Model calibration with uncertain inputs

Richard E. Chandler and Zoi Apostolidi

Department of Statistical Science, University College London
(r.chandler@ucl.ac.uk)

SECURE annual workshop, Glasgow, 20th September 2016

UCL

Models

- Environmental change often studied using models / simulators that convert inputs $\mathbf{X}$ to outputs $\mathbf{Y}$:

$$\begin{array}{c} \text{Parameters } (\boldsymbol{\theta}) \\ \downarrow \\ \text{Inputs } (\mathbf{X}) \longrightarrow \boxed{\text{Model}} \longrightarrow \text{Outputs } (\mathbf{Y}) \end{array}$$

- 'Parameters' control precise relationship between inputs and outputs: usually determined by calibrating model using available data

Models

- Environmental change often studied using models / simulators that convert inputs $\mathbf{X}$ to outputs $\mathbf{Y}$:

$$
\begin{array}{c}
\text{Parameters } (\boldsymbol{\theta}) \\
\downarrow \\
\text{Inputs } (\mathbf{X}) \longrightarrow \boxed{\text{Model}} \longrightarrow \text{Outputs } (\mathbf{Y})
\end{array}
$$

- 'Parameters' control precise relationship between inputs and outputs: usually determined by calibrating model using available data

- Deterministic models: $\mathbf{Y} = f(\mathbf{X}; \boldsymbol{\theta})$
- Models including stochasticity: $\mathbf{Y} \sim \pi(\mathbf{X}; \boldsymbol{\theta})$

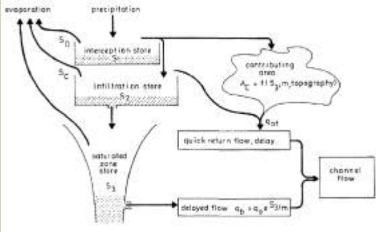
Example: calibrating a rainfall-runoff model (I)

- Rainfall-runoff models typically convert series of precipitation and evapotranspiration (input, **X**) to channel flow (output, **Y**)

Example: TOPMODEL

Specimen parameters

- Storage volumes: S_1, S_2, S_3
- Flow attenuation coefficients: q_{of}, q_o
- Etc. etc.



TOPMODEL schematic (from Beven and Kirkby, Hydrol. Sci. Bull., 1979)

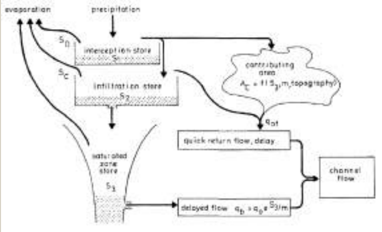
Example: calibrating a rainfall-runoff model (I)

- Rainfall-runoff models typically convert series of precipitation and evapotranspiration (input, $\mathbf{X}$) to channel flow (output, $\mathbf{Y}$)

Example: TOPMODEL

Specimen parameters

- Storage volumes: S_1, S_2, S_3
- Flow attenuation coefficients: q_{of}, q_o
- Etc. etc.

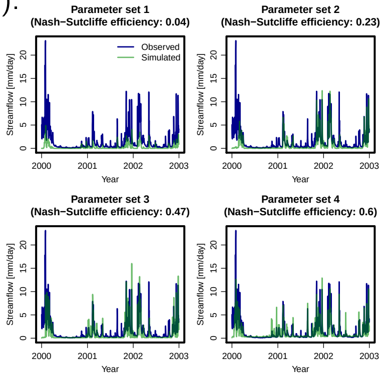


TOPMODEL schematic (from Beven and Kirkby, Hydrol. Sci. Bull., 1979)

- Calibration usually by optimising objective function $Q(\theta, \mathbf{X}, \mathbf{Y})$ say
  - E.g. least-squares, Nash-Sutcliffe efficiency etc.

Example: calibrating a rainfall-runoff model (II)

- Calibrating 24 parameters in TOPMODEL for the Severn at Plynlimon using three years of daily data (precipitation, evapotranspiration, streamflow).
- Sample 1000 parameter sets from ranges defined by Clark et al. (*Hydrol. Proc.*, 2011)
- For each parameter set, simulate streamflows given observed precipitation and evapotranspiration; compare with observed flows using Nash-Sutcliffe efficiency



Software: Claudia Vitolo's "fuse" R package: https://github.com/ICHydro/r_fuse

UCL

Calibration: considerations

What makes a "good" calibration procedure?

