To the Physicist Sitting in Darkness

Probabilities are all well and good. And it is a fine thing to get in touch with your beliefs and feelings. Shall we bang ahead in our old-time, loud pious way, and commit new sciences to the game; or shall we sober up, sit down, and think it over first?

The Blessings-of-Subjective-Probability Trust, wisely and cautiously administered, is a Blue Chip. But Bayesians have been playing it badly of late, and must certainly suffer from it, in my opinion; they have been eager to solve every problem, especially the poorly posed ones, and the Physicists who sit in Darkness have begun to notice it – they have noticed it and have begun to show alarm. They have become suspicious of posteriors on empirically vacuous reals, not to mention function spaces; they have begun to resist the kindly extraction of priors. More – they have begun to examine them! This is not well. The Blessings of Bayesianism are all right, and a good RC commercial property; there could not be better, in a dim light. In the right kind of light, and at the proper distance, with the goods a little out of focus, they are a desirable enticement to the Physicists who sit in darkness.

Probability theory eases the stress of decision making. And improves the outcome, but not if we adulterate it. For the Empirically Adequate and the Large Number Statistic, it is pie. But in cutting edge science, and in extrapolation, here the Physicist sitting in darkness is (almost) sure to say: “These is something curious about this – curious and unaccountable.” … There have been lies yes, but told in a good cause, it might have worked; yet we have passed on a Shadow from one who hadn’t it to sell, and long term infrastructure investments are being made.
Predictability & Understanding of Our Climate Risk: Approximations, Bugs and Insight

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

Not Possible without: H Du, A. Jarman, D Stainforth & E Suckling
What I wish to avoid

An Oxford Bus Stop in the summer of 2010:

X30+N30 predictions are wrong
Sorry for any inconvenience

Why do Predictions go wrong?

a) Technology is Lacking
(Should we hide this?)
b) Science is Lacking
(Can we know this?)
c) Political Shenanigans
(Can we ignore this?)

Is it possible for Probabilistic Predictions to be “wrong”? (yes, sometimes)
But first... a timely example for those of you flying south:
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Area 1: 10% Chance of Tropical Cyclone Formation (click to zoom)
A BROAD AREA OF LOW PRESSURE...ACCOMPANIED BY CLOUDINESS AND THUNDERSTORMS...HAS DEVELOPED ABOUT 250 MILES EAST OF THE SOUTHEASTERN BAHAMAS. ONLY SLOW DEVELOPMENT IS ANTICIPATED AS THE LOW MOVES WESTWARD AT 10 MPH DURING THE NEXT DAY OR TWO. THIS SYSTEM HAS A LOW CHANCE...10 PERCENT...OF BECOMING A TROPICAL CYCLONE DURING THE NEXT 48 HOURS.

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center, Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

800 PM EDT SAT OCT 6 2012

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Medium: 30-50%
- High: >50%
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

800 AM EDT SUN OCT 7 2012

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook
Graphical Tropical Weather Outlook

National Hurricane Center		Miami, Florida

200 PM EDT SUN OCT 7 2012

Satellite Image: 0130 PM EDT

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook

National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

200 AM EDT MON OCT 8 2012
Satellite Image: 0122 AM EDT
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

200 PM EDT MON OCT 8 2012
Satellite Image: 0122 PM EDT

Go to Eastern Pacific Outlook

Oct 2012 IEEE eScience: Science in the Dark Leonard Smith
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook

800 PM EDT MON OCT 8 2012
Satellite Image: 0652 PM EDT
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

200 AM EDT TUE OCT 9 2012
Satellite Image: 0122 AM EDT

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Satellite Image: 0700 AM EDT
800 AM EDT TUE OCT 9 2012

Go to Eastern Pacific Outlook

Leonard Smith
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- **Yellow**: Low <30%
- **Orange**: Medium 30-50%
- **Red**: High >50%

Go to Eastern Pacific Outlook

800 AM EDT TUE OCT 9 2012
Satellite Image: 0700 AM EDT
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

200 PM EDT TUE OCT 9 2012
Satellite Image: 1252 PM EDT

CATS Centre for the Analysis of Time Series
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook

