Bayesian uncertainty analysis for complex physical models

Michael Goldstein
Durham University*

*Work with Ian Vernon, Allan Seheult, Jonathan Cumming, and many others!

Thermohaline circulation

Thermohaline circulation

- Global ocean circulation is driven by winds and the exchange of heat and water vapour at the sea surface.

Thermohaline circulation

- Global ocean circulation is driven by winds and the exchange of heat and water vapour at the sea surface.
- Wind driven surface currents head polewards from the equatorial Atlantic Ocean, cooling all the while. Having lost much of its heat, the surface water becomes so salty (through evaporation) that it is dense enough to sink.

Thermohaline circulation

- Global ocean circulation is driven by winds and the exchange of heat and water vapour at the sea surface.
- Wind driven surface currents head polewards from the equatorial Atlantic Ocean, cooling all the while. Having lost much of its heat, the surface water becomes so salty (through evaporation) that it is dense enough to sink.
- The return flow occurs at the bottom of the North Atlantic, also along the eastern flank of North America. This is the North Atlantic thermohaline circulation (thermo- for heat and -haline for salt, which determine the density of sea water).

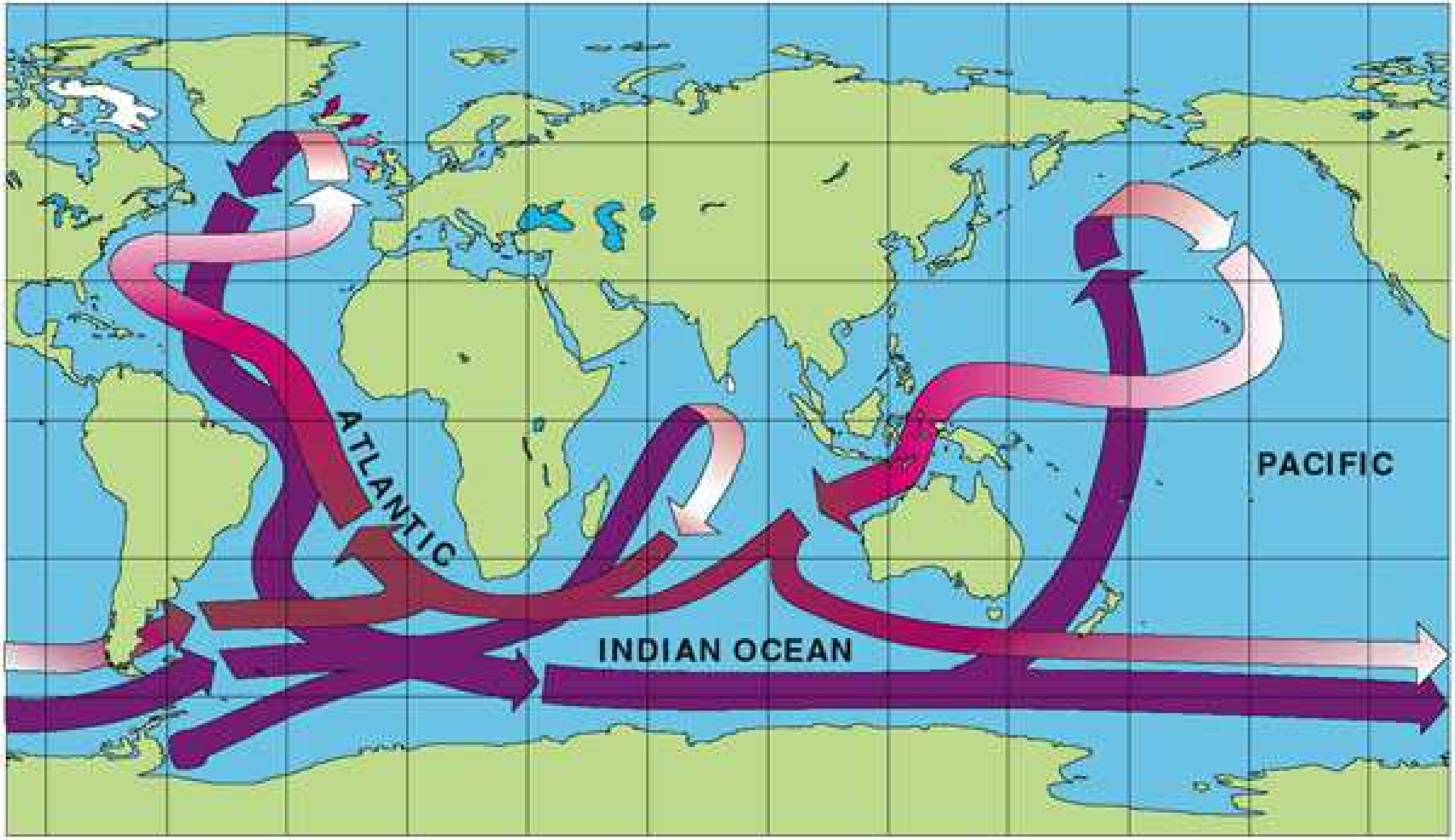
Thermohaline circulation

- Global ocean circulation is driven by winds and the exchange of heat and water vapour at the sea surface.
- Wind driven surface currents head polewards from the equatorial Atlantic Ocean, cooling all the while. Having lost much of its heat, the surface water becomes so salty (through evaporation) that it is dense enough to sink.
- The return flow occurs at the bottom of the North Atlantic, also along the eastern flank of North America. This is the North Atlantic thermohaline circulation (thermo- for heat and -haline for salt, which determine the density of sea water).
- On their journey, the water masses transport heat energy around the globe, which has a large impact on the climate of our planet.

References

Wikipedia (and see also:
Seager,R., 2005,'The Source of Europe's Mild Climate',American Scientist)

Global circulation



Thermohaline shutdown

Many atmosphere-ocean models show a slowdown of thermohaline circulation in simulations of the 21st century with the expected rise in greenhouse gases. [This is due to a combination of effects which reduce the density of surface waters, which makes it harder for them to sink.]
In some models, the THC shuts down rapidly and irreversibly once a critical threshold is passed.

Thermohaline shutdown

Many atmosphere-ocean models show a slowdown of thermohaline circulation in simulations of the 21st century with the expected rise in greenhouse gases. [This is due to a combination of effects which reduce the density of surface waters, which makes it harder for them to sink.]
In some models, the THC shuts down rapidly and irreversibly once a critical threshold is passed.

“ The weight of evidence makes it clear that climate change is a real and present danger. The Exeter conference was told that whatever policies are adopted from this point on, the Earth’s temperature will rise by 0.6F within the next 30 years. Yet those who think climate change just means Indian summers in Manchester should be told that the chances of the Gulf stream - the Atlantic thermohaline circulation that keeps Britain warm - shutting down are now thought to be greater than 50%.”
[**Burying carbon** Leader Column Thursday February 3, 2005 The Guardian]

Climate uncertainty

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Climate uncertainty

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?

Climate uncertainty

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
- What analysis was actually done to reach this conclusion?

Climate uncertainty

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
- What analysis was actually done to reach this conclusion?
- What analysis could possibly be done to justify (or contradict) this conclusion?

Climate uncertainty

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
- What analysis was actually done to reach this conclusion?
- What analysis could possibly be done to justify (or contradict) this conclusion?
- What do we learn about real physical systems from the analysis of (necessarily imperfect) models?

Models and Best Current Judgements

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Models and Best Current Judgements

The chances of the Gulf stream shutting down are now thought to be greater than 50%

[1] Exploring the behaviour of our favourite climate model gives us qualitative and semi-quantitative insights about climate behaviour.

Models and Best Current Judgements

The chances of the Gulf stream shutting down are now thought to be greater than 50%

[1] Exploring the behaviour of our favourite climate model gives us qualitative and semi-quantitative insights about climate behaviour.
However, behaviour of our model is not the same as behaviour of actual climate.
Analysing models helps us to make judgements, but model analyses are not the same as judgements about climate.

Models and Best Current Judgements

The chances of the Gulf stream shutting down are now thought to be greater than 50%

[1] Exploring the behaviour of our favourite climate model gives us qualitative and semi-quantitative insights about climate behaviour.
However, behaviour of our model is not the same as behaviour of actual climate.
Analysing models helps us to make judgements, but model analyses are not the same as judgements about climate.

[2] When we consider what actions we should take, we are concerned with actual climate. 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, [model deficiency, observation error, uncertainty about physical constants, etc.] tell us about actual climate behaviour?

Models and Best Current Judgements

The chances of the Gulf stream shutting down are now thought to be greater than 50%

[1] Exploring the behaviour of our favourite climate model gives us qualitative and semi-quantitative insights about climate behaviour.
However, behaviour of our model is not the same as behaviour of actual climate.
Analysing models helps us to make judgements, but model analyses are not the same as judgements about climate.

[2] When we consider what actions we should take, we are concerned with actual climate. 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, [model deficiency, observation error, uncertainty about physical constants, etc.] tell us about actual climate behaviour?
Such analysis results in our Best Current Judgements as to future climate behaviour, expressed as uncertainties.

Modelling climate

- Climate models are typically specified over the whole of the earth, because the earth is a conservative system with simple solar forcing. However, they can only be solved on large grids, 100s of kms across. Separate models 'downscale' these outputs to the regional level.

Modelling climate

- Climate models are typically specified over the whole of the earth, because the earth is a conservative system with simple solar forcing. However, they can only be solved on large grids, 100s of kms across. Separate models ‘downscale’ these outputs to the regional level.
- Climate models combine conservation equations (eg conservation of momentum) with state equations (eg to express the density of water as a function of temperature and salinity).

Modelling climate

- Climate models are typically specified over the whole of the earth, because the earth is a conservative system with simple solar forcing. However, they can only be solved on large grids, 100s of kms across. Separate models ‘downscale’ these outputs to the regional level.
- Climate models combine conservation equations (eg conservation of momentum) with state equations (eg to express the density of water as a function of temperature and salinity).
- A large climate model couples together modules for an ocean, an atmosphere, a cryosphere (land-ice and sea-ice), a carbon cycle (plankton), a sulphur cycle (emissions, including volcanoes), and land use. In simpler models some of these modules are left out or prescribed; in more complicated models they are dynamic, and interact with each other.

Modelling climate

- Climate models are typically specified over the whole of the earth, because the earth is a conservative system with simple solar forcing. However, they can only be solved on large grids, 100s of kms across. Separate models ‘downscale’ these outputs to the regional level.
- Climate models combine conservation equations (eg conservation of momentum) with state equations (eg to express the density of water as a function of temperature and salinity).
- A large climate model couples together modules for an ocean, an atmosphere, a cryosphere (land-ice and sea-ice), a carbon cycle (plankton), a sulphur cycle (emissions, including volcanoes), and land use. In simpler models some of these modules are left out or prescribed; in more complicated models they are dynamic, and interact with each other.
- The coupling of modules is extremely complicated, because the individual processes tend to be solved on different spatial grids, and tend to evolve at different rates.

Modelling climate

- Climate models are typically specified over the whole of the earth, because the earth is a conservative system with simple solar forcing. However, they can only be solved on large grids, 100s of kms across. Separate models ‘downscale’ these outputs to the regional level.
- Climate models combine conservation equations (eg conservation of momentum) with state equations (eg to express the density of water as a function of temperature and salinity).
- A large climate model couples together modules for an ocean, an atmosphere, a cryosphere (land-ice and sea-ice), a carbon cycle (plankton), a sulphur cycle (emissions, including volcanoes), and land use. In simpler models some of these modules are left out or prescribed; in more complicated models they are dynamic, and interact with each other.
- The coupling of modules is extremely complicated, because the individual processes tend to be solved on different spatial grids, and tend to evolve at different rates.
- **Thermo-haline shutdown?** Tentative prediction shown by simple models and paleo records, but not by many of the large models (so far).

The state of the art

Large climate models take months to run on supercomputers. One of the biggest computers in the world is the Earth Simulator in Japan, which is often used for running climate models.

Leading climate models

One leading climate model at the moment is HadCM3, based at the UK Met Office. One component of this model is HadAM3, the atmospheric module. In a simple experiment to study the effect of CO2-doubling (Murphy et al, 2004, Nature), this is coupled with simple mixed-layer ocean sea-ice models.

Leading climate models

One leading climate model at the moment is HadCM3, based at the UK Met Office. One component of this model is HadAM3, the atmospheric module. In a simple experiment to study the effect of CO2-doubling (Murphy et al, 2004, Nature), this is coupled with simple mixed-layer ocean sea-ice models.
The climate model has about 100 uncertain parameters, including:

1. *Large scale cloud.* Six parameters
2. *Convection.* Six parameters
3. *Sea ice.* Two parameters
4. *Radiation.* Four parameters
5. *Dynamics.* Four parameters
6. *Land surface.* Four parameters
7. *Boundary layer.* Four parameters

Leading climate models

One leading climate model at the moment is HadCM3, based at the UK Met Office. One component of this model is HadAM3, the atmospheric module. In a simple experiment to study the effect of CO2-doubling (Murphy et al, 2004, Nature), this is coupled with simple mixed-layer ocean sea-ice models.
The climate model has about 100 uncertain parameters, including:

1. *Large scale cloud.* Six parameters
2. *Convection.* Six parameters
3. *Sea ice.* Two parameters
4. *Radiation.* Four parameters
5. *Dynamics.* Four parameters
6. *Land surface.* Four parameters
7. *Boundary layer.* Four parameters

We have a few hundred evaluations of the mode, made over a period of about three years. These evaluations are one of the main resources for the UK Climate Projections Programme 2009, which is intended as a fairly definitive statement about how climate change will impact the UK.

