of catchments where the errors fall within certain ranges. For example, if errors of less than 15% are good, then about 45% of catchments have errors (at return periods of between 2 and 50 years) below this threshold (33% when the target site is excluded), which compares with about 80% when calibrated parameters are used. If errors of over 30% are bad, then 23% of catchments have an error larger than this for at least one return period (30% when the target site is excluded), which compares with just 2% for calibrated parameters.

Plots of generalisation errors at different return periods against catchment properties show no obvious relationships, suggesting that no particular types of catchment perform consistently well (or badly) under site-similarity generalisation using the PDM.

Conclusions

This paper investigates a new approach to spatial generalisation of rainfall–runoff model parameters — site-similarity with pooling groups — for use in flood frequency estimation at ungauged sites. The pooling group comprises the ten calibrated catchments closest, in catchment property space, to the target site, where the catchment properties used to define the space differ for each parameter of the model. Accounting for the calibration uncertainty reduces the uncertainty attributed to generalised parameter values. Following a wide-ranging comparison of various options, the specific weighting scheme selected was to create parameter estimates for the target site by linear distance-weighting in conjunction with uncertainty-weighting.

The method devised shows great promise for flood frequency estimation at ungauged sites, because it performs relatively well for a wide range of catchments across Great Britain. It could be particularly useful at sites where some data are available, as these can be incorporated in the method by allowing the target site to be included in its own pooling group, with appropriate weighting (different to that discussed in this paper). A potential disadvantage of site-similarity is that application of the method is more complex than the more standard regression, as it requires the definition of a different pooling group for each parameter and the calculation of weights, rather than simply putting into an equation the values of catchment properties for the ungauged site. However, it is easy to expand the set of calibrated catchments on which the method is based; this could be important in improving the performance of the method for more unusual catchments, as is demonstrated by the poor performance for a Scottish catchment dominated by a large loch (and so with a very low FARL value), due to the lack of sufficiently similar catchments with which to form a pooling group.

In terms of flood frequency estimation, the site-similarity

method developed here generally performed better than using a regression based on a subset of the data. However, that previous work used standard regression whereas the weighting scheme developed here to incorporate calibration uncertainty could also be applied within a weighted regression. Future work will compare the performance of the present site-similarity method with weighted regression based on the same data set.

Calibration uncertainty is important in the methodology applied here, both in adjusting the variation shown by the empirical residuals, so as to represent the true generalisation uncertainty, and in adapting the weights of catchments in the pooling group, so that those with less confidence in calibration contribute less to the generalisation. The calibration uncertainty was estimated through a variation on jack-knifing. Other techniques could be applied, such as bootstrapping, or calibration uncertainty could be related directly to length of record. Merz and Blöschl (2005) found that kriging-with-uncertainty, directly related to record length, performed better than kriging alone when regionalising statistics of flood frequency curves for ungauged sites. In future work, it is planned to assess the effect of alternative methods for estimating calibration uncertainty. In addition, techniques are being developed to incorporate uncertainty bounds on the generalised flood frequency curves.

Acknowledgements

This study was supported by the Department for Environment, Food and Rural Affairs (Defra), project FD2106, and the Scottish Executive.

