Weather roulette

On the socio-economical value of meteorological forecasts: theory and practice

The real value of meteorological forecasts derives from using them. In this talk, we will look at how weather forecasts can be used to make better business decisions today. This involves learning a bit of forecasting theory and a bit of decision theory. And the way we use the forecasts depends on what decisions we are making: is our aim to make money with weather derivatives? or to reduce CO2 emissions by more efficiently using wind and gas power? Meteorological forecasts tend to be of value in the health sector as well, and we will examine the difficulties that appear when we try to use forecasts that range farther into the future, next month rather than next week. Finally, we will touch on the socio-economical value of climate modelling, can see how you can contribute to this effort using your own PC.

On the socio-economical value of meteorological forecasts:

Theory and Practice

Leonard Smith
Centre for the Analysis of Time Series
London School of Economics
Pembroke College, Oxford

www.lsecats.org

Things to remember!

differs from

Data	^	Information
Accuracy	^	Utility
Error	^	Uncertainty
Simulation	^	Forecast
Forecast	^	Decision Support
Probability	^	Probabilistic
(theory)	^	(practice)

Strategic: What do we need in order to…
Tactical: What can we do with what we've got…

Warnings & Alerts given to humans change the way that they will react to future Warnings & Alerts…

Roulston & Smith (2002) The Boy Who Cried Wolf Revisited, *Weather and Forecasting* **19** (2): 391–397

Overview

What is the Aim? ("What is the Product?")
Why Ensemble Forecasts?
Risk Management, Decision Support and Forecasts
How Ensemble Prediction Systems (EPS) Add Value
Socio-economic valuation of forecast information:
    Proofs of Concept, of Information, and of Value

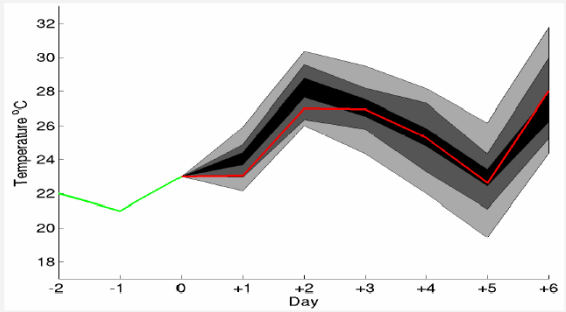
What is the Product? (And what is it worth?)

your time, energy, cash

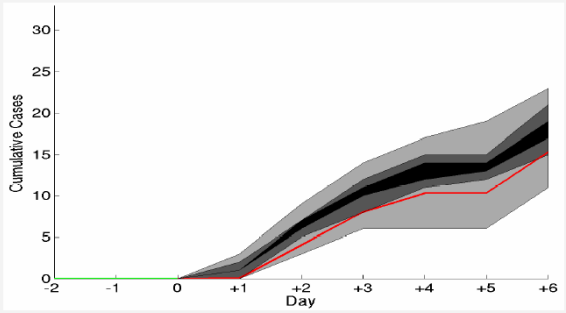
Weather:

Case Load:

LSE-CATS Weather-Health Daily Risk Management Page



Day	-2	-1	0	+1	+2	+3	+4	+5	+6
Temp	21	20	22	21	23	23	27	27	25
Thresh	23	23	23	23	23	23	23	23	23
Cases	0	0	0	-	-	-	-	-	-
Max Exp Cases	0	0	0	3	7	7	6	3	8
Exp Cases	0	0	0	1	5	4	2	1	4



A

©L.A. Smith

Is this relevant?
Is it accurate?
Is it useful?

How does it work?

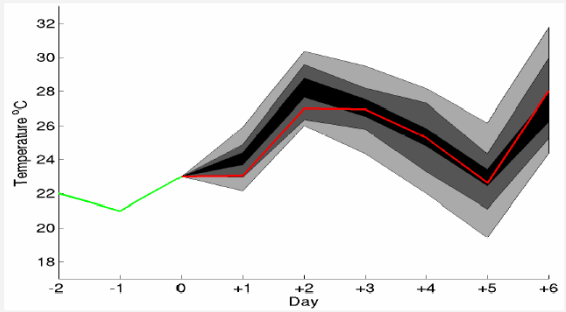
An Example in Healthcare.

What is the Product? (And what is it worth?)

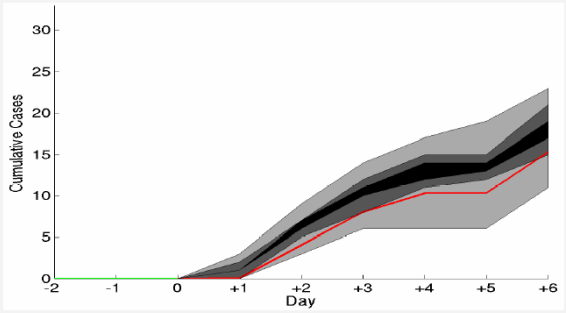
Weather:

Case Load:

LSE-CATS Weather-Health Daily Risk Management Page



Day	-2	-1	0	+1	+2	+3	+4	+5	+6
Temp	21	20	22	21	23	23	27	27	25
Thresh	23	23	23	23	23	23	23	23	23
Cases	0	0	0	-	-	-	-	-	-
Max Exp Cases	0	0	0	3	7	7	6	3	8
Exp Cases	0	0	0	1	5	4	2	1	4



A

An Example in Healthcare.

Is this relevant?

But you have to ask the last two questions down here!

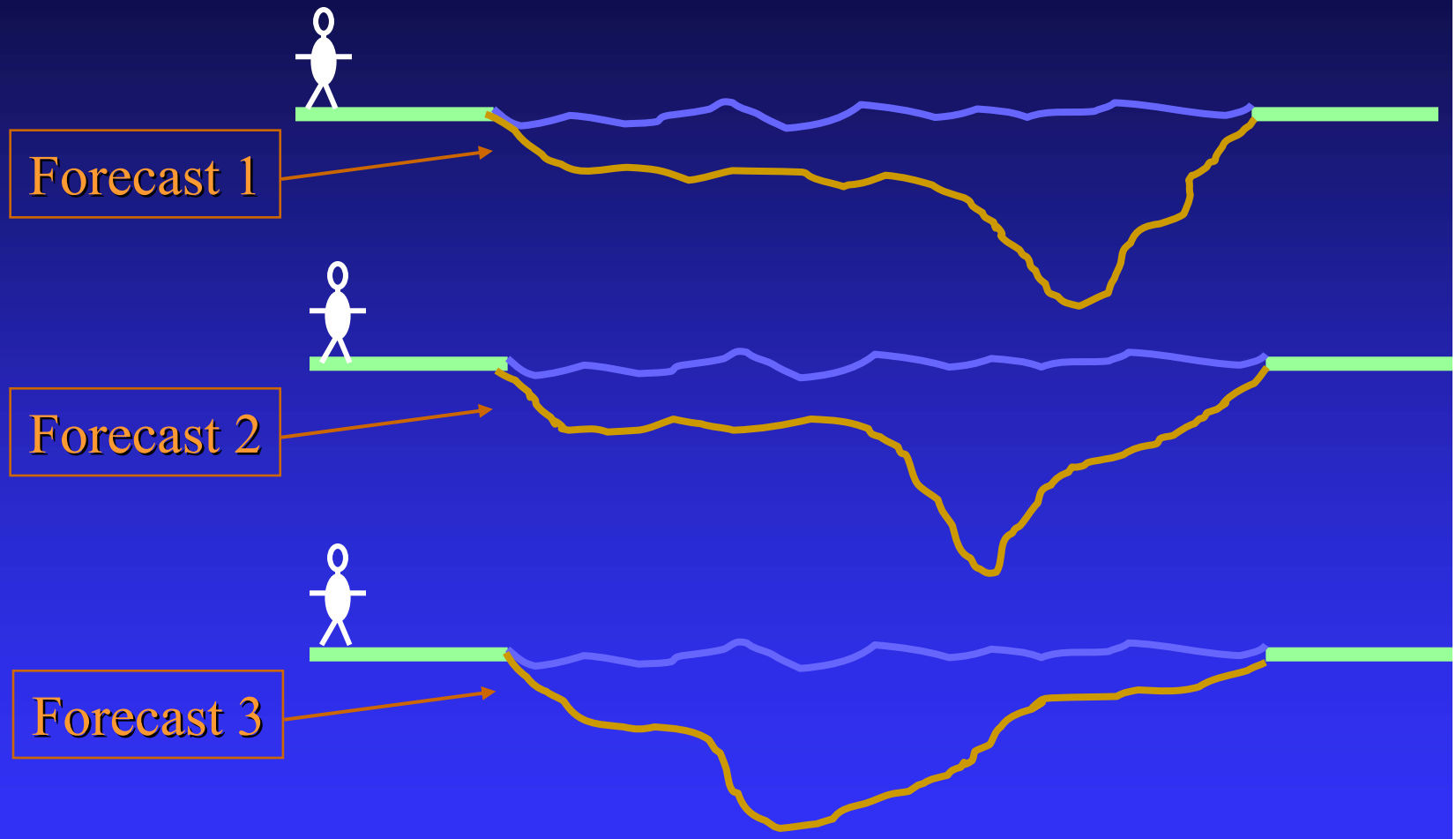
Is it accurate?
Is it useful?

We will come back to a schematic of this figure in the seasonal context. First:
How does it work?

The parable of the three statisticians.

Three non-Floridian statisticians come to a river, they want to know if they can cross safely. (They cannot swim.)

Three non-Floridian statisticians wish to cross a river.
Each has a forecast of depth which indicates they will drown.



So they have an ensemble forecast,with three members

Three non-Floridian statisticians wish to cross a river.
Each has a forecast of depth which indicates they will drown.
So they average their forecasts and decide based on the ensemble mean…

Ensemble mean

Is this a good idea?

ouse Meteo-France

No!

Ensembles may have lots of information, we must be careful not to destroy or discard it!

Note that, as in health risk-management, the statisticians:

- have a nonlinear utility function
  including a very asymmetric risk/utility function
  (overly deep by 2 inches << shallow by 2 inches)
- do not care about the river depth *per se*.

Take Home Message:

If you have an ensemble, use it. (The ensemble mean is meaningless!)

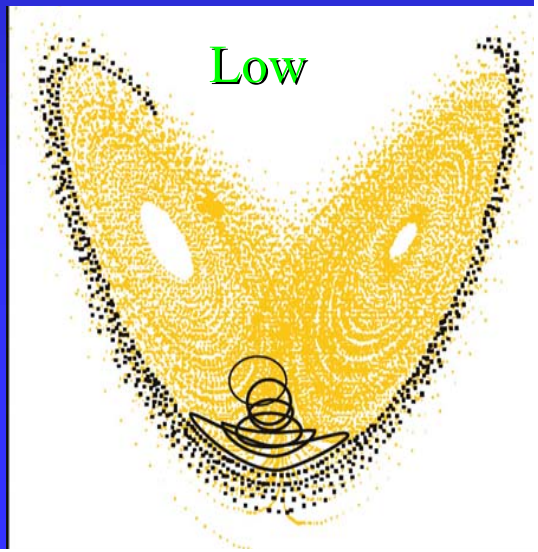
Take Home Message:

If you have an ensemble, use it.

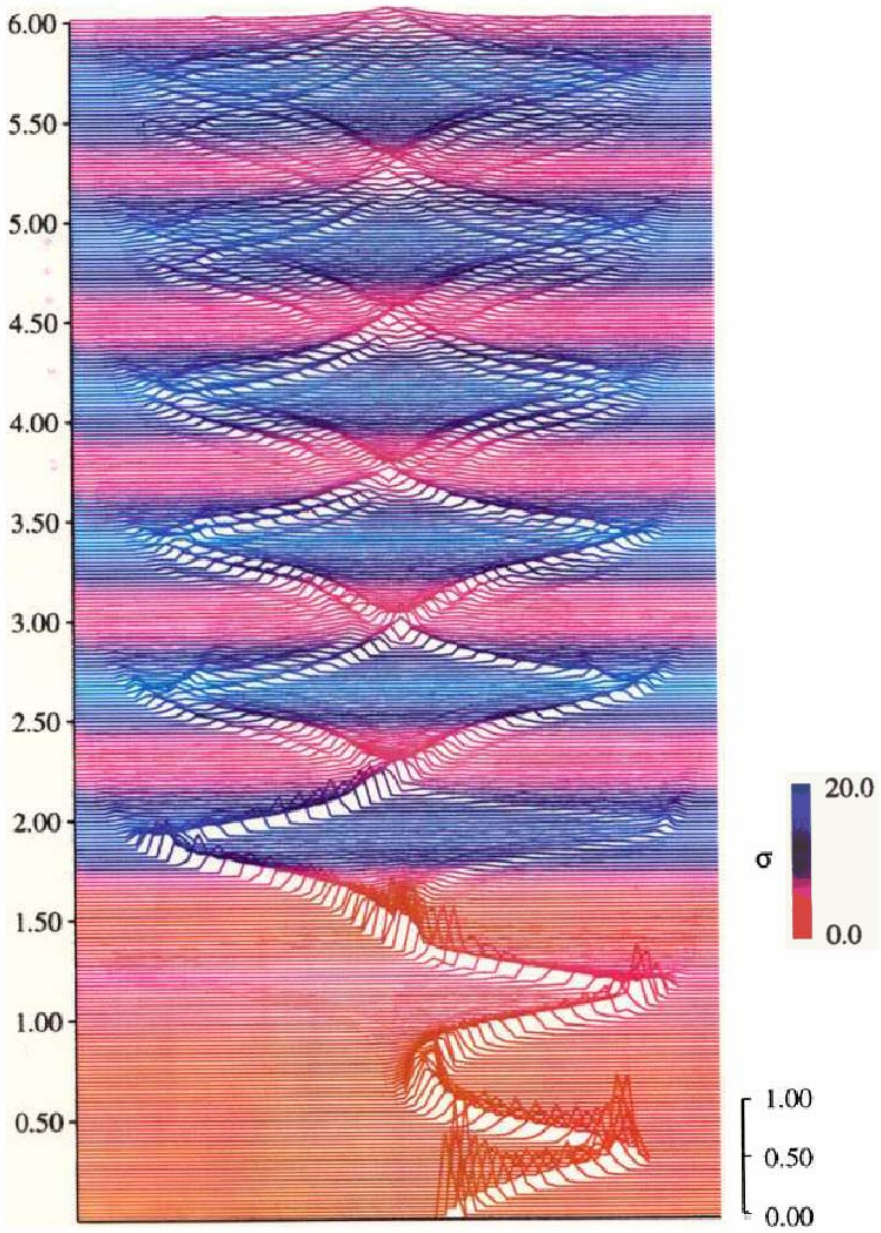
The relevant question is one of decision support, not forecasting.

Predictability: where does it come from?

Pictures from Tim Palmer



We would like to quantify day to day variations in predictability with probability forecasts…



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

The evolution of an ensemble tells us all we can know of the future, give what we know of the present.

This is nothing new, really.

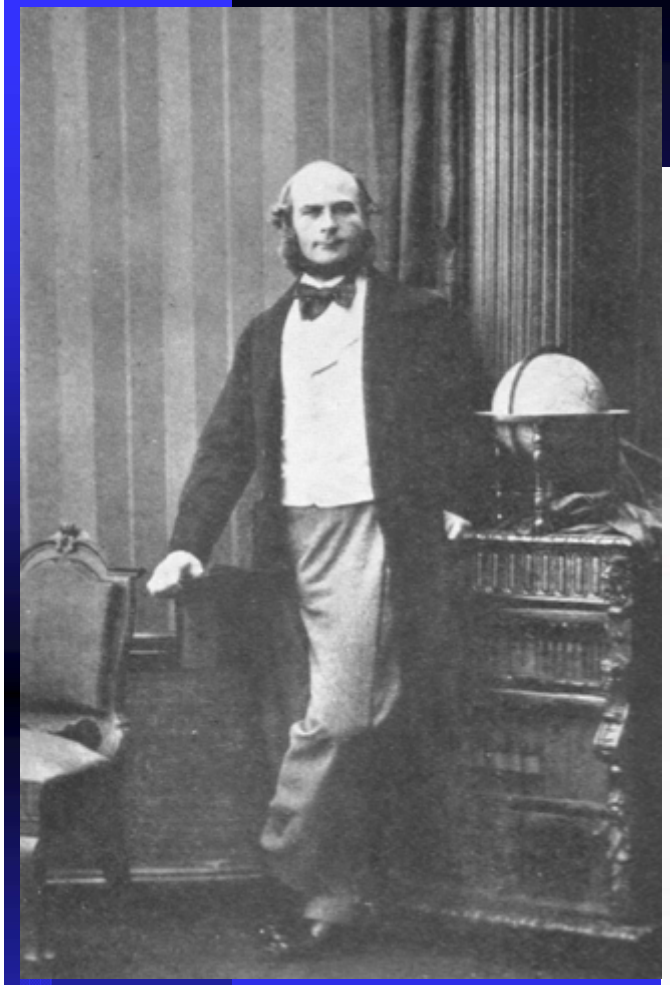
But the key change in thinking is to STOP thinking about "accuracy" and instead think about information.

There can be useful information in an ensemble weather forecast even if it merely confirms that we do NOT yet know what will happen.

They can help us see further ahead into the future.

Health-weather relationships can provide qualitative information, and potentially quantitative information.

But first, one more example of what an ensemble is, and what to keep in mind…



This Galton Board is a mathematical model.

