Análise Estatística de Simuladores
Lecture 4: Calibration

Leonardo Bastos[1] and Richard Wilkinson[2]

[1]Universidade Federal Fluminense
[2]University of Nottingham, UK

19º SINAPE, São Pedro, 26 July, 2010

Contents

This lecture is all about calibration.

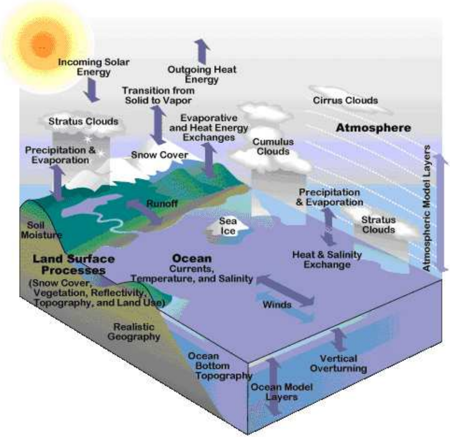
1. Calibration
   - What, why, when
   - Approaches
2. The concept of model error
   - Kennedy and O'Hagan 2001 framework,
   - Toy example: freefall
3. Example: UVic climate model

Calibration

For forwards models we specify parameters θ and i.c.s and the model generates output $\mathcal{D}_{sim}$. Usually, we are interested in the inverse-problem, i.e., observe data $\mathcal{D}_{field}$, want to estimate parameter values.

Different terminology:

- Calibration
- Parameter estimation
- Inverse-problem
- Bayesian inference



Calibration - what?

Calibration is the process of estimating unknown model parameters.

We have three sources of information that we wish to combine

Calibration - what?

Calibration is the process of estimating unknown model parameters.

We have three sources of information that we wish to combine

1. Scientific knowledge captured by the model, η

Calibration - what?

Calibration is the process of estimating unknown model parameters.

We have three sources of information that we wish to combine

1. Scientific knowledge captured by the model, η
2. Empirical information contained in the data, $\mathcal{D}_{field}$

Calibration - what?

Calibration is the process of estimating unknown model parameters.

We have three sources of information that we wish to combine

1. Scientific knowledge captured by the model, η
2. Empirical information contained in the data, $\mathcal{D}_{field}$
3. Expert opinion based upon experience.

We want to combine all three sources to produce the 'best' parameter estimates we can.

Calibration - why?

In physical experiments nature controls all the model parameters.

- e.g., if we drop a ball of the leaning tower of Piza, we do not have to specify what acceleration due to gravity is, nature sets $g = 9.81\text{ms}^{-2}$ for us.

Calibration - why?

In physical experiments nature controls all the model parameters.

- e.g., if we drop a ball of the leaning tower of Piza, we do not have to specify what acceleration due to gravity is, nature sets $g = 9.81\text{ms}^{-2}$ for us.

However, for *in silico* experiments we must tell the computer the value of every physical constant, initial condition, agent specific parameter etc.

Calibration - why?

In physical experiments nature controls all the model parameters.

- e.g., if we drop a ball of the leaning tower of Piza, we do not have to specify what acceleration due to gravity is, nature sets $g = 9.81\text{ms}^{-2}$ for us.

However, for *in silico* experiments we must tell the computer the value of every physical constant, initial condition, agent specific parameter etc.

Any model $\eta(\theta)$ will produce a range of behaviour as we vary θ.

- It is only if we run the model at a 'true' (or optimal) value $\hat{\theta}$ that we will predict the behaviour of the physical system being modelled.

Calibration - why?

In physical experiments nature controls all the model parameters.

- e.g., if we drop a ball of the leaning tower of Piza, we do not have to specify what acceleration due to gravity is, nature sets $g = 9.81\text{ms}^{-2}$ for us.

However, for *in silico* experiments we must tell the computer the value of every physical constant, initial condition, agent specific parameter etc.

Any model $\eta(\theta)$ will produce a range of behaviour as we vary θ.

- It is only if we run the model at a 'true' (or optimal) value $\hat{\theta}$ that we will predict the behaviour of the physical system being modelled.
- For example, if we model an object falling from the tower of Piza, but use $g = 1\text{ms}^{-2}$ in our computer model, we will produce very poor predictions.

Calibration - when?

We can calibrate the model $\eta(\theta)$ whenever we have

- observations of the physical system $\mathcal{D}_{field}$
- an error model relating the model to the system.

Calibration - when?

We can calibrate the model $\eta(\theta)$ whenever we have

- observations of the physical system $\mathcal{D}_{field}$
- an error model relating the model to the system.

