Abstract: Queuing the Wrong U

Just as there are many different types of uncertainty, there are many different types of models. The best technique for quantifying and communicating uncertainty will depend on the nature of that uncertainty: is it mere imprecision in a well-defined number (as with the square-root of two), intractability (as when we know how to compute the answer, but have not yet been able to carry out the calculation), indeterminacy (as when there is no well-defined target about which to be imprecise) or other. The relevance of UQ to a decision maker or scientist will also depend on the type of quantitative model that is considered: is the model intended to explain, or to forecast, or to provide a quantitative analysis of the past?

When a perfect model is available, many of these distinctions collapse. In practice, attempting to quantify one type of uncertainty via a model which may not even display that kind of uncertainty is a nonsense. One must be careful not to confuse the diversity of our models for the uncertainty in our future. Or a well-defined probability forecast for what the next model simulation will report, with a probability forecast for the world. How is UQ to recognize the line between sensitivity analysis and probability forecasting? These questions will be addressed in the context of climate science, and more broadly that of science in support of decision making. The ways and means of UQ are shown to vary with type of model considered, the extent to which that model class is deemed adequate for purpose in a specific application, and whether or not the relevant dominant uncertainty (known from the science, but perhaps absent from the models) has been considered. Uncertainty Quantification may prove to be a very wide field, extending well beyond the bounds of the probability calculus.

LSE CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

www.lsecats.ac.uk

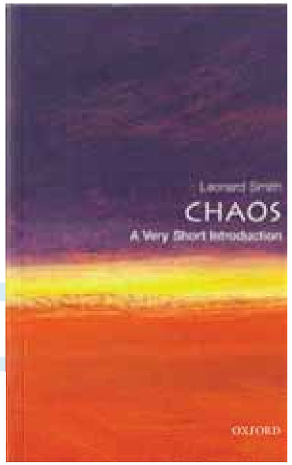
Queuing the Wrong U

Leonard A. Smith
London School of Economics &
Pembroke College, Oxford

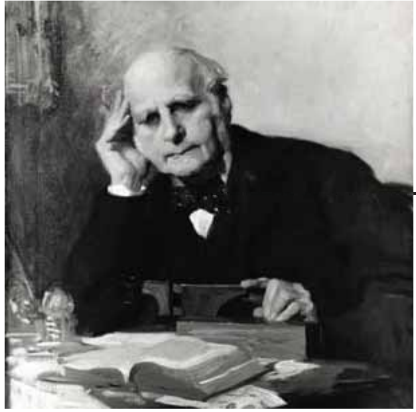
EQUIP

Not Possible without: H Du, D Stainforth & E Suckling

Centre for
Climate Change
Economics and Policy
The Munich Re Programme: *Evaluating the Economics of Climate Risks and Opportunities in the Insurance Sector*



Grantham Research Institute on
Climate Change and
the Environment



Distinguishing Types of Uncertainty

There are many different types of uncertainty.

Best practice requires UQ communicate the RDU (the dominant uncertainty relevant to the question at hand) **not** merely the easiest U to quantify!

Imprecision: There is a well defined value, it is merely unknown:

314th digit of the √2

Intractability: The target may be well-defined, but using/determining it would prevent completion of the required calculations.

Height of the Andes in a GCM

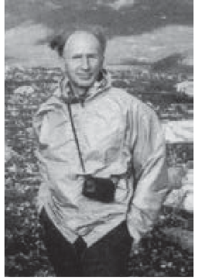
Indeterminacy: There is no true value of be uncertain of.

“The radius” of the Earth (as it not a sphere!)

The value of Newton’s gravitational constant G

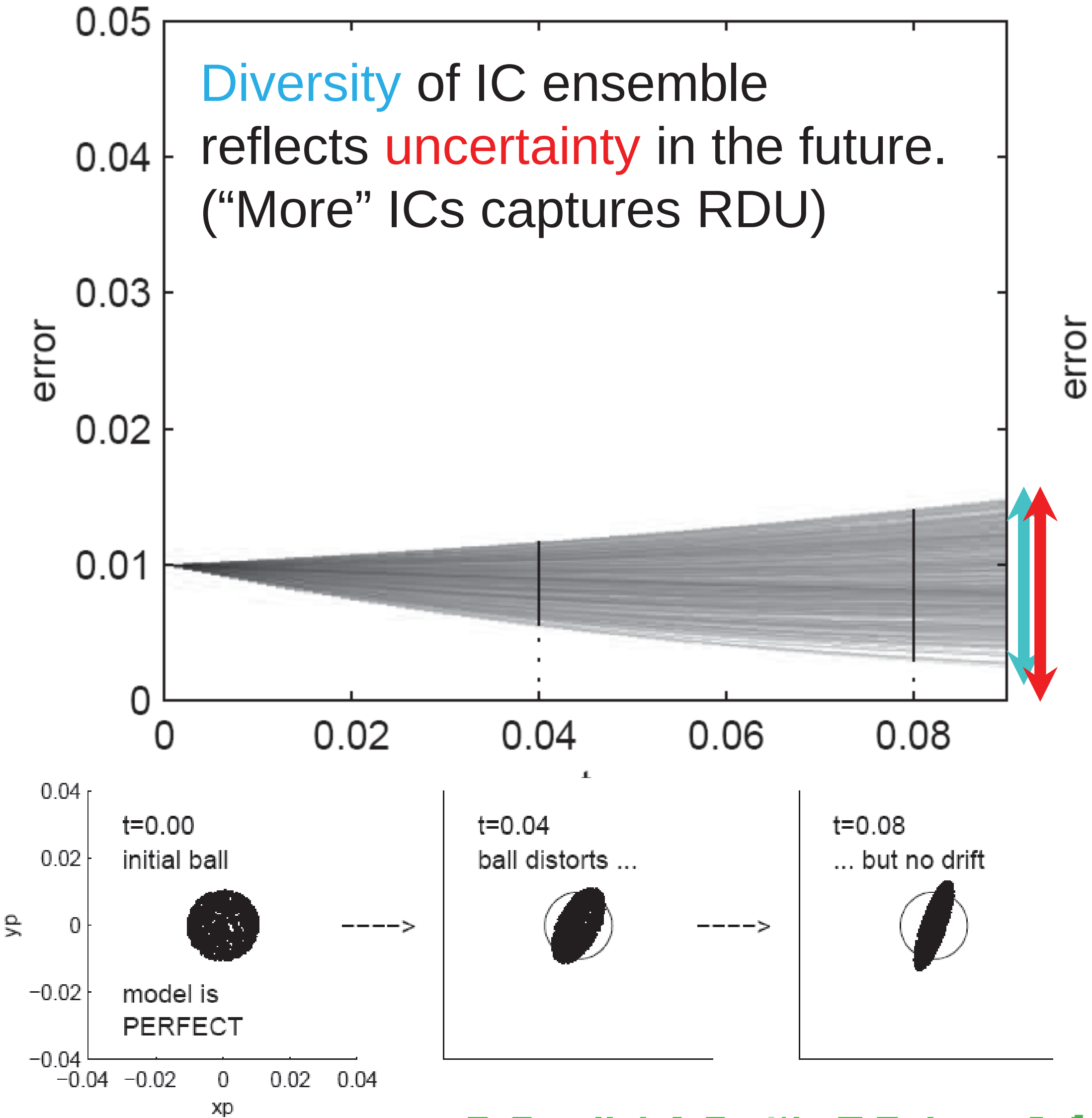
To be of value in application UQ must move well beyond “imprecision”

Relevant Dominant Uncertainty

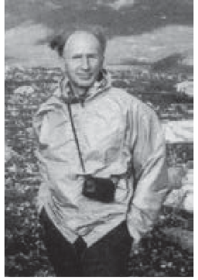


– Initial Condition Uncertainty

(RDU is in IC)

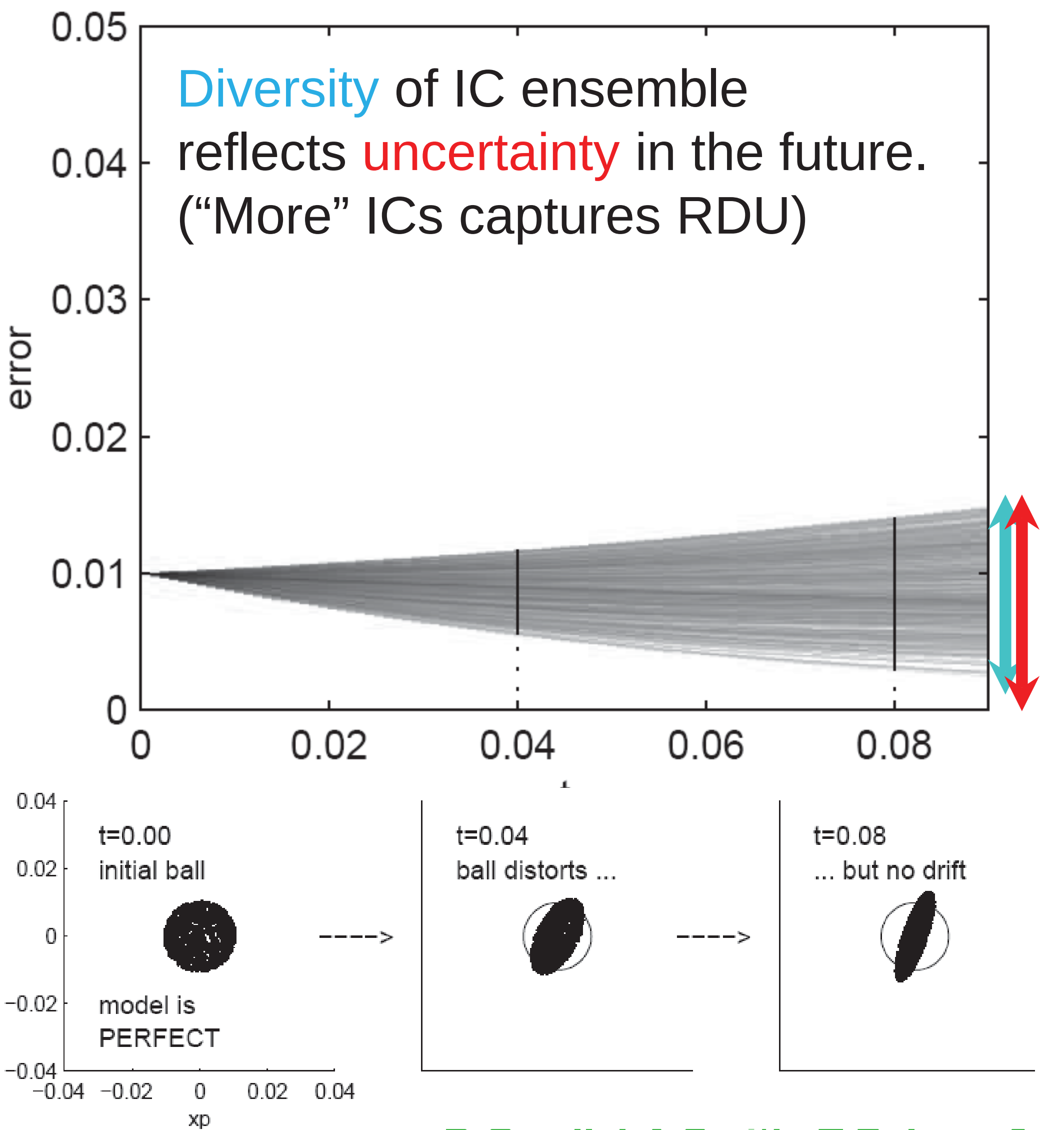


D Orrell, LA Smith, T Palmer & J Barkmeijer (2001) Model Error in Weather Forecasting, Nonlinear Processes in Geophysics 8: 357-371.

