Calibration, Uncertainty and Regional Analysis of Conceptual Rainfall -Runoff models

Howard S. Wheater[1] , Neil McIntyre[1] and Thorsten Wagener[2]

[1] Department of Civil and Environmental Engineering
Imperial College of Science, Technology and Medicine,
London SW7 2BU, UK

[2]Department of Civil and Environmental Engineering
Pennsylvania State University

INTRODUCTION

The majority of continuous-time rainfall runoff models can be classified as conceptual, as discussed in Chapter 1 (see also Wheater et al., 1993; Wheater, 2002). This type of model represents the hydrological processes that are seemingly important in the system using a simplified, conceptual representation (Figure 1). These models have three notable characteristics: a) their model structure is specified a priori; b) the hydrological properties of the catchments are represented as parameters, which are generally assumed to be constant during each model application; c) (at least some of) the model parameters have no direct, physical meaning and are not directly measurable. Therefore model parameters are usually estimated via calibration, using the fit of the model output time-series to observed data to provide a measure of goodness of fit.

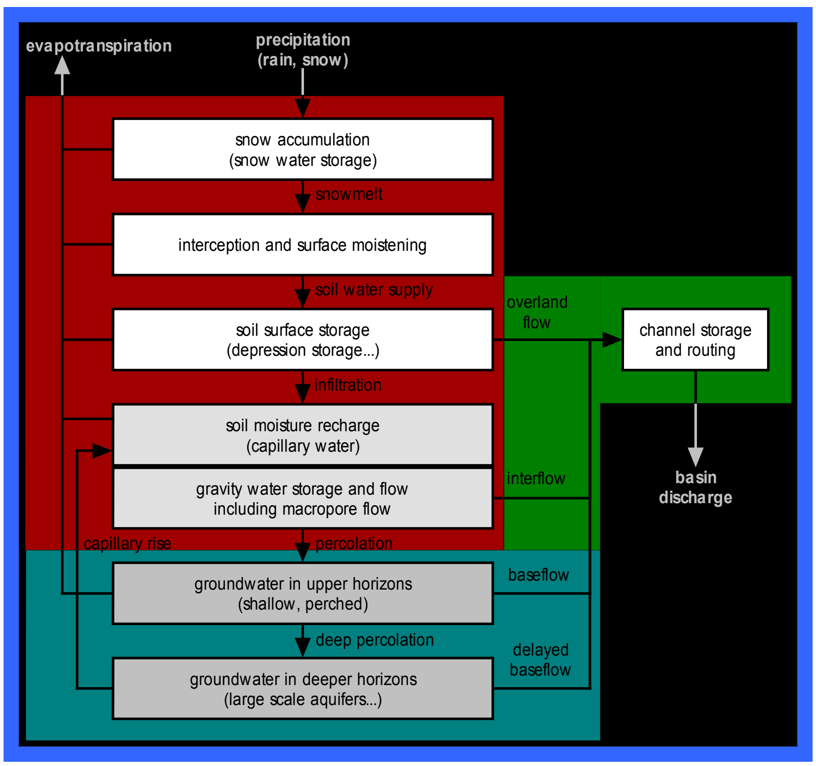

Figure 1 Conceptual model

In this chapter we introduce the issues associated with calibration, and recent developments that allow the associated uncertainty to be specified. Finally the application of models to ungauged catchments (regional analysis) is discussed. For a more extensive treatment of these subjects, the reader is referred to Rainfall-Runoff Modelling in Gauged and Ungauged Catchments by Wagener et al. (2004) and also the AGU Monograph on Calibration of Watershed Models (Duan et al., 2003)).

CALIBRATION ISSUES

Two distinct approaches to calibration can be taken:

a) A manual approach requires the user to adjust parameters interactively in successive model runs. This is a time-consuming and labour-intensive process, dependent on the insight of the modeller (although codified guidance is available for some models (Boyle et al., 2000)). This dependence on the user can be seen as a strength, since it builds on accumulated experience and only intelligent steps through the parameter space will be made. However, it is also a weakness, since the process is subjective, and the parameters derived may be prone to bias. There is also no clear point at which the calibration process can be said to be complete.

In evaluating the quality of model fit, human judgement is used. The eye is a powerful integrator of different signals, and can for example detect that simulation performance over different parts of the hydrograph (e.g. peaks, recessions, low flows) may provide clues to guide the search for optimal values of specific parameters within the model. Alternatively the fit can be measured using one or more objective functions. These commonly aggregate the time-series residuals over the whole calibration period, for example the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), which is a dimensionless form of the sum of squared errors.

b) An automatic approach uses a computer algorithm to search the parameter space, performing multiple trials of the model. This requires that performance is specified by an objective function, and if the model has p parameters, the problem can be posed as the maximisation or minimisation of a (p+1)-dimensional response surface. Classical optimisation methods can be applied, for example gradient-based methods such as the well-known Marquard algorithm and Rosenbrock's rotating coordinate search method (Rosenbrock, 1960), or alternatively, rule-based search methods, such as the Simplex algorithm of Nelder and Mead (1965). Current computing power has allowed more powerful variants to be developed. For example the Shuffled Complex Evolution method, developed at the University of Arizona (SCE-UA) (Duan et al., 1992), combines the Simplex algorithm with a genetic algorithm approach. Multiple Simplexes are propagated through the parameter space, and periodically shuffled to exchange information. The automatic approach has the advantages that the computer does the hard work of exploring the parameter space, rather than the user, and that the procedure is objective; the calibration process would, through thorough searching of the parameter space, ideally be able to define a single well-identified optimum parameter set.

The development of automatic methods has allowed detailed investigation of the issues underling the search for a global optimum set of parameter values in a way that was not possible using a manual approach. Problems in searching the parameter space have long been well-known (e.g. Ibbitt, 1970; Johnston and Pilgrim, 1976). These include:

- Multiple local optima on the objective function surface
- Interdependence of parameters that gives difficulties due to the production of valleys (or ridges) in the objective function
- Insensitive directions in the parameter space, e.g. if a parameter is redundant due to a threshold value
- Search hampered by boundaries in parameter values
- Saddle points, where first derivatives vanish but minima (or maxima) are not reached
- Different scales of parameters, which create difficulties in defining appropriate step lengths in each parameter direction

One important result is non-uniqueness of the identified parameter sets: many combinations of parameter values provide equally good fits to the data (indeed many different model structures may also give similar objective function values (Wheater, 2002)). Beven (1993) defined this problem as 'equifinality', arising from over-parameterisation of models, data limitations and structural faults in the model.

This ambiguity has serious implications for model applications. If parameters cannot be uniquely identified, then they cannot be deterministically linked to catchment characteristics. For example, this in principle precludes the use of models for the prediction of catchment change (where a change in catchment attributes must be represented by an appropriate change in parameter values), or for application to modelling the hydrology of ungauged catchments.

DEVELOPMENTS IN CALIBRATION METHODS

There have been three reactions to this problem of ambiguity (Wagener et al., 2003). These are discussed below, and illustrated using various methods of model performance analysis:

The development and analysis of parsimonious model structures

Firstly there has been a move to simpler models, with fewer parameters, in an attempt to reduce parameter interactions and dependencies and more closely match the complexity of the model to the information content of the data. This reflects the philosophy of metric models, and hence this class of model was termed Hybrid Metric-Conceptual (HMC) by Wheater et al. (1993).