200 AM EDT WED OCT 10 2012
Satellite Image: 1252 AM EDT
Area 2: 0% Chance of Tropical Cyclone Formation (click to zoom)
A weak area of low pressure located a few hundred miles east-northeast of the northwestern Bahamas continues to produce a few disorganized thunderstorms. The low is expected to merge with a frontal zone later today...and significant development is unlikely. This system has a low chance...near 0 percent...of becoming a tropical cyclone during the next 48 hours as it accelerates northeastward.
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

200 AM EDT WED OCT 10 2012
Satellite Image: 1252 AM EDT

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

800 AM EDT WED OCT 10 2012
Satellite Image: 0652 AM EDT

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- **Yellow** Low <30%
- **Orange** Medium 30-50%
- **Red** High >50%

Go to Eastern Pacific Outlook
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

2 (50%)

Go to Eastern Pacific Outlook

200 AM EDT THU OCT 11 2012

Satellite Image: 1252 AM EDT

Leonard Smith
Graphical Tropical Weather Outlook

National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook

800 AM EDT THU OCT 11 2012

Satellite Image: 0652 AM EDT
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Tropical Depression SIXTEEN (click for details)

As of 1100 AM EDT Thu Oct 11
Advisory # 1
Maximum Sustained Winds: 30 knots; 35 mph
Minimum Central Pressure: 1009 mb
Located at: 25.4N 72.6W
Movement: south at 1 knots; 1 mph

800 AM EDT THU OCT 11 2012  Satellite Image: 0652 AM EDT

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Low <30%
- Medium 30-50%
- High >50%
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%

Go to Eastern Pacific Outlook

200 PM EDT THU OCT 11 2012
Satellite Image: 0122 PM EDT
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- **Yellow**: Low <30%
- **Orange**: Medium 30-50%
- **Red**: High >50%

Satellite Image: 0652 PM EDT

800 PM EDT THU OCT 11 2012

Go to Eastern Pacific Outlook

Leonard Smith
Graphical Tropical Weather Outlook
National Hurricane Center  Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours.

- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Graphical Tropical Weather Outlook
National Hurricane Center Miami, Florida

Outlined areas denote current position of systems discussed in the Tropical Weather Outlook. Color indicates probability of tropical cyclone formation within 48 hours:
- Yellow: Low <30%
- Orange: Medium 30-50%
- Red: High >50%
Collect all the forecasts of 10% probability of an event, and compute the fraction of them in which the event occurred!

This science is in the light not because the forecasts tend to be consistent: but merely because they can be evaluated with relevant observations!
The constants they are a chang’n (in the light)

Science often underestimates its own uncertainty. Whether this is due to carelessness to intellectual phase locking, science in the light can advance when new evidence contradicts old belief.
Science in the Dark does not imply scientists have gone over to the dark side.

Scientists are forced to “violate” traditional best practice guidance if such violations are imposed by the nature of the question being addressed. One cannot wait 50 years for out-of-sample observations. It is a brute fact that a climate model’s lifetime is less than its forecast lead-time! The physics underlying CO₂ induced warming remains as solid as science gets.
It was a dark and stormy night…

…in 1707, when a dispute broke out between Sir Cloudesley Shovell, the master of HMS Association and his navigator regarding where they were and when they would reach land. Legend has it the sailor was hung, but no record survives, in part because…

…on the Oct 22 the ships hit “unexpected” rocks; four ships and 1400 sailors were lost, including Sir Shovell, and the Association.

Dangers of Forecasting (c)
Political Shenanigans: The forecaster may be vulnerable

[Web Link](http://www.northwindpictures.com/results.asp?link=J405&co=NORPIC10)
Fast-forward to 1859

In October 1859, the steam clipper *Royal Charter* founded in a violent storm off the coast of Anglesey in North Wales with the loss of over 450 lives. As a direct result of this disaster, FitzRoy introduced the first British storm warning service for shipping in February 1861, making use of the electric telegraph. FitzRoy was the main influence in the early development of the Met Office, which was then primarily intended to improve safety at sea.

Fitzroy exploited **new technology**, the telegraph, to provide information on current meteorological conditions, but he was not allowed to “foretell” the future. (He coined the word “forecast” but was not allowed to issue them!)

He partnered with Le Verrier in France, who was famous for predicting the location (and existence) of the planet Neptune. Fitzroy met resistance in the UK.
Objection has been taken to such forecasts, because they cannot be always exactly correct,—for all places in one district. It is, however, considered by most persons that general, comprehensive expressions, in aid of local observers, who can form independent judgments from the tables and their own instruments, respecting their immediate vicinity, though not so well for distant places, may be very useful, as well as interesting: while to an unprovided or otherwise uninformed person, an idea of the kind of weather thought probable cannot be otherwise than acceptable, provided that he is in no way bound to act in accordance with any such views, against his own judgment.

