Assessing climate uncertainty: models, meaning and methods

Michael Goldstein
Department of Mathematical Sciences,
Durham University *

*Thanks to EPSRC, NERC for funding, and to my collaborators in this work: Allan Seheult, Peter Craig, Jonty Rougier, Jonathan Cumming, Ian Vernon, Leanna House, Danny Williamson, Alan Roberts and the MUCM crew
Mathematical models, implemented as computer simulators, offer a powerful methodology for studying complex physical systems.

“Whenever possible, everything is done in a supercomputer,” he said.

“Look at Formula One - it’s getting rid of all of its wind tunnels and replacing them with supercomputers. It’s the same in the aerospace industry as well. It means you can all the modelling in the supercomputer and then do just one real world test.”

[Jon Lockley, manager of the Oxford Supercomputing Centre: quoted on BBC web-site]
Mathematical models, implemented as computer simulators, offer a powerful methodology for studying complex physical systems.

“Whenever possible, everything is done in a supercomputer,” he said.

“Look at Formula One - it’s getting rid of all of its wind tunnels and replacing them with supercomputers. It’s the same in the aerospace industry as well. It means you can all the modelling in the supercomputer and then do just one real world test.”

[Jon Lockley, manager of the Oxford Supercomputing Centre: quoted on BBC web-site]

Using complex simulators to make statements about physical systems raises many practical, technical and conceptual questions about the uncertainty involved in moving from the model to the system.
Models

Mathematical models, implemented as computer simulators, offer a powerful methodology for studying complex physical systems.

“Whenever possible, everything is done in a supercomputer,” he said.

“Look at Formula One - it’s getting rid of all of its wind tunnels and replacing them with supercomputers. It’s the same in the aerospace industry as well. It means you can all the modelling in the supercomputer and then do just one real world test.”

[Jon Lockley, manager of the Oxford Supercomputing Centre: quoted on BBC web-site]

Using complex simulators to make statements about physical systems raises many practical, technical and conceptual questions about the uncertainty involved in moving from the model to the system.

There is a general methodology being developed to deal with such problems.

A good resource: the Managing Uncertainty in Complex Models web-site

http://www.mucm.ac.uk/

(for references, papers, toolkit, etc.)
While different physical models may vary in many aspects, the formal structures involved in analysing the physical system through computer simulators are very similar (which is why there is a common underlying methodology).
While different physical models may vary in many aspects, the formal structures involved in analysing the physical system through computer simulators are very similar (which is why there is a common underlying methodology).

Each simulator can be conceived as a function $f(x)$, where $x$ is the (often high dimensional) input vector, representing unknown properties of the physical system; $f(x)$ is a (often high dimensional) output vector representing various aspects of the behaviour of the system.
General form

While different physical models may vary in many aspects, the formal structures involved in analysing the physical system through computer simulators are very similar (which is why there is a common underlying methodology).

Each simulator can be conceived as a function $f(x)$, where $x$ is the (often high dimensional) input vector, representing unknown properties of the physical system; $f(x)$ is a (often high dimensional) output vector representing various aspects of the behaviour of the system.

For example, in a climate model, $x$ might denote physical parameters determining the behaviour of various physical processes (relating to clouds, ice, convection, boundary layer, radiation and so forth) which are needed to construct a description of climate behaviour incorporating human interventions (such as levels of CO2 emissions in the future). A typical element of $f(x)$ might be, for example, the global mean temperature in 100 years time.
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.

the “appropriate” (in some sense) choice, $x^*$, for the system properties $x$, 
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.

the “appropriate” (in some sense) choice, $x^*$, for the system properties $x$,

how informative $f(x^*)$ is for actual system behaviour, $y$. In particular, how much confidence we may place in predictions of unobserved system behaviour, $y_p$, from the corresponding components $f_p(x^*)$, 
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.

the “appropriate” (in some sense) choice, $x^*$, for the system properties $x$,

how informative $f(x^*)$ is for actual system behaviour, $y$. In particular, how much confidence we may place in predictions of unobserved system behaviour, $y_p$, from the corresponding components $f_p(x^*)$,

the use that we can make of historical observations $z$, observed with error on a subset $y_h$ of $y$, both to test and to constrain the model,
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.
the “appropriate” (in some sense) choice, \( x^* \), for the system properties \( x \),
how informative \( f(x^*) \) is for actual system behaviour, \( y \). In particular, how much confidence we may place in predictions of unobserved system behaviour, \( y_P \), from the corresponding components \( f_P(x^*) \),
the use that we can make of historical observations \( z \), observed with error on a subset \( y_h \) of \( y \), both to test and to constrain the model,
the optimal assignment of any decision inputs, \( d \), in the model.
Aims of analysis

Interest in the model usually centres on gaining general qualitative insights plus some of the following.

the “appropriate” (in some sense) choice, $x^*$, for the system properties $x$,

how informative $f(x^*)$ is for actual system behaviour, $y$. In particular, how much confidence we may place in predictions of unobserved system behaviour, $y_p$, from the corresponding components $f_p(x^*)$,

the use that we can make of historical observations $z$, observed with error on a subset $y_h$ of $y$, both to test and to constrain the model,

the optimal assignment of any decision inputs, $d$, in the model.

In the climate model, $y_h$ corresponds to historical climate observations recorded over space and time, $y_p$ to future climate, and the “decisions” might correspond to different policy relevant choices such as carbon emission scenarios.
In the simplest version of this problem, where observations are made without error and the model is a precise reproduction of the system, we can write

\[ z = f_h(x^*) \]

where \( f_h(x) \) is the sub-vector of outputs of \( f(x) \) corresponding to the subset \( y_h = z \).
Inverse problems

In the simplest version of this problem, where observations are made without error and the model is a precise reproduction of the system, we can write

\[ z = f_h(x^*) \]

where \( f_h(x) \) is the sub-vector of outputs of \( f(x) \) corresponding to the subset \( y_h = z \).

We invert \( f_h \) to find \( x^* \) (either uniquely or as a family of solutions)
Inverse problems

In the simplest version of this problem, where observations are made without error and the model is a precise reproduction of the system, we can write

$$z = f_h(x^*)$$

where $f_h(x)$ is the sub-vector of outputs of $f(x)$ corresponding to the subset $y_h = z$.

We invert $f_h$ to find $x^*$ (either uniquely or as a family of solutions).

