Journal of Hydrology (2007) 333, 517–531

available at www.sciencedirect.com

ScienceDirect

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

Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment

M. Goswami *, K.M. O'Connor, K.P. Bhattarai

Department of Engineering Hydrology, National University of Ireland, Galway, Ireland

Received 3 September 2005; received in revised form 31 July 2006; accepted 20 September 2006

KEYWORDS
Ungauged catchment;
Regional analysis;
Multi-model;
Combination techniques

Summary Flow simulation in ungauged catchments is presently regarded as one of the most challenging tasks in surface water hydrology. Many of the ungauged catchments are located in the headwaters of rivers in mountainous regions of the world having enormous potential for sustainable water resource development. However, due to inaccessibility, rugged and inhospitable terrain, and historical lack of foresight concerning the need to have these headwaters adequately gauged, their potential is not readily realizable. Many downstream sites also suffer from non-availability of site-specific data as even in countries having extensive networks of gauged stations data may not be available at sites where these are most needed. As predictive tools for water resources, water quality, natural hazard mitigation and water availability assessment are generally data-driven, the lack of adequate hydrometric records poses difficult problems for planners, engineers, managers, and stake-holders alike.

In this study, a methodology is developed for flow simulation in ungauged catchments using a regionalisation and multi-model approach involving a suite of rainfall–runoff models and combination techniques. Daily observed hydrometeorological data for 12 French catchments are used for illustrating the procedures. Following a preliminary investigation of the regional homogeneity of that group of catchments, three regional flow simulation techniques are applied. Although all 12 catchments are gauged, initially each catchment is successively considered as being ungauged for the purpose of flow simulation in that catchment, their actual discharges being subsequently used for evaluating the performance of the flow estimation procedures for the catchment. The Nash-Sutcliffe efficiency index (R^2) is used for assessing and ranking the relative performances of the regionalisation–model couples to identify the most appropriate couple for the region. The final step of applying that couple to a truly ungauged (13th) catchment in the region is described. Results are presented and conclusions drawn on the efficacy of the regional-multi-model

* Corresponding author. Tel.: +353 91 493172; fax: +353 91 495565.
E-mail address: monomoy.goswami@nuigalway.ie (M. Goswami).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2006.09.018

approach. Of the couples considered, the pooling method of regionalisation coupled with the conceptual soil moisture accounting and routing (SMAR) model is deemed to be the best for simulating flow in an ungauged catchment in the region.

Introduction

The need for methods to deal with ungauged catchments

With the growing demand to harness untapped potential of river water resources in many parts of the world, the need to devise new approaches and methodologies for assessment of water resources from these sources is also increasing. In this context, a typical problem often being faced is to simulate flow in an ungauged or poorly gauged catchment, referred to collectively in this paper by the term 'ungauged catchment'. Usually such ungauged catchments are located in headwater regions. Absence of any historical data records or mere inadequacy of whatever information is available for catchments in headwaters is generally caused by inaccessibility, inhospitable terrains, and historical lack of foresight of planners and developers to have potential sites gauged for harnessing the water resource of such catchments in the future. Apart from the headwater regions, many potential sites even in the downstream reaches also suffer from insufficiency of site-specific records of data. In some countries, having wide and extensive networks of gauge stations, data in many cases do not exist at locations where these are most needed. Lack of data, both qualitative as well as quantitative, often inhibits the undertaking of scientific analyses for such catchments which are required for purposes such as assessing the water resource, ensuring its long term availability, forecasting its occurrence over short lead-times, predicting its future occurrence, and developing its source.

Literature review

Following the endorsement of the ''PUB (Prediction in Ungauged Basins) Science and Implementation Plan'' by the IAHS (International Association of Hydrological Sciences) Bureau and the IAHS General Assembly in 2003, and the adoption of the ''IAHS Decade on PUB: 2003–2012'', scientific approaches and systematic efforts are being currently orchestrated by many research groups and individuals for the ''prediction of stream flow, sediment and water quality variables at multiple scales, which is not based on the availability of measured data of these variables'', but ''requires the development of new predictive approaches that are based on deep understanding of hydrological functioning at multiple space–time scales.'' (Sivapalan and Schaake, 2003). Transfer of hydrological information from one or more gauged catchments to a contiguous ungauged catchment by extrapolation from the gauged data, observation by remote sensing, hydrological model simulation, and integrated meteorological and hydrological modelling are recognised as potential predictive approaches. In this study, some procedures for flow simulation in ungauged catchments by regionalisation and rainfall–runoff model simulation are discussed.

Ideally, a physically based distributed model like the Système Hydrologique Européen (Abbott et al., 1986) should be most suited for application to ungauged catchments but, unfortunately, the need to calibrate such models with a large volume of different types of data has largely restricted their application. Hence their initially perceived superiority in this context has not been realised. In earlier works, multiple regressions involving empirical and black-box models were used to relate dominant hydrological behaviour with rainfall and topographic factors. The method advocated by Nash (1960) was based on estimating the values of two parameters (lag and shape, in terms of dimensionless moment ratios) of the Gamma Instantaneous Unit Hydrograph (IUH) model for an ungauged catchment by using regional relations involving topographical characteristics derived from a study of 60 gauged British catchments. The Nash study was revisited, refined, and greatly extended in the Flood Studies Report (NERC, 1975) and its subsequent revision (Flood Estimation Handbook, 1999). The technique of multiple regressions has been applied extensively, primarily to estimate event and unit hydrograph characteristics, a good description of which can be found in Sefton and Howarth (1998).

Given that parameters of a conceptual model are the synthesis of the physiographic and hydrometeorological characteristics of a catchment, attempts are made to relate the parameter values of a single hydrological model to some catchment-specific characteristics for application to ungauged catchments. Vandewiele and Elias (1995) used a monthly water balance model on Belgian catchments, and applied the techniques of kriging and maximum likelihood estimates concluding that kriging produced better estimates of parameters for use in ungauged catchments. Sefton and Howarth (1998) and Littlewood (2003) attempted to derive relationship between the dynamic response characteristics (DRC) and land cover, soil and climatic variables using a variant of the IHACRES model. Dunn and Lilly (2001) used the hydrology of soil type (HOST) system of soil classification in the UK to develop a spatially distributed DIY model, originally developed as a distributed conceptual model for the river Ythan in Scotland (Dunn et al., 1998), with the objective of transferring calibrated parameter sets from one catchment to another. Merz and Blöschl (2004) used eight regionalisation methods to estimate model parameters for ungauged catchments, and found that methods based on spatial proximity performed better than the regression methods based on catchment attributes. Hundecha and Bárdossy (2004) applied a regionalisation scheme using a linear transfer function to relate parameter values to land use, soil types, size, slope and shape of catchments. The IAHS publication by Diekkrüger et al. (1999) included a number of relevant studies involving regionalisation in hydrology. However, reliability of all studies for flow assessment in ungauged catchments by regional analysis depends on the type, quantity and quality of available data at gauged sites in the region, and the degree of similarity of the gauged and the ungauged sites (Littlewood, 2003). All such studies were generally ''limited in terms of statistical accuracy'' (Sefton and Howarth, 1998), producing rather poor results. Andréassian et al. (2003) cited over-parameterisation, dependency on input data bias, and lack of a systematic link between parameter precision and model efficiency as the three main factors which complicate the regionalisation of conceptual rainfall—runoff models.