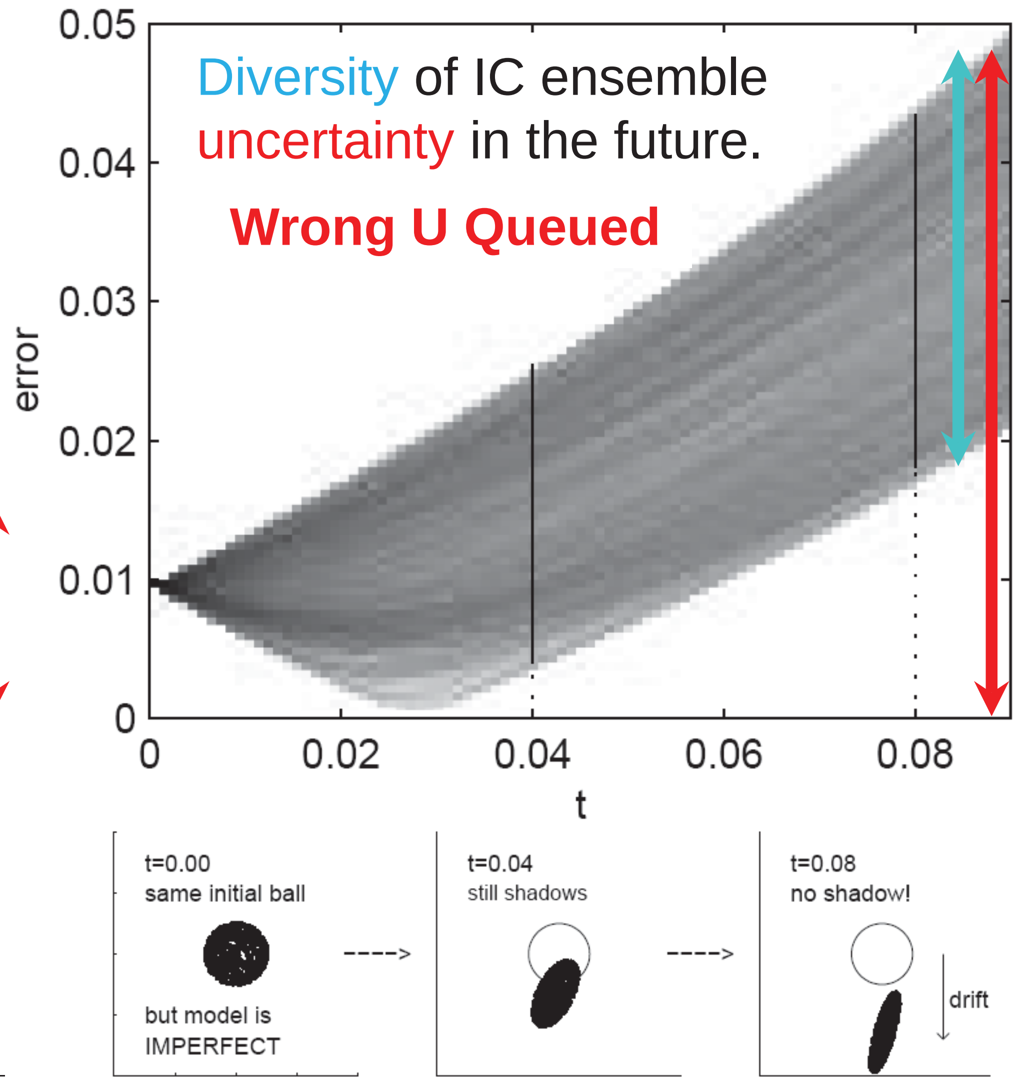


Relevant Dominant Uncertainty

Initial Condition Uncertainty
(RDU is in IC)



Structural Uncertainty
(RDU is Model Inadequacy)



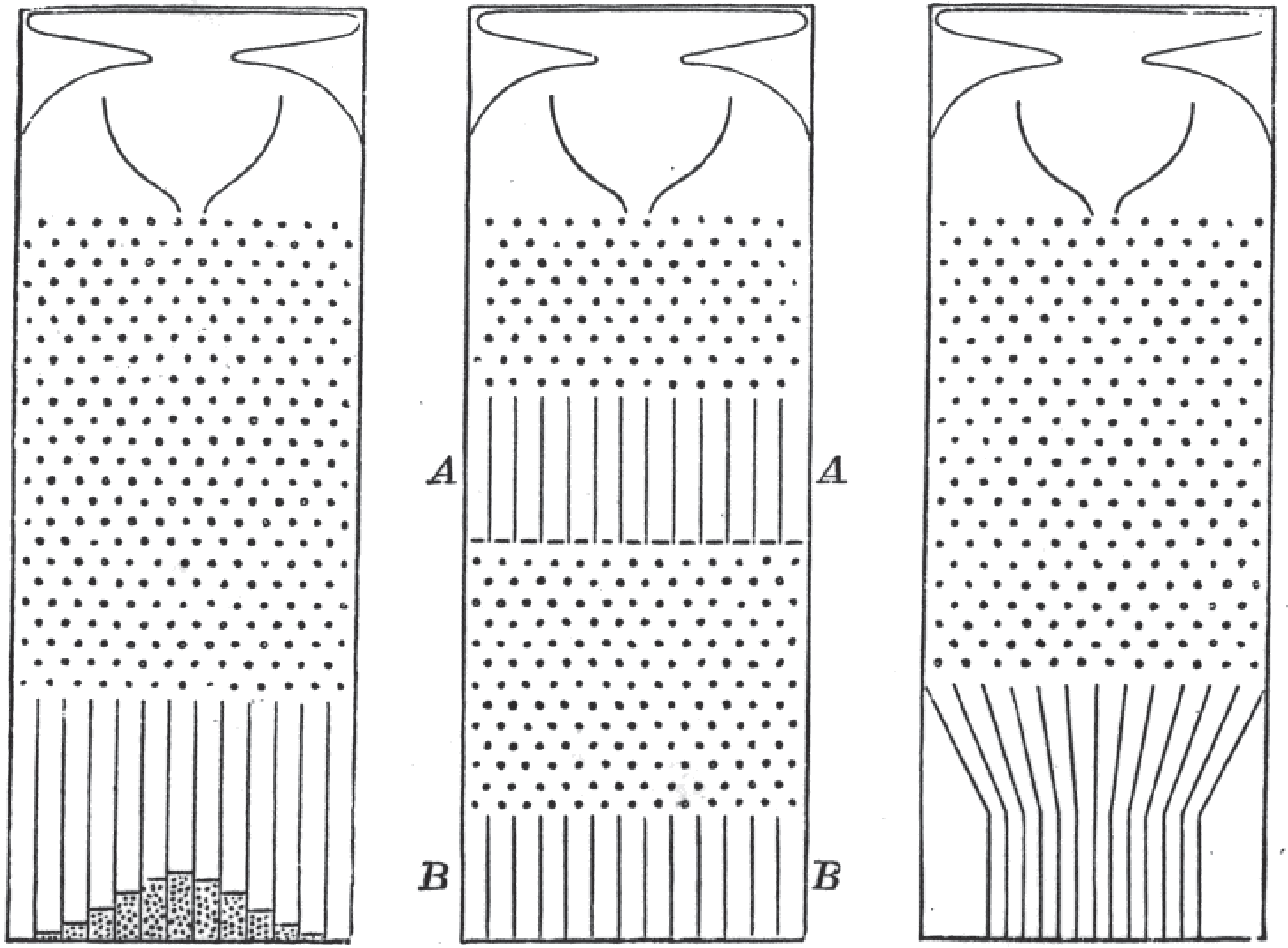
D Orrell, LA Smith, T Palmer & J Barkmeijer (2001) Model Error in Weather Forecasting, Nonlinear Processes in Geophysics 8: 357-371.

CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

The Galton Board (Galton 1889)

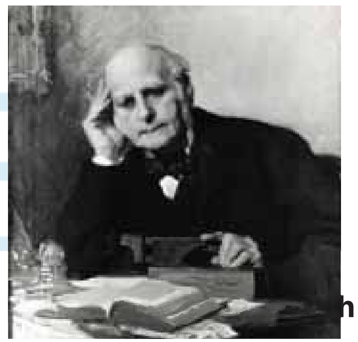
(quincunx)

FIG .7. FIG .8. FIG .9.



Each pellet has a 50/50 chance of going to the right (left) of each nail.

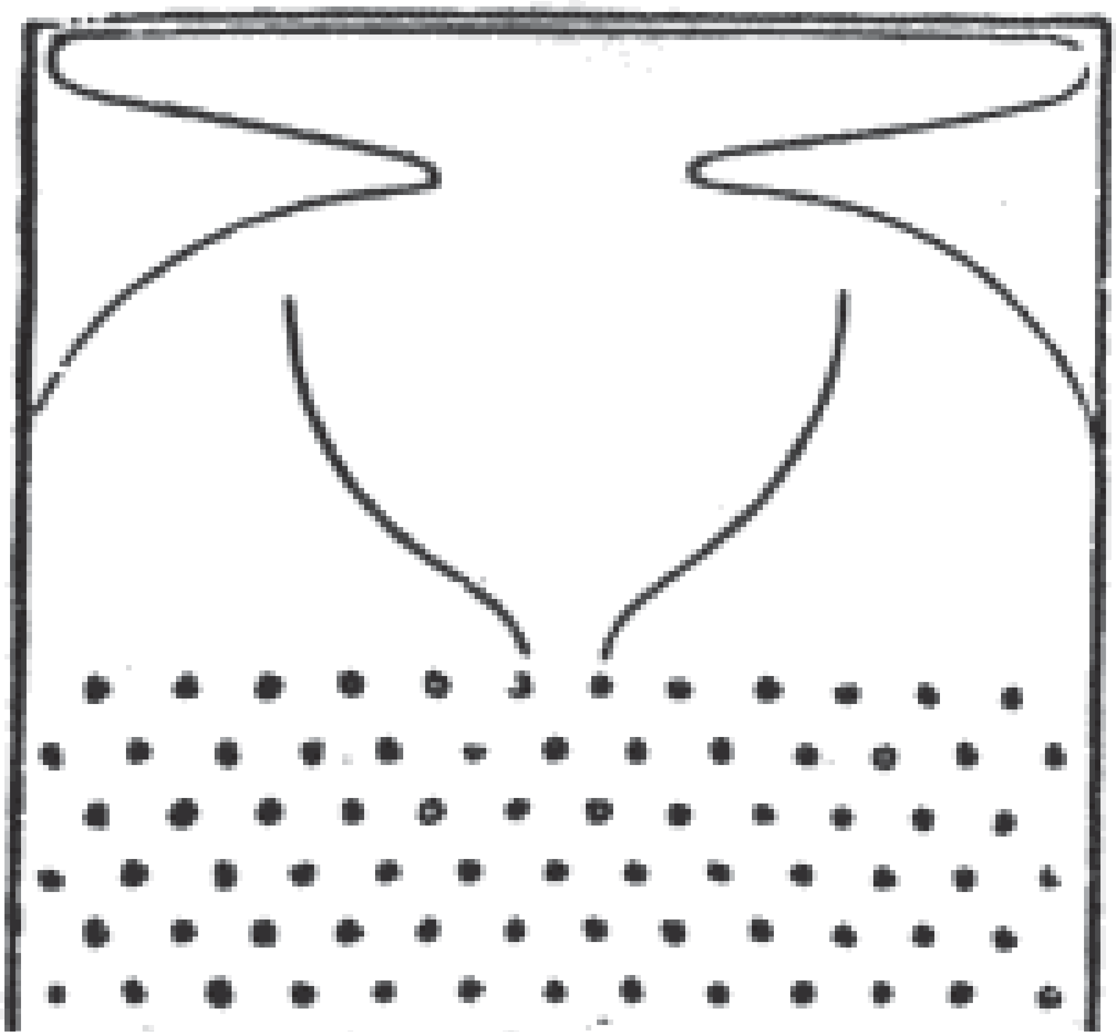
A mathematical result which is easier to match if you pour the shot in all at the same time…