We predict future system behaviour, $y_p$, exactly from the components $f_p(x^*)$ which correspond to the elements of $y_p$. 
Inverse problems

In the simplest version of this problem, where observations are made without error and the model is a precise reproduction of the system, we can write

\[ z = f_h(x^*) \]

where \( f_h(x) \) is the sub-vector of outputs of \( f(x) \) corresponding to the subset \( y_h = z \).

We invert \( f_h \) to find \( x^* \) (either uniquely or as a family of solutions).

We predict future system behaviour, \( y_p \), exactly from the components \( f_p(x^*) \) which correspond to the elements of \( y_p \).

If the future output depends on decision inputs, then we may optimise \( f_p(x^*, d) \) over choices of \( d \).
Inverse problems

In the simplest version of this problem, where observations are made without error and the model is a precise reproduction of the system, we can write

\[ z = f_h(x^*) \]

where \( f_h(x) \) is the sub-vector of outputs of \( f(x) \) corresponding to the subset \( y_h = z \).

We invert \( f_h \) to find \( x^* \) (either uniquely or as a family of solutions).

We predict future system behaviour, \( y_p \), exactly from the components \( f_p(x^*) \) which correspond to the elements of \( y_p \).

If the future output depends on decision inputs, then we may optimise \( f_p(x^*, d) \) over choices of \( d \).

COMMENT: Inverting and optimising complicated high dimensional functions is a challenging technical problem.
In practice, the observations $z$ are made with error, and we must separate the uncertainty representation into two relations:

$$z = y_h \oplus e,$$

$$y = f(x^*)$$

where $e$ has some appropriate probabilistic specification, possibly involving parameters which require estimation.
(Notation: $U \oplus V$ denotes the sum $U + V$ of two random quantities, $U, V$ which are either independent, if there is a full probabilistic specification, or uncorrelated if there is only a second order specification.)
Statistical inversion

In practice, the observations $z$ are made with error, and we must separate the uncertainty representation into two relations:

$$z = y_h \oplus e,$$

$$y = f(x^*)$$

where $e$ has some appropriate probabilistic specification, possibly involving parameters which require estimation.

(Notation: $U \oplus V$ denotes the sum $U + V$ of two random quantities, $U, V$ which are either independent, if there is a full probabilistic specification, or uncorrelated if there is only a second order specification.)

COMMENT: High dimensional statistical inversion combines challenging numerical and statistical problems. Multiple solutions potentially a big issue.
Condition uncertainty

Most computer models depend on many further quantities, $C$ say, which are unknown but require specification for the computer model to be assessed. Examples of such conditions relate to Initial conditions, Boundary conditions, Forcing functions [and, for practical reasons, many parametric inputs to the model will often be treated as fixed]
Condition uncertainty

Most computer models depend on many further quantities, $C$ say, which are unknown but require specification for the computer model to be assessed. Examples of such conditions relate to Initial conditions, Boundary conditions, Forcing functions [and, for practical reasons, many parametric inputs to the model will often be treated as fixed] So, our actual function is $f(x, C)$, with unknown “best” choices $x^*, C^*$. 
Condition uncertainty

Most computer models depend on many further quantities, $C$ say, which are unknown but require specification for the computer model to be assessed. Examples of such conditions relate to Initial conditions, Boundary conditions, Forcing functions [and, for practical reasons, many parametric inputs to the model will often be treated as fixed]
So, our actual function is $f(x, C)$, with unknown “best” choices $x^*, C^*$. For illustration, suppose that we carry out our computer experiment with a single fixed choice $C^+$. [Averaging over a choice of $C_i$ values is similar.]
**Condition uncertainty**

Most computer models depend on many further quantities, \( C \) say, which are unknown but require specification for the computer model to be assessed. Examples of such conditions relate to Initial conditions, Boundary conditions, Forcing functions [and, for practical reasons, many parametric inputs to the model will often be treated as fixed]

So, our actual function is \( f(x, C) \), with unknown “best” choices \( x^*, C^* \).

For illustration, suppose that we carry out our computer experiment with a single fixed choice \( C^+ \). [Averaging over a choice of \( C_i \) values is similar.]

We can assess the additional uncertainty introduced by our lack of knowledge of the appropriate choice of \( C \) in two ways.
Assessing condition uncertainty

Conduct a targeted experiment. Select $C_1, \ldots, C_n$, plausible choices of $C$. Evaluate choices of $f(x_i, C_j)$ for selected $x_i$. [Many samples if possible, otherwise some is better than none!]
Assessing condition uncertainty

Conduct a targeted experiment. Select $C_1, \ldots, C_n$, plausible choices of $C$.
Evaluate choices of $f(x_i, C_j)$ for selected $x_i$.
[Many samples if possible, otherwise some is better than none!]
(i) Quantify “order of magnitude” condition uncertainty

giving approximate relation

$$f(x, C^*) = f(x, C^+) \oplus \epsilon_C$$

$[\text{Var} (\epsilon_C)$ expresses order of magnitude variation over different choices of $C$. The covariance structure reflects similarities in the effect of condition uncertainty on different outputs.]$
Assessing condition uncertainty

Conduct a targeted experiment. Select $C_1, \ldots, C_n$, plausible choices of $C$. Evaluate choices of $f(x_i, C_j)$ for selected $x_i$. [Many samples if possible, otherwise some is better than none!]

(i) Quantify “order of magnitude” condition uncertainty giving approximate relation

$$f(x, C^*) = f(x, C^+) \oplus \epsilon_C$$

$[\text{Var}(\epsilon_C)]$ expresses order of magnitude variation over different choices of $C$. The covariance structure reflects similarities in the effect of condition uncertainty on different outputs.]

(ii) Quantify actual condition uncertainty Based on the experiment, model variability in $f(x, C)$ as a function of $x$, so

$$f(x, C^*) = f(x, C^+) \oplus \epsilon_C(x)$$

$[\text{There are natural “exchangeability” representations to help us to do this.}]$
Function emulation

Statistical inversion, for high dimensional problems, is even more challenging if the function $f(x)$ is expensive, in time and computational resources, to evaluate for any choice of $x$. [For example, large climate models.]
Function emulation

Statistical inversion, for high dimensional problems, is even more challenging if the function $f(x)$ is expensive, in time and computational resources, to evaluate for any choice of $x$. [For example, large climate models.]

In such cases, $f$ must be treated as uncertain for all input choices except the small subset for which an actual evaluation has been made.
Statistical inversion, for high dimensional problems, is even more challenging if the function $f(x)$ is expensive, in time and computational resources, to evaluate for any choice of $x$. [For example, large climate models.]

In such cases, $f$ must be treated as uncertain for all input choices except the small subset for which an actual evaluation has been made. Therefore, we must construct of a description of the uncertainty about the value of $f(x)$ for each $x$.

