

The reliability of an 'off-the-shelf' conceptual rainfall runoff model for use in climate impact assessment: uncertainty quantification using Latin hypercube sampling

Conor Murphy, Rowan Fealy, Ro Charlton and John Sweeney

Irish Climate Analysis and Research Units (ICARUS), Department of Geography, NUI Maynooth,
Co. Kildare, Ireland

Email: conor.m.murphy@nuim.ie

Revised manuscript received 12 April 2005

Much uncertainty is derived from the application of conceptual rainfall runoff models. In this paper, HYSIM, an 'off-the-shelf' conceptual rainfall runoff model, is applied to a suite of catchments throughout Ireland in preparation for use in climate impact assessment. Parameter uncertainty is assessed using the GLUE methodology. Given the lack of source code available for the model, parameter sampling is carried out using Latin hypercube sampling. Uncertainty bounds are constructed for model output. These bounds will be used to quantify uncertainty in future simulations as they include error derived from data measurement, model structure and parameterization.

Key words: Ireland, climate change, uncertainty, Latin hypercube sampling, hydrologic model, transferability

Introduction

Evidence for global warming over the last century is now overwhelming. Global mean surface temperatures have increased by $0.6 \pm 0.2^\circ\text{C}$ over the course of the twentieth century, with the increase in Northern Hemisphere surface temperatures being greater than in any other century in the last 1000 years, while the 1990s were the warmest decade of the millennium (McCarthy *et al.* 2001). Globally averaged surface temperatures are projected to increase under all International Panel on Climate Change (IPCC) emissions scenarios during the present century as a consequence of increases in atmospheric CO_2 concentrations. Indeed, in their Third Assessment Report (TAR), the IPCC dramatically revised upwards the top range limit of their climate predictions of global warming between now and

2100 from $1\text{--}3.5^\circ\text{C}$ in the previous report to $1.4\text{--}5.8^\circ\text{C}$. There is broad agreement that such anthropogenic climate change is likely to have a large impact on water resources, with the availability of water to meet future demands and the magnitude and frequency of future extreme events being uncertain. As such it is important to attempt to model the impacts of such changes so as informed policy decisions can be made.

Uncertainty in modelling

When modelling the effects of climate change on water resources, there is a cascade of uncertainty that begins when future economic story lines are translated into future emission scenarios and ends with impact modelling (Wilby 2005). Conceptual rainfall runoff (CRR) models have been the most

widely used for this purpose (Cunnane and Regan 1994; Arnell and Reynard 1996; Sefton and Boorman 1997; Pilling and Jones 1999 2002; Arnell 2003; Charlton and Moore 2003). Constraints are placed on such an approach by a lack of knowledge of the workings of the hydrological system, a lack of data and by the volume of complex computations required to simulate every process within the hydrological sphere. Consequently, CRR models incorporate large simplifications in order to represent catchment hydrology. One of the major consequences of such simplifications is the generation of uncertainty within the modelling framework. Oberkampf *et al.* (2002) divide uncertainty into aleatory and epistemic uncertainty. The former describes the inherent variation associated with the physical system or environment and is irreducible. Epistemic uncertainty is a potential inaccuracy in any phase or activity of the modelling process that is due to lack of knowledge and is thus referred to as cognitive, subjective and reducible. Such uncertainty is seen in the process of parameter estimation with well-known limitations attributable to parameter identifiability, parameter stability, uncertainty and the equifinality of outputs arising from different combinations of model parameters. In many cases where 'off-the-shelf' CRR models are applied, they are done so without prior knowledge of the uncertainty associated with model output due to parametric sources. Where uncertainty is catered for, the vast majority of work has used probabilistic methods to represent sources of uncertainty and then sampling methods, such as Monte Carlo sampling, to propagate these sources through the model (Khu and Werner 2003).

Uncertainty evaluation

Uncertainty evaluation generally holds that all acceptable parameters or models of a system be retained until they are disproved and consists of analysing the range of parameter sets that are acceptable for a specific application (Wagener 2003). These plausible models are used to construct uncertainty bounds or confidence limits for model output. One established method for uncertainty analysis is the Generalized Likelihood Uncertainty Estimation (GLUE) procedure (Beven and Binley 1992). GLUE recognizes the need to quantify the reliability of model simulations and starts with the recognition that many model structures or parameter sets within a given model framework will simulate a required output. Given this concept of equifinality,

it follows that no single optimum set of model parameters can be readily identified (Beven 1993). Consequently, it is only possible to assign a likelihood value to each parameter set, indicating that it can predict the system and that the set of parameters provides an acceptable or behavioural simulation of the observed flow (Beven and Binley 1992). The GLUE procedure has five main steps (Beven and Binley 1992):

- 1 The definition of a likelihood measure, chosen on the basis of an objective function to determine model performance.
- 2 The definition of a prior distribution for each parameter.
- 3 A large number of parameter sets are sampled using Monte Carlo Random Sampling (MCRS) from the prior distributions (Brazier *et al.* (2000) used up to three million; Beven and Freer (2001) used 60 000).
- 4 Each parameter set is classified as behavioural or non-behavioural through assessing whether it performs above or below a pre-defined threshold.
- 5 Predictive model runs generate results from each of the parameter sets that yield acceptable calibration simulations. These combined simulations are in turn used to determine the weighted mean discharge and simulation probability bounds (Melching 1995).

There are a number of problems commonly encountered when conducting an uncertainty analysis. For instance, it is impossible to extract the amount of uncertainty associated with different sources, such as model structure, the parameters or input data (Wagener 2003). Furthermore, such sampling requires considerable computing resources, and in situations where the end user was not involved in model construction access to the model source code may be problematic. Thus the ability to run the model for large samples without intervention is inhibited. This can be problematic when using MCRS where many thousands or even tens of thousands of model runs are required to adequately sample the entire parameter space.

Uncertainty analysis and LHS

In order to overcome this obstacle, quasi-stratified sampling in the form of Latin hypercube sampling (LHS) (McKay *et al.* 1979) can be applied. LHS selects n different values from each of k variables $x_1 \dots x_k$, by attributing a pdf to each variable and dividing the range into n non-overlapping intervals on the

basis of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The n values thus obtained for x_1 are paired in a random manner with the n values of x_2 . These n pairs are combined in a random manner with the n values of x_3 to form n triplets, and so on until n k -tuplets are formed (Iman and Helton 1988).