Like the storm signals, such notices should be merely cautionary—to denote anticipated disturbance somewhere over these islands, without being in the least degree compulsory, or interfering arbitrarily with the movements of vessels or individuals.

Certain it is, that although our conclusions may be incorrect—our judgment erroneous—the laws of nature, and the signs afforded to man, are invariably true. Accurate interpretation is the real deficiency.

Fitzroy, 1862

Fitzroy aimed to provide decision makers with more information: not orders or decisions made.
Fitzroy

Le Verrier

Galton

Einstein

Whitehead

Lorenz

Richardson
Le Verrier & the “Discovery” of Neptune and Vulcan

Le Verrier found discrepancies between the observations of Uranus and the workings of Newton’s Laws; he knew these must be caused by a previously unknown planet, and in 1846 predicted where that planet should be, leading to the discovery of “Neptune”.

In 1859 Le Verrier found discrepancies between the observations of Mercury and the workings of Newton’s Laws; he knew these must be caused by a previously unknown planet, and predicted another new planet. This lead to the discovery of “Vulcan” by Lescarbault on 26 March 1859. Lescarbault was awarded the Légion d’honneur.

There is no planet Vulcan! Newton’s Laws fail near the sun.
“The advantage of confining attention to a definite group of abstractions, is that you confine your thoughts to clear-cut definite things, with clear-cut definite relations. … The disadvantage of exclusive attention to a group of abstractions, however well-founded, is that, by the nature of the case, you have abstracted from the remainder of things.

... it is of the utmost importance to be vigilant in critically revising your *modes* of abstraction.

*Sometimes it happens that the service rendered by philosophy is entirely obscured by the astonishing success of a scheme of abstractions in expressing the dominant interests of an epoch.*”


Whitehead was criticising straightjacket of Newtonian science; I will suggest that computer simulation may impede more than just the progress of science.

But first, let's bring in Galton…
Sir Francis Galton FRS (16 February 1822 – 17 January 1911), cousin of Douglas Strutt Galton, cousin of Charles Darwin, was an English Victorian polymath: anthropologist, eugenicist, tropical explorer, geographer, inventor, meteorologist, proto-geneticist, psychometrician, and statistician. He was knighted in 1909.

http://galton.org/meteorologist.html
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)

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?
Laplace's Demon (1814)
1) Perfect Equations of Motion (PMS)
2) Perfect noise-free observations
3) Unlimited computational power

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

http://2darts.com/2dtuts/articles/50-t Terrifying-creatures/
Lewis Fry Richardson’s Successful Failure (in the light)

Given the equations and an initial condition, one can compute a forecast.

Richardson made the first numerical weather forecast for 20 May 1910. He did this by hand, while working as an ambulance driver in WWI.

While the forecast was wildly inaccurate, comparison with what actually had happened, it allowed progress leading to modern systems.

Of course, even today's best models do not resolve everything we see from space.
Richardson realised that his algorithm could be run in parallel by a large number of “computers” sitting in a theatre...

He suggested that 64,000 computers could forecast the weather in real time.
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.

Probability Forecasts: Chaos

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?)
Probability Forecasts: “Simple” “chaotic” Physical System

A Big Surprise in the Moore-Spiegel Circuit (by Reason Machette)
Figure 7: Ensemble predictions using (a) model 1 and (b) model 2. The
Figure 7: Ensemble predictions using (a) model 1 and (b) model 2. The 2: Ensemble predictions using (a) model 1 and (b) model 2. The
Laplace's Demon (1814)

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

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

Do the pro’s fare better?
Met Office seasonal forecasts to be scrapped

The Met Office is to stop publishing seasonal forecasts, after it came in for criticism for failing to predict extreme weather.

It was berated for not foreseeing that the UK would suffer this cold winter or the last three wet summers in its seasonal forecasts.

The forecasts, four times a year, will be replaced by monthly predictions.

The Met Office said it decided to change its forecasting approach after carrying out customer research.

Explaining its decision, the Met Office released a statement which said: "By their nature, forecasts become less accurate the further out we look."

