Abstract

The design and interpretation of model simulations for climate services differ significantly from experimental design for the advancement of the fundamental research on predictability that underpins it. Climate services consider the sources of best information available today; this calls for a frank evaluation of model skill in the face of statistical benchmarks defined by empirical models. The fact that Physical simulation models are thought to provide the only reliable method for extrapolating into conditions not previously observed has no bearing on whether or not today's simulation models outperform empirical models. Evidence on the length scales on which today's simulation models fail to outperform empirical benchmarks is presented; it is illustrated that this occurs even on global scales in decadal prediction. At all timescales considered thus far (as of July 2012), predictions based on simulation models are improved by blending with the output of statistical models. Blending is shown to be more interesting in the climate context than it is in the weather context, where blending with a history-based climatology is straightforward. As GCMs improve and as the Earth's climate moves further from that of the last century, the skill from simulation models and their relevance to climate services is expected to increase. Examples from both seasonal and decadal forecasting will be used to discuss a third approach that may increase the role of current GCMs more quickly. Specifically, aspects of the experimental design in previous hind cast experiments are shown to hinder the use of GCM simulations for climate services. Alternative designs are proposed. The value in revisiting Thompson's classic approach to improving weather forecasting in the fifties in the context of climate services is discussed.

LSE CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

www.lsecats.ac.uk

Increasing the Relevance GCM Simulations for Climate Services

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

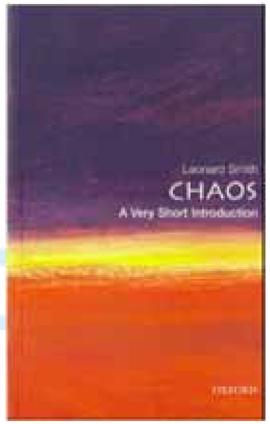
EQUIP

Not Possible without: H Du, A. Lopez, 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



Data Data Everywhere, and Not a Bit to Bank On

It seems we are surrounded by model output… much of it from models which appear unlikely to provide reliable probability forecasts for the questions we must answer.

I restrict attention to “decision-relevant” PDFs.

http://4umi.com/coleridge/rime/1

Data Data Everywhere, and Not a Bit to Bank On

In terms of climate services alone (not research) where should resources be focused?

On spatial and lead-time scales where empirical models outperform simulation models (historically), might their use in climate services improve the product delivered?

I will consider seasonal (3 month) predictions.

Then decadal (1-10 year) predictions.

And avoid 50+ year high resolution predictions.

All in 11 min.

http://4umi.com/coleridge/rime/1

Data Data Everywhere, and Not a Bit to Bank On

In terms of climate services alone (not research) where should resources be focused?

On spatial and lead-time scales where empirical models outperform simulation models (historically), might their use in climate services improve the product delivered?

I will consider seasonal (3 month) predictions.

Then decadal (1-10 year) predictions.

And avoid 50+ year high resolution predictions.