We may wish to calibrate if we want to

- Predict future behaviour
- Find physical values of unknown parameters
- Reduce and estimate the appropriate degree of uncertainty about our predictions

Calibration approaches

There are three main approaches to calibration

Calibration approaches

There are three main approaches to calibration

1. Bayesian

Calibration approaches

There are three main approaches to calibration

1. Bayesian
2. Maximum likelihood
    - We find θ that maximises $\pi(D|\theta)$ where $\pi(\cdot|\cdot)$ is a likelihood function relating observations to model predictions.

Calibration approaches

There are three main approaches to calibration

1. Bayesian
2. Maximum likelihood
   - We find θ that maximises $\pi(D|\theta)$ where $\pi(\cdot|\cdot)$ is a likelihood function relating observations to model predictions. For example, if we assume

     $$\mathcal{D}_{field} = \eta(\hat{\theta}) + e$$

     where $e \sim N(0, \sigma^2_{obs})$ then

     $$\pi(D|\hat{\theta}) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{obs}}} \exp\left(-\frac{(D_i - \eta_i(\hat{\theta}))^2}{2}\right)$$

Calibration approaches

There are three main approaches to calibration

1. Bayesian
2. Maximum likelihood
    - We find θ that maximises $\pi(D|\theta)$ where $\pi(\cdot|\cdot)$ is a likelihood function relating observations to model predictions. For example, if we assume

      $$\mathcal{D}_{field} = \eta(\hat{\theta}) + e$$

      where $e \sim N(0, \sigma^2_{obs})$ then

      $$\pi(D|\hat{\theta}) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{obs}}} \exp\left(-\frac{(D_i - \eta_i(\hat{\theta}))^2}{2}\right)$$

3. General ad hoc approaches
    - For example, choosing θ to minimize the sum of squares difference between the data and the model prediction, i.e., minimize

      $$S = \sum_i (D_i - \eta_i(\hat{\theta}))^2$$

      or some other measure.

Bayesian calibration approach

Probability (Bayesian) can be used to represent uncertainty.

- Under minimal rationality assumptions, probability can be shown to be the only rational way to represent uncertainty.

Bayesian calibration approach

Probability (Bayesian) can be used to represent uncertainty.

- Under minimal rationality assumptions, probability can be shown to be the only rational way to represent uncertainty.
- Probability is subjective probability - distributions represent degrees of belief of individuals. There is no escaping this interpretation in many applications!

Bayesian calibration approach

Probability (Bayesian) can be used to represent uncertainty.

- Under minimal rationality assumptions, probability can be shown to be the only rational way to represent uncertainty.
- Probability is subjective probability - distributions represent degrees of belief of individuals. There is no escaping this interpretation in many applications!
- All uncertainty quantities θ can be given distributions $\pi(\theta)$ that represent our (an experts?) uncertainty about the value - this doesn't mean that they are random quantities, just that we don't know their value.
    - Even unknown function will be described by probability distributions across a class of unknown functions

Bayesian calibration approach

Probability (Bayesian) can be used to represent uncertainty.

- Under minimal rationality assumptions, probability can be shown to be the only rational way to represent uncertainty.
- Probability is subjective probability - distributions represent degrees of belief of individuals. There is no escaping this interpretation in many applications!
- All uncertainty quantities θ can be given distributions $\pi(\theta)$ that represent our (an experts?) uncertainty about the value - this doesn't mean that they are random quantities, just that we don't know their value.
  - Even unknown function will be described by probability distributions across a class of unknown functions
- Bayesian approach uses the principle of conditionality, and always (where possible) conditions on our data.

Bayesian Inference

The basics of Bayesian inference are very simple.

- If quantity θ is unknown we describe our uncertainty by prior distribution $\pi(\theta)$.
- Suppose we have data D from model $\pi(D|\theta)$
- Then conditional on observing this data, our posterior distribution is

$$\pi(\theta|D) = \frac{\pi(\theta)\pi(D|\theta)}{\pi(D)}$$

In general we just use the relationship

$$\text{posterior } \propto \text{prior} \times \text{likelihood (model)}$$

Given a model and prior the entire Bayesian statistical approach is fully specified. In practice, computational difficulties make life difficult, and we must resort to Monte Carlo methods.

Bayesian calibration

The result of a Bayesian calibration is a posterior distribution $\pi(\theta \mid \mathcal{D})$. This summarizes all the available information.

Bayesian calibration

The result of a Bayesian calibration is a posterior distribution $\pi(\theta \mid \mathcal{D})$. This summarizes all the available information.

- Sometimes the posterior is reported as just a mode (maximum a posteriori - MAP)

$$\theta_{MAP} = \arg\max_{\theta} \pi(\theta \mid \mathcal{D})$$

Bayesian calibration

The result of a Bayesian calibration is a posterior distribution $\pi(\theta \mid \mathcal{D})$. This summarizes all the available information.