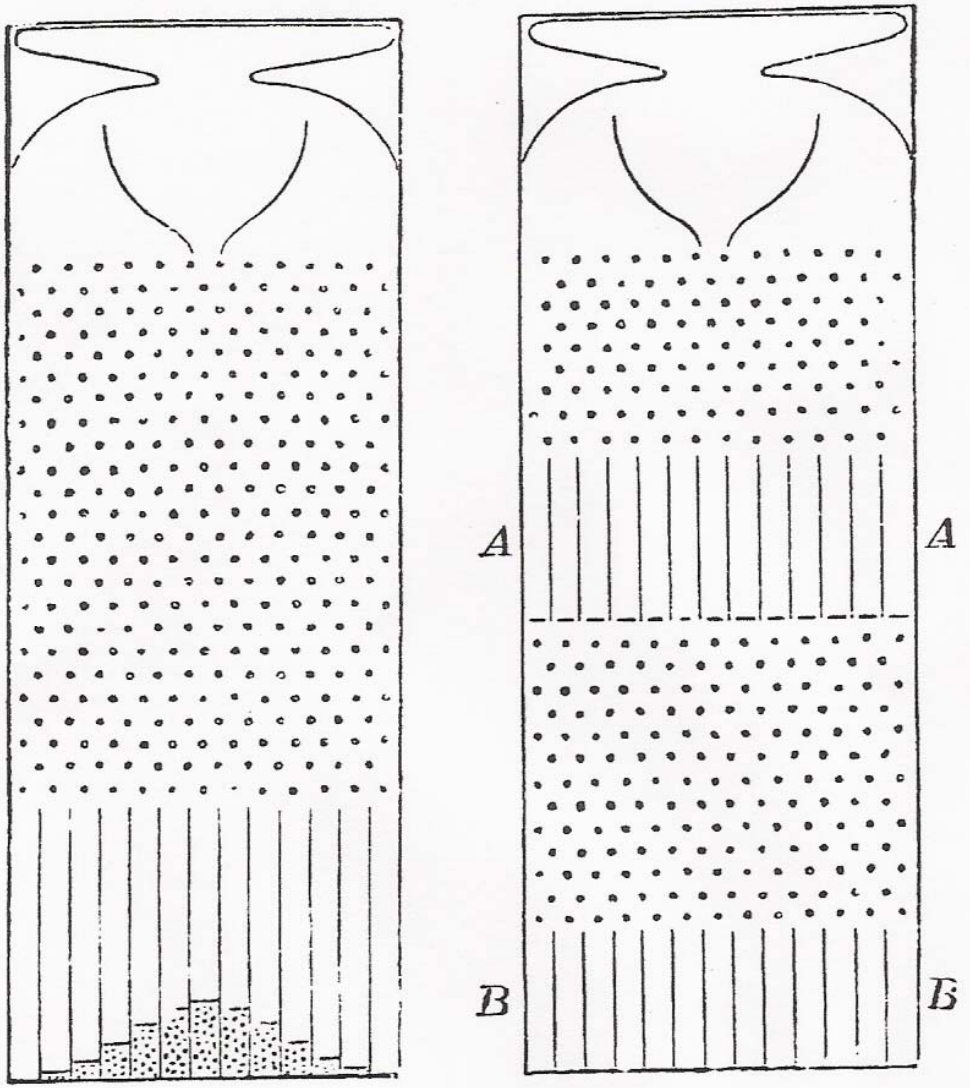
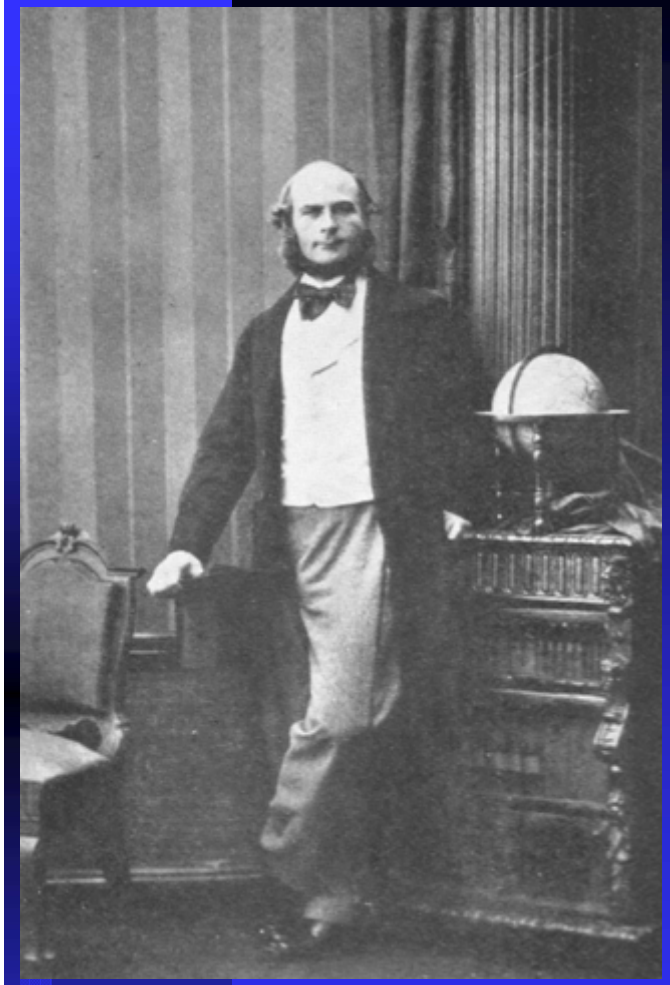


Figure 9.2 A schematic drawing of Galton's Quincunx, from Galton (1889a, p. 63).

2 Feb 2006



While this is Not A Galton (NAG) Board.
It is neither stochastic or chaotic; but at least it is!

This is a NAG Board

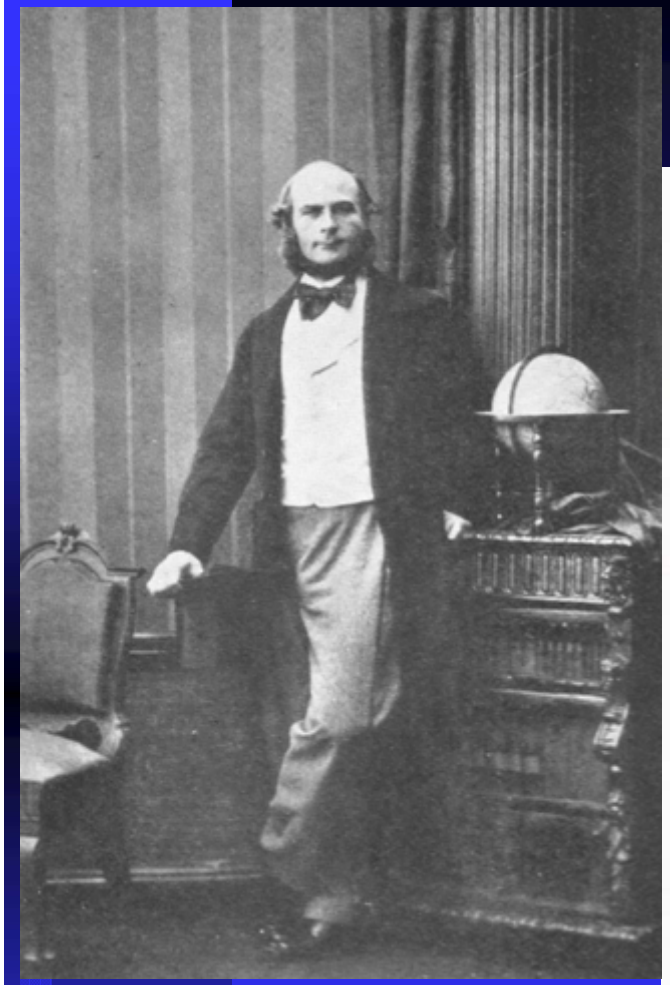
The NAG Board (Not a Galton Board)

Uncertainty in the NAG board corresponds to predicting with a collection (ensemble) of golf balls…

Ensembles inform us of uncertainty growth *within our model!*

But reality is not a golf-ball; this EPS must deal with model inadequacy.

Nevertheless, weather EPS are useful!
Operational Day ~10 Weather Ensembles:
US and European Services: 1992
Canada: Now



I term this a thought experiment because, while Galton clearly in several places described the variant of the Quincunx that performed the experiment, there is no indication that he actually built the apparatus. And having tried to build such a machine, I can testify that it is exceedingly difficult to make one that will accomplish the task in a satisfactory manner.

Stigler, 1999

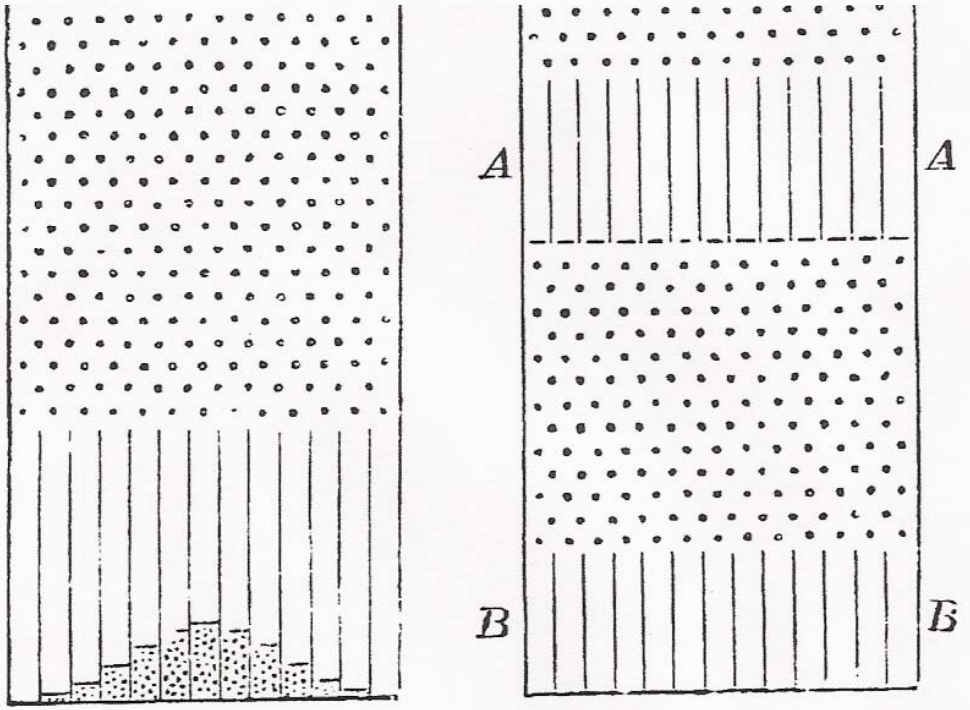
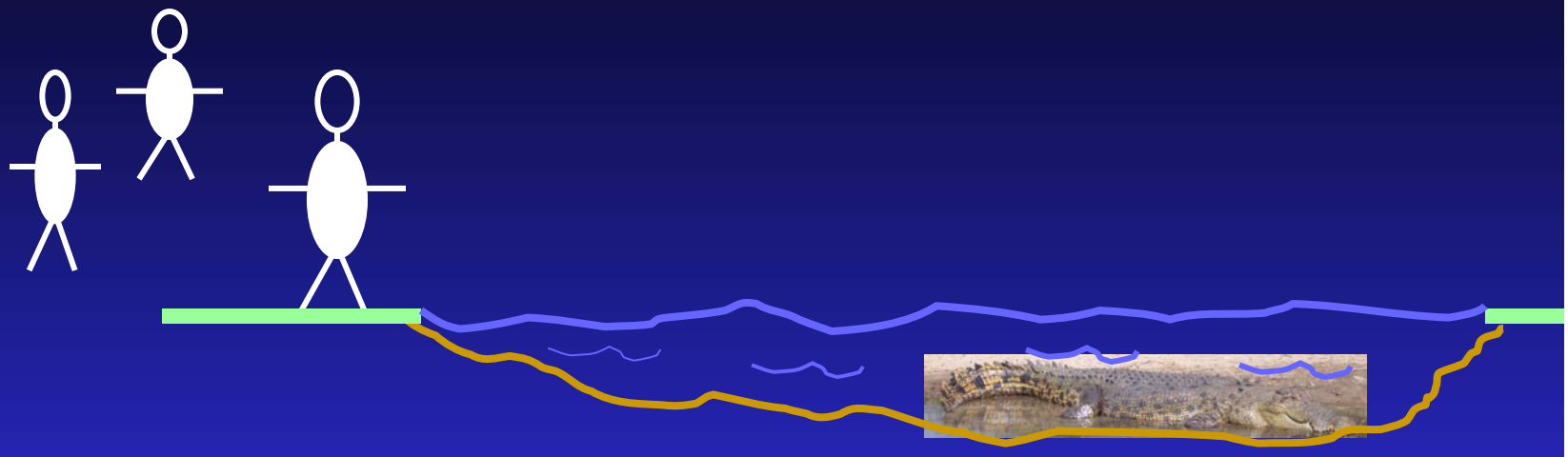


Figure 9.2 A schematic drawing of Galton's Quincunx, from Galton (1889a, p. 63).

2 Feb 2006

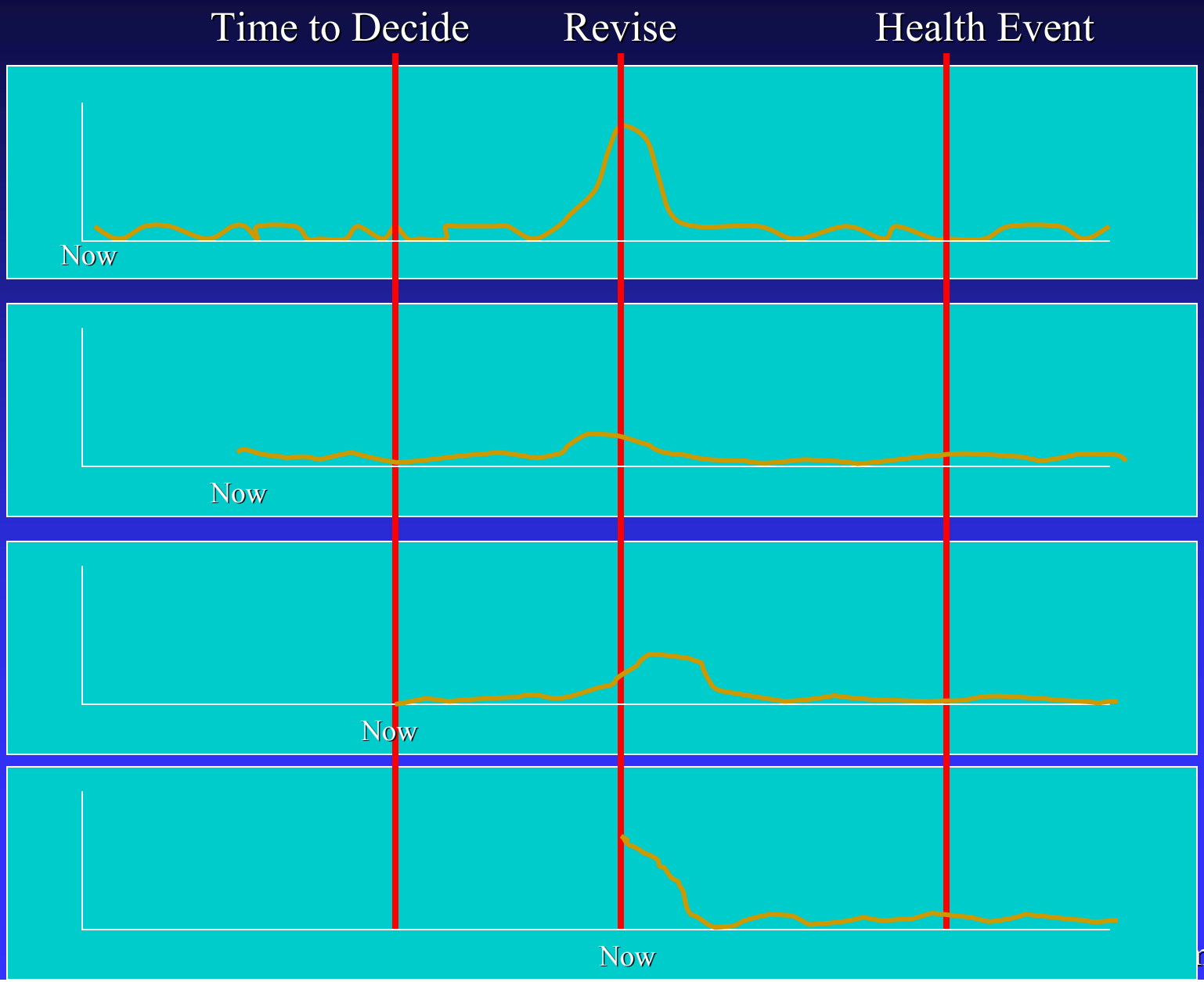
Model Inadequacy and our three non-Floridian statisticians.



