Modelling and Simulation in Modern Science

or

The Hitch-hiker's guide to Model Land

or

The Cat That Looks Most Like A Dog

or

Escaping from Model Land

Draft compiled September 16, 2019

Contents

1 Model worlds and the real world **5**
1.1 One real world; infinitely many model worlds 5
1.2 Motivation: Why do we construct model worlds? What is it all for? 6
1.3 Some consequences: teaser examples, for later in-depth treatment (one per chapter???) 6
1.4 Solutions: teaser/summary 6

2 What do model worlds tell us? **7**
2.1 Abstraction and simplification 7
2.2 Model world quantities 9
2.2.1 How the real world influences the model world, and vice-versa
2.2.2 Real world -> model world
2.2.3 Model world -> real world (eg implied volatility, climate sensitivity)
2.2.4 No true parameter set exists
2.3 Complexity. Subtractability. Noise/music. 12
2.4 Dynamical (process-based) models versus statistical models 12
2.4.1 ?explainability? and interpretability
2.4.2 Horses and Hawkmoths
2.4.3 Note on ?machine learning??

3 Butterflies and Hawkmoths **13**
3.1 Introduction 14
3.2 Butterfly Effect. Initial condition ensembles. Solved. 14
3.3 Hawkmoth Effect. Model structure ensembles? Unsolved. 14
3.4 Ensembles 14
3.4.1 Independence?
3.4.2 Model groupthink
3.4.3 Comment on ?sensitivity analysis?
3.4.4 So what does the ensemble range mean?
3.4.5 A model asylum

3.5 Consequences of the Hawkmoth Effect 15
3.5.1 Things which are not included may not be negligible
3.5.2 Models cannot be monotonically improved
3.6 Mathematical treatment (easily skippable!) 15

4 The Cat that looks most like a Dog 17
4.1 The cat that looks most like a dog 19
4.2 On Bayesian methods 19
4.2.1 Bayes theorem does not apply
4.3 Methodology is epistemology 19
4.4 Many other statstical methods have the same underlying problem (a reference class problem) 19

5 Escaping from Model Land 21
5.1 Statements made about real world need to refer to the real world 22
5.2 Evaluation against past data 24
5.2.1 Evaluation metrics (eg do we require propriety if models are never perfect?)
5.2.2 Big Data
5.2.3 Small Data (accounting for small-number problems)
5.2.4 No data? The reference class problem
5.3 Use of expert judgement 24

6 Making decisions based on models 25
6.1 The aim: getting ?the right answer? or supporting a decision? 25
6.1.1 Properties of good decisions
6.1.2 Sometimes the skill of the model is unimportant (Mali farmers; Start; horoscopes; biodynamic planting?)
6.1.3 Sometimes the confidence, even where reported honestly, is recalibrated by the audience (?ninety per cent? =/= 90%)
6.2 Evaluation of models, evaluation of forecasts 25

7 Conclusions and consequences 27
7.1 The unreasonable effectiveness of mathematics in the natural sciences? 28
7.2 Modelling for decision support 28
7.3 Alternative strategies, find positive/constructive note to end on. 28

Bibliography 29

Chapter 1

Model worlds and the real world

This will be an introductory chapter

1.1 One real world; infinitely many model worlds

Though there is only one real world, there are many model worlds.

Model worlds include any and all representations, from the material or mathematical to the conceptual or artistic.

Model worlds may be complete and internally self-consistent, or more likely they may be incomplete and self-contradictory.

Each model is different, by virtue of abstracting different elements of the original real world.

The model-world, contained and defined by the model, then functions as a surrogate object for the purpose of recording or drawing certain conclusions about the real world, when the characteristics of the real world are inaccessible or unobservable.

