HS1.1 Next hydrological decade

Can we test model hypotheses of flow and transport in assessing the hydrological impacts of change?

Keith Beven

Lancaster Environment Centre

The Lancaster Environment Centre

LANCASTER UNIVERSITY

Summary

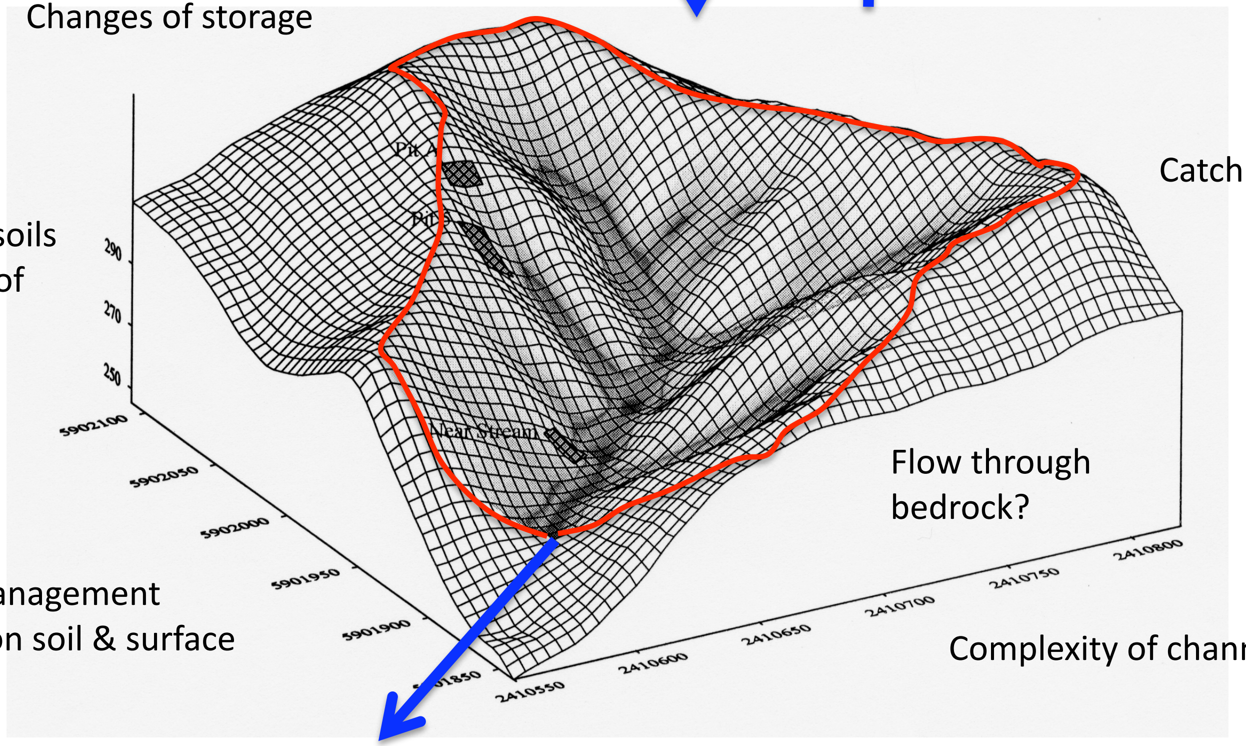
- There is an increasing demand for the prediction of future change on water resources and quality
- But hydrology is an inexact science subject to epistemic as well as aleatory errors in both boundary conditions and state variables
- This means it is difficult to distinguish effects of past changes as well as predict current responses and impacts of future changes
- It also means it is difficult to test models as hypotheses about system response i.e to do hydrological science
- The next decade must address this problem primarily by being pro-active in developing new measurement methods

Hydrology as one of the inexact sciences

Pattern of rainfall inputs

Pattern of actual evapotranspiration

Changes of storage



Catchment area

Heterogeneity of soils & representation of flow processes

Flow through bedrock?