SEE ALSO
- Met Office rethink on forecasts

http://news.bbc.co.uk/2/hi/uk_news/8551416.stm
Met Office document shows it only renamed its seasonal forecasts

A Freedom of Information request submitted to the Met Office by Autonomous Mind reveals the Met Office did not tell the truth when it said it had scrapped its seasonal forecast.

Despite repeatedly trailing the line that the Met Office no longer issues seasonal forecasts because the public says they are not of use (a separate blog post on that public view later today, with some new information that has come to light... Update: part two can now be read here), the reality is that the department’s Chief Executive, John Hirst, engaged in a smoke and mirrors exercise in an attempt at reputation management.

At a Board meeting of the Met Office on 26 January 2010, (original Minutes extract received under FOI: 0012014 AM Attachment) a recommendation was tabled by Hirst to rename the forecasts and locate them in a different part of the department’s website, and that Hirst:

[Link to article: http://autonomousmind.wordpress.com/2011/02/01/met-office-document-shows-it-only-renamed-its-seasonal-forecasts/]
These forecasts are not quite in the dark.

Skill? Value? Expectation?

Time to Value?

http://www.metoffice.gov.uk/media/pdf/n/3/A3-plots-temp-OND.pdf
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
The NAG Board
(Not A Galton Board 2000)
150th Birthday of RMS
How might ensembles help us understand uncertainty? Consider the Not A Galton (NAG) Board.

In the NAG board, probability forecasting corresponds to predicting with a collection (ensemble) of golf balls...

Ensembles inform us of uncertainty growth within our model! (Telling us about the next golf ball.)
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 the one (and only) red rubber ball?

While we never see similar initial states, we can still learn from our mistakes!(in this weather-like case)
Science can anticipate surprises beyond model-land

Interpreting even weather-like distributions is a challenge!

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

Scientific insight can help.

But the best we can hope for is sensible, consistency in distribution between our models (“the details do not matter”).

And to anticipate “Big Surprises” (things are models cannot do)
It would be interesting to trace how the idea that climate models could provide 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?

A very schematic schematic reflecting phenomena the model “includes”. (Note the turtle)

The detail you see above is what is missing in HadCM3: the large squares reflect model grid resolution, the detail reflects the difference between the observed surface height and the model surface height, “constant” within a grid point,
The real concern is that an adequately parameterised process might significantly shift the range.

Discussions of broadening imply confidence in the location.

Do we have that confidence?

The IPCC rejects the diversity of ensembles directly reflecting the pdf of GMT, it follows that “downscaling” cannot provide local probabilities.
Climate in Practice: In-sample examples.

This graph tends to leave the impression GCMs do rather well.

FAQ 8.1, Figure 1. Global mean near-surface temperatures over the 20th century from observations (black) and as obtained from 38 simulations produced by 14 different climate models driven by both natural and human-caused factors that influence climate (yellow). The mean of all these runs is also shown (thick red line). Temperature anomalies are shown relative to the 1901 to 1950 mean. Vertical grey lines indicate the timing of major volcanic eruptions. (Figure adapted from Chapter 9, Figure 9.5. Refer to corresponding caption for further details.)
As we refocus from climate-past to the climate-future, how do we cope with such systematic errors, even as we work to reduce them?

Anomalies may be fine for mitigation. They are a nonsense for adaptation. (and the laws of physics.) (and biology.) (Ice melts at zero C, plants die at ....)

Moving to anomaly space requires
Letting go of the “Laws of Physics”.

Note model anomalies are not exchangeable even after 100 years!
A report of Working Group I of the Intergovernmental Panel on Climate Change

Summary for Policymakers

PROJECTIONS OF SURFACE TEMPERATURES

Not necessarily wider: they may narrow and shift under better models...

The effects of uncertainty in the knowledge of Earth system processes can be partially quantified by constructing ensembles of models that sample different parametrizations of these processes. However, some processes may be missing from the set of available models, and alternative parametrizations of other processes may share common systematic biases. Such limitations imply that distributions of future climate responses from ensemble simulations are themselves subject to uncertainty (Smith, 2002), and would be wider were uncertainty due to structural model errors accounted for.

This risk of overconfidence is well known and well founded.