Model worlds are constructed in lots of different ways:

- physical model of flood plain

- representational model eg OS map
- biological model eg mouse, e. coli
- conceptual model eg wave/particle
- statistical model
- simulation model eg climate, traffic

1.2 Motivation: Why do we construct model worlds? What is it all for?

1.3 Some consequences: teaser examples, for later in-depth treatment (one per chapter???)

1.4 Solutions: teaser/summary

Chapter 2

What do model worlds tell us?

Model: a thing or concept believed to have a correspondence with the real world, such that one may draw conclusions about properties of the real world by observing the model.

Model world: the world-in-the-model: the set of observable states of the model

[Does model-world have a hyphen or not?]

[define parameter, variable, etc]

2.1 Abstraction and simplification

Though there is only one real world, there are many model worlds. Model worlds include any and all representations, from the material or mathematical to the conceptual or artistic. Model worlds may be complete and internally self-consistent, or more likely they may be incomplete and self-contradictory. Each model is different, by virtue of abstracting different elements of the original real world. The model-world, contained and defined by the model, then functions as a surrogate object for the purpose of recording or drawing certain conclusions about the real world, when the characteristics of the real world are inaccessible or unobservable.

Some degree of abstraction and simplification is generally necessary for a model to be of value. An exact replica of the object of study would likely be no more accessible than the object itself, as is memorably described in Borges' 1946 short story "On Exactitude In Science" (quoted below). The important question is not the degree to which the map resembles the territory, but the useful information content of the map in terms of the kinds of questions we might wish to answer with it. If I am asking which is the best way to walk up a mountain, I may be interested in a contour map; if I am looking for a certain kind of fossil then I would prefer a geological map; if I would like to know where to find something to eat on my return, neither of the above will be of much use. Efforts to produce a larger, better, more detailed map sometimes arise from a genuine need for detail and sometimes instead from a lack of clarity about the question.

On Exactitude in Science **Box 2.1**

'... *In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography*."

– *Suárez Miranda, Viajes de varones prudentes, Libro IV, Cap. XLV, Lérida, 1658*

Jorge Luis Borges, 1946, translated by XX

Abstracting the essential qualities of a situation in order to construct a model is something that we do routinely and unconsciously as a means of understanding the world from a limited perspective. There is a model, of sorts, in our expectation that the colour of a raincoat does not affect its waterproofing, or that the sun will rise in the east again tomorrow. These models might be based on long observation that A-follows-B or on a detailed logical explanation, or they may be some kind of unrationalised intuition. Scientific models are constructed in much the same way, either as a statistical relationship between observations of one set of events and another, or as a logical deduction given a set of starting points. Scientific models also arise from intuition, though rationalisations of that intuition are demanded as part of today's scientific method and I will return to discussion of so-called "expert judgement" later.

2.2 Model world quantities

2.2.1 How the real world influences the model world, and vice-versa

Quantities in the model world are usually named after their real-world counterparts, but just as the map is not the territory, neither are the labels the objects. Thus, care is needed to ensure that even those model world quantities (variables) thought to have a strong relationship with the real world quantities (measurables) are not accidentally or implicitly given real-world status.

2.2.2 Real world -> model world

In some cases, the undisputed reality of real-world quantities such as the freezing point of water or the acceleration due to gravity give them a very high epistemic status in the mind of the scientific modeller, such that they are given fixed, known values. By contrast, it is acceptable (often unavoidable) to calibrate lower-status parameters such as a cloud fraction or leaf area index by altering or tuning their values (within a range deemed to be reasonable) to give the best possible correspondence of the output with observation. This makes sense in that we "know the answer" for the real-world freezing point of water in a way that we don't for other parameters, but we are still mistaking the map for the territory: why do we think we "know the answer" in the model world? It could be the case that the freezing point of water in our model world "is" actually 1 degree different from the real-world value, in the sense that a freezing point of water chosen to be 1 degree rather than zero could give output which corresponds more closely to observation.