Land Management Effects on soil & surface

Complexity of channels

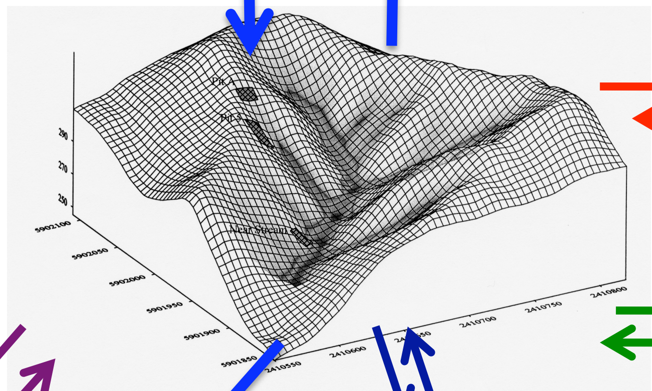
Discharge estimates (requires "rating curve")

The Lancaster Environment Centre

LANCASTER UNIVERSITY

Hydrology as one of the inexact sciences

Atmospheric Models
Weather Forecasting
Climate Change



Water Quality Models
Nutrients
FIOs

Landform Development Models
Sediments
Geomorphology

Ecological Impacts
WFD
Bidiversity
Ecosystem Services

Social Impacts
Flooding / Flood Warning
Planning
Agricultural Policy
Water Resources

Hydrology as one of the inexact sciences

The Water Balance Equation

$$Q = R - E_a - \Delta S$$

At the catchment scale all of the terms are constructed (virtual) variables subject to both epistemic and aleatory uncertainties……and there may be other inputs and outputs impossible to measure

So how to do hypothesis testing in the inexact sciences?

- Martyn Clark, Dmitri Kavetski and Fabrizio Fenicia, Pursuing the method of multiple working hypotheses for hydrological modeling in WRR 2011
- They suggest a formal Bayesian statistical approach to allow for sources of uncertainty in assessing both model structures and parameter sets.
- Bayes ratios can be used to suggest whether one model should be accepted over another
- But requires strong assumptions about the nature of the errors and their information content

Types of error and why they are important

- Formal statistical approach to likelihoods (generally) assumes that the (transformed) errors are additive and random (*aleatory error*) conditional on the model being correct, and that every residual contributing to likelihood is informative in shaping the posterior probability distribution

- But in environmental modelling, many sources of error (in model structure, input data, parameter values,….) are a result of lack of knowledge (*epistemic error*) which will result in *non-stationarity* of error characteristics.