Global Climate Projections

Figure SPM.6. Projected surface temperature changes for the early and late 21st century relative to the period 1980–1999. The central and right panels show the AOGCM multi-model average projections for the B1 (top), A1B (middle) and A2 (bottom) SRES scenarios, averaged over the probabilities of the studies for the SRES scenarios. Therefore the

The IPCC itself might say this a bit louder/earlier: What space-time scales are realistic as a function of lead-time? (Focus on robust, but discuss inappropriate use.)
Dangers in the Dark: The value of qualitative insight is at risk of being discarded in favour of quantitative mis-information.

- Varies, but typically no single source dominates.

How might we avoid miscommunication in this case?

Source: Met Office
How can we know our simulation models are inadequate?

Science is more than simulations

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?

Can we provide Prob(Big Surprise) with lead time?
Model-based probability forecasts are incomplete without a quantitative measure of the likelihood of model irrelevance.

If precip over the Amazon (or Okefenokee) is badly simulated, the biomass will be badly simulated, this missing/extra feedback may lead to model irrelevance… First local, then global.

Timescales for such things may be sound science!
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 magnitude remains unknown)

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

In weather forecasting, we can see when are models become silly, but in climate forecasting we are in the dark.

If our models agreed (in distribution) we would we have more confidence?
Is it plausible to provide a PDF of hottest or stormiest summer day in 2080's Oxford???
Data Data Everywhere, and Not a Bit to Bank On

It seems we are surrounded by model output… but we know that the models are unlikely to be adequate to inform the questions we must answer.

What is the rational path forward when the best available model is known not to be adequate for purpose?
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

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 CO₂ Concentration on the Climate of a General Circulation Model

SUDERU 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)
Political Shenanigans
Senator Inhofe’s List

U.S. Senate Report
Over 400 Prominent Scientists Disputed
Man-Made Global Warming Claims in 2007
Scientists Debunk “Consensus”

Statistician Lenny Smith of the London School of Economics, who co-authored a study on the uncertainties of climate models for the Tyndall Centre for Climate Change Research in Oxford, dubbed climate modeling "naive realism." "Our models are being over-interpreted and misinterpreted," Smith said, according to a New Scientist article from August 16, 2007. "They are getting better, I don't want to trash them per se. But as we change our predictions, how do we maintain the credibility of the science?" Smith explained. "We need to drop the pretence that they are nearly perfect," he added. The article noted that Smith believes that the "over-interpretation of models is already leading to poor financial decision-making." The article continued: "[Smith] singled out for criticism the British government's UK Climate Impacts Programme and Met Office. He accused both of making detailed climate projections for regions of the UK when global climate models disagree strongly about how climate change will affect the British Isles." (LINK)
Derailing Science Ourselves: The Penguin Effect

The challenge of climate change will be with us for some time.

Can we maintain parallel streams: pure research to apply in 2050, and applied research to improve the modelling position we are in when we get there?

When selecting a thesis problem: do you suggest something important, like understanding cloud dynamics (better)?

Or to be the first person in the world to include the penguin effect in a global model? (and thereby all but assured a job at a rival modelling centre?)

(Similar effects plague economics and statistics)

THERE IS NO PENGUIN EFFECT
(My prior on this effect is zero)
It is a joke regarding climate, but sadly not career paths!
Science in the Dark does not imply scientists have gone over to the dark side.

Scientists must “violate” traditional best practice guidance if such violations are imposed by the nature of the question being addressed. One cannot wait 50 years for out-of-sample observations. It is a brute fact that a climate model’s lifetime is less than its forecast leadtime! The physics underlying \( \text{CO}_2 \) induced warming is as solid as science gets.
Science in the Dark does not imply scientists have gone over to the dark side

But science done without the benefit of out-of-sample evaluation cannot call on the unreasonable effectiveness of science done in the light.

Science which cannot be falsified without risking catastrophe needs to communicate clearly the limits of its (current) insights, or it risks the credibility of all science.
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 we have is not adequate from quantitative prediction: 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?

Might this lead to better decision making?
We are walking in Florida. You find you have just been bitten on the hand by a snake. We did not see the snake.

If it was the deadly carbonblack snake, the bite will kill you in a painful way, unless you cut off your hand within 15 secs.

Did you cut off your hand? You have 5 seconds left.

I have a hatchet. It was the deadly carbonblack snake, the bite will kill you in a painful way, unless you cut off your hand within 15 secs.

You find you have just been bitten on the hand by a snake.

We are walking in Florida.

The Risk is Real

Declaring a costly choice could prove a costly choice.