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

encapsulates some, but not all, of the uncertainty associated with projecting future climate;

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

encapsulates some, but not all, of the uncertainty associated with projecting future climate;

is dependent on the method used, including assumptions and choices made, meaning that a different method would produce different results;

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

encapsulates some, but not all, of the uncertainty associated with projecting future climate;

is dependent on the method used, including assumptions and choices made, meaning that a different method would produce different results;

is based on the current evidence (i.e. models and observations), and new evidence in the future may lead to the results being modified;

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

encapsulates some, but not all, of the uncertainty associated with projecting future climate;

is dependent on the method used, including assumptions and choices made, meaning that a different method would produce different results;

is based on the current evidence (i.e. models and observations), and new evidence in the future may lead to the results being modified;

does not reduce uncertainty, just makes it more transparent;

UK Climate Projections Programme (from the web-site)

What are probabilistic climate projections?

Before using the UKCP09 probabilistic climate projections, it is important to understand what they are and what they are not.

A probabilistic climate projection:

IS NOT an objective probability, where a situation is well understood, where all outcomes can be accounted for or where probabilities can be revised based on observed outcomes (such as tossing a coin or rolling a dice);

IS rather a subjective probability, providing an estimate based on the available information and strength of evidence (similar to horse-racing odds);

encapsulates some, but not all, of the uncertainty associated with projecting future climate;

is dependent on the method used, including assumptions and choices made, meaning that a different method would produce different results;

is based on the current evidence (i.e. models and observations), and new evidence in the future may lead to the results being modified;

does not reduce uncertainty, just makes it more transparent;

aims to allow users to make more robust decisions.

Relating a computer simulator to a physical system

- Basic ingredients:

Relating a computer simulator to a physical system

- Basic ingredients:

  x^*: system properties (unknown)

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)

Relating a computer simulator to a physical system

- Basic ingredients:

  x^*: system properties (unknown)

  y: system behaviour (influenced by x^*)

  z: partial observation of y (with error)

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)
  z: partial observation of y (with error)
- Ideally, we would like to construct a deterministic computer model F, embodying the **laws of nature**, which satisfies

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

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)
  z: partial observation of y (with error)
- Ideally, we would like to construct a deterministic computer model F, embodying the **laws of nature**, which satisfies

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

- In practice, however, the our actual model F is inadequate:
  - F simplifies the physics;
  - F approximates the solution of the physical equations

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)
  z: partial observation of y (with error)
- Ideally, we would like to construct a deterministic computer model F, embodying the **laws of nature**, which satisfies

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

- In practice, however, the our actual model F is inadequate:
  - F simplifies the physics;
  - F approximates the solution of the physical equations
- This raises the basic question as to what does the imperfect F tell us about the system values (x^*, y)?

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)
  z: partial observation of y (with error)
- Ideally, we would like to construct a deterministic computer model F, embodying the **laws of nature**, which satisfies

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

- In practice, however, the our actual model F is inadequate:
  - F simplifies the physics;
  - F approximates the solution of the physical equations
- This raises the basic question as to what does the imperfect F tell us about the system values (x^*, y)?
- How about several different (imperfect) models for (x^*, y)?

Relating a computer simulator to a physical system

- Basic ingredients:
  x^*: system properties (unknown)
  y: system behaviour (influenced by x^*)
  z: partial observation of y (with error)
- Ideally, we would like to construct a deterministic computer model F, embodying the **laws of nature**, which satisfies

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

- In practice, however, the our actual model F is inadequate:
  - F simplifies the physics;
  - F approximates the solution of the physical equations
- This raises the basic question as to what does the imperfect F tell us about the system values (x^*, y)?
- How about several different (imperfect) models for (x^*, y)?
- In particular, input and output very high dimensional and evaluating $F(x)$ for any x may be VERY expensive.

Relating the model and the system

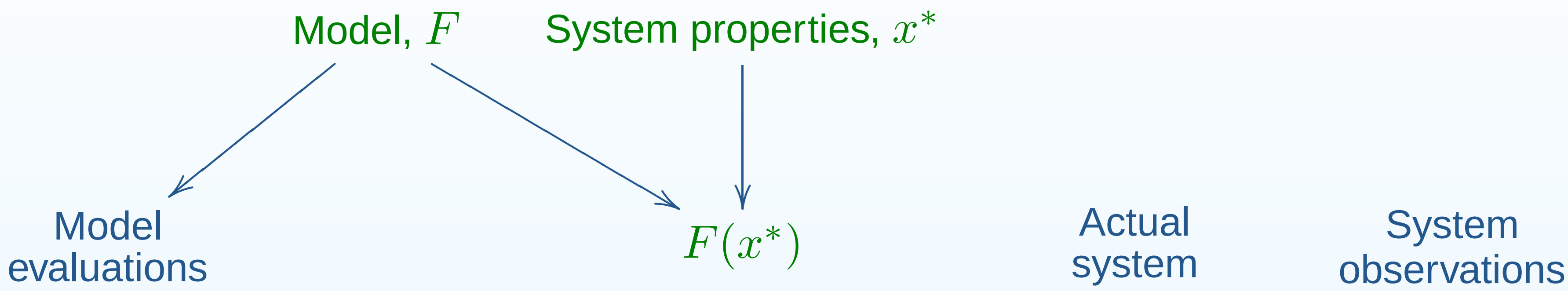
Model evaluations

Actual system

System observations

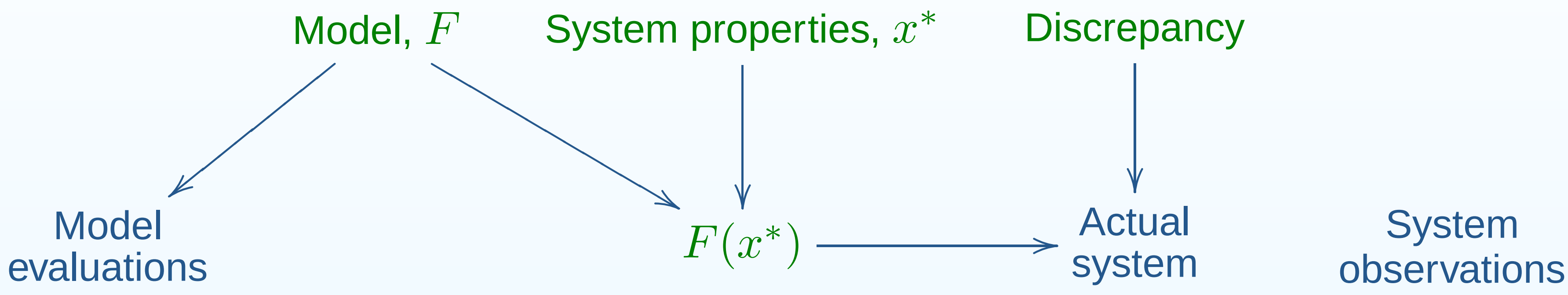
1. We start with a collection of model evaluations, and some observations on the actual system
2. We link the evaluations to the system properties
3. We link the system evaluation to the actual system behaviour
4. We incorporate measurement error into the observations
5. We add initial and boundary condition and forcing function uncertainty.

Relating the model and the system



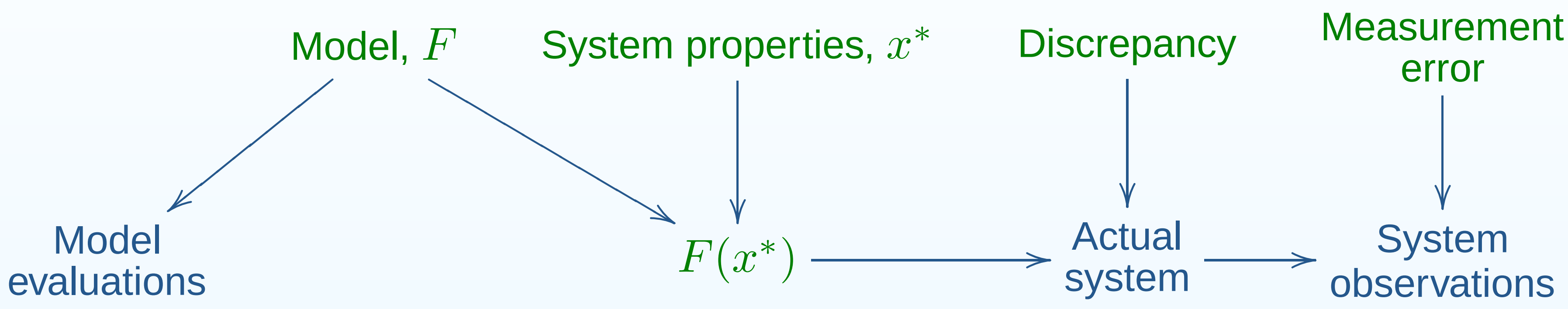
1. We start with a collection of model evaluations, and some observations on the actual system
2. We link the evaluations to the system properties
3. We link the system evaluation to the actual system behaviour
4. We incorporate measurement error into the observations
5. We add initial and boundary condition and forcing function uncertainty.

Relating the model and the system



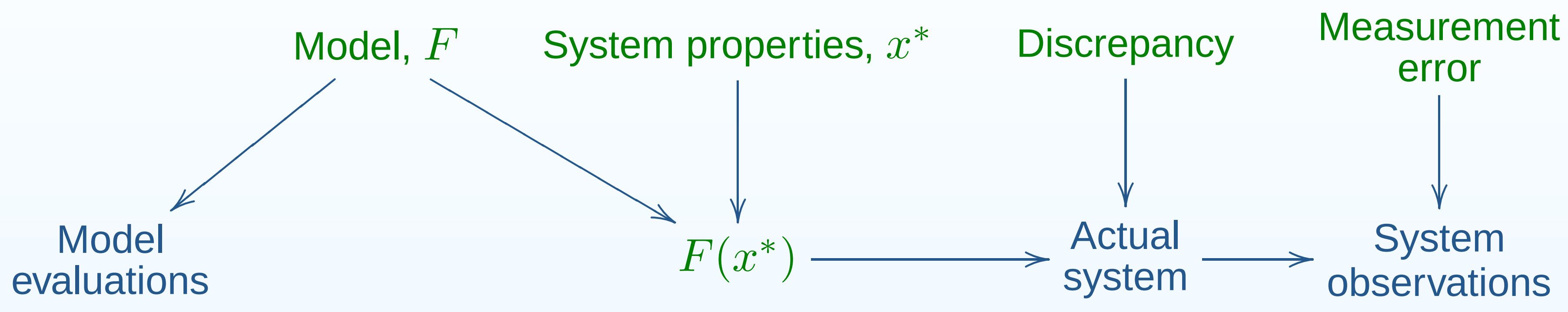
1. We start with a collection of model evaluations, and some observations on the actual system
2. We link the evaluations to the system properties
3. We link the system evaluation to the actual system behaviour
4. We incorporate measurement error into the observations
5. We add initial and boundary condition and forcing function uncertainty.

Relating the model and the system



1. We start with a collection of model evaluations, and some observations on the actual system
2. We link the evaluations to the system properties
3. We link the system evaluation to the actual system behaviour
4. We incorporate measurement error into the observations
5. We add initial and boundary condition and forcing function uncertainty.

Relating the model and the system



1. We start with a collection of model evaluations, and some observations on the actual system
2. We link the evaluations to the system properties
3. We link the system evaluation to the actual system behaviour
4. We incorporate measurement error into the observations
5. We add initial and boundary condition and forcing function uncertainty.

Representing beliefs about f using emulators

An *emulator* is a probabilistic belief specification for a deterministic function. Our emulator for component i of F might be

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

Representing beliefs about f using emulators

An *emulator* is a probabilistic belief specification for a deterministic function.
Our emulator for component i of F might be

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

where $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are known functions of x,
[$\sum_j \beta_{ij}\, g_{ij}(x)$ expresses global variation in F]

Representing beliefs about f using emulators

An *emulator* is a probabilistic belief specification for a deterministic function.
Our emulator for component i of F might be

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

where $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are known functions of x,
[$\sum_j \beta_{ij}\, g_{ij}(x)$ expresses global variation in F]
$u(x)$ is a weakly stationary stochastic process (maybe Gaussian)
so $\mathrm{Var}(u_i(x))$ is the same for all x and $\mathrm{Cov}(u_i(x), u_i(x'))$ depends only on the distance between x and x'.
[$u(x)$ expresses local variation in F]

Representing beliefs about f using emulators

An *emulator* is a probabilistic belief specification for a deterministic function.
Our emulator for component i of F might be

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

where $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are known functions of x,
[$\sum_j \beta_{ij}\, g_{ij}(x)$ expresses global variation in F]
$u(x)$ is a weakly stationary stochastic process (maybe Gaussian)
so $\mathrm{Var}(u_i(x))$ is the same for all x and $\mathrm{Cov}(u_i(x), u_i(x'))$ depends only on the distance between x and x'.
[$u(x)$ expresses local variation in F]
We fit the emulator, given a collection of model evaluations, using our favourite statistical tools - generalised least squares, maximum likelihood, Bayes - with a generous helping of expert judgement.