The Galton Board (Galton 1889)

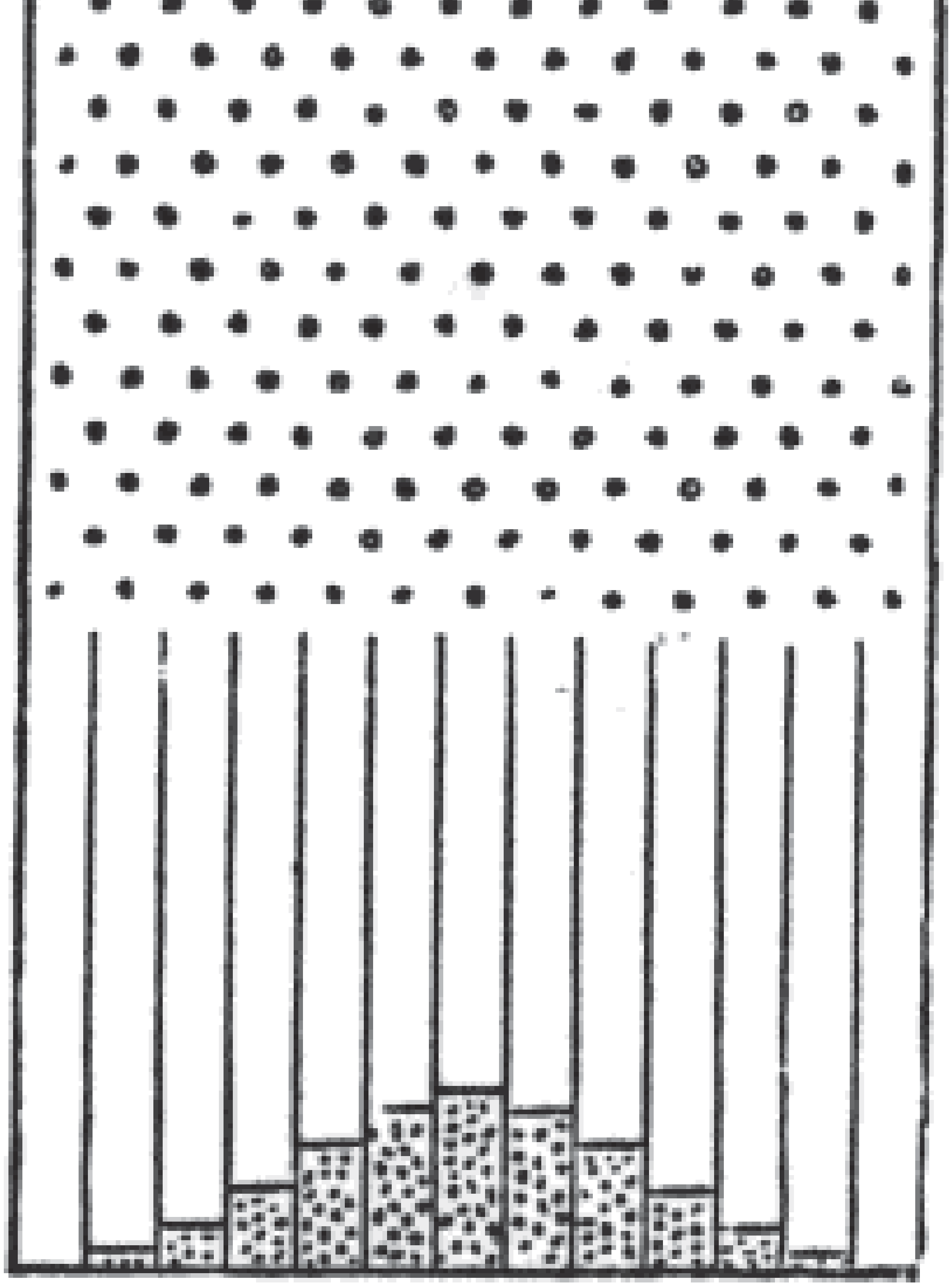
(quincunx)

FIG. 7.



Each pellet has a 50/50 chance of going to the right (left) of each nail.

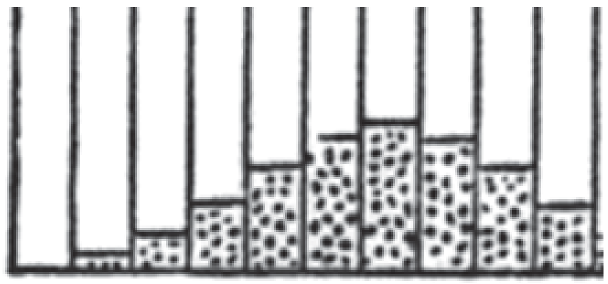
A mathematical result which is easier to match if you pour the shot in all at the same time…



Each pellet has a 50/50 chance of going to the right (left) of each nail.

A mathematical result which is easier to match if you pour the shot in all at the same time…

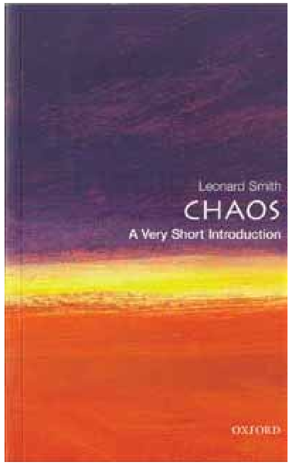
But would any randomness remain if we knew exactly where each pellet started?



The NAG Board

Not A Galton Board (2000)

Constructed for the 150th Birthday of the RMS



A given golf ball does **not** have a 50/50 chance of going to the right (left) of the next nail.

LA Smith (2007) Chaos A Very Short Introduction. OUP.

Quantifying Uncertainty in the Initial Condition

In the NAG board, UQ due to IC uncertainty corresponds to considering a collection (ensemble) of golf balls…

IC UQ via ensembles informs us of uncertainty within our model!

If ICU is not the RDU, then this is merely a sensitivity analysis of the model, and need not quantify the uncertainty in our future.

The uncertainty in the next golf ball is quantified...

Model Diversity does not Q the relevant U

Ensembles inform us of uncertainty growth within our model(s)!

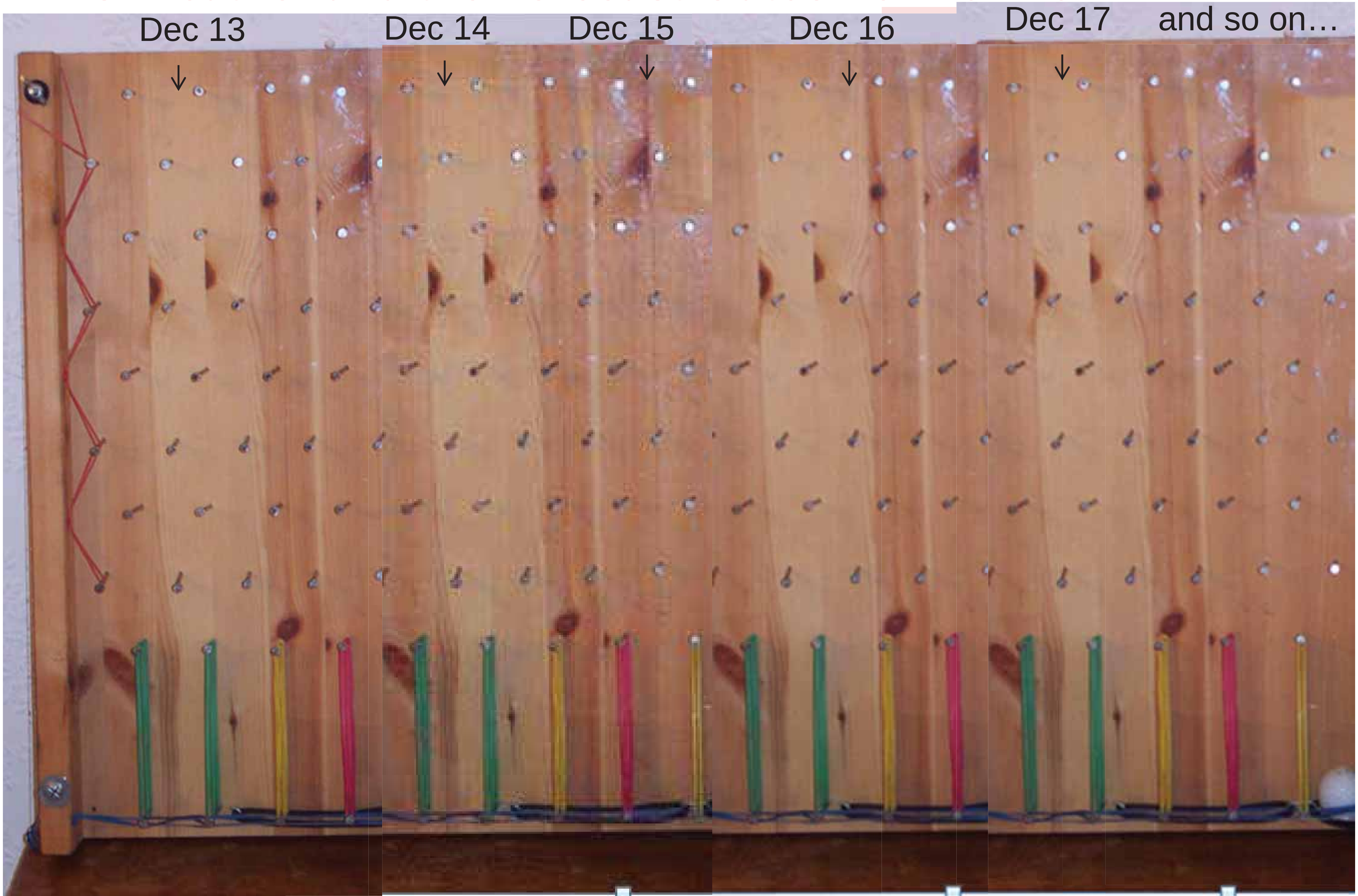
But ***reality*** is not a golf ball…

… ***reality*** is a red rubber ball.

What exactly does the distribution of 1024 golf balls tell us about where the one (and only) red rubber ball falls?

While we never see similar initial states, we can still learn from our mistakes!(in this <u>weather-like</u> case)

NAG Weather and the Forecast-Outcome Archive



Quantifying Uncertainty in the Initial Condition

In the NAG board, UQ due to IC uncertainty corresponds to considering a collection (ensemble) of golf balls…

The uncertainty in the next golf ball is quantified, but there comes a point where better quantifying that uncertainty does not better quantify the uncertainty most relevant to the target: the red ball.

Science can anticipate RDU surprises beyond model-land

Interpreting even weather-like distributions is a challenge!

J Bröcker & LA Smith (2008) From Ensemble Forecasts to Predictive Distribution Functions *Tellus A* 60(4): 663 D

Scientific speculation can help:
rubber balls will bounce more, will carry less angular momentum, will…
These scientific insights can qualitatively inform the use of the quantitative golf ball distributions.

Science can anticipate RDU surprises beyond model-land

Climate predictions require extrapolating out of the observed golf ball-rubber ball archive: into the known-to-be-different (?perhaps still fluid?) unknown.