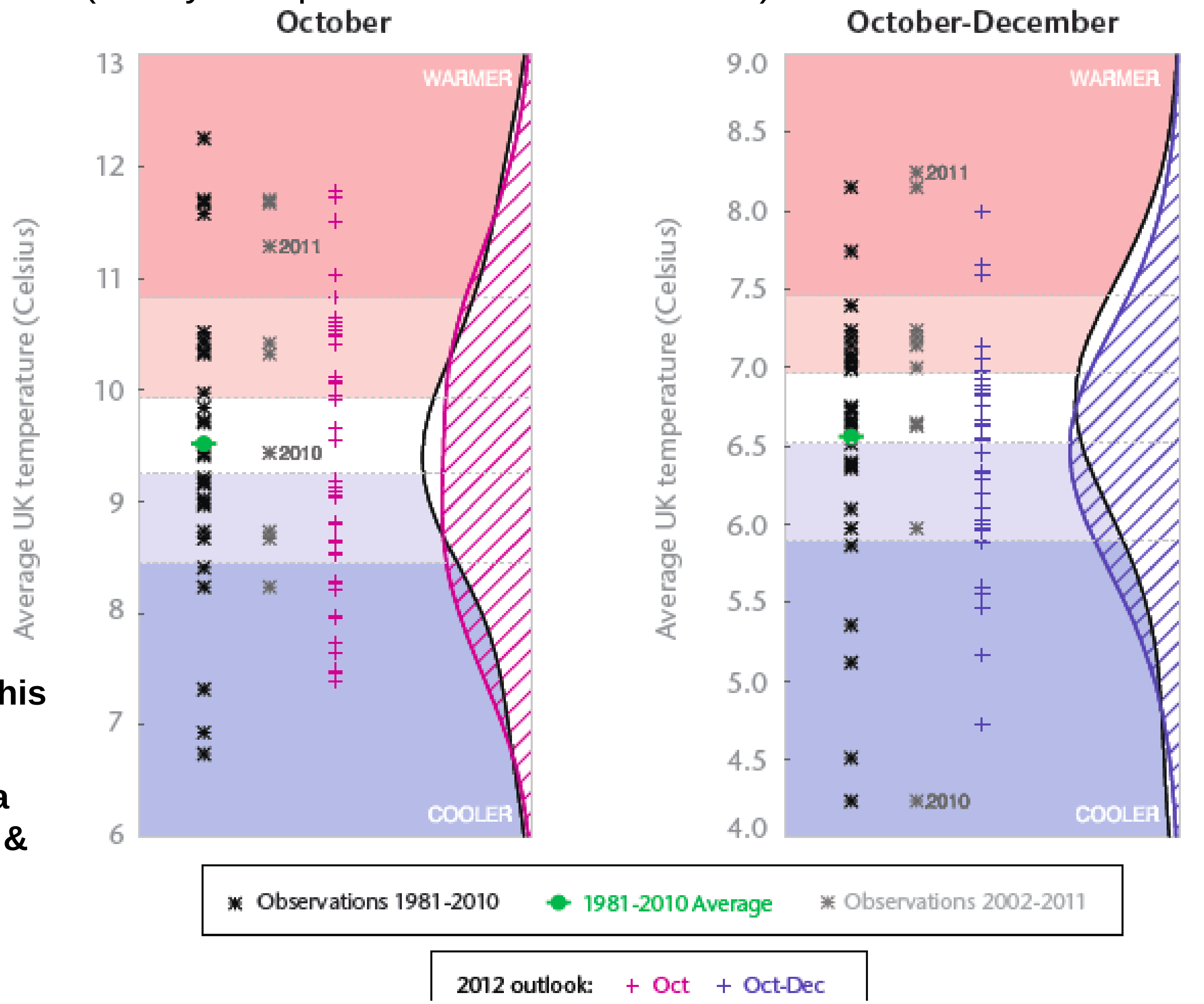
All in 09 min.

At seasonal lead-times…
… the Outlook is good!

http://4umi.com/coleridge/rime/1

Seasonal Forecast accompanied by guidance.

(A very nice presentation of information)



Historical Obs
Climate Distribution

Ensemble Members
Forecast PDF

(and **Averages**,
along with enough
information to make
it clear you do not
want to “use” them.)

How does one get this PDF forecast from:

A small ensemble, a limited climatology & an imperfect model

CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

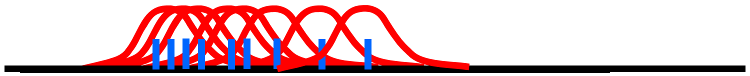
http://www.metoffice.gov.uk/media/pdf/n/3/A3-plots-temp-OND.pdf

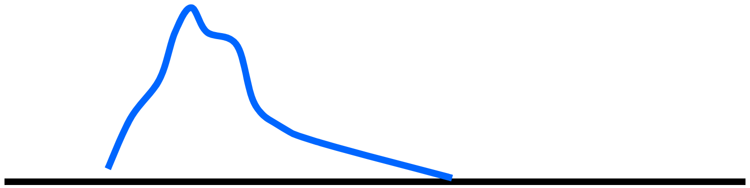
Ensembles Members In - Predictive Distributions Out

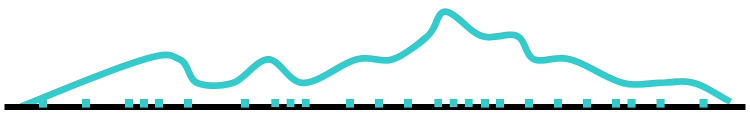
(1) Ensemble Members to Model Distributions