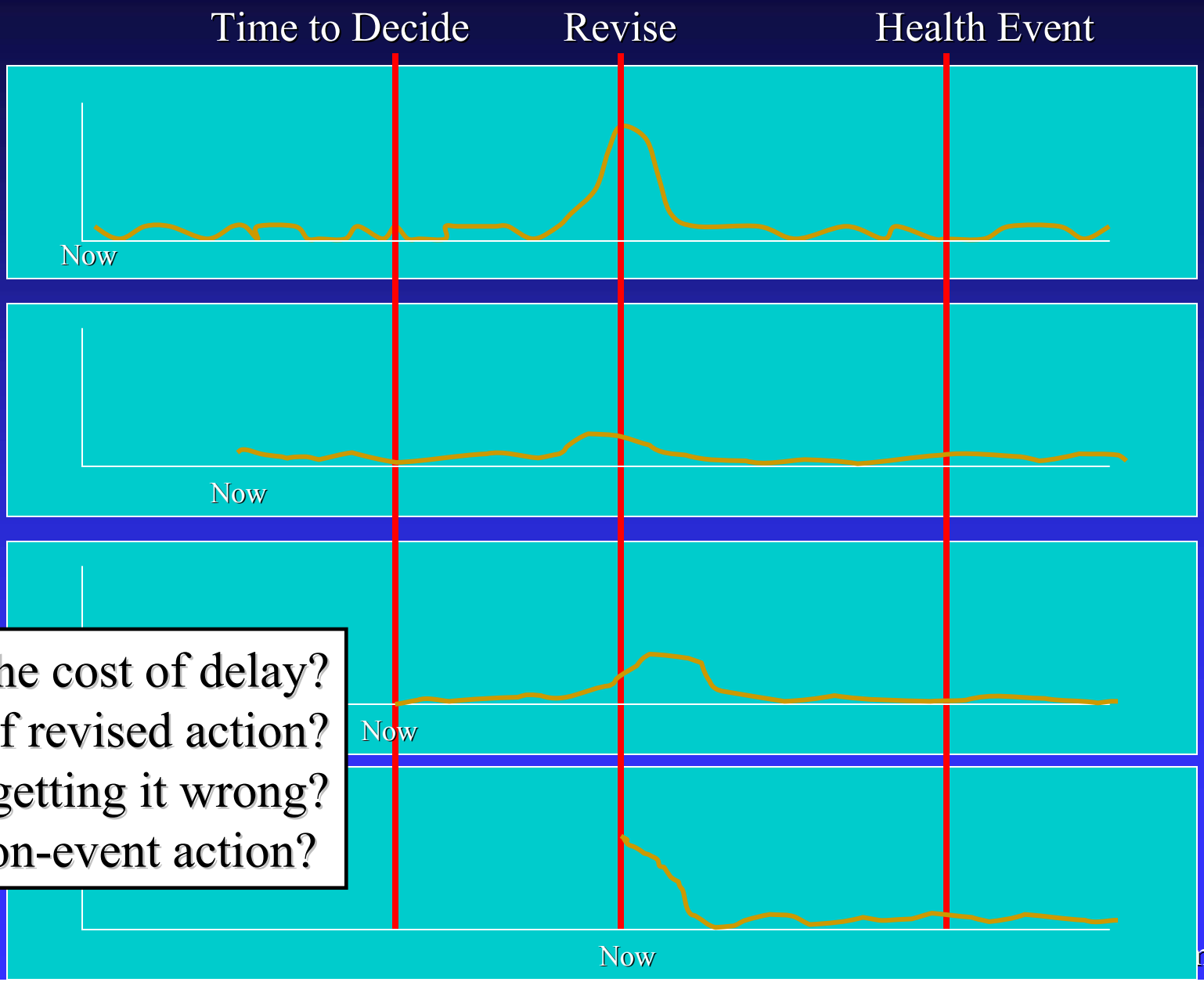
As it turns out, the river is rather shallow.
Model inadequacy covers things in the system but left out of the model.

The real question was could they make it across, the depth of the river was only one component…

Decision Support and Forecasts

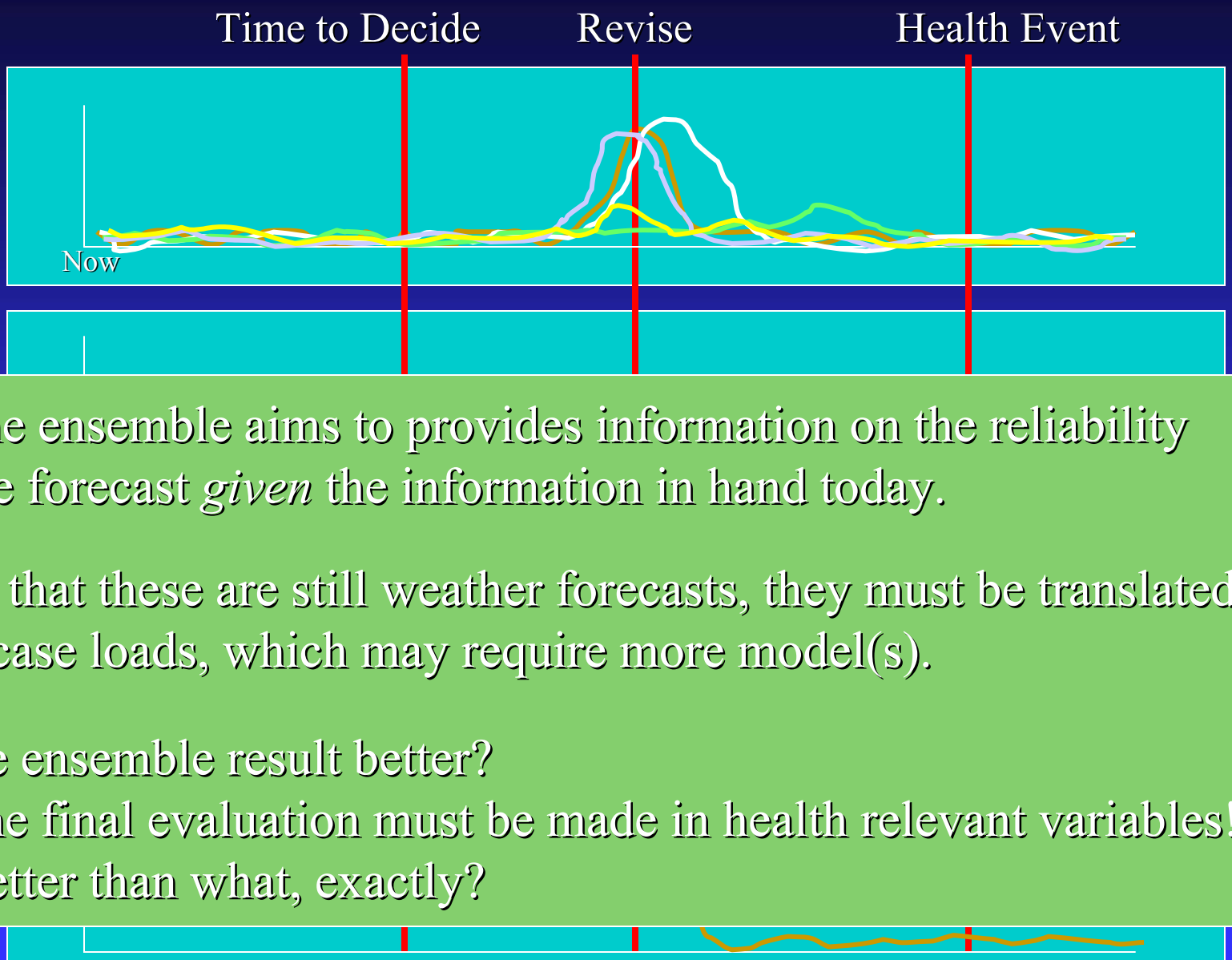


Decision Support and Forecasts



What is the cost of delay?
of revised action?
of getting it wrong?
of a non-event action?

Decision Support and Ensemble Forecasts

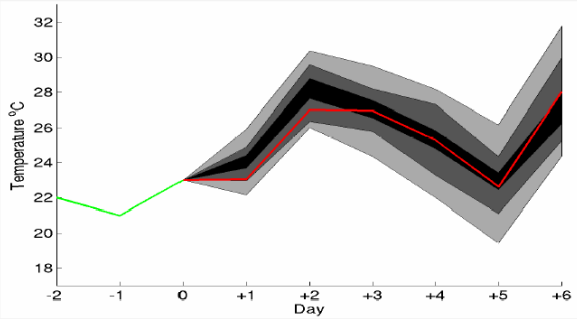


So the ensemble aims to provides information on the reliability of the forecast *given* the information in hand today.

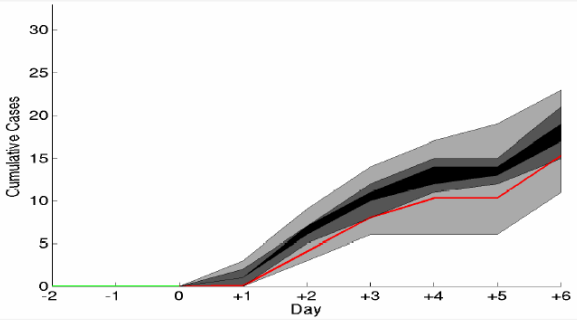
Note that these are still weather forecasts, they must be translated into case loads, which may require more model(s).

Is the ensemble result better?
a) The final evaluation must be made in health relevant variables!
b) Better than what, exactly?

LSE-CATS Weather-Health Daily Risk Management Page

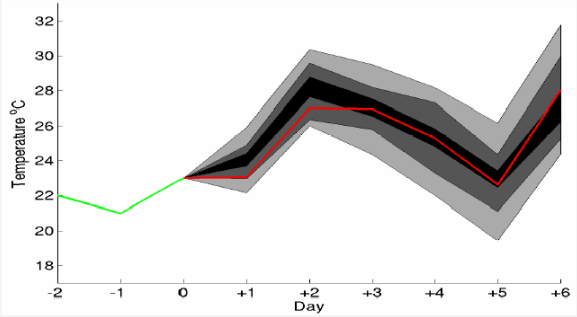


Day	-2	-1	0	+1	+2	+3	+4	+5	+6
Temp	21	20	22	21	23	23	27	27	25
Thresh	23	23	23	23	23	23	23	23	23
Cases	0	0	0	-	-	-	-	-	-
Max Exp Cases	0	0	0	3	7	7	6	3	8
Exp Cases	0	0	0	1	5	4	2	1	4

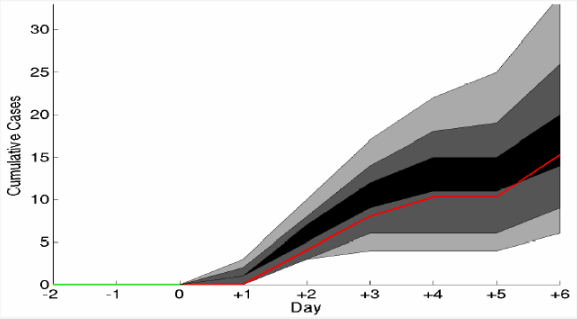


A

LSE-CATS Weather-Health Daily Risk Management Page



Day	-2	-1	0	+1	+2	+3	+4	+5	+6
Temp	21	20	22	21	23	23	27	27	25
Thresh	23	23	23	23	23	23	23	23	23
Cases	0	0	0	-	-	-	-	-	-
Max Exp Cases	0	0	0	3	7	7	6	3	9
Exp Cases	0	0	0	1	5	4	2	1	5



B

In this kind of decision support ensembles can yield:

we see the uncertainty in likely case load in both A & B

but we also see the substantially greater risk in B

This information depends on the full EPS, not just the probability of weather on each day.

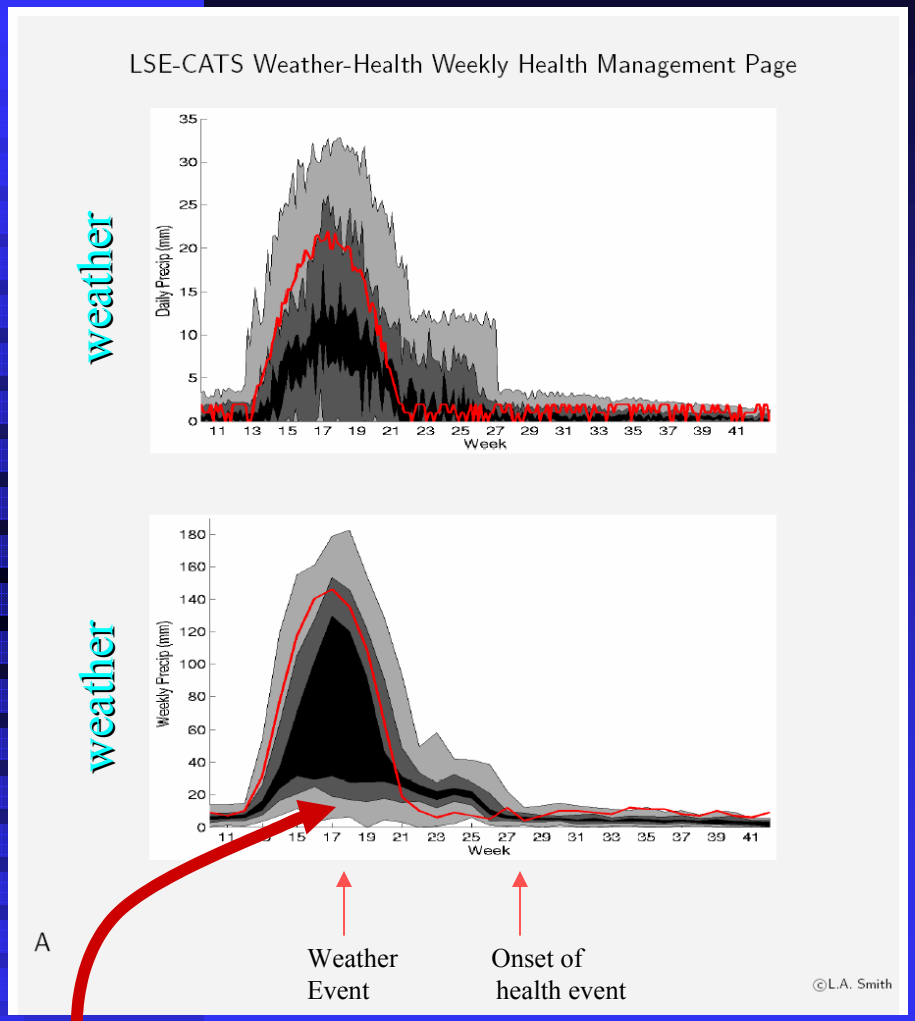
Take Home Messages:

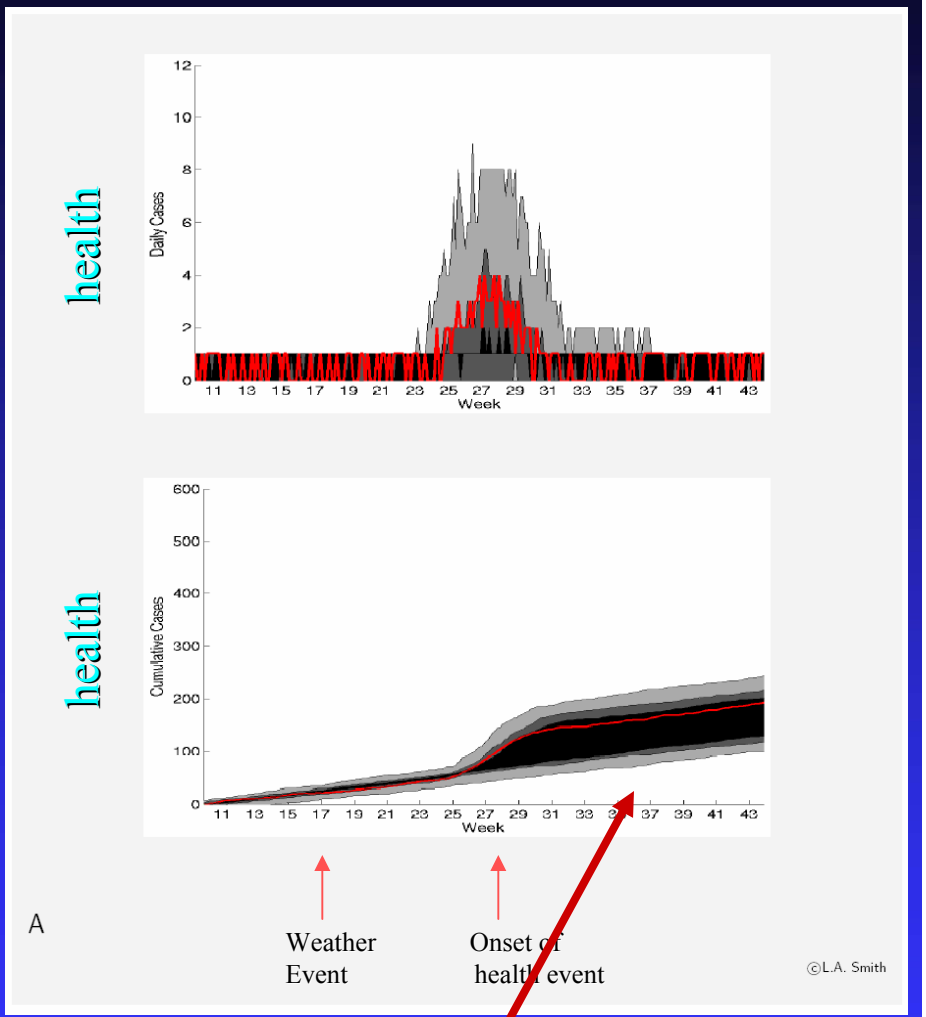
If you have an ensemble, use it.