The best we can hope for is sensible consistency in distribution between our models ("the details do not matter").

Scientific insight/speculation can anticipate "**Big Surprises**" (things our models cannot do)

Probabilistic Forecasts: IPCC Sixty-Forty Rule

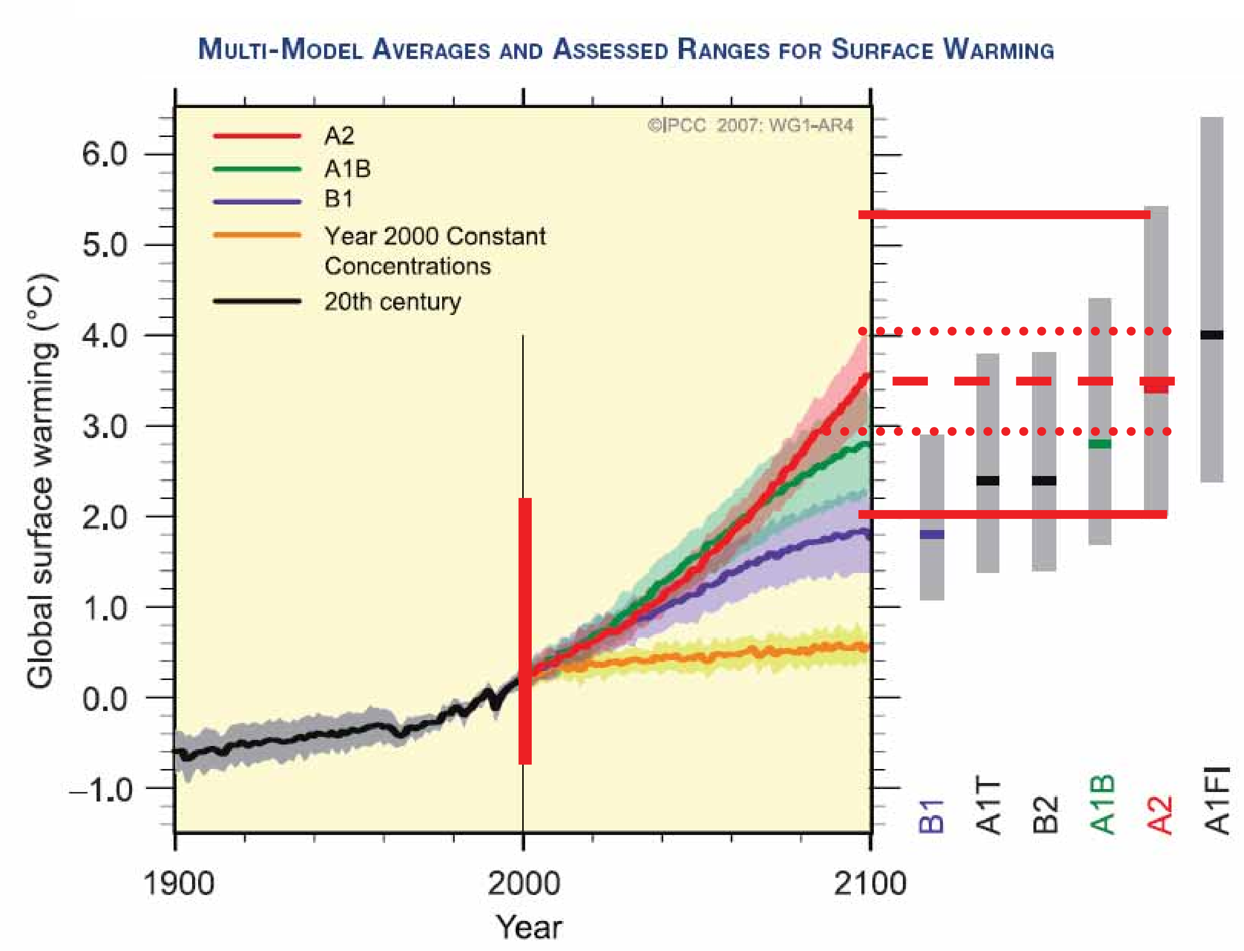
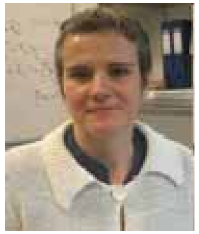


Figure SPM.5. *Solid lines are multi-model global averages of surface warming (relative to 1980–1999) for the scenarios A2, A1B and B1, shown as continuations of the 20th century simulations. Shading denotes the ±1 standard deviation range of individual model annual averages. The orange line is for the experiment where concentrations were held constant at year 2000 values. The grey bars at right indicate the best estimate (solid line within each bar) and the* ***likely*** *range assessed for the six SRES marker scenarios. The assessment of the best estimate and* ***likely*** *ranges in the grey bars includes the AOGCMs in the left part of the figure, as well as results from a hierarchy of independent models and observational constraints. {Figures 10.4 and 10.29}*

– The conditional forecasts (projections) are the grey bars (right); they differ from the ensemble distributions left and centre.



Of course, an adequately parameterised process might significantly shift the range.

Discussions of "broadening" imply confidence the RDU will not shift in the location.

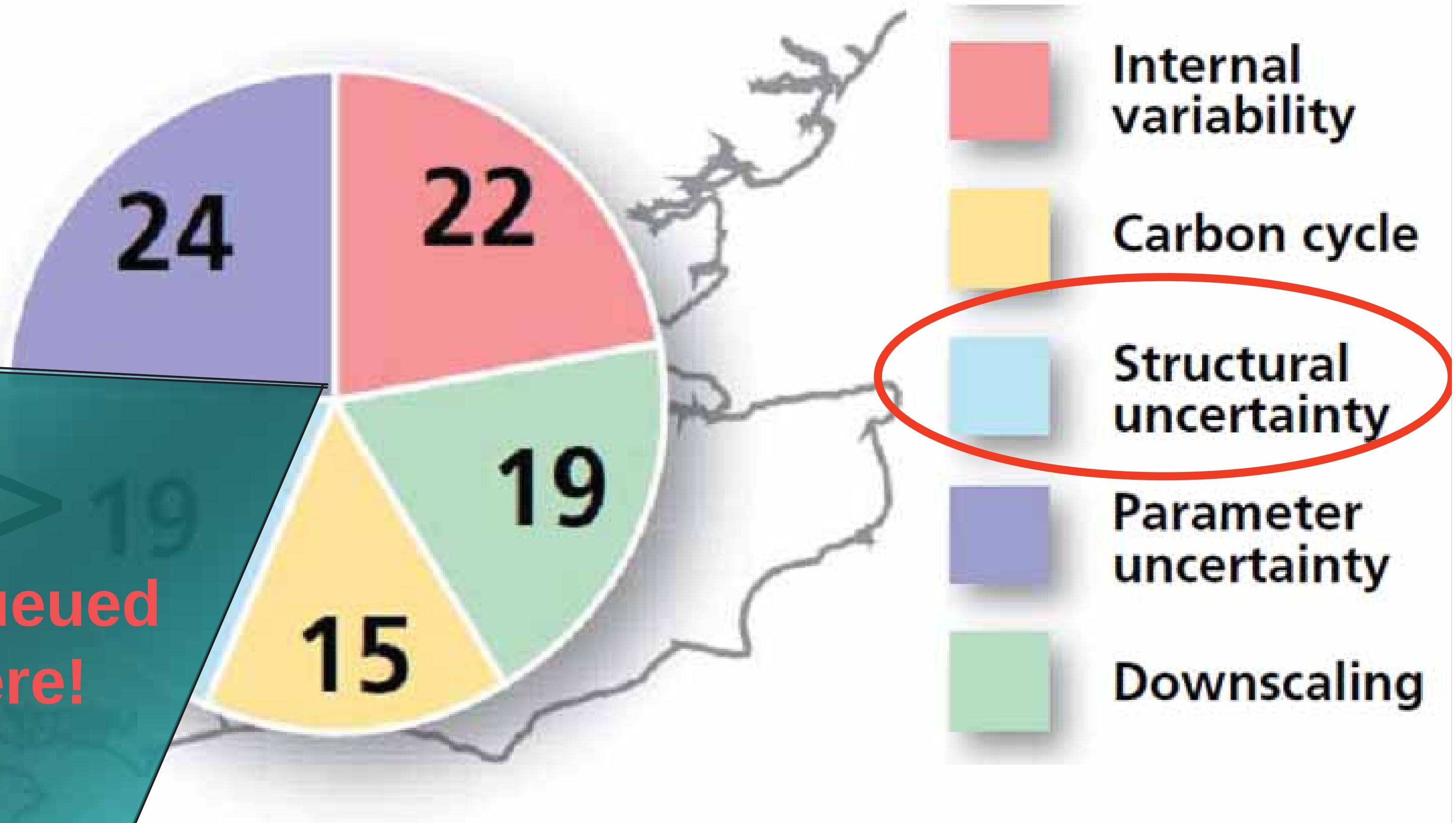
Do we have that confidence Globally? Regionally?

The IPCC rejects the claim that diversity of ensembles directly reflects our uncertainty in GMT.

How important are different sources of uncertainty?

The value of qualitative insight is at risk of being discarded in favour of quantitative mis-information…

- Varies, but typically no single source dominates.



>> 19

…if we have Queued the Wrong U here!

Uncertainties in winter precipitation changes for the 2080s relative to 1961-90, at a 25km box in SE England

Source: Met Office

State-of-the-art GCMs share deficiencies due to technology

Science is more than simulations

Missing 2km tall walls of rock!

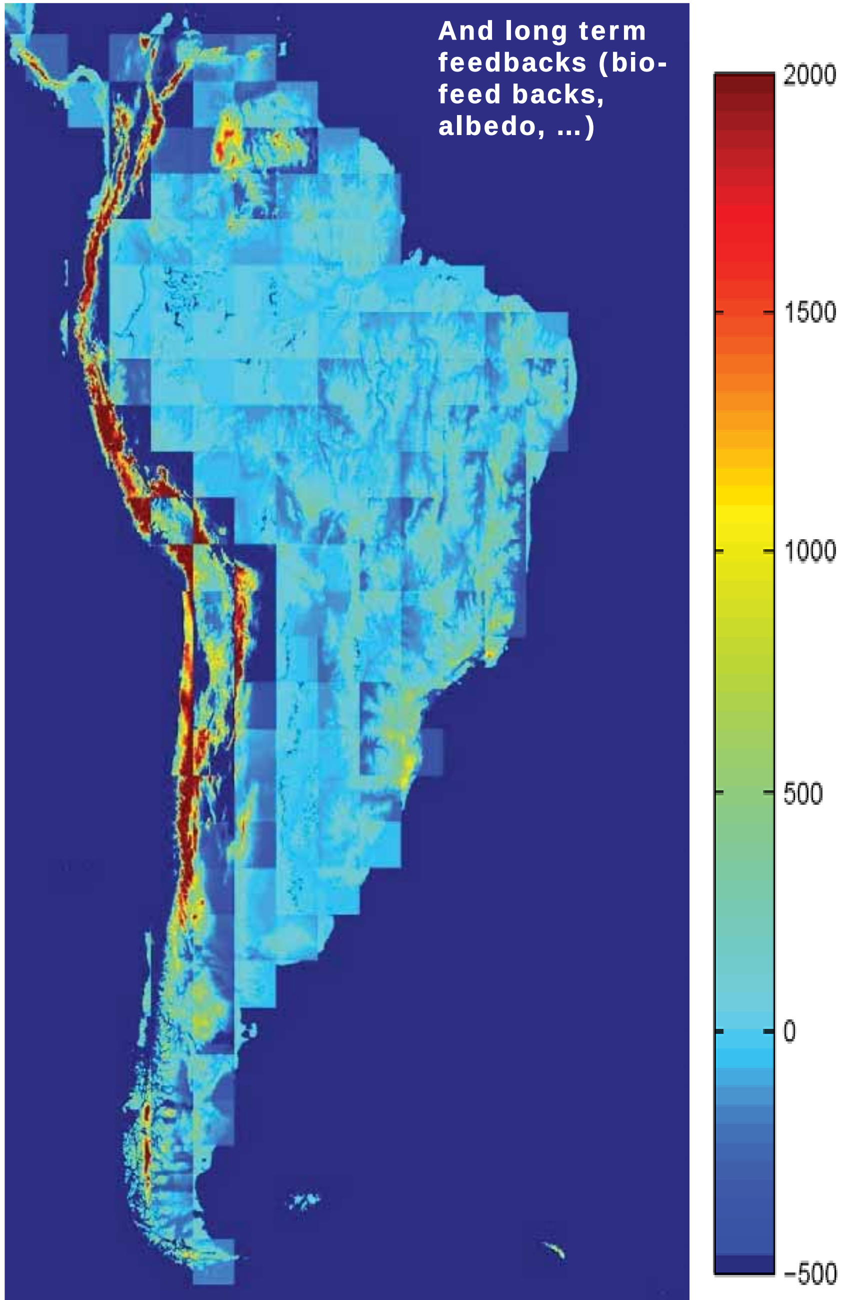
When does
“Sit and Think” trump
“Simulate and Count”?

