Data analysis, data fusion and uncertainty evaluation

Alistair Forbes
Mathematical and Scientific Computing Group
National Physical Laboratory, Teddington, UK

GRUAN, 28th February 2008, Lindenberg

Traceability and uncertainty evaluation

Traceability: express measurement results in terms of standard units

Measurement result: value and associated uncertainty

Mutual Recognition Arrangement (MRA): interoperability of measurement systems, database of laboratory capabilities

Metrology community: GUM *Guide to expression of uncertainty in measurement* ISO 1995

Uncertainty expressed in terms of probability distributions, value = distribution mean, uncertainty = distribution standard deviation.

Build up uncertainty of measurement in terms of uncertainties associated with influence quantities

Well characterised systems: physics, sensor behaviour, environment.

Testing community

ISO 5725 (1995) Accuracy (trueness and precision) of measurement methods

Trueness: expected difference between measurement value and ‘true’ value

Precision: expected spread of results

Repeatability conditions: minimise the effect of influence factors

Reproducibility conditions: maximise the effect of influence factors (lab, instrument, operator, environment, ...)

Sensor characterisation and sensor use

Intercomparison exercises: validate uncertainty budgets — characterise measurement systems

Data fusion and reconciliation

Primary data fusion: need different sensors in order to determine system parameter values

Secondary data fusion: use multiple measurements to reduce uncertainties, provide validaton, diagnostic information

Data fusion algorithms: use sensor characterisation/uncertainties to guide fusion algorithms, provide uncertainties associated with system parameters (maximum likelihood estimation)

Correlation due to common systematic effects

Heterogeneous sampling with respect to location, time

Hybrid sensor systems: potentially provide uncertainties much lower than any component sensor (e.g., calibration)

Use data redundancy to indentify (early) poorly performing components: effective redundancy and network vulnerability, are discrepencies explainable by natural variation, sensor uncertainties

Use data redundancy to improve sensor characterisation (re-analysis): validation — characterisation

Experimental design

Many issues: location, time, sensors, analysis methods, budget

Maximise the information gain: reduce uncertainties

Minimise decision costs, risks associated with incorrect inferences

Use economic model to balance measurement costs and decision costs: measurement costs scale with 1/uncertainty^2, decision costs scale with 1/uncertainty

Two types of experiment: measurement system characterisation, climate monitoring

Measurement system characterision: maximise opportunity for influence quantities to affect the result (variable timings, locations)

Climate monitoring: minimise ... (fixed timings, locations, but include special campaigns to provide validation, reduce risks)

Nested measurement strategies: combine regular, coarse grids with more random subgrids

(Unwelcome) heterogeneous measurement strategies forced by constraints, instrument operation