In comparing uncertainty analysis methods, Yu *et al.* (2001) and Melching (1992) investigated the efficiency of different sampling methods used in uncertainty analysis, including MCRC and LHS, and compared the convergence rates of each method. Yu *et al.* (2001) indicate that only LHS produces results similar to the Monte Carlo approach. The authors concluded that LHS could generate representative samples more efficiently than MCRC due to characteristic uniform sampling of the parameter space. Recommendations for the necessary number of model runs vary in the literature. Iman and Helton (1985) suggest sampling two to five times the number of varied model parameters. Selection of runs in this way is rather subjective. Melching (1995) proposed to define the number of model runs necessary by checking for convergence of statistical measures of model output on the number of executed model runs. The analysis of convergence rates for the HYSIM model show that the generation of 100 samples is adequate.

Model and methodology

HYSIM is a CRR model, which uses rainfall and potential evaporation data, on a daily time-step, to simulate river flow using parameters for hydrology and hydraulics that define the river basin and channels in a realistic way. Although spatially lumped and hydrologically conceptual in nature, the model contains many parameters that can be measured from physical reality. HYSIM has been used for a variety of hydrological applications,

including assessing the impacts of climate change on the water resources (Pilling and Jones 1999 2002; Charlton and Moore 2003). The mathematical model is built around two sub-routines. The first of these simulates the hydrology of the catchment, while the second simulates the hydraulics. The complete flow diagram of the structure of the model is given in Figure 1. In relation to the hydrology routine, seven natural stores are represented. The main components of the model are the upper and lower soil reservoirs, with the works of Brooks and Corey (1964) employed to represent the variation of effective permeability and capillary suction with changes in moisture content. A full description of the model and its structure is given in Manley (1977 1993).

Parameters within HYSIM can be broken down into two groups, the physical parameters and the process or 'free' parameters (Sorooshian and Gupta 1995). The former represent physically measurable properties of the watershed, such as channel slope and land-use characteristics, whereas process parameters represent watershed characteristics that are not directly measurable, such as the lateral interflow rate. The process parameters and a description of each are given (Table 1). There are two approaches to fitting the model that can be taken. The first is that of specification of the physically measurable parameters, while the second involves the optimization of process parameters. In relation to the process parameters, the idea of specification can only be extended to defining parameter ranges (minimum and maximum values) and calibration is employed to refine these estimates (Sorooshian and Gupta 1995).

A split sample procedure was adopted for calibration and validation. The first 30 years of the baseline data set (1961–90) were used for calibration. This period was selected so that the model could be trained on as much variability in streamflow as possible. Validation was conducted for the period

Table 1 HYSIM process parameters and description

Parameter	Description
Permeability at horizon boundary (PHB)	The rate at which moisture moves between the two soil horizons
Permeability at base of lower horizon (PBL)	The rate at which moisture leaves the soil layers
Interflow runoff from upper horizon at saturation (IU)	Direct, or lateral runoff from the upper soil horizon
Interflow runoff from the lower horizon at saturation (IL)	Direct runoff from the lower horizon

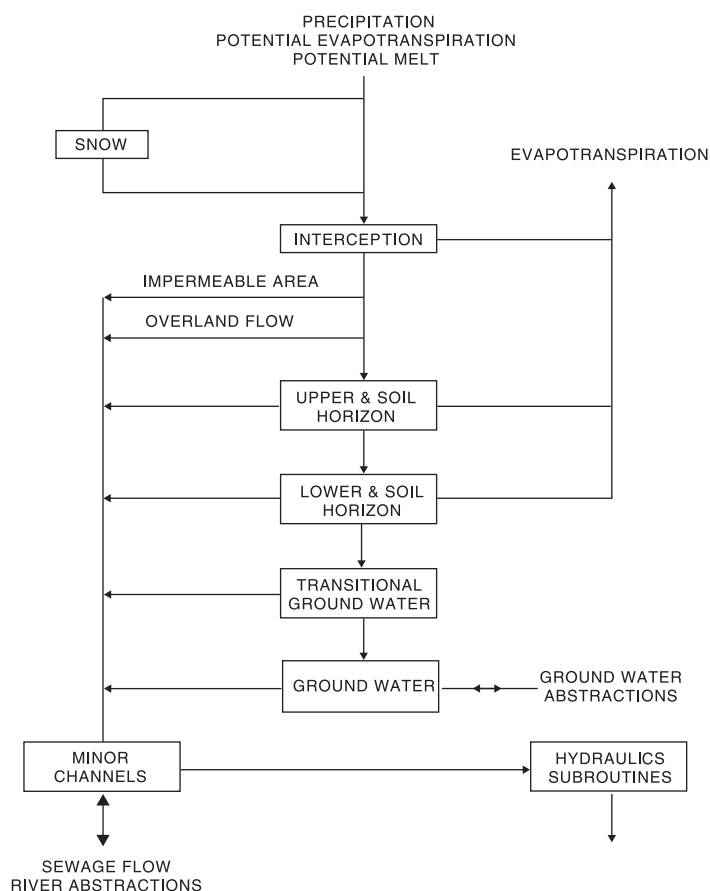


Figure 1 HYSIM model structure

1991–2000. This decade has been the warmest globally, with 1998 being the warmest year on the global instrumental record. In Ireland the warmest year was recorded in 1997. Furthermore the ten years 1991–2000 present some of the largest flood peaks on record in Ireland, such as the November 2000 floods in the Suir catchment. Thus the 1990s provide a good test of model performance, with conditions being more akin to those expected under climate change than at any other period in the baseline data set.

In total ten catchments throughout Ireland are modelled in this work. The catchments studied and their locations are given in Table 2. For ease of presentation only the River Suir will be dealt with here. The methodology is divided into a number of steps beginning with an analysis of the sensitivity of model output to individual parameters. Parameter

specification is aided by the incorporation of a GIS. Problems within the calibration procedure are highlighted and the final step in the methodology is that of uncertainty quantification and the construction of uncertainty bounds for model output.