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


Thank you


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


also now: Chapter 9 of *Predictability of Weather and Climate* (eds T. Palmer and R Hagedorn). Cambridge, UK. Cambridge University Press.


The probability of the real world falling above the 90% line of the UKCP PDF can be much much greater than 10%.

The shortcoming of climate models are more clearly acknowledged in the peer reviewed literature than in the UKCP user guidance.

It is important to stress that our approach to the specification of discrepancy can only be expected to capture a subset of possible structural modelling errors and should be regarded as a lower bound. This is because models tend to share certain common systematic biases, which can be found in diverse elements of climate including multiannual means of basic quantities such as surface temperature,
Model Imperfections

Moving Forward: Plausible Planets or Implausible Earths?

How can we best develop our models as the available computational power increases?

A) Simulate potentially real planets that get more and more Earth-like while omitting any Earth-relevant process for which the model cannot provide coherent physical drivers on Earth-like scales. (no suggestion of linear superposition intended!)
   
   Does water vapour come after mountains?
   
   Does vegetation come after water vapour?
   
   Do we avoid the penguin effect? (until it is simulated realistically)

B) Via an hodgepodge of unphysical/unbiological simulations resembling no planet that could possibly exist, but “including” every phenomena we can think of that might be important (including penguins), and hoping the simulated planets will suddenly become Earth-like at some resolution in an ill-defined higgledy-piggledy way.

One might argue physical intuition is more effective in evaluating plausible planets, as there is physics to intuit in that case. (and at least a few examples.)
Take home questions

How might we better communicate model diversity given the possibility that we cannot get probabilities useful as such!

Do we have a single example of a nontrivial system where anyone has succeeded (and willing to bet on their model-based probabilities?)

At what lead times do inadequacies drive (or fail to drive) feedbacks yielding local impacts? extremes? global impacts?

How far to one go with a simulation model (when to stop: in time? space?)

How can we best deal with models behaving badly?

What prevents the provision of Prob(Big Surprise) with lead time?

How can we improve the communication of insights from simulations without falling afoul of forecasting good practice?

How to distinguish the value of improvement from the utility of prediction?

Might the provision of probability be maladaptive?

How might we better communicate the inadequacy as well as imprecision

Is the value of qualitative insight at risk of being discarded in favour of quantitative mis-information?
Precision or Early Warning an Event is Likely?

Smoothing out the highs and lows of running a business
As in daily life, one of the unruliest factors in running a business can be the weather. It can bring a highly functioning organization to a dead stop in a matter of seconds. Rack up billions in costs, in the blink of an eye. It's completely uncontrollable. And relentless. It is not, however, unpredictable. Increasingly, precise, short-term, extremely local forecasts can help companies sidestep major disasters as well as plan for the more routine weather events that haemorrhage energy and resources.

Where Might Climate Science Better Use eScience?
Better Handling of Data Streams

Coordinated Observations

12/18/2004
PARASOL
1:33

1/18/2004
CALIPSO
1:31:15

CloudSat
1:31

4/28/2006

5/4/2002
Aqua
1:30

OCO
1:26

7/15/2004
Aura
1:38

CloudSat – 3-D cloud climatology
CALIPSO – 3-D aerosol climatology

Aerosols, polarization

TES – T, P, H2O, O3, CH4, CO
MLS – O3, H2O, CO
HIRLMS – T, O3, H2O, CO2, CH4
OMI – O3, aerosol climatology

AIRS – T, P, H2O, CO2, CH4
MODIS – cloud, aerosols, albedo

OCO - - CO2
O2 A-band
ps, clouds, aerosols

Leonard Smith
An Oxford Bus Stop, Summer 2010:

How did we get to zipcode PDFs from here?

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

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,

The Effects of Doubling the CO\textsubscript{2} Concentration on the Climate of a General Circulation Model\textsuperscript{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 9 August 1974)

It is important to stress that our approach to the specification of discrepancy can only be expected to capture a subset of possible structural modelling errors and should be regarded as a lower bound. This is because models tend to share certain common systematic biases, which can be found in diverse elements of climate including multiannual means of basic quantities such as surface temperature,

Why do these words get lost in the graphics?
What is climate?

Climate is what you expect,  
Weather is what you get.  