- Sometimes the posterior is reported as just a mode (maximum a posteriori - MAP)

$$\theta_{MAP} = \arg\max_{\theta} \pi(\theta \mid \mathcal{D})$$

- and perhaps a credibility interval. For example, the two-tailed 95% highest probability credibility interval is defined by the two end points

$$\theta_l = \{\theta : \pi(\theta|\mathcal{D}) = 0.025\} \qquad \theta_u = \{\theta : \pi(\theta|\mathcal{D}) = 0.975\}$$

Bayesian calibration

The result of a Bayesian calibration is a posterior distribution $\pi(\theta \mid \mathcal{D})$. This summarizes all the available information.

- Sometimes the posterior is reported as just a mode (maximum a posteriori - MAP)

$$\theta_{MAP} = \arg\max_{\theta} \pi(\theta \mid \mathcal{D})$$

- and perhaps a credibility interval. For example, the two-tailed 95% highest probability credibility interval is defined by the two end points

$$\theta_l = \{\theta : \pi(\theta|\mathcal{D}) = 0.025\} \qquad \theta_u = \{\theta : \pi(\theta|\mathcal{D}) = 0.975\}$$

- i.e., there is 95% posterior probability that $\theta \in [\theta_l, \theta_u]$ with θ_{MAP} the most likely value.

Model error

Relating simulators to the physical system

In a principled statistical analysis (as opposed to an ad-hoc one) we must carefully relate the simulator prediction to the system.

Relating simulators to the physical system

In a principled statistical analysis (as opposed to an ad-hoc one) we must carefully relate the simulator prediction to the system.

This is often done by specification of the measurement error on the data:

$$\mathcal{D}_{field} = \eta(\hat{\theta}) + e$$

where $e \sim N(0, \sigma^2_{obs})$ then

$$\pi(D|\theta) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{obs}}} \exp\left(-\frac{(D_i - \eta_i(\hat{\theta}))^2}{2}\right)$$

Relating simulators to the physical system

In a principled statistical analysis (as opposed to an ad-hoc one) we must carefully relate the simulator prediction to the system.

This is often done by specification of the measurement error on the data:

$$\mathcal{D}_{field} = \eta(\hat{\theta}) + e$$

where $e \sim N(0, \sigma^2_{obs})$ then

$$\pi(D|\theta) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{obs}}} \exp\left(-\frac{(D_i - \eta_i(\hat{\theta}))^2}{2}\right)$$

This type of calibration is assuming a **perfect model hypothesis** - i.e.,

that our model is capable of exactly reproducing reality.

Model error

Once we have found $\pi(\theta|\mathcal{D}, \eta)$, what do we think this represents?

Model error

Once we have found $\pi(\theta|\mathcal{D}, \eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?

Model error

Once we have found $\pi(\theta|\mathcal{D},\eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?
- Or is θ interpretted statistically? i.e., θ is the value that best explains the data given the model - cf. the coefficients in a linear regression.

Model error

Once we have found $\pi(\theta|\mathcal{D},\eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?
- Or is θ interpretted statistically? i.e., θ is the value that best explains the data given the model - cf. the coefficients in a linear regression.

e.g., ocean physics models must be run with viscosity several orders of magnitude too large - treacle oceans - because there are errors in the climate model.

Model error

Once we have found $\pi(\theta|\mathcal{D}, \eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?
- Or is θ interpretted statistically? i.e., θ is the value that best explains the data given the model - cf. the coefficients in a linear regression.

e.g., ocean physics models must be run with viscosity several orders of magnitude too large - treacle oceans - because there are errors in the climate model.

When can we interpret the value found for θ as a physical value?

Model error

Once we have found $\pi(\theta|\mathcal{D}, \eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?
- Or is θ interpretted statistically? i.e., θ is the value that best explains the data given the model - cf. the coefficients in a linear regression.

e.g., ocean physics models must be run with viscosity several orders of magnitude too large - treacle oceans - because there are errors in the climate model.

When can we interpret the value found for θ as a physical value?

- If the model is a perfect representation of the system

Model error

Once we have found $\pi(\theta|\mathcal{D}, \eta)$, what do we think this represents?

- Does θ have a physical interpretation, i.e., are we estimating physical parameters?
- Or is θ interpretted statistically? i.e., θ is the value that best explains the data given the model - cf. the coefficients in a linear regression.

e.g., ocean physics models must be run with viscosity several orders of magnitude too large - treacle oceans - because there are errors in the climate model.

When can we interpret the value found for θ as a physical value?

- If the model is a perfect representation of the system
- When the model is imperfect, but we have a description (that we believe) of the discrepancy between model and system.

'All models are wrong, but some are useful' George Box

Lets acknowledge that most models are imperfect.