Sensitivity analysis

In order to gain a more complete understanding of the degree to which model output is sensitive to the parameter values a sensitivity analysis was conducted. This is most readily achieved by varying one parameter at a time and running the model with the rest of the parameters remaining fixed (Lenhart *et al.* 2002). The major limitation of this approach is that it provides only a very local estimate of sensitivity within the parameter space (Beven 2000). However, it does remain quite powerful in gaining

Table 2 Catchment characteristics

Catchment	Area (km)	Gauge	Data (days)	Mean rainfall (mm)	Mean ET (mm)	Mean discharge (cumecs)	Land use	Soil texture
Suir	3556.00	Clonmel	14 610	2.7	1.27	48.2	Pasture	Loam
Blackwater	3245.70	Ballyduff	14 610	3.1	1.5	62.3	Pasture	Loam
Boyne	2670.50	Slane	14 610	2.4	1.22	35.4	Pasture	Clay loam
Moy	1980.87	Rahans	9 862	3.9	1.22	57.9	Peat bogs	Loam
Barrow	2956.00	Levitstown	11 688	2.5	1.27	20.9	Pasture	Sandy loam
Brosna	1082.50	Ferbane	14 610	2.4	1.22	17.1	Pasture	Loam
Inny	1072.50	Ballymahon	10 227	2.6	1.22	18.7	Pasture	Loam
Suck	1050.00	Bellagill	9 498	2.8	1.22	25.2	Pasture	Loam
Bonet	371.57	Dromahair	14 516	3.3	1.2	11.2	Natural	Clay loam
Ryewater	213.90	Leixlip	14 610	2.2	1.5	2.3	Pasture	Clay loam

an insight into the functioning of the model and can adequately determine whether parameters are described in enough detail for the modelling procedure at hand.

The most sensitive parameters were those most strongly linked with the water balance. Of the components of the water balance, the soil parameters show themselves to be the most sensitive, with vegetation characteristics also displaying a high sensitivity. The findings of the analysis agree with previous work conducted by Manley (1978), in which the soil parameters such as the pore size distribution index and the process parameters are highlighted as the most sensitive within the model framework. Consequently, the parameterization procedure was constructed in order to adequately account for the most sensitive model inputs.

Parameterization

Parameter specification

The first method to consider when parameterizing the model was that of catchment delineation (Figure 2a). Soil hydrological properties were calculated from the General Soils Map of Ireland (Gardiner and Radford 1980) (Figure 2b). Textures were found for the top two soil horizons, as HYSIM deals with both the upper and lower layers of the soil. Each association within the catchment was examined and the proportions of the soil type and its location within the catchment were considered. The dominant soil texture was calculated by establishing the per cent sand, silt and clay in each

soil association with the derived texture being used to calculate the soil parameters.

Vegetation parameters were obtained using the CORINE dataset (Coordination of Information on the Environment) (O'Sullivan 1994). CORINE provides comprehensive data on biophysical land occupation that are consistent and comparable across Europe at a scale of 1:100 000 (see Figure 2c). Due to the lumped nature of the model, the land use with the highest percentage was used to derive the land use parameters. Groundwater parameters were estimated using the Aquifer Map of Ireland (Geological Survey of Ireland 2003) (Figure 2d).

Parameter estimation

Within HYSIM calibration is catered for by a multi-parameter optimization procedure. HYSIM employs the Rosenbrock method, a local search algorithm using a direct search method. Blackie and Eeles (1985) provide details on the functioning of the Rosenbrock method. In order to calibrate the model the extremes error of estimate objective function was used. This calculation gives equal weight to extreme values, be they high or low flows. Additional goodness of fit measures were also employed. These comprised the Nash-Sutcliffe efficiency criterion (NS), the Coefficient of Determination (R^2), the Mean Actual Error (MAE) and the Percent Bias (PBIAS). Both correlation and relative error measures were included as the use of correlation-based measures alone can be over sensitive to extreme values and are insensitive to additive and

