GLOSSARY

Acceptable simulation – see *behavioural simulation*.

Aleatory Uncertainty - The way in which a quantity varies in some random stochastic way in a system. Often used in contrast to *epistemic uncertainty*.

Approximate Bayesian Computation (ABC) - ABC methods bypass the evaluation of the *formal likelihood* function by the use of an *informal likelihood* measure or summary statistic. In this way, ABC methods widen the realm of models for which statistical inference can be considered. ABC methods are mathematically well-founded, but they inevitably make assumptions and approximations whose impact needs to be carefully assessed.

Autocorrelation (serial correlation) - correlation between the elements of a time series and other elements from the same time series when separated by a given interval.

Autocorrelated Errors - A time series of model residuals that exhibit correlation at successive time steps. In distributed models, errors might be correlated in both space and time.

Autoregression (AR) model predicts future timesteps based on autocorrelation with elements the observations at previous time steps within a given interval.

Bayes Equation - Equation for calculating a posterior probability given a prior probability and a *likelihood function*. Used in the GLUE methodology to calculate posterior model likelihood weights from subjective prior weights and a *likelihood measure* chosen for model evaluation

Bayesianism / Bayes Explanation Approach - In which subjective, empirical (inductive) and theoretical (deductive) information can be combined. Fits in well with current ideas about the sociological context of science and the way it is done in practice.

Behavioural Simulation - A simulation that gives an acceptable reproduction of any observations available for model evaluation. Simulations that are not acceptable are non-behavioural (also called *acceptable simulation*).

Boundary Conditions - Constraints and values of variables required to run a model for a particular flow domain and time period. May include input variables such as rainfall and temperature.

Conditional - The premise that any modelling application and associated assessment of uncertainty is conditional on the information available, method(s) used as well as any assumptions and choices made throughout the analyses.

Condition Tree – A method for determining and recording the *conditional* nature of modelling and uncertainty analyses (see Beven and Alcock, 2012) which can be used as an audit train for a given application.

Conditioning - The process of refining a model structure, or a distribution(s) of parameter values of a model structure by comparison with observations.

Conditioned Uncertainty Estimation – model uncertainty estimation where simulations are conditioned (see *conditioning*) using observations.

Copula - A method of transformation from a space of scaled unit axes to a complex multivariate distribution with dependencies.

Covariance Matrix - A way of expressing the statistical uncertainty for a set of multiple parameters or variables as a square matrix of coefficients. The diagonal elements in the matrix represent the variance of each individual member of the set; the off-diagonal elements the covariation between pairs of members. The higher the degree of covariance, the greater the interaction between pairs of parameters or variables. The covariance can be scaled to represent correlation between the members.

Disinformation – Where errors and uncertainties in observations lead to disinformative inferences associated with model conditioning.

Epistemic Uncertainty - The way in which the response of a system varies in ways that cannot be simply described by random stochastic variation. Often used in contrast to *aleatory uncertainty*. Also known as Knightian uncertainties (after Frank Knight (1885-1872) who himself referred to "true uncertainties" that could not be insured against, as opposed to risk that could be assessed probabilistically, see Knight, 1921)

Equifinality - 1 The concept that there may be many models of a system that are acceptably consistent with the observations available, derived from the General Systems theory of Ludvig von Bertalanffy (1968) and adopted in environmental modelling by Beven (1993, 2006).
Equifinality 2 - The adaption of the von Bertalanffy concept to geomporphology by Culling (1957) that that similar landforms might arise from different processes and histories.
Equifinality 3 - The nonlinear dynamical systems version later expressed by Culling (1987, rejecting his earlier view). Culling distinguished between strict equifinality where a perturbed system will return to its original form after some transition time and weaker forms in which equifinality implies only persistence or stability of some property of the system in its trajectory in state space (as might be observed, for example, if there is some attractor in state space).

Likelihood (Formal) - A quantitative measure of the acceptability of a particular model or parameter set in reproducing the system response being modelled based on a formal parametric function to represent the structure of the errors.

Forward Uncertainty Estimation – The propagation of *a priori* defined uncertainty estimates (e.g. using sampling distributions for model parameters and/or inputs).

Fuzzy Logic - A system of logical rules involving variables associated with a continuous fuzzy measure (normally in the range 0 to 1) rather than the binary measure (right/wrong, 0 or 1) of traditional logic. Rules are available for operations such as addition and multiplication of fuzzy measures and for variables grouped in **fuzzy sets**. Such rules can be used to reflect imperfect knowledge of how a variable will respond in different circumstances in terms of the possibilities of potential outcomes.

