

On Hawkmoths and Horses: Epistemic issues in modelling complex systems



Erica Thompson

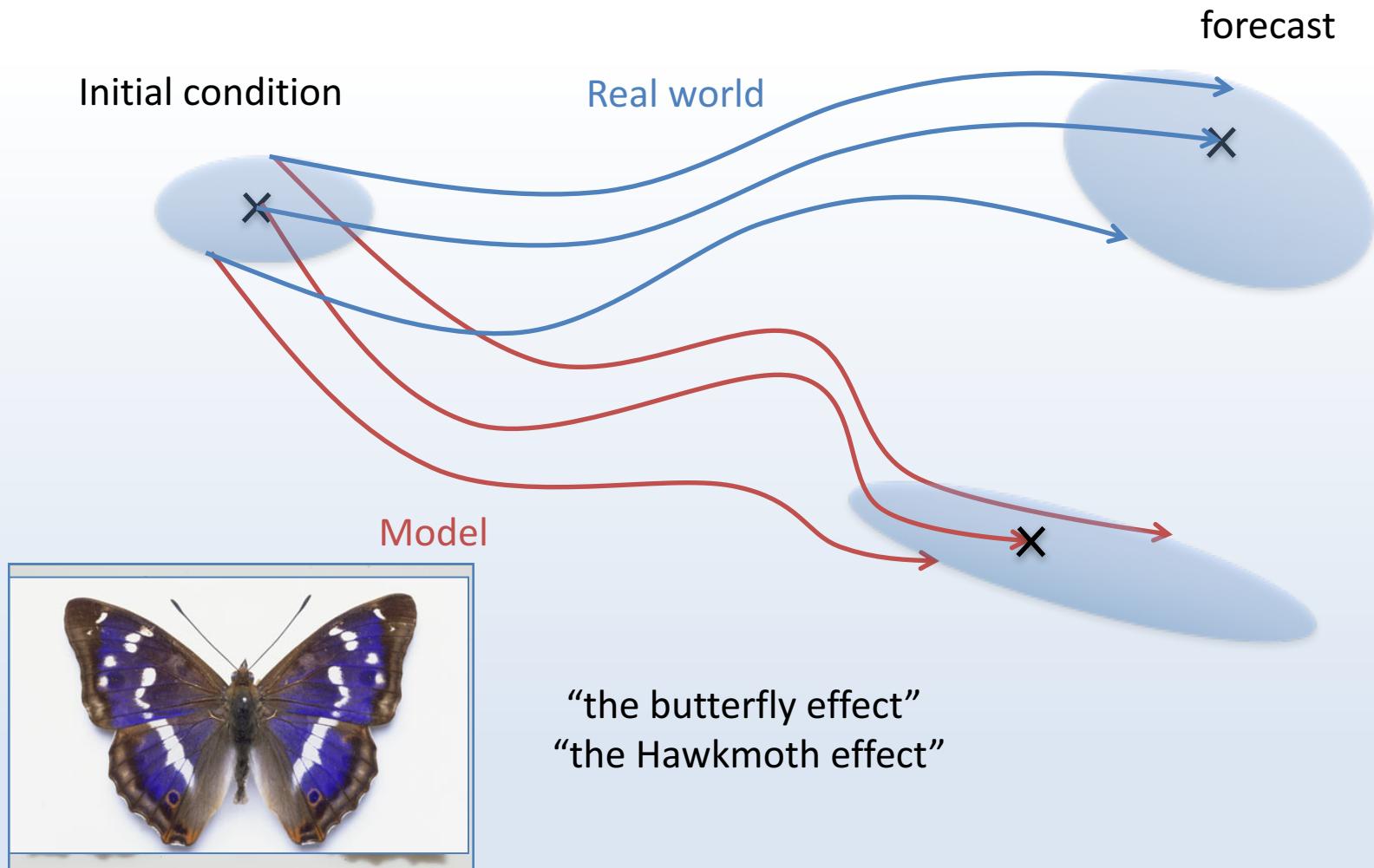
Centre for the Analysis of Time Series

London School of Economics



- What is the Hawkmoth Effect?
- What is a good model?
- Example: climate models
- $P(\text{Horse})$
- Are all models Horses?
- What does that mean?

The Hawkmoth Effect



The Hawkmoth Effect



- Your equations can be *arbitrarily close* to the “correct” equations, and still get completely the wrong answer

real world: $y = f(x)$

model: $y = f(x) + \varepsilon(x)$

- Trajectories (may) differ by *much more* than ε

What is a good model?



- Right answer?
- Right answer for the right reasons?
- What do we mean by the right reasons?
- Right dynamics?

	Right reasons	Wrong reasons
Right answer	Good model	Horse
Wrong answer	Unlucky?	Dead Horse

Model development:

bad → good → better → best ??