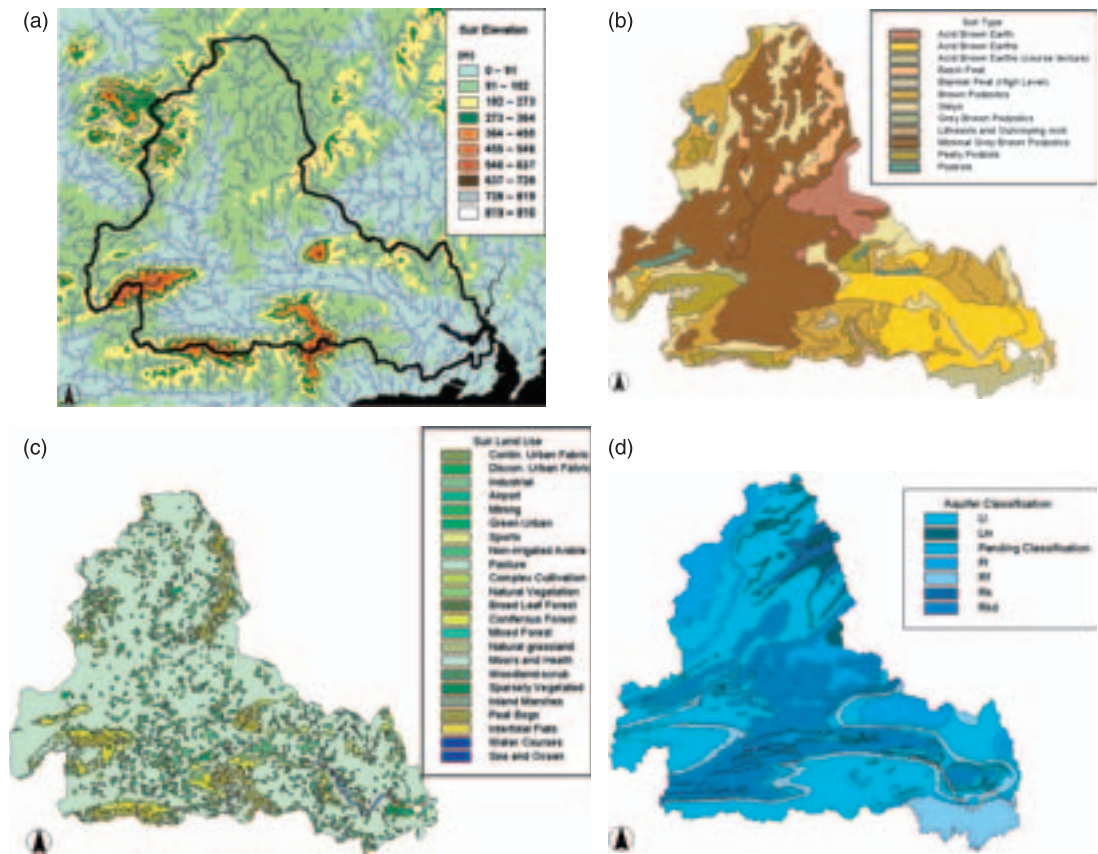


Figure 2 Parameter specification from spatial datasets. (a) Elevation, (b) soil type, (c) vegetation, (d) groundwater

proportional differences between model predictions and observations (Legates and McCabe 1999). Only results for NS and PBIAS will be given here. For NS, values of 1 indicate a perfect fit, while a PBIAS of 0 per cent is ideal.

Once the optimum parameter set was realized for each catchment, the Rosenbrock algorithm was restarted using different parameter values in order to establish whether the results relate to a local or global optimum. By repeating the process from a new starting point a larger area around the optimum may be searched (Blackie and Eeles 1985). When different starting points were used, different end values were encountered due to problems related to the parameter response surface. Sorooshian and Gupta (1995) highlight a number of difficulties associated with the parameter response surfaces that are generic to CRR models. These include the presence of several major regions of attraction into which the

search algorithm may converge, with each of these regions containing numerous local minima. Where parameters exhibit varying degrees of sensitivity a great deal of interaction and compensation may be evident (Sorooshian and Gupta 1995). These obstacles make it very difficult if not impossible for a local search strategy such as the Rosenbrock method to progress towards a global optimum. Such obstacles are displayed in the contour plots produced for selected process parameters where the concept of parameter interaction in the form of long narrow ridges is present in Figure 3a and multimodality is shown in Figure 3b, making progression from these areas very difficult.

Such problems in the response surface limit the identifiability of the process parameters within the HYSIM structure. In addition, this lack of identifiability can be compounded by input data error. Very large parameter ranges can also pose problems

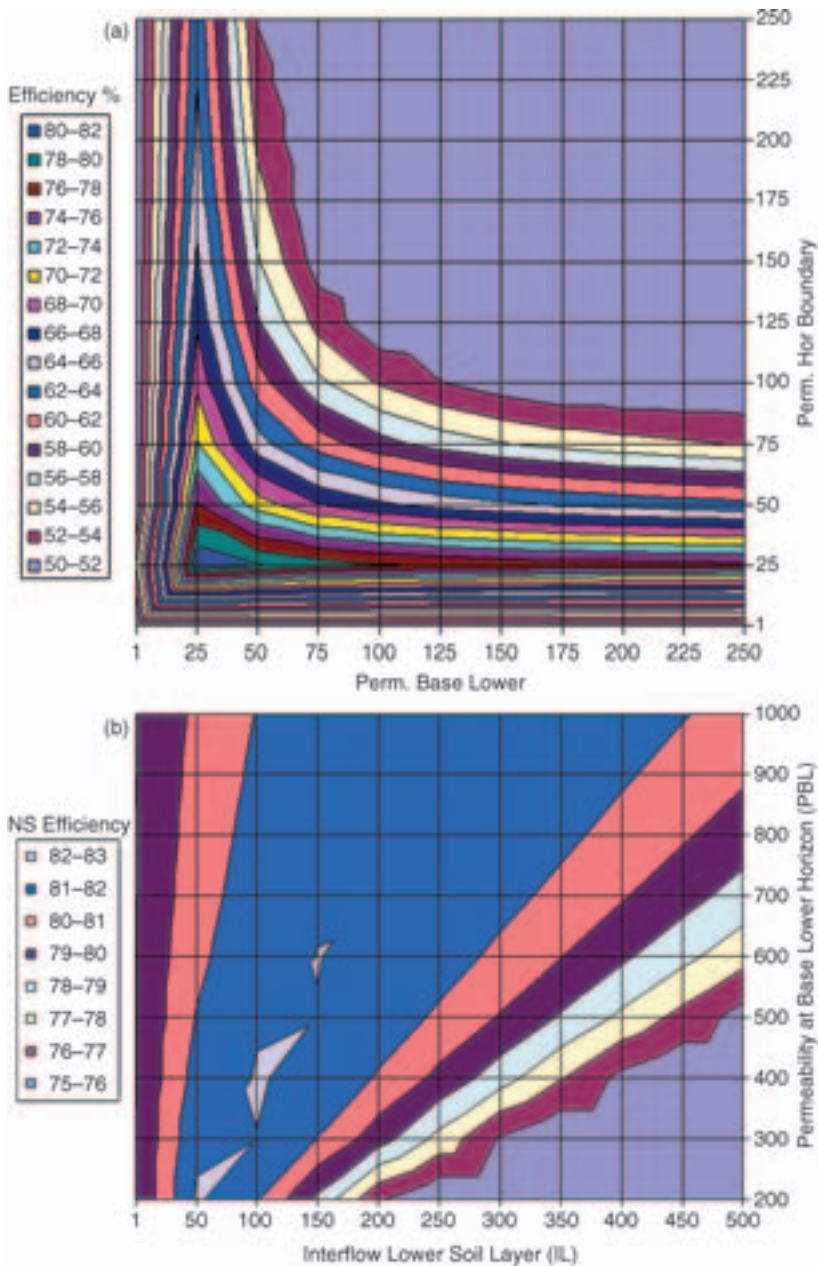


Figure 3 Problems within the parameter response surface. (a) Long narrow ridges, (b) multiple peaks

for parameter identification, especially where source code is not available to automate sampling of the parameter space, as is the case here. When examining the identifiability of the process parameters for the Suir catchment, values were manually incremented across the feasible parameter space for the

calibration period. From Figure 4 it can be seen that there are difficulties in relation to the identifiability of the parameters representing the permeability at the horizon boundary (PHB) and the permeability at the base of the lower horizon (PBL). High values in the efficiency criterion are evident for these parameters