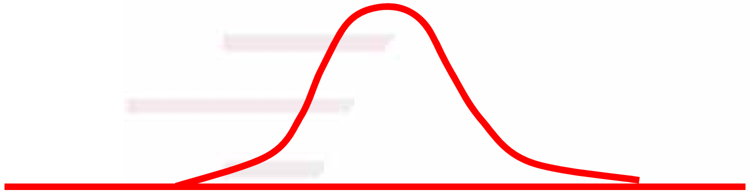
J Bröcker, LA Smith (2008) From Ensemble Forecasts to Predictive Distribution Functions *Tellus A* 60(4): 663.

K is the kernel, with parameters σ,δ (*at least*)









$$P_1(x) = \sum_{i=1}^{n_{eps}} K(x, s_i^1)/n_{eps}$$

$$P_{clim} = \sum_{i=1}^{n_{clim}} K(o_i)/n_{clim}$$

Kernel & blend parameters are fit *simultaneously* to avoid adopting a wide kernel to account for a small ensemble.

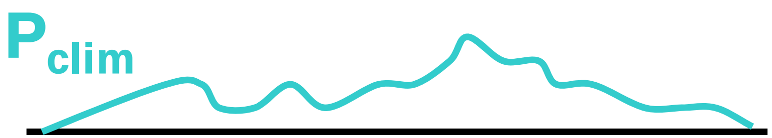
One would always dress (K) and blend (α) a finite ensemble, even with a perfect model and perfect IC ensemble.

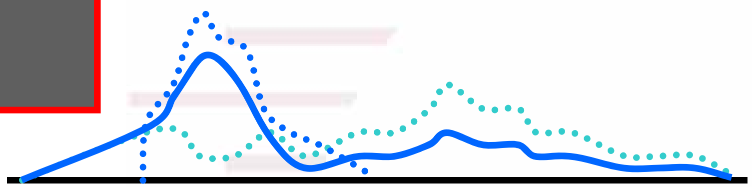
Forecast busts and lucky strikes remain a major problem when the archive is small.

Ensembles Members In - Predictive Distributions Out

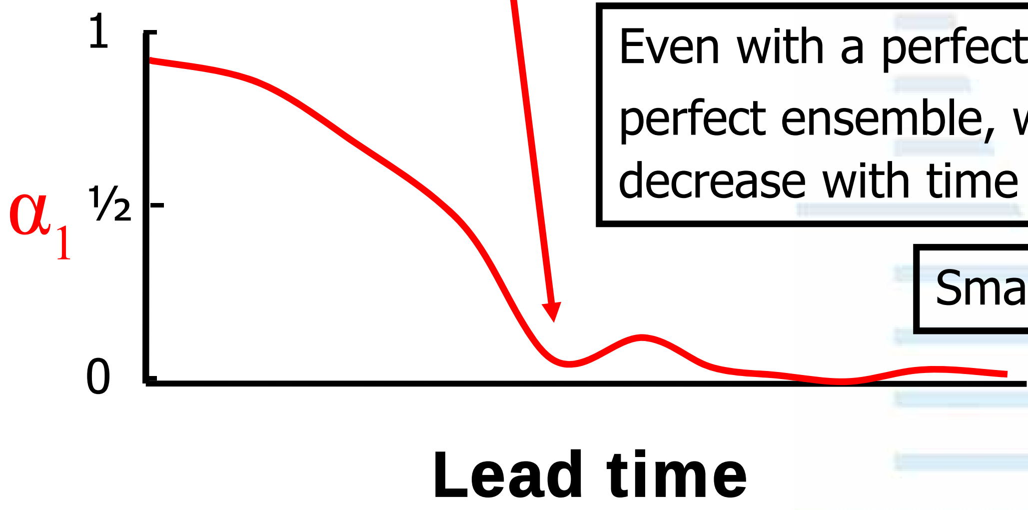
For a fixed ensemble size α decreases with time

And if $\alpha_1 \approx 0$, can there be any *operational* justification for running the prediction system.