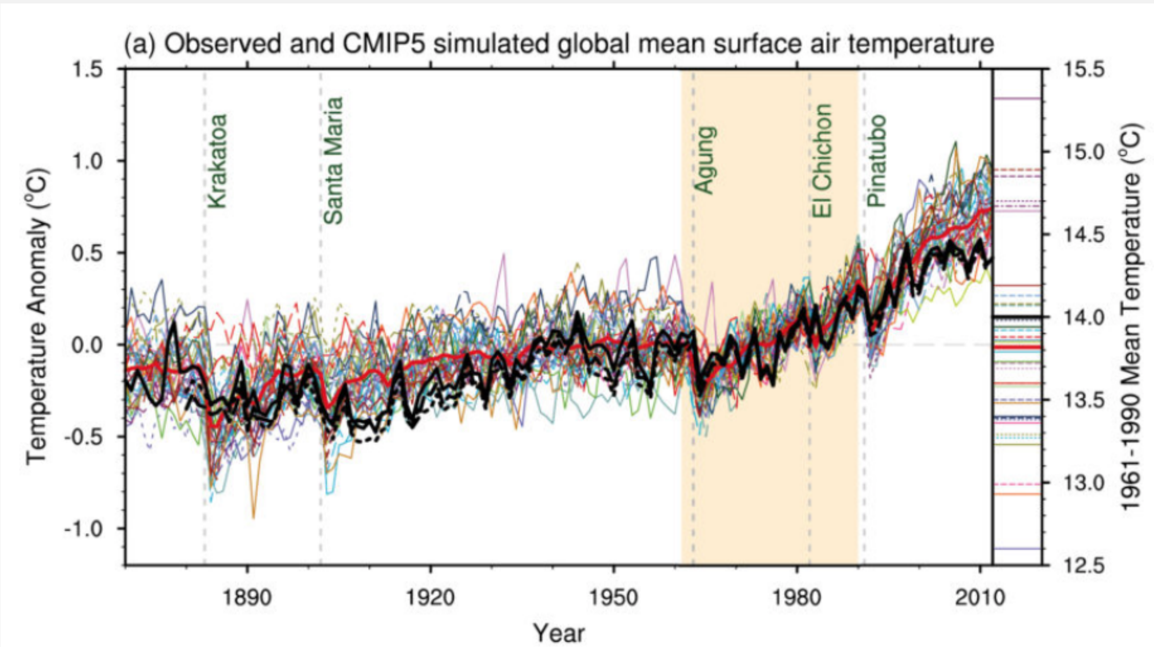
The freezing point of water, as far as I am aware, is hard-coded into all global climate models as zero degrees Celsius. But after many real-years of work, thousands of model-years of simulation and calibrating many other parameters, models still differ by a few degrees in their global mean temperature (IPCC fig 9.8, 2013). Most presentations subtract off the difference, implying that the net effects of all processes must be linear with regard to temperature, and effectively varying the freezing point of water by a couple of degrees Celsius.

2.2.3 Model world -> real world (eg implied volatility, climate sensitivity)

At other times, the relationship goes the other way: quantities which have been given a definition or meaning solely as a result of their usefulness in the model world become applied to the real world as well.

In climate research, one example is the "climate sensitivity": the amount by which the global mean temperature increases given a doubling in the radiative forcing applied. The concept is already theory-laden: it is assumed that the relationship is logarithmic and only the constant of proportionality need be determined. Climate sensitivity allows the comparison of multiple models across a limited range of hypothetical scenarios. In the real world, while there is indeed a question of interest regarding to what extent the global mean temperature will increase given a doubling in forcing, it could be argued that this measure is given undue weight relative to other quantities which could be more policy-relevant but are less natural measurables of the model

Climate models have different global mean temperatures **Box 2.2**



Global mean temperature as simulated by 36 global climate models (thin coloured lines), the average of those 36 (thicker red line) and observations (thicker black lines), normalised by subtracting the 1961-1990 mean from each time series (ie, they all have an average of zero over 1961-1990). The right hand axis, which is not quite to the same scale as the other, shows the absolute value of the 1961-1990 mean temperature for each of the models: the amount that was subtracted in the normalisation. These differ by nearly 3°C.

world. Climate sensitivity is also insensitive to the systematic differences in global mean temperature mentioned above, and it sidesteps discussions of potential nonlinear feedbacks at higher forcings.