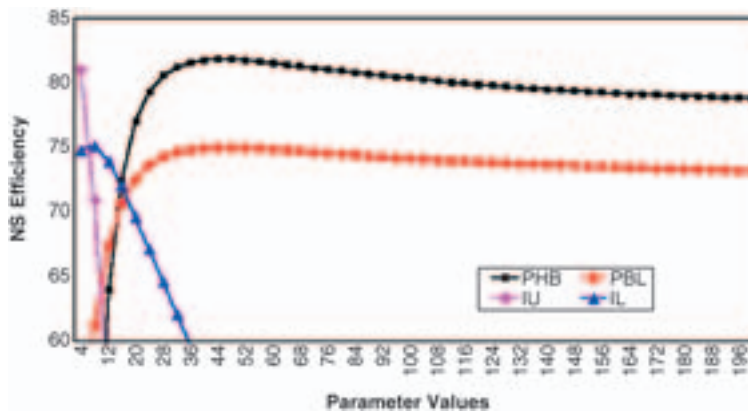


Figure 4 Identifiability of process parameters for the Suir at Clonmel

across a large portion of the parameter space and thus it is possible for unrealistic values to be returned by the optimization procedure. Consequently, no confidence can be placed in the optimum parameter set realized by the Rosenbrock method and uncertainty is introduced to the output from the HYSIM model. As a result, the concept of the existence of an optimal parameter set was rejected and replaced with the concept of equifinality. Equifinality rejects the existence of an optimal parameter set in favour of multiple possibilities for producing simulations that are acceptable (Beven and Freer 2001).

Taking these results into consideration and given the inability to automate the modelling procedure, a technique based on the GLUE methodology is adopted, with MCRS being replaced with Latin Hypercube Sampling. This approach has been widely used in environmental modelling studies. However, examples using lumped conceptual hydrological models are less numerous, with the majority of applications conducted in the context of complex, spatially distributed models (Christiaens and Feyen 2001).

Results

In implementing the GLUE procedure for the Suir at Clonmel, the NS efficiency criterion was adopted as the likelihood measure, with behavioural parameter sets taken as those with efficiency value above 0.7. A uniform distribution was attributed to each process parameter (as proposed by Beven and Freer 2001) and samples were generated using LHS. Based on the analysis of convergence rates, 100 different parameter combinations were generated. These parameter sets were run for the calibration

Table 3 Range from which behavioural parameter sets were sampled

Par. range	Lower	Upper
PHB	6.77	99.22
PBL	10.06	99.67
IU	1.47	48.13
IL	0.70	49.55

period of 1961–90, and of these 50 were retained as behavioural. Efficiency values for the entire calibration series ranged from 0.701 to 0.825. In order to validate these parameter sets a blind simulation was conducted on each set for the validation period 1991–2000. From the 50 behavioural parameter sets obtained during calibration, all were retained as acceptable sets in representing the period 1991–2000. For the validation period, model efficiency ranged from 0.702 to 0.852.

From an analysis of the retained sets from the validation period, it was found that large variations in the parameter values resulted in similar efficiency values. The range of values for each parameter from which acceptable sets were generated is given in Table 3.

The large range in the values of each process parameter, along with the fact that many of the acceptable parameter sets produce a similar efficiency value, is indicative of the problems within the parameter response surface. Such problems are attributable to parameter identifiability and the equifinality of outputs arising from different combinations of model parameters.

Transferability of parameter sets

In order to assess how applicable the HYSIM model is for operational use in climate impact assessment, the transferability of parameter sets between wet and dry periods is assessed for the validation period. Wilby (2005) demonstrated how different calibration periods affect the transferability of parameter sets and concluded that the transferability of model parameters depends on the representativeness of the calibration period. In order to ascertain the representativeness and thus the range of conditions provided by the 1961–90 calibration period, the transferability of parameter sets over wet and dry periods was assessed for the validation years. The ten most skilful parameter sets were extracted and run for both the calibration and validation periods as well as for individual years within the validation period.

The adopted calibration period is extensive (30 years) and covers both wet and dry periods. On a decadal timescale the 1970s are representative of a relatively dry decade, while the 1980s are considered to be wet. The driest year within this period is 1971, with a cumulative annual discharge of 9786 cumecs, whereas the wettest year, occurring in 1982, produced a corresponding cumulative sum of 22 175 cumecs. Therefore the calibration period provides a wide range of flow conditions on which to train the model. Furthermore, the calibration of the model over 30 years allows the incorporation of natural variability into the training period. Kiely (1999) showed that 1975 was representative of a change point in Irish streamflow records associated with a shift in the North Atlantic Oscillation (NAO). The NS value and PBIAS of the ten most skilful parameter sets in capturing the entire calibration period are shown in Table 4. These sets show good results, especially in terms of NS values, with nine out of the ten sets having a PBIAS of less than 10 per cent.

Table 4 also shows the results obtained for the validation period using the best sets derived from calibration. For each there is an increase in the efficiency skill and a decrease in the PBIAS. As mentioned, the ten years used for validation (1991–2000) comprise the warmest decade on the instrumental record and thus provide the best available surrogate for expected future conditions as a result of climate change. The results achieved when transferring the best calibration parameter sets to the validation period indicate that the calibration period provides a representative sample of the range of hydrological conditions for the Suir.