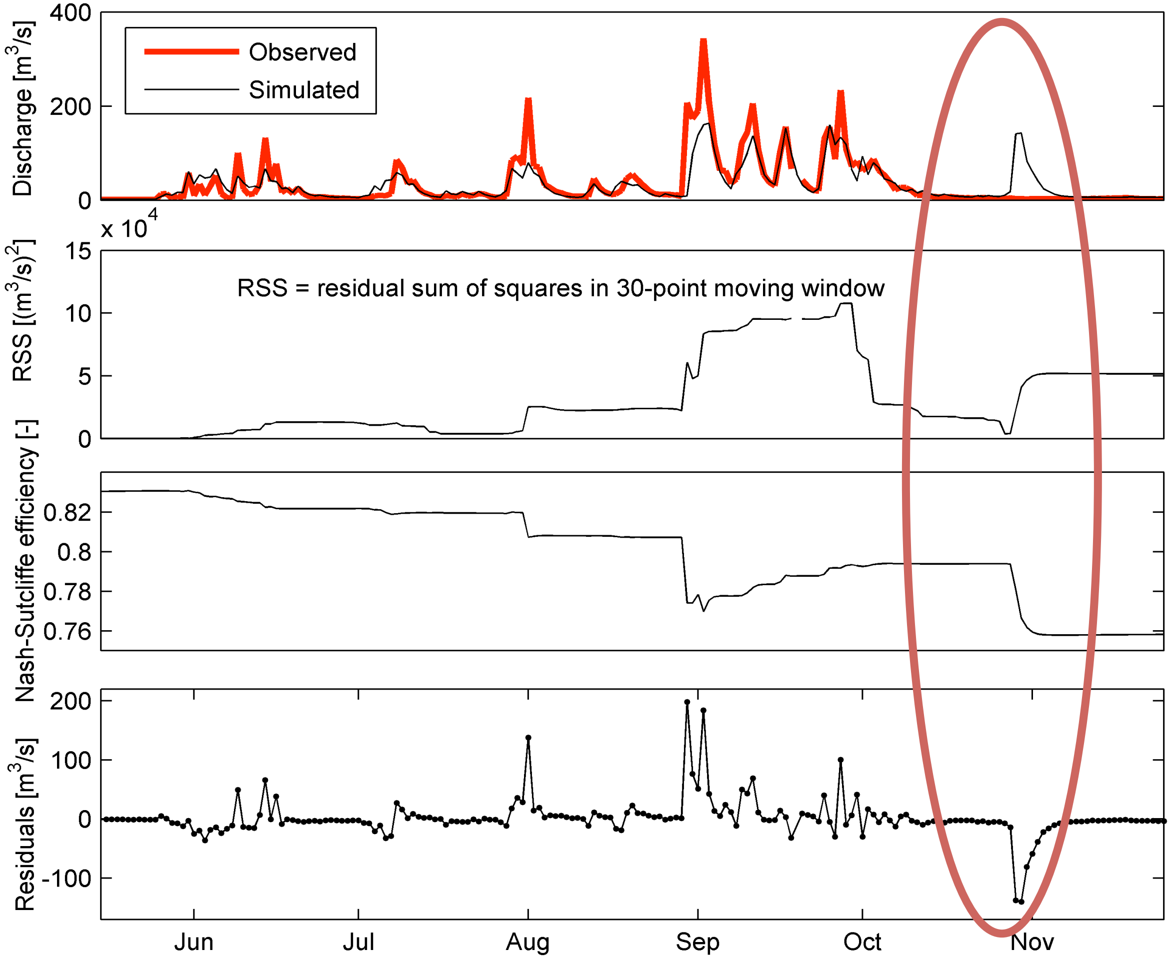
- And in hydrology, data may sometimes be *disinformative* (Beven and Westerberg, HP, 2011; Beven et al. HESS 2011)

Types of error and why they are important

- Errors in the input and boundary condition data (A/E/D)
- Errors in the model structure (E/D?)
- Errors in estimates of parameter values (A/E)
- Commensurability of modelled and observed variables and parameters (A/E/D)
- Errors in the observations used to calibrate or evaluate models (A/E/D)
- Errors of omission (sometimes known omissions) (E/D?)
- The unknown unknowns (D?, becoming E/D)

Disinformation in calibration data

Application of WASMOD to Pasa La Ceiba, Honduras (from Ida Westerberg, Uppsala)