A second example is from economics: the “implied volatility” of an option is the value of the “volatility” parameter which would result in the observed market price of that option, given a certain pricing model. When first defined, this was a derived property believed to elicit information about traders’ views of the future volatility of an instrument. Now, the implied volatility, another extremely theory-laden concept, has taken on a life of its own as an entity about which people have opinions.

2.2.4 No true parameter set exists

2.3 Complexity. Subtractability. Noise/music.

2.4 Dynamical (process-based) models versus statistical models

2.4.1 ?explainability? and interpretability

2.4.2 Horses and Hawkmoths

2.4.3 Note on ?machine learning??

Chapter 3

Butterflies and Hawkmoths

Chapter overview

In this chapter I...

3.1 Introduction **14**
3.2 Butterfly Effect. Initial condition ensembles. Solved. **14**
3.3 Hawkmoth Effect. Model structure ensembles? Unsolved. **14**
3.4 Ensembles **14**
3.4.1 Independence?
3.4.2 Model groupthink
3.4.3 Comment on ?sensitivity analysis?
3.4.4 So what does the ensemble range mean?
3.4.5 A model asylum
3.5 Consequences of the Hawkmoth Effect **15**
3.5.1 Things which are not included may not be negligible
3.5.2 Models cannot be monotonically improved
3.6 Mathematical treatment (easily skippable!) **15**

3.1 Introduction

3.2 Butterfly Effect. Initial condition ensembles. Solved.

3.3 Hawkmoth Effect. Model structure ensembles? Unsolved.

3.4 Ensembles

3.4.1 Independence?

3.4.2 Model groupthink

3.4.3 Comment on ?sensitivity analysis?

3.4.4 So what does the ensemble range mean?

3.4.5 A model asylum

3.5 Consequences of the Hawkmoth Effect

3.5.1 Things which are not included may not be negligible

3.5.2 Models cannot be monotonically improved

3.6 Mathematical treatment (easily skippable!)

Andronov/Pontryagin 1937, Anosov, etc. Shadowing timescales and numerical approximations. Solution of PDEs. Ergodicity, existence/timescales of attractors, ?statistical stability??

Chapter 4

The Cat that looks most like a Dog

In this chapter...

4.1 The cat that looks most like a dog **19**
4.2 On Bayesian methods **19**
4.2.1 Bayes theorem does not apply
4.3 Methodology is epistemology **19**
4.4 Many other statstical methods have the same underlying problem (a reference class problem) **19**

4.1 The cat that looks most like a dog

4.2 On Bayesian methods

4.2.1 Bayes theorem does not apply

4.3 Methodology is epistemology

4.4 Many other statstical methods have the same underlying problem (a reference class problem)

Chapter 5

Escaping from Model Land

Chapter overview

In this chapter...

5.1 Statements made about real world need to refer to the real world **22**
5.2 Evaluation against past data **24**
5.2.1 Evaluation metrics (eg do we require propriety if models are never perfect?)
5.2.2 Big Data
5.2.3 Small Data (accounting for small-number problems)
5.2.4 No data? The reference class problem
5.3 Use of expert judgement **24**

5.1 Statements made about real world need to refer to the real world

Disentangling the meaning of statements made using models about the real world can be difficult. Compare the following three key findings from the Summary for Policymakers of the most recent report from the Intergovernmental Panel on Climate Change. All of these key findings from the same short document refer to conclusions informed by output from the same models, and they are all presented as real-world statements which are relevant for policymakers. Yet they take quite different approaches to reporting model output.