Table 4 Performance of the best ten behavioural parameter sets

Top parameter sets	Calibration		Validation	
	Efficiency	PBIAS	Efficiency	PBIAS
32	82.2	8.6	85.0	4.1
84	82.2	8.9	85.2	4.5
11	82.0	9.1	85.1	4.5
71	81.9	8.8	84.0	4.3
35	81.6	9.6	83.8	4.9
48	81.4	10.1	84.4	5.2
17	81.2	8.9	84.1	3.9
56	81.0	9.6	83.5	5.3
8	80.8	8.4	84.7	4.0
18	80.7	9.5	82.1	5.1

In order to assess the transferability of parameter sets between wet and dry years, the ten most skilful sets were run for individual years within the validation period. The driest years within this period are 1991, 1992 and 1999, while the wettest include the years 1994, 1997, 1998 and 2000. Figure 5 shows the NS values achieved for each parameter set over this period. The model performs well for the validation period, with efficiency values of above 0.70 for all years, with the exception of 1992 and 1997. However, actual error measurements are quite good for these years. With the exception of 1992, the behavioural sets perform quite well for the dry years 1991 and 1999, with the majority of sets providing skill levels of 0.80 or above. Similar skill is also evident for the wet years of 1998 and 2000, with these two years providing some of the largest floods on record. In terms of the warmest years, model results are also quite satisfactory. Variation in model skill for individual years can be derived from differences in the information content of the calibration period (Wilby 2005). However, the good results achieved here for both wet and dry years would suggest that the parameter sets have been trained on a period representing the full variation of streamflow for the Suir. Indeed the results obtained for both wet and dry periods lend confidence to the study of extremes under future climatic conditions.

Uncertainty bounds

Given the uncertainty generated from parametric sources, the 50 behavioural parameter sets were

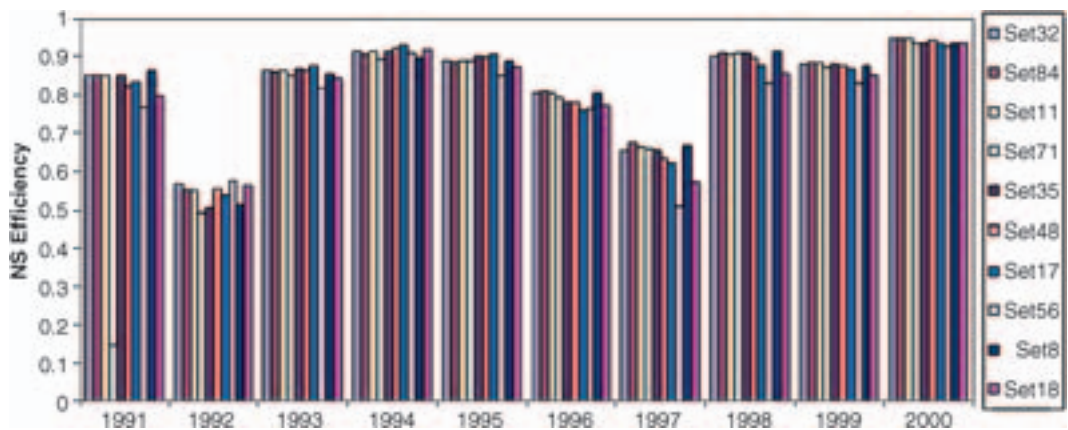


Figure 5 Transferability of parameter sets between years in the validation period

used to calculate the probability distribution of model output for each time step (daily). Uncertainty bounds for model output were generated using the 5th and 95th percentiles. From the bounds constructed for the period of October to December 2000 (Figure 6), the measured data are contained quite well within the 90 per cent uncertainty bounds. This is reassuring, as the peak recorded on the 6 November 2000 is the largest flood peak in the baseline dataset for the Suir at Clonmel. These uncertainty bounds incorporate the epistemic uncertainty derived from data measurement, model structure and parameterization.

Uncertainty bounds were also considered for four selected years within the validation period. Included

are the dry years of 1991 and 1999, as well as the wet years 1994 and 1998 (Figure 7). On a seasonal basis the uncertainty bounds capture the recorded flow quite well during times of high flow, with good results being achieved during winter and spring. However, there are sections of the hydrograph where the recorded flow lies outside the 90 per cent uncertainty bounds. This is most evident for the winter and spring of 1991 where, for various storm events, the uncertainty bounds fail to capture the recorded flow adequately. Given that 1991 lies at the start of the data series the problems for this particular year may be related to the insufficient initialization or spin up of the model. The failure of the uncertainty bounds to capture the recorded flow