The present authors are not aware of other published work on the use of the multi-model approach for the simulation of flow in ungauged catchments. However, Shamseldin et al. (1997) and Shamseldin and O'Connor (1999), for example, used a multi-model approach in applying the simple average method, the weighted average method and the neural network method to combine the outputs of rainfall—runoff models to obtain better consensus discharge estimates for gauged catchments. Xiong et al. (2001) used a Takagi—Sugeno model (Takagi and Sugeno, 1985) in a flood forecasting study, combining the forecasts of five different rainfall—runoff models. See and Abrahart (2001) used an amalgamation of neural network, fuzzy logic, statistical, and persistence forecasts to produce a single predicted output following a multi-model data-fusion approach to hydrological forecasting. Working within a multi-model framework, Aspinall (2004) used multiple models of land use patterns to draw inference from a set of time-variant models. With this growing evidence of the advantages of the multi-model approach, its application to flow simulation in an ungauged catchment is considered worthy of investigation in this study.

Scope of the study

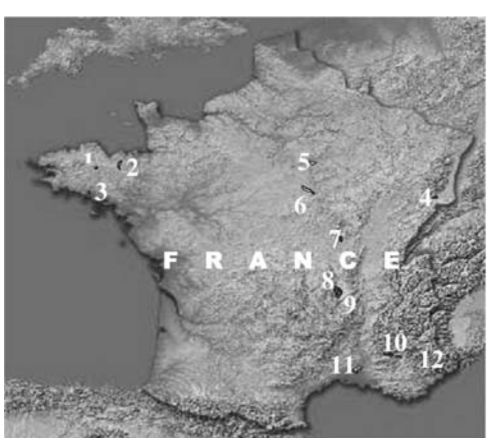
The focus of this study is the development of a methodology for simulating the flow in an ungauged catchment for which its rainfall and evaporation data together with the data of a number (12) of gauged catchments in the region are available. It is based on a comparison of the results of the application of selected flow regionalisation methods coupled with rainfall—runoff models to the available gauged catchments. Each catchment is treated in turn as if it was ungauged in order to identify the best couple for simulating flows in the region, having validated the procedures by checking the simulated discharges against their actual values. The selected optimum couple can be applied to simulate the discharge in a truly ungauged (13th) catchment in the region by using the data of all (12) gauged catchments to calibrate the couple for that catchment using its actual rainfall and evaporation data. Although the simulation of flow for an actual ungauged (e.g. 13th) catchment in the region is not carried out in this study, the procedure for doing so is described. Naturally, the simulated discharges so obtained for such an ungauged can not be validated.

In this study, three methods of regional analysis are applied for simulation of flow in ungauged catchments using records of observed data from neighbouring catchments. The procedure involves an amalgamation or combination of data and also of model results (in a multi-model context) obtained for the gauged catchments for application to the ungauged catchment under consideration. In this context, the term 'region' is not restricted to geographical proximity. Contiguous or local catchments are generally deemed to be included in the region. The theory and the applicability of the regionalisation approaches for flow simulation in ungauged catchments are demonstrated using the data from 12 French catchments considered in two combinations, i.e. the whole set of gauged catchments and a sub-set of three catchments.

Six black-box models and one conceptual model are used in this study as individual substantive rainfall—runoff models. For simulating flow by combining outputs from different models, three combination techniques are applied. Results are presented and conclusions are drawn on the potential of the seven individual models, the multi-model approach, and three regionalisation schemes.

The catchments and the data characteristics

Twelve French catchments, ranging in area from 32.1 km^2 to 371 km^2, are considered. The locations of these catchments, denoted by their respective station codes, are shown in Fig. 1. A summary of the catchment characteristics is given in Table 2. Topographically, the mean altitude, the altitude at the highest point, and the altitude at the outlet of the three catchments in the north-west, namely J2034010,



No.	Station	Name
1	J2034010	Le Guindy à Plouguiel
2	J3024010	Le Guillec à Trézilidé
3	J4124420	La Rivière de Pont-l'Abbé à Plonéour-Lanvern
4	A1522020	La Lauch à Guebwiller
5	H5723011	L'Orgeval à Boissy-le-Châtel
6	H3613020	Le Lunain à Épisy
7	H2001020	L'Yonne à Corancy
8	K0744010	L'Anzon à Débats-Rivière-d'Orpra
9	K0753210	Le Lignon du Forez à Boën
10	V6035010	Le Toulourenc à Malaucène
11	Y3514020	Le Vistre à Bernis
12	Y5615030	Le Loup à Villeneuve-Loubet

Figure 1 Location of 12 catchments in France.

Table 1 Individual models and references

Model	Acronym	Type	Reference for description
Parametric simple linear model	P-SLM	System theoretic	Kachroo and Liang (1992), Goswami et al. (2002)
Non-parametric simple linear model	NP-SLM	System theoretic	Kachroo and Liang (1992), Goswami et al. (2002)
Parametric linear perturbation model	P-LPM	System theoretic	Kachroo and Liang (1992), Goswami et al. (2002)
Non-parametric linear perturbation model	NP-LPM	System theoretic	Kachroo and Liang (1992), Goswami et al. (2002)
Linearly varying gain factor model	LVGFM	System theoretic	Ahsan and O'Connor (1994), Goswami et al. (2002)
Artificial neural network model	ANN	Data-driven	Shamseldin (1997), Goswami et al. (2002)
Soil moisture accounting and routing model (original form)	SMAR	Conceptual	Kachroo (1992)
Variant of SMAR model for application to conservative system (with groundwater modification)	SMARG	Conceptual	Goswami et al. (2002)
Variant of SMAR model for application to non-conservative system (loss/gain module before groundwater storage module)	SMAR-NC1	Conceptual	Goswami and O'Connor (2005)
Variant of SMAR model for application to non-conservative system (loss/gain module after groundwater storage module)	SMAR-NC2	Conceptual	Goswami and O'Connor (2005)

Table 2 Catchment characteristics

Station code no.	Area (km^2)	Length of longest stream (km)	Altitude at outlet (m)	Altitude at highest point (m)	Mean altitude (m)
J2034010	125	40.737	20	300	83
J3024010	43	22.652	35	120	85
J4124420	32.1	20.456	15	158	84
A1522020	68.1	13.572	290	1420	775
H5723011	104	12.818	77	185	148
H3613020	252	51.447	51	201	131
H2001020	98	29.615	332	900	592
K0744010	181	29.192	410	1345	755
K0753210	371	66.707	470	1628	863
V6035010	150	39.242	311	1900	850
Y3514020	291	38.770	15	212	79
Y5615030	279	46.648	2	1760	837

J3024010 and J4124420 are of the same order of magnitude, whereas Y3514020 in the group of three catchments in the south-east is located at a much lower altitude in comparison with V6035010 and Y5615030 in that group. A1522020 is located at a high altitude in the north-east. K0753210, and its sub-catchment K0744010, in the south-central region, have mean altitudes of 755 and 592 m, respectively. Among the three catchments in the central region, H5723011 and H3613020 are located at nearly the same altitude, whereas H2001020 is at a higher altitude.

Seven years of concurrent daily rainfall, evaporation and discharge data for each of the 12 catchments are used. These data were generously provided by Météo France and the Direction de l'eau, through Dr. Vazken Andréassian, of Cemagref, Paris, for application in the MOPEX (model parameter estimation experiment) research project and made available to the present authors for their contribution to the 2004 MOPEX Workshop held in Paris (Andréassian et al., 2006). Summary characteristics of these hydrological data are given in Tables 3a and 3b.