An early example is the Probability-Distributed Model (PDM) of Moore and Clarke (1981). IHACRES (see the workshop paper by Croke and Jakeman) is another. In the PDM model, the spatial variation of soil moisture storage capacity is represented by a prescribed probability distribution. The Pareto distribution is used here. It has the distribution function,

$$\mathrm{F(c)} = 1 - (1 - \frac{\mathrm{c}}{cmax})^{b} \qquad \text{Equation.1}$$

where c/*cmax* is the degree of saturation and the parameter *b* controls the degree of spatial variability. When the soil storage at any point on its distribution is filled, effective rainfall is generated. The actual evaporation is assumed to be proportional to the wetness of the catchment. This model element has two model parameters, the maximum storage capacity, *cmax*, and the degree of spatial variability of the soil moisture capacity in the catchment, *b*. The routing component of the model can take various forms. One consists of two linear reservoirs in parallel, one representing a relatively quick catchment response and the other a slower response. All the effective rainfall is split between these two reservoirs, defined by parameter *%(q)* (proportion of the total effective rainfall going to the quick response reservoir). The two components of outflow are aggregated into total streamflow. This model element has three model parameters, i.e. two residence time of the each reservoirs, *rtq* and *rts*, and *%(q)*, thus giving 5 parameters for the model as a whole.

Models of this type can typically be considered to consist of a non-linear loss function, which accounts for rainfall losses, and a routing component, comprising a set of linear or non-linear stores in parallel. The Imperial College Rainfall-Runoff Modelling Toolbox (RRMT) (Wagener and Lees, 2001a, Wagener et al., 2002), written in the MATLAB environment, supports the implementation of this type of model. It allows a wide range of alternative moisture accounting modules and routing modules to be selected from a library and easily combined to form a chosen model structure (Figure 2).

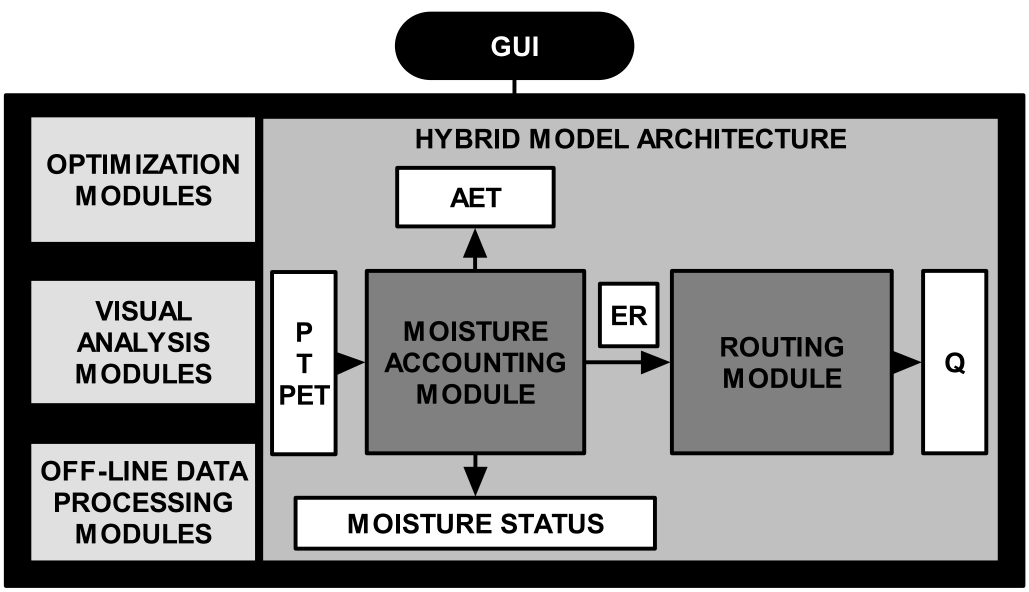

Figure 2 The Rainfall-Runoff Modelling Toolbox (RRMT)

Current computing power allows such models to be run many thousands of times on an ordinary PC, and hence simple, but powerful methods to analysis model structure and performance can be easily implemented. Figure 3 shows typical ‘dotty plots’ from a Monte Carlo analysis, where each dot represents a simulation. Values of the individual model parameters are drawn from pre-specified distributions by random sampling. The upper plots show the value of the objective function (which is being maximised) on the vertical axis, plotted against the associated value of a given parameter on the horizontal axis. The left hand image shows a clearly-defined optimum parameter value; the parameter is clearly ‘identifiable’. In contrast the right hand image shows that any value for the parameter can (in conjunction with different combinations of the other parameters) yield comparable results. The parameter is non-identifiable, at least in a univariate sense.

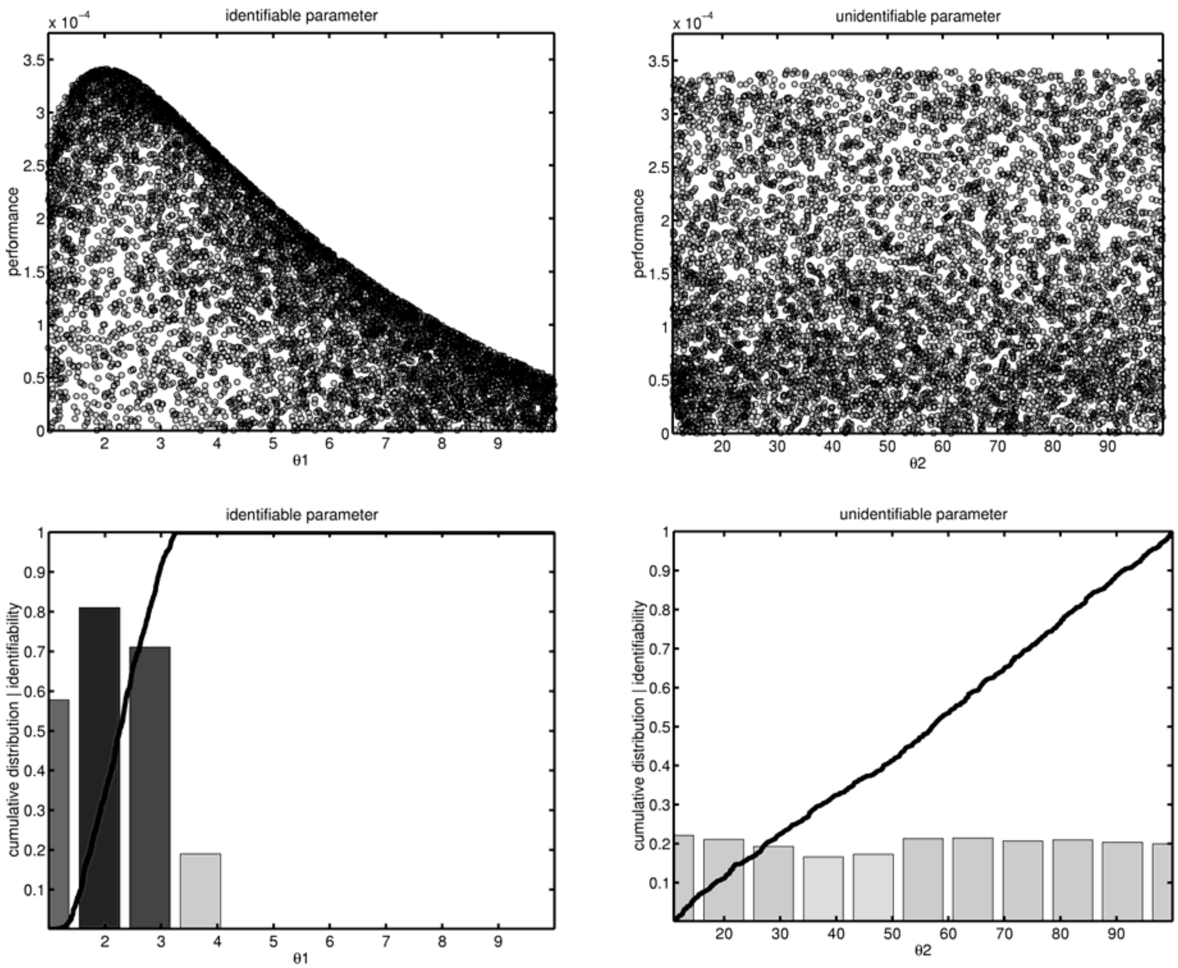

Figure 3. Example of a well and a poorly identified parameter. The top row shows scatter plots of parameter versus measure of performance. It has to be considered that these projections into a single parameter dimension can, however, hide some of the structure of the response surface (Beven, 1998). The bottom row shows the cumulative distribution of the best performing 10% of parameter sets and the corresponding gradients within each segment of the parameter range