Figure 6 Uncertainty bounds for the largest flood peak in the baseline dataset

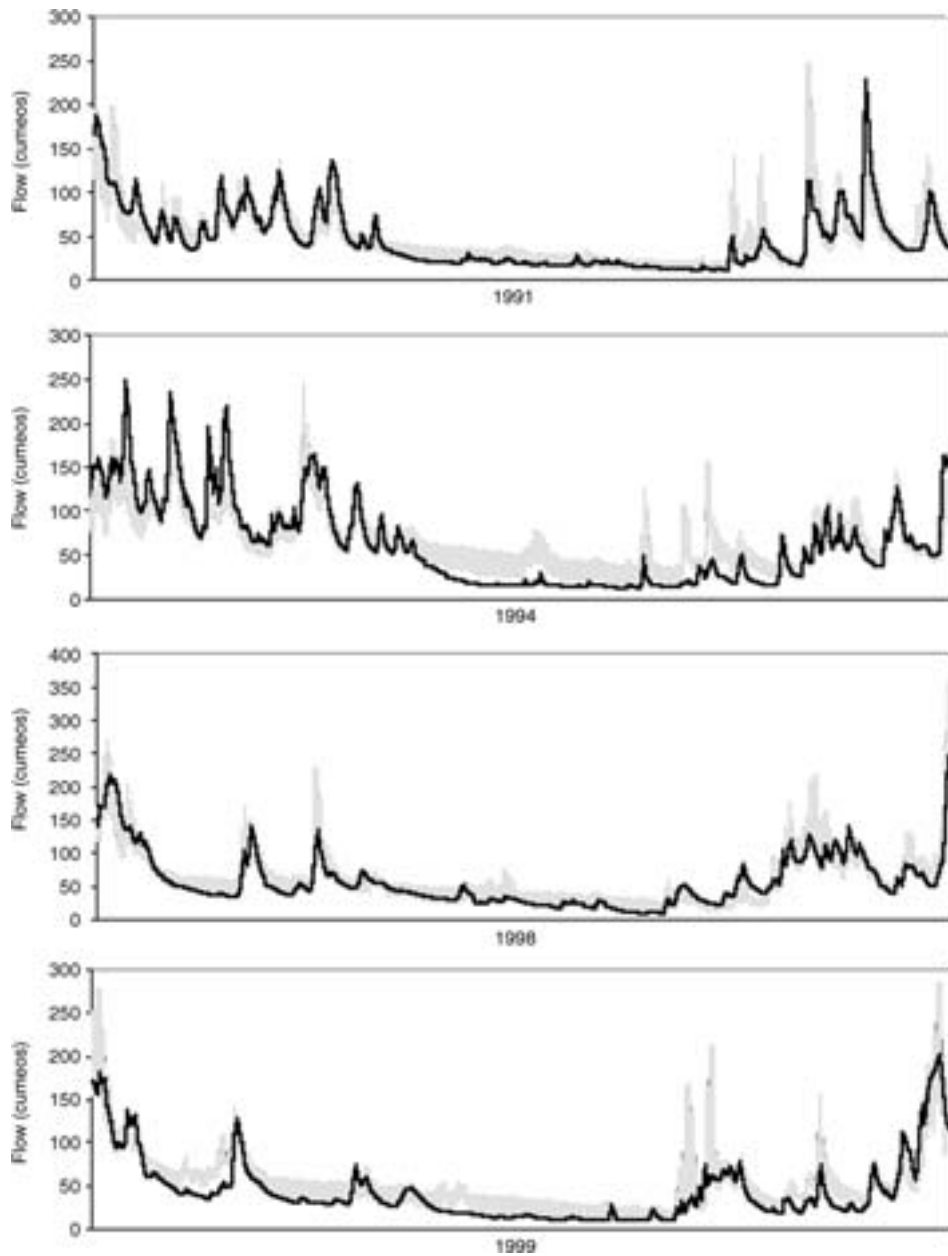


Figure 7 Uncertainty bounds for four selected years

in other areas such as in spring 1999 and summer 1994 may be explained by the use of large scale maps for the calculation of some of the physical parameters such as the soil textural attributes and land use characteristics. Further research is required in this area.

Generally the summer simulations are the most problematic. Typical R^2 values indicated that only around 50 per cent of the variance in observed flows is captured by many of the parameter sets. This fact is evident for the summer periods of both wet and dry years, where summer storms are difficult

to capture. Indeed the inability of CRR models to adequately capture catchment response to storm events following a prolonged dry period seems to be a generic problem. Cameron *et al.* (1999) and Lamb (1999) have indicated the inability of TOP-MODEL to reproduce flood events following an extended dry period. Uhlenbrook and Sieber (2005) have also reported similar difficulties for the process-based catchment model TAC^D, with the response to summer convective storms proving difficult to replicate. Such problems are also evident within the HYSIM model. For each year in Figure 7 the first major storm events after the summer season show the widest uncertainty bounds. Summer storms are also difficult to capture, with some of the larger summer peaks rising beyond the uncertainty bounds. This may be due to a scale mismatch of using such spatially lumped models to represent streamflow from summer convective storms, which can produce high volumes of precipitation in a small space of time. Given the intensity of these events, there is likely to be a larger proportion of effective precipitation as more rainfall is transported by overland flow rather infiltrating the soils. The spatial variability and localized nature of such storm events also provide problems, as rain gauges may fail to capture an event that is present on the flow record, highlighting the importance of high density rain gauge networks.

Conclusion

Large amounts of uncertainty can be derived from the use of CRR models in environmental impact assessment. As such it is important to attempt to analyse and quantify this uncertainty. Such assessments are limited when using off-the-shelf CRR models, as the lack of access to model source code can limit the ability to produce many runs of the model. However, the adoption of an efficient sampling strategy such as Latin hypercube sampling can substantially reduce the number of model runs required to conduct a successful uncertainty assessment. This procedure provides an efficient and feasible sampling methodology in the absence of the ability to consider large Monte Carlo random samples. The uncertainty bounds constructed incorporate the error derived from model structure, data measurement, parameterization and lack of knowledge in the process parameters, and can thus be used to quantify uncertainty in model simulations beyond the baseline calibration

period. The use of multiple parameter sets in future simulations caters for the need to represent change in the form of ranges so as the full span of possible future conditions can be taken onboard. Due to the uncertainty involved in extracting the optimal parameter set within HYSIM the use of a single parameter set derived from automatic calibration may provide misleading results for future hydrological conditions as a result of problems within the parameter space. Finally, given the application of the GLUE methodology, the procedure adopted allows easy updating of the likelihood weights and thus uncertainty bounds as more data become available.

In terms of the representativeness of the calibration period, the transferability of parameter sets highlights the variation in streamflow conditions that the model has been trained on. Given that the validation period provides the closest surrogate possible for future conditions, the good results achieved give greater confidence to the transference of behavioural parameter sets from current to changed climate conditions. However, there is the caveat that land use and soil textural characteristics will remain the same under a changed climate.

This research provides a means by which uncertainty analysis can be conducted for rainfall-runoff models in which the automation of sampling procedures is not possible. The reduced number of samples required for LHS provides a useful alternative to MCRS. Furthermore, a methodology is described in which the possible ranges of sensitive parameters within the HYSIM framework can be reduced in order to avoid the use of unrealistic parameter values. However, uncertainty derived from problems associated with the calibration of CRR models is just one aspect of the cascade of uncertainty involved when assessing climate change impacts. Other sources of uncertainty are derived from the construction of future emissions scenarios, the sensitivity of different Global Climate Models (GCMs), as well as in the method employed in relating large-scale output from GCMs to point sources (e.g. downscaling). The uncertainty derived from rainfall-runoff model calibration remains important and can be comparable in magnitude to the uncertainty arising from future greenhouse gas emission scenarios (Wilby 2005).

Acknowledgements

This research forms part of the Environmental RTDI Programme 2000–2006, developed and managed by the

Environmental Protection Agency (Ireland) and funded by the National Development Programme. Thanks to Professor Rob Wilby for his helpful comments. Thanks are also extended to the referees who provided constructive recommendations.