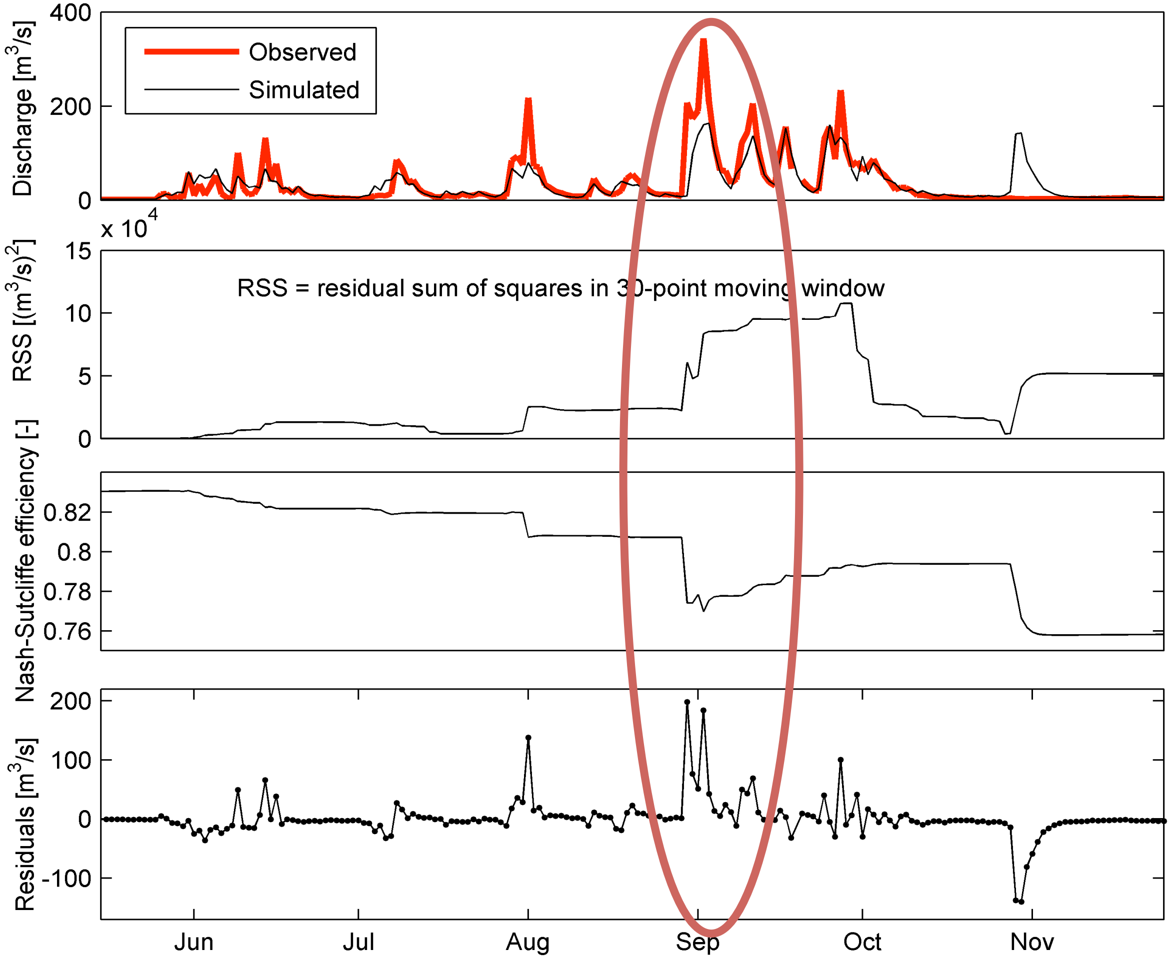


The Lancaster
Environment Centre

LANCASTER UNIVERSITY

Disinformation in calibration data

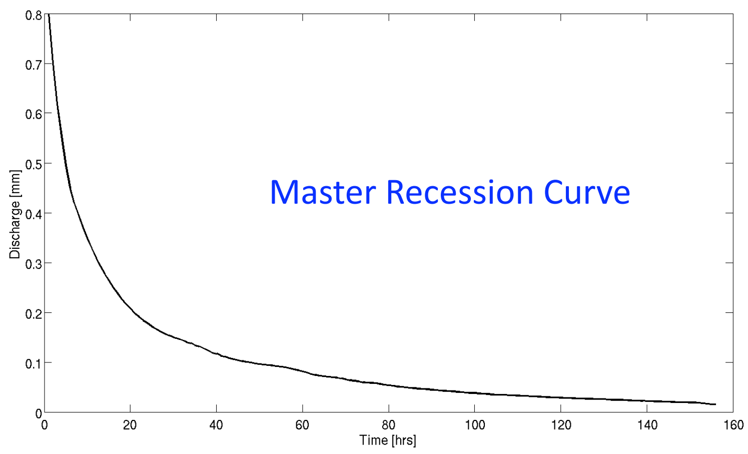
Application of WASMOD to Pasa La Ceiba, Honduras (from Ida Westerberg, Uppsala)

