The Role of Uncertainty in Design of Objective Functions

B.F.W. Croke

Integrated Catchment Assessment and Management Centre, The Fenner School of Environment and Society, and Department of Mathematics, The Australian National University, Canberra, Australia.
Email: barry.croke@anu.edu.au

Keywords: *Model evaluation, uncertainty, objective functions, Prediction in Ungauged Basins*

EXTENDED ABSTRACT

The role of performance indicators is to give an accurate indication of the fit between a model and the system being modelled. This is done through comparison with observations from a particular viewpoint. Ideally, the performance indicator(s) employed should reflect the purpose of the modelling exercise (i.e. indicate how well a model answers the specific question being asked of it). Consequently a standard performance indicator may not always be the correct choice; for example, a study investigating low flows should not necessarily employ the same performance indicator as one investigating flood events.

All measurements have an associated uncertainty which determines the significance that should be given to the measurement. Therefore, by definition, performance indicators should take into account errors in the observed quantities being modelled as well as in the model predictions (due to errors in inputs, model parameters and model structure). Failure to adequately account for variations in the errors in the observed and modelled quantities means that the objective function is only giving a measure of how well the modelled values represent the observed values, not how well the model is representing the system being modelled (unless the uncertainties in the observed and modelled values are sufficiently small).

Commonly-used objective functions comparing observed and modelled flows (e.g. the Nash-Sutcliffe efficiency - NSE) do not explicitly take into account the uncertainties in the model input, or in the recorded flows. Rather, the uncertainty is often assumed to be homoscedastic (i.e. constant irrespective of magnitude). As a result high flow events, which are often the most uncertain, are given too much weight. There have been several attempts to overcome this limitation, from not using the highest n% of observed flows, to transforming the flow prior to calculating the NSE. Transforming the flow still assumes a particular distribution of uncertainties (for example, when using the logarithm of flows, the uncertainties are assumed to be a constant multiple of the flow). While this is better than the linear form of the NSE, the relative uncertainties in the high flow values are likely to be underestimated (e.g. when the rating curve is a power law, the log transformed NSE assumes negligible uncertainty in the power).

This paper addresses how objective functions can be modified to include the influence of uncertainty in model inputs and outputs. This includes discussion of the propagation of errors through functions (including the influence of thresholds) and the combination of errors from different inputs.

Modifications to commonly used objective functions (RMSE, NSE, χ^2, r^2) are presented using both the ratio and optimal weighted average approaches. It is recommended that the optimal weighted average approach is used for most applications, with the advantage of both ensuring that the uncertainties are homoscedastic as well as producing the optimal signal-to-noise in the resulting objective function.. The exception is for situations where the magnitude of the uncertainty is important; in which case the normalisation employed in the optimal weighted average approach should not be used.

How the uncertainties in streamflow data can be handled is also discussed, using a power law formulation for the rating curve. While the power law form is used as an example, the technique discussed is applicable for any form employed, though care needs to be taken regarding any thresholds introduced.

Finally, an example of the technique is presented using synthetic streamflow data with induced uncertainty in the parameters of the rating curve, showing that the modified NSE giving reduced uncertainty in the parameter values of the IHACRES non-linear module. Further work is underway to include the effect of propagation of uncertainty in the rainfall through the model.

1. INTRODUCTION

Most goodness-of-fit indicators do not take into account variations in the errors (i.e. they assume that the errors are homoscedastic). If the uncertainties in the inputs (including flow) are not adequately represented in the objective function/s, then the evaluation of the model's performance may be biased, resulting in sub-optimal parameter sets (tracking the uncertainties in the data rather than the catchment response), and increased uncertainty in any model parameter regionalisation scheme for assisting with estimating flows in ungauged basins. This strongly suggests that all datasets need to include a realistic estimate of their uncertainty, and that this uncertainty needs to be taken into account when designing an objective function. This is particularly the case for stream gauges that do not have good control structures – a common problem in parts of Australia as well as in developing countries.

The uncertainties in the inputs can be included into objective functions either analytically (by modifying the functional form of the objective function) or stochastically. For linear systems, only the standard deviation is required for analytical approaches. For non-linear systems, or when there are thresholds included in the model, both analytical and stochastic approaches require the information on the distributions of the uncertainties.