Prediction: Calibrated models deliver (a) outputs that are 'close' to actual values in situations outside calibration setting (b) credible, decision-relevant uncertainty assessments

Explanation / understanding: Calibrated parameters correspond to what you think they represent e.g. for data generated from model, calibration procedure should recover 'true' parameter values (minimum requirement) — asymptotically for stochastic models

Calibration: considerations

What makes a "good" calibration procedure?

Prediction: Calibrated models deliver (a) outputs that are 'close' to actual values in situations outside calibration setting (b) credible, decision-relevant uncertainty assessments
Explanation / understanding: Calibrated parameters correspond to what you think they represent e.g. for data generated from model, calibration procedure should recover 'true' parameter values (minimum requirement) — asymptotically for stochastic models

How might you screw up?

- Inappropriate choice of objective function $Q(\theta, \mathbf{X}, \mathbf{Y})$ (more subtle than you might think!)
- Systematic errors in data on $\mathbf{X}$ or $\mathbf{Y}$
- Other data issues …?

Imperfect inputs (and outputs)

- Data $\mathbf{X}^*$ and $\mathbf{Y}^*$ often do not correspond to $\mathbf{X}$ and $\mathbf{Y}$ in model
- Possible reasons:
  - Observations subject to measurement error;
  - Observations interpolated to model grid from irregular network of recording stations;
  - Data derived indirectly from measurements of related quantities using imperfect models (e.g. remote sensing technology);
  - etc.

Imperfect inputs (and outputs)

- Data $\mathbf{X}^*$ and $\mathbf{Y}^*$ often do not correspond to $\mathbf{X}$ and $\mathbf{Y}$ in model
- Possible reasons:
  - Observations subject to measurement error;
  - Observations interpolated to model grid from irregular network of recording stations;
  - Data derived indirectly from measurements of related quantities using imperfect models (e.g. remote sensing technology);
  - etc.

Does this affect anything?

- If there are systematic errors e.g. $\mathbf{X}^*$ systematically higher or lower than $\mathbf{X}$: **yes, obviously**
- If there are no systematic errors e.g. $\mathbf{X}^* \neq \mathbf{X}$ but $\overline{\mathbf{X}}^* \approx \overline{\mathbf{X}} \ldots$?

Three 'imperfect data' scenarios

1. **Measurement error in output:** $\mathbf{Y}^* = \mathbf{Y} + \varepsilon$ where ε is zero-mean 'measurement error'
2. **Measurement error in input:** $\mathbf{X}^* = \mathbf{X} + \delta$ where δ is zero-mean 'measurement error'
   - Error δ independent of actual input $\mathbf{X}$;
   - Recorded input $\mathbf{X}^*$ more variable than actual input $\mathbf{X}$
3. **Imprecisely-administered input:** $\mathbf{X} = \mathbf{X}^* + \delta$ where δ is zero-mean 'dose error'
   - Error δ independent of recorded input $\mathbf{X}^*$;
   - Recorded input $\mathbf{X}^*$ less variable than actual input $\mathbf{X}$

Example: linear relationship between scalar quantities

- Suppose $\mathbf{Y} = (Y_1 \ \ldots \ Y_n)'$, $\mathbf{X} = (X_1 \ \ldots \ X_n)'$ and model is $Y_i = \alpha + \beta X_i$ with parameters $\boldsymbol{\theta} = (\alpha \ \beta)'$
  - Deterministic linear relationship between input and output
- Consider calibration by least-squares regression

Example: linear relationship between scalar quantities

- Suppose $\mathbf{Y} = (Y_1 \; \dots \; Y_n)'$, $\mathbf{X} = (X_1 \; \dots \; X_n)'$ and model is $Y_i = \alpha + \beta X_i$ with parameters $\boldsymbol{\theta} = (\alpha \; \beta)'$
  - Deterministic linear relationship between input and output
- Consider calibration by least-squares regression