Dynamic Identifiability Analysis (DYNIA) (Wagener et al., 2003) extends the analysis of parameter identifiability to investigate time-dependence. A measure of parameter identifiability is required, and the lower plots in Figure 3 illustrate one approach. The cumulative distribution of the best performing 10% of parameter sets is plotted, and the corresponding gradients within each segment of the parameter range are represented by a histogram. Grey-scale coding indicates the value of the gradient – darker colouring indicates a steeper gradient, and thus greater identifiability. We can now take a moving window through the output time-series (i.e. streamflow in this case) and evaluate the identifiability as a function of time. An example is shown in Figure 4 for the parameter CMAX (the maximum soil moisture storage capacity in the PDM model), evaluated using the NSE criterion. It can be seen that there are periods in the data series where the identifiability of the parameter increases (shown by the dark shading). It can also be seen that there are tensions in the model. There are periods where the identified parameter value is in the region of 100; there are other periods, notably where there is a significant hydrograph response following a dry period, where a higher value (around 500) is preferred.

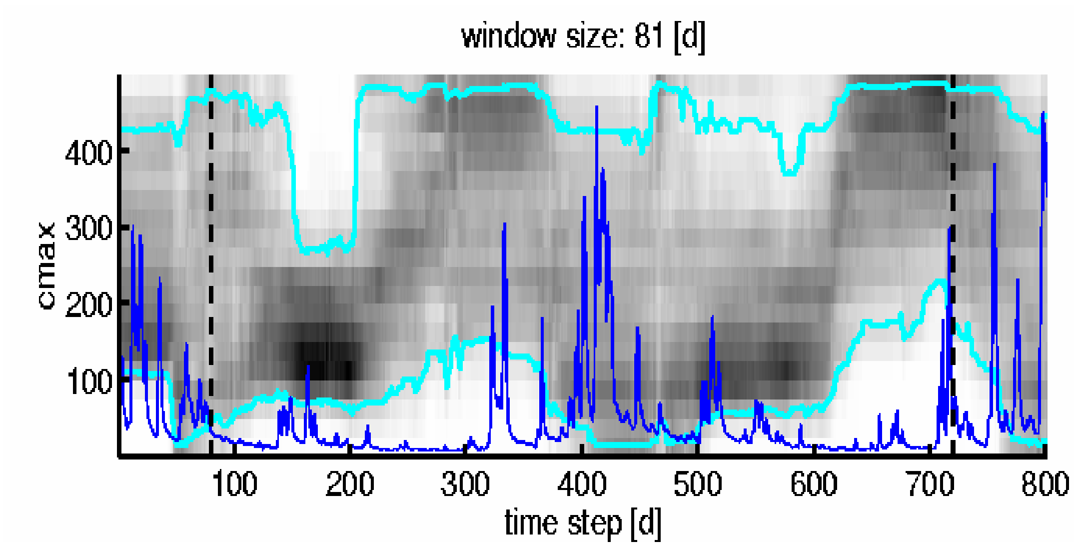

Figure 4 DYNamic Identifiability Analysis (DYNIA)

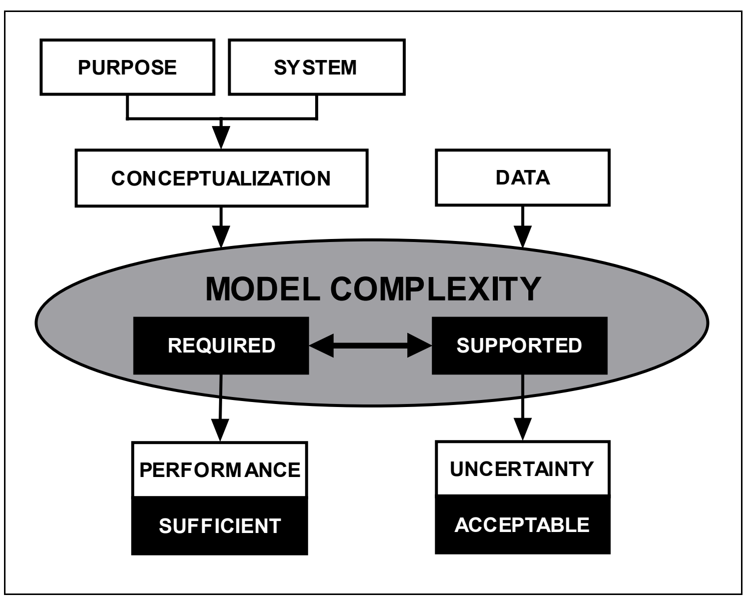

Figure 5 A framework for modelling

The analysis of model structure and parameter identifiability opens the way to a new framework for modelling, in which the attributes of the model are tuned to the modelling task (Figure 5). If an application is for example to select a model for regionalisation, then it is important to maximise parameter identifiability, and this may require simplification of model structure, possibly with some loss of simulation performance.

The use of multiple performance criteria

Secondly, there has been a move to extract more information from the available data. Until recently, automatic optimisation methods have used a single objective function (in contrast to the subjective manual assessment of goodness of fit, which intuitively inspects many different facets of model performance). The ambiguity that can arise from this is illustrated in Figure 6, where a set of simulations with the same NSE value are seen to have markedly different low flow properties. Research has shown that the information retrieved from a single objective function is sufficient to identify between three and five parameters (Beven, 1989, Jakeman and Hornberger, 1993), but it was realised that more information can be obtained from the available streamflow time-series by considering different aspects of the flow time-series (Wheater et al., 1986, Gupta et al., 1998, Boyle et al., 2000).