Example: climate models

- How would we know climate model is giving the “right answer”?

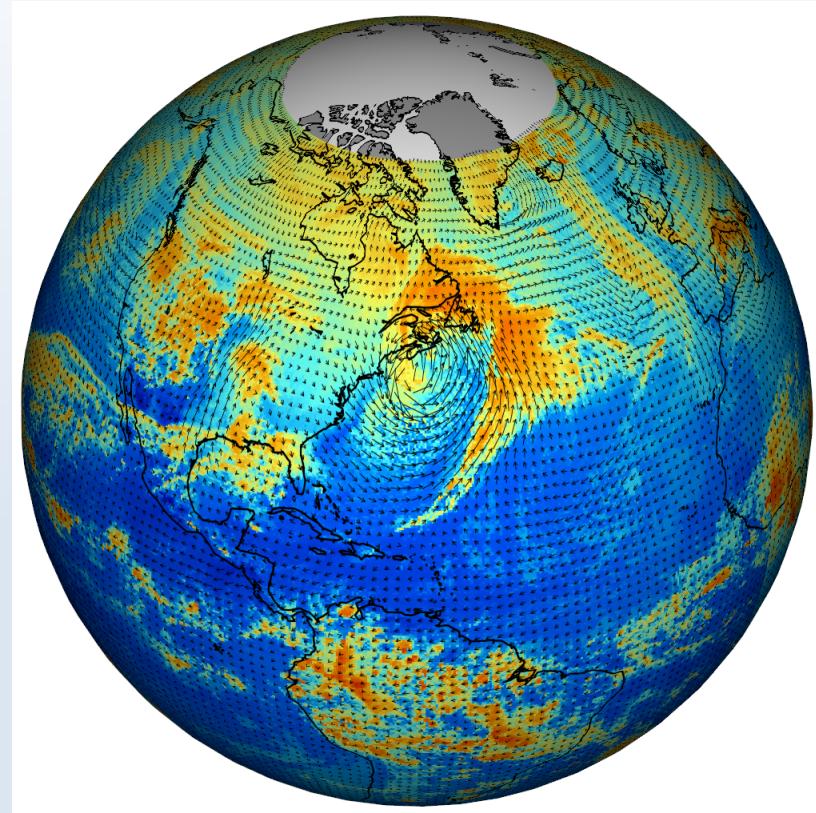
Agreement with other models?

Agreement with physical intuition?

Makes pretty pictures?

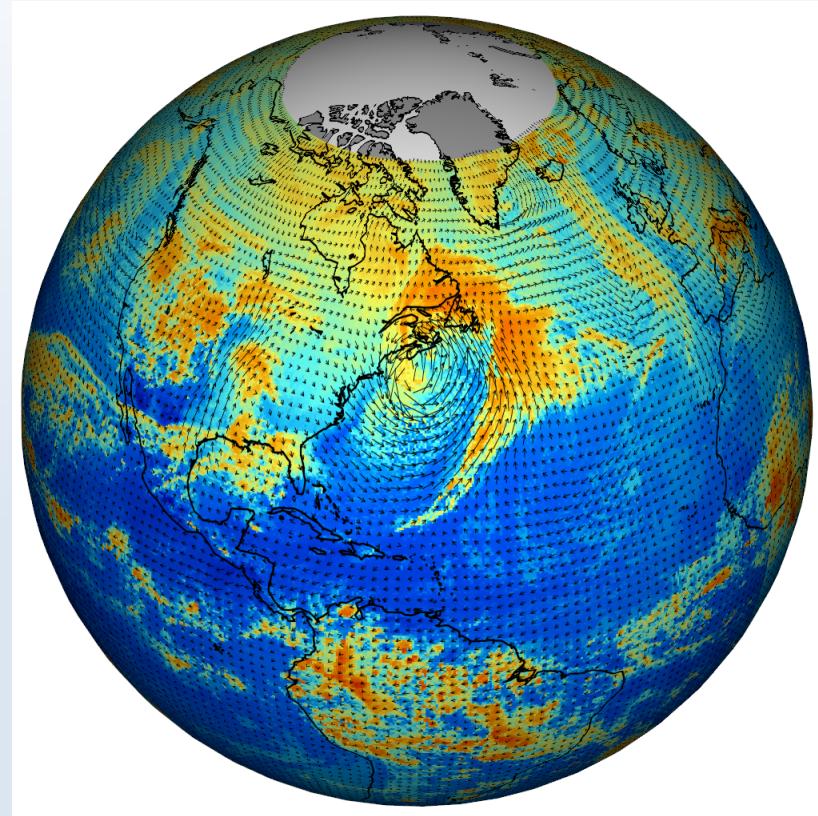
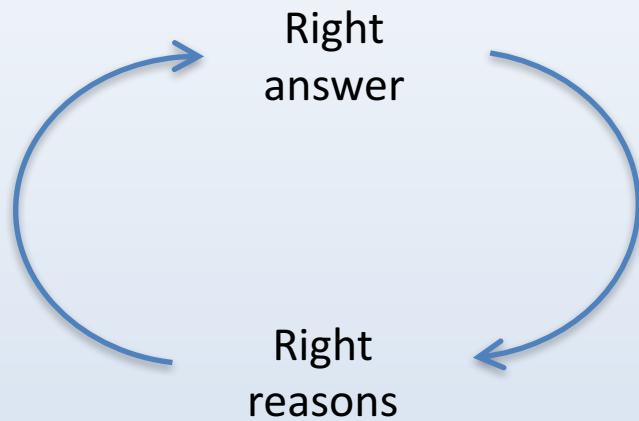
Looks plausible?