Such a representation is often termed an emulator of the function - the emulator both suggests an approximation to the function and also contains an assessment of the likely magnitude of the error of the approximation.
Form of the emulator

We may represent beliefs about component $f_i$ of $f$, using an emulator of the form,

$$f_i(x) = \sum_j \beta_{ij} g_{ij}(x) \oplus u_i(x)$$
Form of the emulator

We may represent beliefs about component $f_i$ of $f$, using an emulator of the form,

$$f_i(x) = \sum_j \beta_{ij} g_{ij}(x) \oplus u_i(x)$$

where $B = \{\beta_{ij}\}$ are unknown scalars, $g_{ij}$ are known deterministic functions of $x$, 
Form of the emulator

We may represent beliefs about component \( f_i \) of \( f \), using an emulator of the form,

\[
f_i(x) = \sum_j \beta_{ij} g_{ij}(x) \oplus u_i(x)
\]

where \( B = \{\beta_{ij}\} \) are unknown scalars, \( g_{ij} \) are known deterministic functions of \( x \), \( u_i(x) \) is a weakly second order stationary stochastic process, with (for example) correlation function

\[
\text{Corr}(u_i(x), u_i(x')) = \exp\left(-\left(\frac{\|x - x'\|}{\theta_i}\right)^2\right)
\]
Form of the emulator

We may represent beliefs about component $f_i$ of $f$, using an emulator of the form,

$$f_i(x) = \sum_j \beta_{ij} g_{ij}(x) \oplus u_i(x)$$

where $B = \{\beta_{ij}\}$ are unknown scalars, $g_{ij}$ are known deterministic functions of $x$, $u_i(x)$ is a weakly second order stationary stochastic process, with (for example) correlation function

$$\text{Corr}(u_i(x), u_i(x')) = \exp\left(-\left(\frac{\|x - x'\|}{\theta_i}\right)^2\right)$$

We fit the emulators, given a collection of carefully chosen model evaluations, using our favourite statistical tools, with a substantial component of expert judgement, supported by a careful diagnostic analysis.
Linked emulators

If the simulator is really slow to evaluate, then we emulate by jointly modelling the simulator with a fast approximate version, $f'$, plus older generations of the simulator which we’ve already emulated and so forth.
Linked emulators

If the simulator is really slow to evaluate, then we emulate by jointly modelling the simulator with a fast approximate version, $f'$, plus older generations of the simulator which we’ve already emulated and so forth. So, for example, based on many fast simulator evaluations, we build emulator

$$f'_i(x) = \sum_j \beta'_{ij} g_{ij}(x) \oplus u'_i(x)$$
**Linked emulators**

If the simulator is really slow to evaluate, then we emulate by jointly modelling the simulator with a fast approximate version, $f'$, plus older generations of the simulator which we’ve already emulated and so forth. So, for example, based on many fast simulator evaluations, we build emulator

$$f'_i(x) = \sum_j \beta'_{ij} g_{ij}(x) \oplus u'_i(x)$$

We use this form as the prior for the emulator for $f_i(x)$. Then a relatively small number of evaluations of $f(x)$, using relations such as

$$\beta_{ij} = \alpha_i \beta'_{ij} + \gamma_{ij}$$

lets us adjust the prior emulator to an appropriate posterior emulator for $f_i(x)$. 
Linked emulators

If the simulator is really slow to evaluate, then we emulate by jointly modelling the simulator with a fast approximate version, \( f' \), plus older generations of the simulator which we’ve already emulated and so forth. So, for example, based on many fast simulator evaluations, we build emulator

\[
f'_i(x) = \sum_j \beta'_{ij} g_{ij}(x) \oplus u'_i(x)
\]

We use this form as the prior for the emulator for \( f_i(x) \). Then a relatively small number of evaluations of \( f_i(x) \), using relations such as

\[
\beta_{ij} = \alpha_i \beta'_{ij} + \gamma_{ij}
\]

lets us adjust the prior emulator to an appropriate posterior emulator for \( f_i(x) \). [This approach exploits the heuristic that we need many more function evaluations to identify the qualitative form of the model (i.e. choose appropriate forms \( g_{ij}(x) \), etc) than to assess the quantitative form of all of the terms in the model - particularly if we fit meaningful regression components.]
Internal uncertainty analysis

Specification of
a prior distribution for $x^*$,
a likelihood for the observational error $e$,
a probabilistic emulator for $f$,
and some type of treatment of condition uncertainty,
Internal uncertainty analysis

Specification of
a prior distribution for $x^*$,
a likelihood for the observational error $e$,
a probabilistic emulator for $f$,
and some type of treatment of condition uncertainty,
updated by observation of the data $z$, using Bayes theorem,
Internal uncertainty analysis

Specification of
a prior distribution for $x^*$,
a likelihood for the observational error $e$,
a probabilistic emulator for $f$,
and some type of treatment of condition uncertainty,
updated by observation of the data $z$, using Bayes theorem,
gives a natural Bayesian treatment of the statistical inverse problem.
Internal uncertainty analysis

Specification of
a prior distribution for $x^*$,
a likelihood for the observational error $e$,
a probabilistic emulator for $f$,
and some type of treatment of condition uncertainty,
updated by observation of the data $z$, using Bayes theorem,
gives a natural Bayesian treatment of the statistical inverse problem.

We term this an **internal uncertainty analysis** - i.e. an analysis of all of the uncertainties arising directly in the Bayesian specification of the problem. It is a state of the art way of carrying out an uncertainty analysis based on a computer simulator for a physical system.
**Internal uncertainty analysis**

Specification of

- a prior distribution for $x^*$,
- a likelihood for the observational error $e$,
- a probabilistic emulator for $f$,

and some type of treatment of condition uncertainty,

updated by observation of the data $z$, using Bayes theorem,

gives a natural Bayesian treatment of the statistical inverse problem.

We term this an **internal uncertainty analysis** - i.e. an analysis of all of the uncertainties arising directly in the Bayesian specification of the problem. It is a state of the art way of carrying out an uncertainty analysis based on a computer simulator for a physical system.

The analysis is conceptually straightforward, though often technically challenging, requiring particular care when constructing the emulator for the function, dealing with condition uncertainty and handling the computational difficulties arising from high dimensional and often highly multimodal likelihood functions for high dimensional input and output spaces.
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation). Forecasting
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).
Forecasting
Control (or, at least influencing) the system
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).
Forecasting
Control (or, at least influencing) the system

COMMENT Within the Bayesian approach, we have two choices.
(i) Full Bayes analysis, with complete joint probabilistic specification of all of the uncertain quantities in the problem
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).