‘All models are wrong, but some are useful’ George Box

Lets acknowledge that most models are imperfect.
Consequently,

- predictions will be wrong, or will be made with misleading degree of confidence

'All models are wrong, but some are useful' George Box

Lets acknowledge that most models are imperfect.
Consequently,

- predictions will be wrong, or will be made with misleading degree of confidence
- solving the inverse problem $y = \eta(\theta) + e$ may not give sensible results.
  - e is measurement error
  - $\eta(\theta)$ is our computer model
  - y is our data

‘All models are wrong, but some are useful’ George Box

Lets acknowledge that most models are imperfect.
Consequently,

- predictions will be wrong, or will be made with misleading degree of confidence
- solving the inverse problem $y = \eta(\theta) + e$ may not give sensible results.
  - e is measurement error
  - $\eta(\theta)$ is our computer model
  - y is our data
- If the model is imperfect, the residuals $D - \eta(\theta)$ will be correlated, even if the real measurement error process is iid.

Can we

- account for the error?
- correct the error?

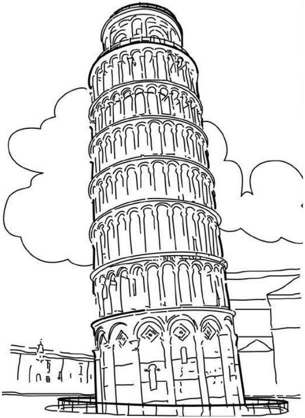
Can we

- account for the error?
- correct the error?

Kennedy and O'Hagan (2001) suggested we introduce reality ζ into our statistical inference

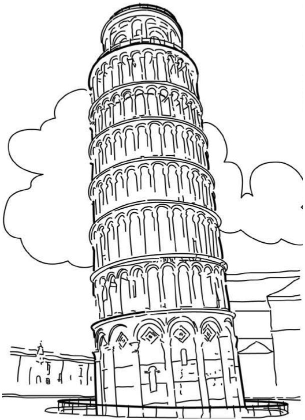
- Reality $\zeta = \eta(\hat{\theta}) + \delta$, the best model prediction plus model error $\delta(x)$.
- Data $y = \zeta + e$ where e represents measurement error

Toy Example: Freefall



Consider an experiment where we drop a weight from a tower and measure its position x_t every Δt seconds.

Toy Example: Freefall

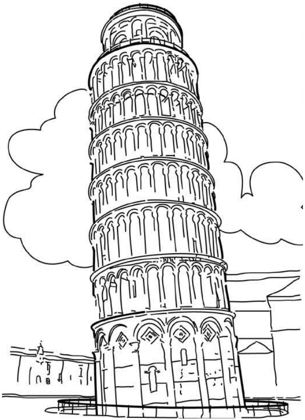


Consider an experiment where we drop a weight from a tower and measure its position x_t every Δt seconds.

We observe reality ζ with independent identically distributed errors

- Noisy observation: $D_t \sim N(x_t, \sigma^2_{obs})$

Toy Example: Freefall



Consider an experiment where we drop a weight from a tower and measure its position x_t every Δt seconds.

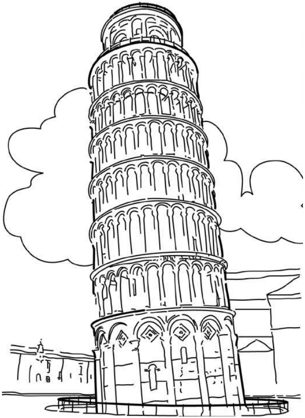
We observe reality ζ with independent identically distributed errors

- Noisy observation: $D_t \sim N(x_t, \sigma^2_{obs})$

i.e.,

$$D_t = \zeta_t + e_t \qquad \text{where } e_t \sim N(0, \sigma^2_{obs})$$

Toy Example: Freefall

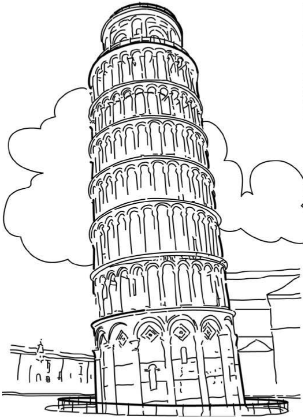


Assume that the 'true' dynamics (reality) are given by

$$\frac{\mathrm{dv}}{\mathrm{dt}} = g - kv$$

gravity minus a Stokes' drag term for air-resistance.

Toy Example: Freefall



Assume that the 'true' dynamics (reality) are given by

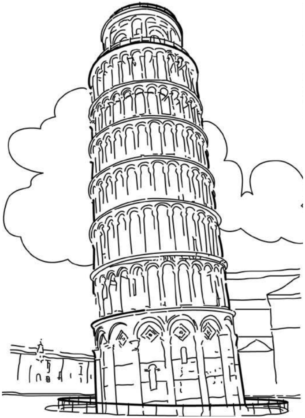
$$\frac{\mathrm{dv}}{\mathrm{dt}} = g - kv$$

gravity minus a Stokes' drag term for air-resistance.