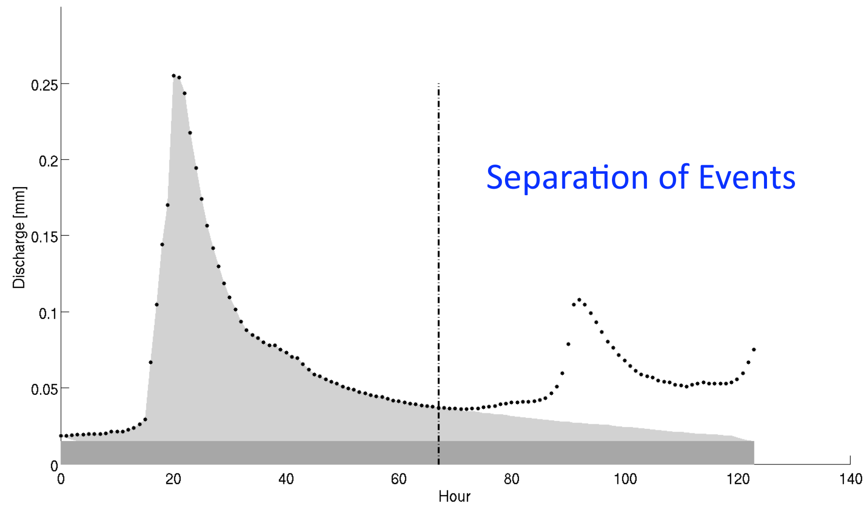


Identification of disinformation

First criterion: Event mass balance consistency (expectation that event runoff coefficient Q / R will be less than one)

But…difficulty of separating events

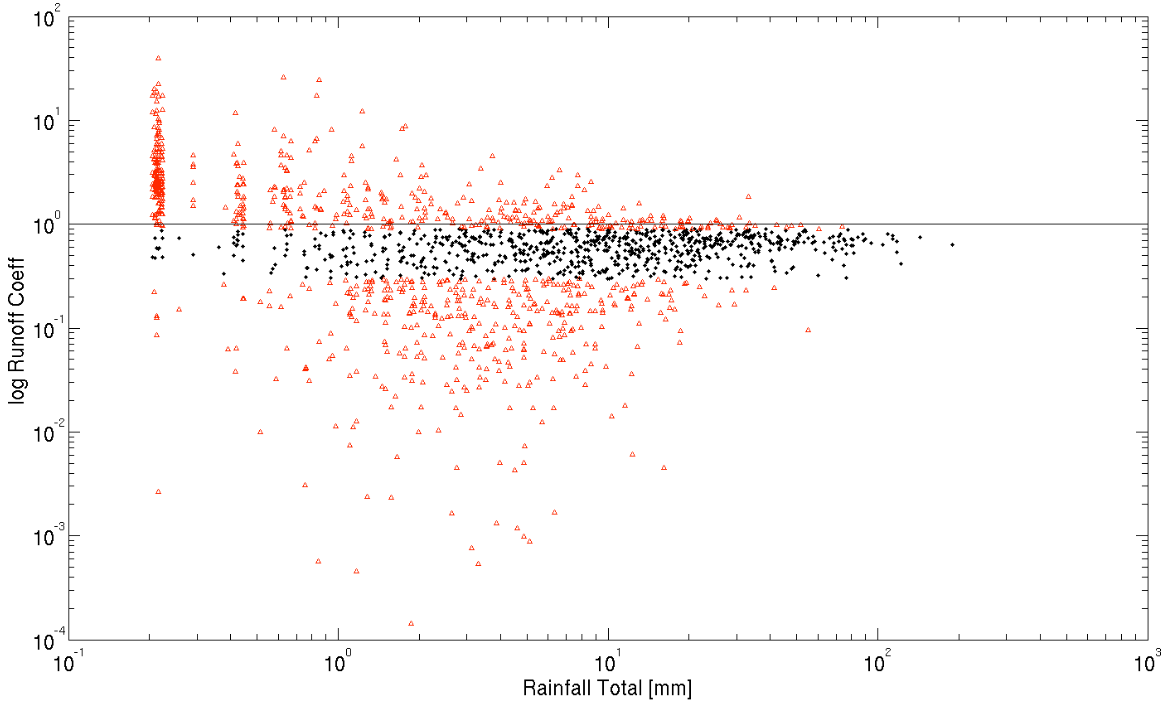




and impact of an inconsistent event on model results might persist for following events, gradually decaying

Identification of disinformation

Results of runoff coefficient determination for River Tyne at Station 23006 – plotted against rainfall totals over catchment area as estimated from 5 gauges (black – range 0.3 to 0.9)