Robert Heinlein (1973)

The climate system is a complex, interactive system consisting of the atmosphere, land surface, snow and ice, oceans and other bodies of water, and living things. The atmospheric component of the climate system most obviously characterizes climate; climate is often defined as ‘average weather’. Climate is usually described in terms of the mean and variability of temperature, precipitation and wind over a period of time, ranging from months to millions of years (the classical period is 30 years).

climate—"The synthesis of the weather" (C. S. Durst); the long-term manifestations of weather, however they may be expressed. More rigorously, the climate of a specified area is represented by the statistical collective of its weather conditions during a specified interval of time (usually several decades).

Climate is a distribution of multivariate time series!  
(It’s not just a number or two)  
And for policy and (most) decision support:  
“All Climate is Local”
The Advantage of being Local

While there is a common desire for non-locality, I know of no examples where a nonlocal score can be expected to have any advantages in probability forecasting.

“Should” this probability forecast score more highly for the red verification than the green (because there is more probability mass “near” the red verification)?

Is this not merely a case of trying to evaluate a single probability forecast in isolation?

Is there an example where this can happen repeatedly, which is not a case where smoothing the forecast pdf and using a local score would be advantageous?
Parameter Estimation via Probability Forecasts

Consider parameter estimation in the case of the logistic map \( F(x) = a x (1-x) \)

We can evaluate the skill of our entire ensemble prediction system as a whole by examining the skill of our forecast as a function of \( a \).

Unless the prediction system is perfect, this need not return the true value of \( a \).

If the model is structurally in error, there is no True value of \( a \).
Where should decision makers draw the line?

Climate models are based on well-established physical principles and have been demonstrated to reproduce observed features of recent climate (see Chapters 8 and 9) and past climate changes (see Chapter 6). There is considerable confidence that Atmosphere-Ocean General Circulation Models (AOGCMs) provide credible quantitative estimates of future climate change, particularly at continental and larger scales. Confidence in these estimates is higher for some climate variables (e.g., temperature) than for others (e.g., precipitation). This summary highlights areas of progress since the TAR:

Clear, plain spoken discussion of what today’s models cannot capture quantitatively would be of great value.
So what about UKCP probabilities?
What is the chance of falling above the 90% line of UKCP PDFs?

The probability of the real world falling above the 90% line of the UKCP PDF can be much greater than 10%.

The shortcoming of climate models are more clearly acknowledged in the peer reviewed literature than in the UKCP user guidance.

It is important to stress that our approach to the specification of discrepancy can only be expected to capture a subset of possible structural modelling errors and should be regarded as a lower bound. This is because models tend to share certain common systematic biases, which can be found in diverse elements of climate including multiannual means of basic quantities such as surface temperature,
So what about UKCP probabilities?
What is the chance of falling above the 90% line of UKCP PDFs?

The probability of the real world falling above the 90% line of the UKCP PDF can be much much greater than 10%.

The shortcoming of climate models are more clearly acknowledged in the peer reviewed literature than in the UKCP user guidance.

It is important to stress that our approach to the specification of discrepancy can only be expected to capture a subset of possible structural modelling errors and should be regarded as a lower bound. This is because models tend to share certain common systematic biases, which can be found in diverse elements of climate including multiannual means of basic quantities such as surface temperature,
Statistical post-processing: These are anomalies, not temperatures. Parameterization of cloud formation is a bit of a distraction when we are missing two kilometre tall walls of rock...
Does a big model add?

?Confidence?

?Insight?

?Numbers?

http://www.metoffice.gov.uk/climatechange/guide/what/
Probability and Weather-like Simulations

Here we usually have a “unconditioned distribution” called “climatology”

\[ P_{\text{clim}}(x|\text{archive}) \]

It is “unconditioned” inasmuch as it does not take into account our knowledge of the current state of the system.

In addition we have a number of model simulations that are conditioned on current observations, \( s \), and of course on the model used.

\[ P_i(x|s, M_i) \]

Given that we know the model is inadequate, this notation is already odd…

But in the weather case, we use the forecast-verification archive to determine when the model has less skill than climatology, and thereby cope with model inadequacy.
Probability in Climate-like Simulations

Here the problem is one of extrapolation, we have no relevant forecast-verification archive and we do not know the climatology (that being our aim!)

Arguably, the best we can hope for is agreement between our models (in distribution) which suggests the details do not (better: did not) matter.


We do not have that sort of agreement today, even for the planet’s temperature.