A simulation experiment

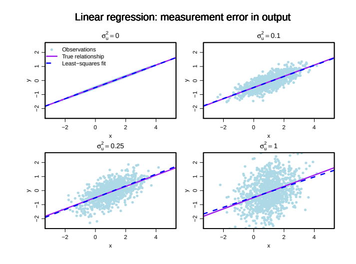
- Generate 10 000 datasets $(\mathbf{X}, \mathbf{Y}, \mathbf{X}^*, \mathbf{Y}^*)$ under each of several imperfect data scenarios
- 1 000 observations per dataset, $X_i \sim N(1, 1)$
- Linear relationship as above with $\alpha = -0.5, \beta = 0.4$
- Estimate α and β for each dataset using least-squares regression of $\mathbf{Y}^*$ on $\mathbf{X}^*$
- Calculate means of least-squares estimates over all datasets

Simulations: linear relationship, scenario 1

Measurement error in output: $\mathbf{Y}^* = \mathbf{Y} + \varepsilon$

- $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ for $\sigma_\varepsilon^2 = 0, 0.1, 0.25, 1$
- Means of least-squares estimates over 10 000 simulated datasets:

	σ_ε^2			
	0	0.1	0.25	1
α	-0.50	-0.50	-0.50	-0.50
β	0.40	0.40	0.40	0.40



Specimen datasets with true and fitted regression lines

True values $\alpha = -0.5, \beta = 0.4$: **no problems here**

Simulations: linear relationship, scenario 2

Measurement error in input: $\mathbf{X}^* = \mathbf{X} + \delta$

- $\delta_i \sim N(0, \sigma_\delta^2)$ for $\sigma_\delta^2 = 0, 0.1, 0.25, 1$
- Means of least-squares estimates over 10 000 simulated datasets:

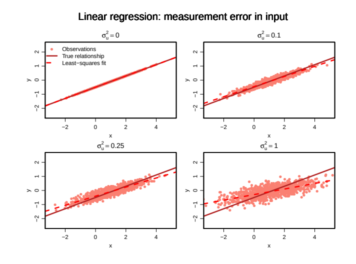
	σ_δ^2			
	0	0.1	0.25	1

Simulations: linear relationship, scenario 2

Measurement error in input: $\mathbf{X}^* = \mathbf{X} + \delta$

- $\delta_i \sim N(0, \sigma_\delta^2)$ for $\sigma_\delta^2 = 0, 0.1, 0.25, 1$
- Means of least-squares estimates over 10 000 simulated datasets:

	σ_δ^2 0	0.1	0.25	1
α	-0.50	-0.46	-0.42	-0.30
β	0.40	0.36	0.32	0.20



Specimen datasets with true and fitted regression lines

True values $\alpha = -0.5, \beta = 0.4$: **something wrong** ...

Simulations: linear relationship, scenario 3

Imprecisely-administered input: $\mathbf{X} = \mathbf{X}^* + \delta$

- $\delta_i \sim N(0, \sigma^2_\delta)$ for $\sigma^2_\delta = 0, 0.1, 0.25, 1$
- Means of least-squares estimates over 10 000 simulated datasets:

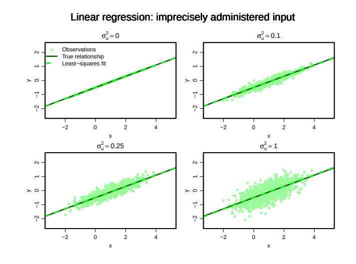
	σ^2_δ			
	0	0.1	0.25	1

Simulations: linear relationship, scenario 3

Imprecisely-administered input: $\mathbf{X} = \mathbf{X}^* + \delta$

- $\delta_i \sim N(0, \sigma_\delta^2)$ for $\sigma_\delta^2 = 0, 0.1, 0.25, 1$
- Means of least-squares estimates over 10 000 simulated datasets:

	σ_δ^2 0	0.1	0.25	1
α	-0.50	-0.50	-0.50	-0.50
β	0.40	0.40	0.40	0.40



Specimen datasets with true and fitted regression lines

True values $\alpha = -0.5, \beta = 0.4$: **everything OK here**

Linear relationship: lessons from simulation

- **Non-differential measurement error on Y** (i.e. $\mathbf{Y}^* = \mathbf{Y}$ on average): least-squares estimates unbiased
- **Non-differential measurement error on X**: least-squares estimates biased
  - Phenomenon called regression dilution bias
- **Imprecisely-administered inputs X**: least-squares estimates unbiased
  - Known as Berkson error model