This gives single time step updates

$$\begin{aligned}
x_{t+1} &= \zeta_x(x_t, v_t) \\
&= x_t + \frac{1}{k}(\frac{g}{k} - v_t)(\mathrm{e}^{-\mathrm{k}\Delta \mathrm{t}} - 1) + \frac{\mathrm{g}\Delta \mathrm{t}}{\mathrm{k}} \\
v_{t+1} &= \zeta_x(x_t, v_t) \\
&= (v_t - \frac{g}{k})\mathrm{e}^{-\mathrm{k}\Delta \mathrm{t}} + \frac{\mathrm{g}}{\mathrm{k}}
\end{aligned}$$

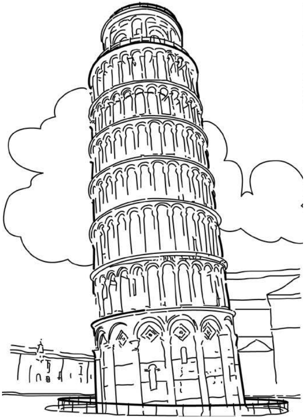
Toy Model Error Example: Freefall



Suppose we are given a computer model based on

$$\frac{\mathrm{dv}}{\mathrm{dt}} = g$$

Toy Model Error Example: Freefall



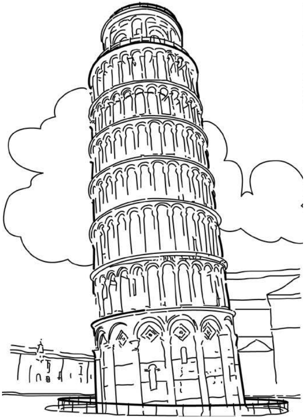
Suppose we are given a computer model based on

$$\frac{\mathrm{dv}}{\mathrm{dt}} = g$$

Which gives predictions at the observations of

$$\begin{aligned} x_{t+1} &= \eta_x(x_t, v_t) \\ &= x_t + v_k \Delta t + \frac{1}{2} g (\Delta t)^2 \\ v_{t+1} &= \eta_v(x_t, v_t) \\ &= v_t + g \Delta t \end{aligned}$$

Toy Model Error Example: Freefall



Suppose we are given a computer model based on

$$\frac{\mathrm{dv}}{\mathrm{dt}} = g$$

Which gives predictions at the observations of

$$\begin{aligned} x_{t+1} &= \eta_x(x_t, v_t) \\ &= x_t + v_k \Delta t + \frac{1}{2} g (\Delta t)^2 \\ v_{t+1} &= \eta_v(x_t, v_t) \\ &= v_t + g\Delta t \end{aligned}$$

Because we have ignored air-resistance in our model, we have model error!

Model Error Term

In this toy problem, the true discrepancy function can be calculated.

- It is a two dimensional function

  $$\delta = \begin{pmatrix} \delta_x \\ \delta_v \end{pmatrix} = \zeta - \eta$$

  giving the difference between the reality and the prediction from our model.

Model Error Term

In this toy problem, the true discrepancy function can be calculated.

- It is a two dimensional function

$$\delta = \begin{pmatrix} \delta_x \\ \delta_v \end{pmatrix} = \zeta - \eta$$

giving the difference between the reality and the prediction from our model.

If we expand $\mathrm{e}^{-\mathrm{k}\Delta\mathrm{t}}$ to second order we find

$$\delta(x, v, t) = \begin{pmatrix} \delta_x \\ \delta_v \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{-gk(\Delta t)^2}{2} \end{pmatrix} - v_t \begin{pmatrix} \frac{k(\Delta t)^2}{2} \\ k\Delta t(1 - \frac{k\Delta t}{2}) \end{pmatrix}$$

Model Error Term

In this toy problem, the true discrepancy function can be calculated.

- It is a two dimensional function

$$\delta = \begin{pmatrix} \delta_x \\ \delta_v \end{pmatrix} = \zeta - \eta$$

giving the difference between the reality and the prediction from our model.

If we expand $\mathrm{e}^{-\mathrm{k}\Delta\mathrm{t}}$ to second order we find

$$\delta(x, v, t) = \begin{pmatrix} \delta_x \\ \delta_v \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{-gk(\Delta t)^2}{2} \end{pmatrix} - v_t \begin{pmatrix} \frac{k(\Delta t)^2}{2} \\ k\Delta t(1 - \frac{k\Delta t}{2}) \end{pmatrix}$$

This is solely a function of v.

- Note, to learn δ we only have the observations $D_1, \ldots, D_T$ of $x_1, \ldots, x_T$ - we do not observe v.