The generally accepted standard objective function in hydrology is the Nash-Sutcliffe efficiency (NSE - Nash and Sutcliffe, 1970) and is referred to in the literature in a number of ways, including *NSE*, *E*, R^2 and *D*. Following the convention used by Nash and Sutcliffe, the NSE will be referred to as R^2_{NS}. with the *NS* subscript added for clarity.

While R^2_{NS} is a widely accepted performance indicator in hydrology, it assumes that the uncertainties are homoscedastic – i.e. that the magnitude of the uncertainties is independent of the quantity being measured. Thus it does not take into account errors in the observed or modelled flows which are highly heteroscedastic. This is a major limitation as R^2_{NS} is dominated by the mismatch between observed and modelled values at high flows, even though these have the highest uncertainty due primarily to uncertainty in the rating curve. Chiew and Siriwardena (2005) opted to ignore the highest five flow values in calculating their objective functions in order to minimise the impact of the errors in the extreme high flows.

One way to address this issue is through Monte Carlo techniques, and while this is a simple approach to the problem, there is a significant increase in the run-time. Another option is to modify an objective function through either transforming the data (e.g. using the logarithm) or through introduction of weights so that the heteroscedasticity of the uncertainties is reduced. Lichty *et al*. (1968) used the logarithm of the observed and modelled flow peaks in calculating the objective function. The latest version of the IHACRES rainfall-runoff modelling software (Croke et al., 2006) allows a range of transformations to aid in calibration and testing of the model. However, it should be remembered that such transformations are *ad hoc* and are based on assumptions of the nature of the heteroscedasticity rather than from an analysis of the errors (e.g. a log transformation assumes that the uncertainties are approximately constant in a multiplicative sense.

Objective functions can be further modified for specific purposes by introducing a significance for each time step. For example, for flood studies, higher significance might be placed on high flow events, whereas studies focusing on baseflows may need to put increased significance on the low flows. The introduction of a significance term in the weights is subjective, and will not be explored further in this paper.

Flow measurements are based on rating curves to convert stage height (*h*) into discharge rate (*Q*), typically using a power law form:

$$Q = a\left(h - h_o\right)^b \tag{1}$$

which has 3 parameters *a*, *b* and h_o. A rating curve often comprises segments, each of the form shown in equation 1, due primarily to the complex nature of the cross section. The log transformation assumes that the *b* and h_o parameters do not contribute significantly to the uncertainty and as a result, the uncertainties at very high and very low flows are underestimated.

A comparison of the performance of different models using objective functions is especially difficult due to the difference in the number of parameters, and data used. In terms of the number of parameters, one approach is to scale the objective function by the degrees of freedom; for example, (n-1)/(n-p-1), where *n* is the sample size, and *p* is the number of parameters. Care should be taken regarding the value of *n*. The naïve approach would be to use the number of time steps. However, this doesn't take into consideration the amount of information contained in the time series. For a catchment with relatively few rainfall events in a given period, the information contained within the time series will be less than that for a catchment with more frequent rainfall events. Thus n should reflect more the number of events (n_e) than the number of time steps (n_t). A better estimate for n would logically be αn_e, where α is an estimate of the amount of information the typical event will provide.

Another naïve view is that models which utilise additional datasets are expected to perform better than models that use minimal data. This is true providing that the information contained in the additional datasets is significantly larger than the uncertainty that is added. It is important to consider the information-to-noise ratio when evaluating the advantages of using additional datasets. For example, Andréassian *et al.* (2007) presented results of a study of the effectiveness of subdividing a catchment and running their model on each part separately, concluding that the additional information on the spatial distribution of rainfall had negligible impact on the model's performance (information-to-noise ratio was approximately 1).

Aside from time series based forms, performance indicators can also be based on transformations of the observed and modelled values. Examples of these include cumulative probability distributions (flow duration curves used in hydrology), cross correlation functions and power spectra (e.g. Croke, 2005). Such performance indicators have the same problems discussed above for the untransformed series, but can yield information useful in assessing the performance of a model, or in comparing the performance of two or more models. These transformations can also be useful in developing graphical performance indicators.

An alternative may be to adopt a wavelet approach, where the fit to the data is measured for a range of scales across all available time periods, for example, the DYNIA approach of Wagener *et al.* (2003). This produces a 2-D image representation of the model performance, thus giving the user much more information at the cost of potentially making comparisons between models more difficult. In all cases, objective functions should take into account the errors in the quantities being compared.