Fuzzy Measure - A degree of membership of a quantity to a fuzzy set (see *fuzzy logic*).

Global optimum - A set of parameter values that gives the best fit possible to a set of observations.

Homoscedastic errors - A time series of model residuals that exhibit a stationary variance over a simulation period (see also *Heteroscedastic errors* and *Autocorrelated Errors*).

Heteroscedastic Errors - A time series of model residuals that exhibit a changing variance over a simulation period (see also *Homoscedastic errors* and *Autocorrelated Errors*).

Incommensurate/incommensurability - Used here to refer to variables or parameters with the same name that refer to different quantities because of a change in scale.

Independence - Two variables are independent if a change in the value of one variable has no effect on the effect of the other.

Initial Conditions - The auxiliary conditions required at the start of a run of a model to define all initial model states.

Joint probability - A statistical measure of the likelihood of two events occurring together and at the same point in time.

Latin Hypercube sampling – a sampling approach aimed at reducing the number of near-random samples made (Iman and Conover, 1982). Each parameter axis is split into the same number of equally probable intervals before sampling commences. See also Sobol sampling**.**

Likelihood (Informal) - A quantitative, but subjectively chosen, measure of the acceptability of a particular model or parameter set in reproducing the system response being modelled.

Likelihood Measure - A quantitative measure of the acceptability of a particular model or parameter set in reproducing the system response being modelled.

Limits of Acceptability - an approach which provides an alternative framework for testing and rejecting models as hypotheses of how a system works that is not based on sums of residual measures treating errors as purely statistical variables.

Linearity - A model (or model component) is linear if the outputs are in direct proportion to the inputs.

Local Optimum - A local peak in the parameter response surface where a set of parameter values gives a better fit to the observations than all parameter sets around it, but not as good a fit as the *global optimum*.

Low discrepancy sampling – A sequence of **n-dimensional** samples which fills the **n-sample-space** more uniformly than an uncorrelated random sequence.

Marginal distribution - In a multivariate distribution, the marginal distribution obtained by integrating over all but the dimension associated with the particular variable for which the marginal distribution is required. It is the distribution of that variable conditioned on the distributions of all the other variables in the multivariate distribution function.

Markov Chain Monte Carlo (MCMC) – A simulation-based density-dependent guided *Monte Carlo* search. MCMC are iterative schemes for sequentially simulating from a probability distribution where each draw is a random function of the previous draw and are ultimately aimed at converging to *posterior parameter (variable) distributions* by accepting or rejecting proposed values (e.g. proposed samples of parameter sets) based upon observations. See *Markov Process.*

Markov Process - A process that satisfies the Markov property that prediction for the future of the process can be based solely on its present state (as well as it could be if we had knowledge the process's full history).

Model - A set of constructs, derived from explicit assumptions, about the how a system responds to specified inputs. Quantitative models are normally expressed as sets of assumptions and mathematical equations and implemented as computer codes.

Model Space - A hyperspace defined by the ranges of feasible models and parameter values, with dimensions for each parameter within each model.

Monte Carlo Simulation - Simulation involving multiple runs of a model using different randomly chosen sets of parameter values.

Nonlinear - A model is nonlinear if the outputs are **not** in direct proportion to the inputs but may vary with intensity or volume of the inputs or with antecedent conditions.

Nonparametric Method - A method of estimating distributions without making any assumptions about the mathematical form of the distribution

Nonstationarity - A system in which the characteristics are expected to change over time; a model in which the parameters are expected to change over time.

Non-uniqueness - A expression of the problem of identifying parameter values in a model, given limited observational data (see also *equifinality*)

Normally Distributed -A variable is normally distributed if its distribution can be adequately fitted by the Normal or **Gaussian distribution** function that is symmetrical about the mean, bell-shaped and with infinite tails.

Objective Function (*Performance measure, goodness-of-fit*) - A measure of how well a simulation fits the available observations.

Over-parameterisation - Problem induced by trying to calibrate the parameter values of a model that has too many parameters than can be supported by the information content of the calibration data

Parameter - A constant (but see *nonstationarity*) that must be defined before running a model simulation.

Parameter Space - A space defined by the ranges of feasible model parameters, with one dimension for each parameter.

Parsimony - The concept, sometimes known as Occam's razor, that a model should be no more complex than necessary to predict the observations sufficiently accurately to be useful.