Example: When we know moist air must go over or around in (and only in) the real world!

If our models cannot reproduce today’s driving meteorological phenomena, can we expect them to get second order feedbacks “well enough”?

At what lead times do inadequacies in downstream flow (or precipitation) result in feedbacks with beyond local impacts? alter extremes? &c?

Science can estimate Prob(Big Surprise) as a function of lead time.

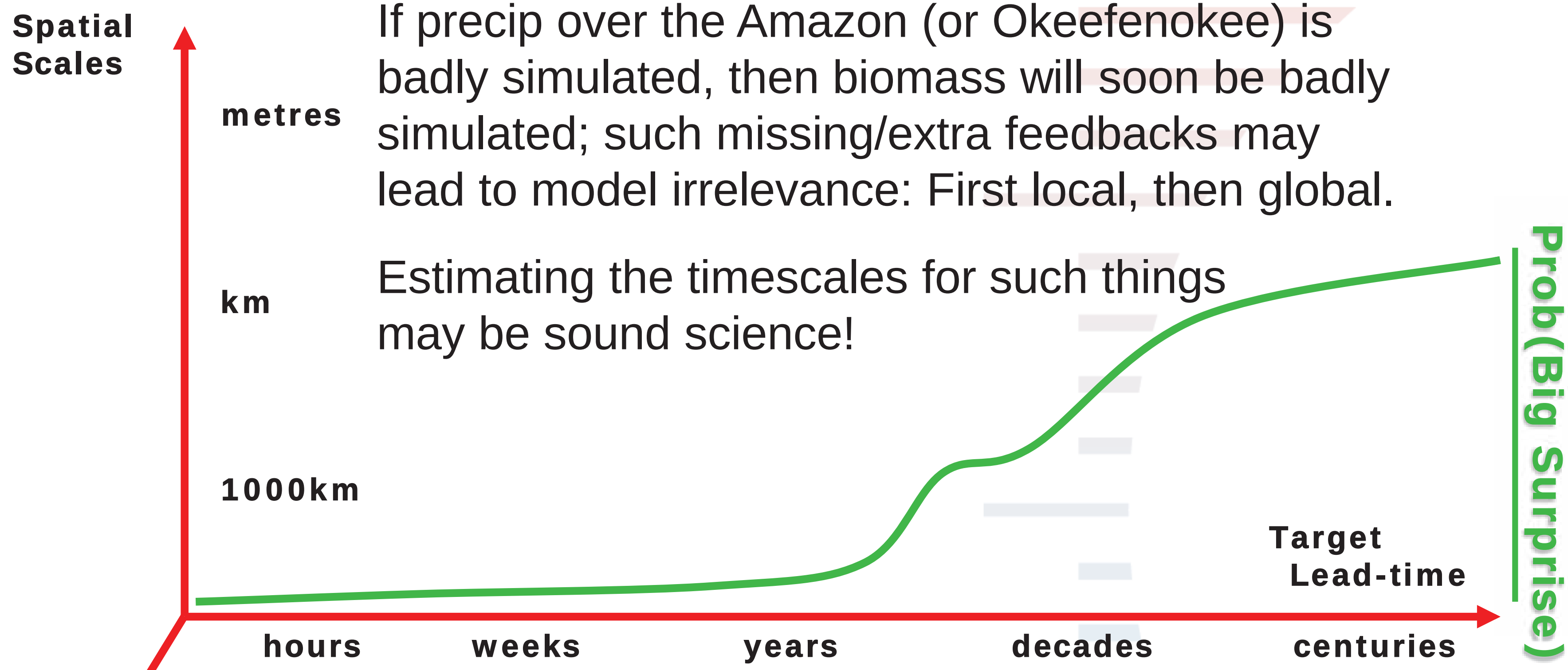


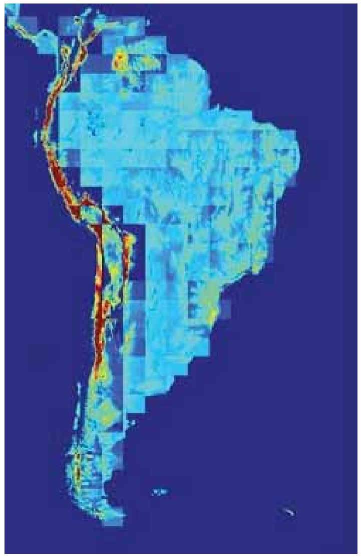
Observed Height – HadCM3 Height

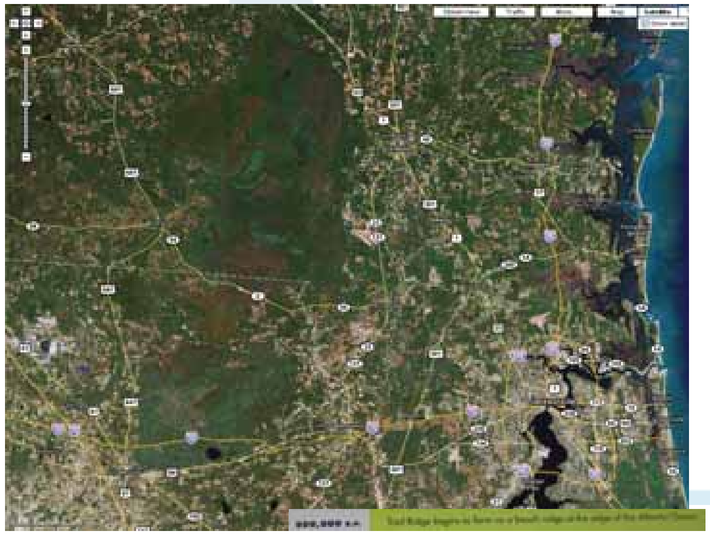
Model-based probability forecasts are incomplete without a quantitative measure of the likelihood of model irrelevance.

If precip over the Amazon (or Okeefenokee) is badly simulated, then biomass will soon be badly simulated; such missing/extra feedbacks may lead to model irrelevance: First local, then global.

Estimating the timescales for such things may be sound science!







A genuine expert can always foretell a thing that is 500 years away easier than he can a thing that's only 500 seconds off.

- *A Connecticut Yankee in King Arthur's Court*

What is a "Big Surprise"?

Big Surprises arise when something our simulation models cannot mimic turns out to have important implications for us.

Often we can identify cases where we are "leaking probability" when a fraction of our model runs explore conditions which we know they cannot simulate realistically. (Science can warn of "known unknowns" even when the corresponding futures remains unknown)

Big Surprises invalidate (not update) model-based probability forecasts changing the *I* in P(x|*I*). (Arguably "Bayes" does not apply: this is not a question of probability theory.)

A genuine expert can always foretell a thing that is 500 years away easier than he can a thing that's only 500 seconds off.

- *A Connecticut Yankee in King Arthur's Court*

Convergent Probabilities

In weather forecasts we can see the lead-time when our models become misinformative, but in climate forecasting we are in the dark.

If our models agreed (in distribution) why we would we have more confidence?

Arguably only when the next U in the UQ queue can be argued small in theory: that is when all significant RDUs have been well quantified **and we no longer *expect* our probabilities to be evolving.**

And if the best available model is not adequate for purpose

Accept (for a moment) that Model Inadequacy may make probability forecasting irrelevant in just the same way that chaos made the RMS/least-squares error of point forecasts irrelevant.

When the best model is not adequate for quantitative prediction or a RDU has not been quantified:
What is the role of quantitative modelling & simulation in decision support? In explanation?

How can we better extract insight and information from big models and ensemble forecasts without taking them too seriously?

How do we do UQ in this case?

Might this lead to better decision making?

Queuing the Right Uncertainty

"*When in doubt*, distrusting the indications, or inferences from them (duly considered on purely scientific principles, and checked by experience), the words "*Uncertain*," or "*Doubtful*," may be used, without hesitation.

Fitzroy, 1862

Best practice requires UQ communicate the RDU (the dominant uncertainty relevant to the question at hand), **not** the easiest uncertainty to quantify!

Conditioning on "the evidence considered" (rather than all the available evidence) will mislead decision makers and risk the credibility of science-based policy when an RDU is known not to have been "considered".

To be of value in application UQ must move well beyond "imprecision".

Where models disagree (in distribution), one must communicate the probability all available models are misinformative. Similarly when they agree!

What are the alternatives to using science to quantify the Prob(Big Surprise)? Are there not dangers to the credibility of science when it is suppressed?

Thank you

LA Smith (2002) What Might We Learn from Climate Forecasts? *Proc. National Acad. Sci.* USA 4 (99): 2487-2492.
LA Smith & N Stern (2011) Uncertainty in science and its role in climate policy *Phil. Trans. R. Soc. A (2011),* **369**
K Bevan, W Buytaert & L A Smith (2012) On virtual observatories and modelled realities *Hydrol. Process.*, 26: 1905
J Bröcker & LA Smith (2008) From Ensemble Forecasts to Predictive Distribution Functions *Tellus A* 60(4): 663 D Orrell, LA Smith, T Palmer & J Barkmeijer (2001) Model Error in Weather Forecasting, *Nonlinear Processes in Geophysics* 8: 357-371.

CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

http://www2.lse.ac.uk/CATS/publications/Publications_Smith.aspx

Publications http://www2.lse.ac.uk/CATS/publications/Publications_Smith.aspx

LA Smith, (2002) What Might We Learn from Climate Forecasts? *Proc. National Acad. Sci.* USA 4 (99): 2487-2492.

Smith, LA and Stern, N (2011) Uncertainty in science and its role in climate policy *Phil. Trans. R. Soc. A (2011)*, **369**, 1-24

K Bevan, W Buytaert & L A Smith (2012) On virtual observatories and modelled realities *Hydrol. Process.*, 26: 1905–1908

R Hagedorn and LA Smith (2009) Communicating the value of probabilistic forecasts with weather roulette. *Meteorological Appl* 16 (2): 143-155.

K Judd, CA Reynolds, LA Smith & TE Rosmond (2008) The Geometry of Model Error. *Journal of Atmospheric Sciences* 65 (6), 1749-1772.