$$M_1 = \alpha_1 P_1 + (1-\alpha_1) P_{clim}$$



Even with a perfect model and perfect ensemble, we expect α to decrease with time for small n_{eps}

Small :: $n_{eps} << n_{clim}$

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

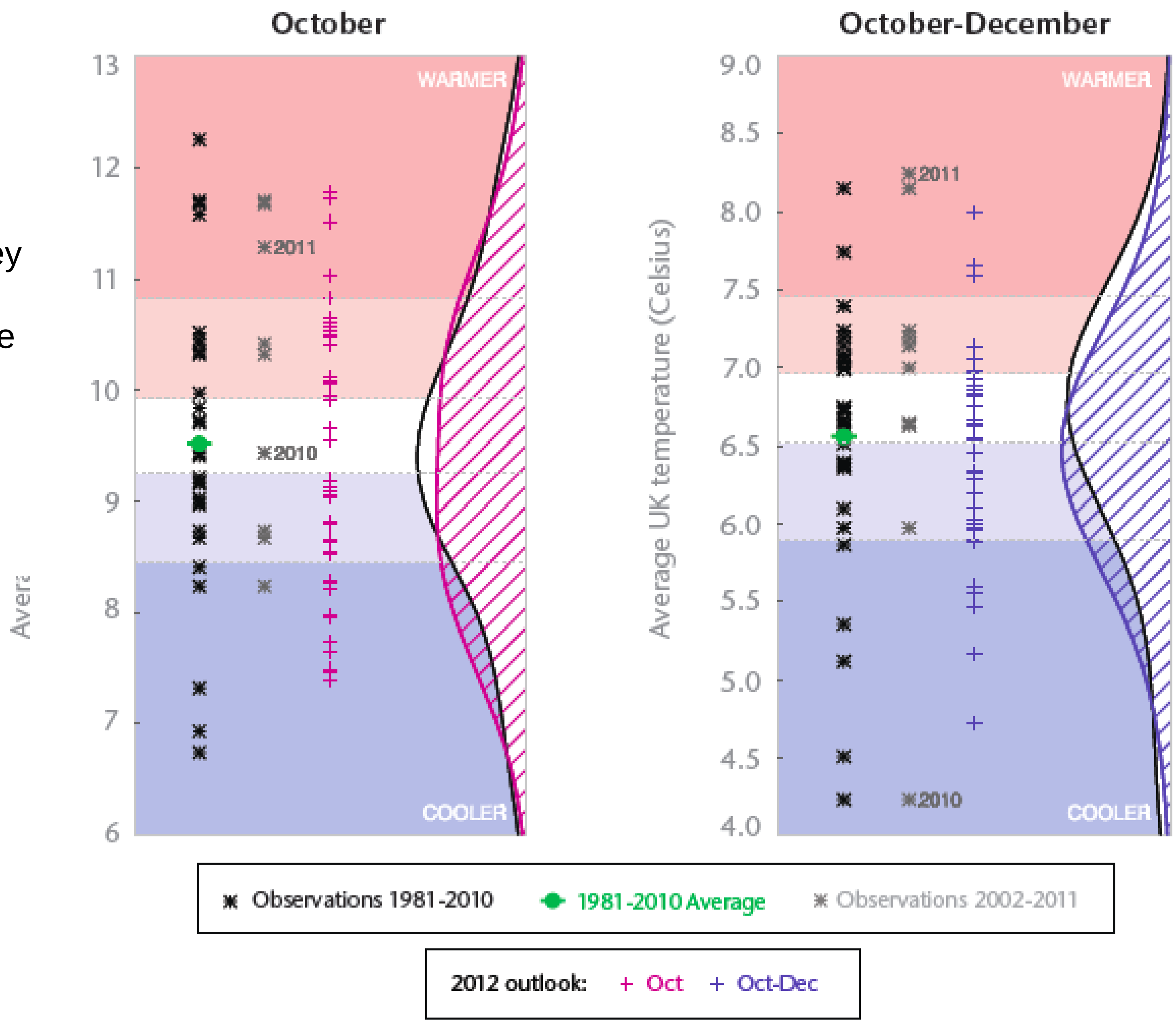
Distinguishing Value and Skill in the MetOffice Outlook

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

Are these potentially decision relevant?
YES!

