NLFIT - George Kuczera

PUB-IAHS Workshop

Uncertainty Analysis in Environmental Modelling

6th – 8th July 2004

a) Introduction:

The NLFIT software implements Bayesian nonlinear regression. It has the following features:

•Offers two search methods to find most probable parameters: i) an interactive Gauss-Marquardt search method which is well suited for smooth problems; ii) SCE method for non-smooth problems.

•Performs joint or multi-response fitting using a weighting scheme based on informativeness of each response time series

•Accepts prior information on model parameters

•Offers a general error model which describes time dependence and non-stationary variances

•Uses interactive mouse-based graphical user interface to drive program, report results and explore diagnostics

•Parameter uncertainty can be described by second-order Hessian approximation and Metropolis algorithm. Reports confidence and prediction limits.

•Virtually any model can be hooked to NLFIT using two Fortran subroutines that describe the model and data

b) Advantages

•The approach cannot hide from the evidence. Assumptions about how errors propagate are explicit and can be tested using diagnostic methods.

•General error models employs Box-Cox transformation and ARMA process to relax the rarely justified least squares assumptions of independence and constant error variance.

•Reasonably robust. When diagnostics are satisfied, predictive uncertainty is reasonably well described.

c) Disadvantages

•Assumes inference is conditioned on inputs. In other words, no uncertainty in inputs is assumed. This affects predictive uncertainty and causes bias in parameters. Input uncertainty can be handled approximately by creating additional parameters to model input uncertainty. However, this is awkward in the NLFIT software framework – the BATEA approach is superior.

•Lumps model and output error as an additive error. Though model error is not well understood, it is suggested at time scales at which quickflow operates, it is unlikely that the model error manifests itself as an additive process. At longer aggregation time scales, by virtue of central limit theorem, accumulation of errors will be better approximated by additive error.

•Requires moderate Fortran programming skill to hook two subroutines that describe the model and data to NLFIT.

d) Assumptions

NLFIT assumptions are consistent with those of multi-response regression methods.

e) Most appropriate application areas

NLFIT has been applied to a variety of conceptual models: continuous soil moisture accounting models (eg, Sacramento, TOPMODEL, SFB); event-based rainfall-runoff models (eg, RORB, kinematic wave); water quality models (eg, CATPRO) and water demand modeling.

Diagnostics suggest NLFIT assumptions are best approximated when input data are of reasonable quality and short aggregation time scales are avoided.

f) Reading list

Kuczera, G. Improved parameter inference in catchment models, Parts I and II, Water Resources Research, 19(5), 1151-1172, 1983.

Mroczkowski, M., Raper, G.P. and Kuczera, G., The quest for more powerful validation of conceptual catchment models, Water Resources Research, 33(10), 2325-2336., 1997.

Kuczera, G and Parent, E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, Journal of Hydrology, 211(1-4), 69-85, 1998.

g) Software availability

Freely available to researchers and not-for-profit users.

Fortran source code can be obtained by emailing author at george.kuczera@newcastle.edu.au.

NLFIT uses a simple graphical user interface which is operating system dependent. Presently it supports Windows 95/98/NT/2000/XP using Intel/Compaq Visual Fortran and any system running X Windows.

h) Web links or other information

www.eng.newcastle.edu.au/~cegak

i) Figures – none given

j) Delegates Comments (please add !!)