J Bröcker & LA Smith (2008) From Ensemble Forecasts to Predictive Distribution Functions *Tellus A* 60(4): 663. Abstract

DA Stainforth, MR Allen, ER Tredger & LA Smith (2007) Confidence, uncertainty and decision-support relevance in climate predictions, *Phil. Trans. R. Soc. A*, **365**, 2145-2161. Abstract

LA Smith (2006) Predictability past predictability present. MS Roulston, J Ellepola & LA Smith (2005) Forecasting Wave Height Probabilities with Numerical Weather Prediction Models, *Ocean Engineering* **32** (14-15), 1841-1863. Abstract

DA Stainforth, T Aina, C Christensen, M Collins, DJ Frame, JA Kettleborough, S Knight, A Martin, J Murphy, C Piani, D Sexton, L Smith, RA Spicer, AJ Thorpe, M.J Webb, MR Allen (2005) Uncertainty in the Predictions of the Climate Response to Rising Levels of Greenhouse Gases *Nature* 433 (7024): 403-406.

A Weisheimer, LA Smith & K Judd (2005) A New View of Forecast Skill: Bounding Boxes from Seasonal Forecasts, *Tellus* **57** (3) 265-279 MAY.

K Judd & LA Smith (2004) Indistinguishable States II: The Imperfect Model Scenario. *Physica D* **196**: 224-242.

MG Altalo & LA Smith (2004) Using ensemble weather forecasts to manage utilities risk, *Environmental Finance* October 2004, 20: 8-9.

MS Roulston & LA Smith (2004) The Boy Who Cried Wolf Revisited: The Impact of False Alarm Intolerance on Cost-Loss Scenarios, *Weather and Forecasting* 19 (2): 391-397.

LA Smith (2003) Predictability Past Predictability Present. In 2002 ECMWF Seminar on Predictability. pg 219-242. ECMWF, Reading, UK.
also now: Chapter 9 of *Predictability of Weather and Climate* (eds T. Palmer and R Hagedorn). Cambridge, UK. Cambridge University Press.

MS Roulston, DT Kaplan, J Hardenberg & LA Smith (2003) Using Medium Range Weather Forecasts to Improve the Value of Wind Energy Production, *Renewable Energy* 29 (4) April 585-602.

MS Roulston and LA Smith (2002) 'Weather and Seasonal Forecasting' in *Climate Risk and the Weather Market* pg 115-126. Risk Books, London.

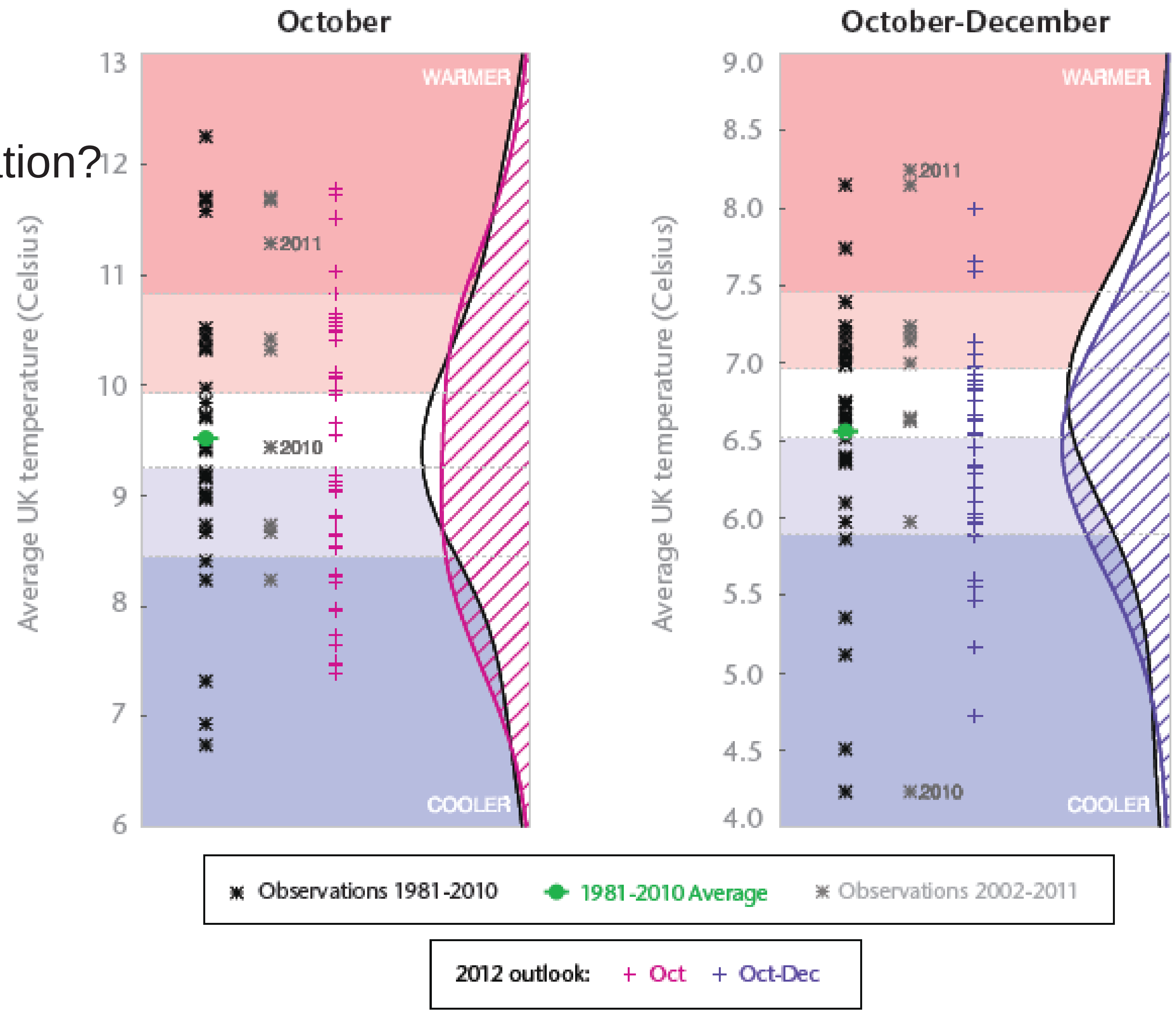
MS Roulston & LA Smith (2002) Evaluating probabilistic forecasts using information theory, *Monthly Weather Review* 130 6: 1653-1660.

D Orrell, LA Smith, T Palmer & J Barkmeijer (2001) Model Error in Weather Forecasting, *Nonlinear Processes in Geophysics* 8: 357-371.

The UK MetOffice Queuing the Right U.

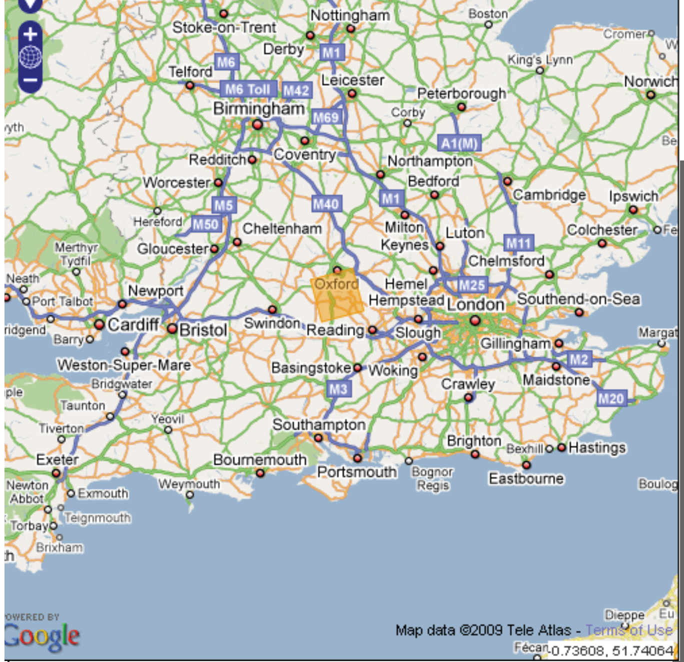
Skill?
Value?
Expectation?

Time to
Value?



http://www.metoffice.gov.uk/media/pdf/n/3/A3-plots-temp-OND.pdf

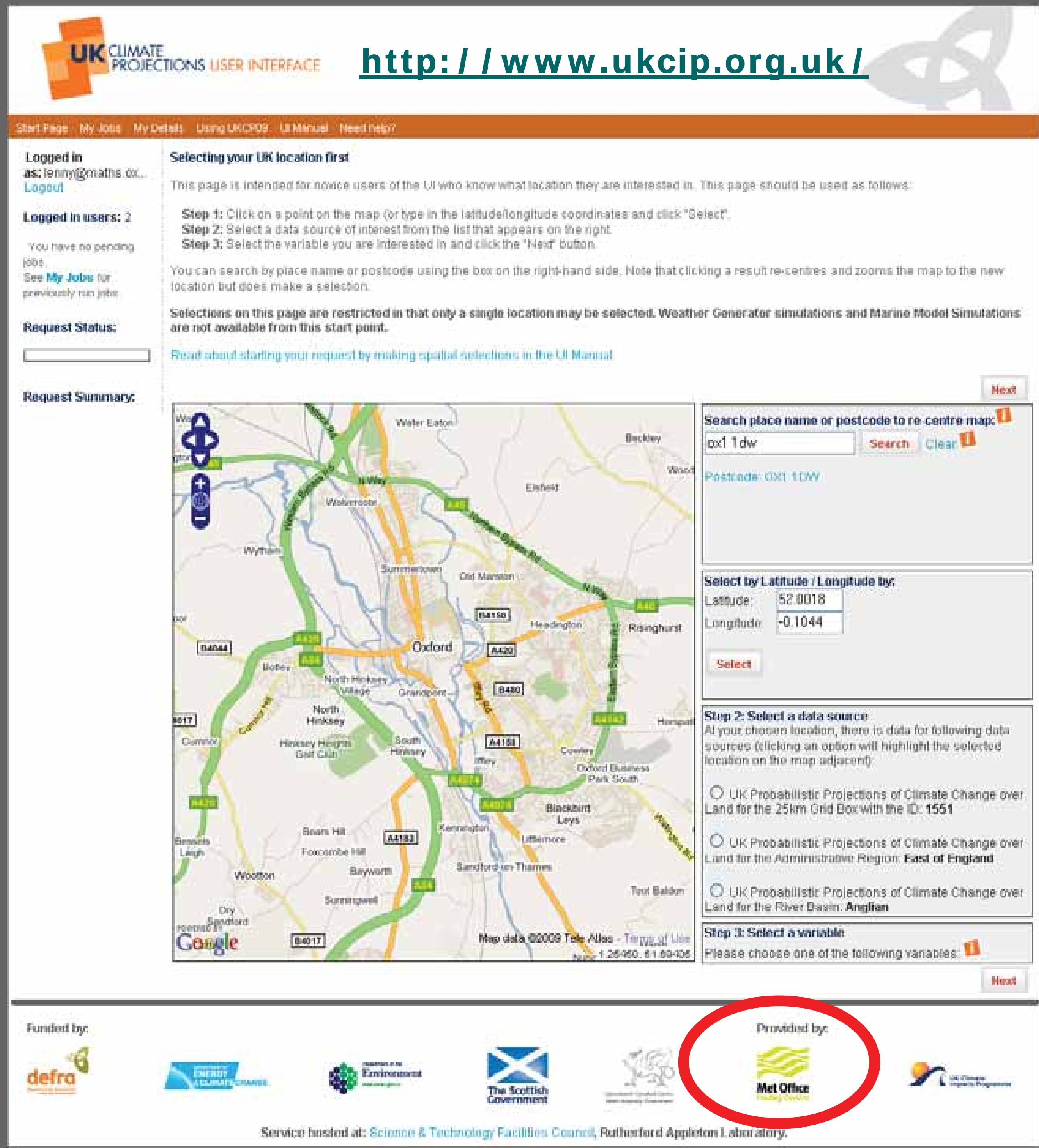
Is it plausible to provide a PDF of hottest or stormiest summer day in 2080's Oxford???