The Lancaster
Environment Centre

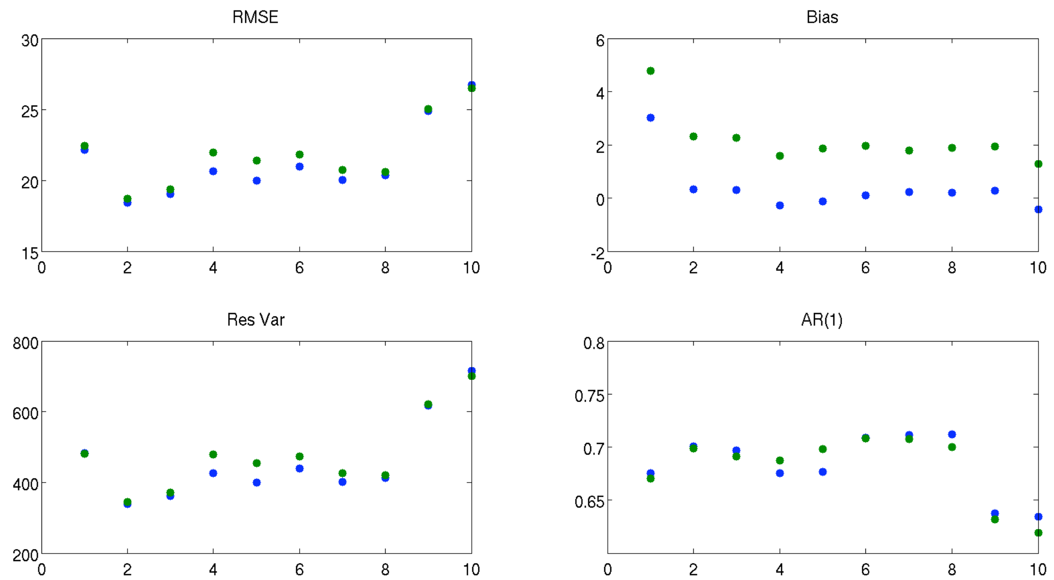
Results from Beven et al. HESS, 2011

LANCASTER UNIVERSITY

Testing parameter sets within a model structure

2 parameter sets with similar RMSE (first 5 years calibration, last 5 years evaluation)

	Cmax	Bexp	alpha	ks	kf
P1	201.98	0.27	0.97	0.02	0.48
P2	245.23	0.61	0.92	0.07	0.49