This paper will investigate the use of weights in modifying a range of objective functions. The following definitions will be used:

- observed value at timestep i: $x_{o,i}$
- modelled value at timestep i: $x_{m,i}$
- model residual at timestep i: $e_i = x_{o,i} - x_{m,i}$
- observed deviation from mean: $d_{o,i} = x_{o,i} - \bar{x}_o$
- modelled deviation from mean: $d_{m,i} = x_{m,i} - \bar{x}_m$

2. THEORY

Typically, an objective function gives an aggregated measure of how well a model matches the data. The objective function may produce a single value (e.g. Nash-Sutcliffe efficiency (NSE), root mean square error (RMSE), Chi-squared) or can be constructed to give a measure of how the fit changes through the data set (e.g. Lane, 2007). Typically, objective functions assume that the uncertainties are homoscedastic; that is, the uncertainties are independent of the value. As discussed above, this is rarely the case.

While the uncertainty in a calculated value depends on the uncertainty in all the values used in the calculation, only the uncertainties that are significant contributors need to be quantified. Consequently, the first step is to provide an order of magnitude estimate of the various uncertainties. The formula used is then studied to determine which of the inputs will significantly contribute to the uncertainty in the calculated value. The uncertainties of these quantities then need to be more accurately estimated.

If a calculation is likely to be significantly dependent on the estimated uncertainty of the inputs (e.g. SIMEX, Sharma and Shahadat, 2007), then the influence of the probable range of uncertainties needs to be considered. If this significantly affects the result, then this needs to be considered. While consideration of the uncertainty in the uncertainty may be regarded as frivolous, if it significantly affects the result, it needs to be considered (though not necessarily calculated).

2.1. Error propagation

Generally, the propagation of an uncertainty through a function is considered only at the first derivative – that is, the response is assumed to be linear over the interval of interest. The Taylor series expansion of $y=f(x)$ about the mean of x is:

$$y = f(\bar{x}) + \sum_n \left. \frac{d^{(n)} f}{dx^{(n)}} \right|_{\bar{x}} \frac{(x-\bar{x})^n}{n!} \tag{2}$$

The number of terms needed to reproduce the function f will depend on the degree of non-linearity, the accuracy needed in the expansion and the range of x values over which the expansion is required to meet that accuracy. Considering an ensemble of N measurements of x and considering only the first derivative gives:

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2$$

$$y_i \approx f(\bar{x}) + \left.\frac{df}{dx}\right|_{\bar{x}} (x_i - \bar{x}), \qquad f(\bar{x}) \approx \bar{y} \tag{3}$$

$$\therefore \Delta y \approx \left.\frac{df}{dx}\right|_{\bar{x}} \Delta x$$

However, if $f(x)$ is non-linear over the range Δx then we may need to either consider higher order terms, or use an effective gradient over the region of interest (e.g. fitting a line to the function over the interval $x \pm \Delta x$). Considering the second derivative, the mean value of y will not be given by $f(\bar{x})$, rather the mean will be given by:

$$\bar{y} \approx f(\bar{x}) + \left.\frac{d^2 f}{dx^2}\right|_{\bar{x}} \frac{1}{2}\sigma_x^2 \tag{4}$$

assuming N is sufficiently large. Consequently, the variance of y will be given by:

$$\sigma_y^2 \approx \left(\left.\frac{df}{dx}\right|_{\bar{x}}\right)^2 \sigma_x^2 + \left.\frac{df}{dx}\right|_{\bar{x}} \left.\frac{d^2 f}{dx^2}\right|_{\bar{x}} {}^3M_x - \frac{1}{4}\left(\left.\frac{d^2 f}{dx^2}\right|_{\bar{x}}\right)^2 \left\{\sigma_x^4 - {}^4M_x\right\} \tag{5}$$

$${}^nM_x = \sum_{i=1}^{N}\left\{\frac{(x_i - \bar{x})^n}{N-1}\right\}$$

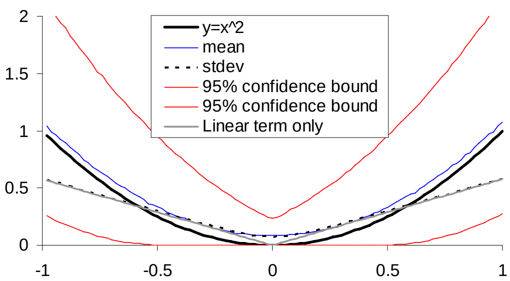

Figure 1. Monte Carlo estimation of the mean, standard deviation and 95% confidence bounds for $y=x^2$ when the uncertainty in x is a uniform distribution with width 1 (n=8000 for each point).