Climatologically, the three catchments in the north-west are in the humid seaboard climatic zone, the three in the south-east are in the Mediterranean zone, the one in the north-east is in the semi-continental zone, and the remaining five are characterized as being in an intermediate climatic zone. Of the 12 catchments, A1522020 is the wettest and H3613020 is the driest. The group of three south-eastern catchments have considerable evaporation, with evaporation exceeding rainfall for almost 80% of the time. From Table 3a, it may be observed that the group of

Table 3a Characteristics of hydrological data (period 1/8/1995–31/7/2002)

Station code no.	Discharge Q (mm/day)				Rainfall R (mm/day)				Evaporation E (mm/day)				% days $R > E$
	Q_{max}	Q_{min}	Q_{mean}	SD_Q	R_{max}	R_{min}	R_{mean}	SD_R	E_{max}	E_{min}	E_{mean}	SD_E	
J2034010	11.8	0.10	0.89	0.90	45.8	0.0	2.63	4.56	3.7	0.50	1.94	1.11	37.3
J3024010	14.4	0.30	1.45	1.28	49.9	0.0	2.78	4.66	3.6	0.34	1.88	1.17	38.7
J4124420	8.94	0.08	1.43	1.43	55.5	0.0	3.38	5.96	4.4	0.48	1.97	1.20	38.1
A1522020	24.4	0.14	2.04	2.63	84.8	0.0	4.56	8.61	4.26	0.28	2.01	1.41	40.0
H5723011	16.0	0.02	0.66	1.34	47.3	0.0	2.20	4.50	4.30	0.34	2.06	1.39	30.7
H3613020	3.0	0.00	0.34	0.25	46.5	0.0	2.21	4.41	4.26	0.28	2.01	1.38	31.6
H2001020	33.1	0.17	2.36	2.35	59.4	0.0	3.56	6.48	4.26	0.28	2.03	1.34	37.0
K0744010	16.5	0.00	1.02	1.20	73.1	0.0	2.68	5.16	4.34	0.24	1.99	1.35	34.1
K0753210	13.7	0.07	1.21	1.25	77.6	0.0	2.77	5.35	4.34	0.24	1.99	1.35	34.6
V6035010	14.35	0.00	0.83	1.48	102.7	0.0	2.90	7.89	6.18	0.62	2.96	1.89	20.1
Y3514020	11.92	0.07	0.71	1.04	65.4	0.0	2.32	7.19	6.48	0.80	3.18	1.90	15.4
Y5615030	62.28	0.06	1.58	3.34	129.3	0.0	3.17	9.96	5.70	1.18	3.07	1.51	17.1

Subscripts 'max', 'min' and 'mean' indicate maximum, minimum and simple average of Q (discharge), R (Rainfall) and E (Evaporation) data in the available data series, and SD indicates respective standard deviations.

Table 3b Characteristics of hydrological data (period 1/8/1995–31/7/2002)

Station code no.	$\frac{E_{mean}}{R_{mean}}$	$R_{mean} - E_{mean}$ (mm/day)
J2034010	0.74	0.69
J3024010	0.68	0.90
J4124420	0.58	1.41
A1522020	0.44	2.55
H5723011	0.94	0.14
H3613020	0.91	0.20
H2001020	0.57	1.53
K0744010	0.74	0.69
K0753210	0.72	0.78
V6035010	1.02	−0.06
Y3514020	1.37	−0.86
Y5615030	0.97	0.10

three catchments in the north-west, i.e. J2034010, J3024010 and J4124420 have very little variability in their hydrometeorological data values.

Hydrometeorological regions and regional homogeneity

From Table 3b, it is seen that the ratio of mean values of evaporation to rainfall (E_{mean}/R_{mean}) is considerably greater than unity (at 1.37) for the Y3514020 catchment in the south-east, very close to unity for the two neighbouring catchments, i.e. V6035010 and Y5615030, and less than unity for remaining nine catchments. These statistics, while indicating higher evaporation levels in the three catchments in the south-east of the country, also suggest that in the context of regional homogeneity, the catchment Y3514020 might perhaps be justifiably excluded from the regional analysis. Topographically, Y3514020 is located at the lowest altitude amongst the three catchments in the south-east. For A1522020, the value of the difference $R_{mean}-E_{mean}$ is the highest (2.55 mm/day), indicating that the rainfall in this catchment is greatly in excess of the evaporation in comparison with the other 11 catchments. The climatological and topographical characteristics indicate that the three catchments in the north-west of France, namely J2034010, J3024010 and J4124420, may be considered to constitute a hydrometeorologically homogeneous sub-group. K0753210, and its sub-catchment, K0744010, which are both located at similar altitudes, also display uniformity in their hydrometeorological characteristics and hence may be considered a homogeneous sub-group.

Although it is recognised that the catchments A1522020 and Y3514020 in the group of 12 catchments may be considered as outliers, the data from all 12 catchments are used in this heuristic study. In a separate exercise to further investigate the effectiveness of the methods tested, the more homogeneous sub-group of the three north-western catchments, namely J2034010, J3024010 and J4124420, is also considered. Clearly, while it would be desirable from the perspective of drawing a generalized conclusion on the performance of the regional methods to include more catchments in the sample, the use of 12 catchments is deemed sufficient for the purpose of the present exercise.

Methodology

For assessing the relative performance of the methods for flow estimation in ungauged catchments, each of the group of the 12 gauged catchments is used initially in turn as if it was ungauged but having rainfall and evaporation data. Subsequently, after simulation of its flows by application of the regionalisation method and calibration of the model, it is recognised as being gauged (having measured flow data) for the sole purpose of evaluating the efficiency of the procedure of flow simulation in that catchment. A catchment, when thus considered, is called 'pseudo-ungauged'. The ultimate step of actually predicting the discharge in an ungauged (e.g. 13th) catchment within the region using the finally selected regionalisation–model couple, is not undertaken in this study.

Firstly, a naïve no-model regional method, to be used as a base-line for evaluating the efficiency of the simulation methods on each pseudo-ungauged catchment, is adopted. For this purpose, the no-model simulated discharge depths (i.e. the discharge expressed as an equivalent depth over the catchment area in the given time step),of each pseudo-ungauged catchment is considered to be the average of the synchronous discharge depths of the whole set of gauged catchments (excluding its own measured discharge as if that was unknown). The discharge depths (in mm) are used instead of discharge (in $m^3 s^{-1}$) in order to offset the scale effect of catchment size. Clearly, no rainfall–runoff modelling exercise is involved in estimating this no-model discharge. For a large region such as that encompassing the 12 catchments located throughout France, the averaging of discharge depths is rather a crude approximation to that of a pseudo-ungauged catchment in that region due to the wide variability of hydrometeorological and physiographical conditions of the catchments. The regionalisation procedures proposed in this study are developed for application to hydrometeorologically homogeneous regions and are applied to the two groups, consisting of the whole set of 12 and a sub-group of only three catchments, for the purpose of demonstrating the methodology.

For regionalisation of data to simulate discharges for the pseudo-ungauged catchments to enable calibration of the six rainfall–runoff models, the three methods described in the following sub-sections are applied and their simulation efficiencies compared with those of the no-model case. With the exception of the pooling method, for which a regionally calibrated model based on a group of gauged catchments is applied with the rainfall and evaporation data of the pseudo-ungauged catchment (and ultimately for a truly ungauged catchment) for which the simulation is required, these methods do not involve the transfer of parameter values of any model from a gauged catchment to the pseudo-ungauged one, and no attempt is made to link the parameter values of the rainfall–runoff models to physical catchment descriptors.