Forecasting

Control (or, at least influencing) the system

COMMENT Within the Bayesian approach, we have two choices. (i) Full Bayes analysis, with complete joint probabilistic specification of all of the uncertain quantities in the problem

or

(ii) Bayes linear analysis, based on a prior specification of the means, variances and covariances of all quantities of interest.
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).

Forecasting

Control (or, at least influencing) the system

COMMENT Within the Bayesian approach, we have two choices.

(i) Full Bayes analysis, with complete joint probabilistic specification of all of the uncertain quantities in the problem or

(ii) Bayes linear analysis, based on a prior specification of the means, variances and covariances of all quantities of interest.

Full Bayes analysis is more informative if done extremely carefully, both in terms of the prior specification and the analysis. Bayes linear analysis is partial but much easier and faster and typically more robust.
Some objectives of internal analysis

Model calibration (assessing whether the model can be calibrated, whether there are no/many solutions, possibly as a prelude to data assimilation).
Forecasting
Control (or, at least influencing) the system

COMMENT Within the Bayesian approach, we have two choices.
(i) Full Bayes analysis, with complete joint probabilistic specification of all of the uncertain quantities in the problem
or
(ii) Bayes linear analysis, based on a prior specification of the means, variances and covariances of all quantities of interest.

Full Bayes analysis is more informative if done extremely carefully, both in terms of the prior specification and the analysis. Bayes linear analysis is partial but much easier and faster and typically more robust.

In particular, there is a very fast forecasting methodology which does not require pre-calibration, which is very useful for exploring the implications of different scenarios.
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$.
We may have three concerns about such a procedure.
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$
We may have three concerns about such a procedure.
(i) Bayes calibration analysis may be very difficult/non-robust;
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$

We may have three concerns about such a procedure.
(i) Bayes calibration analysis may be very difficult/non-robust;
(ii) we may not believe in a unique true input value for the model;
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$.

We may have three concerns about such a procedure.

(i) Bayes calibration analysis may be very difficult/non-robust;
(ii) we may not believe in a unique true input value for the model;
(iii) we may be unsure whether there are any good choices of input parameters (due to model deficiencies)
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$
We may have three concerns about such a procedure.
(i) Bayes calibration analysis may be very difficult/non-robust;
(ii) we may not believe in a unique true input value for the model;
(iii) we may be unsure whether there are any good choices of input parameters
(due to model deficiencies)
A conceptually simple alternative is “history matching”, i.e. finding the collection
of all input choices $x$ for which you judge $\|y_h - f_h(x, C^*)\|$ to be acceptably
small, in some natural probabilistic metric.
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$
We may have three concerns about such a procedure.
(i) Bayes calibration analysis may be very difficult/non-robust;
(ii) we may not believe in a unique true input value for the model;
(iii) we may be unsure whether there are any good choices of input parameters
(due to model deficiencies)
A conceptually simple alternative is “history matching”, i.e. finding the collection
of all input choices $x$ for which you judge $\|y_h - f_h(x, C^*)\|$ to be acceptably
small, in some natural probabilistic metric.
History matching is (relatively) fast and straightforward, based on iteratively
sampling the parameter space, emulating, restricting the input space,
resampling and re-emulating within the reduced space, and so forth. It fits
naturally within the Bayes linear approach, and it is simple to introduce the
notion of “tolerance of model discrepancy” to denote the extent to which
different aspects of the model may fail to reproduce the corresponding aspects
of the physical system.
Calibration or history matching?

Model calibration aims to identify the “true” input parameters choices $x^*$
We may have three concerns about such a procedure.
(i) Bayes calibration analysis may be very difficult/non-robust;
(ii) we may not believe in a unique true input value for the model;
(iii) we may be unsure whether there are any good choices of input parameters
(due to model deficiencies)
A conceptually simple alternative is “history matching”, i.e. finding the collection
of all input choices $x$ for which you judge $\|y_h - f_h(x, C^*)\|$ to be acceptably
small, in some natural probabilistic metric.
History matching is (relatively) fast and straightforward, based on iteratively
sampling the parameter space, emulating, restricting the input space,
resampling and re-emulating within the reduced space, and so forth. It fits
naturally within the Bayes linear approach, and it is simple to introduce the
notion of “tolerance of model discrepancy” to denote the extent to which
different aspects of the model may fail to reproduce the corresponding aspects
of the physical system.
[Even if calibrating, history match first, to check model/reduce search space]
Limitations of internal analysis

The internal analysis would be correct if the model was a precise representation of the system, so that the only unknowns were the appropriate choices at which to evaluate the simulator.
Limitations of internal analysis

The internal analysis would be correct if the model was a precise representation of the system, so that the only unknowns were the appropriate choices at which to evaluate the simulator. A physical model is a description of the way in which system properties (the inputs to the model) affect system behaviour (the output of the model). This description involves two basic types of simplification.
Limitations of internal analysis

The internal analysis would be correct if the model was a precise representation of the system, so that the only unknowns were the appropriate choices at which to evaluate the simulator.

A physical model is a description of the way in which system properties (the inputs to the model) affect system behaviour (the output of the model). This description involves two basic types of simplification.

(i) we approximate the properties of the system (as these properties are too complicated to describe fully and anyway we don’t know them)
Limitations of internal analysis

The internal analysis would be correct if the model was a precise representation of the system, so that the only unknowns were the appropriate choices at which to evaluate the simulator.

A physical model is a description of the way in which system properties (the inputs to the model) affect system behaviour (the output of the model). This description involves two basic types of simplification.

(i) we approximate the properties of the system (as these properties are too complicated to describe fully and anyway we don’t know them)

(ii) we approximate the rules for finding system behaviour given system properties (because of necessary mathematical simplifications, simplifications for numerical tractability, and because we do not fully understand the physical laws which govern the process).
Limitations of internal analysis

The internal analysis would be correct if the model was a precise representation of the system, so that the only unknowns were the appropriate choices at which to evaluate the simulator. A physical model is a description of the way in which system properties (the inputs to the model) affect system behaviour (the output of the model). This description involves two basic types of simplification.

(i) we approximate the properties of the system (as these properties are too complicated to describe fully and anyway we don’t know them)

(ii) we approximate the rules for finding system behaviour given system properties (because of necessary mathematical simplifications, simplifications for numerical tractability, and because we do not fully understand the physical laws which govern the process).