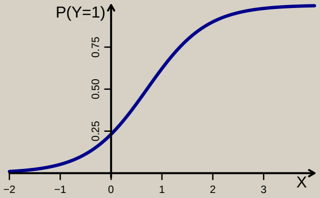
Also:

- Imperfect data contributes to uncertainty in calibration.
- Differential measurement error (i.e. systematic biases in $\mathbf{X}^*$ or $\mathbf{Y}^*$) usually induces bias

Nonlinear models

A simple nonlinear simulation experiment

- Setup as before but using logistic regression: $Y_i = 0$ or 1, with $P(Y_i = 1) = p_i$ and
$$\log\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta X_i \Rightarrow p_i = [1 + \exp(-\alpha - \beta X_i)]^{-1}$$
- Calibration using maximum likelihood estimation
- Other changes from 'linear' experiment:
  - $\alpha = -1.2$, $\beta = 1.7$
  - 'Measurement error in output' $\Rightarrow Y_i$ misclassified with probability π (**NB** differential error)



Simulation results: logistic regression

Means of maximum likelihood estimates over 10 000 datasets

Measurement error in output (i.e. misclassification)

	π			
	0	0.05	0.1	0.2
α	-1.21	-0.98	-0.80	-0.53
β	1.71	1.38	1.13	0.76

Measurement error in input

	σ^2_δ			
	0	0.1	0.25	1
α	-1.21	-1.00	-0.78	-0.28
β	1.71	1.48	1.23	0.68

Imprecisely-administered input

	σ^2_δ			
	0	0.1	0.25	1
α	-1.21	-1.14	-1.07	-0.83
β	1.71	1.62	1.51	1.17

Recall true values: $\alpha = -1.2$, $\beta = 1.7$ $\Rightarrow$ bias now exists in all scenarios

Does it matter?

- **No problem** if aim is just to predict $\mathbf{Y}^*$ from $\mathbf{X}^*$ under identical conditions
- **Problem** if aim is to:
  - Predict $\mathbf{Y}$ from $\mathbf{X}^*$ (possible bias in nonlinear settings; incorrect assessments of uncertainty);
  - Predict $\mathbf{Y}$ or $\mathbf{Y}^*$ from $\mathbf{X}$
  - Quantify drivers of changes in output $\mathbf{Y}$
  - Quantify effects of changes in input $\mathbf{X}$ e.g. 'effect on streamflow of 20% increase in precipitation'
- **Note also:** biases can be propagated to other parameters not directly connected to components of $\mathbf{X}$ subject to measurement error (e.g. correlated covariates in regression models; intercepts in previous simulation examples)

Confronting the problem: some possibilities

Method of moments: find explicit expression for bias (not always possible $\Rightarrow$ limited applicability), estimate required correction

Orthogonal regression: for linear regression ($\Rightarrow$ limited applicability again) with errors in covariates

- Also called Deming regression

Unbiased estimating equations

SIMEX (SIMulation-based EXtrapolation)

Bayesian methods

Note:
Measurement error characteristics — e.g. error variances — are usually known to some extent

Unbiased estimating equations

- **Idea**: calibration often done by solving estimating equation $\mathbf{g}\left(\theta; \mathbf{X}^*, \mathbf{Y}^*\right) = \mathbf{0}$.
  - $\mathbf{g}\left(\theta; \mathbf{X}^*, \mathbf{Y}^*\right)$ is estimating function.

Unbiased estimating equations

- **Idea**: calibration often done by solving estimating equation $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*) = \mathbf{0}$.
    - $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*)$ is estimating function.
- If 'target' value of $\boldsymbol{\theta}$ is $\boldsymbol{\theta}_0$ e.g. if $\mathbf{Y} = f(\mathbf{X}, \boldsymbol{\theta}_0)$ then unbiased estimating equation has $\mathbb{E}[\mathbf{g}(\boldsymbol{\theta}_0; \mathbf{X}^*, \mathbf{Y}^*)] = \mathbf{0}$
    - Necessary (but not sufficient) for decent estimation of $\boldsymbol{\theta}$ in large samples (Jesus & Chandler, 2011, *Interface Focus*).
    - **NB** expectation implies probability distribution — uncontroversial if multiple sets of observations $(\mathbf{X}^*, \mathbf{Y}^*)$ are possible given same set of model quantities $(\mathbf{X}, \mathbf{Y})$.