Regional averaging of discharge for model calibration

In this method, the regional average discharge series, as computed for the no-model method, is used for rainfall–runoff simulation in each pseudo-ungauged catchment by calibration of the chosen hydrological models. Thus for a pseudo-ungauged catchment in a group of N catchments, this discharge series is taken as the average of the discharge depths of the other $N-1$ gauged catchments. The rainfall and the evaporation data series used for model calibration are those of the pseudo-ungauged catchment. The models are calibrated using six out of the seven years of record, the first year being considered as the warm-up period. The method is initially applied to all 12 catchments in the group and subsequently to the sub-group of the three north-western catchments which are deemed to be strongly regionally homogenous.

Regional pooling of data for model calibration

In this method, the observed hydrological data series are combined by putting the m years of data of each of the gauged catchments ($N-1$ in a group of N) in series, end to end. The rainfall–runoff models are then calibrated by maximising the R^2 value over the calibration period using the pooled input of the $N-1$ gauged catchments in the region as inputs to the model. Finally, using the regionally calibrated model with the rainfall and evaporation data of the pseudo-ungauged catchment as its inputs, the discharge series for that pseudo-ungauged catchment is simulated. It is assumed in this procedure that the concurrent rainfall and discharge data from all catchments in a hydrometeorologically homogeneous region are expected to produce similar response characteristics representative of the whole region. An appropriate hydrological model fitted to the long data series obtained by pooling is therefore considered to be applicable to any ungauged catchment in the region. Clearly there will always be discontinuities in the pooled data series made up of data from catchments having different hydrometeorological and physiographical conditions when placed end to end. In a hydrometeorologically homogeneous region, the effect of such discontinuities would be less. In order to offset such effects, a warm-up period of one year is considered at the beginning of the data series from each catchment used for pooling. Thus, in calibrating any given model, while the simulated discharge values are computed using the complete pooled length of data, the objective function is evaluated and the efficiency values obtained excluding the warm-up periods.

Transposition of nearest neighbour discharge data for model calibration

This nearest neighbour approach is appropriate when very few catchments in a homogeneous region in the neighbourhood of a pseudo-ungauged catchment are gauged. The flow data series of the nearest gauged catchment, measured in volume of flow per unit time, i.e. $m^3 s^{-1}$, are scaled up or down in the proportion of catchment areas depending on whether the pseudo-ungauged catchment is larger or smaller in area than the gauged catchment considered. In effect, this means that the discharge depths are averaged. Taking the rainfall and the evaporation data series as those observed for the pseudo-ungauged catchment, and the flow data series as that obtained by scaling the data series of the nearest-neighbour gauged catchment, the hydrological models are calibrated (for the last six years in this study, taking the first year as the warm-up period). This method is a special case of regional averaging for estimation of discharge where data from only one gauged catchment are used. In this study, the transposition method is applied only to the sub-group of the three north-western catchments.

Validation of the simulation procedures

In each of the three cases above, the corresponding validation efficiency for each pseudo-ungauged catchment is expressed in terms of the Nash-Sutcliffe R^2 value. This is obtained by taking the last 6 of the 7 years of data for the pseudo-ungauged catchment as the validation period (taking the first year as the warm-up period) based on the mean square error between its simulated and measured discharge values. Clearly, in application to a truly ungauged catchment, no such validation is possible.

Models and efficiency criterion

The individual models

The full names of the seven individual substantive models (or variants), used in this study and their acronyms are given in Table 1. One of these is a conceptual model, the rest being black-box. The emphasis in this paper being on the applicability of the regionalisation methods, the descriptions of the models are not included. Suggested references for their detailed descriptions are also given in Table 1. It may be observed from Table 3b that for A1522020, unlike the other catchments, the mean effective rainfall depth (2.55 mm) is substantially higher than that of the mean discharge (2.04 mm). This indicates that, apart from evaporation loss, there is additional loss of water from the system which is unaccounted for in the given discharge data. Such systems are 'apparently non-conservative'. In regional analyses, the flow series estimated for pseudo-ungauged catchments including data from A1522020 (considering it as a gauged catchment) are biased by the latter's apparent non-conservative system behaviour. In transposition of discharge data from a nearest-neighbour catchment, the simulated discharge also displays similar behaviour when used with observed input data for some pseudo-ungauged catchments due to the observed rainfall being greatly in excess of the simulated discharge. In such situations, either the SMAR-NC1 or the SMAR-NC2 variant of the SMAR model is used instead of the rigidly conservative SMARG. With the exception of the naïve NP-SLM and P-SLM, all of the other individual models applied in this study were developed at the Department of Engineering Hydrology at the National University of Ireland, Galway, Ireland. The Galway Flow Modelling and Forecasting System (GFMFS), a software package incorporating a suite of different hydrological models and techniques, also developed in Galway, is used.

The multi-model approach

In the context of regionalisation, irrespective of the degree of homogeneity of a region, the gauged catchments in the region and the associated rainfall–runoff relations i.e. the models, are not equally valuable as sources of information for application to ungauged catchments. Combining these various sources of information in an efficient consensus form may be a useful means of utilizing all relevant and available information. The multi-model approach is an effective means of such combination. In this approach, it is recognised that (i) the ''plurality of models and modelling approaches may be valid for the same catchment and application'' (Sivapalan et al., 2003), (ii) each model has its inherent strengths and weaknesses, (iii) each makes use of different information, processing different forms of knowledge, and (iv) it is possible to use a number of models simultaneously whereby the strengths of individual models are pooled and perceptible weaknesses de-emphasised to produce a consensus output.

For obtaining consensus outputs by combination, the neural network method (NNM), the weighted average method (WAM), and the simple average method (SAM) are used. These techniques are described by Shamseldin et al. (1997) and Shamseldin and O'Connor (1999).

The performance evaluation criterion

The GFMFS has provision for 14 different model performance evaluation criteria. However, for brevity, and despite its known shortcomings (Kachroo and Natale, 1992), only the dimensionless efficiency index R^2 (Nash and Sutcliffe, 1970) is used in this paper for judging the relative performance of the individual models and the techniques. This criterion, which is based on the mean square error that penalises the model much more for large errors than for small errors, irrespective of the magnitude of the variable at which such errors occur, is widely used by catchment modellers. Since the objective of the study is to assess relative performance of the regionalisation methods without emphasis on any particular application of flow analysis e.g. low flow study, flood forecasting, water resource assessment etc., use of the R^2 index only for performance evaluation is considered satisfactory for the purpose of the study. Whereas $R^2 = 100\%$ would denote an ideal or perfect fit, it is generally agreed that for simulation of flows in gauged catchments, $R^2 > 90\%$ is indicative of a very good model fit, while that in the range of 80–90% is a fairly good fit, with a range of 60–80% considered as being unsatisfactory. In many studies involving flow simulation in ungauged catchments, mean and median values of the R^2 index are reported (Sefton and Howarth, 1998; Merz and Blöschl, 2004). Sefton and Howarth (1998) used R^2 value of 50% as the threshold for rejection of a model or procedure in the context of ungauged catchments.