The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

A Schematic Seasonal Example

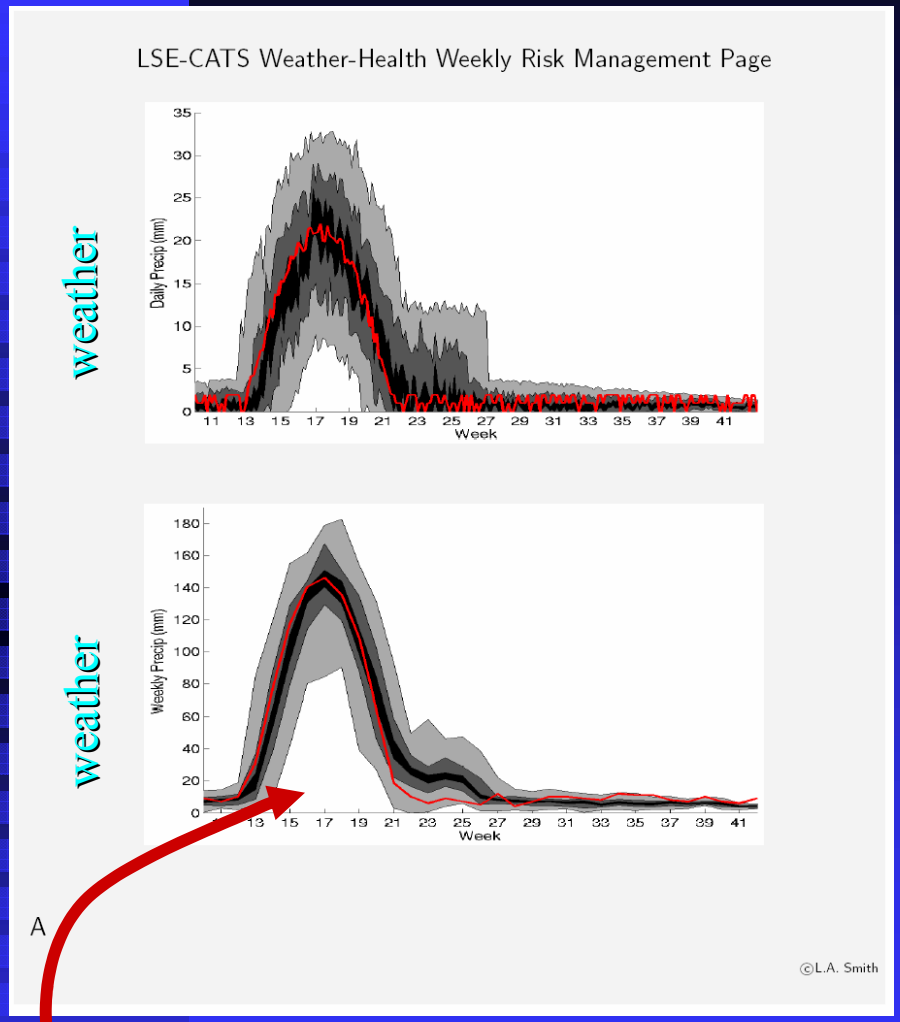


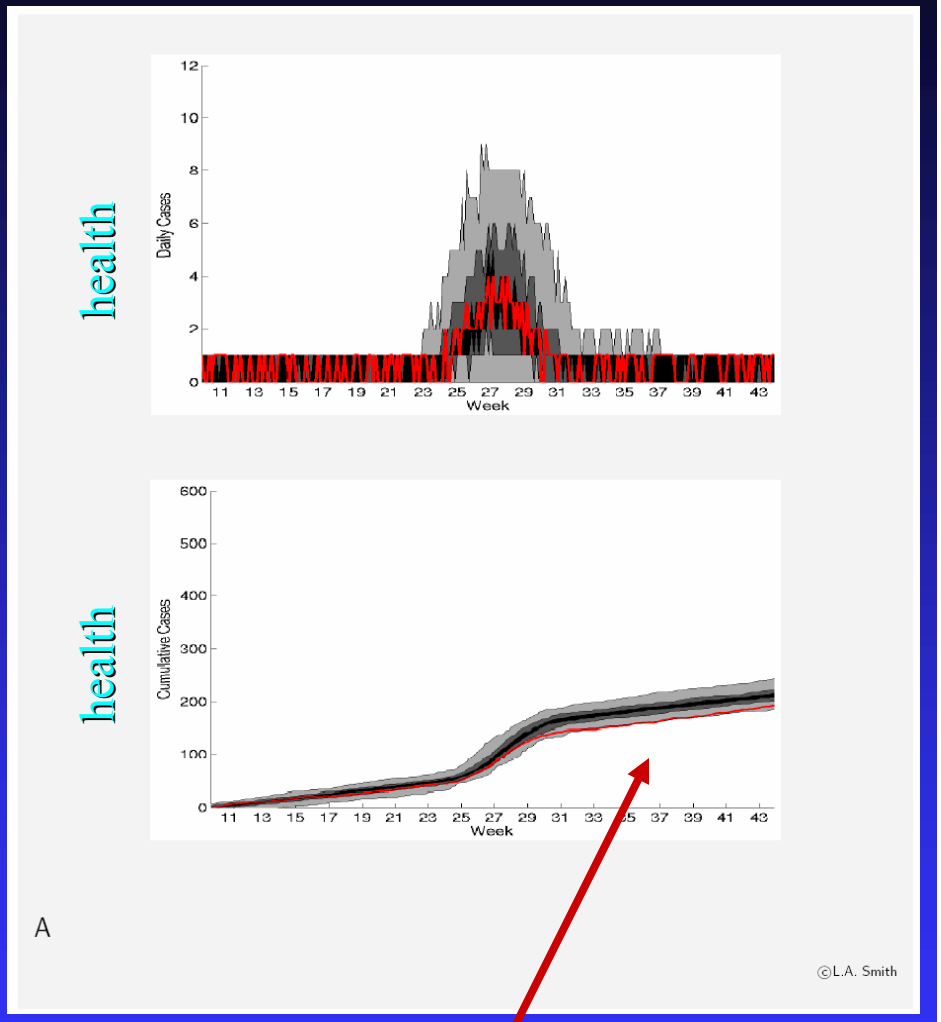


Note the delay between Precip on Onset

In this case we do not know if there will be on event, but at least we know that we do not know!

A Schematic Seasonal Example





In this case we know *the model* does "expect" an event.

Take Home Messages:

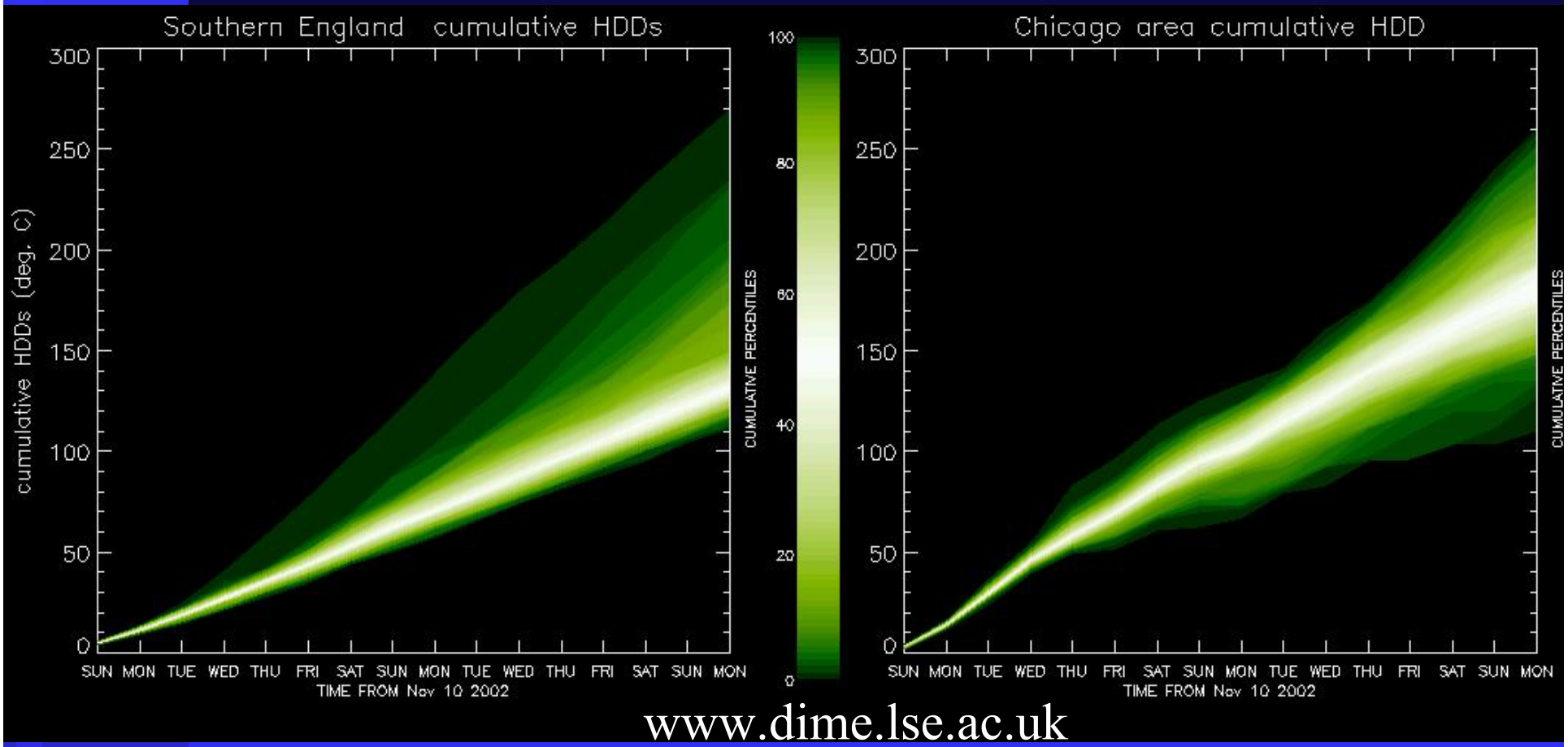
If you have an ensemble, use it.

The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

A good EPS can also indicate what the likely alternatives are, and thus assist in decision support.

Scenario-based cumulative HDD forecasts.



www.dime.lse.ac.uk

Note:

1) High Impact Forecasts need not include Severe Weather!

2) This cumulative information on total energy requirements is simply not in a single hi-res forecast.

Take Home Messages:

If you have an ensemble, use it.

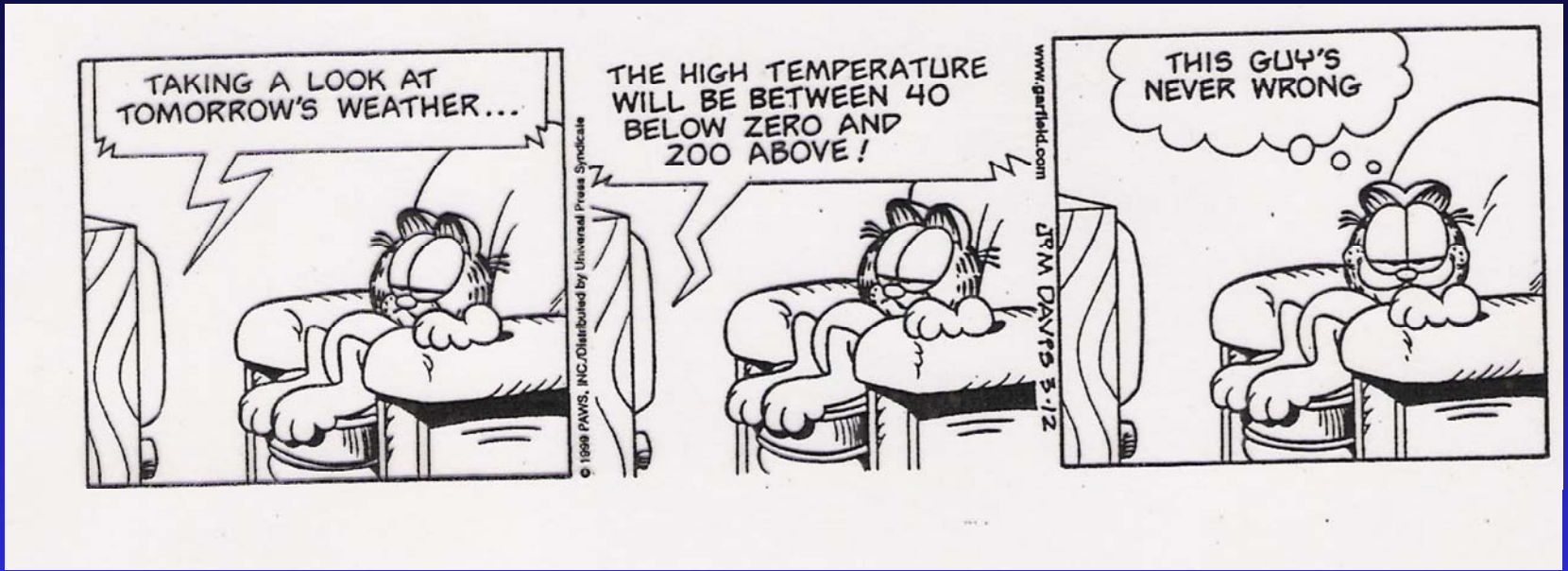
The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

A good EPS can also indicate what the likely alternatives are, and thus assist in decision support.

Focus on information content, *not* on meteorological accuracy.

Tactical usually means forecast usage



Aim: Deployable Probabilistic Forecasts with
Accountability, Resolution, and Relevance

Informative

Assigns non-trivial probabilities (to what happened, not what is the chance x happened

Suffers only from sampling finite N effects

But how do you evaluate a probability forecast?

What is a probability forecast?

What is a Probability Forecast?

Given:

- a complete, finite set of mutually exclusive events
- some symmetry assumptions

Then we can construct (empirically) useful probability forecasts.

What is a Probability Forecast?

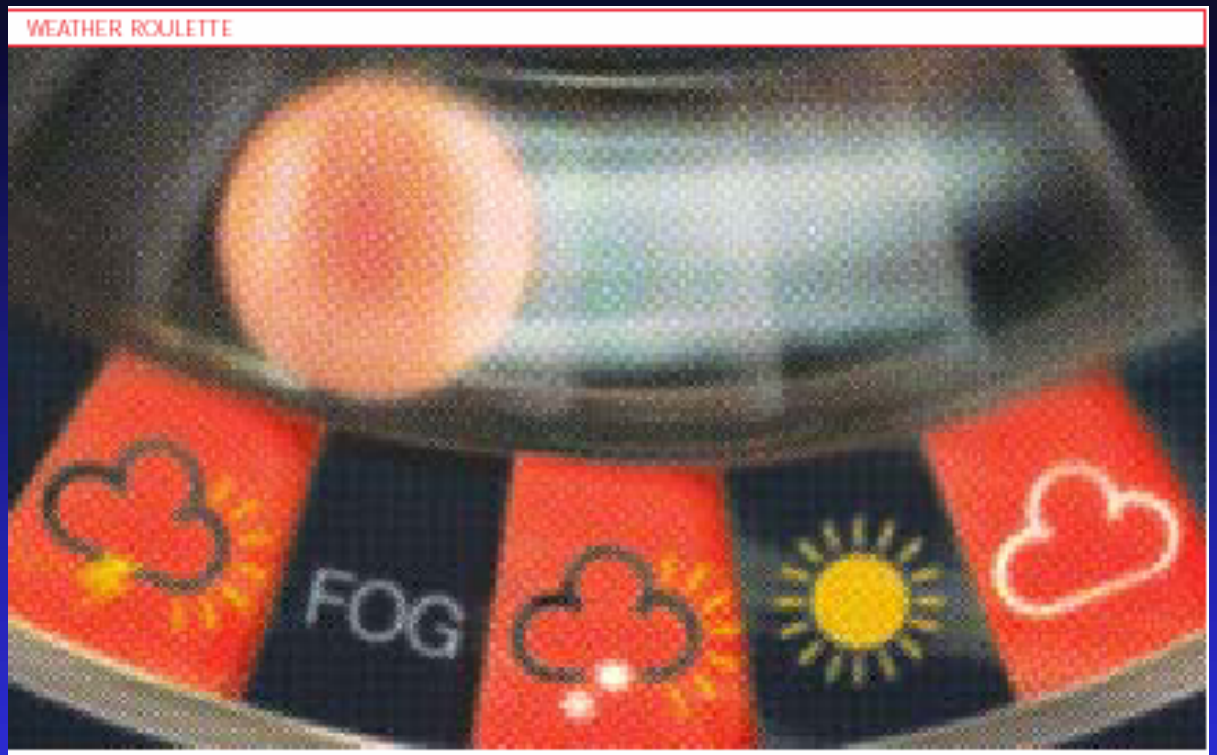
These are good assumptions for rolling dice:

Not so good for rolling gold bars!

Probabilities assigned to random events are rather different than probabilities which reflect only our ignorance.

It is best to bet on (or advise regarding) only the former!

Probability forecasts do not have to be accurate to be useful!



Weather roulette

Wager £100 each day on the temperature at Heathrow, betting an amount proportional to your predicted probability of that outcome (Kelly Betting).

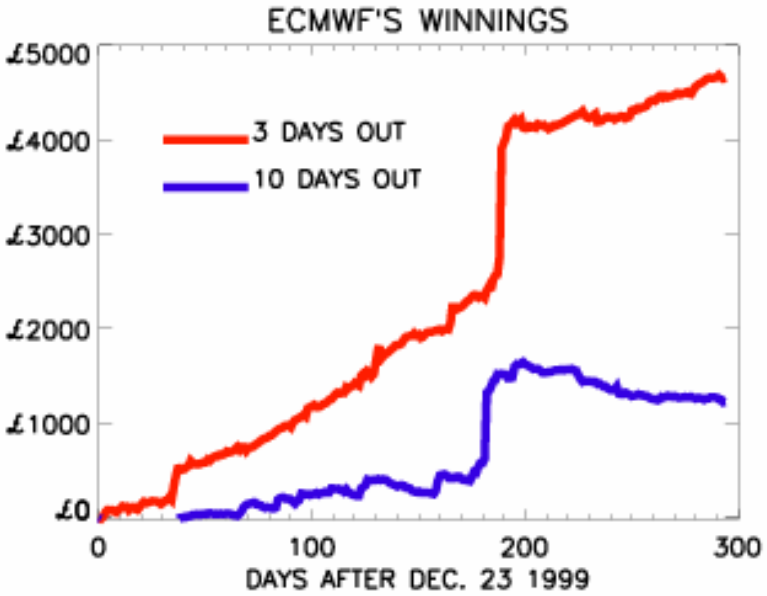
How would a probability forecast based on the ECMWF EPS fare against a house that set its odds using climatology?