Unbiased estimating equations

- **Idea**: calibration often done by solving estimating equation $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*) = \mathbf{0}$.
  - $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*)$ is estimating function.
- If 'target' value of $\boldsymbol{\theta}$ is $\boldsymbol{\theta}_0$ e.g. if $\mathbf{Y} = f(\mathbf{X}, \boldsymbol{\theta}_0)$ then unbiased estimating equation has $\mathbb{E}[\mathbf{g}(\boldsymbol{\theta}_0; \mathbf{X}^*, \mathbf{Y}^*)] = \mathbf{0}$
  - Necessary (but not sufficient) for decent estimation of $\boldsymbol{\theta}$ in large samples (Jesus & Chandler, 2011, *Interface Focus*).
  - **NB** expectation implies probability distribution — uncontroversial if multiple sets of observations $(\mathbf{X}^*, \mathbf{Y}^*)$ are possible given same set of model quantities $(\mathbf{X}, \mathbf{Y})$.
- So bias-correct the estimating equation, and corrected calibration follows.

Example: bias-correcting the linear regression model

- Consider $Y_i = \alpha + \beta X_i + \varepsilon_i$ $(i = 1, \ldots, n)$ with measurement error in input: $X_i^* = X_i + \delta_i$, $\delta_i \sim N(0, \sigma_\delta^2)$.

Example: bias-correcting the linear regression model

- Consider $Y_i = \alpha + \beta X_i + \varepsilon_i$ $(i = 1, \ldots, n)$ with measurement error in input: $X_i^* = X_i + \delta_i$, $\delta_i \sim N(0, \sigma_\delta^2)$.
- Least-squares estimates solve $\mathbf{M}' (\mathbf{Y}^* - \mathbf{M}\theta) = \mathbf{0}$ where $\mathbf{M}' = \begin{pmatrix} 1 & \cdots & 1 \\ X_1^* & \cdots & X_n^* \end{pmatrix}$.
  - Solution is $\hat{\theta} = (\mathbf{M}'\mathbf{M})^{-1} \mathbf{M}'\mathbf{Y}^*$
  - Estimating function is $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*) = \mathbf{M}' (\mathbf{Y}^* - \mathbf{M}\theta)$.

Example: bias-correcting the linear regression model

- Consider $Y_i = \alpha + \beta X_i + \varepsilon_i \ (i = 1, \ldots, n)$ with measurement error in input: $X_i^* = X_i + \delta_i$, $\delta_i \sim N(0, \sigma_\delta^2)$.
- Least-squares estimates solve $\mathbf{M}'(\mathbf{Y}^* - \mathbf{M}\theta) = \mathbf{0}$ where $\mathbf{M}' = \begin{pmatrix} 1 & \cdots & 1 \\ X_1^* & \cdots & X_n^* \end{pmatrix}$.
  - Solution is $\hat{\theta} = (\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}'\mathbf{Y}^*$
  - Estimating function is $\mathbf{g}(\theta; \mathbf{X}^*, \mathbf{Y}^*) = \mathbf{M}'(\mathbf{Y}^* - \mathbf{M}\theta)$.
- Can show $\mathbb{E}[\mathbf{g}(\theta_0; \mathbf{X}^*, \mathbf{Y}^*)] = -\begin{pmatrix} 0 & 0 \\ 0 & n\sigma_\delta^2 \end{pmatrix}\theta_0 = -\mathbf{V}\theta_0$, say — biased unless $\sigma_\delta^2 = 0$.
- So modify estimating equation to $\mathbf{M}'(\mathbf{Y}^* - \mathbf{M}\theta) + \mathbf{V}\theta = \mathbf{0}$
  - Corresponding estimator is $\tilde{\theta} = (\mathbf{M}'\mathbf{M} - \mathbf{V})^{-1}\mathbf{M}'\mathbf{Y}^*$

Estimating equations: comments

- Easily implemented but possibly limited to fairly simple statistical models / calibration techniques in practice:
  - Linear models with weighted least squares ✓
  - Generalised linear models with iterative (re)weighted least squares — requires approximations, bias is quadratic in $\boldsymbol{\theta}$, solution not straightforward ☹
  - Other nonlinear least-squares techniques . . . ?