- Future Climate Change Only
- Future Absolute Climate Values

Variable

- Change in mean temperature (°C)
- Change in mean daily maximum temperature (°C)
- Change in mean daily minimum temperature (°C)
- Change in temperature of the coolest day (°C)
- Change in temperature of the warmest day (°C)
- Change in temperature of the coldest night (°C)
- Change in temperature of the warmest night (°C)
- Change in precipitation (%)
- Change in precipitation on the wettest day (%)
- Change in mean sea level pressure (hPa)
- Change in total cloud (%)
- Change in relative humidity (%)
- Change in specific humidity (%)
- Change in net surface longwave flux (W m-2)
- Change in net surface shortwave flux (W m-2)
- Change in total downward surface shortwave flux (W m-2)



Smith

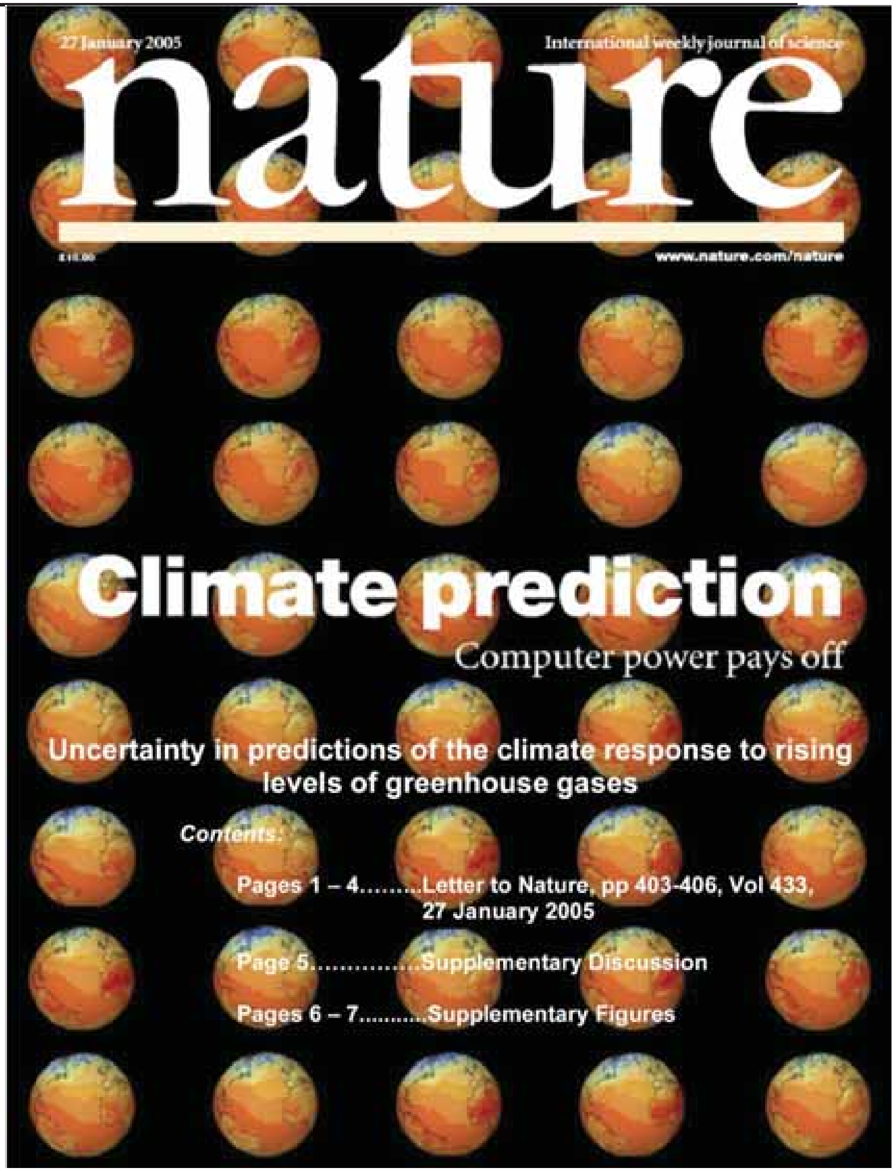
What is the aim of Climate Modelling?

It would be interesting to trace how the idea that climate models could provided quantitative insight came about.

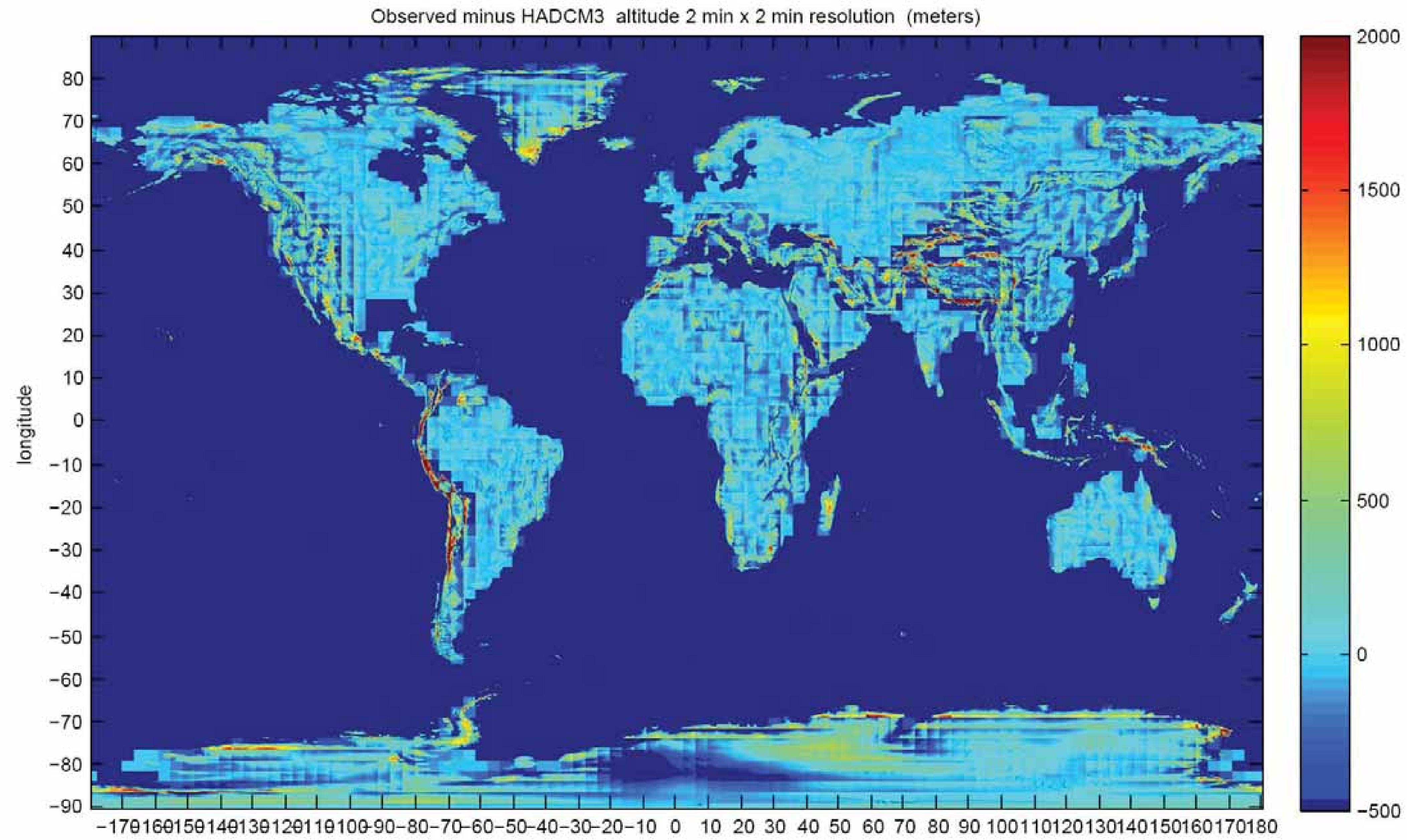
Weather models are simplified climate models: you need not turn on ocean currents in the first few days, or ice in the first few weeks, or forest in the first few years…

But climate models must run faster than real-time, and so are simplified in implementation:
do we have the technology to run high fidelity climate models?

Why do we hide behind clouds when we cannot realistically simulate rock?



DA Stainforth, T Aina, C Christensen, M Collins, DJ Frame, JA Kettleborough, S Knight, A Martin, J Murphy, C Piani, D Sexton, L Smith, RA Spicer, AJ Thorpe, M.J Webb, MR Allen (2005) Uncertainty in the Predictions of the Climate Response to Rising Levels of Greenhouse Gases *Nature* 433 (7024): 403-406.


Observed minus HADCM3 altitude 2 min x 2 min resolution (meters)
longitude
80
70
60
50
40
30
20
10
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−170 −160 −150 −140 −130 −120 −110 −100 −90 −80 −70 −60 −50 −40 −30 −20 −10 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
2000
1500
1000
500
0
−500

?What year did climate prediction move beyond understanding to quantitative forecasting'

The basic insight here is not new

When in doubt, **distrusting the indications, or inferences from them (duly considered on purely scientific principles, and checked by experience), the words "*Uncertain*," or "*Doubtful*," may be used, without hesitation.**

Fitzroy, 1862

Dr. Platzman

I may add to this another point mentioned by Dr. Charney, a somewhat philosophical comment concerning model experiments. I think that I agree with Dr. Charney's suggestion that machines are suitable for replacing model experiments. But I think it is also necessary to remember that there are in general two types of physical systems which one can think of modeling. In one type of system one has a fairly good understanding of the dynamical workings of the system, involved. Under those conditions the machine modeling is not only practical but probably is more economical in a long run. Typical examples of this kind, I think, are problems where you are concerned, let's say, with wave action in harbors, in general a whole class of engineering problems of that kind. But there is another class of problem where we are still far from a good understanding of the dynamical properties of the system. In that case laboratory models, I think, are very effective and have a very important place in the scheme of things.

PROCEEDINGS
OF
THE INTERNATIONAL SYMPOSIUM
ON NUMERICAL WEATHER
PREDICTION IN TOKYO

NOVEMBER 7–13, 1960

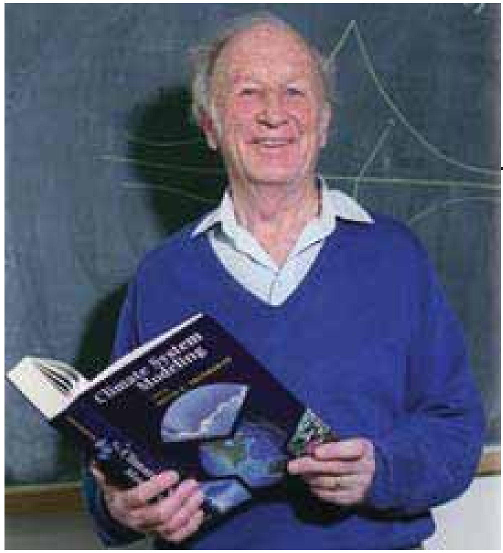
Because of the various simplifications of the model described above, it is not advisable to take too seriously the quantitative aspect of the results obtained in this study. Nevertheless, it is hoped that this study not only emphasizes some of the important mechanisms which control the response of the climate to the change of carbon dioxide, but also identifies the various requirements that have to be satisfied for the study of climate sensitivity with a general circulation model.