Representing beliefs about f using emulators

An *emulator* is a probabilistic belief specification for a deterministic function.

Our emulator for component i of F might be

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

where $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are known functions of x,
[$\sum_j \beta_{ij}\, g_{ij}(x)$ expresses global variation in F]
$u(x)$ is a weakly stationary stochastic process (maybe Gaussian)
so $\mathrm{Var}(u_i(x))$ is the same for all x and $\mathrm{Cov}(u_i(x), u_i(x'))$ depends only on the distance between x and x'.
[$u(x)$ expresses local variation in F]
We fit the emulator, given a collection of model evaluations, using our favourite statistical tools - generalised least squares, maximum likelihood, Bayes - with a generous helping of expert judgement.
So, we need careful experimental design to choose which evaluations of the model to make, and detailed diagnostics, to check emulator validity.

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic “emulator” f for the computer function F

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic "emulator" f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic “emulator” f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic “emulator” f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

This full probabilistic description provides a formal framework to synthesise expert elicitation, historical data and a careful choice of simulator runs.

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic “emulator” f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

This full probabilistic description provides a formal framework to synthesise expert elicitation, historical data and a careful choice of simulator runs.
We may then use our collection of computer evaluations and historical observations to make a Bayesian analyssis of the physical process

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic "emulator" f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

This full probabilistic description provides a formal framework to synthesise expert elicitation, historical data and a careful choice of simulator runs.
We may then use our collection of computer evaluations and historical observations to make a Bayesian analyssis of the physical process

- to determine "correct" settings for simulator inputs (calibration);

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic “emulator” f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

This full probabilistic description provides a formal framework to synthesise expert elicitation, historical data and a careful choice of simulator runs.
We may then use our collection of computer evaluations and historical observations to make a Bayesian analyssis of the physical process

- to determine “correct” settings for simulator inputs (calibration);
- to assess the future behaviour of the system (forecasting).

Uncertainty analysis for complex models

Aim: to tackle problems arising from all of the uncertainties inherent in imperfect computer models of highly complex physical systems, using a Bayesian formulation. This involves

- a prior probability distribution for system properties x^*
- a probabilistic "emulator" f for the computer function F
- a probabilistic discrepancy measure relating $F(x^*)$ to system behaviour y
- a likelihood function relating historical data z to y

This full probabilistic description provides a formal framework to synthesise expert elicitation, historical data and a careful choice of simulator runs.
We may then use our collection of computer evaluations and historical observations to make a Bayesian analyssis of the physical process

- to determine "correct" settings for simulator inputs (calibration);
- to assess the future behaviour of the system (forecasting).
- to "optimise" the performance of the system

Bayes linear approach

For very large scale problems a full Bayes analysis is very hard because
(i) it is difficult to give a meaningful full prior probability specification over high dimensional spaces;
(ii) the computations, for learning from data (observations and computer runs) and choosing informative runs, may be technically difficult;
(iii) the likelihood surface is extremely complicated, and any full Bayes calculation may be extremely non-robust.

Bayes linear approach

For very large scale problems a full Bayes analysis is very hard because
(i) it is difficult to give a meaningful full prior probability specification over high dimensional spaces;
(ii) the computations, for learning from data (observations and computer runs) and choosing informative runs, may be technically difficult;
(iii) the likelihood surface is extremely complicated, and any full Bayes calculation may be extremely non-robust.
However, the idea of the Bayesian approach, namely capturing our expert prior judgements in stochastic form and modifying them by appropriate rules given observations, is conceptually appropriate (and there is no obvious alternative).

Bayes linear approach

For very large scale problems a full Bayes analysis is very hard because
(i) it is difficult to give a meaningful full prior probability specification over high dimensional spaces;
(ii) the computations, for learning from data (observations and computer runs) and choosing informative runs, may be technically difficult;
(iii) the likelihood surface is extremely complicated, and any full Bayes calculation may be extremely non-robust.
However, the idea of the Bayesian approach, namely capturing our expert prior judgements in stochastic form and modifying them by appropriate rules given observations, is conceptually appropriate (and there is no obvious alternative).
The Bayes Linear approach is (relatively) simple in terms of belief specification and analysis, as it is based only on the mean, variance and covariance specification which, following de Finetti, (see de Finetti “Theory of Probability”, Wiley, 1974), we take as primitive.
For a full account, see
Michael Goldstein and David Wooff (2007) Bayes Linear Statistics: Theory and Methods, Wiley.

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

[2] an approximation to a full Bayes analysis;

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y, z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y, z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z, y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

[2] an approximation to a full Bayes analysis;

[3] a generalisation of conditioning;

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$
\begin{aligned}
\mathrm{E}_z(y) &= \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),\\
\mathrm{Var}_z(y) &= \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)
\end{aligned}
$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

[2] an approximation to a full Bayes analysis;

[3] a generalisation of conditioning;

[4] the “appropriate” analysis given a partial specification based on expectation (with methodology for modelling, interpretation and diagnostic analysis);

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

[2] an approximation to a full Bayes analysis;

[3] a generalisation of conditioning;

[4] the "appropriate" analysis given a partial specification based on expectation (with methodology for modelling, interpretation and diagnostic analysis);

[5] mathematically - a treatment where we consider subspaces, rather than random quantities, as fundamental;

Bayes linear adjustment

Bayes Linear adjustment of the mean and the variance of y given z is

$$\mathrm{E}_z(y) = \mathrm{E}(y) + \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}(z - \mathrm{E}(z)),$$
$$\mathrm{Var}_z(y) = \mathrm{Var}(y) - \mathrm{Cov}(y,z)\mathrm{Var}(z)^{-1}\mathrm{Cov}(z,y)$$

$\mathsf{E}_z[y]$, $\mathsf{Var}_z[y]$ are the expectation and variance for y adjusted by z.

Bayes linear adjustment may be viewed as:

[1] an estimation procedure;

[2] an approximation to a full Bayes analysis;

[3] a generalisation of conditioning;

[4] the “appropriate” analysis given a partial specification based on expectation (with methodology for modelling, interpretation and diagnostic analysis);

[5] mathematically - a treatment where we consider subspaces, rather than random quantities, as fundamental;

[6] a formal subjective framework accounting for all uncertainties (recognising Bayesian analysis itself as a model).

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h
Calibration: we learn about x^* using simulator evaluations (from which we build the emulator) and z.

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h

Calibration: we learn about x^* using simulator evaluations (from which we build the emulator) and z.

History matching: we rule out regions of x space which are unlikely to give rise to observed history z.

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h

Calibration: we learn about x^* using simulator evaluations (from which we build the emulator) and z.

History matching: we rule out regions of x space which are unlikely to give rise to observed history z.

Using the emulator we can obtain, for each set of inputs x, the mean and variance, $\mathrm{E}(F_h(x))$ and $\mathrm{Var}(F_h(x))$.

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h

Calibration: we learn about x^* using simulator evaluations (from which we build the emulator) and z.

History matching: we rule out regions of x space which are unlikely to give rise to observed history z.

Using the emulator we can obtain, for each set of inputs x, the mean and variance, $\mathrm{E}(F_h(x))$ and $\mathrm{Var}(F_h(x))$.

If $x = x^*$, then

$\mathrm{Var}(z_i - \mathrm{E}(F_i(x))) = \mathrm{Var}(F_i(x)) + \mathrm{Var}(\epsilon_i) + \mathrm{Var}(e_i)$.

Model Calibration: History matching via Implausibility

$z_h = y_h + e_h$, $y_h = F_h(x^*) + \epsilon_h$, for outputs F_h with observed history z_h

Calibration: we learn about x^* using simulator evaluations (from which we build the emulator) and z.

History matching: we rule out regions of x space which are unlikely to give rise to observed history z.

Using the emulator we can obtain, for each set of inputs x, the mean and variance, $\mathrm{E}(F_h(x))$ and $\mathrm{Var}(F_h(x))$.

If $x = x^*$, then
$\mathrm{Var}(z_i - \mathrm{E}(F_i(x))) = \mathrm{Var}(F_i(x)) + \mathrm{Var}(\epsilon_i) + \mathrm{Var}(e_i)$.

We can therefore calculate, for each output $F_i(x)$, the “implausibility” if we consider the value x to be the best choice x^*, which is

$$I_{(i)}(x) = |z_i - \mathrm{E}(F_i(x))|^2 / [\mathrm{Var}(F_i(x)) + \mathrm{Var}(\epsilon_i) + \mathrm{Var}(e_i)]$$

[Large values of $I_{(i)}(x)$ suggest that it is implausible that $x = x^*$.]

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.
With this information, we can then refocus our analysis on the 'non-implausible' regions of the input space, by

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.

With this information, we can then refocus our analysis on the 'non-implausible' regions of the input space, by

(i) making more simulator runs

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.

With this information, we can then refocus our analysis on the 'non-implausible' regions of the input space, by

(i) making more simulator runs

(ii) refitting our emulator

over such sub-regions and repeating the analysis.

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.

With this information, we can then refocus our analysis on the 'non-implausible' regions of the input space, by

(i) making more simulator runs

(ii) refitting our emulator

over such sub-regions and repeating the analysis.

This process is a form of iterative global search aimed at finding all choices of x^* which would give good fits to historical data.

[Note the strong relationship between this statistical approach and more traditional optimisation methods for solving high dimensional ill-posed inverse problems.]

Using Implausibility measures

The implausibility calculation can be performed univariately, or by multivariate calculation over sub-vectors. The implausibilities are then combined, such as by using $I_M(x) = \max_i I_{(i)}(x)$, and can then be used to identify regions of x with large $I_M(x)$ as implausible, i.e. unlikely to be good choices for x^*.

With this information, we can then refocus our analysis on the 'non-implausible' regions of the input space, by

(i) making more simulator runs

(ii) refitting our emulator

over such sub-regions and repeating the analysis.

This process is a form of iterative global search aimed at finding all choices of x^* which would give good fits to historical data.

[Note the strong relationship between this statistical approach and more traditional optimisation methods for solving high dimensional ill-posed inverse problems.]

We may find no good choices at all which give good fits and that is a clear sign of problems with our physical simulator or with our data.

Forecasting

AIM: to predict those values y_p for which we do not yet have observations.

Forecasting

AIM: to predict those values y_p for which we do not yet have observations.
From the computer model emulator, we have the mean and variance of $F(x)$ for each input choice x.
Therefore, we compute the mean and variance of $F^* = F(x^*)$ by conditioning on and then integrating with respect to a prior distribution on x^*.

Forecasting

AIM: to predict those values y_p for which we do not yet have observations.
From the computer model emulator, we have the mean and variance of $F(x)$ for each input choice x.
Therefore, we compute the mean and variance of $F^* = F(x^*)$ by conditioning on and then integrating with respect to a prior distribution on x^*.
Given $\mathrm{E}(F^*)$, $\mathrm{Var}(F^*)$, and the discrepancy and observation error variances $\mathrm{Var}(\epsilon)$, $\mathrm{Var}(e)$, compute joint mean and variance of collection y_p, z,
[from $(y_p = F_p^* + \epsilon_p, \;\; z = y_h + e)$].

Forecasting

AIM: to predict those values y_p for which we do not yet have observations.

From the computer model emulator, we have the mean and variance of $F(x)$ for each input choice x.

Therefore, we compute the mean and variance of $F^* = F(x^*)$ by conditioning on and then integrating with respect to a prior distribution on x^*.

Given $\mathrm{E}(F^*)$, $\mathrm{Var}(F^*)$, and the discrepancy and observation error variances $\mathrm{Var}(\epsilon)$, $\mathrm{Var}(e)$, compute joint mean and variance of collection y_p, z,

[from $(y_p = F_p^* + \epsilon_p, \;\; z = y_h + e)$].

We now evaluate the adjusted mean and variance for future values y_p adjusted by z using the Bayes linear adjustment formulae.

This analysis gives system forecasts without model calibration, and therefore is tractable even for large systems.

EXAMPLE:Understanding the Universe

EXAMPLE:Understanding the Universe

Major advances in cosmology in the last 100 years (mainly thanks to Einstein)

- Universe began in hot dense state: The Big Bang
- Since then Universe has been expanding rapidly

EXAMPLE:Understanding the Universe

Major advances in cosmology in the last 100 years (mainly thanks to Einstein)

- Universe began in hot dense state: The Big Bang
- Since then Universe has been expanding rapidly

Cosmologists have spent much time and money researching the beginning, the evolution, the current content, and the ultimate fate of the Universe.

EXAMPLE:Understanding the Universe

Major advances in cosmology in the last 100 years (mainly thanks to Einstein)

- Universe began in hot dense state: The Big Bang
- Since then Universe has been expanding rapidly

Cosmologists have spent much time and money researching the beginning, the evolution, the current content, and the ultimate fate of the Universe.