And we have a hodge-podge of various types of “uncertainties” mixed together and shown to policy makers on the same graph:
Reference: Probability level

Probability level

Describes the strength of evidence associated with a given value within a probabilistic climate projection. Probabilistic climate projections fall under subjective probability as the probabilities are a measure of the degree to which a particular level of future climate change is consistent with the evidence considered. In the case of UKCP09, the evidence comes from observations and outputs from a number of climate models, all with their associated uncertainties.

In detail

One of the main advances associated with UKCP09 is that it provides probabilistic climate projections. This means that different future climate outcomes described by a probabilistic projection have different strengths of evidence associated with them. As such, probability levels associated with a given change should be interpreted as indicating the relative likelihood of the projected change being at or less than the given change.

For example, if a projected temperature change of +4.5°C is associated with the 90% at a particular location in the 2080s for the UKCP09 medium emission scenario, this should be interpreted as it is projected that there is a 90% likelihood that temperatures at that location will be up to and including 4.5°C warmer than temperatures in the 1961–90 baseline period. Conversely, there is a 10% likelihood that those temperatures will be at or greater than 4.5°C warmer than the baseline period.
Misleading argument 5: ‘Global warming computer models which predict the future climate are unreliable’

Computer models which predict the future climate are unreliable and based on a series of assumptions

What does the science say?

Modern climate models have become increasingly accurate in reproducing how the real climate works. They are based on our understanding of basic scientific principles, observations of the climate and our understanding of how it functions.

By creating computer simulations of how different components of the climate system - clouds, the Sun, oceans, the living world, pollutants in the atmosphere and so on - behave and interact, scientists have been able to reproduce the overall course of the climate in the last century. Using this understanding of the climate system, scientists are then able to project what is likely to happen in the future, based on various assumptions about human activities.

It is important to note that computer models cannot exactly predict the future, since there are so many unknowns concerning what might happen. Scientists model a range of future possible climates using different scenarios of what the world will ‘look like’. Each scenario makes different assumptions about important factors such as how the world’s population may increase, what policies might be introduced to deal with climate change and how much carbon dioxide and other greenhouse gases humans will pump into the atmosphere. The resulting projection of the future climate for each scenario, gives various possibilities for the temperature but within a defined range.

While climate models are now able to reproduce past and present changes in the global climate rather well, they are not, as yet, sufficiently well-developed to project accurately all the detail of the impacts we might see at regional or local levels. They do, however, give us a reliable guide to the direction of future climate change. The reliability also continues to be improved through the use of new techniques and technologies.
Bayesian probability calculus

Main article: Bayes' Theorem

Given some data and some hypothesis, the posterior probability that the hypothesis is true is proportional to the product of the likelihood multiplied by the prior probability. For simplicity, the "prior probability" is often abbreviated as the "prior" and the "posterior probability" as the "posterior". The likelihood brings in the effect of the data, while the prior specifies the belief in the hypothesis before the data was observed.

More formally, Bayesian inference uses Bayes' formula for conditional probability:

\[
P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)}
\]

where

- \(H\) is a hypothesis, and \(D\) is the data.
- \(P(H)\) is the prior probability of \(H\), the probability that \(H\) is correct before the data \(D\) was seen.
- \(P(D|H)\) is the conditional probability of seeing the data \(D\) given that the hypothesis \(H\) is true. \(P(D | H)\) is called the likelihood.
- \(P(D)\) is the marginal probability of \(D\).
- \(P(H|D)\) is the posterior probability, the probability that the hypothesis is true, given the data and the previous state of belief about the hypothesis.

\(P(D)\) is the prior probability of witnessing the data \(D\) under all possible hypotheses. Given any exhaustive set of mutually exclusive hypotheses \(H_i\), we have:

\[
P(D) = \sum_i P(D, H_i) = \sum_i P(D|H_i) P(H_i).
\]

We can consider \(i\) here to index alternative worlds, of which there is exactly one which we inhabit, and \(H_i\) is the hypothesis that we are in the world \(i\).

\(P(D|H_i)\) is then the probability that we are in the world \(i\) and witness the data. Since the set of alternative worlds was assumed to be mutually exclusive and exhaustive, the above formula is a case of the law of total probability.

\(P(D)\) is the normalizing constant, which in many cases need not be evaluated. As a result, Bayes' formula is often simplified to:

\[
P(H|D) \propto P(D|H) \cdot P(H)
\]

where \(\propto\) denotes proportionality.