References

- Abdulla, F.A. and Lettenmaier, D.P., 1997. Development of regional parameter estimation equations for a macroscale hydrologic model. *J. Hydrol.*, **197**, 230–257.
- Acreman, M. and Wiltshire, S.E., 1987. Identification of regions for regional flood frequency analysis, (abstract). *EOS Trans. AGU*, **68**, 1262.
- Acreman, M. and Wiltshire, S.E., 1989. The regions are dead; long live the regions. Method of identifying and dispensing with regions for flood frequency analysis. In: *FRIENDS in Hydrology*, L. Roalds et al. (Eds.), *IAHS Publication no. 187*, 175–188.
- Boorman, D.B., Hollis, J.M. and Lilly, A., 1995. *Hydrology of soil types: a hydrologically based classification of soils in the United Kingdom*. IH Report No. 126, Institute of Hydrology, Wallingford, UK.
- Burn, D.H., 1990. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resour. Res.*, **26**, 2257–2265.
- Burn, D.H. and Boorman, D.B., 1992. *Catchment classification applied to the estimation of hydrological parameters at ungauged sites*. IH Report 118, Institute of Hydrology, Wallingford, UK.
- Calver, A., Lamb, R., Kay, A. L. and Crewett, J., 2001. *The continuous simulation method for river flood frequency estimation*. Defra Project FD0404 Final Report, Centre for Ecology and Hydrology, Wallingford, UK.
- CEH and BGS, annually. *Hydrological Data UK: Hydrometric Register and Statistics*. Centre for Ecology and Hydrology and the British Geological Survey, Wallingford, UK.
- Fuller, R.M., 1993. The land cover map of Great Britain. *Earth Space Rev.*, **2**, 13–18.
- Holmes, M.G.R., Young, A.R., Gustard, A.G. and Grew, R., 2002. A region of influence approach to predicting flow duration curves in ungauged catchments. *Hydrol. Earth Syst. Sci.*, **6**, 721–731.
- Houghton-Carr, H., 1999. Restatement and application of the Flood Studies Report rainfall-runoff method. In: *Flood Estimation Handbook IV*, Institute of Hydrology, Wallingford, UK.
- Institute of Hydrology, 1999. *Flood Estimation Handbook* (5 volumes). Institute of Hydrology, Wallingford, UK.
- Jones, S.B., 1983. *The estimation of catchment average point rainfall profiles*. IH Report No. 87, Institute of Hydrology, Wallingford, UK.
- Kokkonen, T.S., Jakeman, A.J., Young, P.C. and Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: model regionalisation from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. *Hydrol. Process.*, **17**, 2219–2238.
- Lamb, R., Crewett, J. and Calver, A., 2000. *Relating hydrological model parameters and catchment properties to estimate flood frequencies from simulated river flows*. In: Proc. BHS 7th National Hydrology Symposium, September 2000, Newcastle, UK, 3.57–3.64.
- Lee, H., McIntyre, N., Wheater, H. and Young, A., 2005. Selection of conceptual models for regionalisation of the rainfall-runoff relationship. *J. Hydrol.*, **312**, 125–147.
- McIntyre, N., Lee, H., Wheater, H. and Young, A., 2004. Tools and approaches for evaluating uncertainty in streamflow predictions in ungauged UK catchments. In: *Complexity and Integrated Resources Management*, C. Pahl-Wostl, S. Schmidt, A.E. Rizzoli and T. Jakeman, (Eds.). Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society (iEMSs), **3**, 1171–1176.
- Merz, R. and Blöschl, G., 2004. Regionalisation of catchment model parameters. *J. Hydrol.*, **287**, 95–123.
- Merz, R. and Blöschl, G., 2005. Flood frequency regionalisation—spatial proximity vs. catchment attributes. *J. Hydrol.*, **302**, 283–306.
- Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrolog. Sci. J.*, **30**, 273–297.
- Moore, R.J., 1999. Real-time flood forecasting systems: Perspectives and prospects. In: *Floods and landslides: Integrated Risk Assessment*, R. Casale and C. Margottini (Eds.), Chapter 11, Springer, Berlin, Germany. 147–189.
- Moore, R.J., 2007. The PDM rainfall-runoff model. *Hydrol. Earth Syst. Sci.*, **11**, 483–499.
- Morris, D.G. and Flavin, R.W., 1990. *A digital terrain model for hydrology*. Proc. 4th International Symposium on Spatial Data Handling, Zurich, Switzerland. **1**, 250–262.
- NERC, 1975. *Flood Studies Report*, 5 vols. Natural Environment Research Council, London, UK.
- Parajka, J., Merz, R. and Blöschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. *Hydrol. Earth Syst. Sci.*, **9**, 157–171.
- Pilgrim, D.H., Cordery, I. and French, R., 1969. Temporal patterns of design rainfall for Sydney. *Civil Eng. Trans., Inst. Eng., Australia*, **CE11**, 9–14.

- Post, D.A. and Jakeman, A.J., 1999. Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model. *Ecol. Model.*, **123**, 91–104.
- Sefton, C.M. and Howarth, S.M., 1998. Relationships between dynamic response characteristics and physical catchment descriptors of catchments in England and Wales. *J. Hydrol.*, **211**, 1–16.
- Seibert, J., 1999. Regionalisation of parameters for a conceptual rainfall-runoff model. *Agr. Forest Meteorol.*, **99**, 279–293.
- Servat, E. and Dezetter, A., 1993. Rainfall-runoff modelling and water resources assessment in northwestern Ivory Coast. Tentative extension to ungauged catchments. *J. Hydrol.*, **148**, 231–248.
- Shao, J. and Tu, D., 1995. *The Jackknife and Bootstrap*. Springer, New York, USA.
- Thompson, N., Barrie, I.A. and Ayles, M., 1982. *The Meteorological Office Rainfall and Evaporation Calculation System: MORECS (July 1981)*. Hydrological Memorandum No. 45, Met Office, Bracknell, UK.
- Xu, C.Y., 1999. Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resour. Manage.*, **13**, 353–368.
- Yokoo, Y., Kazama, S., Sawamoto, M. and Nishimura, H., 2001. Regionalization of lumped water balance model parameters based on multiple regression. *J. Hydrol.*, **246**, 209–222.