1. **Temperature projections** give a 5-95% model range and downgrade the uncertainty assessment "after accounting for additional uncertainties or different levels of confidence in models". This is a meaningful statement about the authors' expectations of a real future climate (given the hypothetical changes in forcing):

> *Increase of global mean surface temperatures for 2081-2100 relative to 1986-2005 is projected to likely be in the ranges derived from the concentration-driven CMIP5 model simulations, that is, 0.3C to 1.7C (RCP2.6), 1.1C to 2.6C (RCP4.5), 1.4C to 3.1C (RCP6.0), 2.6C to 4.8C (RCP8.5).* SPM E.1, page 20

Here, **model output has been used to inform expert opinion**, with expert judgement used to quantify the plausible chance of non-modelled outcomes occurring.

2. **AMOC projections** give a model spread and best estimate, attaching no likelihood to the numerical values reported from models but a very likely assessment of the sign of the change (not conditional on forcing). The multi-model mean is explicitly identified as a best estimate:

> *It is very likely that the Atlantic Meridional Overturning Circulation (AMOC) will weaken over the 21st century. Best estimates and ranges [footnote: The*

> *ranges in this paragraph indicate a CMIP5 model spread.] for the reduction are 11% (1 to 24%) in RCP2.6 and 34% (12 to 54%) in RCP8.5. SPM E.4, page 24*

Here, **model output has been presented as a proxy for expert opinion**. The unstated assumption is that the model average is the best estimate. It would be helpful for the authors to clarify whether the direct use of model output as a best estimate is because it is genuinely thought that the models capture all relevant processes, or because there is no extant information on which to base further expert judgement about the inadequacies of the models.

3. **Arctic sea ice projections** give a range from the lowest to highest RCP scenarios, reporting only the model outcomes but then attaching a medium confidence label ? when surely, this should be extremely high confidence that the models give these reductions. There is an unstated implication that we have medium confidence of the real-world outcome lying between these bounds:

> *Year-round reductions in Arctic sea ice extent are projected by the end of the 21st century from multi-model averages. These reductions range from 43% for RCP2.6 to 94% for RCP8.5 in September and from 8% for RCP2.6 to 34% for RCP8.5 in February (medium confidence)* SPM E.5, page 24

Here, **model output has been presented as a proxy for expert opinion**, but in addition a confidence assessment of only "medium" has been used to weaken the statement without giving any information, either qualitative or quantitative, about the degree to which it is weakened. There is no sense in which anyone could use this statement to generate a real-world expectation beyond the sign of the change.

These three ad-hoc methods of dealing with the difficult step of translating model-world values into something fit for a "Summary for Policy Makers" demonstrate the need for a more transparent guidance on the use and interpretation of model output. In my view the first is reasonably well-considered and documented, the second is disingenuous and requires explicit statement and clarification of the underlying assumption, and the third is positively misleading.

5.2 Evaluation against past data

5.2.1 Evaluation metrics (eg do we require propriety if models are never perfect?)

5.2.2 Big Data

5.2.3 Small Data (accounting for small-number problems)

5.2.4 No data? The reference class problem

5.3 Use of expert judgement

Chapter 6

Making decisions based on models

6.1 The aim: getting ?the right answer? or supporting a decision?

6.1.1 Properties of good decisions

6.1.2 Sometimes the skill of the model is unimportant (Mali farmers; Start; horoscopes; biodynamic planting?)

6.1.3 Sometimes the confidence, even where reported honestly, is recalibrated by the audience (?ninety per cent? =/= 90%)

6.2 Evaluation of models, evaluation of forecasts

Chapter 7

Conclusions and consequences

7.1 The unreasonable effectiveness of mathematics in the natural sciences?

7.2 Modelling for decision support

7.3 Alternative strategies, find positive/constructive note to end on.

Bibliography

[1] C. Schär, P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller. The role of increasing temperature variability in European summer heatwaves. *Nature*, 427:332–336, January 2004.