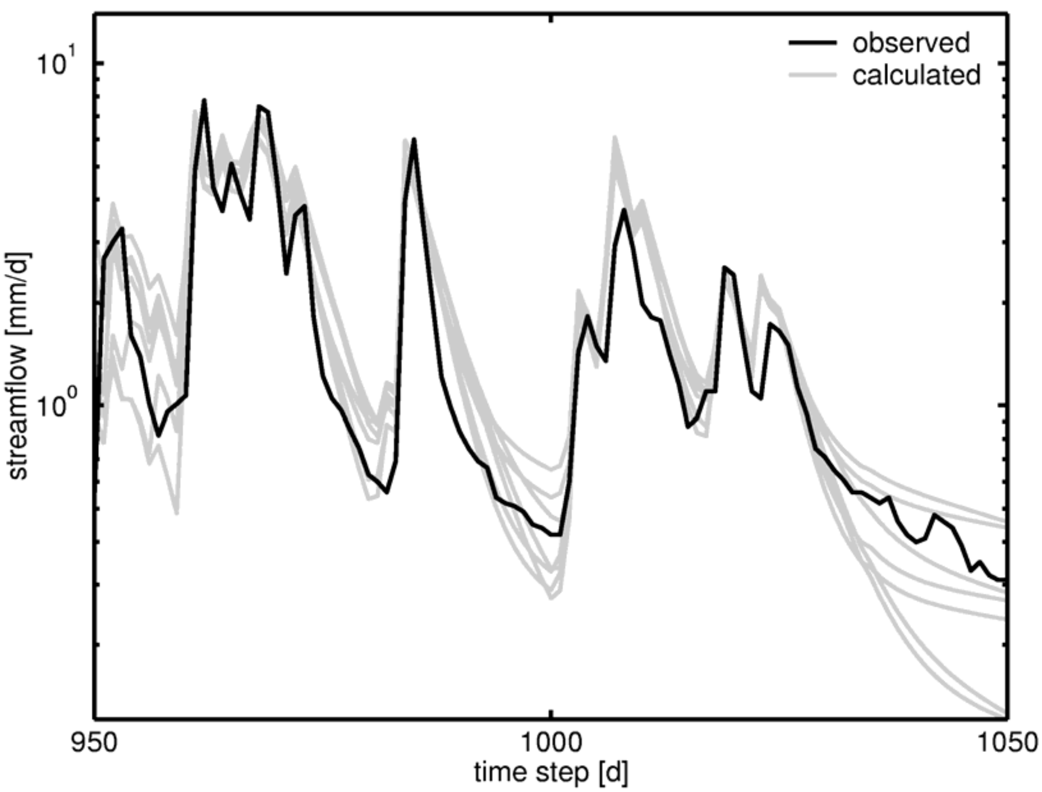

Figure 6 Simulated hydrographs with the same objective function value

Using multiple objective functions can provide information that improves the identifiability of certain parameters (Wheater et al., 1986). It can also illustrate tensions in the model. Figure 7 illustrates results from a set of Monte Carlo simulations, as discussed above. Two objective functions are calculated, one (NSE* =1-NSE) based on the sum of squared errors (and hence most strongly influenced by high flow performance), the other (FSB) based on the errors in low flow performance. The aim in this case is to minimise the objective functions. In the left-hand diagram it can be seen that the best low flow performance does not coincide with the best high flow performance, i.e. different parameter sets are required to optimise for low flows in comparison with high flows. A trade-off is needed – the user must decide where the modelling priorities lie, and choose an appropriate parameter set. In the right-hand diagram, the model structure is capable of maximising both high and low flow performance with minimal trade-off required.

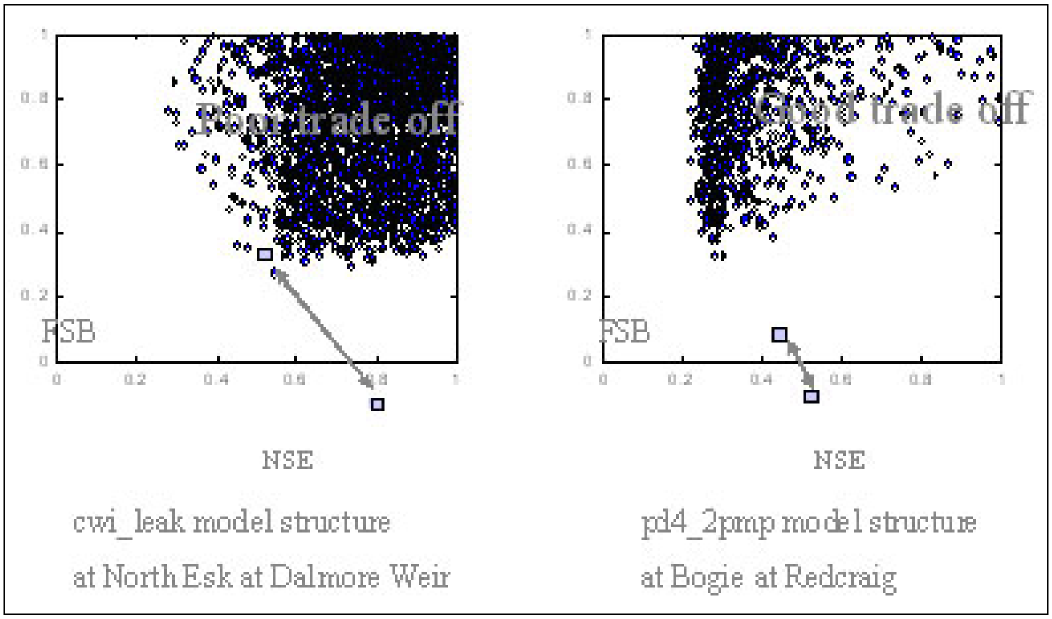

Figure 7 Multiple Objective Functions

Abandoning the concept of a unique best-fit model

The logical conclusion from the widespread observation that there is equifinality in model parameters and in model structures is to abandon the idea that a uniquely identifiable model exists. Rather, there is a population of models (i.e. structures and parameter sets) that can be defined according to their consistency with the available data. Spear and Hornberger (1980) developed the concept of Regionalised Sensitivity Analysis (RSA). They classified their model realisations as "behavioural" or "non-behavioural", depending on consistency with available data, and used this classification to explore parameter sensitivity. Parameters were sensitive if there was a significant difference between the set of behavioural and non-behavioural parameters. This is illustrated in Figure 8. The cumulative distribution of the values of parameter θ_i for the model realisations considered ‘behavioural’ is denoted ($F(\theta_i|B)$) and the distribution for the "non-behavioural" populations is denoted ($F(\theta_i|\underline{B})$). In the figure, parameter θ_i is sensitive (the populations are clearly different) and parameter θ_j is insensitive.

This approach was extended by Freer et al. (1996); instead of 2 classes, the model realisations are split into 10 groups of equal number, ranked according to their objective function performance, and the cumulative distributions can be plotted to indicate parameter sensitivity.

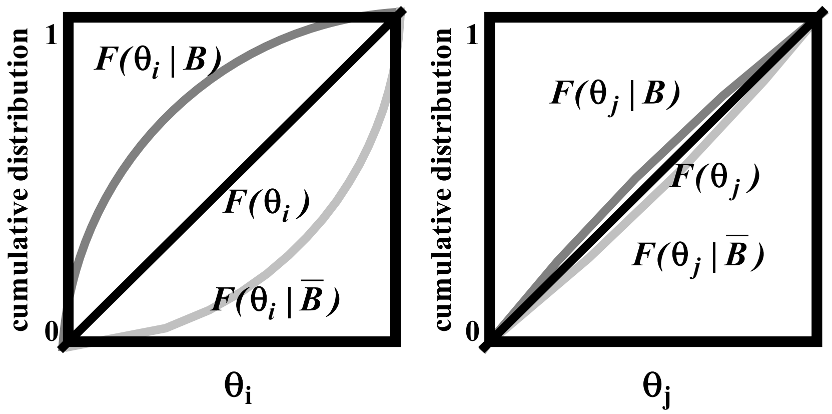

Figure 8. Cumulative distribution of initial (F(θi)), "behavioural" (F(θi|B)) and "non-behavioural" (F(θi|B)) populations for a sensitive parameter θi, and a (conditionally) insensitive parameter θi.

This analysis can be extended to estimate the uncertainties that arise in model structure, parameter values and data. A popular approach is the Generalised Likelihood Uncertainty Estimation (GLUE) procedure (Beven and Binley, 1992; Freer et al., 1996). Recalling the basic assumption that for a given data set a unique best-fit model (in terms of structure and parameter values) does not exist, the likelihood that a particular model and parameter set represents the data can be estimated, using some appropriate performance criterion.

As above, for a given model structure, parameter sets are generated (either based on random sampling of the feasible parameter space assuming a uniform distribution, or some known or assumed prior distribution) and the model is run using a Monte Carlo procedure. The simulations are classified as behavioural or non-behavioural, and the latter are rejected. The objective function values of the behavioural set of simulations can be used as 'likelihood' measures, which are scaled and used to weight the predictions associated with individual behavioural parameter sets. The modelling uncertainty is then propagated into the simulation results as confidence limits of any required percentile (Figure 9).

There are clearly limitations in this approach. The decision on whether a simulation is behavioural or not requires judgement to define a threshold value of objective function. And the various sources of uncertainty (input and output data error, model structural error, parameter estimation error and random error) are lumped into the single entity of parameter error. Nevertheless, the ability to specify the uncertainty associated with a simulation is a fundamentally important step. Given the availability of such tools, it will soon no longer be adequate to present a simulation output as a single best estimate, with no attempt to specify the associated confidence intervals.