The linear case (equation 3) applies for any distribution and any definition of the uncertainty (e.g. 1σ, 95% confidence). However, introducing the second derivative term in the Taylor expansion introduces higher moments (3rd and 4th moment about the mean), thus requiring information regarding the distribution of the uncertainties (e.g. uniform, normal, logarithmic, etc). Whether the higher order terms are needed depends on both the degree of non-linearity of the function f and the magnitude of the uncertainty in x (see Figure 1).

The presence of a threshold inside the uncertainty bounds will require separate treatment of each side of the threshold, with the results then combined to give the final distribution. This is the case where the relationship at the threshold is continuous or discontinuous as the Taylor series expansion will not be valid on both sides of the threshold. The existence of thresholds may mean that a Monte Carlo approach is the only viable option (see Figure 2).

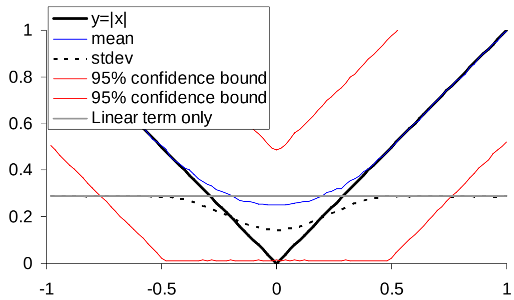

Figure 2. Monte Carlo estimation of the mean, standard deviation and 95% confidence bounds for $y=|x|$ when the uncertainty in x is a uniform distribution with width 1 (n=8000 for each point).

2.2. Combination of errors

Let z be the sum of x and y. The variance of z is:

$$\sigma_z^2 = \sigma_x^2 + 2\sigma_{xy} + \sigma_y^2 \tag{6}$$

where σ_{xy} is the covariance of x and y. The uncertainty in z is then given by:

$$\Delta z^2 = \Delta x^2 + 2\varsigma_{\Delta x \Delta y}\Delta x \Delta y + \Delta y^2$$

$$\varsigma_{\Delta x \Delta y} = \frac{\sum_{i=1}^{n}\varepsilon_{x,i}\varepsilon_{y,i}}{\sqrt{\sum_{i=1}^{n}\varepsilon_{x,i}^2 \sum_{i=1}^{n}\varepsilon_{y,i}^2}} \tag{7}$$

$$\Delta x = \sum_{i=1}^{n}\varepsilon_{x,i}^2 \Big/ (N-1)$$

where $\zeta_{\Delta x \Delta y}$ is the covariance of the uncertainties (not the actual values of x and y). When multiple measures of x and y are not available, an estimate of $\zeta_{\Delta x \Delta y}$ is needed based on an understanding of the measurements made (e.g. two independent observations of the same quantity are likely to have a covariance of 0, even though the covariance of the observations will be close to 1).

3. MODIFIED FORM OF OBJECTIVE FUNCTIONS

As has been stated, objective functions need to inform how well the model is fitting the system being modelled. In the presence of heteroscedastic uncertainties, an objective function must either account for the heteroscedasticity, or the values used in the objective function must be transformed so that the uncertainties are homoscedastic. Two approaches will be considered here.

3.1. Ratio method

This method of modifying the objective function is based on scaling the model residuals by the uncertainty in the residual e_i. This means that the uncertainty in the scaled residual is constant (homoscedastic) provided that the estimated uncertainty reasonably accounts for the actual variation in uncertainty. If the magnitude of the estimated uncertainty is sufficiently correct, then the uncertainty in the scaled residual is approximately 1. While it is not necessary that the magnitude of the uncertainty be correctly estimated in order for the scaled residuals to be homoscedastic, for comparison between models and/or sites, a consistent estimate of the uncertainty is needed.