Expected form of the discrepancy

Forget the previous slide. If we wanted a statistical model for δ what would we choose?

Expected form of the discrepancy

Forget the previous slide. If we wanted a statistical model for δ what would we choose?

There are three variables in this problem, displacement, velocity and time (x, v, t) so we might think to model δ as a function of these three terms.

Expected form of the discrepancy

Forget the previous slide. If we wanted a statistical model for δ what would we choose?

There are three variables in this problem, displacement, velocity and time (x, v, t) so we might think to model δ as a function of these three terms.

However, the principal of universality says that nature is consistent throughout all space and time (background independence), so with a little thought we might reason that δ should be independent of x and t.

Expected form of the discrepancy

Forget the previous slide. If we wanted a statistical model for δ what would we choose?

There are three variables in this problem, displacement, velocity and time (x, v, t) so we might think to model δ as a function of these three terms.

However, the principal of universality says that nature is consistent throughout all space and time (background independence), so with a little thought we might reason that δ should be independent of x and t.

We could use something very flexible and generic, such as a Gaussian process (which can fit any functional form)

$$\delta(x, v, t) \sim GP(m(v), \sigma^2 c(v, v'))$$

Expected form of the discrepancy

With input from an experienced user of our model, it is feasible we might be able to get other information such as that δ approximately scales with v, or at least that the error is small at low speeds and large at high speeds.

Expected form of the discrepancy

With input from an experienced user of our model, it is feasible we might be able to get other information such as that δ approximately scales with v, or at least that the error is small at low speeds and large at high speeds.

This might suggest either using a linear mean in our Gaussian process ($m(x, v, t) = \alpha v$) or using a simpler model for δ such as a linear regression

$$\delta(x, v, t) = \alpha v + \epsilon \qquad \text{where } \epsilon \sim N(0, \sigma^2_{model})$$

Expected form of the discrepancy

With input from an experienced user of our model, it is feasible we might be able to get other information such as that δ approximately scales with v, or at least that the error is small at low speeds and large at high speeds.

This might suggest either using a linear mean in our Gaussian process ($m(x, v, t) = \alpha v$) or using a simpler model for δ such as a linear regression

$$\delta(x, v, t) = \alpha v + \epsilon \qquad \text{where } \epsilon \sim N(0, \sigma^2_{model})$$

- See Wilkinson and Oakley 2010 for a method to estimate a functional form for δ.

Summary of the Bayesian calibration framework

- Assume that we observe reality plus measurement error.

$$D(t) = \zeta(t) + e(t)$$

Often, $e(\cdot)$ will be a white noise process with known mean and variance!

Summary of the Bayesian calibration framework

- Assume that we observe reality plus measurement error.

$$D(t) = \zeta(t) + e(t)$$

  Often, $e(\cdot)$ will be a white noise process with known mean and variance!

- Introduce a model error (discrepancy) term. Assume that reality is the best model prediction plus an error

$$\zeta(t) = \eta(t, \hat{\theta}) + \delta(t).$$

  Note δ does not depend on θ.

Summary of the Bayesian calibration framework

- Assume that we observe reality plus measurement error.

$$D(t) = \zeta(t) + e(t)$$

  Often, $e(\cdot)$ will be a white noise process with known mean and variance!

- Introduce a model error (discrepancy) term. Assume that reality is the best model prediction plus an error

$$\zeta(t) = \eta(t, \hat{\theta}) + \delta(t).$$

  Note δ does not depend on θ.

- Argue that $\eta(\cdot, \hat{\theta})$ and $\delta(\cdot)$ are independent. Kennedy and O'Hagan 2001 use Gaussian processes to model both the model η and the error δ.

Calibrating the UVic Climate Model

Example: UVic Earth System Climate Model

With Nathan Urban (Penn State)

UVic ESCM is an intermediate complexity model with a general circulation ocean and dynamic/thermodynamic sea-ice components coupled to a simple energy/moisture balance atmosphere. It has a dynamic vegetation and terrestrial carbon cycle model (TRIFFID) as well as an inorganic carbon cycle.

Example: UVic Earth System Climate Model

With Nathan Urban (Penn State)

UVic ESCM is an intermediate complexity model with a general circulation ocean and dynamic/thermodynamic sea-ice components coupled to a simple energy/moisture balance atmosphere. It has a dynamic vegetation and terrestrial carbon cycle model (TRIFFID) as well as an inorganic carbon cycle.

- Inputs: Q_{10} = respiration sensitivity to temperature (carbon source) and $K_c = CO_2$ fertilization of photosynthesis (carbon sink).