WEATHER ROULETTE

TEMPERATURE AT HEATHROW

TABLE MAXIMUM: £100

1982–99 CLIMATOLOGICAL ODDS



TEMPERATURE (°C)

25	26	27	28	29
20	21	22	23	24
15	16	17	18	19
10	11	12	13	14
5	6	7	8	9
0	1	2	3	4
−5	−4	−3	−2	−1

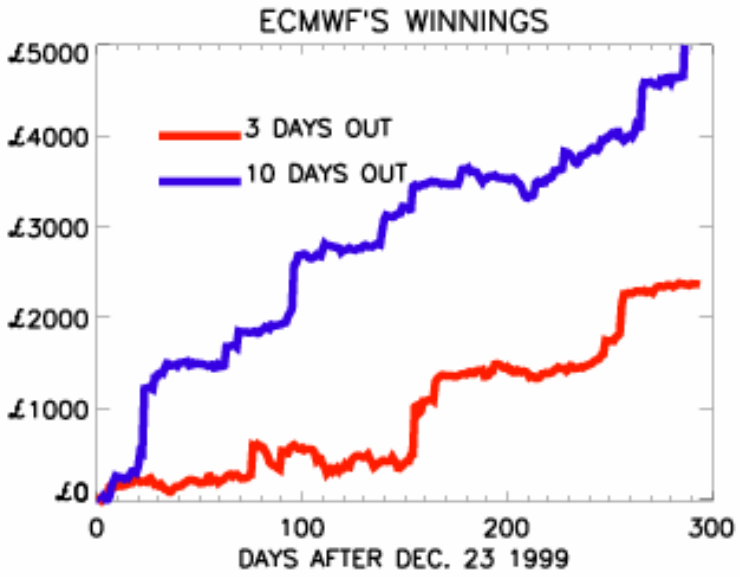
WEATHER ROULETTE

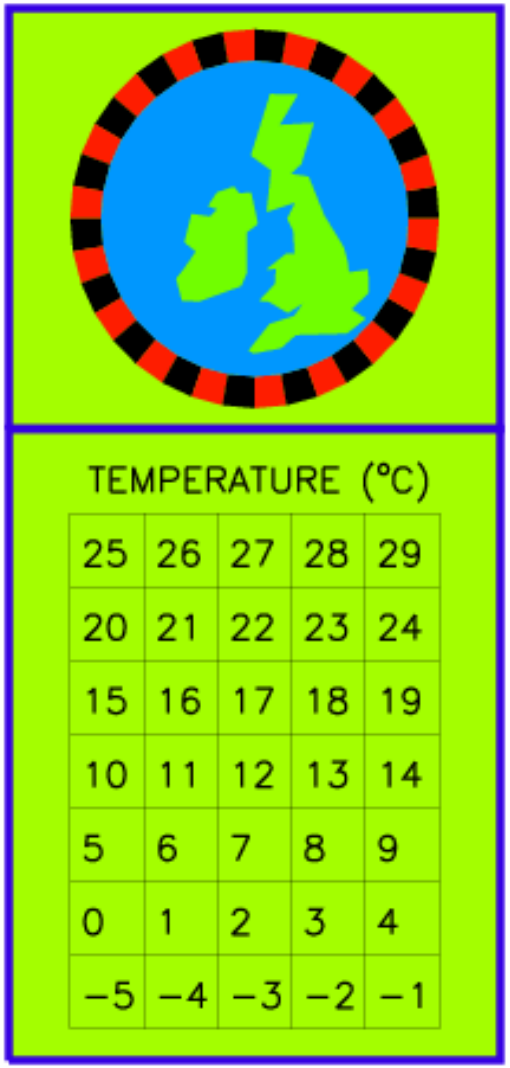
TEMPERATURE AT HEATHROW

TABLE MAXIMUM: £100

ODDS SET BY HIGH RES. FORECAST

BETS PLACED ACCORDING TO ENSEMBLE





We can measure this kind of skill with scores like –log(p).

BUT: we should do the evaluation close to the user!

The proof of the pudding is in the eating.

The proof of predictability is in the utility.

Establishing positive socio-economic benefit from an EPS usually takes four steps.

- Proof of Concept
  Timescales make sense (re: decision support)
  Historical/Theoretical causal connections OK.
  Models work on toy targets (internal consistency
- Proof of Information
  Forecasts contain relevant information for relevant *empirical* targets.
  Risk management scenario viable.
- Proof of Value
  End-to-end hindcasts on *actual target data*.
  Is the insight demonstrated worth more than the full cost of the decision support system?
- Real-time Demonstration System deployed and proven in real time.

Scientifically, success at each stage is interesting, valuable, and exciting.
From a users point of view, anything less than PoV is incomplete.

Take Home Messages:

If you have an ensemble, use it.

The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

A good EPS can also indicate what the likely alternatives are, and thus assist in decision support.

Focus on information content, not on meteorological accuracy.

Require "verification" on relevant, semi-independent, real target, observations!

Take Home Messages:

If you have an ensemble, use it.

The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

A good EPS can also indicate what the likely alternatives are, and thus assist in decision support.

Focus on information content, not on meteorological accuracy.

Require "verification" on relevant, semi-independent, real target, observations!

The goal is utility, not optimality. (Decision aid, not decision made)

Take Home Messages:

If you have an ensemble, use it.

The relevant question is one of decision support, not forecasting.

Ensembles are always valuable in nonlinear models, when they warn you that the model does NOT know what will happen.

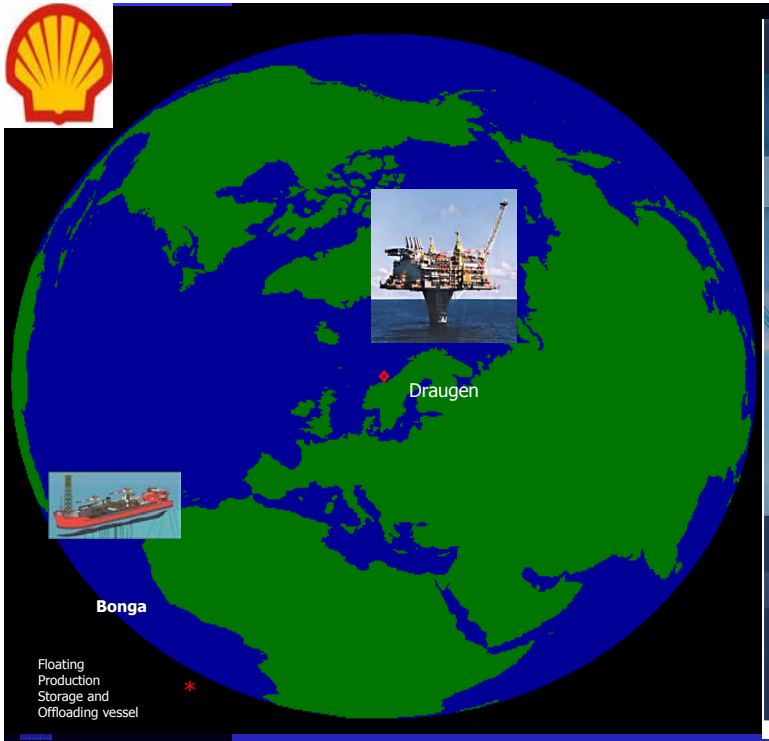
A good EPS can also indicate what the likely alternatives are, and thus assist in decision support.

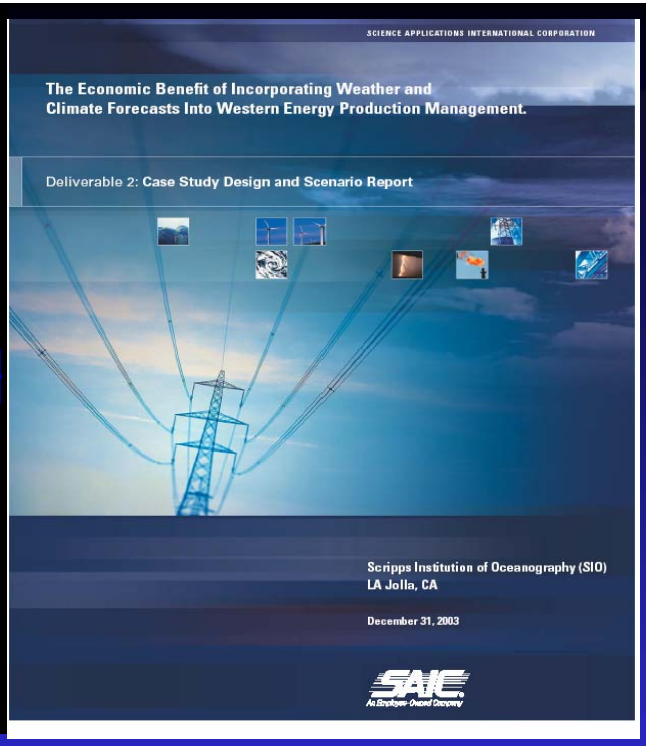
Focus on information content, not on meteorological accuracy.

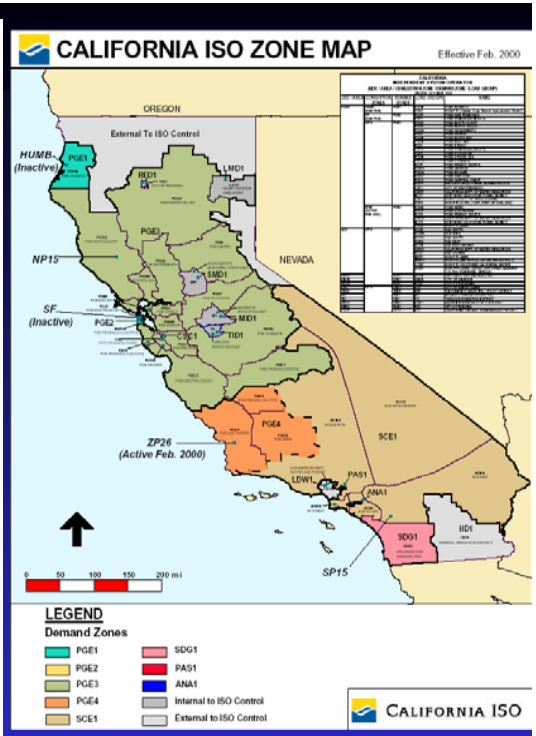
Require "verification" on relevant, semi-independent, real target, observations!

The goal is utility, not optimality. (Decision aid, not decision made)

If one forecast is good, then 50 forecasts will be better!
(but not 50 times better)







Seeing Through Weather Models

EDF ENERGY

CALIFORNIA ISO

SAIC

ALSTOM

DEPARTMENT OF THE NAVY ONR Science & Technology

METRA
METRA ENERGY

THE WEATHER CHANNEL
weather.com

LONDON ELECTRICITY

National Grid

British Energy

Risk Management Solutions

RMS™

PERGAMON

Renewable Energy 28 (2003) 585–602

RENEWABLE ENERGY

www.elsevier.com/locate/renene

Using medium-range weather forcasts to improve the value of wind energy production

M.S. Roulston [a,b,*], D.T. Kaplan [c], J. Hardenberg [a,b], L.A. Smith [a,b]

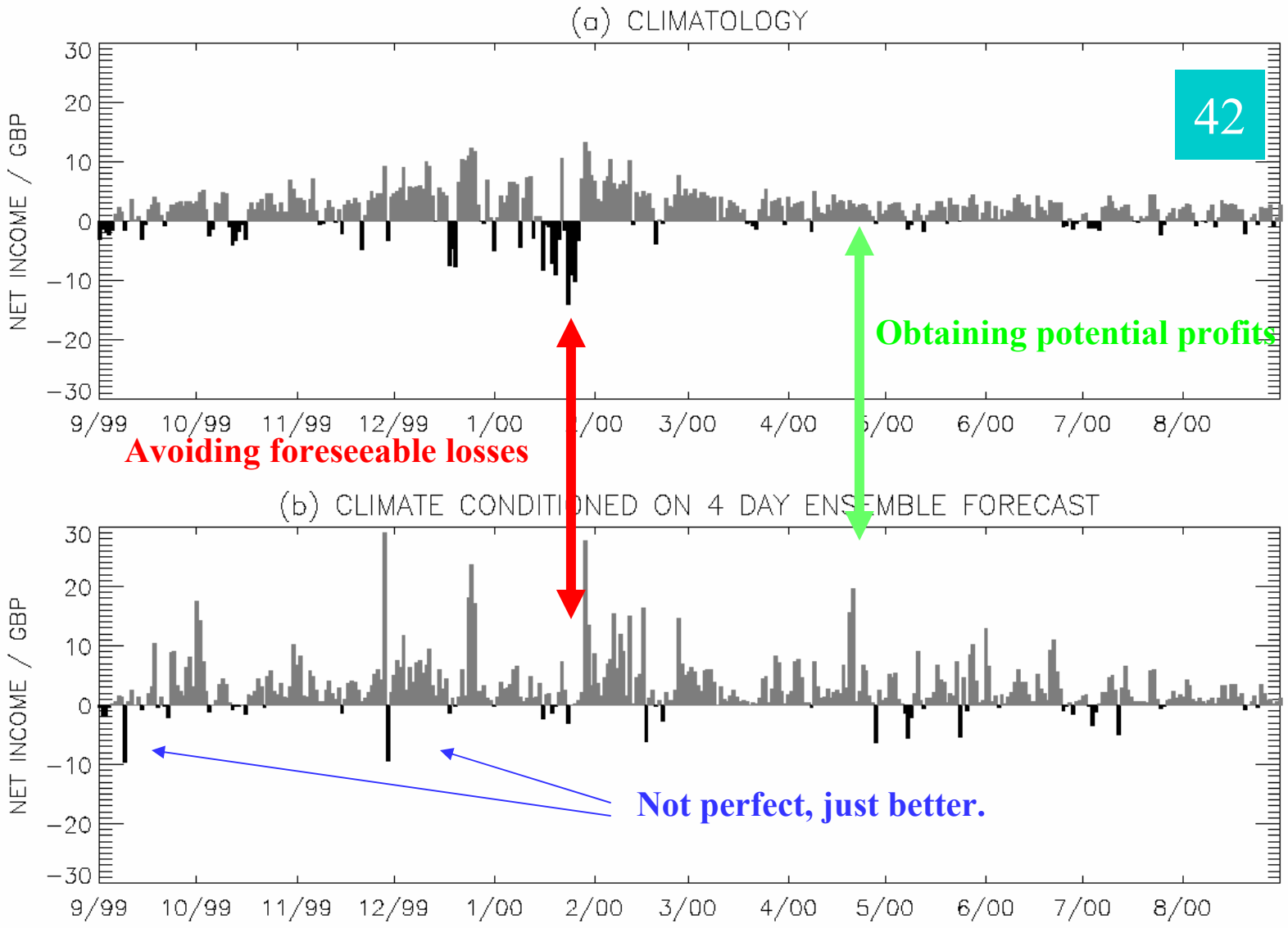
[a] *Pembroke College, Oxford University, Oxford, UK*
[b] *Centre for the Analysis of Time Series, London School of Economics, London, UK*
[c] *Macalester College, Minnesota, USA*

Received 12 February 2002; accepted 30 April 2002

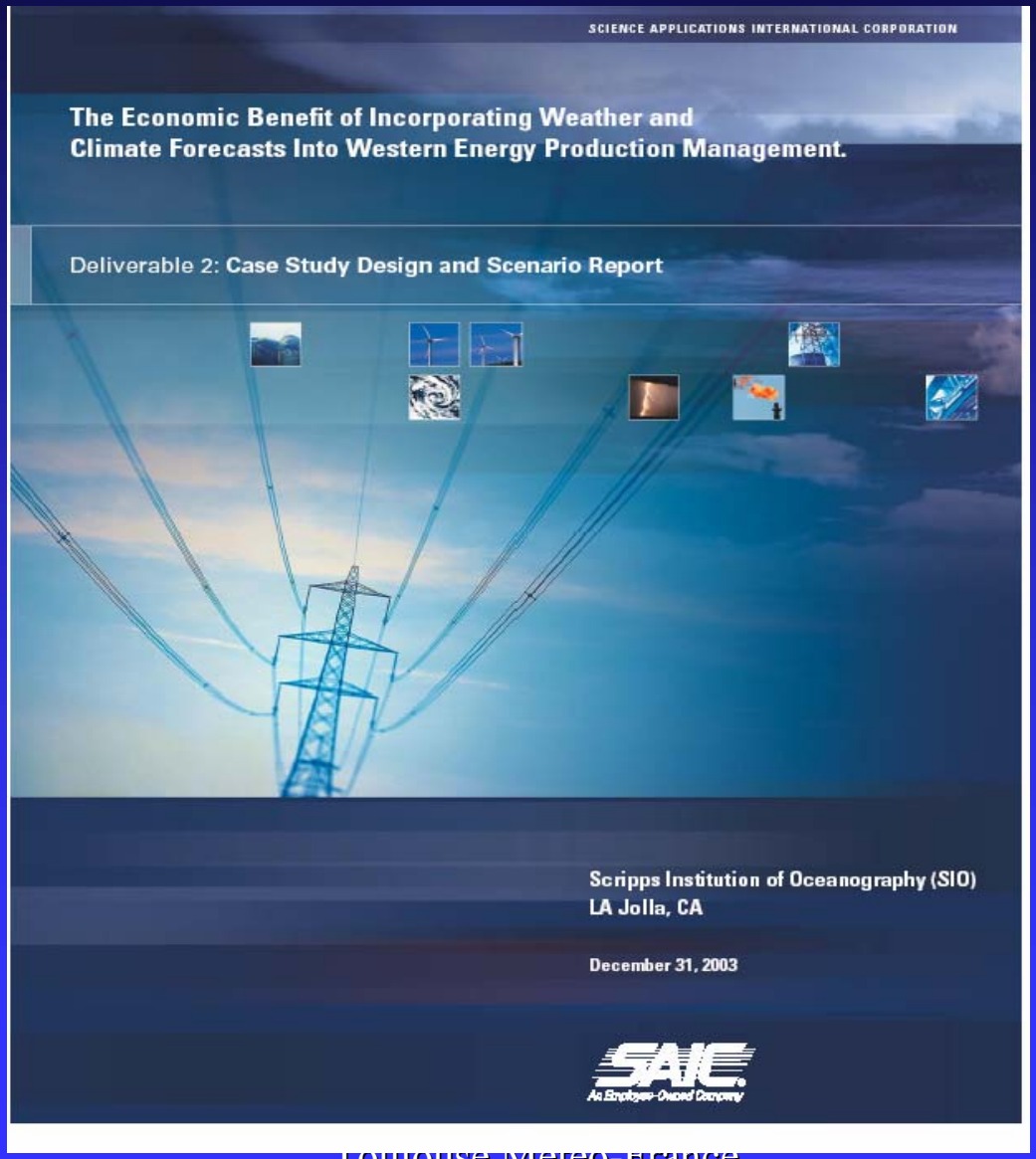
Abstract

The value of different strategies for consolidating the information in European Centre for Medium Range Weather Forecasting (ECMWF) forecasts to wind energy generators is investigated. Simulating the performance of generators using the different strategies in the context of a simplified electricity market revealed that ECMWF forecasts in production decisions improved the performance of generators at lead times of up to 6 days. Basing half-hourly production decisions on a production forecast generated by condtioning the climate on the ECMWF operational ensemble forecast yields the best results of all the strategies tested.