References

- Arnell N W** 2003 Relative effects of multi-decadal climatic variability and changes in the mean and variability of climate due to global warming: future streamflow in Britain *Journal of Hydrology* 270 195–213
- Arnell N W and Reynard N S** 1996 The effect of climate change due to global warming on river flows in Great Britain *Journal of Hydrology* 183 397–424
- Beven K** 1993 Prophecy, reality and uncertainty in distributed hydrological modelling *Advances in Water Resources* 16 41–51
- Beven K J** 2000 *Rainfall–runoff modelling: the primer* Wiley & Sons Ltd, New York
- Beven K and Binley A** 1992 The future of distributed models-model calibration and uncertainty prediction *Hydrological Processes* 6 279–98
- Beven K and Freer J** 2001 Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology *Journal of Hydrology* 249 11–29
- Blackie J R and Eeles C W O** 1985 Lumped catchment models in **Anderson M G and Burt T P** eds *Hydrological forecasting* Wiley & Sons, New York 311–46
- Brazier R E, Beven K J, Freer J and Rowan J S** 2000 Equifinality and uncertainty in physically based soil erosion models: application of the GLUE methodology to WEPP – The Water Erosion Prediction Project – for sites in the UK and USA *Earth Surface Processes and Landforms* 25 825–45
- Brooks R H and Corey A T** 1964 Hydraulic properties of porous media *Colorado State University Hydrological Paper* 3
- Cameron D S, Beven K J, Tawn J, Blazcova S and Naden P** 1999 Flood frequency estimation for a gauged upland catchment with uncertainty *Journal of Hydrology* 219 169–87
- Charlton R and Moore S** 2003 The impact of climate change on water resources in Ireland in **Sweeney J** ed *Climate change: scenarios and impacts for Ireland* EPA Publication 81–102
- Christiaens K and Feyen J** 2001 Analysis of uncertainties associated with different methods to determine soil hydraulic properties and their propagation in the distributed hydrological MIKE SHE model *Journal of Hydrology* 246 63–81
- Cunnane C and Regan S** 1994 Hydrology and freshwater resources in **McWilliams B E** ed *Climate change: studies of the implications for Ireland* Department of the Environment, Stationery Office, Dublin 89–108
- Gardiner M J and Radford T** 1980 *Ireland general soil map* National Soil Survey, Dublin
- Geological Survey of Ireland** 2003 *Draft national aquifer map*
- Iman R L and Helton J C** 1985 *A comparison of uncertainty and sensitivity analysis techniques for computer models* Technical Report SAND84-1461 Sandia National Laboratories, Albuquerque NM
- Iman R L and Helton J C** 1988 An investigation of uncertainty and sensitivity analysis techniques for computer models *Risk Analysis* 8 71–90
- Khu S T and Werner M G F** 2003 Reduction of Monte-Carlo simulation runs for uncertainty in hydrological modelling *Hydrology and Earth Systems Sciences* 7 680–92
- Kiely G** 1999 Climate change in Ireland from precipitation and stream flow observations *Advances in Water Resources* 23 141–51
- Lamb R L** 1999 Calibration of a conceptual rainfall–runoff model for flood frequency estimation by continuous simulation *Water Resources Research* 35 3103–14
- Legates D R and McCabe G J** 1999 Evaluating the use of ‘goodness of fit’ measures in hydrologic and hydroclimatic model validation *Water Resources Research* 35 233–41
- Lenhart T, Eckhardt K, Fohrer N and Frede G H** 2002 Comparison of two different approaches of sensitivity analysis *Physics and Chemistry of the Earth* 27 645–54
- Manley R E** 1977 The soil moisture component of mathematical catchment simulation models *Journal of Hydrology* 35 341–56
- Manley R E** 1978 Calibration of hydrological model using optimization technique *Journal of the Hydraulics Division, ASCE* 104 HY2 189–202
- Manley R E** 1993 *HYSIM reference manual* R E Manley Consultancy, Cambridge
- McCarthy J J, Canziani O F, Leary N A, Dokken D J and White K S** eds 2001 *Climate change 2001: impacts, adaptation, and vulnerability* Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) Cambridge University Press, Cambridge
- McKay M D, Conover W J and Beckman R J** 1979 A comparison of three methods for selection values of input variables in the analysis of output from a computer code *Technometrics* 2 239–45
- Melching C S** 1992 A comparison of methods for estimating variance of water resources model predictions in **Kuo J-T and Lin G-F** eds *Stochastic hydraulics, Proceedings of the Sixth International Association for Hydraulic Research Symposium on Stochastic Hydraulics* Taipei, Taiwan Water Resources Publications
- Melching C S** 1995 Reliability estimation in **Singh V P** ed *Computer models of watershed hydrology* Water Resources Publications, Highlands Ranch CO
- O’Sullivan G** ed 1994 *CORINE land cover project* Ireland Project Report December 1994 Ordnance Survey of Ireland and Ordnance Survey of Northern Ireland Belfast, Dublin
- Oberkampf W L, Deland S M, Rutherford B M, Diegert K V and Kenneth A F** 2002 Error and uncertainty in modelling and simulation reliability *Engineering and System Safety* 75 333–57

- Pilling C G and Jones J A A** 1999 High resolution climate change scenarios: implications for British runoff *Hydrological Processes* 13 2877–95
- Pilling C G and Jones J A A** 2002 The impact of future climate change on seasonal discharge, hydrological processes and extreme flows in the Upper Wye experimental catchment, mid-Wales *Hydrological Processes* 16 1201–13
- Sefton C E M and Boorman D B** 1997 A regional investigation into climate change impacts on UK streamflows *Journal of Hydrology* 195 26–44
- Sorooshian S and Gupta V K** 1995 Model calibration in **Singh V P** ed *Computer models of watershed hydrology* Water Resources Publications, Highlands Ranch CO
- Uhlenbrook S and Sieber A** 2005 On the value of experimental data to reduce the prediction uncertainty of a process-orientated catchment model *Environmental Modelling and Software* 20 19–32
- Wagner T** 2003 Evaluation of catchment models *Hydrological Processes* 17 3375–8
- Wilby R L** 2005 Uncertainty in water resource model parameters used for climate change impact assessment *Hydrological Processes* 19 3201–19
- Yu P S, Yang T Ch and Chen S J** 2001 Comparison of uncertainty analysis methods for a distributed rainfall runoff model *Journal of Hydrology* 244 43–59