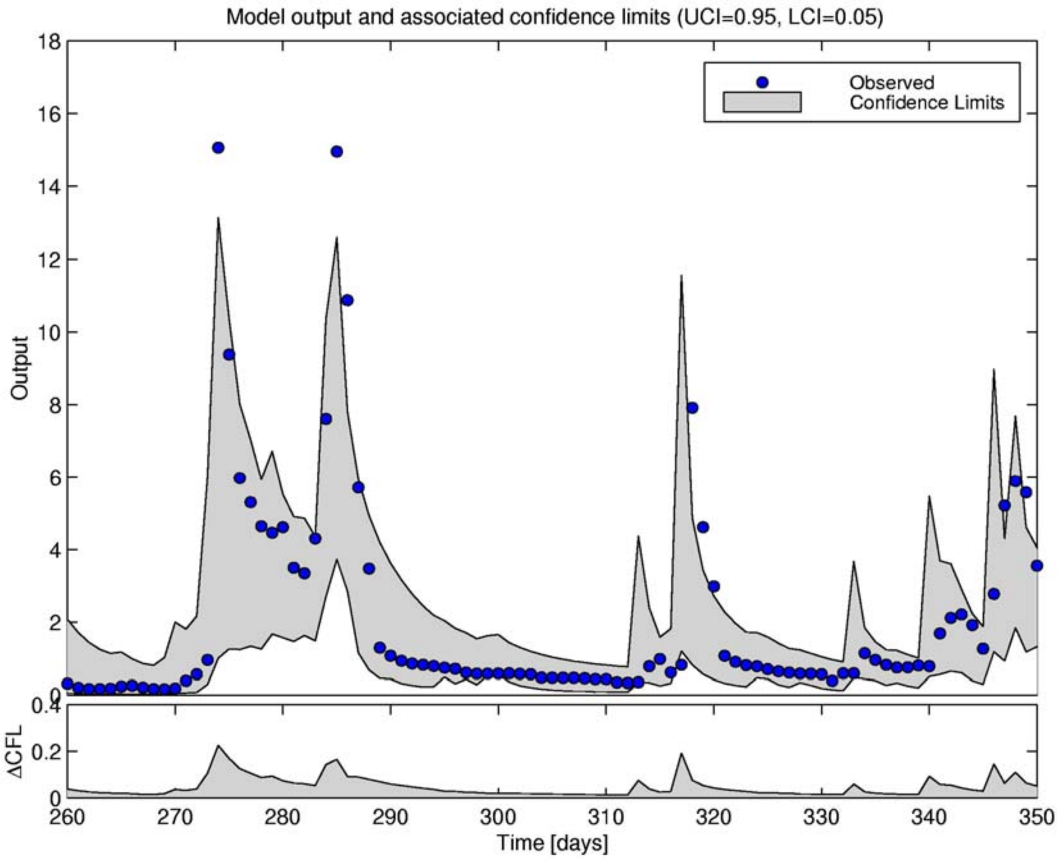

Figure 9 Simulation with 95% confidence limits

TOOLBOXES

It will be seen that important steps have been taken in our understanding of hydrological models, and that the ability to use simple Monte Carlo simulation methods has provided a powerful set of tools to support hydrological modelling through analysis of model structures, parameter identifiability and uncertainty. These techniques are now readily accessible through recently developed toolboxes.

The Imperial College Rainfall Runoff Modelling Toolbox (RRMT) (Wagener and Lees, 2001a) has already been introduced. The toolkit contains a large selection of parsimonious conceptual rainfall-runoff models, and also optimisation tools such as SCE-UA. A second Imperial College toolbox, also available in the MATLAB environment, is the Monte Carlo Analysis Toolbox (MCAT) (Wagener and Lees 2001b), Figure 10. This supports the tools for model sensitivity and uncertainty analysis discussed above, based largely on generalised sensitivity analysis and GLUE. Recently a capability for Dynamic Identifiability Analysis (DYNIA) has been added to the RRMT.

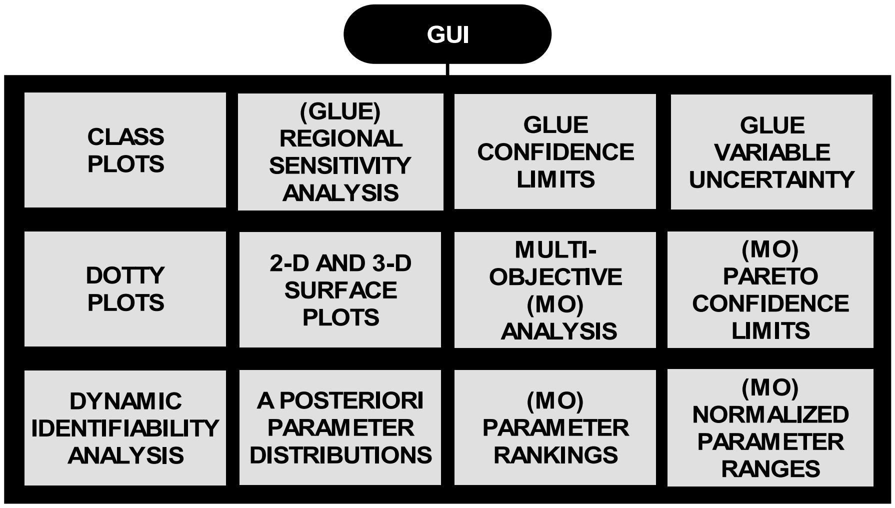

Figure 10 the Monte Carlo Analysis Toolbox (MCAT)

REGIONALISATION

One of the major challenges to hydrologists is the prediction of the hydrology of ungauged catchments. However, using parsimonious models and the tools described above, important progress is being made to regionalise continuous simulation rainfall-runoff models.

The procedure of regionalisation can be summarised as shown in Figure 11 (Wagener et al. 2004). Data from gauged catchments are used to calibrate a 'local model', and hence derive a set of model parameters for each catchment. These parameters are used, with a set of catchment characteristics, to generate a 'regional model', most commonly by regression analysis. The regional model can then be simply applied to ungauged catchments – if the catchment characteristics are known, the regional model generates the parameters to run the local model structure to represent the ungauged catchment. Many issues arise, for example what is the most appropriate local model given the range of catchment types, sizes, climate characteristics; what is the most approporiate procedure for development of the regional model. These are discussed in more detail by Wagener at al., 2004, but also provide considerable scope for new research.

A summary case study is presented below (after Wheater et al., 2002), to illustrate the basic procedure, using daily modelling of 23 small to medium sized catchments from the Thames basin, UK. Mean daily discharges for the ten year period 1st Jan 1990 to 31st Dec 1999 were used, together with mean daily precipitation (mm/d) for 36 raingauges across the study area, and weekly potential evaporation estimates. Physical catchment descriptors were avilable from the Flood Estimation Handbook, FEH (Institute of Hydrology, 1999) (see Table 1).