Now know that the observable Universe is composed of billions of galaxies each made up of 10 million - 10 trillion stars

EXAMPLE:Understanding the Universe

Major advances in cosmology in the last 100 years (mainly thanks to Einstein)

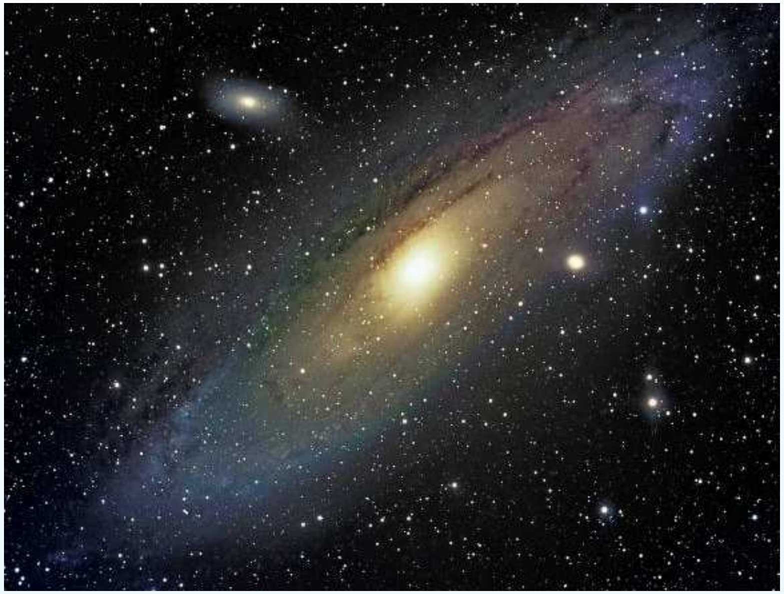
- Universe began in hot dense state: The Big Bang
- Since then Universe has been expanding rapidly

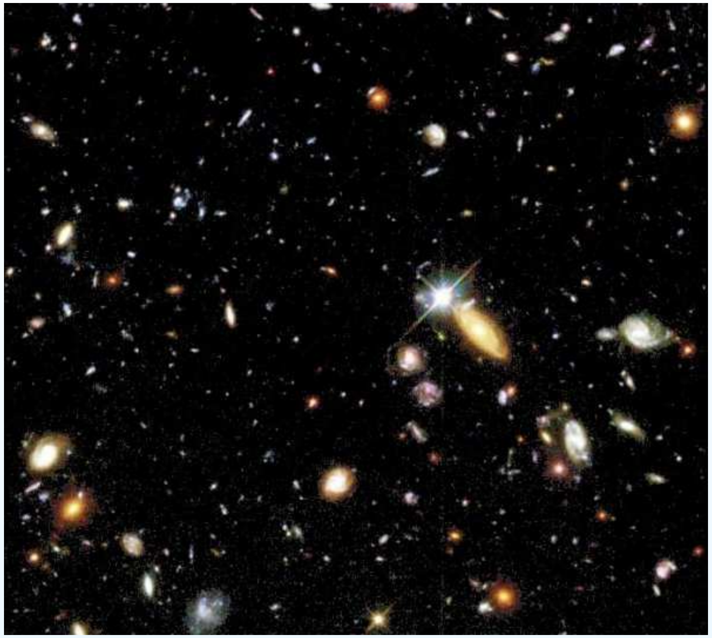
Cosmologists have spent much time and money researching the beginning, the evolution, the current content, and the ultimate fate of the Universe.

Now know that the observable Universe is composed of billions of galaxies each made up of 10 million - 10 trillion stars

How did these galaxies form?

Andromeda Galaxy and Hubble Deep Field View





- Andromeda Galaxy: closest large galaxy to our own milky way, contains 1 trillion stars.
- Hubble Deep Field: furthest image yet taken. Covers 2 millionths of the sky but contains over 3000 galaxies.

Dark Matter and the Evolution of the Universe

Recent observations of galaxies have suggested that only 3 percent of the entire energy content of the universe is the normal matter which forms stars, planets and us.

Dark Matter and the Evolution of the Universe

Recent observations of galaxies have suggested that only 3 percent of the entire energy content of the universe is the normal matter which forms stars, planets and us.

A further 23 percent is ’Dark Matter’ (and the rest is Dark Energy).

Dark Matter and the Evolution of the Universe

Recent observations of galaxies have suggested that only 3 percent of the entire energy content of the universe is the normal matter which forms stars, planets and us.

A further 23 percent is ’Dark Matter’ (and the rest is Dark Energy).

Dark Matter cannot be ’seen’ as it does not give off light (or anything else). However it does have mass and therefore affects stars and galaxies via gravity.

Dark Matter and the Evolution of the Universe

Recent observations of galaxies have suggested that only 3 percent of the entire energy content of the universe is the normal matter which forms stars, planets and us.

A further 23 percent is 'Dark Matter' (and the rest is Dark Energy).

Dark Matter cannot be 'seen' as it does not give off light (or anything else). However it does have mass and therefore affects stars and galaxies via gravity.

In order to study the effects of Dark Matter cosmologists try to model **Galaxy formation**

- Inherently linked to amount of Dark Matter
- Of fundamental interest as tests cosmologists' knowledge of a wide range of complicated physical phenomena

Simulating Galaxy Evolution: Two Stage approach

The simulation is performed in two parts:

Simulating Galaxy Evolution: Two Stage approach

The simulation is performed in two parts:

[1] First an N-Body simulation is run to determine the behaviour of fluctuations of mass in the early Universe, and their subsequent growth into millions of galaxy sized lumps of mass in the following 12 billion years.
[A very heavy simulation which takes 3 months, done on a supercomputer and cannot be easily repeated.]

Simulating Galaxy Evolution: Two Stage approach

The simulation is performed in two parts:

[1] First an N-Body simulation is run to determine the behaviour of fluctuations of mass in the early Universe, and their subsequent growth into millions of galaxy sized lumps of mass in the following 12 billion years.
[A very heavy simulation which takes 3 months, done on a supercomputer and cannot be easily repeated.]

[2] These results on the behaviour of the massive lumps are then used by a more detailed Galaxy Formation simulation (called GALFORM) which models the far more complicated interactions of normal matter: gas cloud formation, star formation and the effects of black holes at the centre of galaxies.

Simulating Galaxy Evolution: Two Stage approach

The simulation is performed in two parts:

[1] First an N-Body simulation is run to determine the behaviour of fluctuations of mass in the early Universe, and their subsequent growth into millions of galaxy sized lumps of mass in the following 12 billion years.
[A very heavy simulation which takes 3 months, done on a supercomputer and cannot be easily repeated.]

[2] These results on the behaviour of the massive lumps are then used by a more detailed Galaxy Formation simulation (called GALFORM) which models the far more complicated interactions of normal matter: gas cloud formation, star formation and the effects of black holes at the centre of galaxies.

The first simulation is done on a volume of size (500 Mega-Parsec)^3 or (1.63 billion light-years)^3

Simulating Galaxy Evolution: Two Stage approach

The simulation is performed in two parts:

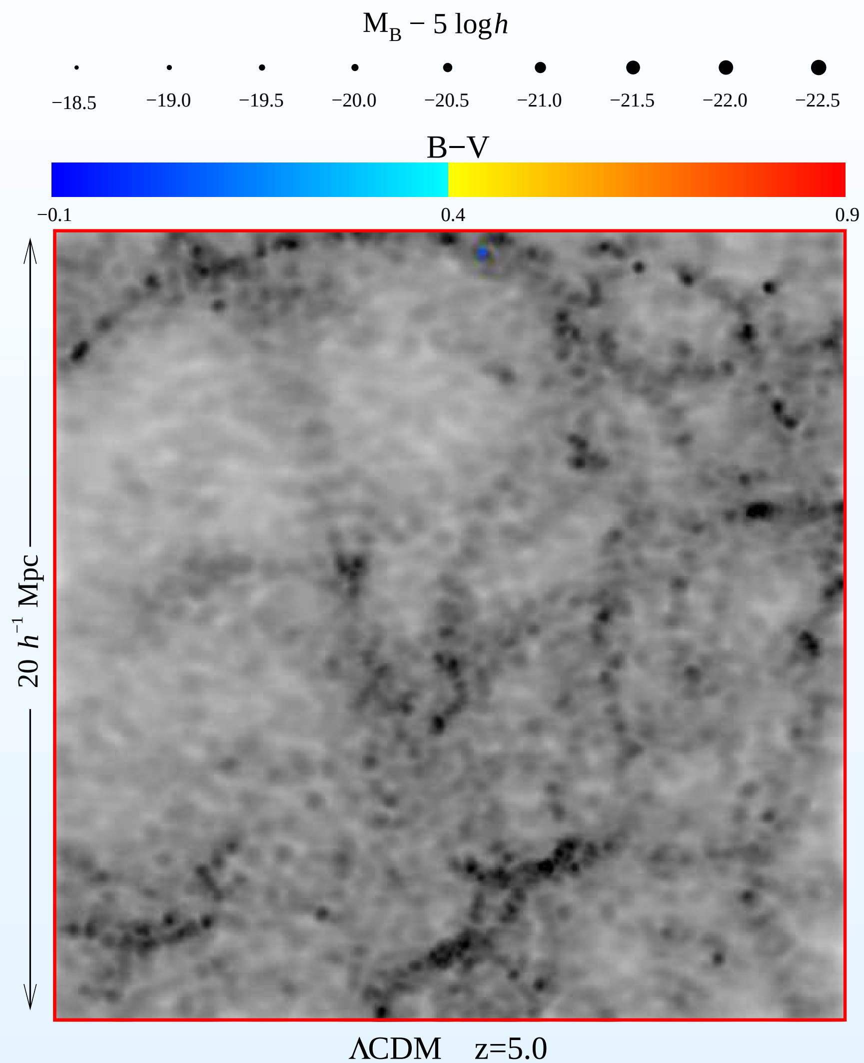
[1] First an N-Body simulation is run to determine the behaviour of fluctuations of mass in the early Universe, and their subsequent growth into millions of galaxy sized lumps of mass in the following 12 billion years.
[A very heavy simulation which takes 3 months, done on a supercomputer and cannot be easily repeated.]

[2] These results on the behaviour of the massive lumps are then used by a more detailed Galaxy Formation simulation (called GALFORM) which models the far more complicated interactions of normal matter: gas cloud formation, star formation and the effects of black holes at the centre of galaxies.

The first simulation is done on a volume of size (500 Mega-Parsec)^3 or (1.63 billion light-years)^3

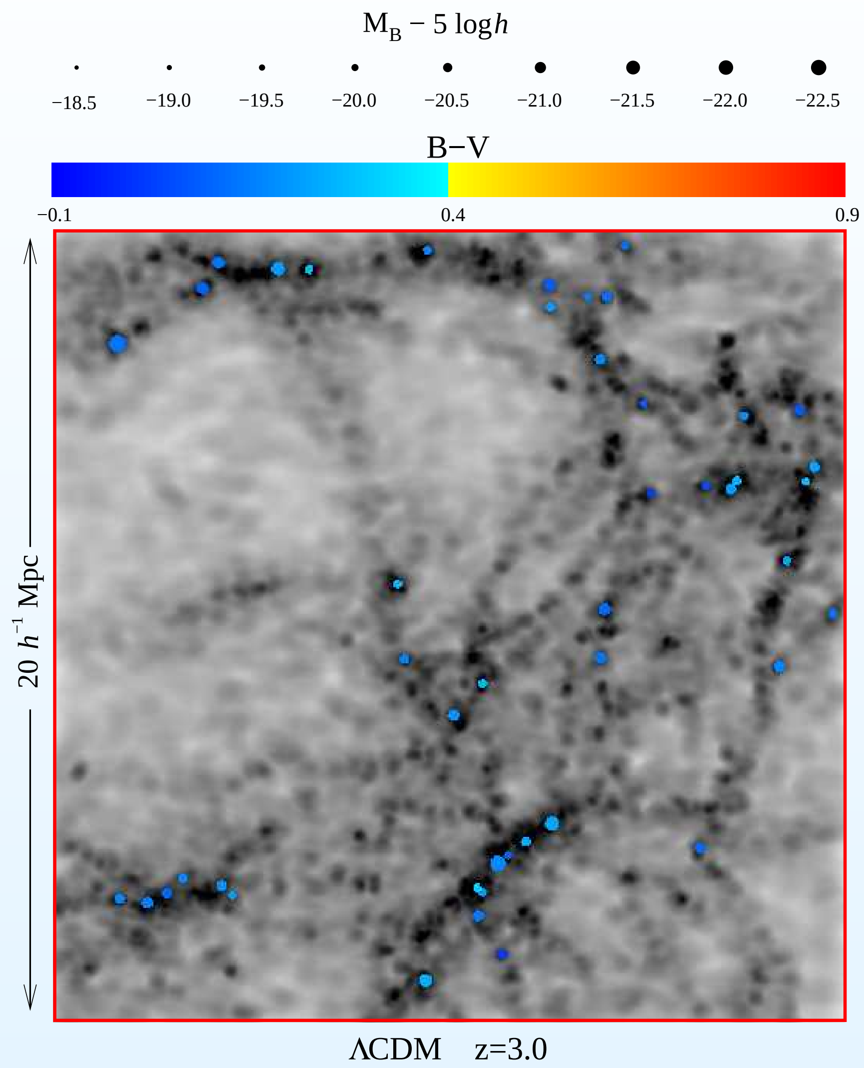
This volume is split into 512 sub-volumes which are independently simulated using the second model GALFORM. This simulation is run on upto 256 parallel processors, and takes 20-30 minutes per sub-volume per processor

