Multicriterion Model Evaluation based on Recursive Partitioning of the Input Space (MMERPIS) – Ion Iorgulescu & Keith Beven

PUB-IAHS Workshop
Uncertainty Analysis in Environmental Modelling
6th – 8th July 2004

a) Introduction:

Principle: Perform a multi-criterion model evaluation (& optimisation in the 'optimality' framework) on the terminal partitions identified by the non-parametric direct mapping of rainfall-runoff relationships of Iorgulescu & Beven. With respect to other existing multi-criterion methods, the partitioning algorithm we use offers some interesting and desirable properties. In particular it provides effective information extraction where 1) numerically relevant partitioning variables and threshold are identified from the data, 2) complexity-control of the partitioning process, 3) insights in error structure, and 4) conservative bounds on measurement error. The latter are defined by a measure of within-partition deviation (sum of squares or absolute deviations as informed by the within-partition distribution). These are held as justifiable thresholds for behavioural models; i.e. we ask the hydrological model to provide better predictions than the partition average (or median) discharge.

Procedure ('optimality' framework; see Figure): (1) Divide the available data in three sets: calibration (50%), complexity-control (20%) and validation (30%); (2) Build a regression tree by growing on the calibration set and pruning on the complexity-control set; (3) Estimate the parameters of the hydrological model using a multi-criterion method (Maximin or constrained) on the terminal partitions using the calibration set; (4) Reject model if any (95%) partition-based criterion < threshold on calibration set (under-fit test); (5) Reject model if any (95%) partition-based criterion < threshold on complexity-control set (over-fit test); (6) Re-estimate parameters using both calibration and complexity-control sets; (7) Asses and report performance on the validation set.

b) Advantages

- Provides justifiable criteria and thresholds for behavioural models.
- Controls model complexity (for both under-fit and over-fit).
- Able to extract from the data and use information on the error structure (heteroscedascity & outliers).
- Limits interference in parameter estimation.
- Easy to understand and to implement.

c) Disadvantages

- Not yet mature, still in stage of concepts and initial applications.
- Oriented towards behavioural model identification rather than uncertainty assessment.
- Sub-optimality.

d) Assumptions

Measures of within-partition deviation from the partition modal value of the response variable represent justifiable thresholds for deciding the 'behaviourability' of a model.

e) Most appropriate application areas

Successful non-parametric direct mapping (seems to be the case for rainfall-runoff on the catchments studied to date; research by Selle & Huwe, EGU 2004 suggests this could also the case for subsurface flows)

Where information in data dominates prior knowledge about model and especially the error structure.

f) Reading list

Iorgulescu, I. and K. Beven, Non-parametric direct mapping of rainfall-runoff relationships: an alternative to data analysis and modelling, Water Resour. Res., in press.

g) Software availability

No. We are very keen to do collaborative research.

h) Web links or other information

none given

i) Figures

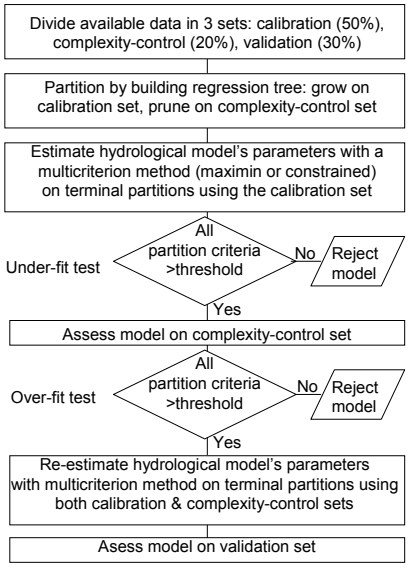

Figure 1: Proposed methodology ('optimality' framework)

j) Delegates Comments (please add !!)