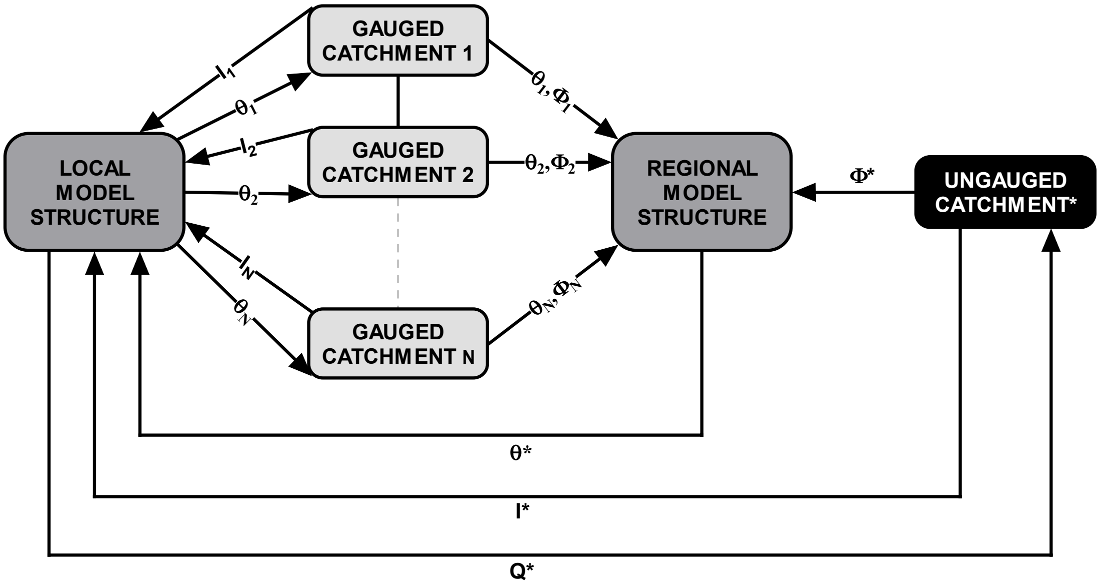

Figure 11 Model regionalisation procedure

CATCHMENT CHARACTERISTIC	UNIT	DESCRIPTION
AREA	km^2	Catchment drainage area
LDP	km	Longest drainage path
BFIHOST	-	Baseflow index derived using the HOST classification
SPRHOST	%	Standard percentage runoff derived using the HOST classification
FARL	-	Index of flood attenuation due to reservoirs and lakes
PROPWET	-	Index of proportion of time that soils are wet
DPLBAR	km	Index describing catchment size and drainage path configuration
DPSBAR	mkm^{-1}	Index of catchment steepness
ASPBAR	-	Index representing the dominant aspect of catchment slopes
ASPVAR	-	Index describing the invariability in aspect of catchment slopes
RMED-1D	mm	Median annual maximum 1-day rainfall
RMED-2D	mm	Median annual maximum 2-day rainfall
RMED-1H	mm	Median annual maximum 1-hour rainfall
SAAR	mm	1961-90 standard-period average annual rainfall
$SAAR_{4170}$	mm	1941-70 standard-period average annual rainfall
$URBEXT_{1990}$	-	FEH index of fractional urban extent for 1990
URBCONC	-	Index of concentration of urban and suburban land cover
URBLOC	-	Index of location of urban and suburban land cover

Table 1. Description of catchment characteristics (after Institute of Hydrology, 1999).

Several model structures were explored from the RRMT toolbox. Here we report results from the Catchment Wetness Index (CWI) loss function used in the IHACRES model (Jakeman *et al.*, 1990), combined with a parallel linear two-store model (2PAR), to represent the quick and slow flow components respectively.

Local Calibration

Table 2 lists the calibrated parameter sets (derived using Monte Carlo sampling using the root mean squared error (RMSE) criterion) and the RMSE and NSE criteria for all catchments (two catchments have been excluded due to shorter data length and calibration problems).

	Catchment Number	*tau*	*Refp*	*mf*	*rt(q)*	*rt(s)*	*%(q)*	RMSE	NSE
	C02	38.06	1.33	0.44	1.84	218.23	0.78	1.175	0.730
	C03	2.82	4.48	0.93	2.78	497.51	0.77	0.796	0.740
CLAY	C01	0.55	9.71	0.48	1.64	251.43	0.71	1.131	0.740
	C09	6.26	2.77	1.23	3.01	191.97	0.81	0.373	0.790
	C23	6.98	2.66	1.35	7.42	53.61	0.89	0.657	0.730
	C22	29.04	1.44	1.02	7.10	56.02	0.62	0.290	0.870
MIXED 1	C04	15.38	3.41	0.62	3.90	241.36	0.56	0.244	0.800
	C08	26.81	1.79	1.65	3.27	106.56	0.15	0.070	0.890
	C14	30.83	1.10	1.52	21.51	135.08	0.35	0.108	0.920
	C07	26.95	1.76	1.14	7.98	91.16	0.25	0.191	0.880
	C05	35.21	1.47	0.88	2.56	173.76	0.40	0.239	0.760
MIXED 2	C21	12.76	2.90	0.76	11.72	59.10	0.36	0.183	0.900
	C12	20.66	1.81	1.38	92.42	108.46	0.23	0.108	0.910
	C13	25.92	2.27	0.88	52.18	216.52	0.81	0.169	0.890
	C18	12.32	1.97	1.25	24.89	35.98	0.31	0.247	0.900
	C06	29.28	1.22	1.19	2.12	106.64	0.31	0.149	0.810
MIXED 3	C20	36.00	1.90	0.66	13.46	22.30	0.24	0.412	0.770
	C19	18.11	2.89	0.66	32.69	68.38	0.61	0.273	0.870
	C15	4.01	1.43	17.27	51.31	73.29	0.95	0.218	0.750
CHALK	C17	21.39	1.41	17.59	18.62	69.40	0.37	0.241	0.820
	C16	2.68	1.52	15.08	38.31	53.60	0.07	0.265	0.780

Table 2 Parameter values and model fits for the CWI_2PAR model.

The resulting model fits are all considered good, with NSE values ranging from 0.73 to 0.92 (average 0.82). The reduction in performance for the top five catchments can be clearly seen and is likely to be due to the use of daily data that is relatively coarse to represent the rapid flow responses on these impermeable catchments.

Regional analysis

Relationships between model parameters and individual FEH catchment characteristics were investigated using simple correlation analysis, although SMDBAR and ALTBAR values were only available for 12 of the 21 catchments. The parameter *tau,* the time constant of catchment wetness decline, is related to SMDBAR (correlation coefficient 0.60). The *refp* parameter is poorly correlated with all the catchment characteristics across most response periods. Parameter *mf*, which determines the difference in evapotranspiration (ET) between summer and winter, is, not surprisingly, strongly influenced by SMDBAR (correlation coefficient 0.98, see Figure 12). Parameter *mf* also shows a relationship with BFIHOST, which may be indicative of the influence of soil type and geology on water losses from the catchment.

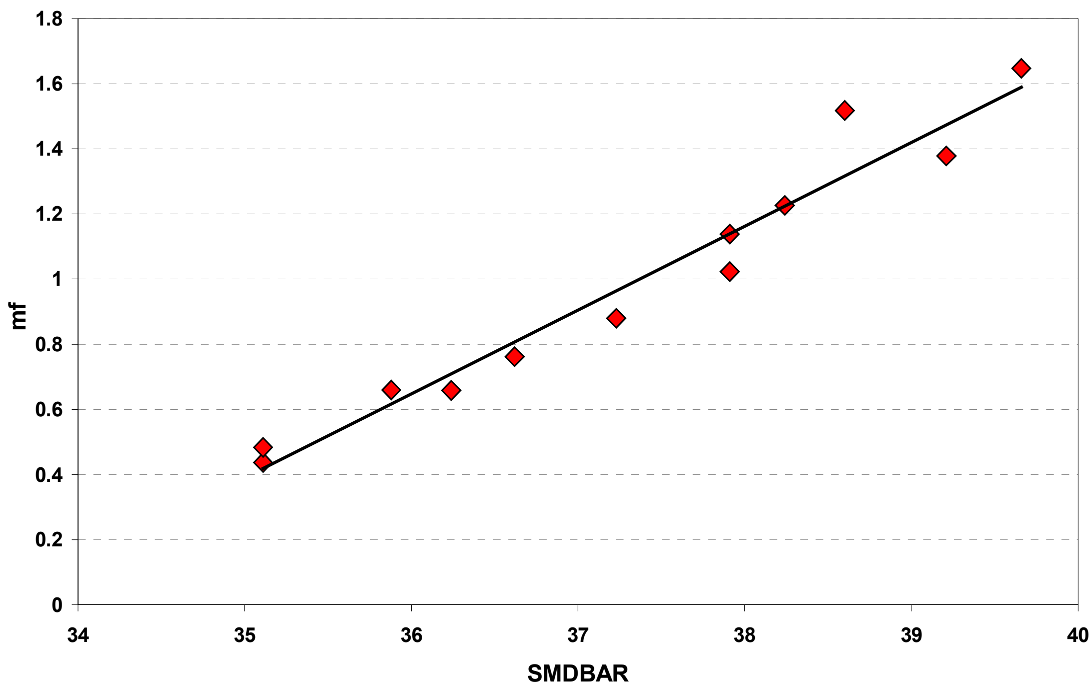

Figure 12. Overall RMSE best-fit mf versus SMDBAR.