Universe at $<$ 100 million years



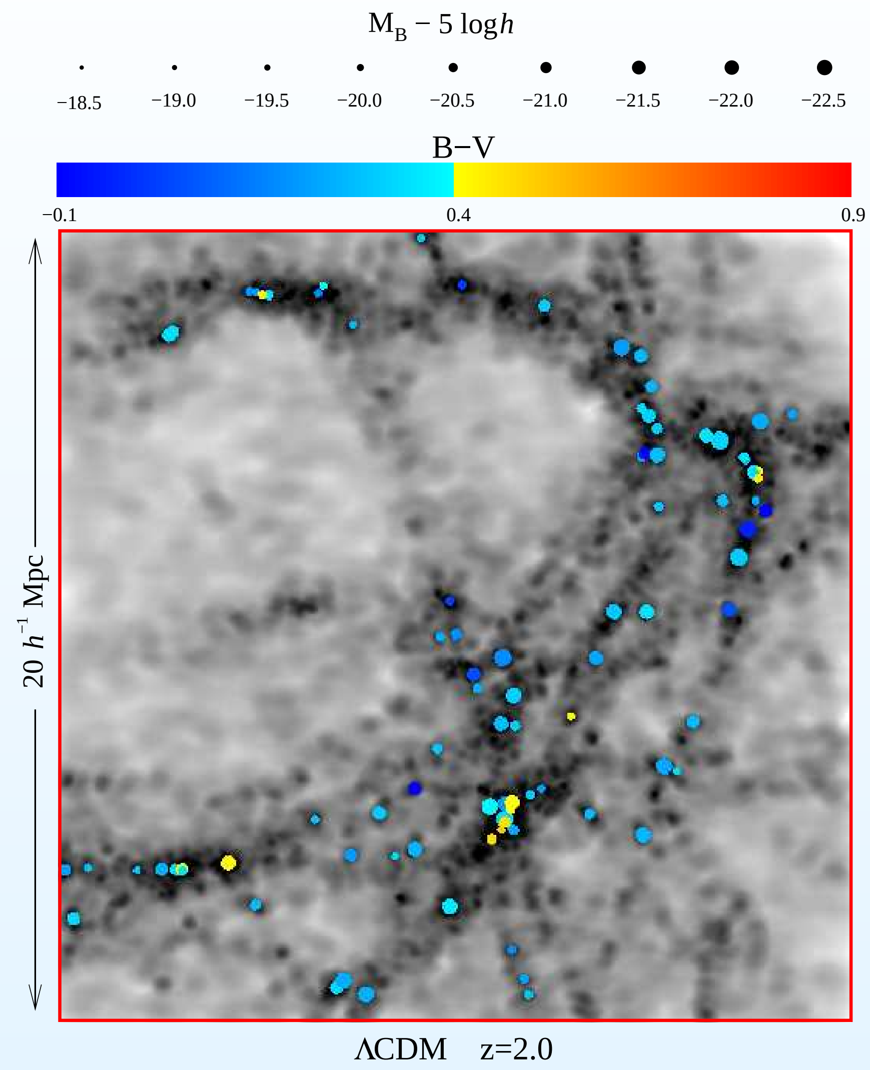
Benson, Frenk, Baugh, Cole & Lacey (2001)

Universe at $\sim$ 1 billion years



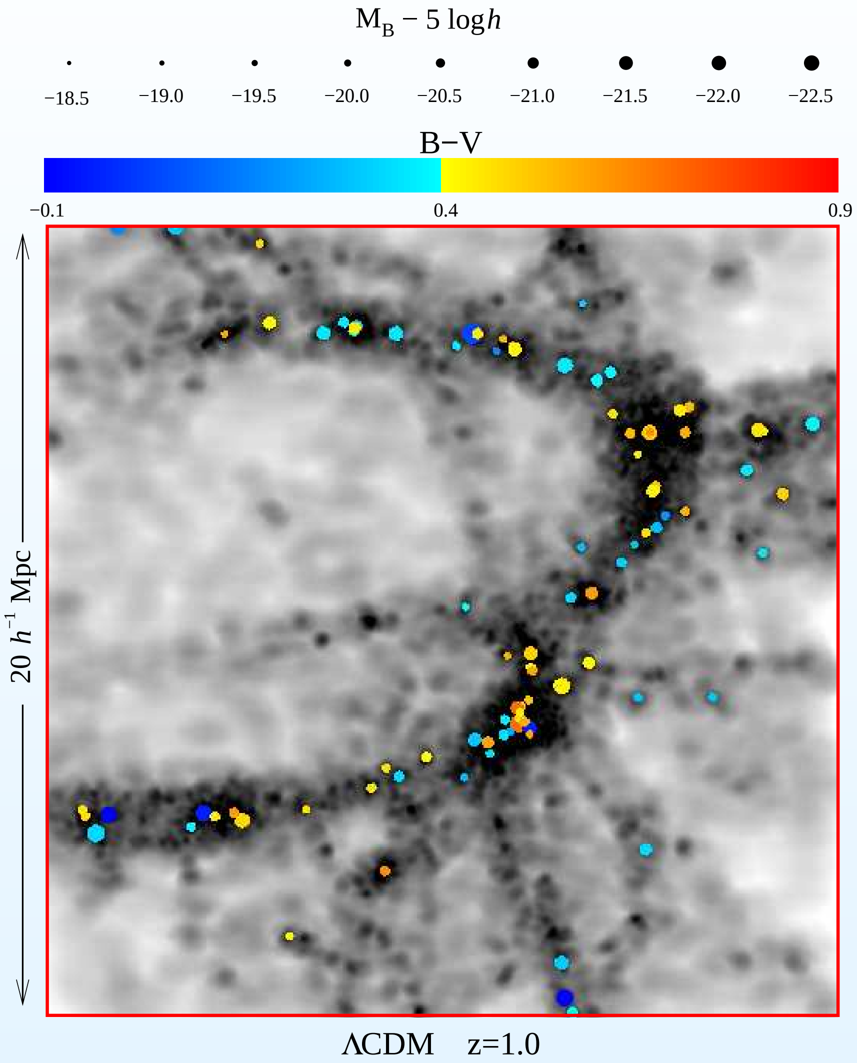
Benson, Frenk, Baugh, Cole & Lacey (2001)

Universe at $\sim$ 2 billion years



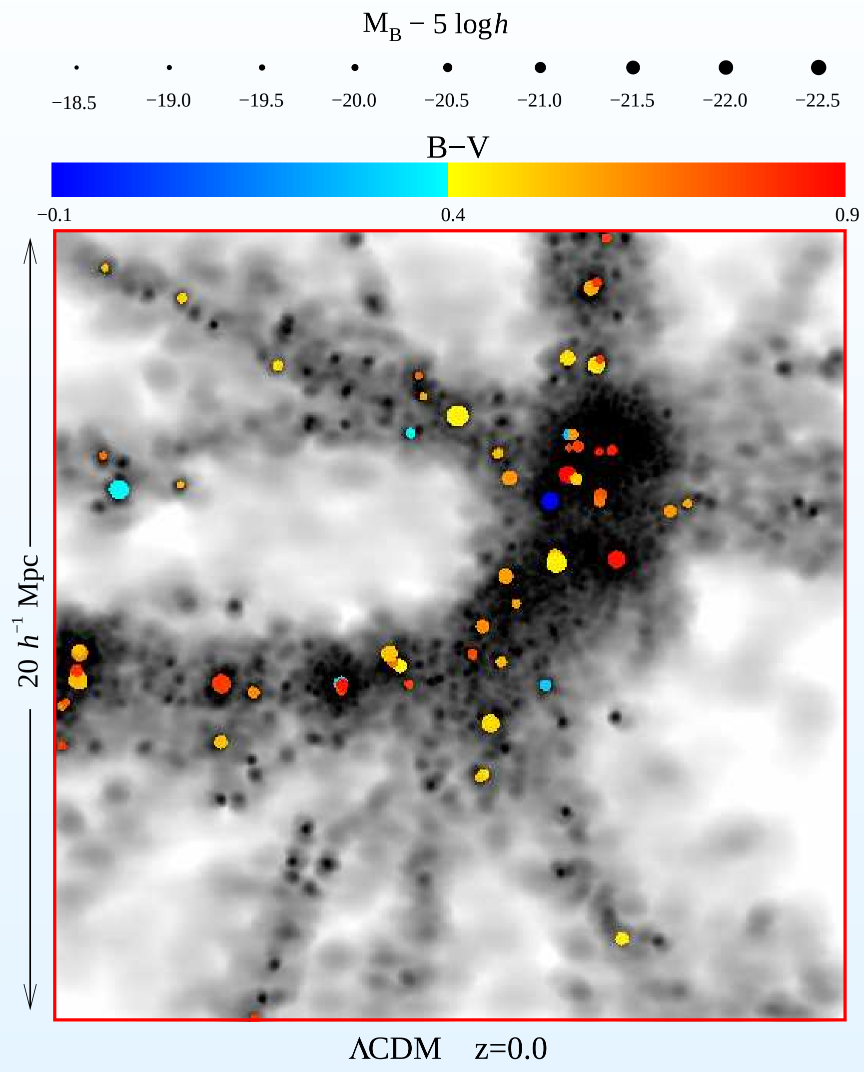
Benson, Frenk, Baugh, Cole & Lacey (2001)

Universe 4 billion years



Benson, Frenk, Baugh, Cole & Lacey (2001)

Universe 13 billion years (Today)



Benson, Frenk, Baugh, Cole & Lacey (2001)

Galform: Inputs and Outputs

Outputs: Galform provides many outputs but we start by looking at the bj and K luminosity functions

- bj luminosity function: the number of blue (i.e. young) galaxies of a certain luminosity per unit volume
- K luminosity function: the number of red (i.e. old) galaxies of a certain luminosity per unit volume

These outputs can be compared to observational data

Galform: Inputs and Outputs

Outputs: Galform provides many outputs but we start by looking at the bj and K luminosity functions

- bj luminosity function: the number of blue (i.e. young) galaxies of a certain luminosity per unit volume
- K luminosity function: the number of red (i.e. old) galaxies of a certain luminosity per unit volume

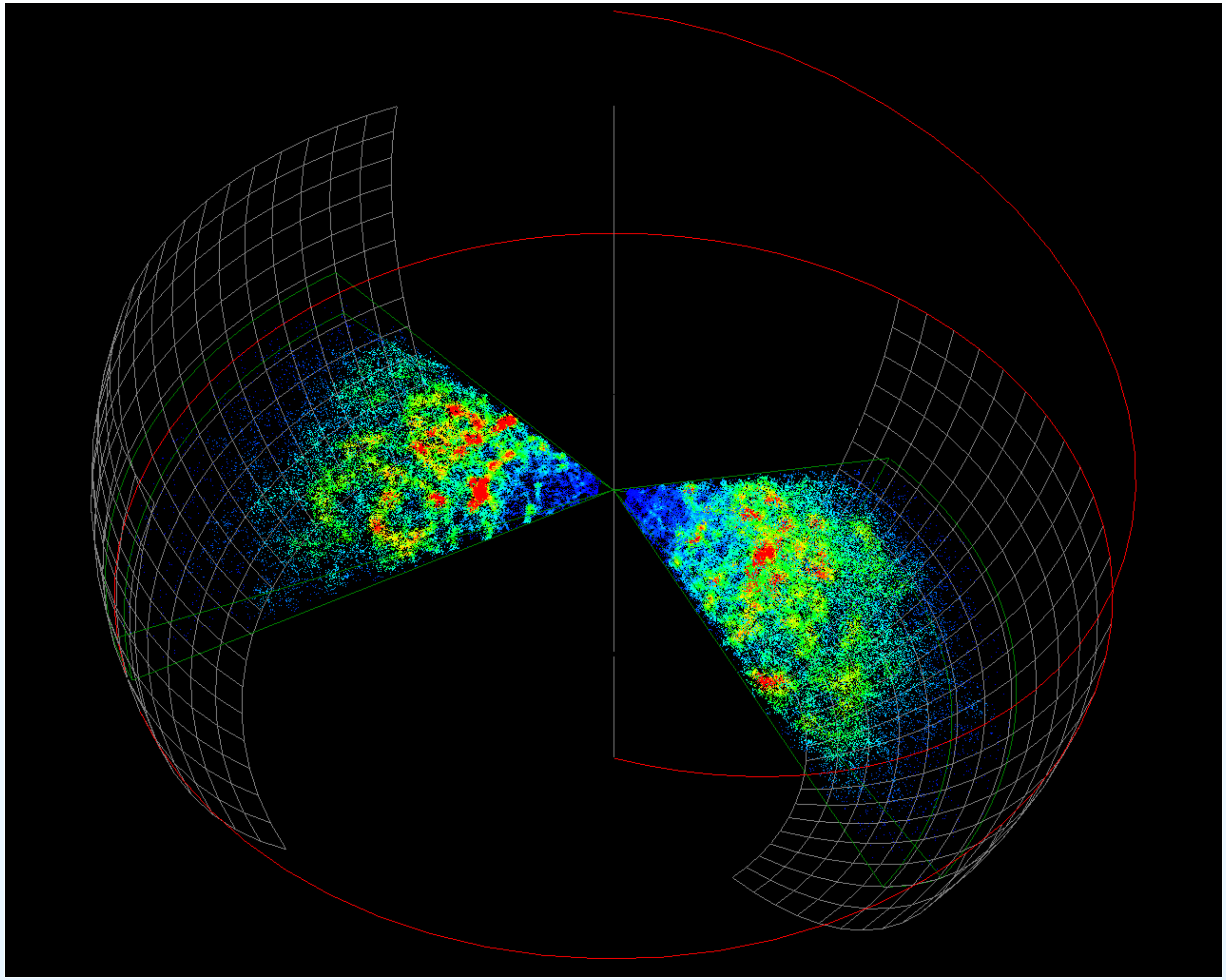
These outputs can be compared to observational data

Inputs: 17 input variables reduced to 8 after expert judgements. These include:

- **vhotdisk**: relative amount of energy in the form of gas blown out of a galaxy due to star formation
- **alphacool**: regulates the effect the central black hole has in keeping large galaxies 'hot'
- **yield**: the metal content of large galaxies

and five others: alphahot, stabledisk, epsilonStar, alphareheat and vhotburst

Observational Data: Galaxy Surveys



Earth at centre of image. Data taken by telescopes looking in two seperate directions. Galaxies observed up to a distance of 1.2 billion light years.