Neither of these approximations invalidates the modelling process. Problems only arise when we forget these simplifications and confuse uncertainty analysis of the model with the corresponding uncertainty analysis for the physical system itself.
External uncertainty analysis

Models do not generate statements about reality, however carefully they are analysed. Such statements require *external uncertainty* analysis, taking account the mismatch between the simulator and the physical system.
External uncertainty analysis

Models do not generate statements about reality, however carefully they are analysed. Such statements require external uncertainty analysis, taking account the mismatch between the simulator and the physical system.

One of the simplest, and most popular, approaches to external uncertainty is to suppose that there is an appropriate choice of system properties $x^*$ (currently unknown), so that $f(x^*)$ contains all the information in the simulator about the system:

$$y = f(x^*) \oplus \epsilon_D$$
External uncertainty analysis

Models do not generate statements about reality, however carefully they are analysed. Such statements require external uncertainty analysis, taking account the mismatch between the simulator and the physical system.

One of the simplest, and most popular, approaches to external uncertainty is to suppose that there is an appropriate choice of system properties $x^*$ (currently unknown), so that $f(x^*)$ contains all the information in the simulator about the system:

$$y = f(x^*) \oplus \epsilon_D$$

$\epsilon_D$, often termed the model or structural discrepancy, has some appropriate probabilistic specification, possibly involving parameters which require estimation. [In practice, often, $\epsilon_D$ absorbs condition uncertainty, $\epsilon_C$, though it is better to keep them separate.]
External uncertainty analysis

Models do not generate statements about reality, however carefully they are analysed. Such statements require external uncertainty analysis, taking account the mismatch between the simulator and the physical system.

One of the simplest, and most popular, approaches to external uncertainty is to suppose that there is an appropriate choice of system properties $x^*$ (currently unknown), so that $f(x^*)$ contains all the information in the simulator about the system:

$$y = f(x^*) \oplus \epsilon_D$$

$\epsilon_D$, often termed the model or structural discrepancy, has some appropriate probabilistic specification, possibly involving parameters which require estimation. [In practice, often, $\epsilon_D$ absorbs condition uncertainty, $\epsilon_C$, though it is better to keep them separate.]

Much of the information in $\epsilon_D$ is conveyed by the covariance structure we assign, namely our view as to how simulator/system mismatches in the past are likely to translate into mismatches in the future.
Value of an external uncertainty analysis

Basic principle: it is always better to recognise than to ignore uncertainty, even if modelling and analysis of uncertainty is difficult and partial.
Value of an external uncertainty analysis

Basic principle: it is always better to recognise than to ignore uncertainty, even if modelling and analysis of uncertainty is difficult and partial.

For any non-trivial problem, even using the simple form for discrepancy is a big improvement on ignoring external uncertainty altogether.
Value of an external uncertainty analysis

Basic principle: it is always better to recognise than to ignore uncertainty, even if modelling and analysis of uncertainty is difficult and partial.

For any non-trivial problem, even using the simple form for discrepancy is a big improvement on ignoring external uncertainty altogether.

A Bayesian analysis based on carrying out an internal analysis with an added external discrepancy term would be very much the state of the art. (Indeed, mostly, substantially beyond the state of the art.)
Value of an external uncertainty analysis

Basic principle: it is always better to recognise than to ignore uncertainty, even if modelling and analysis of uncertainty is difficult and partial.

For any non-trivial problem, even using the simple form for discrepancy is a big improvement on ignoring external uncertainty altogether.

A Bayesian analysis based on carrying out an internal analysis with an added external discrepancy term would be very much the state of the art. (Indeed, mostly, substantially beyond the state of the art.)

The problem with an internal uncertainty analysis is that it fails to distinguish between the simulator and the system.

How much have we addressed this question by adding an external structural discrepancy into our treatment of uncertainty?
The meaning of an external uncertainty analysis

A Bayesian analysis based on the relations

\[ z = y_h \oplus e, \ y = f(x^*, C^*) \oplus \epsilon_D \]

results in a collection of uncertainty statements about \( x^* \) and \( y \).
The meaning of an external uncertainty analysis

A Bayesian analysis based on the relations

\[ z = y_h \oplus e, \quad y = f(x^*, C^*) \oplus \epsilon_D \]

results in a collection of uncertainty statements about \( x^* \) and \( y \).

What do these uncertainty statements mean?
The meaning of an external uncertainty analysis

A Bayesian analysis based on the relations

\[ z = y_h \oplus e, \quad y = f(x^*, C^*) \oplus \epsilon_D \]

results in a collection of uncertainty statements about \( x^* \) and \( y \).

What do these uncertainty statements mean?
While we may talk casually about the probability of some important event, when we are being careful then we need to be more precise about meaning.
The meaning of an external uncertainty analysis

A Bayesian analysis based on the relations

\[ z = y_h \oplus e, \ y = f(x^*, C^*) \oplus \epsilon_D \]

results in a collection of uncertainty statements about \( x^* \) and \( y \).

What do these uncertainty statements mean?

While we may talk casually about the probability of some important event, when we are being careful then we need to be more precise about meaning.

This quote from the BBC web-site is typical:

‘Fortunately, rapid climate change is one area that the UK has taken the lead in researching, by funding the Rapid Climate Change programme (RAPID), the aim of which is to determine the probability of rapid climate change occurring.’
The meaning of an external uncertainty analysis

A Bayesian analysis based on the relations

\[ z = y_h \oplus e, \quad y = f(x^*, C^*) \oplus \epsilon_D \]

results in a collection of uncertainty statements about \( x^* \) and \( y \).

What do these uncertainty statements mean?
While we may talk casually about the probability of some important event, when we are being careful then we need to be more precise about meaning.

This quote from the BBC web-site is typical:
‘Fortunately, rapid climate change is one area that the UK has taken the lead in researching, by funding the Rapid Climate Change programme (RAPID), the aim of which is to determine the probability of rapid climate change occurring.’

This means what exactly?
RAPID-WATCH

What are the implications of RAPID-WATCH observing system data and other recent observations for estimates of the risk due to rapid change in the MOC? In this context risk is taken to mean the probability of rapid change in the MOC and the consequent impact on climate (affecting temperatures, precipitation, sea level, for example). This project must:
RAPID-WATCH

What are the implications of RAPID-WATCH observing system data and other recent observations for estimates of the risk due to rapid change in the MOC? In this context risk is taken to mean the probability of rapid change in the MOC and the consequent impact on climate (affecting temperatures, precipitation, sea level, for example). This project must:
* contribute to the MOC observing system assessment in 2011;
* investigate how observations of the MOC can be used to constrain estimates of the probability of rapid MOC change, including magnitude and rate of change;
* make sound statistical inferences about the real climate system from model simulations and observations;
* investigate the dependence of model uncertainty on such factors as changes of resolution;
* assess model uncertainty in climate impacts and characterise impacts that have received less attention (eg frequency of extremes).