The residence time constant of the quickflow reservoir, *rt(q)*, is consistently well identified for all objective functions by BFIHOST, which reflects the underlying geology of the catchment; as BFIHOST increases, so does the residence time. *rt(q)* was also found to be, to some extent, related to the urban measures of concentration (URBCONC) and location (URBLOC). This may be coincidental. *rt(q)* is also consistently related to the catchment slope measure (DPSBAR) and the median annual maximum rainfall for duration of 1 hour (RMED-1H).

The *rt(s)* parameter, the time constant for the slow flow reservoir, is again strongly related to BFIHOST, here with an inverse relationship. *rt(s)* is also related to PROPWET, a measure of the soil moisture status over time, to the average aspect of the catchment and also to the average altitude, ALTBAR.

Finally, relationships between the percentage quickflow of total flow (*%(q)*) and catchment characteristics were investigated. As expected, BFIHOST shows the strongest relationship with this parameter (Fig. 13), which also shows some level of correlation with SMDBAR and ASPBAR.

The most significant relationships between the model parameters and catchment descriptors, indicated by the Pearson correlation coefficient, *r*, in brackets, are shown in Table 3.

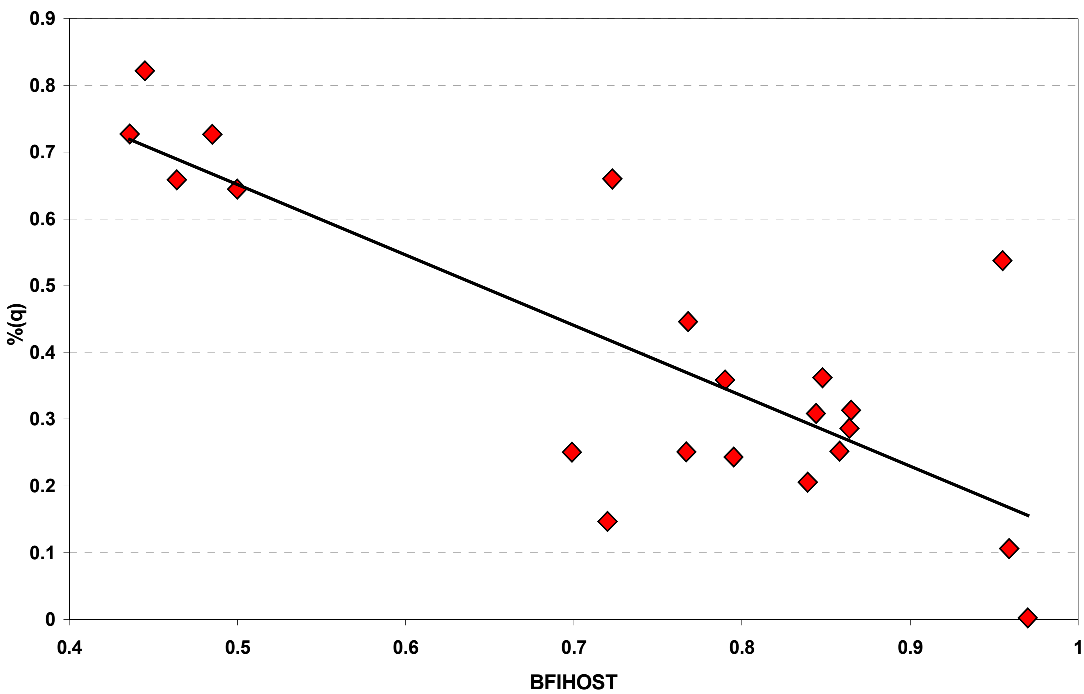

Figure 13. RMSE best-fit %(q) versus BFIHOST for all catchments

MODEL PARAMETER	**MOST SIGNIFICANT CATCHMENT DESCRIPTOR (CORRELATION COEFFICIENT)**			
tau	DPSBAR (-0.27)	BFIHOST (0.19)	RMED-1H (0.13)	ASPVAR (-0.12)
refp	BFIHOST (-0.53)	PROPWET (0.42)	FARL (-0.42)	ASPVAR (0.32)
mf	BFIHOST (0.53)	AREA (-0.32)	FARL (0.54)	PROPWET (-0.05)
rt(q)	BFIHOST (0.53)	DPSBAR (0.47)	PROPWET (-0.38)	FARL (0.27)
rt(s)	BFIHOST (-0.62)	PROPWET (-0.38)	DPSBAR (-0.38)	ASPBAR (0.37)
%(q)	BFIHOST (-0.54)	ASPBAR (0.32)	RMED-1H (-0.23)	
volc	BFIHOST (0.58)	DPSBAR (0.54)	ASPBAR (-0.44)	SAAR (0.25)

Table 3. The most significant relationships between CWI_2PAR parameters and catchment descriptors. (Note - Urban measures are excluded.)

Multiple regression regional analysis

A multiple regression analysis was carried out using simple routines in MATLAB. The procedure allows the simple exclusion of outlier values, e.g. the *mf* values found for the three chalk catchments (C15, 16, 17). A test catchment was selected, namely the River Wey at Farnham (C06), and removed from the multiple regression data. The resulting relationships are as follows:

tau **= -0.4648 (DPSBAR) - 37.5915 (ASPVAR) - 4.2586 (RMED-1H) + 99.5594**

(R^2 = 0.514)

$refp$ = -1.1997 (BFIHOST) - 12.7610 (PROPWET) + 0.2643 (ASPVAR) - 0.2687 (FARL) + 7.0517 (R^2 = 0.342)

mf = -13.3104 (PROPWET) - 0.3199 (BFIHOST) + 0.00013 (AREA) + 4.9492 (FARL) + 0.8636 (R^2 = 0.647)

$rt(q)$ = 29.0766 (BFIHOST) - 0.1514 (DPSBAR) -14.1469 (PROPWET) - 3.8213

(R^2 = 0.741)

$rt(s)$ = -315.0290 (BFIHOST) + 1016.5670 (PROPWET) + 0.01487 (ASPBAR) - 0.0619 (DPSBAR) + 25.6877 (R^2 = 0.685)

$\%(q)$ = -1.0597 (BFIHOST) + 0.00019 (ASPBAR) - 0.0031 (RMED-1H) + 1.3342

(R^2 = 0.733)

$volc$ = -0.00698 (BFIHOST) – 0.00004 (DPSBAR) – 0.000004 (ASPBAR) + 0.01065 (R^2 = 0.754)

where *volc* is a factor introduced to ensure that the total volume of modelled effective rainfall equals the total volume of observed streamflow. This parameter is calculated during the calibration stage, but has to be estimated during prediction and therefore regionalisation.

Table 4 gives a comparison of the calibrated and regionally estimated parameter values for the Farnham catchment (C06); Table 5 and Figure 14 present the associated performance.