Galaxy Formation: Main Issues

Basic Questions

- Do we understand how galaxies form?
- Could the galaxies we observe have been formed in the presence of large amounts of dark matter?

Galaxy Formation: Main Issues

Basic Questions

- Do we understand how galaxies form?
- Could the galaxies we observe have been formed in the presence of large amounts of dark matter?

Fundamental Sources of Uncertainty

- We only observe the galaxies in our ‘local’ region of the Universe: it is possible that they are not representative of the whole Universe.
- The output of the simulation is a ‘possible’ Universe which should have similar properties to ours, but is not an exact copy.
- The output of the simulation is 512 different computer models for “slices” of the universe which are **exchangeable** with each other and (hopefully) with slices of our universe.
- We are uncertain which values of the input parameters should be used when running the model

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data

We do this in stages, as follows:

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data

We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data

We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest
- choose a subset of the outputs for which we have system observations

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an 'acceptable' fit to the output data

We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest
- choose a subset of the outputs for which we have system observations
- emulate these outputs

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data

We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest
- choose a subset of the outputs for which we have system observations
- emulate these outputs
- calculate implausibility over the selected input volume

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.

We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data

We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest
- choose a subset of the outputs for which we have system observations
- emulate these outputs
- calculate implausibility over the selected input volume
- discard all x input points that have implausibility greater than a certain cutoff

History matching and Galaxy Formation

We want to history match the Galaxy Formation model Galform using the emulation and implausibility techniques that we have outlined.
We want to reduce the volume of input parameter space as much as we can by discarding all points that we are (reasonably) sure will not give an ’acceptable’ fit to the output data
We do this in stages, as follows:

- design a set of runs of the simulator within the input volume of interest
- choose a subset of the outputs for which we have system observations
- emulate these outputs
- calculate implausibility over the selected input volume
- discard all x input points that have implausibility greater than a certain cutoff

This process is then repeated. This is **refocusing**. As we are now in a reduced input volume, outputs may be of simpler form and therefore easier to emulate.
As we have reduced the variation in the ouputs arising from the most important inputs, this also allows us to assess variation due to secondary inputs.

Galform: Stage 1

Following the cosmologist own attempt to history match Galform, we chose to run only the first 40 sub-volumes (out of 512) and examine their mean. The simulator function $f_i(x)$ is now taken to be the mean of the luminosity outputs over the first 40 sub-volumes.

Galform: Stage 1

Following the cosmologist own attempt to history match Galform, we chose to run only the first 40 sub-volumes (out of 512) and examine their mean. The simulator function $f_i(x)$ is now taken to be the mean of the luminosity outputs over the first 40 sub-volumes.

Design: Ran a 1000 point Latin Hypercube design across the key input parameters

Galform: Stage 1

Following the cosmologist own attempt to history match Galform, we chose to run only the first 40 sub-volumes (out of 512) and examine their mean. The simulator function $f_i(x)$ is now taken to be the mean of the luminosity outputs over the first 40 sub-volumes.

Design: Ran a 1000 point Latin Hypercube design across the key input parameters

Outputs: Decided to choose 11 outputs from the luminosity functions as they could be emulated accurately

Galform: Stage 1

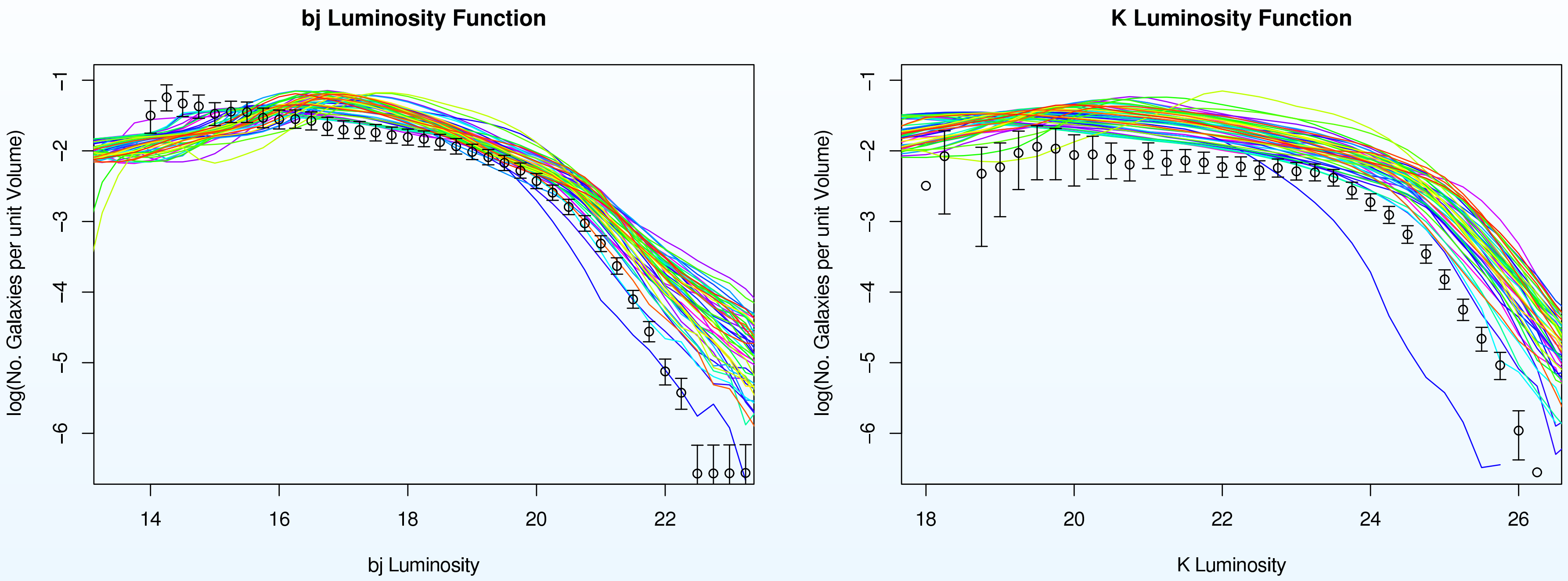
Following the cosmologist own attempt to history match Galform, we chose to run only the first 40 sub-volumes (out of 512) and examine their mean. The simulator function $f_i(x)$ is now taken to be the mean of the luminosity outputs over the first 40 sub-volumes.

Design: Ran a 1000 point Latin Hypercube design across the key input parameters

Outputs: Decided to choose 11 outputs from the luminosity functions as they could be emulated accurately

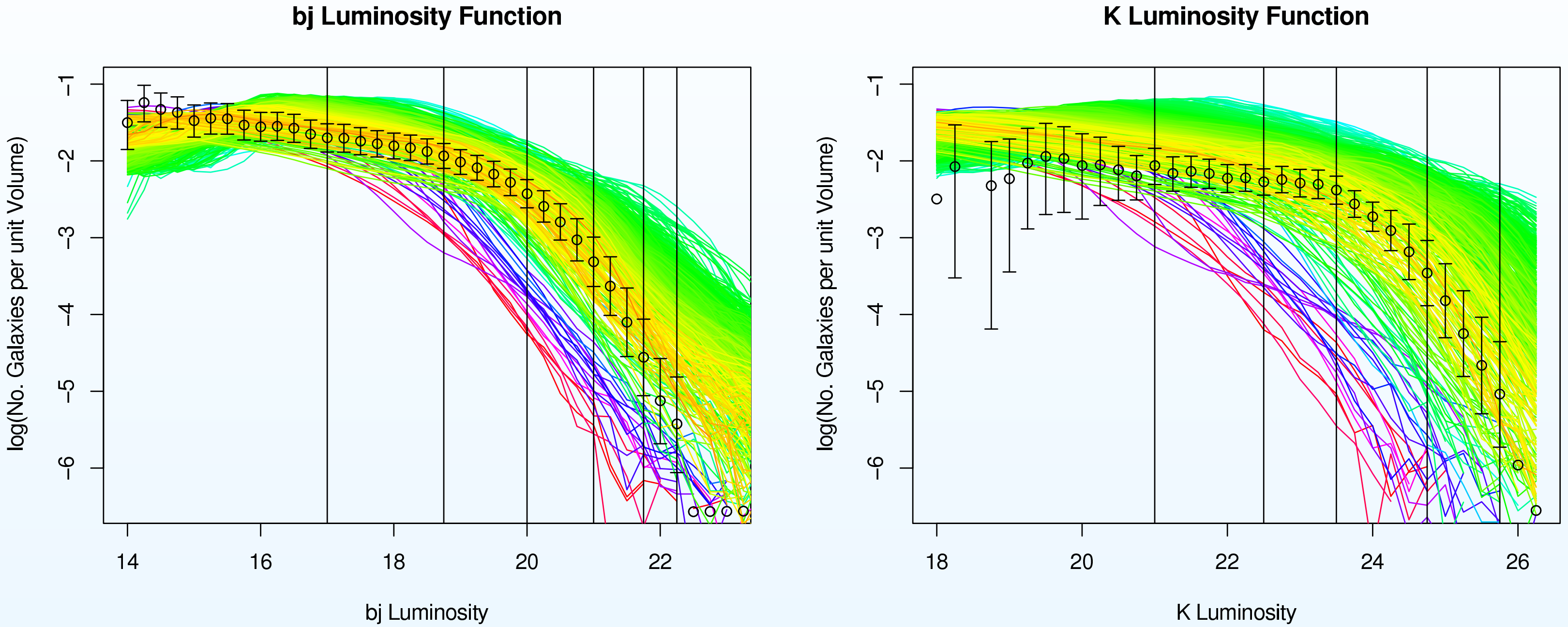
Active Variables: For each output we choose 5 active variables x^A, i.e. those inputs which are the most important for explaining variation in the output.

Galform Outputs: The Luminosity Functions



- Bj Luminosity: young (blue) galaxies
- K Luminosity: old (red) galaxies
- Circles are observed
- Coloured lines are example outputs for different input parameter choices

11 Output points Chosen



Outputs chosen to be informative enough to allow us to cut down the parameter space, but simple enough to be emulated easily.

Galform: Stage 1

We then emulate each of the 11 outputs univariately using:

$$f_i(x) = \sum_j \beta_{ij}\, g_{ij}(x^A) + u_i(x^A) + \delta_i(x)$$

where now $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are now monomials in x^A of order 3 or less, and $u(x^A)$ is a gaussian process. The nugget $\delta_i(x)$ models the effects of inactive variables as random noise.

Galform: Stage 1

We then emulate each of the 11 outputs univariately using:

$$f_i(x) = \sum_j \beta_{ij}\, g_{ij}(x^A) + u_i(x^A) + \delta_i(x)$$

where now $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are now monomials in x^A of order 3 or less, and $u(x^A)$ is a gaussian process. The nugget $\delta_i(x)$ models the effects of inactive variables as random noise.

The $u_i(x)$ have covariance structure given by:

$$\mathrm{Cov}(u_i(x_1^A), u_i(x_2^A), \;) \;=\; \sigma^2 \exp[-\theta_i |x_1^A - x_2^A|^2]$$

Galform: Stage 1

We then emulate each of the 11 outputs univariately using:

$$f_i(x) = \sum_j \beta_{ij}\, g_{ij}(x^A) + u_i(x^A) + \delta_i(x)$$

where now $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are now monomials in x^A of order 3 or less, and $u(x^A)$ is a gaussian process. The nugget $\delta_i(x)$ models the effects of inactive variables as random noise.

The $u_i(x)$ have covariance structure given by:

$$\mathrm{Cov}(u_i(x_1^A), u_i(x_2^A), \ ) \ = \ \sigma^2 \exp[-\theta_i |x_1^A - x_2^A|^2]$$

Note that different outputs have different active variables: all of the 8 input variables were found to be active for at least one of the outputs.

Galform: Stage 1

We then emulate each of the 11 outputs univariately using:

$$f_i(x) = \sum_j \beta_{ij}\, g_{ij}(x^A) + u_i(x^A) + \delta_i(x)$$

where now $B = \{\beta_{ij}\}$ are unknown scalars, g_{ij} are now monomials in x^A of order 3 or less, and $u(x^A)$ is a gaussian process. The nugget $\delta_i(x)$ models the effects of inactive variables as random noise.

The $u_i(x)$ have covariance structure given by:

$$\mathrm{Cov}(u_i(x_1^A), u_i(x_2^A), \ ) \ = \ \sigma^2 \exp[-\theta_i |x_1^A - x_2^A|^2]$$

Note that different outputs have different active variables: all of the 8 input variables were found to be active for at least one of the outputs.