3.2. Optimal weighting method

An alternative to the ratio method is to adopt optimal weighting in order to obtain the best signal-to-noise ratio in the objective function.

The optimal weighting of the average a of 2 values b and c with uncertainties Δb and Δc is given by:

$$a = \frac{\omega_b b + \omega_c c}{\omega_b + \omega_c},$$
$$\omega_b = \frac{b}{(\Delta b)^2}, \quad \omega_c = \frac{c}{(\Delta c)^2} \tag{8}$$

Table 1. Modified versions of RMSE, NSE, χ^2 and Coefficient of determination

Objective function	Formula	Ratio method	Optimal weighting
RMSE	$\sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}}$	$\sqrt{\frac{\sum_{i=1}^{n} e_i^2 / \alpha_i}{n}}$	$\sqrt{\sum_{i=1}^{n} \frac{1}{\alpha_i \gamma_{e_i^2}} e_i^2 \Big/ \sum_{i=1}^{n} \frac{1}{\alpha_i \gamma_{e_i^2}}}$
NSE	$1 - \frac{\sum_i^n e_i^2}{\sum_i^n d_{o,i}^2}$	$1 - \frac{\sum_i^n e_i^2 / \alpha_i}{\sum_i^n d_{o,i}^2 / \lambda_i}$	$1 - \left[\frac{\sum_i^n \frac{1}{\alpha_i \gamma_{e_i^2}} e_i^2}{\sum_i^n \frac{1}{\lambda_{o,i} \gamma_{f_{o,i}^2}} d_{o,i}^2}\right]\left[\frac{\sum_{i=1}^n \frac{1}{\lambda_{o,i} \gamma_{f_{o,i}^2}}}{\sum_{i=1}^n \frac{1}{\alpha_i \gamma_{e_i^2}}}\right]$
χ^2	$\sum_{i=1}^{n} \frac{e_i^2}{x_{o,i}}$	$\sum_{i=1}^{n} \xi_i \frac{(e_i)^2}{x_{o,i}}$	$\left(\sum_{i=1}^{n} \xi_i^2 \left(\frac{(e_i)^2}{x_{o,i}}\right)^2\right) \Big/ \sum_{i=1}^{n} \xi_i^2 \frac{(e_i)^2}{x_{o,i}}$
Coefficient of determination r^2	$\left[\frac{\sum_{i=1}^{n} d_{o,i} d_{m,i}}{\sqrt{\sum_{i=1}^{n} d_{o,i}^2}\sqrt{\sum_{i=1}^{n} d_{m,i}^2}}\right]^2$	$\left[\frac{\sum_{i=1}^{n} \left(d_{o,i}/\sqrt{\lambda_{o,i}}\right)\left(d_{m,i}/\sqrt{\lambda_{m,i}}\right)}{\sqrt{\sum_{i=1}^{n} d_{o,i}^2/\lambda_{o,i}}\sqrt{\sum_{i=1}^{n} d_{m,i}^2/\lambda_{m,i}}}\right]^2$	$\left[\frac{\sum_{i=1}^{n} d_{o,i} d_{m,i}}{\sqrt{\sum_{i=1}^{n} \frac{1}{\lambda_{o,i}\gamma_{f_{o,i}^2}} d_{o,i}^2}\sqrt{\sum_{i=1}^{n} \frac{1}{\lambda_{m,i}\gamma_{f_{m,i}^2}} d_{m,i}^2}}\right]^2$

$$\alpha_i = (\Delta e_i)^2 = (\Delta x_{o,i})^2 + (\Delta x_{m,i})^2, \qquad \lambda_i = (\Delta d_i)^2 = (\Delta x_{o,i})^2 + (\Delta \bar{x}_o)^2 + (\Delta x_{m,i})^2 + (\Delta \bar{x}_m)^2$$

$$\xi_i = \frac{1}{\sqrt{\left(\left[1 - \left(\frac{x_{m,i}}{x_{o,i}}\right)^2\right]\Delta x_{o,i}\right)^2 + \left(\left[2 - \frac{2x_{m,i}}{x_{o,i}}\right]\Delta x_{m,i}\right)^2}}, \qquad \gamma_{x_i^2} = 1 + \frac{\left.\frac{df}{dx}\right|_{\bar{x}} \left.\frac{d^2 f}{dx^2}\right|_{\bar{x}} {}^3M_x + \frac{1}{4}\left(\left.\frac{d^2 f}{dx^2}\right|_{\bar{x}}\right)^2 \left\{\sigma_x^4 - {}^4M_x\right\}}{2x\Delta x}$$