And you could have run a more profitable UK wind farm in the 1999 under NETA rules.



And from an American ISO's point of view?
PoV Case Study: Electricity Demand

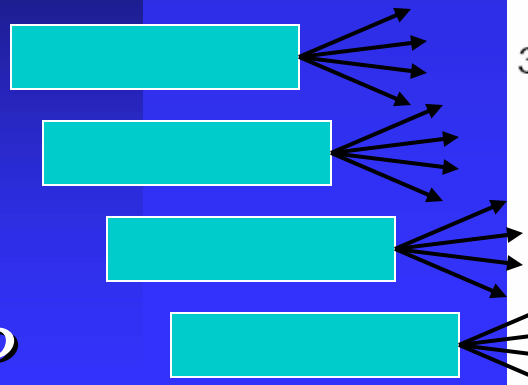


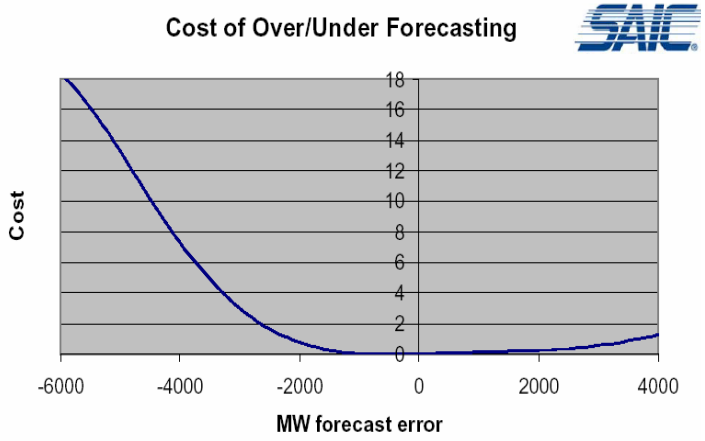
Cal ISO wants to move beyond Expected Utility in Practice!

AVN_mos

ETA_mos

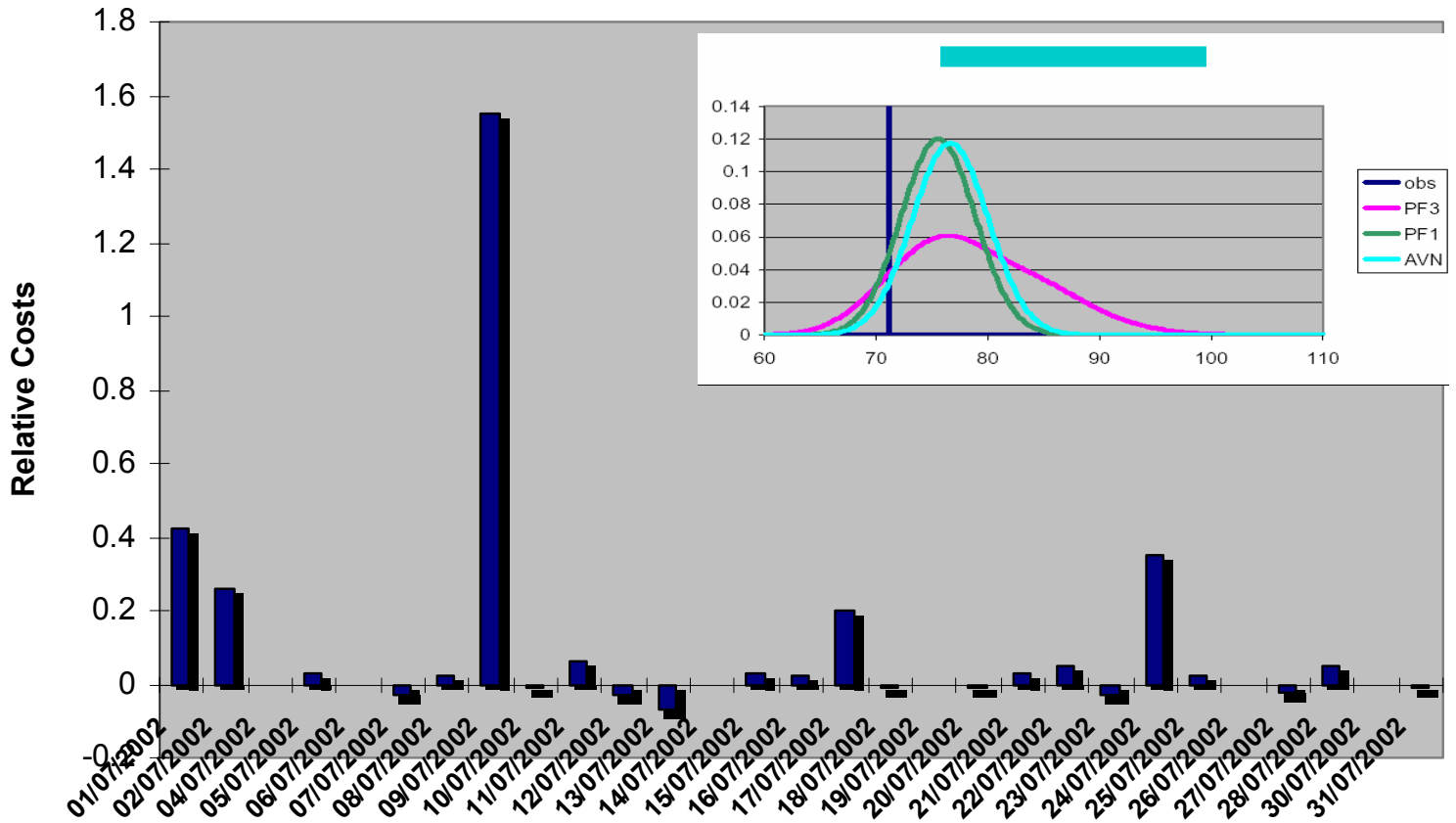
MRF





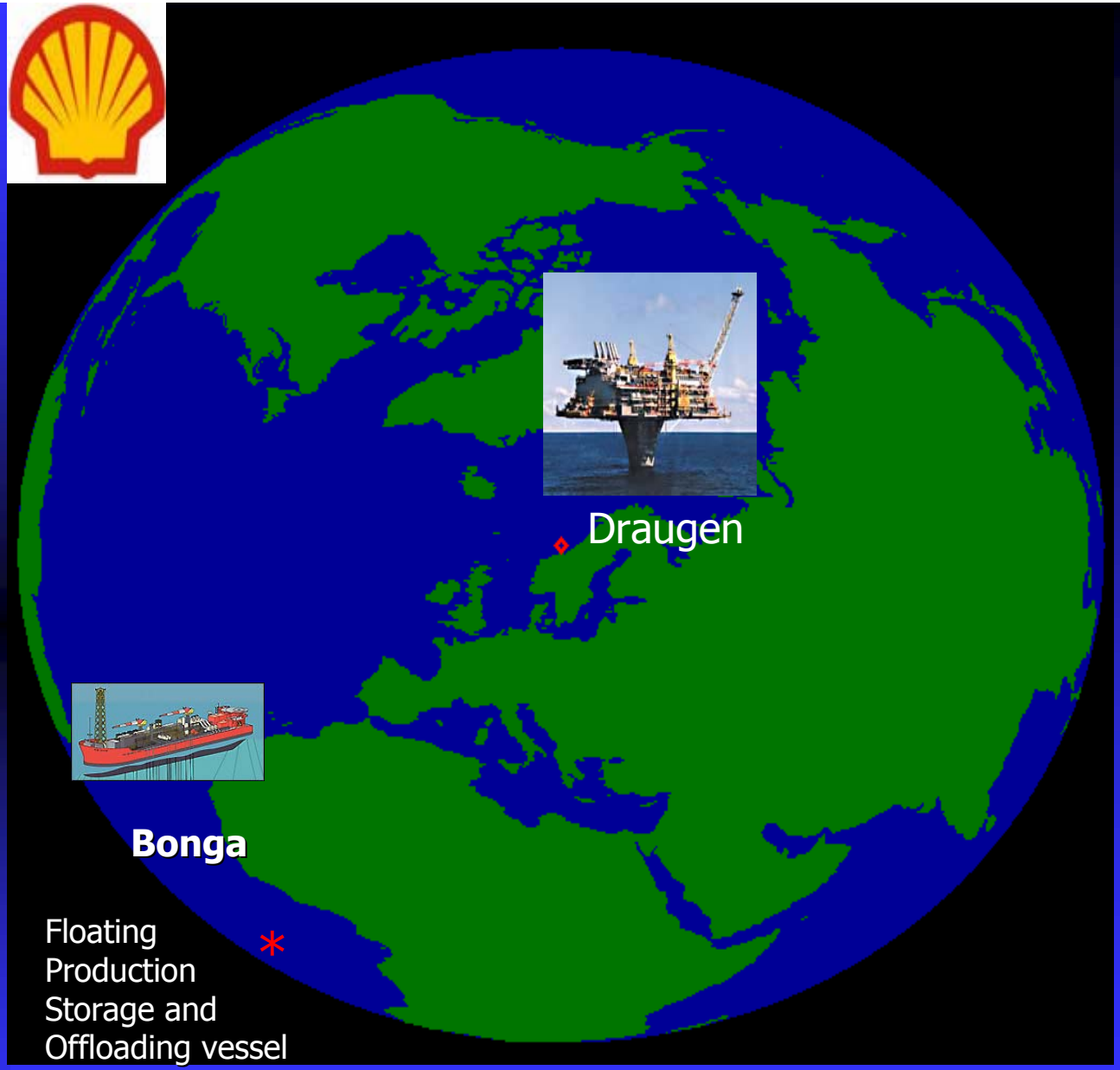
2 Feb 2006

In this case, interpreting the ensemble as a probability (and maximizing Expected Utility) is far from optimal: is it then rational to interpret the forecast distribution as a probability forecast?



From Smith, Altalo & Ziehmann (2004)

Figure 6: Relative costs of PF1 forecasts versus the Cal ISO surrogate forecasts for days in July 2002, a positive value represents a savings of using PF1. Note the significant savings on July 9th.



Shell is incorporating these ideas into their safety forecasts.

And what about big storms?

Well, I hope I've shown you that there is real socio-economic value in everyday weather!

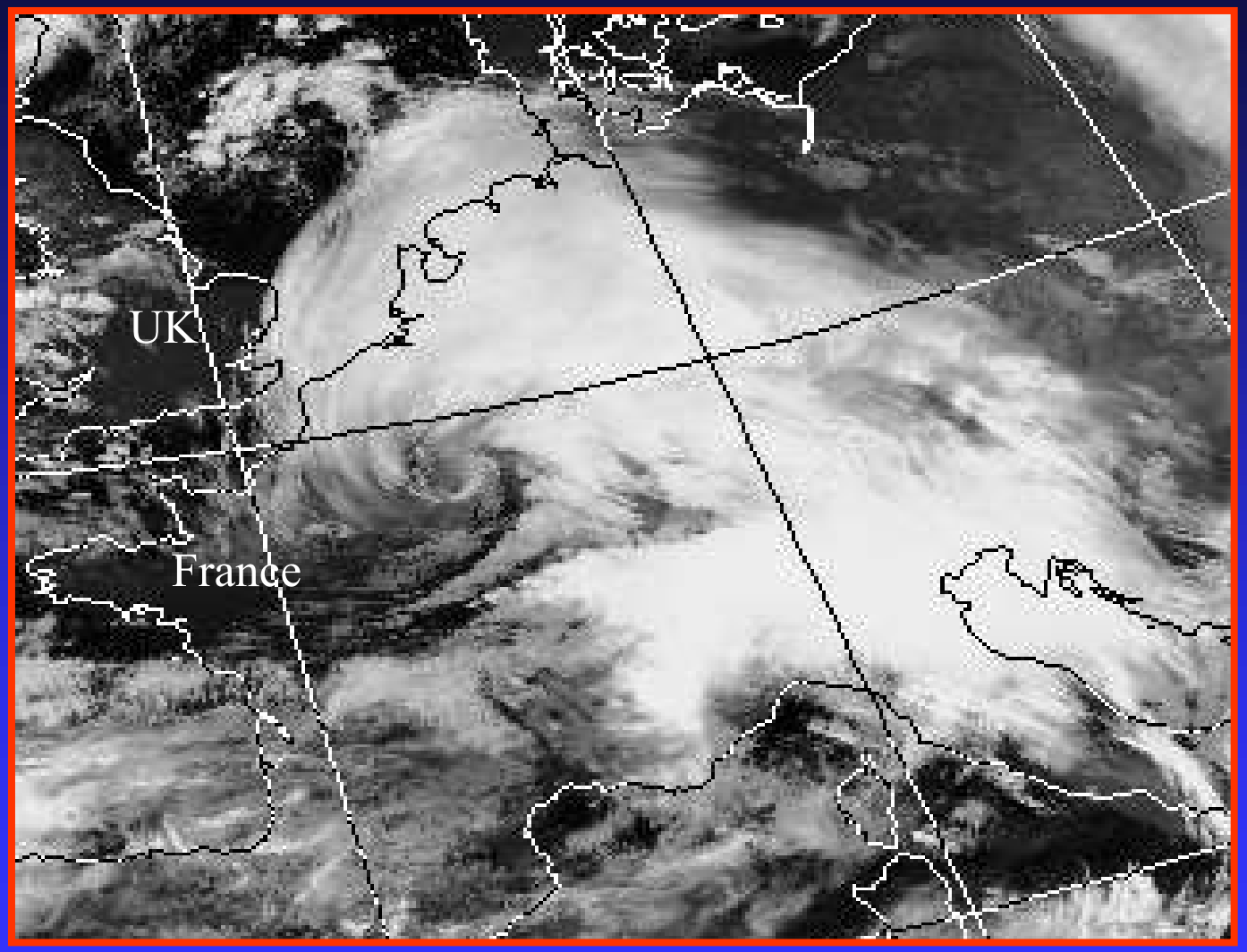
But big storms and climate change get more press, what can we say about them?

Tropical Storm/Hurricane "Cone of Uncertainty"

The Presentation of Uncertainty (The Weather Channel):

. How should I interpret this distribution?

