Hydrology and Earth System Sciences Discussions

Interactive
Comment

Interactive comment on "Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation" *by* Y. Tang et al.
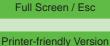
R. Clarke (Editor)

clarke@iph.ufrgs.br

Received and published: 9 January 2007

Editor's concluding comment.

The issue of sensitivity, and how to evaluate the uncertainty in model predictions, is likely to become ever more important, particularly where, as is increasingly the case, hydrological models are linked to models of atmospheric behaviour for the purpose of gaining insight into the hydrological effects of changing climate conditions. The material in the paper, comparing four methods for assessing sensitivity, is a valuable contribution to the assessment of uncertainty, as noted by Reviewer #1. The authors state that

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU

sensitivity analysis consists of two parts: (i) a strategy for sampling the model parameter space, and (ii) a numerical or visual measure which quantifies the impacts of sampled parameters on the model output of interest, although this breakdown into two parts does not consider model error, arising from the fact that any model is no more than a simplified representation of the complex reality: the authors also restrict themselves to a single widely-used soil-moisture accounting model. Amongst the procedures used by the authors are techniques well-known to experimentalists in other fields, namely the analysis of variance (ANOVA) and sampling methods based on Latin hypercubes, also known as orthogonal Latin squares. Indeed the whole issue of sensitivity of models and their parameters, by running ensembles of model predictions, is strongly reminiscent of the problems faced by statisticians half a century ago, when an "experimental design" was to be selected for the comparison of "treatments" (for which read "hydrological models"), using limited "experimental material" (for which read "number of watersheds, computer time and programming effort") with "observations" (for which read "hydrological model predictions") subject to uncontrolled and uncontrollable variability (for which read "uncertain and possibly unrepresentative measurements of precipitation and other meteorological inputs"). The degree of replication (number of experimental units on which each "treatment" was tested) has as its present-day analogy the number of watersheds on which hydrological models are to be compared, and as in the past, whilst, a high degree of replication is desirable, in practice the degree of replication is in major part determined by practical limitations of computer time and programming effort. Much of the effort by statisticians collaborating with experimentalists has been, and still is, directed at using available experimental material to best effect, and there appears to be great scope for fruitful collaboration between hydrologists, who need to do experiments that identify the best models for particular purposes, and statisticians familiar with concepts of experimental design. Indeed, it is possible to pursue the analogy with experimental design, and the statistical models underlying such designs, still further; for example, the effects of watersheds (replicates) on which hydrological models are tested could be regarded as fixed ("Fixed-effects" Model: ANOVA Model I)

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

or drawn from a population of watersheds ("Components of Variance Model": ANOVA Model II) from which the watersheds used in the study can be considered a random sample. "Mixed-effects" models – say with models fixed, but watersheds random – are also possible.

To the editor given the pleasure of handling the paper, it appears possible for the issue of sensitivity to be formalized in terms of an ANOVA structure. This could be done in many ways, one of which could be the following.

Assume that M models ($i = 1 \ldots M$) are to be tested using data (precipitation, streamflow, meteorological variables) from W watersheds ($j = 1 \ldots W$). Draw K samples of length (say) 10 consecutive years from each watershed's record; fit the M models to each sample drawn; let *Yijk* be the quantity of interest obtained when model i is fitted to the k-th sample from watershed j. (The quantity *Yijk* might be simply a measure of model goodness of fit, or some quantity predicted by the model, or any other quantity of interest). A statistical model could be formulated for the *MWK* values *Yijk*, of the simple form

$$Yijk = \mu + m_i + w_j + \varepsilon_{ijk} \qquad \text{(I)}$$

where m_i and w_j measure the parts of *Yijk* explained by model i and watershed j respectively. The "residual" ε_{ijk} measures the discrepancy between the predicted μ + m$_i$ + w_j and the observed *Yijk*. With both m_i and w_j regarded as fixed, least-squares could be used to estimate μ, m$_i, w_{jj}$ and the variance of residuals σ_ε^2. Hence the standard errors of the estimated model effects $\overset{\wedge}{\mu} + \overset{\wedge}{m_i}$ can be calculated from which the models can be compared The ANOVA structure for this simple model is

	Degrees of freedom	Expected Mean Square
Variation between watersheds	$W-1$	$\sigma_\varepsilon^2 + \Sigma w_j^2/(W-1)$
Variation between models	$M-1$	$\sigma_\varepsilon^2 + \Sigma m_i^2/(M-1)$
Residual variation	$WM(K-1)$	σ_ε^2

It is not essential for the residuals ε_{ijk} to be Normally-distributed; an interaction term

mw_{ij} could also be included. The computational work could be reduced in a number of ways, such as through the use of balanced incomplete block designs with the watersheds as blocks; then not all models need be tested on each watershed. The design could be made more efficient by making the 10-year sequences common to all the W watersheds used (so that the first of the K random samples might be, say, the years 1942-51 on all watersheds; the second of the K random samples perhaps from the years 1959-1968, again for all watersheds;...). It might even be possible to obtain estimates of model performance on ungauged catchments, by re-writing the statistical model (I) above in the form

$$Yijk = \mu + m_i + \beta_1.X_{1j} + \beta_2.X_{2j} + \ldots + \varepsilon_{ijk} \qquad \text{(II)}$$

where X_1, X_2, ... are watershed characteristics; the parameters β_1, β_2 ... could be estimated from the W gauged watersheds, and their estimates used together with estimates of μ and m_i to give an estimate of how model i would perform on an ungauged watershed with given characteristics $X_1, X_2 \ldots$.

Clearly the computational load increases with W, M and K, but there could well be scope for choosing them in some near-optimal fashion; a further possibility might be to combine results obtained by different groups of researchers in the form of meta-analyses used by statisticians in medical fields.

The above comments are speculative, but are given in the hope that they will stimulate the authors to pursue the question of assessing sensitivity beyond the rather restricted format given in their paper, which is nevertheless a useful contribution in an area of enormous importance.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 3333, 2006.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

EGU