For objective functions that use the sum of square residual (e.g. RMSE, NSE), this differs from the ratio method only through the introduction of the normalisation term (ω_b + ω_c - see Table 1). The advantage of the normalisation term is that the mean value of the objective function is not affected by the introduction of the weights. The disadvantage is that the resulting objective function is sensitive only to the variability in the errors, and not the mean error over the entire data period (if weights are equal, then equation 8 reduces to a standard mean).

The result is that the optimal weighting approach gives a modified objective function that is insensitive to the scale of the weighting factors. Thus, only the variation through the data series is needed. Thus the optimal weighting approach only requires an estimate of the relative uncertainty. Consequently the optimal weighting approach is easier to implement, but cannot be used to discriminate between different models.

Objective functions that use absolute values (e.g. Legates and McCabe, 1999), need special care (see Figure 2). Which approach is better (i.e. ratio or optimal weighting) will vary according to what the objective function is attempting to measure, but generally, the optimal weighting method is preferred.

4. UNCERTAINTY IN STREAMFLOW

Care must be taken when evaluating the uncertainty in recorded streamflow values. estimation of streamflow is often based on measurement of stage height and a rating curve relating stage height with discharge. Assuming a power law form for the rating curve:

$$Q_{o,i} = a(h - h_o)^b \tag{9}$$

Ignoring higher order terms, the uncertainty in the observed flow is given by:

$$\begin{aligned}\left(\frac{\Delta Q_{o,i}}{Q_{o,i}}\right)^2 &= \left(\frac{\Delta a}{a}\right)^2 + \left((h - h_o)\Delta b\right)^2 \\ &+ \left(\frac{b}{h - h_o}\right)^2 \left[(\Delta h)^2 + (\Delta h_o)^2\right]\end{aligned} \tag{10}$$

assuming that the uncertainties in a, b, h, and h_o are independent (there are likely to be significant covariances between a, b, and h_o). If the uncertainty in the rating curve is sufficiently large (often the case), then most of the information contained in the shape of the stage height data is masked by the uncertainty in the rating curve. This can be overcome by considering the ratio between observed flows at timesteps i and j:

$$r_{o,ij} = \frac{Q_{o,j}}{Q_{o,i}} = \left(\frac{h_j - h_o}{h_i - h_o}\right)^b \tag{11}$$

Assuming that the true values of the parameters a, b, and h_o do not vary significantly between timesteps i and j, the uncertainty in the ratio $r_{o,ij}$ is:

$$\begin{aligned}\left(\frac{\Delta r_{o,ij}}{r_{o,ij}}\right)^2 &= \left(\ln\left(\frac{h_j - h_0}{h_i - h_0}\right)\Delta b\right)^2 \\ &+ \left(b\frac{(h_j - h_0)(h_j - h_i)}{(h_i - h_0)^3}\Delta h_0\right)^2 \\ &+ \left(b\frac{(h_j - h_0)^2}{(h_i - h_0)^3}\Delta h_i\right)^2 \\ &+ \left(b\frac{h_j - h_0}{(h_i - h_0)^2}\Delta h_j\right)^2\end{aligned} \tag{12}$$

If the $h_j \sim h_i$, then $R_{o,i}$=1, and the first two terms in equation 12 can be ignored and the uncertainty in the ratio depends only on the uncertainty in the stage heights for the two timesteps. For high values of $r_{o,ij}$, the uncertainties in b and h_0 become important, but the uncertainty in a does not contribute (compare with the log-transformed NSE, which assumes that all the error is in the a).

5. APPLICATION

To test the use of the optimal weighting average approach, a Monte Carlo trial of the influence of heteroscedatic uncertainty in synthetic flows has been carried out using the IHACRES rainfall-runoff model (Croke *et al*. 2006), using rainfall timeseries for the Murrindindi River catchment, located in northern Victoria, Australia. The "observed" streamflow was generated using the non-linear parameter values: c=100; τ_w = 2; f = 0; l=0 and p=1, and linear module parameter values: τ_q = 1.2; τ_s = 60; v_s=0.3. Uncertainty was introduced into the "observed" flows using a uniform distribution scaled by a separate preset magnitude of uncertainty for each timestep. In order to minimise the effect of uncertainty in the parameter values, all model parameters were fixed at the values used to derive the original flow time series, except for c and τ_w.