Out of sample: can't test it



Example: climate models

- How would we know climate model is giving the right answer for the “right reasons”?



P(Horse)

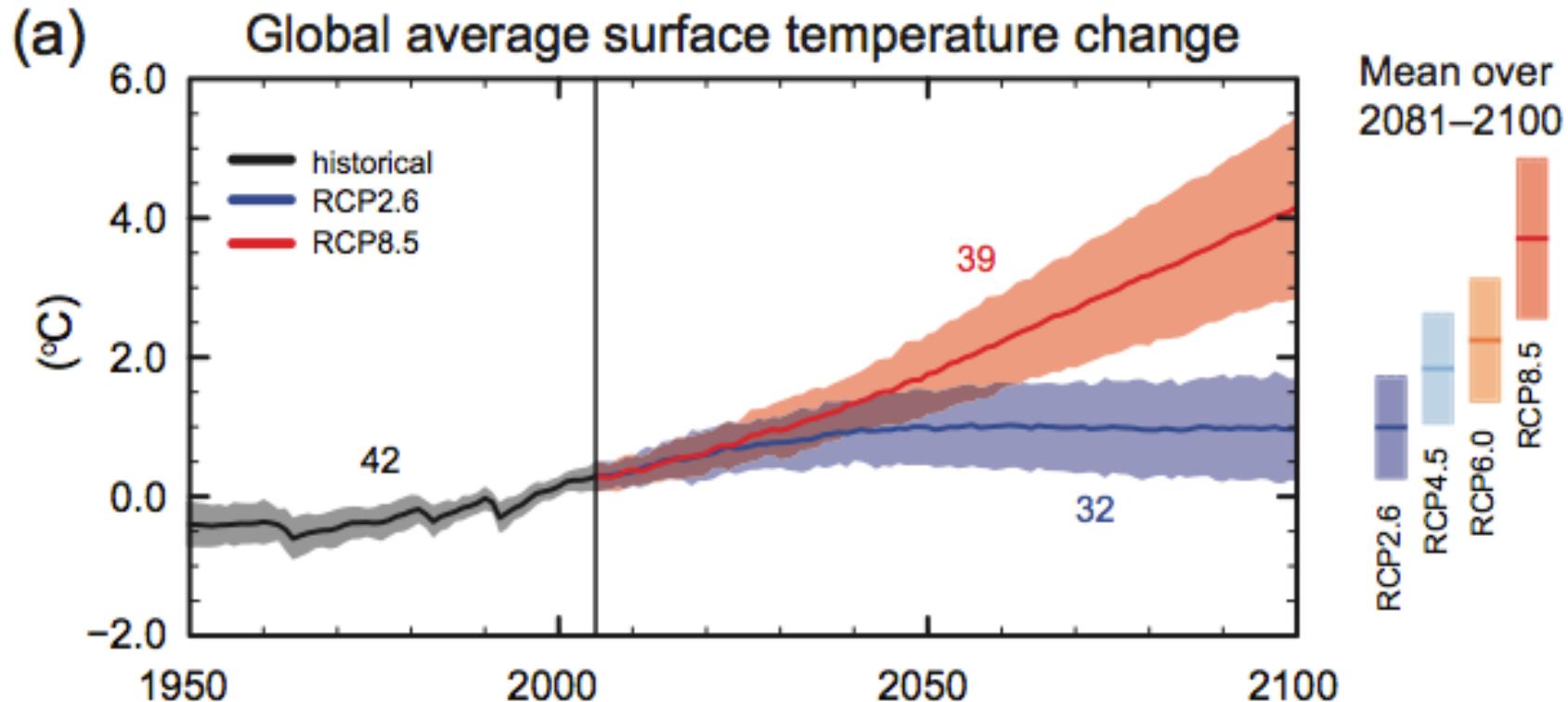


- The IPCC do effectively state the probability that climate models are Horses
- Anyone like to guess that probability?