Results

Results of regional averaging of discharge

Considering each of the 12 catchments as being ungauged in turn, i.e. as a pseudo-ungauged catchment, each corresponding regionally averaged discharge series is generated. The results obtained by calibrating the rainfall–runoff models both individually and in combination for all 12 catchments are given in Table 4. The values of the R^2 index for the no-model case are also given. Although the unrealistic negative no-model R^2 value for the catchment Y3514020 suggests possible heterogeneity, as indicated also in Section ''Hydrometeorological regions and regional homogeneity'', the greater possibility of heterogeneity suggested by the much larger negative value for H36113020 is not apparent from the study described in that section, except for the fact that the catchment is the driest in the group. Although a very poor model performance, compared to that of the other catchments in the region, does not necessarily indicate that a catchment is an outlier, it does make it a suspect.

Amongst the individual models, performances of the P-SLM and the NP-SLM are generally inferior to those of all other models in calibration. The performance of the conceptual SMAR model in calibration is close to that of the LPM forms for all except the Y3514020 and Y5615030 catchments, the performance of the SMAR model for these last two catchments being considerably lower than that of the LPM model forms. However, while the non-linear ANN model is best in nine out of the 12 catchments in calibration, it fails to perform well for the three catchments in the south

Table 4 R^2 (%) for pseudo-ungauged catchments: individual models, and the NNM combination using regionally averaged values (all 12 catchments)

	J2034010	J3024010	J4124420	A1522020	H5723011	H3613020	H2001020	K0744010	K0753210	V6035010	Y3514020	Y5615030
No-model	*23.6*	*34.3*	*39.6*	*15.6*	*7.8*	*−2766*	*10.0*	*37.2*	*44.0*	*15.6*	*−110*	*11.9*
Individual models												
NPSLM												
Calibration	44.7	36.9	42.0	32.6	17.8	19.6	40.2	15.7	14.7	3.7	−11.6	−21.8
Validation	40.8	41.8	48.2	12.7	6.1	−1704	0.1	34.2	35.7	41.1	30.5	21.7
PSLM												
Calibration	37.8	33.7	32.2	23.4	17.8	16.3	36.9	14.7	14.5	3.4	−17.6	−26.4
Validation	51.4	29.9	30.5	−0.7	10.2	−1219	−6.0	32.7	313	42.4	42.2	12.1
NPLPM												
Calibration	60.5	54.3	56.5	54.2	55.3	57.0	63.4	62.0	61.8	54.5	45.0	46.2
Validation	53.5	53.0	56.9	17.2	14.7	−2080	11.8	53.1	56.7	26.3	−24.6	12.8
PLPM												
Calibration	60.4	56.7	58.5	55.4	58.1	59.1	63.3	61.9	61.8	*58.0*	*48.7*	*49.9*
Validation	44.6	43.9	45.0	14.5	12.6	−2060	8.5	44.3	46.2	24.7	−21.2	13.4
LVGFM												
Calibration	61.3	57.4	56.4	56.1	62.8	63.3	66.4	59.5	52.7	46.6	21.1	16.2
Validation	24.2	64.01	64.0	20.7	23.5	−2284	16.7	62.2	60.3	47.5	31.5	21.6
ANN												
Calibration	*62.7*	*58.31*	*60.4*	*56.5*	*62.9*	*64.1*	*69.2*	*67.2*	*64.2*	57.9	7.2	48.6
Validation	59.7	48.82	49.4	17.0	19.5	−2162	18.2	58.3	60.6	25.1	−16.8	12.3
SMAR												
Calibration	60.0	51.33	55.8	50.8	57.6	58.7	68.9	62.0	61.9	42.5	6.2	10.8
Validation	63.8	65.68	66.2	21.7	26.9	−1722	17.9	65.0	67.8	42.7	26.4	15.1
Model combinations												
(All 7 models)												
Calibration	*69.0*	58.22	64.2	*64.1*	*68.2*	*73.3*	72.2	*71.5*	*72.9*	*66.7*	49.9	53.6
Validation	58.7	57.74	58.1	18.6	21.8	−2351	18.0	59.7	58.1	33.5	−35.6	16.8
(Best 6)												
Calibration	68.5	61.60	*64.3*	62.7	66.3	70.2	71.2	69.6	72.1	64.0	*56.0*	60.5
Validation	58.5	55.11	57.6	20.0	23.9	−2293	17.8	60.2	59.8	35.6	−40.8	15.0
(Best 5)												
Calibration	67.4	*62.39*	63.2	62.7	66.9	68.0	*73.2*	69.9	70.3	66.7	54.9	*64.5*
Validation	58.2	54.83	57.5	18.9	24.1	−2260	17.8	59.3	61.5	35.0	−39.4	16.3
(Best 4)												
Calibration	67.0	60.40	63.0	57.8	65.9	67.1	71.1	70.9	71.8	63.7	53.2	54.1
Validation	59.3	55.54	57.8	20.0	25.4	−2238	18.0	59.2	59.7	35.8	−37.4	15.7

(Best 3)												
Calibration	64.0	60.39	63.1	62.9	66.0	67.2	69.7	69.6	69.8	63.6	50.4	54.6
Validation	61.6	55.57	58.5	12.6	25.3	−2237	18.1	60.8	62.5	36.0	−28.4	6.2
(Best 2)												
Calibration	61.6	52.94	63.2	51.3	59.1	68.1	69.3	66.8	70.8	46.3	49.5	24.2
Validation	62.2	57.08	58.9	19.4	23.6	−2268	18.4	58.1	60.3	39.4	−18.1	14.8
% Improvement in validation by the best calibrated model	*149*	*60*	*45*	*19*	*179*	*15*	*78*	*60*	*32*	*115*	*63*	*37*
% Improvement in validation by the SMAR model	*170*	*91*	*67*	*39*	*245*	*38*	*79*	*75*	*54*	*174*	*124*	*27*

east, namely Y3514020, Y5615030 and V6035010. In validation, the performance of the SMAR model is generally best but, for the three south-eastern catchments, the performances of all models are poor.

When combination techniques are applied to the outputs of the individual substantive models, the results show that the performance of the NNM is generally the best, followed by the WAM. SAM, being a special case of WAM with equal weights, generally performs worse than the other two combination techniques. For brevity, only the results of the NNM are provided in Table 4. Although some slight improvement of performance in calibration is achieved by the combination of model outputs, the performance in validation generally remains lower than that achieved by the individual SMAR model, and all performances are still considered unsatisfactory.

The Table 4 also shows that the improvement in performance attained by the SMAR model in validation over that of the no-model case is generally considerably higher than that of the best performing combination technique. Overall, when using the regionally averaged data series, the SMAR model simulates the flow in the pseudo-ungauged catchments better than the other individual models and combination techniques.

Table 5 shows the results of the regional averaging method when applied to the small homogeneous region composed of the sub-group of the three north-western catchments only. Amongst the individual models, the ANN model performs best in all three catchments in calibration and in two out of the three in validation. For the third, namely J2034010, the NP-LPM model performs best in validation. The R^2 values in calibration, obtained by the best combination technique, range from 90% to 94% but the values in validation remain below 80% with an unacceptably low value of 25% for J2034010. The performances of the best calibrated structures of the combination techniques in calibration, for all three catchments, are higher than the corresponding performances of the best individual models. However, in validation, the combination technique outperforms the best individual model in only one case, namely J3024010. In the case of J2034010, for which the no-model efficiency is very low, no individual model or model combination produces good efficiency in validation.