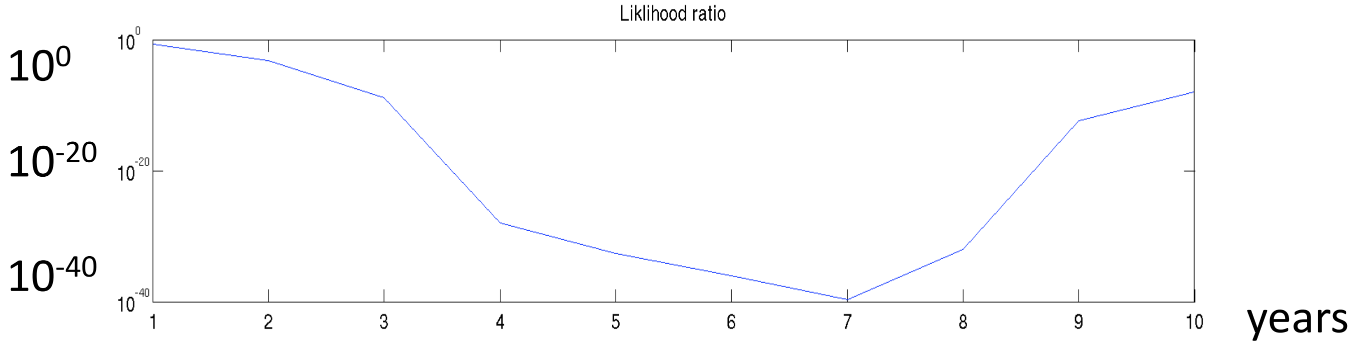


Results from Paul Smith

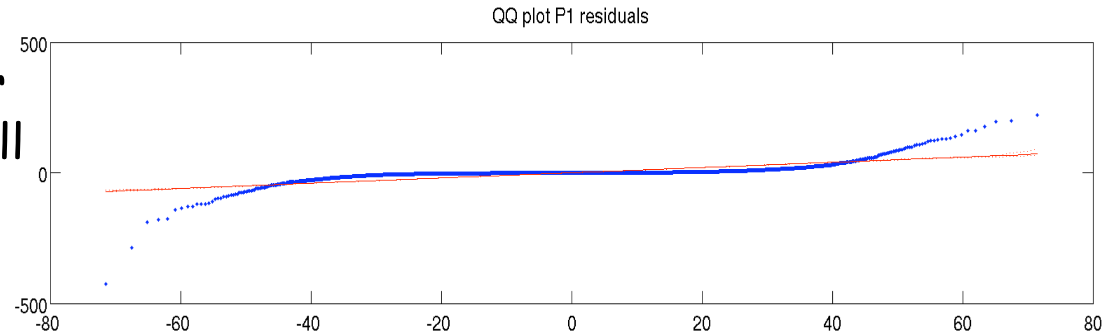
Testing parameter sets within a model structure

Likelihood based on assumption of additive Gaussian errors with mean bias and AR(1) correlation

(a) Likelihood ratio



(b) QQ plot for model P1 over all 10 years



The Lancaster Environment Centre

Results from Paul Smith

LANCASTER UNIVERSITY

Formal likelihoods and hypothesis testing

- Likelihood ratios suggest that information content of data is being over-estimated and models are being over-conditioned (stretching of the likelihood surface)
- Same will apply to testing different model structures given posterior distributions of parameters
- And because inference is being made conditional on model being assumed correct, models are never rejected – Bayes ratios can be used only to suggest if one model might be preferred to another.

Rejection is a good thing!

- But it is useful to be able to reject models – it suggests that some improvements need to be made, either to the forcing/evaluation data or to the model structure

- Limits of acceptability can be used within GLUE to suggest when models should be rejected, even on the basis of a single critical observation

See Beven, CR Geoscience, 2012

Why is this so important in assessing future change?

Type I and Type II errors

- Want to avoid Type II errors of rejecting models that might be useful in prediction because of (epistemic and aleatory) data uncertainties

- And (less seriously) avoid Type I errors (should be corrected as more data are used in evaluation)

What does avoiding Type II errors mean?

- Making proper allowance for errors in forcing data and evaluation data – both epistemic and aleatory errors
- Not over-conditioning - treating epistemic error as if it is random variability is likely to lead to over-conditioning and increased possibility of Type II errors
- Being aware that epistemic errors in prediction will be different but can only be recognised a posteriori (i.e. expect some surprises!)