The Effects of Doubling the CO_2 Concentration on the Climate of a General Circulation Model[1]

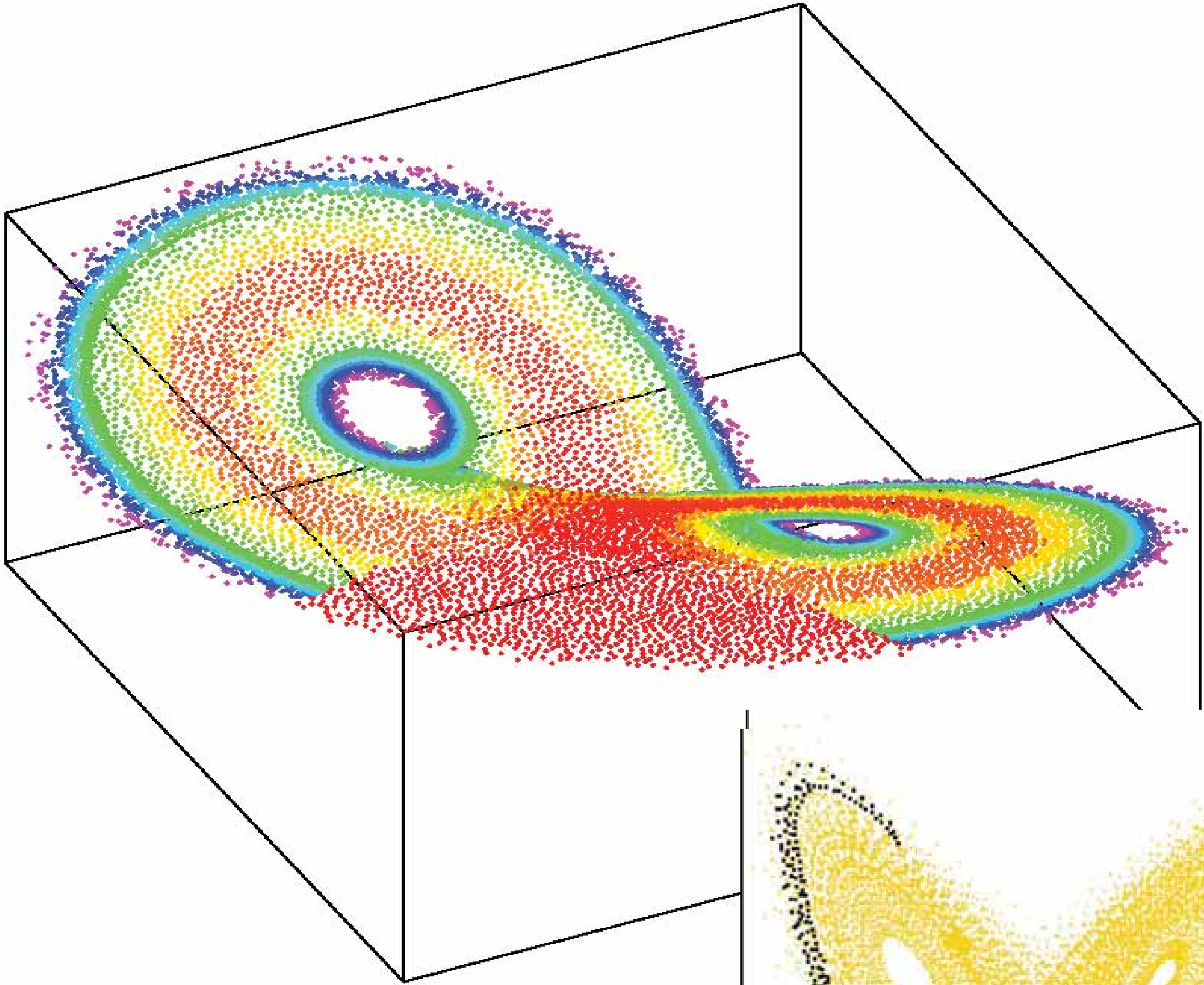
Syukuro Manabe and Richard T. Wetherald

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, N.J. 08540

(Manuscript received 6 June 1974, in revised form 8 August 1974)

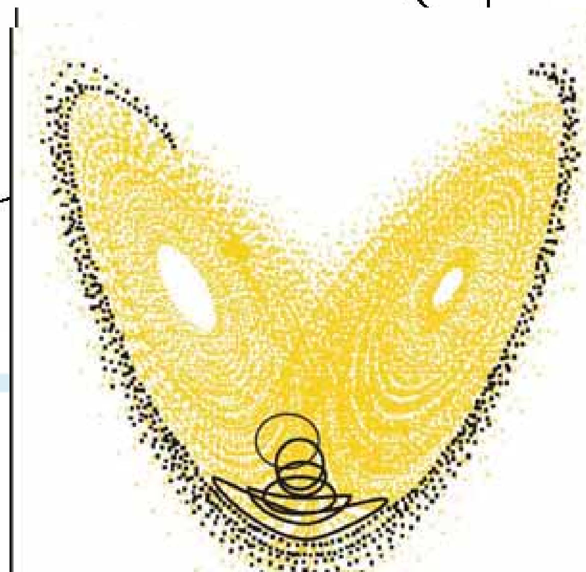


Ed Lorenz: Weather and Chaos (and Error)



Lorenz realised that even for the Apprentice, small uncertainties could grow exponentially fast, leading to "chaos."

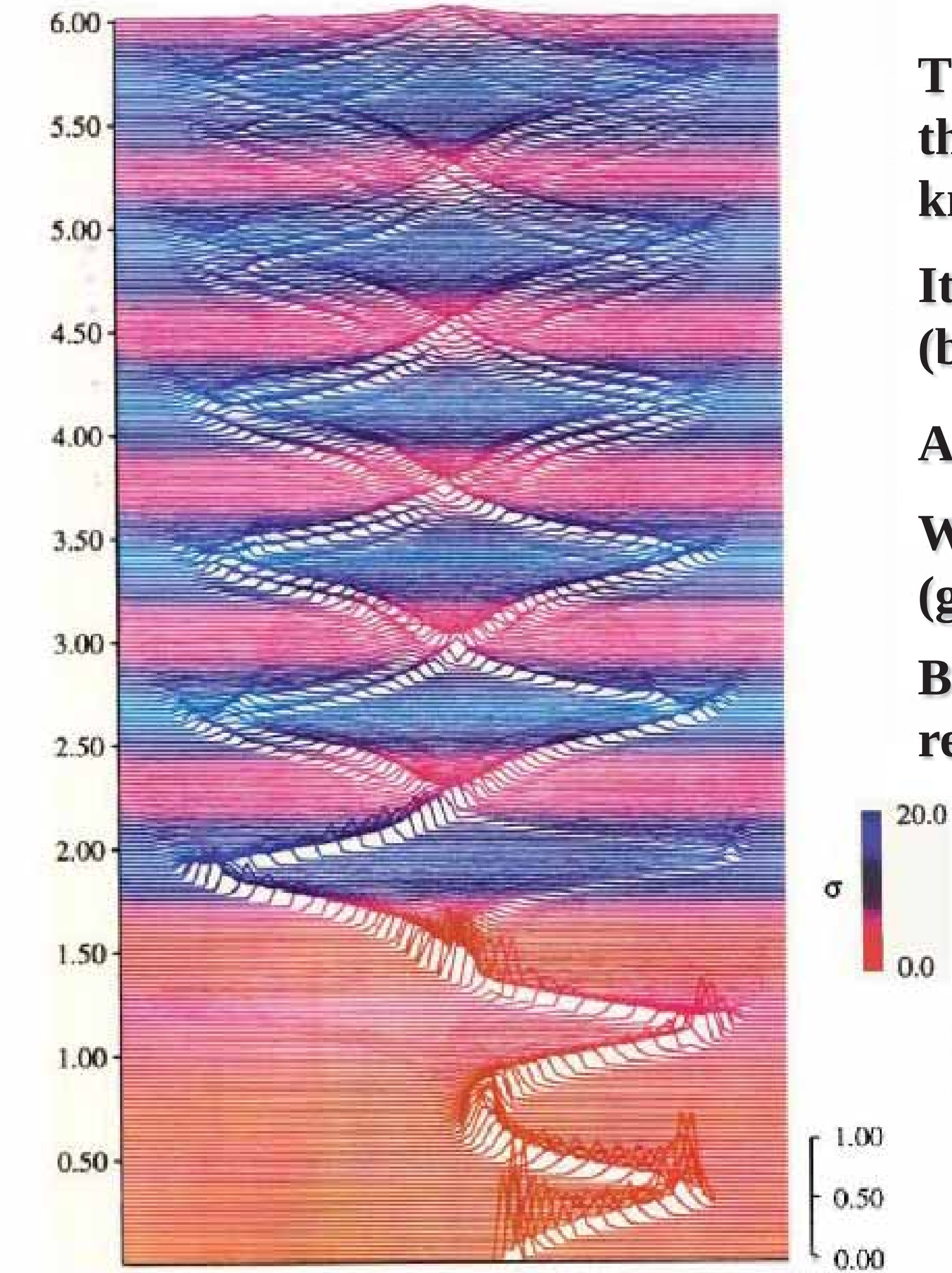
He was also very concerned about the role of model error, which is much harder to solve than that of mere chaos.



Thx to Tim Palmer

LA Smith (1994) Local Optimal Prediction: Exploiting strangeness and the variation of sensitivity to initial condition. *Phil. Trans. Royal Soc. Lond.* A, 348 (1688): 371-381.

Probability Forecasts: Chaos



Smith (2002) Chaos and Predictability in *Encyc Atmos Sci*

The evolution of this probability distribution for the chaotic Lorenz 1963 system tells us all we can know of the future, given what we know now.

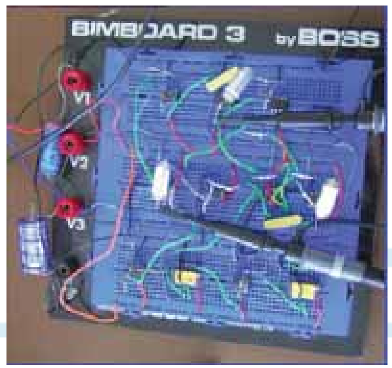
It allows prudent quantitative risk management (by brain-dead risk managers)

And sensible resource allocation.

We can manage uncertainty for chaotic systems (given a perfect model).

But how well do we manage uncertainty in the real world? For GDP? Weather? Climate?

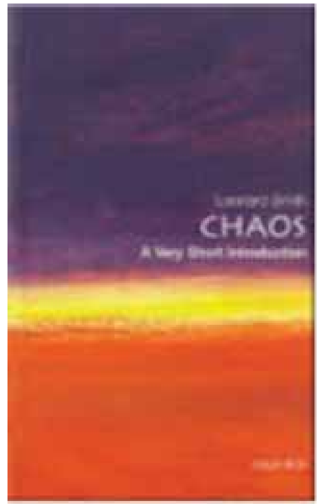
Do we have a single example of a nontrivial system where anyone has succeeded (and willing to offer odds given their model-based PDFs?)



Can we disentangle these “uncertainties”?

Demon’s Apprentice (2007)

1) Perfect Equations of Motion (PMS)
2) ~~Perfect noise-free observations~~
3) Unlimited computational power



Apprentice’s Novice (2012)

1) ~~Perfect Equations of Motion (PMS)~~
2) ~~Perfect noise-free observations~~
3) Unlimited computational power

Enter the NAG Board