24%

	Right reasons	Wrong reasons
Right answer	Good model	Horse
Wrong answer	Unlucky?	Dead Horse

IPCC projections



30+ big numerical models

All values displayed as change since baseline period 1986-2005, hence the scale

Range of model outcomes displayed as shaded area with mean as central projection

Projections

		2046–2065		2081–2100	
	Scenario	Mean	Likely range ^c	Mean	Likely range ^c
Global Mean Surface Temperature Change (°C)^a	RCP2.6	1.0	0.4 to 1.6	1.0	0.3 to 1.7
	RCP4.5	1.4	0.9 to 2.0	1.8	1.1 to 2.6
	RCP6.0	1.3	0.8 to 1.8	2.2	1.4 to 3.1
	RCP8.5	2.0	1.4 to 2.6	3.7	2.6 to 4.8
Global Mean Sea Level Rise (m)^b	Scenario	Mean	Likely range ^d	Mean	Likely range ^d
	RCP2.6	0.24	0.17 to 0.32	0.40	0.26 to 0.55
	RCP4.5	0.26	0.19 to 0.33	0.47	0.32 to 0.63
	RCP6.0	0.25	0.18 to 0.32	0.48	0.33 to 0.63
	RCP8.5	0.30	0.22 to 0.38	0.63	0.45 to 0.82

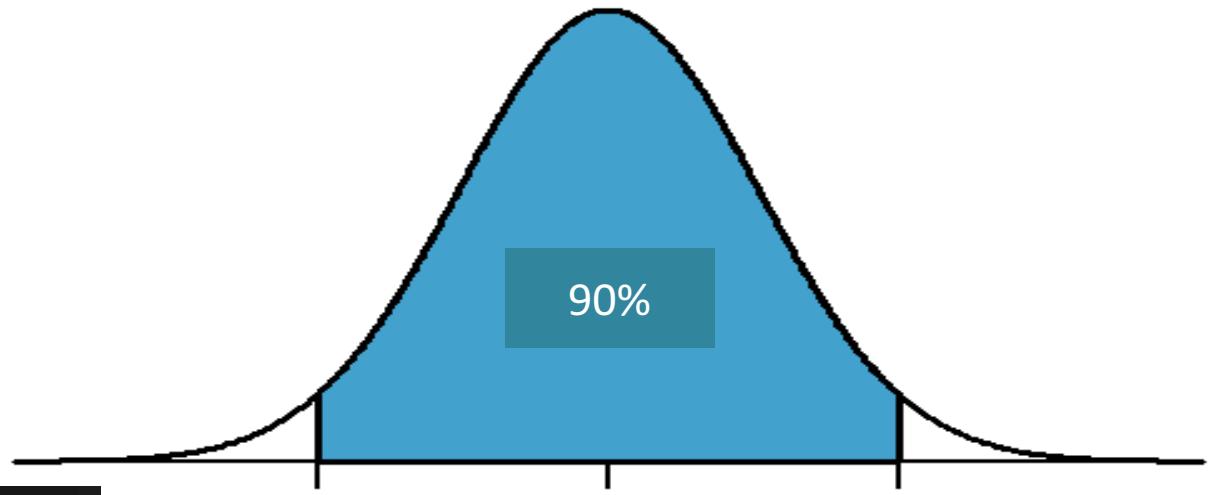
Notes c and d: “Calculated from projections as 5–95% model ranges. These ranges are then assessed to be *likely* ranges after accounting for additional uncertainties or different levels of confidence in models.”

Communication

Table 1. Likelihood Scale

Term*	Likelihood of the Outcome
<i>Virtually certain</i>	99-100% probability
<i>Very likely</i>	90-100% probability
<i>Likely</i>	66-100% probability
<i>About as likely as not</i>	33 to 66% probability
<i>Unlikely</i>	0-33% probability
<i>Very unlikely</i>	0-10% probability
<i>Exceptionally unlikely</i>	0-1% probability

Confidence in models



90%: "Very likely"



66%: "Likely"

90% model range



"likely" assessed range

Are all models Horses?



Two questions

- How do we know that all scientific models are not Horses?
 - Out of sample testing!
- If you have a model, and you test it out of sample, **and it works**, does a bit of equine DNA matter anyway?
 - Statistical models vs dynamical models
 - Perhaps it is still abstracting something informative

What does that mean?



- How can we know when models that cannot be tested out of sample are not Horses?
 - We can't know
 - In some sense they must all be Horses, as they are all prone to Hawkmoths and therefore can never be using “the right reasons”
 - But we can still make useful expert judgements about the quality of predictions made by the models



THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■



Erica Thompson

e.thompson@lse.ac.uk

www.ericathompson.co.uk