Example: UVic Earth System Climate Model

With Nathan Urban (Penn State)

UVic ESCM is an intermediate complexity model with a general circulation ocean and dynamic/thermodynamic sea-ice components coupled to a simple energy/moisture balance atmosphere. It has a dynamic vegetation and terrestrial carbon cycle model (TRIFFID) as well as an inorganic carbon cycle.

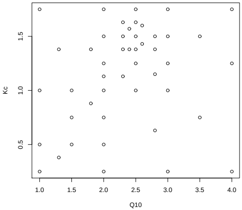
- Inputs: Q_{10} = respiration sensitivity to temperature (carbon source) and $K_c = CO_2$ fertilization of photosynthesis (carbon sink).
- Output: time-series of CO_2 values, cumulative carbon flux measurements, ...

Example: UVic Earth System Climate Model

With Nathan Urban (Penn State)

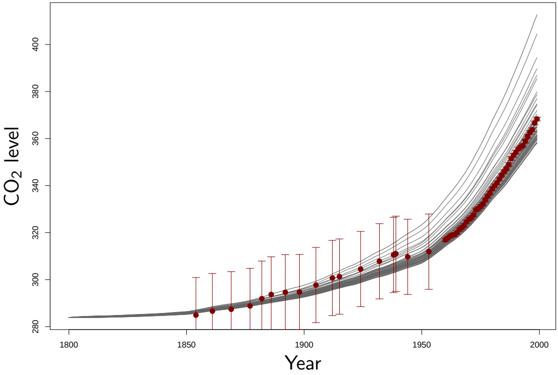
UVic ESCM is an intermediate complexity model with a general circulation ocean and dynamic/thermodynamic sea-ice components coupled to a simple energy/moisture balance atmosphere. It has a dynamic vegetation and terrestrial carbon cycle model (TRIFFID) as well as an inorganic carbon cycle.

- Inputs: Q_{10} = respiration sensitivity to temperature (carbon source) and $K_c = CO_2$ fertilization of photosynthesis (carbon sink).
- Output: time-series of CO_2 values, cumulative carbon flux measurements, ...
- 48 member ensemble, grid design D, output $\mathcal{D}_{sim}$ ($48 \times n$).



The U.Vic Model

- 60 field measurements, $\mathcal{D}_{field}$:
  - 40 instrumental CO_2 measurements from 1960-1999
  - 17 ice core CO_2 measurements
  - 3 cumulative ocean carbon flux measurements



Multivariate Models

Wilkinson 2010

The output from UVic is a temporal spatial field of predicted CO2 values. To emulate this we need an extension of the univariate GP method produced previously.

Multivariate Models
Wilkinson 2010

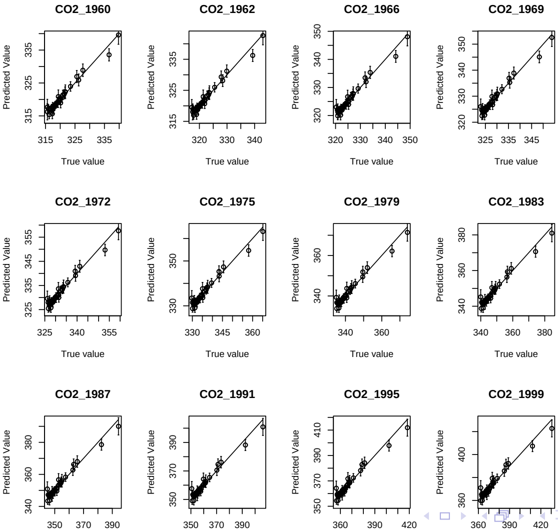
The output from UVic is a temporal spatial field of predicted CO2 values. To emulate this we need an extension of the univariate GP method produced previously. If model outputs are highly correlated we can reduce the dimension of the output by projecting the data onto some lower dimensional manifold $\mathcal{Y}^{pc}$.

The output from UVic is a temporal spatial field of predicted CO2 values. To emulate this we need an extension of the univariate GP method produced previously. If model outputs are highly correlated we can reduce the dimension of the output by projecting the data onto some lower dimensional manifold $\mathcal{Y}^{pc}$. We can then emulate the function that maps the input space Θ to the reduced dimensional output space $\mathcal{Y}^{pc}$.

$$
\begin{tikzcd}
\Theta \arrow[rr, "\eta(\cdot)"] \arrow[dr, "\eta^{pc}(\cdot)"'] & & \mathcal{Y} \arrow[dl, shift right, "\mathrm{PCA}"'] \\
& \mathcal{Y}^{pc} \arrow[ur, shift right, "\mathrm{PCA}^{-1}"'] &
\end{tikzcd}
$$