Some critical questions for the next decade

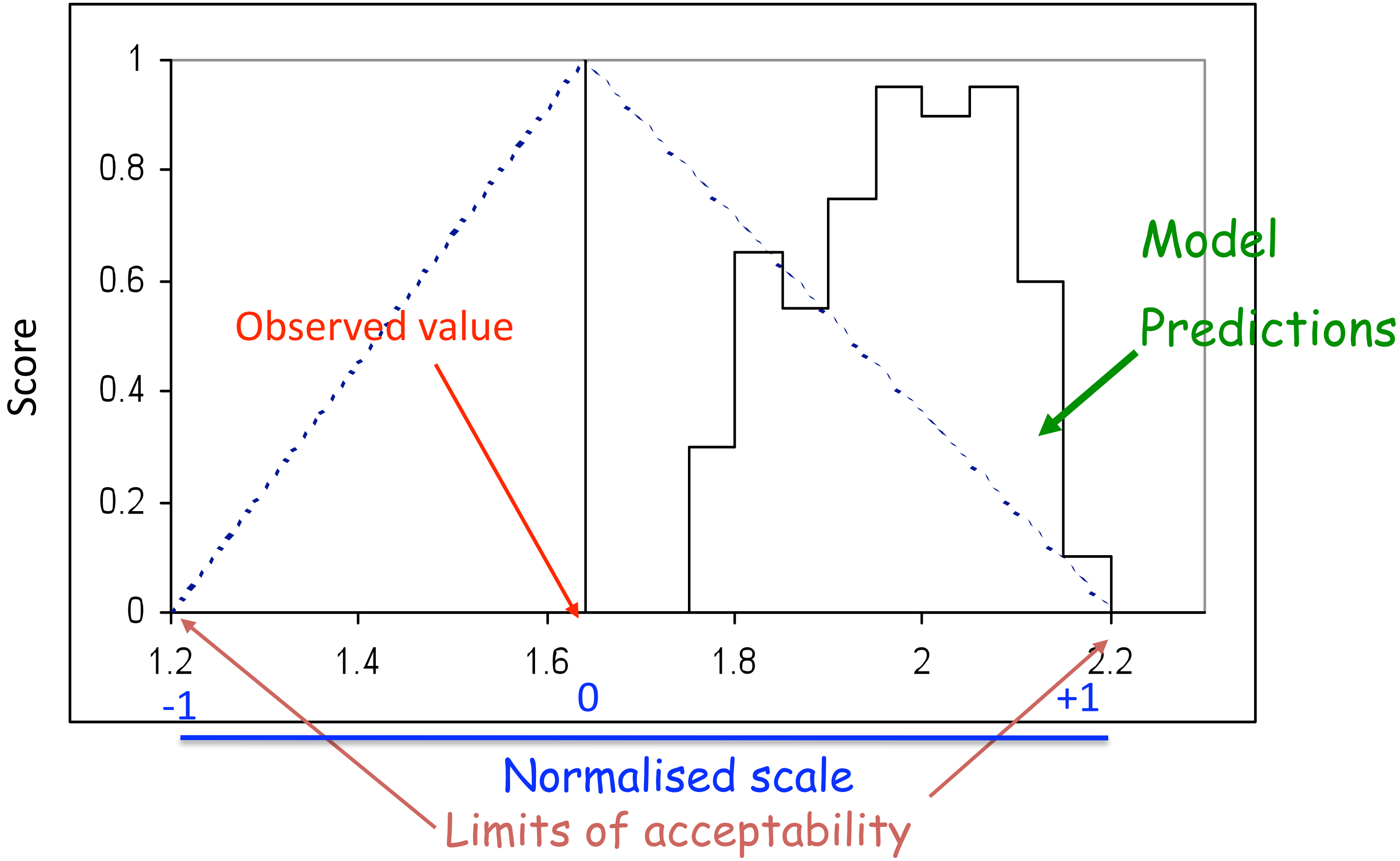
- How can disinformation in hydrological data be identified?
- How can the reduction of information content associated with epistemic uncertainties be reflected in formal or informal likelihoods?
- Can modellers be persuaded to use a rejectionist framework rather than inference conditional on assuming the model is correct (see HS2.18 Thursday Rm34)?
- How can observational techniques be improved so that disinformation and epistemic errors become much less significant?

One potential approach

- Rather than conditioning on assumption that model is correct, take alternative view that all models are wrong until proven useful within the limitations of available data.
- Alternative formal methodology based on assumptions about data errors *prior* to running any models
- Normalise scores based on limits of acceptability in observed variables prior to running the model.
- Test utility in terms of distribution of scores relative to a non-parametric generator of realisations based on data (e.g. fuzzy clustering algorithm)

Normalised deviation scores

Allows different types and uncertainty of observations to be evaluated on a common basis



One potential approach

Should then test

- Proportion of zero scores
- Benefit relative to null data based model
- Posterior analysis of normalised scores

Does not protect against new types of epistemic error in new prediction periods – but to do so requires prior knowledge which (by definition) will not be available (so expect occasional failures!)

That is why we need better observational techniques!!!

New Observational Techniques

- What is the functional requirement?
  - Improved estimates of discharge (good enough and cheap enough that incremental discharges can be determined continuously)
  - Improved estimates of integrated rainfalls at surface (at what scale….. 0 to low order catchment?)
  - Improved estimates of total storage (at what scale… hillslope scale?)
  - More continuous determination of residence time distributions (new isotope methods adequate in time and space?)
- How to turn the functional requirement into a technical specification?

Summary

- There is an increasing demand for the prediction of future change on water resources and quality
- But hydrology is an inexact science subject to epistemic as well as aleatory errors in both boundary conditions and state variables
- This means it is difficult to distinguish effects of past changes as well as predict current responses and impacts of future changes
- It also means it is difficult to test models as hypotheses about system response i.e to do hydrological science
- The next decade must address this problem primarily by being pro-active in developing new measurement methods

For more on catchment change including blog see
www.catchmentchange.net