LOTHAR



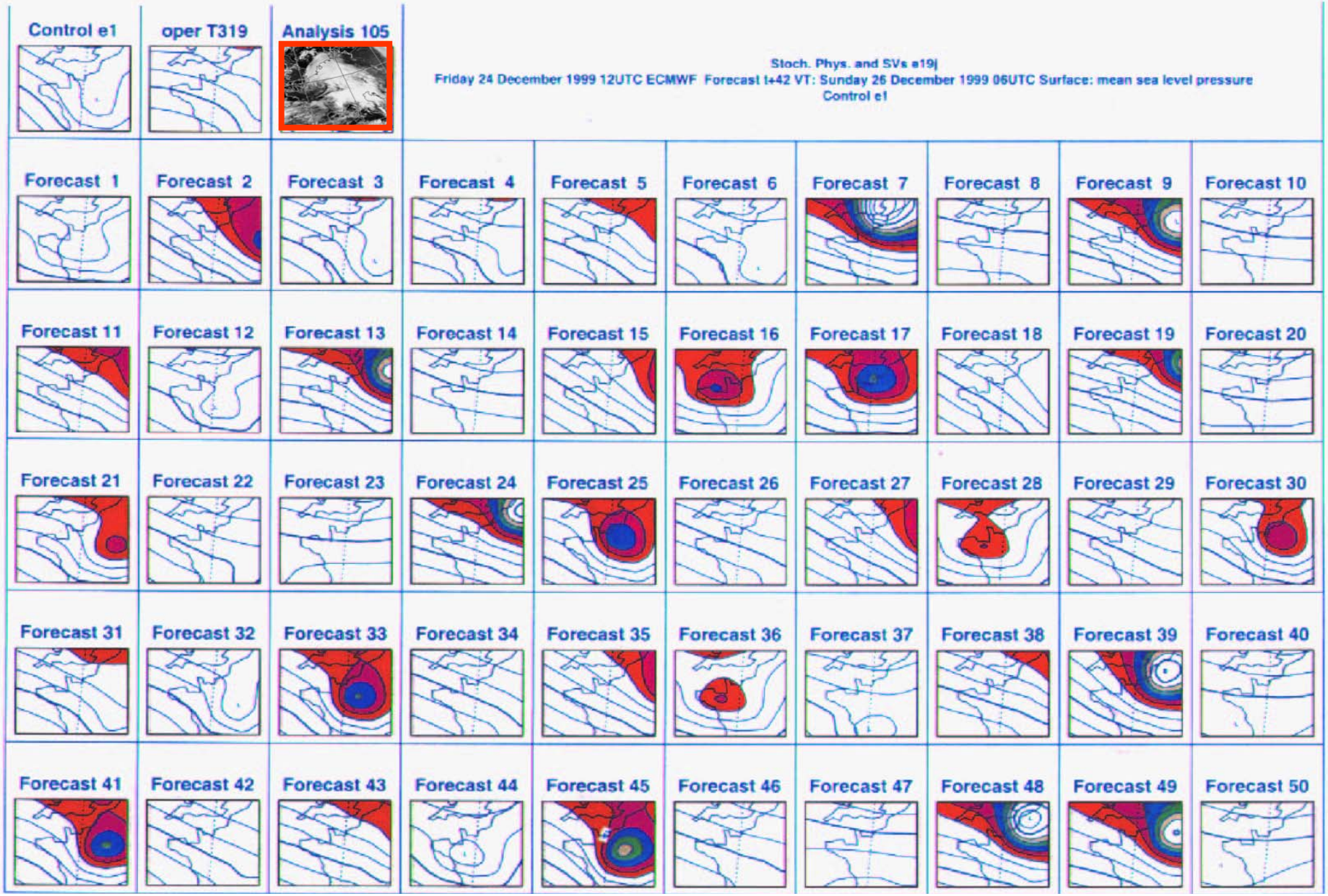
Dundee Satellite Station: 0754 UTC 26 December 1999

Two 42 hour forecasts and the weather for Dec 26 1999

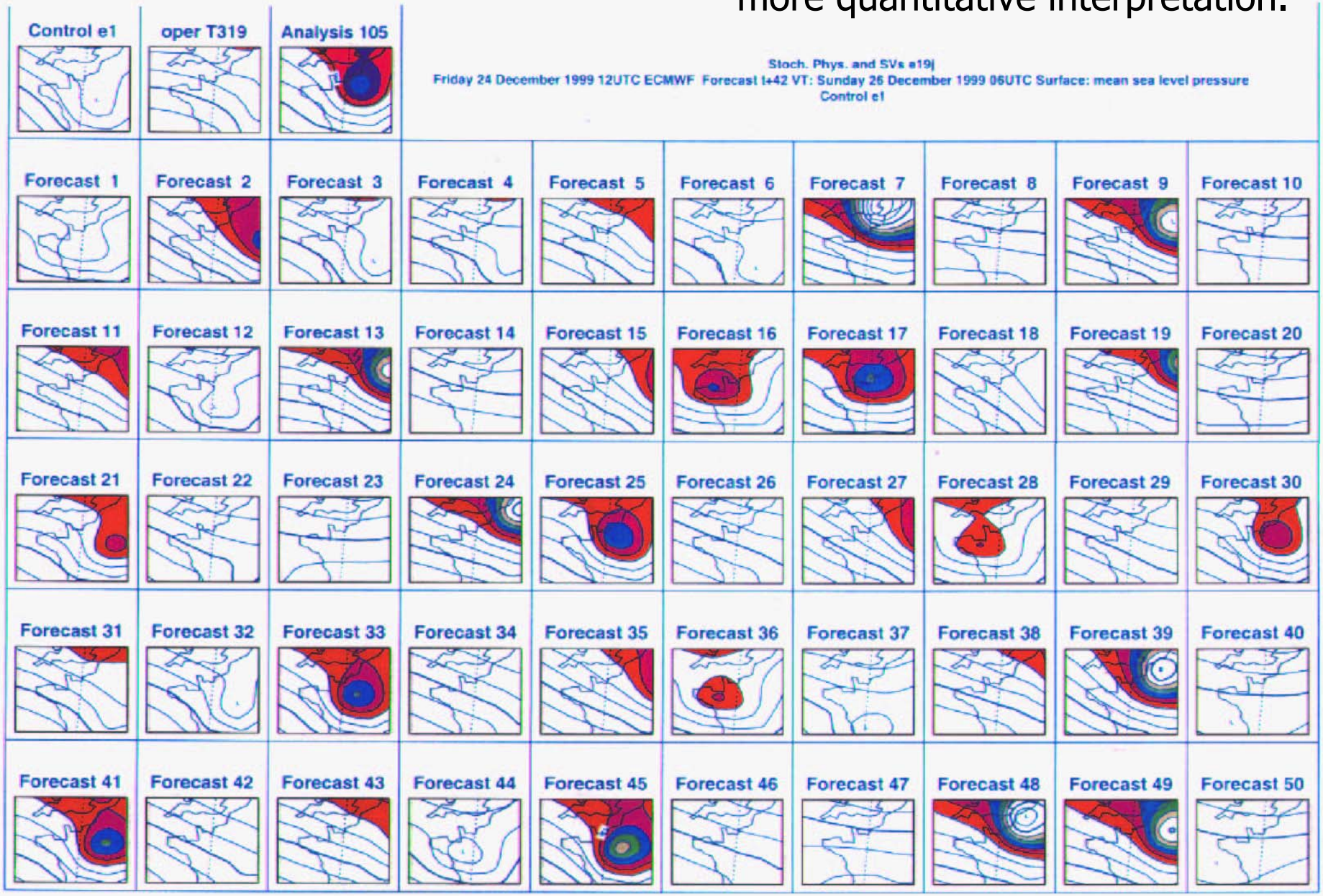
24 December 1999 12UTC ECMWF Forecast
6 December 1999 06UTC Surface: mean sea l
Stoch. Phys. and SVs e19j

Guess which are the "forecasts" and which is "reality".

This is the ECMWF forecast: each is like a golf ball. ***Each of them*** correspond to the 6AM on Dec 26!



This is a set of 52 simultaneous simulations (like the different golf balls); and it is of great value, even if the model cannot support a more quantitative interpretation.

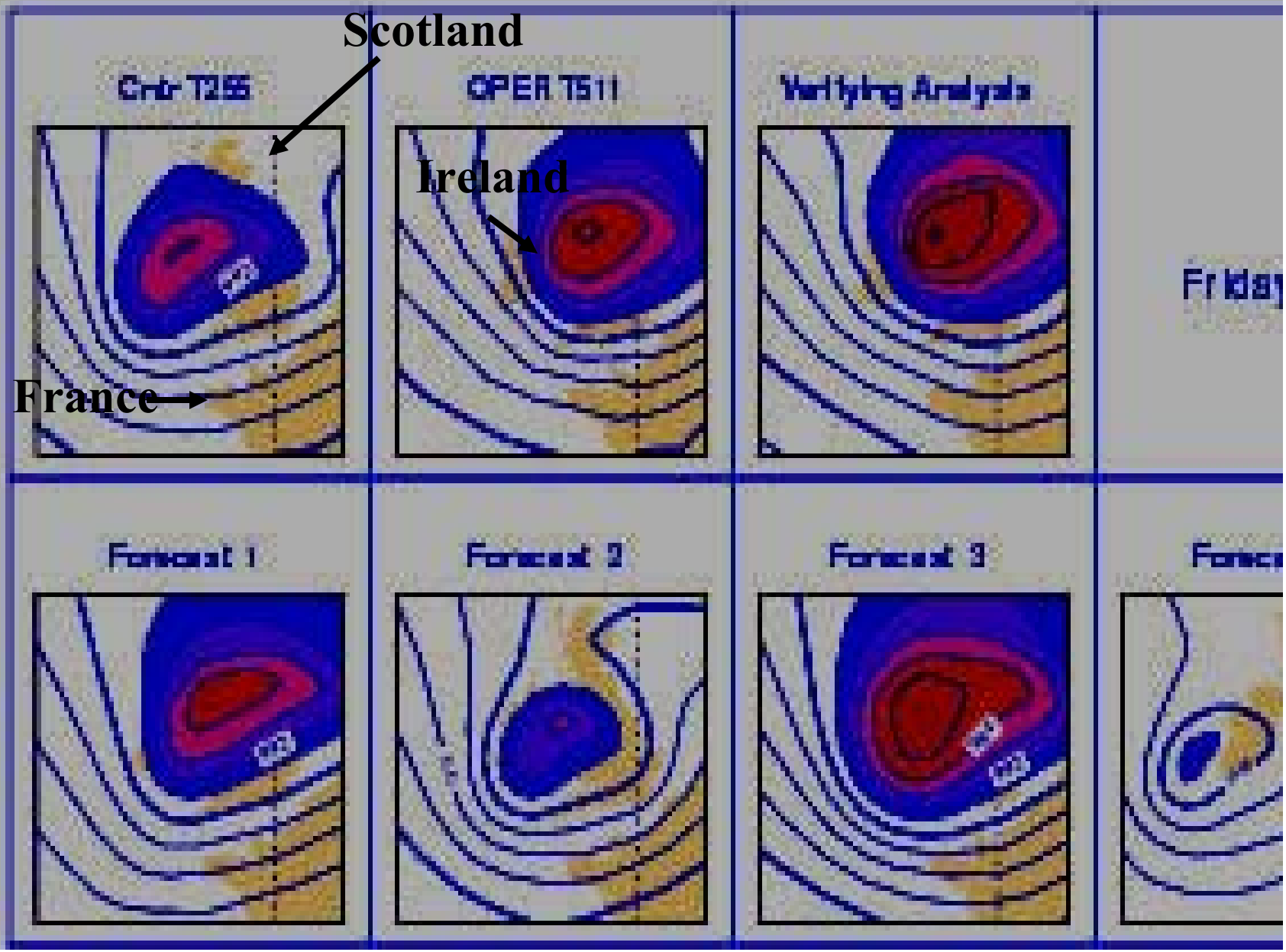


European Wind Storms: December 1999

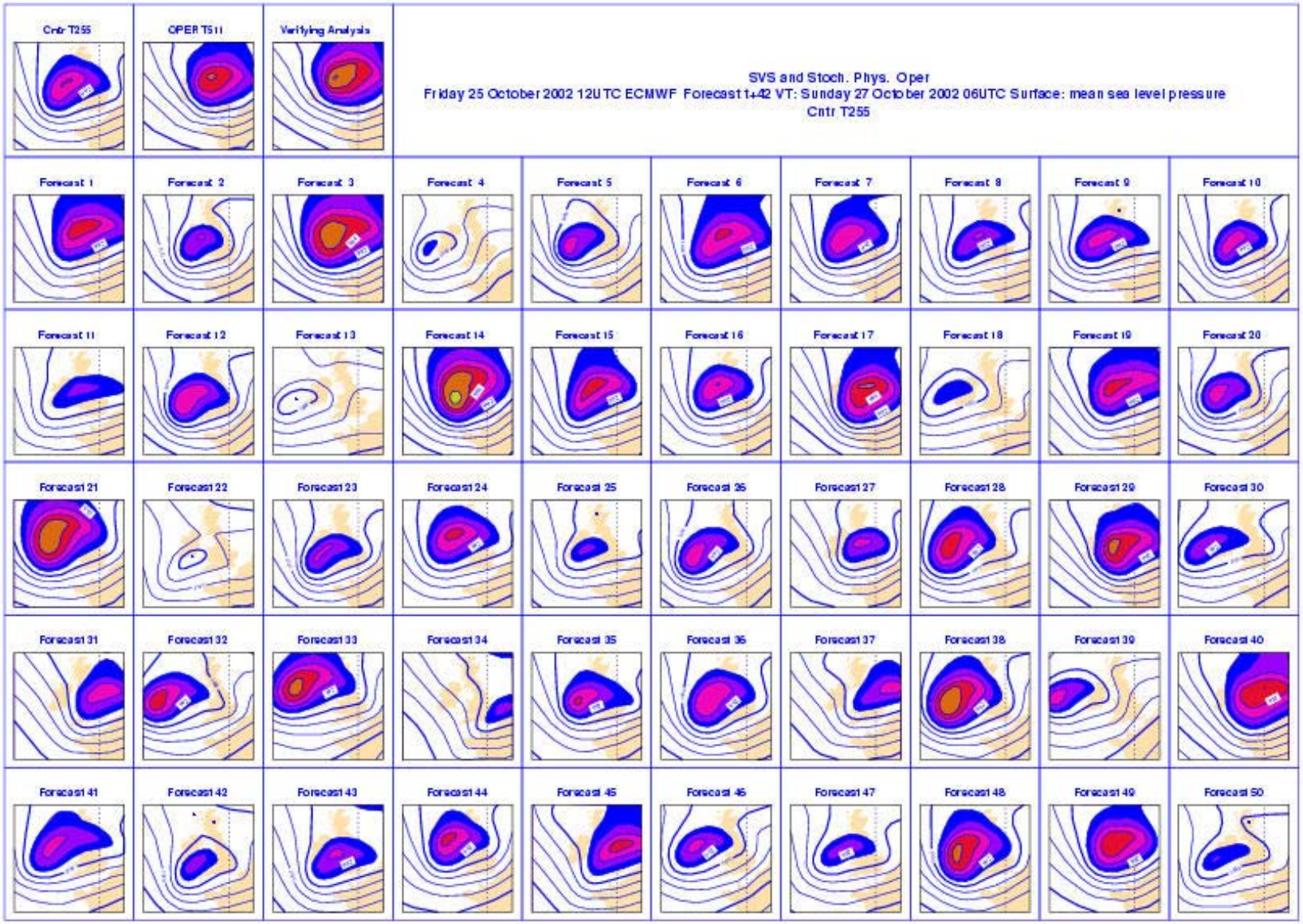
100 lives lost

Distruction of the church in Balliveirs (left) and the devistation of the ancient forest at Versailes (below).

Indistinguishable Forecasts (well, simulations)



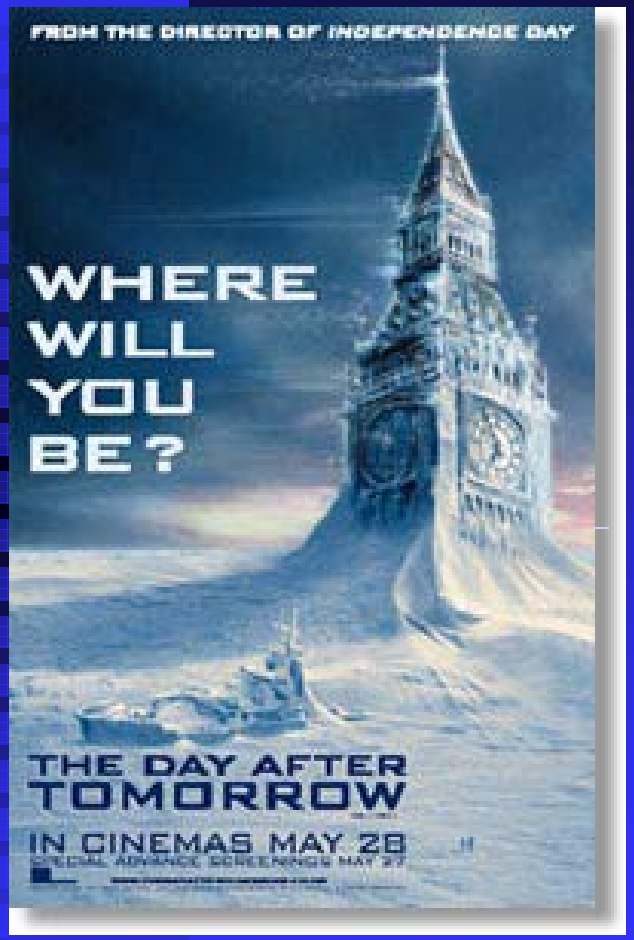
ECMWF 42 hour forecast for the October Storm of 2002



Here each simulation looks physically reasonable; most have storms: I do not want my car sitting under a tree

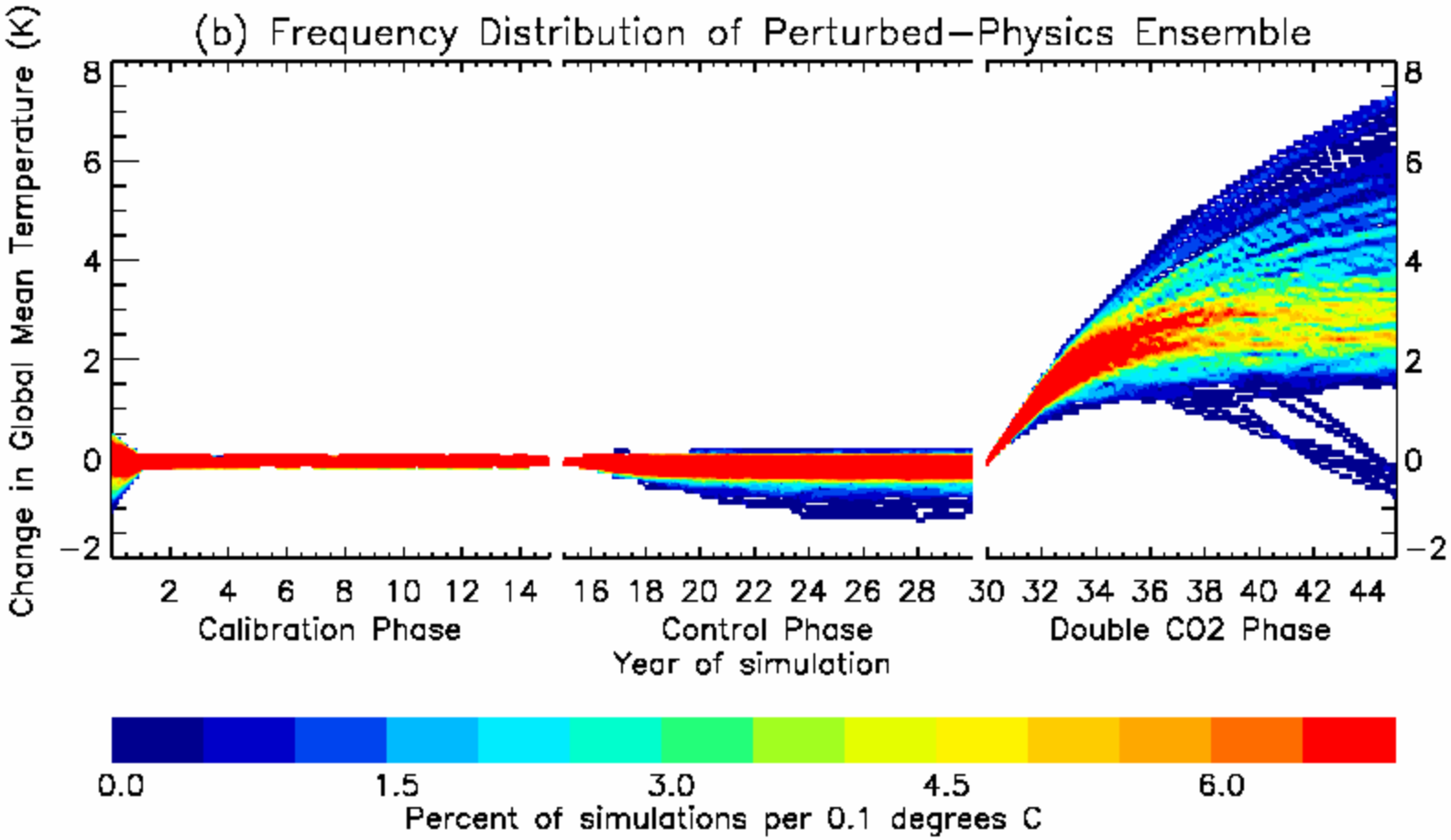
Justifying Action under Uncertainty

How big is a “low probability” to you?