Estimating equations: comments

- Easily implemented but possibly limited to fairly simple statistical models / calibration techniques in practice:
  - Linear models with weighted least squares ✓
  - Generalised linear models with iterative (re)weighted least squares — requires approximations, bias is quadratic in $\boldsymbol{\theta}$, solution not straightforward ☹
  - Other nonlinear least-squares techniques . . . ?
- Uncertainty assessments also available (standard errors etc.)
- Simulation experiments indicate good calibration performance and accurate uncertainty assessments where method is feasible

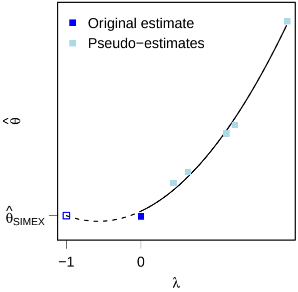
Estimating equations: comments

- Easily implemented but possibly limited to fairly simple statistical models / calibration techniques in practice:
  - Linear models with weighted least squares ✓
  - Generalised linear models with iterative (re)weighted least squares — requires approximations, bias is quadratic in $\boldsymbol{\theta}$, solution not straightforward ☹
  - Other nonlinear least-squares techniques . . . ?
- Uncertainty assessments also available (standard errors etc.)
- Simulation experiments indicate good calibration performance and accurate uncertainty assessments where method is feasible
- Elegant way to discover whether you have a problem

Simulation-based extrapolation (SIMEX)

- **Idea** (Cook & Stefanski, 1994, *J. Amer. Statist. Assoc.*): add extra error to available data, recalibrate model and extrapolate back to zero error
- Various extrapolation schemes available: linear, quadratic etc.
- Uncertainty assessments using bootstrap & jacknife procedures

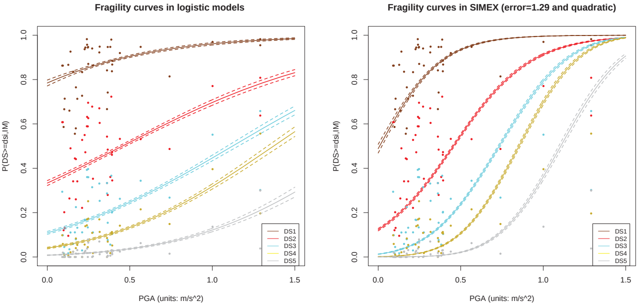
Schematic of idea



- Designed for measurement error in inputs, but (presumably) more widely applicable
- Implemented for various statistical models in `simex` library in R

SIMEX example: seismic fragility curve estimation

Estimated probabilities of buildings exceeding specific damage states as a function of peak ground acceleration (PGA), derived using logistic regression from survey data following the 1980 Irpinia earthquake in southern Italy (source: UCL MSc dissertation by Qiwei Zhang, September 2012)



Left: *naïve estimates.* **Right**: *SIMEX-based estimates*

SIMEX: comments

- Intuitive and easily implemented for wide variety of models
  - `simex` library provides 'off-the-shelf' implementation for linear models and GLMs
- Several choices required, not always obvious:
  - Appropriate extrapolation method;
  - # of 'pseudo-data' points to use;
  - # of simulated data sets for each pseudo-data point
- Can be relatively expensive computationally
- Simulation experiments indicate reasonable performance with relatively small measurement error but poor performance when errors are larger (also expected on theoretical grounds)

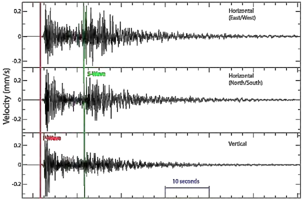
Bayesian methods

- **Idea**: represent all interrelationships in the system via conditional probability distributions, and use probability calculus to infer values of unobserved quantities with assessments of uncertainty.
- Most easily done using Bayesian computational machinery — also allows incorporation of expert judgement.