The project must also demonstrate close partnership with the Hadley Centre.
In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences.
Subjectivist Bayes

In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences. This interpretation has an agreed and testable meaning which is sufficiently precise that it is capable of forming the basis of a discussion about the meaning of the analysis.
In the subjectivist Bayes view, the meaning of any probability statement is straightforward. It is the uncertainty judgement of a specified individual, expressed on the scale of probability by consideration of some operational elicitation scheme, for example by consideration of betting preferences. This interpretation has an agreed and testable meaning which is sufficiently precise that it is capable of forming the basis of a discussion about the meaning of the analysis.

The choice of the subjectivist view of uncertainty does not settle the question as to the meaning of the computer uncertainty analysis, but it does allow us to pose it clearly. So, consider again what we mean by a statement such as ‘Fortunately, rapid climate change is one area that the UK has taken the lead in researching, by funding the Rapid Climate Change programme (RAPID), the aim of which is to determine the probability of rapid climate change occurring.’
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.

There is a big issue of perception here, as most people expect something more authoritative and objective than a probability which is one person’s judgement.
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.

There is a big issue of perception here, as most people expect something more authoritative and objective than a probability which is one person’s judgement.

However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.

There is a big issue of perception here, as most people expect something more authoritative and objective than a probability which is one person’s judgement.

However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.

So, it is not unreasonable that the objective of our analysis should be probabilities which are asserted by at least one person (more would be good!).
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.

There is a big issue of perception here, as most people expect something more authoritative and objective than a probability which is one person’s judgement.

However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.

So, it is not unreasonable that the objective of our analysis should be probabilities which are asserted by at least one person (more would be good!). Is this a sufficient objective?
The meaning of an uncertainty analysis

In the subjectivist interpretation, any probability statement is the judgement of a named individual, so we should speak not of the probability of rapid climate change, but instead of Anne’s probability or Bob’s probability of rapid climate change and so forth.

There is a big issue of perception here, as most people expect something more authoritative and objective than a probability which is one person’s judgement.

However, the disappointing thing is that, in almost all cases, stated probabilities emerging from a complex analysis are not even the judgements of any individual.

So, it is not unreasonable that the objective of our analysis should be probabilities which are asserted by at least one person (more would be good!).

Is this a sufficient objective?

[Alternative meanings of external uncertainty statements are fine - but all of the other meanings that I know of really reduce to modelling constructs which are useful for internal uncertainty analysis stage.]
The question of whether the Bayesian analysis of a simulator does indeed represent the judgements of the expert, is, in a sense, uninteresting. If experts are too busy, too lazy or too uninterested in the problems, then they are always free to equate their beliefs with the results of the computer analysis, however flawed, faulty or misconceived they perceive the analysis to be.
The question of whether the Bayesian analysis of a simulator does indeed represent the judgements of the expert, is, in a sense, uninteresting. If experts are too busy, too lazy or too uninterested in the problems, then they are always free to equate their beliefs with the results of the computer analysis, however flawed, faulty or misconceived they perceive the analysis to be.

When is the probability of an individual scientifically valuable?
The question of whether the Bayesian analysis of a simulator does indeed represent the judgements of the expert, is, in a sense, uninteresting. If experts are too busy, too lazy or too uninterested in the problems, then they are always free to equate their beliefs with the results of the computer analysis, however flawed, faulty or misconceived they perceive the analysis to be.

When is the probability of an individual scientifically valuable?

[1] This individual is sufficiently knowledgeable in the area for his/her judgements to carry weight and
Best current judgements

The question of whether the Bayesian analysis of a simulator does indeed represent the judgements of the expert, is, in a sense, uninteresting. If experts are too busy, too lazy or too uninterested in the problems, then they are always free to equate their beliefs with the results of the computer analysis, however flawed, faulty or misconceived they perceive the analysis to be.

When is the probability of an individual scientifically valuable?

[1] This individual is sufficiently knowledgeable in the area for his/her judgements to carry weight and

[2] the analysis that has led to this judgement has been sufficiently careful and thorough to support this judgement and sufficiently transparent that the reasoning, not simply the conclusions, can be understood and reassessed by similarly knowledgeable experts in the field.
The question of whether the Bayesian analysis of a simulator does indeed represent the judgements of the expert, is, in a sense, uninteresting. If experts are too busy, too lazy or too uninterested in the problems, then they are always free to equate their beliefs with the results of the computer analysis, however flawed, faulty or misconceived they perceive the analysis to be.

When is the probability of an individual scientifically valuable?

[1] This individual is sufficiently knowledgeable in the area for his/her judgements to carry weight and
[2] the analysis that has led to this judgement has been sufficiently careful and thorough to support this judgement and sufficiently transparent that the reasoning, not simply the conclusions, can be understood and reassessed by similarly knowledgeable experts in the field.

So, perhaps we should require that the objective of the analysis is to produce the “best” current judgements of a specified expert, in a sufficiently transparent form that the reasoning which led to these judgements should be open to critical scrutiny.
“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments.
“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis.
“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis. It is a fair and important question for the expert to have to consider and reveal as to just how “second best” the declared judgements are.
"Best expert judgements" are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis. It is a fair and important question for the expert to have to consider and reveal as to just how “second best” the declared judgements are. There are three levels of best current judgements.
“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis. It is a fair and important question for the expert to have to consider and reveal as to just how “second best” the declared judgements are. There are three levels of best current judgements.

**Stage 1** The best current judgements of an individual expert (expressed as probabilities). (This is a *subjective* Bayes analysis)
“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis. It is a fair and important question for the expert to have to consider and reveal as to just how “second best” the declared judgements are. There are three levels of best current judgements.

**Stage 1** The best current judgements of an individual expert (expressed as probabilities). (This is a *subjective* Bayes analysis)

**Stage 2** A careful analysis of the range of uncertainty judgements that it would be reasonable to hold given the differing views of experts. (A *scientific* Bayes analysis)
Best expert judgements

“Best expert judgements” are those which are sufficiently well founded that the expert is not aware of any further calculations that could “feasibly” done which would be judged to lead to substantially improved assessments. If a problem is important enough that the uncertainty analysis will have a large scientific, commercial or public policy implications, then best current judgements set a meaningful, rigorous standard for the analysis.