Here the Adjusted R^2 for the polynomial fits were between 0.79 and 0.92

The Emulators give the expectation $\mathrm{E}(f_i(x))$ and variance $\mathrm{Var}(f_i(x))$ at point x for each output given by $i = 1, ..., 11$.

Model Discrepancy and Measurement Error

Before calculating the implausibility we need to assess the Model Discrepancy and Measurement error.

Model Discrepancy and Measurement Error

Before calculating the implausibility we need to assess the Model Discrepancy and Measurement error.

Model Discrepancy has three components:

- Φ_E: Expert assessment of model discrepancy of full model with 17 parameters and using 512 sub-volumes

Model Discrepancy and Measurement Error

Before calculating the implausibility we need to assess the Model Discrepancy and Measurement error.

Model Discrepancy has three components:

- Φ_E: Expert assessment of model discrepancy of full model with 17 parameters and using 512 sub-volumes

- Φ_{40}: Discrepancy term due to (i) choosing first 40 sub-volumes from full 512 sub-volumes, and (ii) need to extrapolate to our universe. Assess this by repeating 100 runs but now choosing 40 random regions.

  [More carefully, we may construct an **exchangeable** system of emulators to fully account for this discrepancy.]

Model Discrepancy and Measurement Error

Before calculating the implausibility we need to assess the Model Discrepancy and Measurement error.

Model Discrepancy has three components:

- Φ_E: Expert assessment of model discrepancy of full model with 17 parameters and using 512 sub-volumes

- Φ_{40}: Discrepancy term due to (i) choosing first 40 sub-volumes from full 512 sub-volumes, and (ii) need to extrapolate to our universe. Assess this by repeating 100 runs but now choosing 40 random regions.

  [More carefully, we may construct an **exchangeable** system of emulators to fully account for this discrepancy.]

- Φ_{12}: As we have neglected 9 parameters (due to expert advice) we need to assess effect of this (by running latin hypercube design across all 17 parameters)

Model Discrepancy and Measurement Error

Observational Errors composed of 4 parts:

- Normalisation Error: correlated vertical error on all luminosity output points

Model Discrepancy and Measurement Error

Observational Errors composed of 4 parts:

- Normalisation Error: correlated vertical error on all luminosity output points
- Luminostiy Zero Point Error: correlated horizontal error on all luminosity points

Model Discrepancy and Measurement Error

Observational Errors composed of 4 parts:

- Normalisation Error: correlated vertical error on all luminosity output points
- Luminostiy Zero Point Error: correlated horizontal error on all luminosity points
- $k + e$ Correction Error: Outputs have to be corrected for the fact that galaxies are moving away from us at different speeds (light is red-shifted), and for the fact that galaxies are seen in the past (as light takes millions of years to reach us)

Model Discrepancy and Measurement Error

Observational Errors composed of 4 parts:

- Normalisation Error: correlated vertical error on all luminosity output points
- Luminostiy Zero Point Error: correlated horizontal error on all luminosity points
- $k + e$ Correction Error: Outputs have to be corrected for the fact that galaxies are moving away from us at different speeds (light is red-shifted), and for the fact that galaxies are seen in the past (as light takes millions of years to reach us)
- Poisson Error: assumed Poisson process to describe galaxy production (not very accurate assumption!)

Implausibility

With the help of expert judgements, we decide on an Implausibility cutoff: all x points with $I_M(x)$ greater than this cutoff are discarded as they represent points that we expect would give a fit to the data that would be unacceptable to the experts.

Implausibility

With the help of expert judgements, we decide on an Implausibility cutoff: all x points with $I_M(x)$ greater than this cutoff are discarded as they represent points that we expect would give a fit to the data that would be unacceptable to the experts.
For first stage decided on a cutoff of $I_M(x) < 3.5$ With this choice we were able to exclude approximately 92% of the initial input space.

Implausibility

With the help of expert judgements, we decide on an Implausibility cutoff: all x points with $I_M(x)$ greater than this cutoff are discarded as they represent points that we expect would give a fit to the data that would be unacceptable to the experts.

For first stage decided on a cutoff of $I_M(x) < 3.5$ With this choice we were able to exclude approximately 92% of the initial input space.

We want to understand the structure of the implausibility function of input parameter space.

Implausibility

With the help of expert judgements, we decide on an Implausibility cutoff: all x points with $I_M(x)$ greater than this cutoff are discarded as they represent points that we expect would give a fit to the data that would be unacceptable to the experts.

For first stage decided on a cutoff of $I_M(x) < 3.5$ With this choice we were able to exclude approximately 92% of the initial input space.

We want to understand the structure of the implausibility function of input parameter space.

In the first stage we have 8 input parameters that are active, so it is difficult to visualise or even to evaluate the implausibility $I_M(x)$ over a 8D grid of 15^8 points.

Implausibility

With the help of expert judgements, we decide on an Implausibility cutoff: all x points with $I_M(x)$ greater than this cutoff are discarded as they represent points that we expect would give a fit to the data that would be unacceptable to the experts.

For first stage decided on a cutoff of $I_M(x) < 3.5$ With this choice we were able to exclude approximately 92% of the initial input space.

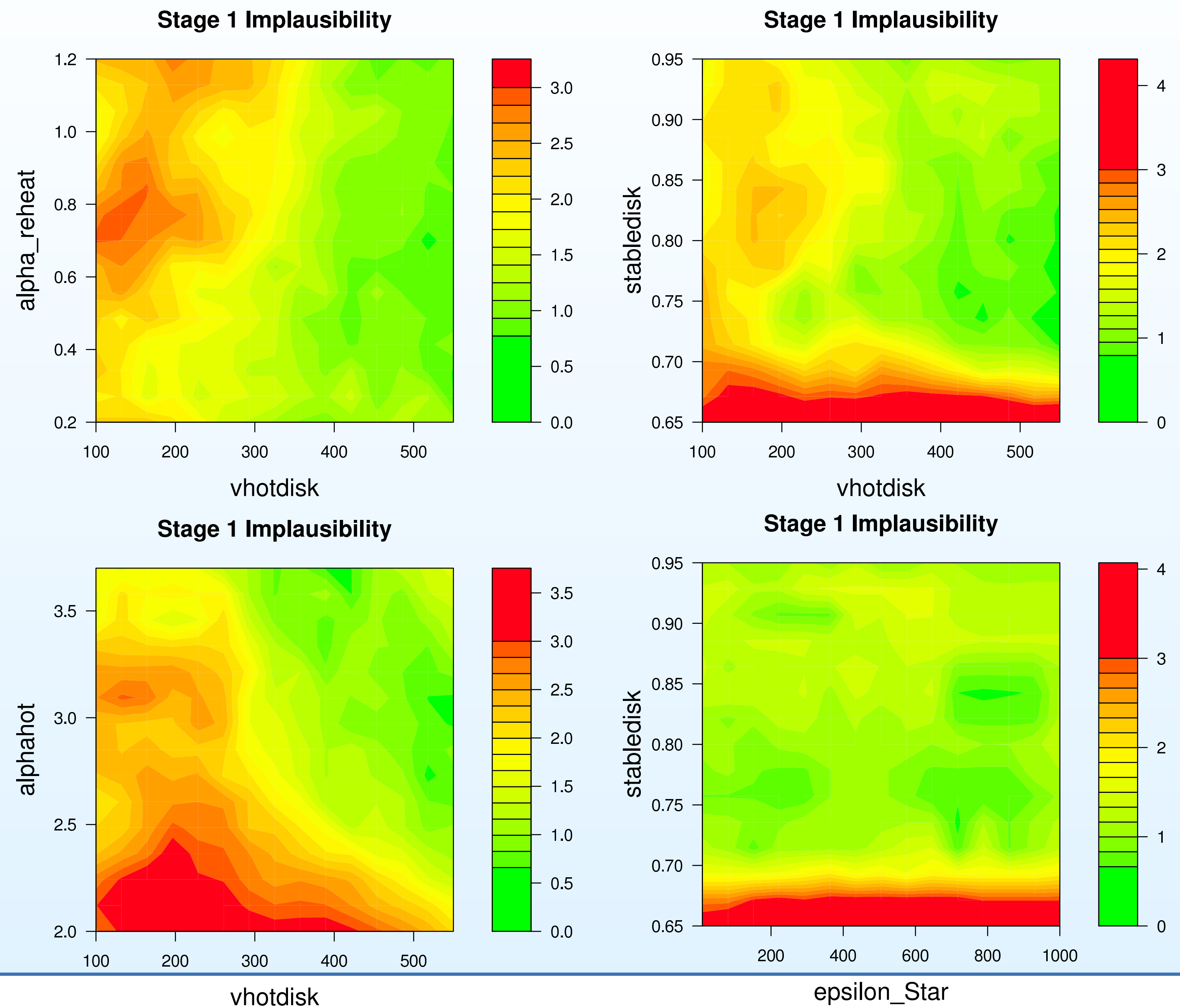
We want to understand the structure of the implausibility function of input parameter space.

In the first stage we have 8 input parameters that are active, so it is difficult to visualise or even to evaluate the implausibility $I_M(x)$ over a 8D grid of 15^8 points.

2-Dimensional Projection For each of the 28 pairs of active variables we minimize the implausibility across the remaining active variables

If a point on these plots is implausible then it will be implausible for any choice of the other active variables.

2D Implausibility Projections: Stage 1 (8%)



Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

We can now refocus by designing a new set of runs of Galform, where these runs are restricted to the reduced or non-implausible volume.

Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

We can now refocus by designing a new set of runs of Galform, where these runs are restricted to the reduced or non-implausible volume.

Do this by generating a large latin hypercube design over the original active variables, and reject any points in this design that have $I_M(x) > 3.5$.

Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

We can now refocus by designing a new set of runs of Galform, where these runs are restricted to the reduced or non-implausible volume.

Do this by generating a large latin hypercube design over the original active variables, and reject any points in this design that have $I_M(x) > 3.5$.

We found that 2004 points survived and these were taken as the next set of runs for Galform.

Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

We can now refocus by designing a new set of runs of Galform, where these runs are restricted to the reduced or non-implausible volume.

Do this by generating a large latin hypercube design over the original active variables, and reject any points in this design that have $I_M(x) > 3.5$.

We found that 2004 points survived and these were taken as the next set of runs for Galform.

It is possible that the reduced volume is now composed of 2 or more unconnected volumes. We need to know this as each volume would most likely have very different behavior and hence in Stage 2 we would emulate each of these volumes separately.

Understanding the Structure of the Non-Implausible Volume

Once the implausibility cutoff $I_M(x) < 3.5$ has been imposed this defines the volume of input parameter space that we are still interested in searching.

We can now refocus by designing a new set of runs of Galform, where these runs are restricted to the reduced or non-implausible volume.

Do this by generating a large latin hypercube design over the original active variables, and reject any points in this design that have $I_M(x) > 3.5$.

We found that 2004 points survived and these were taken as the next set of runs for Galform.

It is possible that the reduced volume is now composed of 2 or more unconnected volumes. We need to know this as each volume would most likely have very different behavior and hence in Stage 2 we would emulate each of these volumes separately.

We used Cluster Analysis on points in the reduced volume to determine connectedness.

In this case the region was found to be simply connected.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.
- As we are now dealing with only 8 percent of the original input parameter volume, the Galform outputs may be smoother.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.
- As we are now dealing with only 8 percent of the original input parameter volume, the Galform outputs may be smoother.
- Therefore our emulation techniques may give better representations for the outputs, and we may be able to include more output variables that were previously hard to emulate with sufficient reliability.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.
- As we are now dealing with only 8 percent of the original input parameter volume, the Galform outputs may be smoother.
- Therefore our emulation techniques may give better representations for the outputs, and we may be able to include more output variables that were previously hard to emulate with sufficient reliability.
- We now use all 8 variables in each emulator.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.
- As we are now dealing with only 8 percent of the original input parameter volume, the Galform outputs may be smoother.
- Therefore our emulation techniques may give better representations for the outputs, and we may be able to include more output variables that were previously hard to emulate with sufficient reliability.
- We now use all 8 variables in each emulator.
- The Adjusted R^2 of the emulators were now between 0.83 and 0.98. This is due to previously masked variance now being resolved as we are emulating a smoother function with more active variables.

Galform: Stage 2

- We now repeat the above process but now we are only studying Galform's behavior over the reduced volume.
- As we are now dealing with only 8 percent of the original input parameter volume, the Galform outputs may be smoother.
- Therefore our emulation techniques may give better representations for the outputs, and we may be able to include more output variables that were previously hard to emulate with sufficient reliability.
- We now use all 8 variables in each emulator.
- The Adjusted R^2 of the emulators were now between 0.83 and 0.98. This is due to previously masked variance now being resolved as we are emulating a smoother function with more active variables.
- We can now calculate the implausibility as before. As our emulator accuracy improves, we may further reduce the input parameter space.