Example: unified earthquake catalogue for the UK (I)

- Preferred measure of earthquake magnitude is related to overall energy release, usually denoted M_W & estimated since 1970s using modern instrumentation as $\tilde{M}_W$ with rough assessment of uncertainty $\tilde{\sigma}_\varepsilon^2$

- Earlier (post-1900) magnitudes estimated from networks of seismometers providing information on different scale M_L. Estimates $\tilde{M}_L$ could depend on network characteristics, so $\tilde{M}_L \neq M_L$.

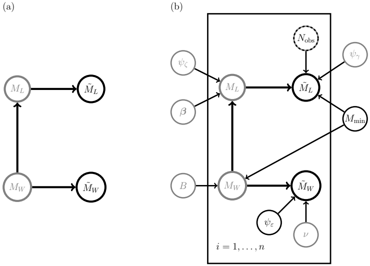


Seismograph trace

- Aim to use recent events for which both $\tilde{M}_W$ and $\tilde{M}_L$ are available, to develop calibration relationship that enables M_W to be estimated from $\tilde{M}_L$

Example: unified earthquake catalogue for the UK (II)

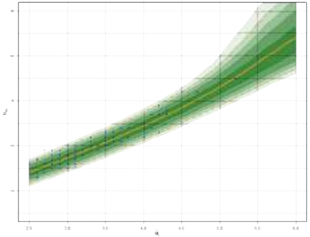
- Build graphical model showing relationships between all quantities required to describe problem



- Use Markov Chain Monte Carlo (MCMC) techniques in Bayesian framework to infer uncertainties based on all observations and expert judgements ('prior distributions')
- **Key output**: posterior predictive distribution for M_W for all events

Example: unified earthquake catalogue for the UK (III)

- Output from Bayesian methodology includes uncertainty distributions for M_W as function of measured $\tilde{M}_L$, for use with pre-1970 events where no observation $\tilde{M}_W$ is available.
- Accounts for measurement error and other features of data e.g. catalogue incompleteness due to non-detection of small events



Predictive distributions of M_W given observation $\tilde{M}_L$ for UK

- **NB** smaller uncertainties in predicting M_W than in predicting imperfect observations $\tilde{M}_W$.

Software: OpenBUGS and R2OpenBUGS

Bayesian methods: comments

- Completely flexible — can handle complex data structures
- Simulation experiments indicate good performance when set up carefully, but . . .
- Considerable skill needed for effective implementation, even with packages like OpenBUGS
  - Need priors for $\mathbf{X}$ and $\mathbf{Y}$; also hyperpriors to allow inference to adapt to data
  - Care needed over independence judgements in informative prior specifications
  - Also with initialisation of MCMC algorithms
- Computationally expensive

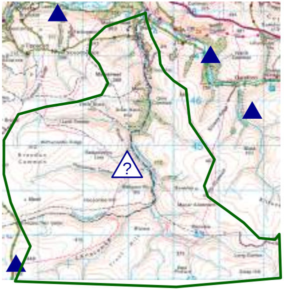
Calibration of non-statistical models

- Most techniques reviewed above are developed primarily for statistical models
- Some additional issues for non-statistical models / simulators:
  - Calibration often more computationally expensive
  - More diverse model / simulator structures — hard to develop general-purpose algorithms
  - Lack of off-the-shelf software for wide-ranging applications

Example: calibrating a rainfall-runoff model

Hypothetical example

- Want to calibrate for catchment with flow gauge but no raingauge
- Standard practice: interpolate precipitation records from neighbouring raingauges
- Result: interpolated precipitation $\mathbf{X}^*$ less variable than true precipitation $\mathbf{X}$ $\Rightarrow$ similar to 'imprecisely-administered input' situation
- How to account for this in calibration?



Non-statistical models: options for calibration with $\mathbf{X}^*, \mathbf{Y}^*$

Bayesian methods including statistical emulation of complex simulators

- Powerful but non-trivial and often computationally challenging

SIMEX-type approach? Not attempted to my knowledge . . .

Anything else?

Concluding questions

- Is this a problem or a red herring?
- If it's a problem, does it worry anybody else?
- If it worries anybody else, does anyone know of **relatively straightforward** options for dealing with it?

☺Thank you for your attention ☺

A useful reference:

Carroll, R.J., D. Ruppert, L.A. Stefanski & C. Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall / CRC Press.