It is a fair and important question for the expert to have to consider and reveal as to just how “second best” the declared judgements are.

There are three levels of best current judgements.

Stage 1 The best current judgements of an individual expert (expressed as probabilities). (This is a subjective Bayes analysis)

Stage 2 A careful analysis of the range of uncertainty judgements that it would be reasonable to hold given the differing views of experts. (A scientific Bayes analysis)

Stage 3 An analysis so clear and compelling that it would command agreement from all knowledgeable experts. (An objective Bayes analysis and the only case where we can talk about, eg THE probability of rapid change. )
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians policy makers don’t know how to frame the right questions
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians policy makers don’t know how to frame the right questions there are no funding mechanisms to support this activity
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians policy makers don’t know how to frame the right questions there are no funding mechanisms to support this activity and it is hard (because different)!
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians policy makers don’t know how to frame the right questions there are no funding mechanisms to support this activity and it is hard (because different)!

For policy development, the basic question is: what does the collection of models, scientific theories, observations and analysis of the likely implications arising from our imperfect knowledge, tell us about actual climate behaviour? Such analysis requires our Best Current Judgements.
How are we doing so far?

Despite all the enormous amounts of very hard science that is being done, and the detailed knowledge that we are acquiring, I’m not sure that anyone is making a careful specification of best current judgements representing current knowledge yet. This is because scientists don’t think this way, nor do most statisticians policy makers don’t know how to frame the right questions there are no funding mechanisms to support this activity and it is hard (because different)!

For policy development, the basic question is: what does the collection of models, scientific theories, observations and analysis of the likely implications arising from our imperfect knowledge, tell us about actual climate behaviour? Such analysis requires our Best Current Judgements. ‘Best’ is a high standard to set for our judgements (though why aim for less?). What we require is care and clarity. These are challenging requirements, but no more challenging, in principle, than the process of collecting climate data and building and analysing climate models themselves. However, this does require a different tool-set and proper resources to carry through.
Linking models to reality

The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, $x$, and system behaviour $y$. 
Linking models to reality

The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, \( x \), and system behaviour \( y \).

More generally, evaluations of a collection of models are jointly informative for the physical system as they are jointly informative for these general relationships.
Linking models to reality

The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, $x$, and system behaviour $y$. More generally, evaluations of a collection of models are jointly informative for the physical system as they are jointly informative for these general relationships. Therefore, our inference from model to reality can proceed in two parts.
Linking models to reality

The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, $x$, and system behaviour $y$. More generally, evaluations of a collection of models are jointly informative for the physical system as they are jointly informative for these general relationships. Therefore, our inference from model to reality can proceed in two parts. [1] We emulate the relationship between system properties and system behaviour (we call this relationship the “reified model” (from reify: to treat an abstract concept as if it were real).
The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, $x$, and system behaviour $y$.

More generally, evaluations of a collection of models are jointly informative for the physical system as they are jointly informative for these general relationships. Therefore, our inference from model to reality can proceed in two parts.

[1] We emulate the relationship between system properties and system behaviour (we call this relationship the “reified model” (from reify: to treat an abstract concept as if it were real).

[2] We decompose the difference between our model and the physical system into two parts.

[A] The difference between our simulator and the reified form.
Linking models to reality

The reason that the evaluations of the simulator are informative for the physical system is that the evaluations are informative about the general relationships between system properties, $x$, and system behaviour $y$. More generally, evaluations of a collection of models are jointly informative for the physical system as they are jointly informative for these general relationships. Therefore, our inference from model to reality can proceed in two parts.

[1] We emulate the relationship between system properties and system behaviour (we call this relationship the “reified model” (from reify: to treat an abstract concept as if it were real).

[2] We decompose the difference between our model and the physical system into two parts.

[A] The difference between our simulator and the reified form.

[B] The difference between the reified form at the physically appropriate choice of $x$ and the actual system behaviour $y$. 
Relating models and the system

Reifying principle [1]
Simulator $F$ is informative for $y$, because $F$ is informative for $F^*$ and $F^*(x^*)$ is informative for $y$. 
Relating models and the system

**Reifying principle [1]**
Simulator $F$ is informative for $y$, because $F$ is informative for $F^*$ and $F^*(x^*)$ is informative for $y$. 

![Diagram]

- Model evaluations
- Model, $F$
- ‘Best’ input, $x^*$
- $F(x^*)$
- Discrepancy
- Actual system
- Measurement error
- System observations
Relating models and the system

Reifying principle [1]
Simulator $F$ is informative for $y$, because $F$ is informative for $F^*$ and $F^*(x^*)$ is informative for $y$. 

```
Model, $F$  ↓  Model evaluations  ↓  $F^*$  ↓  $F^*(x^*)$  ↓  Discrepancy  ↓  Actual system  ↓  System observations
```

Measurement error
Relating models and the system

Reifying principle [1]
Simulator $F$ is informative for $y$, because $F$ is informative for $F^*$ and $F^*(x^*)$ is informative for $y$.

Reifying principle [2]
A collection of simulators $F_1, F_2, \ldots$ is jointly informative for $y$, as the simulators are jointly informative for $F^*$. 

Model evaluations \[\rightarrow\] Model, $F$ \[\rightarrow\] ‘Best’ input, $x^*$ \[\rightarrow\] Discrepancy \[\rightarrow\] Measurement error

Model evaluations \[\rightarrow\] $F^*$ \[\rightarrow\] $F^*(x^*)$ \[\rightarrow\] Actual system \[\rightarrow\] System observations
Linking $F$ and $F^*$ using emulators

Suppose that our emulator for $F$ is

$$f(x) = Bg(x) \oplus u(x)$$
Suppose that our emulator for $F$ is

$$f(x) = Bg(x) \oplus u(x)$$

Our simplest emulator for $F^*$ might be

$$f^*(x, w) = B^*g(x) \oplus u^*(x) \oplus u^*(x, w)$$

where we might model our judgements as $B^* = CB + \Gamma$, correlate $u(x)$ and $u^*(x)$, while $u^*(x, w)$, with additional parameters, $w$, is uncorrelated with remainder.
Linking $F$ and $F^*$ using emulators

Suppose that our emulator for $F$ is

$$f(x) = Bg(x) \oplus u(x)$$

Our simplest emulator for $F^*$ might be

$$f^*(x, w) = B^* g(x) \oplus u^*(x) \oplus u^*(x, w)$$

where we might model our judgements as $B^* = CB + \Gamma$, correlate $u(x)$ and $u^*(x)$, while $u^*(x, w)$, with additional parameters, $w$, is uncorrelated with remainder.