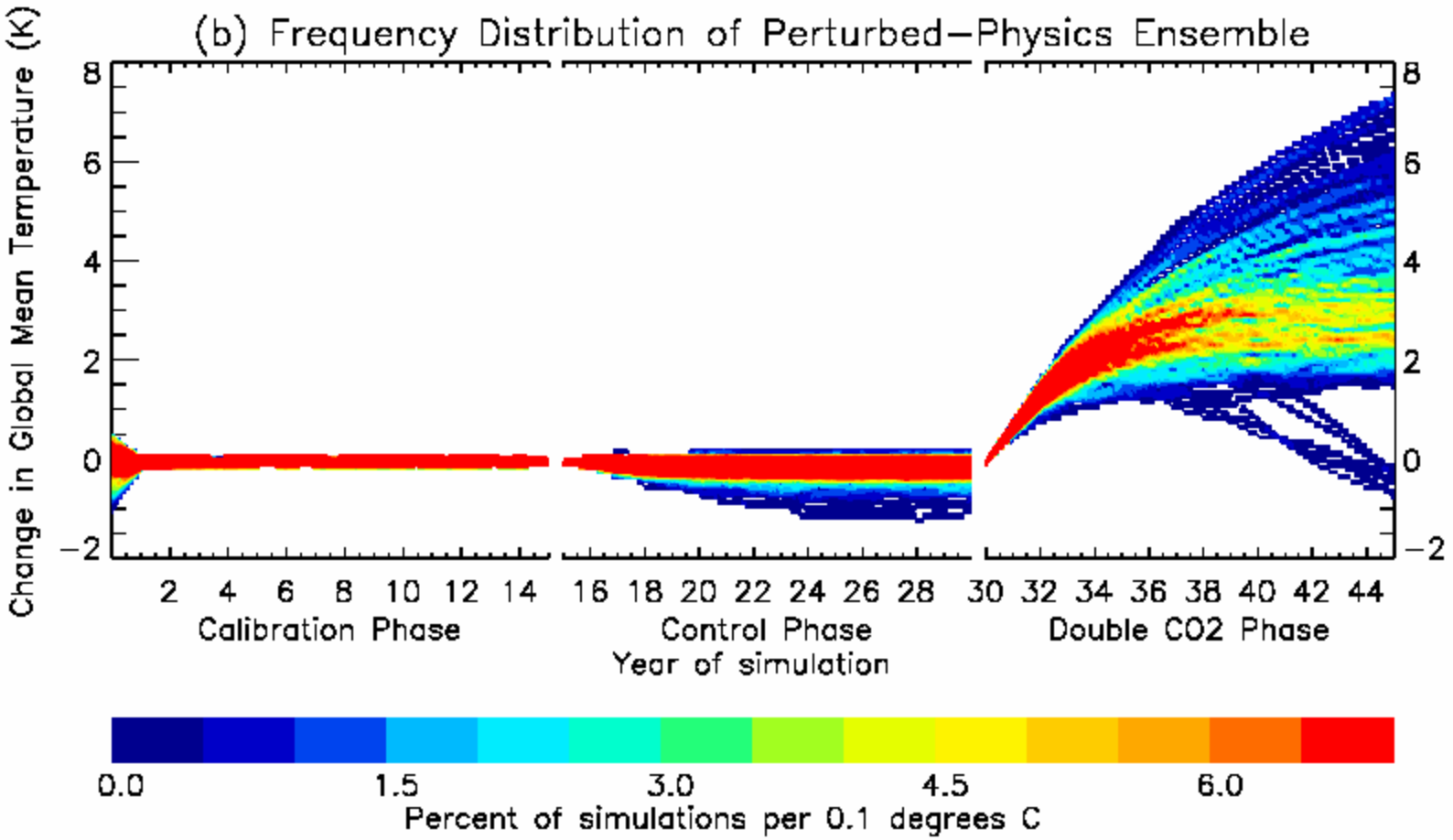
Although we estimate that the chances of a 'Big Chill' in the next hundred years has a *low probability, we don't know how low*, and if it happened it would have a very high impact” UKMO

60%	1
30%-50%	12
10%-25%	8
0%	1



And we can already see that the IPCC range of climate model sensitivity is unreasonably small.

(Stainforth et al, 2004)



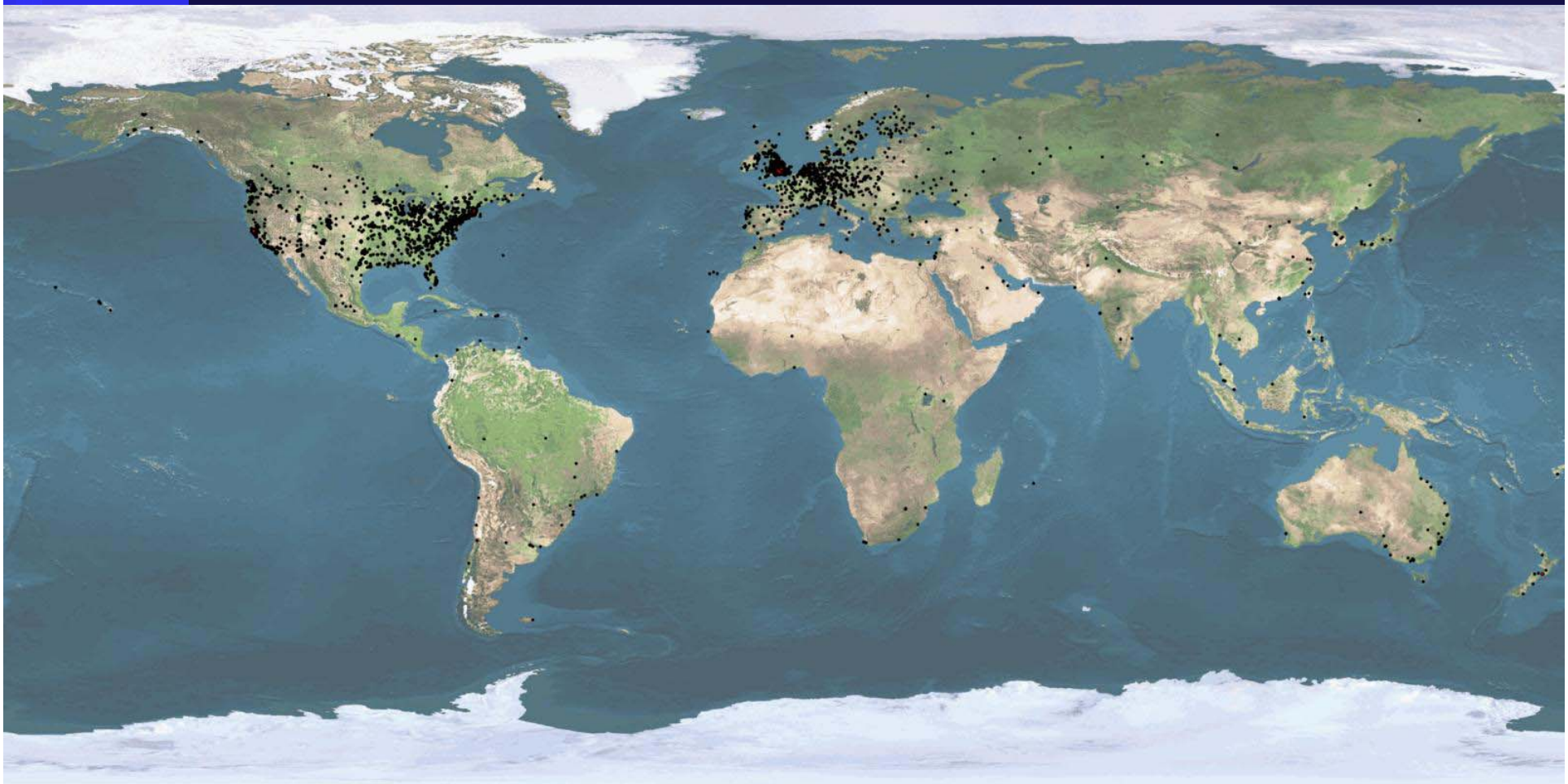
This distribution represents the evolution of several thousand full GCM climate model runs; it is *not* a probability forecast.

It tells us a lot about the model(s)

And a *lower bound* on our uncertainty

climate*prediction*.net

We are currently running the largest GCM ensemble ever!



~50,000 volunteers, 130 countries, ~2M GCM-years

And we can already see that the IPCC range of climate model sensitivity is unreasonably small (Stainforth et al, 2004)

FREE

METRO

Clubbing: This is the only guide you need

Thursday, January 27, 2005

www.metro.co.uk

Page 25

Smoke billows into the sky above the port of Dudinka in Russia yesterday

11°C

That's how much hotter scientists believe the world will get ... and it will be worse in Britain

BY SUZY AUSTIN

THE world is likely to heat up by an average of 11°C by the end of the century, the biggest ever study of global warming showed yesterday.

And the effect could be even more marked in Britain, where temperatures could soar by up to 20°C unless greenhouse gases are cut.

Such a rise – far higher than the 2°C previously forecast – would see Britain endure tropical temperatures, flooding and devastating drought.

It would change the weather patterns of the world, melt the polar ice caps and warm the oceans, causing a surge in sea levels threatening the lives of billions of people.

The findings come from a study which tapped into the processing power of 100,000 home computers in 150 countries.

Researchers racked up the equivalent of 8,000 years of processing time as they ran 60,000 potential scenarios through the network, far more than the 128 scenarios the powerful computers at the Met Office can check in a year.

Each scenario was based on the assumption that carbon dioxide levels had reached double those of pre-Industrial Revolution times by the middle of this century.

Researcher David Stainforth, from Oxford University, said: 'An 11 degree warmed world would be a dramatically different world.

'Warming is not constant at all latitudes and tends to be greater at high latitudes.

'With a world warmed by 11 degrees there would be large areas of high latitude that could be 20 degrees warmer than they are today.

'I think it would probably not be a tropical paradise. The UK would be at the high end of this change, well into the teens as the temperature changes. I don't think we'll be building many snowmen in winter, or going sledging.'

The findings could mean world leaders need to toughen their commitment in the Kyoto agreements to cut CO_2 emissions to 5.2 per cent below 1990 levels by 2012.

The warning came as Tony Blair used the World Economic Forum in Switzerland to call for action on global warming and to pressure America to sign up to Kyoto.

Blair's call – Page 5

Assessing the risk:
a METRO from Feb 2005

Is this an accurate reflection of what is known today?

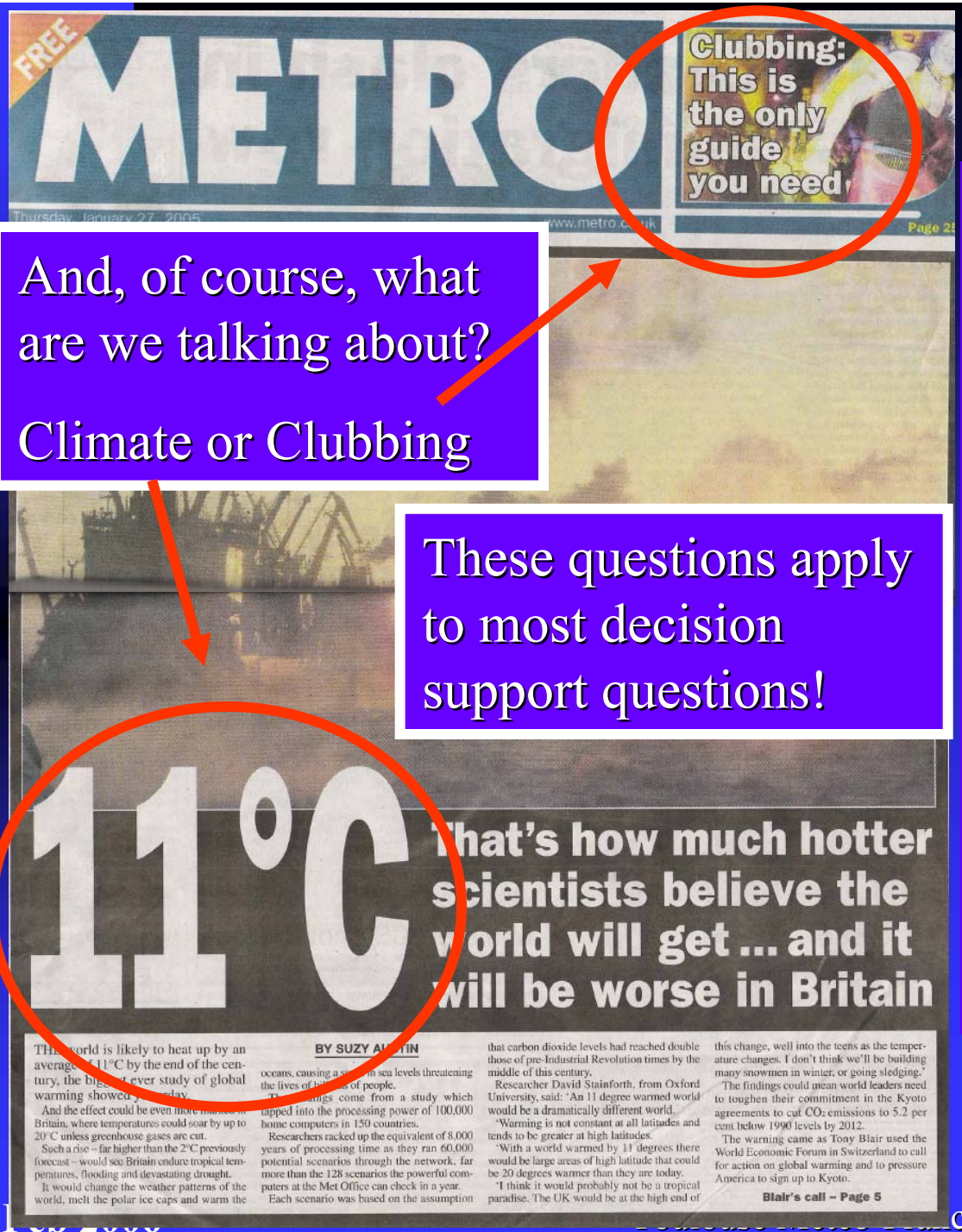
Is what we know today accurate?

How relevant is this information to the decisions we have to make today?

And for long-term decisions:

How robust is the current information?
Is informed expert opinion likely to change in a years time?

A METRO from Feb 2005



And, of course, what are we talking about?

Climate or Clubbing

These questions apply to most decision support questions!

Is this an accurate reflection of what is known today?

Is what we know today accurate?

How relevant is this information to the decisions we have to make today?

And for long-term decisions:

How robust is the current information? Is informed expert opinion likely to change in a years time?

References

L.A. Smith, M. Roulston & J. von Hardenburg (2001) *End to End Ensemble Forecasting: Towards Evaluating The Economic Value of an Ensemble Prediction System* Technical Memorandum **336,** 29 pp. European Centre for Medium Range Weather Forecasts, Shinfeld Road, Reading, UK.

M.S. Roulston, C. Ziehmann & L.A. Smith (2001) *A Forecast Reliability Index from Ensembles: A Comparison of Methods* Tech Report for Deutscher Wetterdienst

M.S. Roulston D.T. Kaplan, J. Hardenberg & L.A. Smith (2003) Using Medium Range Weather Forecasts to Improve the Value of Wind Energy Production. *Renewable Energy* **28** (4) 585–602

LA Smith (2003) Predictability Past Predictability Present. ECMWF Seminar on Predictability. soon to be in a CUP book (ed. Palmer).

LA Smith (2000) *Disentangling Uncertainty and Error*, in Nonlinear Dynamics and Statistics (ed A.Mees) Birkhauser.

LA Smith (2002) *What might we learn from climate forecasts?,* Proc. National Acad. Sci. **99**: 2487-2492.

www.lsecats.org

lenny@maths.ox.ac.uk

Weather roulette