PARAMETER	CALIBRATED	ESTIMATED
tau	29.2828	25.1805
refp	1.2182	1.2975
mf	1.1864	0.8528
rt(q)	2.1170	8.0050
rt(s)	106.6432	106.5143
%(q)	0.3130	0.3959
volc	0.0017	0.0021

Table 4. Calibrated and regionally estimated parameter values for the Farnham catchment (C06).

PERFORMANCE MEASURE	CALIBRATED	ESTIMATED
RMSE	0.149	0.189
NSE	0.81	0.69

Table 5. Model performance measures for the calibrated and estimated model parameters for the Farnham catchment.

Conclusions from regionalisation

This pilot investigation was of limited scope, but nevertheless provides some insight into the hydrological controls on the catchments investigated, as well as confirming the potential of the methodology for more extensive regional analysis.

The regionally estimated parameters, found using the multiple regression method, showed encouraging results with respect to both the proximity of the estimated to the calibrated parameter values, and also in the good model fits to observed streamflow data. Only limited validation was possible given the number of catchments used. The derived regional relationships show a good level of physical justification with regard to the catchment descriptors used and the BFIHOST variable was found to be significant in many of the parameter estimation equations, especially for the routing module parameters. This is indicative of the dominant effect of the soils and geology of the catchment, reflected in the BFIHOST value, on the catchment response characteristics.

CONCLUSIONS

This Chapter has reviewed developments in our understanding of the performance of conceptual rainfall-runoff models. Since they were first developed, some 40 years ago, important progress has been made, and current computing power has enabled the development and application of a powerful range of stochastic, Monte-Carlo based tools for the analysis of model performance. The limitations of these models have mainly been due to the problems of parameter interdependence – models have been too complex to be identified uniquely from the available data.

Various issues have been discussed. It has been shown that automatic methods of model calibration have finally caught up with the experienced user by incorporating multiple objective functions in the evaluation of model performance. For a rainfall-runoff model, different aspects of the streamflow hydrograph can be considered. For other applications, such as sediments or water quality, different output signals can be evaluated. This generally illustrates tensions in our models. The parameters which are optimal for high flow simulation will not be the same as those which are optimal for low flows, and in a water quality model, the hydrological parameters which are optimal for flow may not be same as those that give optimal chemical simulations. This provides important information to the user, who can decide what are the appropriate criteria for a particular application.

A major step forward has been the recognition that alternative models and parameters may be equally likely interpretations of the available data. This has led to the development of methods to quantify model uncertainty, which means that the modeller can provide an informed estimate of the uncertainty associated with a model simulation. This provides a radically different way of communicating information to the users of model results, and enables much more informed decisions to be made, for example about the risk associated with a given management strategy.

Finally, the ultimate challenge to the hydrologist is to predict the response of an ungauged catchment. We have seen that parsimonious models have the capability to deliver convincing results, and such regional methods can be expected to rapidly find their place in hydroloogical analyses (Wheater, 2002).

The modelling methods and tools discussed are available in the Matlab-based RRMT and MCAT Toolboxes. These can be downloaded free for research users from:

http://ewre.cv.imperial.ac.uk. Please note that Matlab software is required for their implementation.

REFERENCES

Beven, K.J. Changing ideas in hydrology: the case of physically-based models. J.Hydrol., 105, 157-172 (1989)

Beven, K.J. Prophecy, reality and uncertainty in distributed hydrological modelling. Adv. In Water Resourc., 16, 41-51 (1993)

Beven, K.J. and Binley, A.M. The future of distributed models: model calibration and predictive uncertainty. Hydrol. Processes, 6, 279-298 (1992)

Boyle, D.P., Gupta, H.V. and Sorooshian, S. Towards improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. Water Resour. Res. 36, 3663-3674 (2000)

Duan, Q., Gupta, V.K. and Sorooshian. S. Effective and efficient global optimisation for conceptual rainfall-runoff models. Water Resour. Res. 28, 1015-1031 (1992)

Duan. Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N.and Turcotte, R. (Eds.) Calibration of Watershed Models. American Geophysical Union Monograph Series, Water Science and Application 6 (2003)

Freer, J., Beven, K. and Abroise, B. Bayesian uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resour. Res. 32, 2163-2173 (1996)

Gupta, H.V., Sorooshian, S. and Yapo, P.O. Towards improved calibration of hydrological models: multiple and non-commensurable measures of information. Water Resour. Res. 34(4), 751-763 (1998)

Ibbitt, R. Systematic Parameter fitting for Conceptual Models of Catchment Hydrology. Unpublished PhD thesis, Univ. of London (1970)

Institute of Hydrology. Flood Estimation Handbook (1999)

Jakeman, A.J. and Hornberger, G.M. How much complewxity is warranted in a rainfall-runoff model? Water Resour. Res., 29, 2637-2649 (1993)

Jakeman, A.J., Littlewood, I.G. and Whitehead, P.G. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. Journal of Hydrology, 117, 275-300 (1990)

Johnston, P.R.and Pilgrim, D.H. Parameter optimization for watershed models. Water Resour. Res. 12(3), 477-186 (1976)

Moore, R.J. and Clarke, R.T. A distribution function approach to rainfall-runoff modelling Water Resour. Res. 17(5), 1376-1382 (1981)
Nash, J.E. and Sutcliffe, J.V. River flow forecasting through conceptual models 1. A discussion of principles. J.Hydrol.10, 282-290 (1970)

Nelder, J.A. and Mead, R. A simplex method for function minimisation Computer journal, 7, 308-313 (1965)

Rosenbrock, H.H. An automatic method for finding the greatest or least value of a function. Computer Journal,3, 175-184 (1960)

Spear, R.C. and Hornberger, G.M. Eutrophication in Peel inlet, II, Identification of critical uncertainties via generalised sensitivity analysis. Water Resour. Res. 14, 43-49 (1980)

Wagener, T. and Lees, M.J. Rainfall-Runoff Modelling Toolbox user manual. Department of Civil and Environmental Engineering, Imperial College London, UK. (2001a)

Wagener, T. and Lees, M.J. Monte-Carlo analysis Toolbox user manual. Department of Civil and Environmental Engineering, Imperial College London, UK. (2001b)

Wagener, T., Lees, M.J. and Wheater, H.S. A toolkit for the development and application of parsimonious hydrological models. In Singh, V.P. and Frevert, D. (eds.) Mathematical models of large watershed hydrology – Volume 1. Water Resources Publishers, USA 87-136 (2002)

Wagener, T., Wheater, H.S. and Gupta, H.V. Identification and evaluation of watershed models. In: Calibration of Watershed Models. Water Science and Application Vol 6, Ed Qingyun Duan, Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. American Geophysical Union., 29-47 (2003a)

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S. and Gupta, H.V. Towards reduced uncertainty in conceptual rainfall-runoff modeling: dynamic identifiability analysis. Hydrological Processes, 17(2), 455-476 (2003b)

Wagener, T., Wheater, H.S. and Gupta, H. Rainfall-Runoff Modeling in Gauged and Ungauged Catchments. Imperial College Press (2004)

Wheater, H.S. Progress in and prospects for fluvial flood modelling. Phil. Trans. R. Soc. Lond. A, 360, 1409-1431 (2002)

Wheater, H.S., Bishop, K.H. & Beck, M.B. "The identification of conceptual hydrological models for surface water acidification". J. Hydrol. Proc. 1, 89-109 (1986)

Wheater, H.S., Jakeman, A.J. and Beven, K.J. "Progress and directions in rainfall-runoff modelling". In: Modelling Change in Environmental Systems, Ed. A.J. Jakeman, M.B. Beck and M.J. McAleer, Wiley, 101-132 (1993).

Wheater, H.S., Boxall, S. and Wagener, S. Regionalisation of rainfall-runoff models: an application to the Thames Basin. Proceedings British Hydrological Society 8th National Hydrology Symposium, Birmingham, 2002, pp 199-206 (2002)