Table 2 shows that the modified NSE gave a much smaller standard deviation between the 100 trials for both parameters being tested. Furthermore, the extreme values were much closer to the nominal values for the modified NSE.

Table 2. Calibrated parameter values for *c* and τ_w using the standard NSE objective function, and the modified NSE (using the optimal weighted average method).

	Standard NSE		Modified NSE	
	τ_W	*c*	τ_W	*c*
mean	2.03	102.7	2.04	101.6
stdev	0.36	12.8	0.11	3.83
max	2.79	129.5	2.31	112.7
min	1.58	79.9	1.81	93.3

Further work is needed in order to propagate the uncertainty in climate data (particularly rainfall) through the model to obtain an estimate of the uncertainty in the modelled flow. This must also include the influence of uncertainty in parameter values. It should be noted that during calibration, some model parameters may not have any uncertainty. This is the case for the non-linear module parameters in the IHACRES modelling methodology. The non-linear module parameters do however have uncertainty during simulation.

6. CONCLUSION

Indicators of model performance must take into account the uncertainties in the quantities being compared in order to give an accurate appraisal of the model's ability to represent the behaviour of the system. Most, if not all, objective functions used either assume homoscedastic uncertainties or make some assumption about the heteroscedasticity in the uncertainties. The nature of data available for rainfall-runoff models means that, generally, these assumptions do not hold, and as a result, the objective functions are poor measures of a model's performance. The suggestion made here is that the objective functions be modified to explicitly include estimates of the uncertainty, preferably using the optimal weighted average approach.

7. REFERENCES

Andréassian, V. Bourqui, M., Loumagne, C. and Thibault, M. (2007). Reducing the uncertainty of streamflow simulation by accounting explicitly for the spatial variability of rainfall in a lumped rainfall-runoff model. Presented in Session HW2004, IUGG General Assembly, Perugia, Italy, July 2007.

Chiew, F.H.S. and Siriwardena, L., 2005. Estimation Of SIMHYD Parameter Values For Application In Ungauged Catchments. In Zerger, A. and Argent, R.M. (eds) *MODSIM 2005 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 2904-2910. ISBN: 0-9758400-2-9. http://www.mssanz.org.au/modsim05/papers/chiew_2.pdf

Croke, B.F.W., Land use impacts on hydrologic response in the Mae Chaem catchment, Northern Thailand, Proceedings of the 2005 International Conference on Simulation and Modelling. V. Kachitvichyanukul, U. Purintrapiban, P. Utayopas, eds., Bangkok, Thailand, January 17-19, 2005, pp434-439, http://www.mssanz.org.au/simmod05/papers/C5-03.pdf.

Croke, B.F.W., Andrews, F., Jakeman, A.J., Cuddy, S.M. and Luddy, A. "IHACRES Classic Plus: A redesign of the IHACRES rainfall-runoff model" Environmental Modelling and Software, 21 (2006) 426-427.

Lane, S.N. (2007), Assessment of rainfall-runoff models based upon wavelet analysis, *Hydrological Processes*, 21, 586-607 (published on-line 29 August, 2006 DOI: 10.1002/hyp.6249).

Legates, D.R. and McCabe, G.J. (1999), Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimaic model validation, *Water Resources Research*, 35(1), 233-242.

Lichty, Dawdy and Bergmann (1968). Rainfall-runoff model for small basin flood hydrograph simulation. In *The use of analog and digital computers in hydrology, Vol II*, Int. Ass. sci. Hydrol., Publ 81, pp356-367.

Nash, J.E. and J.V. Sutcliffe (1970), River flow forecasting through conceptual models, i, a discussion of principles, *Journal of Hydrology*, 10, 282-290.

Sharma, A., and Shahadat, C. (2007). A simulation based approach for mitigating parameter bias in hydrological models due to uncertainty in inputs. Presented in Session HW2004, IUGG General Assembly, Perugia, Italy, July 2007.

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