Results of regional pooling of data

The second method, using pooled data, is applied to the whole sample of 12 catchments and also to the sub-group of the three north-western catchments. A warm-up period of one year is considered at the beginning of the data series from each catchment. This is done in order to take into account the effect of discontinuities caused by joining of data from catchments having hydrometeorological and physiographical heterogeneity. The results are shown in Tables 6 and 7, respectively. In calibration, for the whole group of 12 catchments, the LVGFM performs as the best individual model for eight of the catchments, the ANN model being the best for the remaining four. The individual model performance in calibration improves by a few percent by the application of the combination technique. In validation, the performance of the LVGFM is best for six catchments, the SMAR for five and the ANN for one. In validation also,

Table 5 R^2 (%) for pseudo-ungauged catchments: individual models, and the NNM combination using regionally averaged values (three catchments in NW France)

	J2034010	J3024010	J4124420
No-model	*5.1*	*78.7*	*74.8*
Individual models			
NPSLM			
Calibration	48.2	61.9	56.5
Validation	47.7	65.6	53.6
PSLM			
Calibration	43.6	59.3	46.0
Validation	39.4	53.0	34.0
NPLPM			
Calibration	68.5	85.1	76.1
Validation	47.7	65.6	53.6
PLPM			
Calibration	60.3	75.1	65.2
Validation	24.9	60.9	49.8
LVGFM			
Calibration	85.6	90.5	84.2
Validation	24.8	76.4	76.6
ANN			
Calibration	*86.4*	*91.4*	*85.0*
Validation	29.9	77.3	79.2
SMAR			
Calibration	80.4	89.1	83.9
Validation	42.0	76.5	77.7
Model combinations			
(All 7 models)			
Calibration	88.3	*94.3*	90.3
Validation	25.3	79.1	76.3
(Best 6)			
Calibration	88.1	93.8	90.1
Validation	26.6	78.0	76.1
(Best 5)			
Calibration	*89.8*	93.6	90.6
Validation	25.1	77.5	75.8
(Best 4)			
Calibration	89.8	93.6	*90.6*
Validation	24.6	77.4	76.9
(Best 3)			
Calibration	87.5	94.0	90.1
Validation	29.2	78.6	77.0
(Best 2)			
Calibration	87.9	93.8	89.9
Validation	26.6	78.9	77.4
% Improvement in validation by the best calibrated model	*392*	*0.5*	*3*
% Improvement in validation by the SMAR model	*724*	*−3*	*4*

the performance of the best combination technique is better than that of the corresponding best performing individual model in 10 out of the 12 catchments.

From Tables 4 and 6, it is also observed that, in comparison with the regional averaging method, the regional pooling method produces better R^2 values in both calibration and validation.

For the sub-group of three catchments, amongst the individual models, the ANN model is best in calibration in two catchments and the LVGFM is best in the third. However, in validation, the performance of the SMAR model is higher than that of the other individual models in two catchments, while in the third the ANN model is best. After application of the combination techniques, although the ANN performance in validation in one catchment is increased, the performance of the combination techniques in other two catchments does not increase over the corresponding values attained by the individual SMAR model. In the case of J2034010, for which the no-model efficiency is very low, the method of pooled data produces very high efficiency

Table 6 R^2 (%) for pseudo-ungauged catchments: individual models, and the NNM combination using pooled data (all 12 catchments)

	J2034010	J3024010	J4124420	A1522020	H5723011	H3613020	H2001020	K0744010	K0753210	V6035010	Y3514020	Y5615030
No-model	*23.6*	*34.3*	*39.6*	*15.6*	*7.8*	*−2766*	*10.0*	*37.2*	*44.0*	*15.6*	*−110*	*11.9*
Individual models												
NPSLM												
Calibration	51.2	50.9	51.7	47.6	52.2	52.7	51.9	51.5	51.5	51.5	51.2	47.9
Validation	35.4	50.1	40.8	63.6	28.6	−8.6	46.1	35.6	37.3	44.0	39.3	61.7
PSLM												
Calibration	49.5	49.4	50.0	44.6	50.3	50.3	49.9	49.7	49.5	49.3	49.5	46.3
Validation	29.0	42.4	27.6	65.3	29.1	−9.2	43.3	32.7	33.3	40.5	31.7	62.6
NPLPM												
Calibration	62.0	61.8	62.0	59.2	62.8	64.1	62.1	61.7	61.6	62.2	62.1	61.3
Validation	44.5	58.3	53.9	73.4	41.7	6.2	58.6	60.4	60.6	52.1	44.8	66.7
PLPM												
Calibration	58.8	58.7	59.0	53.9	59.6	59.9	58.6	58.4	58.5	59.0	61.1	57.5
Validation	38.9	52.2	42.1	72.6	41.3	−18.6	52.7	55.1	54.2	48.6	38.5	63.7
LVGFM												
Calibration	75.0	*73.6*	*75.2*	*74.4*	75.0	74.5	*75.1*	*74.8*	*75.0*	75.0	*74.2*	*73.0*
Validation	85.9	73.3	67.5	66.8	42.7	69.9	59.0	57.6	54.5	76.3	86.3	83.7
ANN												
Calibration	*75.8*	73.0	74.1	72.6	*75.1*	*75.3*	74.5	73.0	70.7	*75.5*	72.7	73.0
Validation	78.5	77.6	56.6	70.0	51.1	48.8	65.3	64.7	55.8	70.6	81.1	73.9
SMAR												
Calibration	74.5	73.5	75.1	73.7	74.4	74.3	74.7	74.5	74.7	74.4	74.1	72.6
Validation	67.7	76.5	57.1	74.4	61.4	42.7	72.7	71.6	68.3	73.1	82.3	74.0
Model combinations												
(All 7 models)												
Calibration	77.0	77.4	*78.6*	77.1	78.4	77.7	*79.4*	77.4	*77.9*	77.2	*77.0*	75.4
Validation	80.5	75.3	65.9	73.4	52.9	50.1	66.0	69.2	67.1	75.7	86.9	80.9
(Best 6)												
Calibration	*78.6*	77.4	77.7	76.5	*78.5*	*77.9*	79.2	*78.3*	77.8	*77.3*	76.9	75.4
Validation	76.1	75.8	64.0	74.9	52.5	51.3	66.2	69.8	67.3	76.6	87.1	80.5
(Best 5)												
Calibration	76.9	*77.5*	77.4	*77.3*	78.0	77.4	78.6	77.0	77.6	76.8	76.4	*76.5*
Validation	80.9	76.4	62.5	76.5	53.7	57.6	65.9	69.5	66.5	76.1	87.0	84.9
(Best 4)												
Calibration	76.7	76.8	78.2	76.8	78.0	77.3	78.8	76.8	77.6	76.8	76.5	74.9
Validation	81.0	76.3	64.1	75.6	53.9	53.6	65.9	69.3	67.3	76.1	88.0	80.5

(continued on next page)

Table 6 *(continued)*

	J2034010	J3024010	J4124420	A1522020	H5723011	H3613020	H2001020	K0744010	K0753210	V6035010	Y3514020	Y5615030
(Best 3)												
Calibration	76.0	76.6	77.2	76.1	77.8	76.6	78.8	77.2	76.9	76.8	76.4	74.7
Validation	83.6	77.9	64.6	74.9	53.7	60.0	66.5	67.7	63.7	77.0	87.2	79.0
(Best 2)												
Calibration	75.9	76.6	77.0	76.6	77.8	76.4	78.7	76.7	76.8	76.4	76.0	75.1
Validation	83.4	78.2	62.6	74.5	55.7	57.5	66.2	66.9	63.5	75.5	87.2	80.3
% Improvement in validation by the best calibrated model	*223*	*123*	*66*	*389*	*569*	*102*	*558*	*88*	*53*	*390*	*179*	*614*
% Improvement in validation by the SMAR model	*187*	*123*	*44*	*376*	*684*	*102*	*625*	*93*	*55*	*368*	*175*	*523*

in validation in comparison with the regional averaging method. From Tables 5 and 7 it is observed that, in comparison with the regional averaging method, improvement in validation using pooled data is very high in J2034010. However, in the other two catchments, there is generally no improvement of performance of the regional pooling method, either in calibration or validation, over that of the regional averaging method.