Structured reification improves on this with systematic probabilistic modelling for all those aspects of model deficiency whose effects we are prepared to consider explicitly.
Linking $F$ and $F^*$ using emulators

Suppose that our emulator for $F$ is

$$f(x) = Bg(x) \oplus u(x)$$

Our simplest emulator for $F^*$ might be

$$f^*(x, w) = B^*g(x) \oplus u^*(x) \oplus u^*(x, w)$$

where we might model our judgements as $B^* = CB + \Gamma$, correlate $u(x)$ and $u^*(x)$, while $u^*(x, w)$, with additional parameters, $w$, is uncorrelated with remainder.

Structured reification improves on this with systematic probabilistic modelling for all those aspects of model deficiency whose effects we are prepared to consider explicitly.

Comment: All our calibration and forecasting methodology is unchanged - all that has changed is our description of the joint covariance structure.
A Reified influence diagram

\[ F^{1}_{h:[n]}(x), \ldots, F^{m}_{h:[n]}(x) \]

Evaluations of the simulator at each of \( m \) initial conditions for historical components of simulator
Global information $F_{h:suff}$ (from second order exchangeability modelling). passes to Reified global form and to reified emulator.
A Reified influence diagram

\[
\begin{align*}
\left[ F_{h:n}^1(x), \ldots, F_{h:n}^m(x) \right] &\rightarrow F_{h:suff} \rightarrow F_{h:suff}^* \rightarrow f_h^*(x) \rightarrow F_{h:*}^*(x^*) \rightarrow y_h \rightarrow z \\
\end{align*}
\]

Link with \( x^* \) to reified function, at true initial condition, linked to data \( z \)
Add observation of a related multi-model ensemble (MME) consisting of tuned runs from related models (more exchangeability modelling).
A Reified influence diagram

\[
\left[ F^1_{h:[n]}(x), \ldots, F^m_{h:[n]}(x) \right] \rightarrow F_{h:suff} \rightarrow F^*_h : \text{suff} \rightarrow f^*_h(x) \rightarrow F^*_h(x^*) \rightarrow y_h \rightarrow z
\]

Add a set of evaluations from a fast approximation
Add evaluations of fast simulator for outcomes to be predicted, with decision choices $d$.
A Reified influence diagram

\[
\begin{align*}
\left[ F_{h:[n]}^1(x), \ldots, F_{h:[n]}^m(x) \right] & \rightarrow F_{h:suff} \rightarrow F_{h:suff}^* \rightarrow f_h^*(x) \rightarrow F_{h}^*(x^*) \rightarrow y_h \rightarrow z \\
F_{h:n}'(x) & \rightarrow F_{h:suff}' \\
F_{p:n}'(x, d) & \rightarrow F_{p:suff}' \rightarrow F_{p:suff}'^* \\
\end{align*}
\]

Link to reified global terms for quantities to be predicted
A Reified influence diagram

\[
\left[ F_{h:[n]}(x), \ldots, F_{h:[n]}(x) \right] \rightarrow F_{h:suff} \rightarrow F_{h:suff} \rightarrow f_h(x) \rightarrow F_h(x^*) \rightarrow y_h \rightarrow z
\]

\[
F'_{h:[n]}(x) \rightarrow F'_{h:suff} \rightarrow F_{h:suff} \rightarrow f_h(x) \rightarrow F_h(x^*) \rightarrow y_h \rightarrow z
\]

\[
F'_{p:[n]}(x, d) \rightarrow F'_{p:suff} \rightarrow F_{p:suff} \rightarrow f_p(x, d)
\]

And to reified global emulator, based on inputs and decisions
A Reified influence diagram

And link, through true future values $y_p$, to the overall utility cost $C$ of making decision choice $d^*$ [Attach more models to diagram at $F^*(x^*)$]
Best current judgements for complex systems

To assess best current judgements about complex systems, it is enormously helpful to have an overall framework to unify all the uncertainties. Within this framework, all of the scientific, technical, computational, statistical and foundational issues can be addressed in principle. Such analysis poses serious challenges, but they are no harder than all of the other modelling, computational and observational challenges involved with studying climate.
To assess best current judgements about complex systems, it is enormously helpful to have an overall framework to unify all the uncertainties. Within this framework, all of the scientific, technical, computational, statistical and foundational issues can be addressed in principle. Such analysis poses serious challenges, but they are no harder than all of the other modelling, computational and observational challenges involved with studying climate. However, this is not a part-time activity for a few people. In my view, what is needed is a large scale planned programme of co-operation joining up many interested individuals and groups involved in relevant aspects of the uncertainty analysis into a proper Climate Uncertainty Project.
Best current judgements for complex systems

To assess best current judgements about complex systems, it is enormously helpful to have an overall framework to unify all the uncertainties. Within this framework, all of the scientific, technical, computational, statistical and foundational issues can be addressed in principle. Such analysis poses serious challenges, but they are no harder than all of the other modelling, computational and observational challenges involved with studying climate. However, this is not a part-time activity for a few people. In my view, what is needed is a large scale planned programme of co-operation joining up many interested individuals and groups involved in relevant aspects of the uncertainty analysis into a proper Climate Uncertainty Project. Because judgements about actual climate are urgent, such a project should be started as soon as possible. (It could begin informally.)
Best current judgements for complex systems

To assess best current judgements about complex systems, it is enormously helpful to have an overall framework to unify all the uncertainties. Within this framework, all of the scientific, technical, computational, statistical and foundational issues can be addressed in principle. Such analysis poses serious challenges, but they are no harder than all of the other modelling, computational and observational challenges involved with studying climate. However, this is not a part-time activity for a few people. In my view, what is needed is a large scale planned programme of co-operation joining up many interested individuals and groups involved in relevant aspects of the uncertainty analysis into a proper Climate Uncertainty Project. Because judgements about actual climate are urgent, such a project should be started as soon as possible. (It could begin informally.) That way we can have Best Current Judgements which are scientifically meaningful, reflect the informed range of consensus/disagreement among the expert community and can both reflect and influence future policy and research in climate science.
References


M. Goldstein and J.C. Rougier (2008). Reified Bayesian modelling and inference for physical systems (with discussion), JSPI, 139, , 1221-1239


I. Vernon, M. Goldstein, and R. Bower (2010) Galaxy Formation: a Bayesian Uncertainty Analysis (with discussion), Bayesian Analysis, 5, 619-670