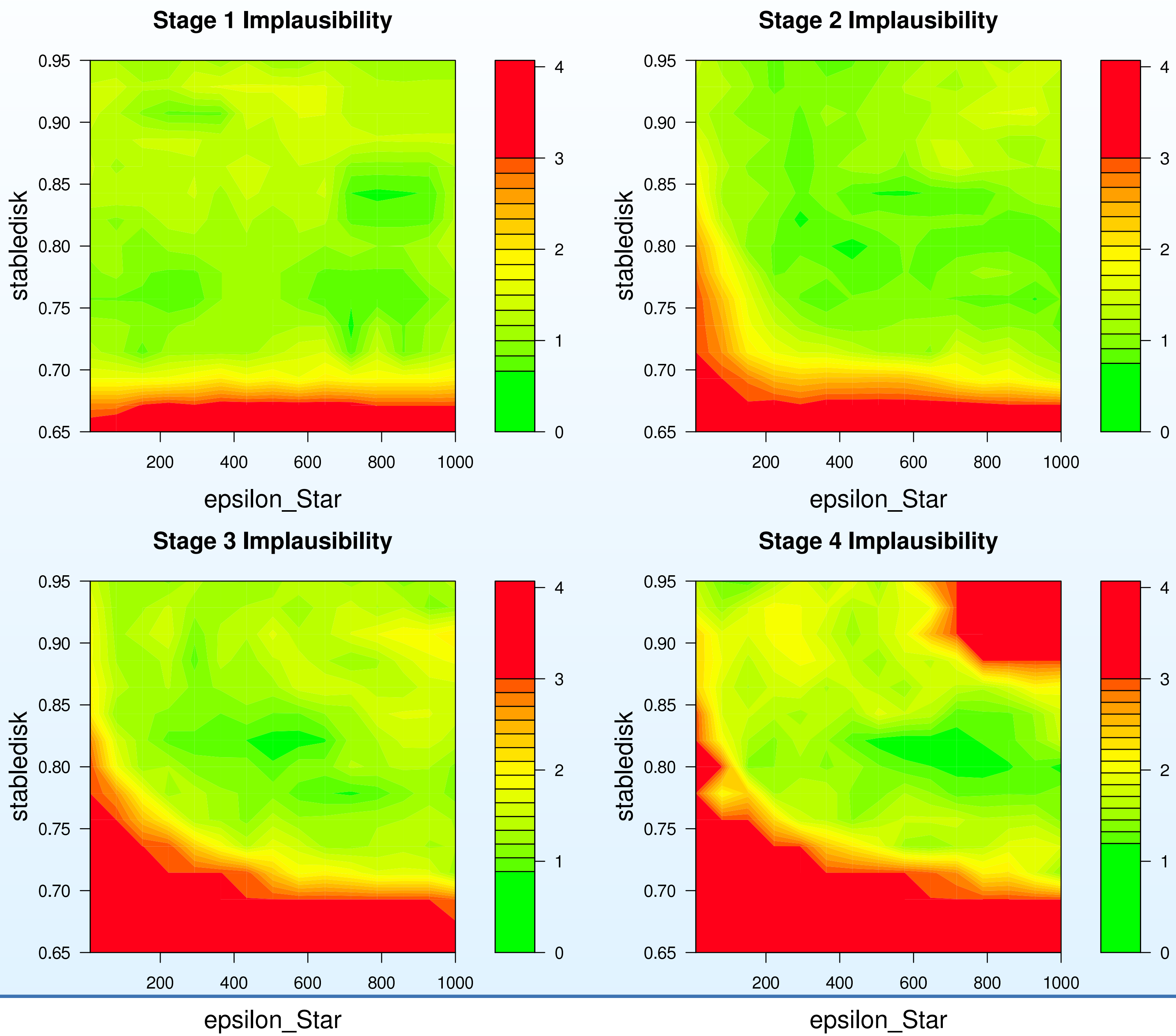
Summary of Results

- We have completed Four Stages:

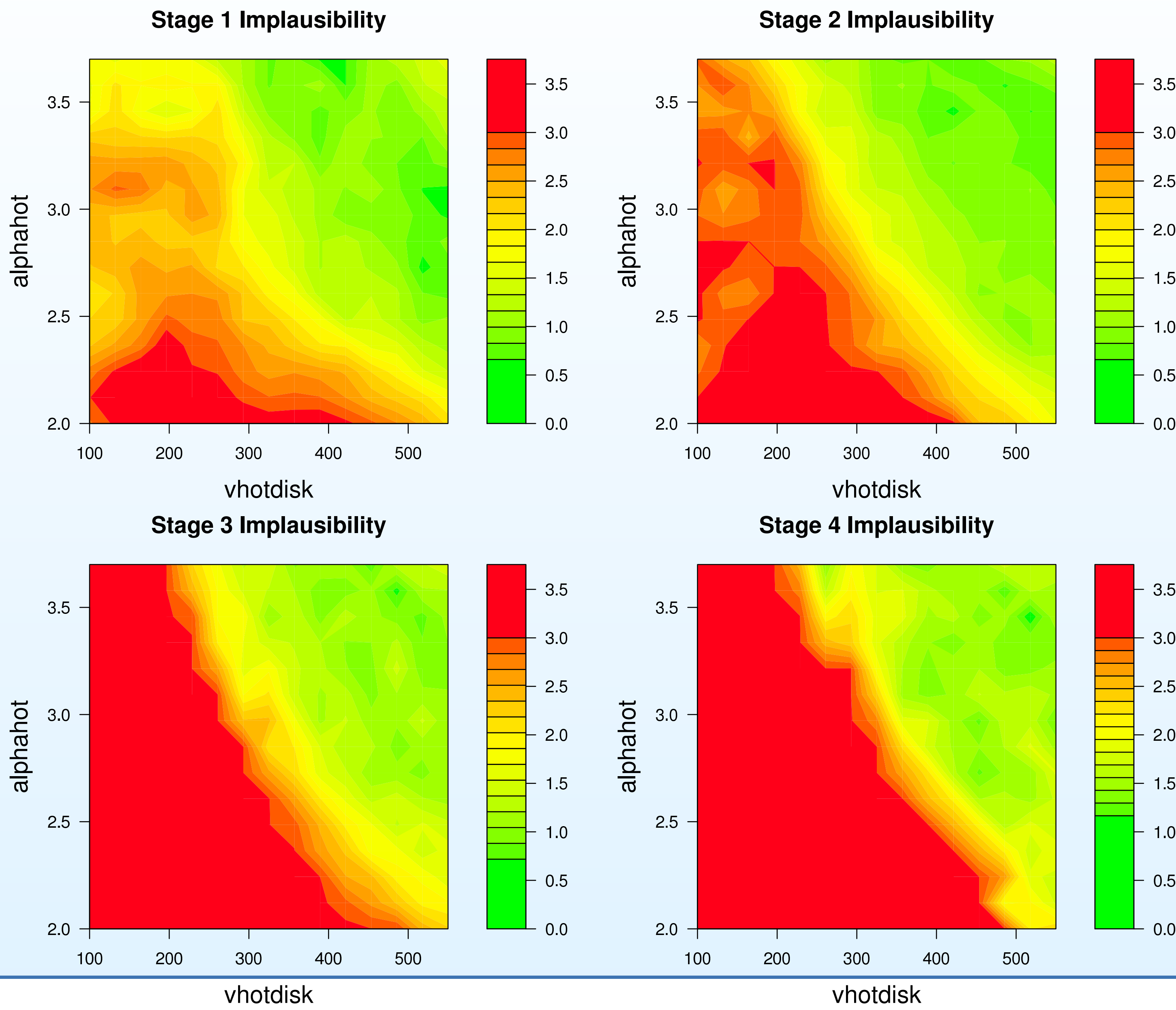
	No. Model Runs	No. Active Vars	Adjusted $R2$	Space Remaining
Stage 1	1000	5	0.58 - 0.90	8.0 %
Stage 2	1916	8	0.83 - 0.98	2.9 %
Stage 3	1487	8	0.79 - 0.99	1.2 %
Stage 4	1899	10	0.75 - 0.99	0.12 %

- In Stages 3 and 4 we used a Multivariate Implausibility measure to help reduce space further.
- In Stage 4 we included 2 more active input variables that had previously been inactive.

2D Implausibility Projections: Stage 1 (8%) to Stage 4 (0.12%)



2D Implausibility Projections: Stage 1 (8%) to Stage 4 (0.12%)



Concluding comments: 1

The approach that we have described appears to be very effective in identifying the sub-space of potential matches to observations of the universe.

Concluding comments: 1

The approach that we have described appears to be very effective in identifying the sub-space of potential matches to observations of the universe.

We are now at the stage of introducing additional physical observations to constrain further the parameter space. In this way, we have found good matches to a variety of other phenomena that the model is intended to explain.

Concluding comments: 1

The approach that we have described appears to be very effective in identifying the sub-space of potential matches to observations of the universe.

We are now at the stage of introducing additional physical observations to constrain further the parameter space. In this way, we have found good matches to a variety of other phenomena that the model is intended to explain.

In general, this approach to statistical ill-posed inverse problems, namely finding the class of all “history matches” which are of sufficiently good quality subject to all of the irreducible uncertainties associated with the model and the data, seems tractable, promising and effective.

Concluding comments: 1

The approach that we have described appears to be very effective in identifying the sub-space of potential matches to observations of the universe.

We are now at the stage of introducing additional physical observations to constrain further the parameter space. In this way, we have found good matches to a variety of other phenomena that the model is intended to explain.

In general, this approach to statistical ill-posed inverse problems, namely finding the class of all “history matches” which are of sufficiently good quality subject to all of the irreducible uncertainties associated with the model and the data, seems tractable, promising and effective.

History matching often is the first stage of a larger methodology which applies the uncertainty models that we have described for activities such as system forecasting and optimisation.

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?

  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?

  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?
  Probably not much.

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?
  Probably not much.
- What analysis could possibly be done to justify (or contradict) this conclusion?

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?
  Probably not much.
- What analysis could possibly be done to justify (or contradict) this conclusion?
  The type of full Bayes linear uncertainty analysis that we've described.

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?
  Probably not much.
- What analysis could possibly be done to justify (or contradict) this conclusion?
  The type of full Bayes linear uncertainty analysis that we've described.
- What do we learn about real physical systems from the analysis of (necessarily imperfect) models?

Concluding comments: 2

The chances of the Gulf stream shutting down are now thought to be greater than 50%

Questions

- What does this statement mean?
  It should be a statement of the uncertainty of the scientist about the physical process given all of the uncertainties inherent in the physical modelling.
- What analysis was actually done to reach this conclusion?
  Probably not much.
- What analysis could possibly be done to justify (or contradict) this conclusion?
  The type of full Bayes linear uncertainty analysis that we've described.
- What do we learn about real physical systems from the analysis of (necessarily imperfect) models?
  If we account for all uncertainties (which may be difficult), then we obtain a full uncertainty analysis for the behaviour of the physical system.

References

P.S. Craig, M. Goldstein, A.H. Seheult, J.A. Smith (1997). Pressure matching for hydocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments (with discussion), in Case Studies in Bayesian Statistics, vol. III, eds. C. Gastonis et al. 37-93. Springer-Verlag.

M. Goldstein and J.C.Rougier (2006). Bayes linear calibrated prediction for complex systems, JASA, 101, 1132-1143

References

P.S. Craig, M. Goldstein, A.H. Seheult, J.A. Smith (1997). Pressure matching for hydocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments (with discussion), in Case Studies in Bayesian Statistics, vol. III, eds. C. Gastonis et al. 37-93. Springer-Verlag.

M. Goldstein and J.C.Rougier (2006). Bayes linear calibrated prediction for complex systems, JASA, 101, 1132-1143

M. Goldstein and J.C.Rougier (2009). Reified Bayesian modelling and inference for physical systems (with discussion), JSPI, 139, 1221-1239.

References

P.S. Craig, M. Goldstein, A.H. Seheult, J.A. Smith (1997). Pressure matching for hydocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments (with discussion), in Case Studies in Bayesian Statistics, vol. III, eds. C. Gastonis et al. 37-93. Springer-Verlag.

M. Goldstein and J.C.Rougier (2006). Bayes linear calibrated prediction for complex systems, JASA, 101, 1132-1143

M. Goldstein and J.C.Rougier (2009). Reified Bayesian modelling and inference for physical systems (with discussion), JSPI, 139, 1221-1239.

Kennedy, M.C. and O'Hagan, A. (2001). Bayesian calibration of computer models (with discussion). Journal of the Royal Statistical Society, B,63, 425-464

Santner, T., Williams, B. and Notz, W. (2003). The Design and Analysis of Computer Experiments. Springer Verlag: New York.

References

P.S. Craig, M. Goldstein, A.H. Seheult, J.A. Smith (1997). Pressure matching for hydocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments (with discussion), in Case Studies in Bayesian Statistics, vol. III, eds. C. Gastonis et al. 37-93. Springer-Verlag.

M. Goldstein and J.C.Rougier (2006). Bayes linear calibrated prediction for complex systems, JASA, 101, 1132-1143

M. Goldstein and J.C.Rougier (2009). Reified Bayesian modelling and inference for physical systems (with discussion), JSPI, 139, 1221-1239.

Kennedy, M.C. and O'Hagan, A. (2001). Bayesian calibration of computer models (with discussion). Journal of the Royal Statistical Society, B,63, 425-464

Santner, T., Williams, B. and Notz, W. (2003). The Design and Analysis of Computer Experiments. Springer Verlag: New York.

And check out the website for the
Managing Uncertainty in Complex Models (MUCM) project
[A consortium of Aston, Durham, LSE, Sheffield and Southampton all hard at work on developing technology for computer model uncertainty problems.]