Results of transposition of data

The results of application of the method of transposition to the sub-group of three north-western catchments are given in Table 8. The performance of the individual models only is investigated in this case. It may be observed from the table that in the case of the catchments J4124420 and J3024010, having areas of the same order of magnitude, i.e. 32.1 and 43 km^2, respectively, the transposition of data method performed well. The R^2 values in validation for these two catchments are even higher than those achieved by regional averaging and regional pooling. However, due to the considerable difference in areas of the two catchments J2034010 and J4124420, when selected for transposition, i.e. 125 km^2 for J2034010 and 32.1 km^2 for J4124420, the performance is generally low, i.e. the method performed poorly.

Discussion

Assessment of regional homogeneity from the study of physiographic and hydrological conditions and a preliminary analysis of the available data are very important for the subsequent model calibration and estimation of flow in an ungauged catchment using regionalisation approaches. The performance of the regionalisation method, both in calibration and validation, reduces considerably as more and more catchments, which are not consistent as regards their rainfall—runoff behaviour with the original members in the region, are included in the regional analysis.

Amongst the three methods of regionalisation involving model calibration, the method of pooling generally provides good estimates of flow in the pseudo-ungauged catchments. Such improvement in performance is attributed to the use, without any averaging, of catchment-specific data series from each member in the group and hence without distortion of the response characteristics of the individual catchments. Thus, the models calibrated to the data series generated by pooling can satisfactorily simulate the response from a pseudo-ungauged catchment in the region. When regional pooling of data is used instead of regional averaging, the median and the mean value of R^2 index (as a %) in the case of all 12 catchments increases from 67.5 and 67.2 to 77.9 and 77.9, respectively, in calibration, and from 27.7 and −166.2 to 72.95 and 70.83, respectively, in validation. The negative mean value can be attributed to the large and unrealistic negative R^2 values in validation for H3613020 and Y3514020. Hence it is seen that regional pooling performs better than regional averaging. However, for a region having a large number of gauged catchments of different sizes, such as the group of 12 catchments, a high performance level by any model using this pooling method is not achieved due to the scale-effect caused by pooling of

Table 7 R^2 (%) for pseudo-ungauged catchments: individual models, and the NNM combination using pooled data (for the sub-group of three catchments in NW France)

	J2034010	J3024010	J4124420
No-model	*5.1*	*78.7*	*74.8*
Individual models			
NPSLM			
Calibration	51.72	57.61	50.60
Validation	46.65	52.22	51.38
PSLM			
Calibration	48.07	56.40	43.83
Validation	57.07	49.93	41.01
NPLPM			
Calibration	68.50	77.17	65.45
Validation	60.67	62.16	67.63
PLPM			
Calibration	62.77	70.32	59.95
Validation	54.68	58.65	55.31
LVGFM			
Calibration	83.29	90.65	*88.29*
Validation	85.39	70.10	60.30
ANN			
Calibration	*85.65*	*91.74*	86.17
Validation	87.37	72.79	71.76
SMAR			
Calibration	83.02	89.95	82.57
Validation	83.67	76.02	80.89
Model combinations			
(All 7 models)			
Calibration	86.90	*93.08*	88.73
Validation	87.45	72.23	64.88
(Best 6)			
Calibration	*87.16*	93.06	88.48
Validation	87.72	72.22	68.21
(Best 5)			
Calibration	86.93	92.98	88.39
Validation	88.21	72.52	66.38
(Best 4)			
Calibration	86.79	92.96	88.40
Validation	88.44	72.39	64.29
(Best 3)			
Calibration	86.70	92.88	*89.05*
Validation	88.52	72.17	58.84
(Best 2)			
Calibration	86.57	92.25	87.25
Validation	87.52	72.89	70.98
% Improvement in validation by the best calibrated model	*1613*	*−8*	*−21*
% Improvement in validation by the SMAR model	*1534*	*−3*	*8*

data from catchments having different areas, and the inherent diversities in the larger region.

Amongst all individual models and the combination techniques, no single model performs consistently best. However, with pooled data, the SMAR model generally performs best in validation in the more homogeneous sub-group of the three north-western catchments. This reflects the ability of the model to perform well in an ungauged catchment within such a homogeneous region. This simple conceptual model, coupled with the method of regional pooling of data, is therefore deemed to be the best for the pseudo-ungauged catchments in a homogeneous region and, by interpolation, to a truly ungauged (as distinct from pseudo-ungauged) catchment within that region.

In the case of the availability of only a few gauged catchments in the region, transposition of data may be used provided the catchments are similar in hydrometeorological and physiographic characteristics, and their areas are of

Table 8 R^2 (%) for pseudo-ungauged catchments with individual models using transposition (for the subgroup of three catchments in NW France)

Index catchment	Ungauged catchment		NP-SLM	P-SLM	NP-LPM	P-LPM	SMAR	LVGFM	ANN
J3024010 (43 km^2)	J2034010 (125 km^2)	Calibration	56.4	33.0	71.8	58.0	76.2	78.4	*83.0*
		Validation	5.6	33.3	22.3	18.1	43.1	27.9	32.4
	J4124420 (32.1 km^2)	calibration	65.0	36.6	78.4	60.8	76. 8	80.0	*82.4*
		Validation	61.2	33.7	82.3	52.1	82.5	78.8	82.2
J4124420 (32.1 km^2)	J2034010 (125 km^2)	Calibration	51.0	46.3	72.9	60.6	77.2	79.8	*83.8*
		Validation	−0.7	30.4	9.5	22.8	32.9	12.7	11.5
	J3024010 (43 km^2)	Calibration	59.6	55.0	81.8	67.1	84.1	86.2	*87.2*
		Validation	60.1	58.4	72.2	66.0	76.8	76.7	80.5
J2034010 (125 km^2)	J4124420 (32.1 km^2)	Calibration	55.1	40.0	69.4	61.2	75.3	*75.6*	75.4
		Validation	37.3	16. 4	55.1	32.3	56.7	55.7	57.3
	J3024010 (43 km^2)	Calibration	58.6	48.2	75.1	66.3	83.0	83.8	*85.8*
		Validation	37.3	20.5	51.2	33.6	54.1	55.1	55.6

the same order of magnitude. Transposition of data from a large gauged catchment to a small pseudo-ungauged catchment or vice versa does not produce good results.

Although warm-up periods of one year were used in the present study, for consistency with earlier results for all the catchments presented at the MOPEX Paris Workshop of 2004, the authors are of the opinion that for the twelve catchments considered in this study much shorter periods, say one to two months, as determined by trial and error, would be quite adequate. This would permit the use of more of the data for calibration and validation.