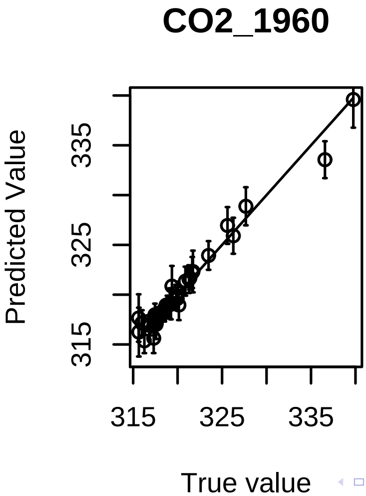
Emulator Diagnostics

Cross-validation plots, instrumental data only (1960:1999), using 3PCs (99.2% of variance explained)



Emulator Diagnostics

Cross-validation plots, instrumental data only (1960), using 3PCs (99.2% of variance explained)

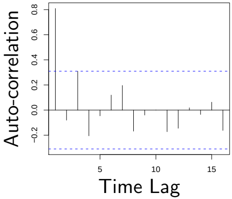


UVic Model Discrepancy

The calibration framework used is:

$$\mathcal{D}_{field}(t) = \eta(\theta, t) + \delta(t) + e(t)$$

The model predicts the underlying trend, but real climate fluctuates around this. We model



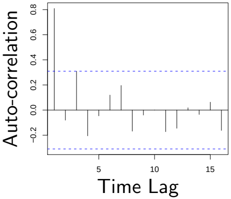
UVic Model Discrepancy

The calibration framework used is:

$$\mathcal{D}_{field}(t) = \eta(\theta, t) + \delta(t) + e(t)$$

The model predicts the underlying trend, but real climate fluctuates around this. We model

- discrepancy as an AR1 process: $\delta(0) \sim N(0, \sigma_\delta^2)$, and $\delta(t) = \rho\delta(t-1) + N(0, \sigma_\delta^2)$.



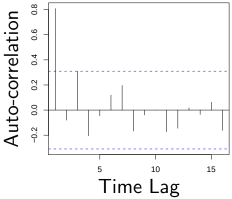
UVic Model Discrepancy

The calibration framework used is:

$$\mathcal{D}_{field}(t) = \eta(\theta, t) + \delta(t) + e(t)$$

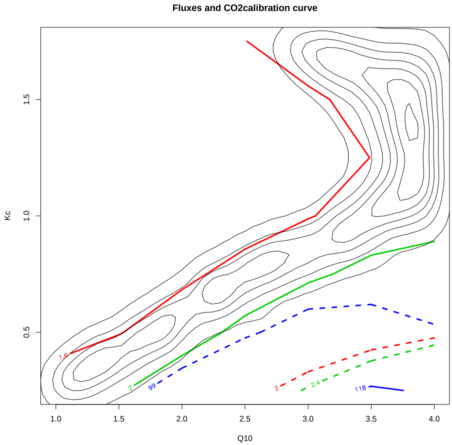
The model predicts the underlying trend, but real climate fluctuates around this. We model

- discrepancy as an AR1 process: $\delta(0) \sim N(0, \sigma_\delta^2)$, and $\delta(t) = \rho\delta(t-1) + N(0, \sigma_\delta^2)$.
- Measurement error as heteroscedastic independent random noise $e(t) \sim N(0, \lambda(t))$.



Results

After several hours of MCMC (Metropolis-within-Gibbs - see lecture 6) we find



The data are equally well explained by a range of parameter values.

Summary

- Calibration is the process of estimating parameters from empirical observations
- The Bayesian approach is ideal for measuring and summarizing any uncertainties.
- When calibrating models it is important to account for the existence of model error
  - we need a statistical relationship between the computer simulator output and the physical system.
- Accounting for model error is a relatively new idea in statistics. See Goldstein *et al.* 2009, Beven 2006, and Poole *et al.* 2000, for a selection of other approaches to calibration and model error.

References

- Kennedy M and O'Hagan A 2001 Bayesian calibration of computer models (with discussion). *Journal of the Royal Statistical Society, Series B* **63**, 425–464.
- Higdon D, Gattiker J, Williams B and Rightley M 2008 Computer model calibration using high-dimensional output. *JASA* **103**, 570-583.
- Wilkinson, to appear 2010, Bayesian calibration of expensive multivariate computer experiments, *Large-scale inverse problems and quantification of uncertainty*, Wiley.
- Beven K 2006 A manifesto for the equifinality thesis. *Journal of Hydrology* **320**, 18–36.
- Poole D and Raftery E 2000 Inference for deterministic simulation models: The Bayesian melding approach. *JASA* **95**, 1244-1255.
- Goldstein M and Rougier JC 2009 Reified Bayesian modelling and inference for physical systems, with discussion and rejoinder. *Journal of Statistical Planning and Inference* **139(3)**, 1221-1239.
- MUCM toolkit, available at `http://mucm.aston.ac.uk/MUCM/MUCMToolkit/`