Performance Measure (*Objective function, goodness-of-fit*) - A measure of how well a simulation fits the available observations.

Possibility - A non-statistical measure of the potential for an outcome or occurrence of an event as an alternative to probability theory. Used in fuzzy set theory, but also has a more general usage (see Klir, 2006)

Posterior distribution - The statistical distribution of a model parameter or output variable after conditioning on the basis of observed data, for example, using a calculated *likelihood* in *Bayes equation*. In a Bayesian learning process, the posterior distribution after one conditioning step may become the prior distribution when new observational data are made available.

Prior distribution - The statistical distribution of a model parameter or output variable assumed or calculated on the basis of only knowledge about the characteristics of the system before data are collected.

Probability - A statistical measure of the potential for an outcome or occurrence of an event. Many statisticians, most notably Lindley (2006), believe that probability is the only way of expressing uncertainty in a potential outcome. There are a variety of foundations for the estimation and interpretation of probabilities of which the main examples are the Frequentist and the Bayesian views. Frequentists hold that probabilities represent the likelihood of outcomes

that would be found if it was possible to take a large number of samples over all potential outcomes. Bayesians (see *Bayesianism*) recognise that this can only ever be an ideal and that prior (often subjective) estimates of probabilities might be useful as an input to estimating probabilities based on limited amounts of evidence.

Random Sample - A set of realisations of a model or variable generated by making choices from a feasible range of possibilities drawn by selecting pseudo-random numbers from specified distributions.

Response Surface - The surface defined by the values of an objective function as it changes with changes in parameter values. May be thought of conceptually as a surface with ''peaks'' and ''troughs'' in the multidimensional space defined by the parameter dimensions, where the ''peaks'' represent good fits to the observations and the ''troughs'' represent poor fits to the observations (see also *Parameter Space*)

Sensitivity analysis - A sensitivity analysis determines how different parameter values affect a particular simulated variable under a given set of assumptions. Variables will be more sensitive to changes in the values of some parameters than to others.

Sobol sampling – Uses Sobol sequences to generate a quasi-random sample that covers the unit hypercube with a more uniformly distributed sample than completely random sampling (see also Latin Hypercube).

Spatial dependence (spatial correlation) - The spatial relationship of variable values (variables defined over space, such as rainfall).

State Space - The space of potential trajectories of a model (or, in a more limited sense, of a particular variable in a model).

Unknowabilty -The concept that because of limitations of current measurement techniques there is much about environmental systems that cannot feasibly be known.

Validation 1 - A process of evaluation of models to confirm that they are acceptable representations of a system. Philosphers of science have some problems with the concept of validation and verification (e.g. Oreskes et al., 1994) and it may be better to use ''evaluation'' or ''confirmation'' rather than validation or verification (which imply a degree of truth in the model).
Validation 2 - Validation is sometimes used in the much more restrictive sense of validation of a computer code (mode) to show that it does produce accurate solutions of the equations on which it is based. For complex models this can also be difficult to show in practice.

Verification - See *Validation*

REFERENCES

Beven K. J., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources, 16:41–51.

Beven, K. J., 2006. A manifesto for the equifinality thesis, *J. Hydrology*, 320, 18-36.

Beven, K. J. and Alcock, R., 2012. Modelling everything everywhere: a new approach to decision making for water management under uncertainty, Freshwater Biology, 56, 124-132, doi:10.1111/j.1365-2427.2011.02592.x

Culling W. E. H., 1957. Multicyle streams and the equilibrium theory of grade. J. Geology, 65:259–274.

Culling W. E. H., 1987. Equifinality: Modern approaches to dynamical systems and their potential for geographical thought. Transactions of the Institute of British Geographers, 12:57–72.

Iman R. L. and Conover W. J., 1982. A distribution-free approach to inducing rank correlation among input variables. Communications in Statistics, Series B, 11:311–334.

Klir, G and Parviz, B (1992) PROBABILITY-POSSIBILITY TRANSFORMATIONS: A COMPARISON, International Journal of General Systems, 21:3, 291-310, DOI: 10.1080/03081079208945083

Knight F. H. Risk, Uncertainty and Profit. Hart, Schaffner and Marx (reprinted by Beard Books, Washinton,DC, 2002), 1921.

Lindley, D., 2006. Understanding Uncertainty, Wiley-Interscience, London.

Oreskes N., Schrader-Frechette K. and Belitz K., 1994. Verification, validation and confirmation of numerical models in the earth sciences. Science, 263:641–646.