Conclusions

Hydrometeorological and topographic characteristics of catchments in a region are of assistance in the preliminary screening of catchments for homogeneity. Assessing homogeneity is very important for estimation of flow in ungauged catchments by regional analysis. Amongst the regionalisation methods tested, the method of pooling generally works best. In comparison with regional averaging, the regional pooling of data consistently produces better estimates of flow in the pseudo-ungauged catchments in a region having a large number of gauged catchments. For a homogeneous region having a small number of gauged catchments, the results obtained by regional pooling of data is comparable to that obtained by regional averaging. In the case of availability of only a few gauged catchments in the region, transposition of data from the gauged to the pseudo-ungauged catchment may be used provided the catchments are similar in their hydrometeorological and physiographic characteristics and their areas do not differ greatly. The method of combination in the multi-model approach improves the performance of individual models considerably in calibration but not in validation. Amongst the models and the techniques tested, the conceptual SMAR model, coupled with the method of regional pooling of data, appears to be the best to simulate flow in an ungauged catchment within the region.

Based on the comparison of the simulation efficiency values of the regionalisation methods and rainfall—runoff models tested on each of the pseudo-ungauged catchments (12 in all), the best method—model couple for the region can be identified. As the final step in the procedure, the selected method—model couple can be subsequently applied to a truly ungauged (13th) catchment in the region, without validation of the simulated discharge. In the cases of regional averaging and transposition, this is achieved by regionalisation of flows involving all 12 (instead of just 11) gauged catchments and calibration of the selected rainfall—runoff model using the rainfall and evaporation data of that ungauged (13th) catchment. In the case of regional pooling, the regionally calibrated model is used to simulate the discharges, without recalibration, with the rainfall and evaporation data of the ungauged (13th) catchment as inputs. This final step was not undertaken in this study which had as its primary concern the development of the methodology for the simulation of discharge in an ungauged catchment.

Directions for further work

Subsequent to this heuristic study, the following aspects are currently under investigation for the purpose of refining the procedures described above.

(i) Based on a more elaborate regional homogeneity study, some catchments considered as being outliers may justifiably be excluded in subsequent analyses.
(ii) As refined no-model discharge estimates, the areally weighted average can be considered instead of simple average.
(iii) Although the method of transposition of data was considered here for a region having a small number of gauged catchments, it would be worth investigating if the transposition (nearest-neighbour) principle applies also to a region having a large number of gauged catchments.
(iv) Clearly, the use of rainfall—runoff models other than those considered in this study is also worth investigating.
(v) As the current version of the GFMFS used in this study is not specifically designed for simulation in ungauged catchments, its modification for this purpose is under consideration.

Acknowledgements

The authors are grateful to the two anonymous reviewers whose detailed comments and constructive suggestions contributed greatly to the improvement and readability of this paper.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connel, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System–Système Hydrologique Européen. 'SHE', 1: history and philosophy of a physically-based distributed modelling system. J. Hydrol. 87, 45—59.

Ahsan, M., O'Connor, K.M., 1994. A simple non-linear rainfall—runoff model with a variable gain factor. J. Hydrol. 155, 151—183.

Andréassian, V., Rojas-Serna, C., Michel, C., Perrin, Ch., Mouelhi, S., Loumagne, C., 2003. Is the regionalization of conceptual rainfall—runoff models impossible? Geographical Research Abstracts. European Geophysical Society 5, 04175.

Andréassian, V., Hall, A., Chahinian, N., Schaake, J., 2006. Large sample basin experiments for hydrological model parameterization: Results of the MOdel Parameter Experiment (MOPEX) Paris (2004) and Foz de Iguaçu (2005) workshops. IAHS Red Book, p. 307.

Aspinall, R., 2004. Modelling land use change with generalized linear models — a multi-model analysis of change between 1860 and 2000 in Gallatin valley, Montana. J. Environ. Manag. 72, 91—103.

Diekkrüger, B., Kirkby, M.J., Schröder, U. (Eds.), 1999. Regionalization in Hydrology. IAHS Publication No. 254, ISSN 0144-7815.

Dunn, S.M., Lilly, A., 2001. Investigating the relationship between a soils classification and the spatial parameters of a conceptual catchment-scale hydrological model. J. Hydrol. 252, 157—173.

Dunn, S.M., McAlister, E., Ferrier, R.C., 1998. Development and application of a distributed catchment-scale hydrological model for the River Ythan, NE Scotland. Hydrol. Process. 12, 401—416.

Flood Estimation Handbook, 1999. Institute of Hydrology, Wallingford, Oxfordshire, UK.

Goswami, M., O'Connor, K.M., 2005. Application of a conceptual rainfall—runoff simulation model to three European catchments characterised by non-conservative system behaviour. In: Proceedings of the International Conference Hydrological Perspectives for Sustainable Development, Roorkee, India, vol. 1, pp. 117—130.

Goswami, M., O'Connor, K.M., Shamseldin A.Y., 2002. Structures & performances of five rainfall—runoff models for continuous river-flow simulation. In: Proc. 1st Biennial Meeting of Int. Env. Modeling and Software Soc., Lugano, Switzerland, vol. 1, pp. 476—481.

Hundecha, Y., Bárdossy, A., 2004. Modeling of the effect of land use changes on the runoff generation of a river catchment through parameter regionalization of a watershed model. J. Hydrol. 292, 281—295.

Kachroo, R.K., 1992. River flow forecasting, Part 5: applications of a conceptual model. J. Hydrol. 133, 141—178.

Kachroo, R.K., Liang, G.C., 1992. River flow forecasting, Part 2: algebraic development of linear modelling techniques. J. Hydrol. 133, 17—40.

Kachroo, R.K., Natale, L., 1992. Non-linear modelling of the rainfall—runoff relation. J. Hydrol. 135, 341—369.

Littlewood, I.G., 2003. Improved unit hydrograph identification for seven Wesh rivers: implications for estimating continuous streamflow at ungauged sites. Hydrol. Sci. J. 48 (5), 743—762.

Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. J. Hydrol. 287, 95—123.

Nash, J.E., 1960. A unit hydrograph study with particular reference to British catchments. Proc. Inst. Civil Eng. 17, 249—282.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, Part 1, A discussion of principles. J. Hydrol. 10, 282—290.

Natural Environment Research Council (NERC), 1975. Flood Studies Report. Nat. Environ. Res. Counc., London, vol. 1 (chapter 6).

See, L., Abrahart, R.J., 2001. Multi-model data-fusion for hydrological forecasting. Comput. Geosci. 27, 987—994.

Sefton, C.E.M., Howarth, S.M., 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. J. Hydol. 211, 1—16.

Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall—runoff modelling. J. Hydrol. 199, 272—294.

Shamseldin, A.Y., O'Connor, K.M., 1999. A real-time combination method for the outputs of different rainfall—runoff models. Hydrol. Sci. J. 44 (6), 895—912.

Shamseldin, A.Y., O'Connor, K.M., Liang, G.C., 1997. Methods for combining the outputs of different rainfall—runoff models. J. Hydrol. 197, 203—229.

Sivapalan, M., Schaake, J., 2003. PUB Science and Implementation Plan. International Association of Hydrological Sciences.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnel, J.J., Mendiondo, E.M., O'Connel, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., Zehe, E., 2003. IAHS Decade on Predictions in Ungauged Catchments (PUB), 2003—2012: Shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48 (6), 857—880.

Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its application to modelling and control. IEEE Trans. Syst. Man. Cybern. 15, 116—132.

Vandewiele, G.L., Elias, A., 1995. Monthly water balance of ungauged catchments obtained by geographic regionalization. J. Hydrol. 170, 277—291.

Xiong, L., Shamseldin, A., O'Connor, K., 2001. A non-linear combination of the forecasts of rainfall—runoff models by the first-order Takagi—Sugeno fuzzy system. J. Hydrol. 245, 196—217.