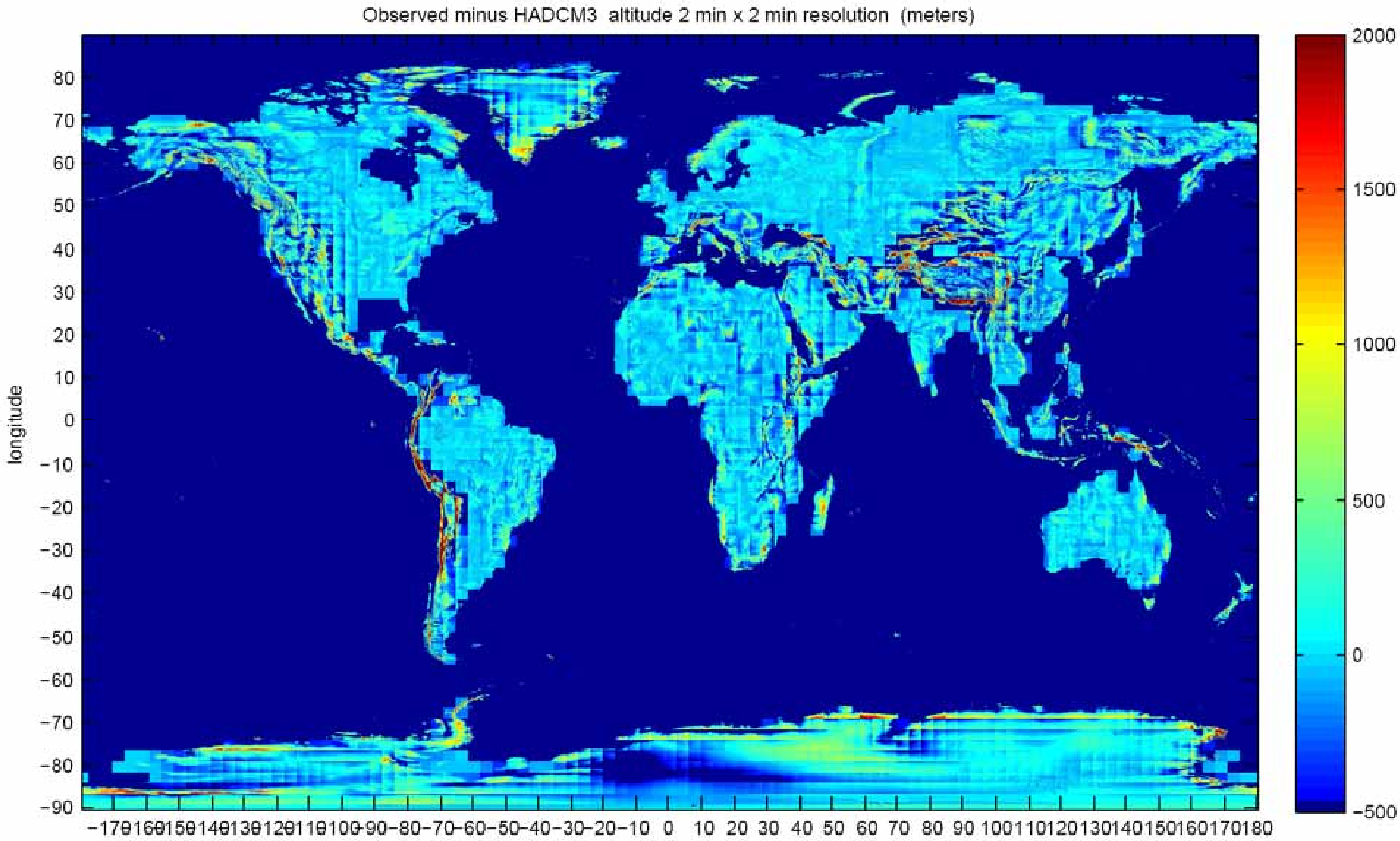
Would we have to wait 100 years to know if they are skilful?
(Not necessarily, and we can compute the time required from their generalized information deficit given only a few forecasts)

H Du & LA Smith (2012) Parameter estimation using ignorance *Phys Rev E 86*



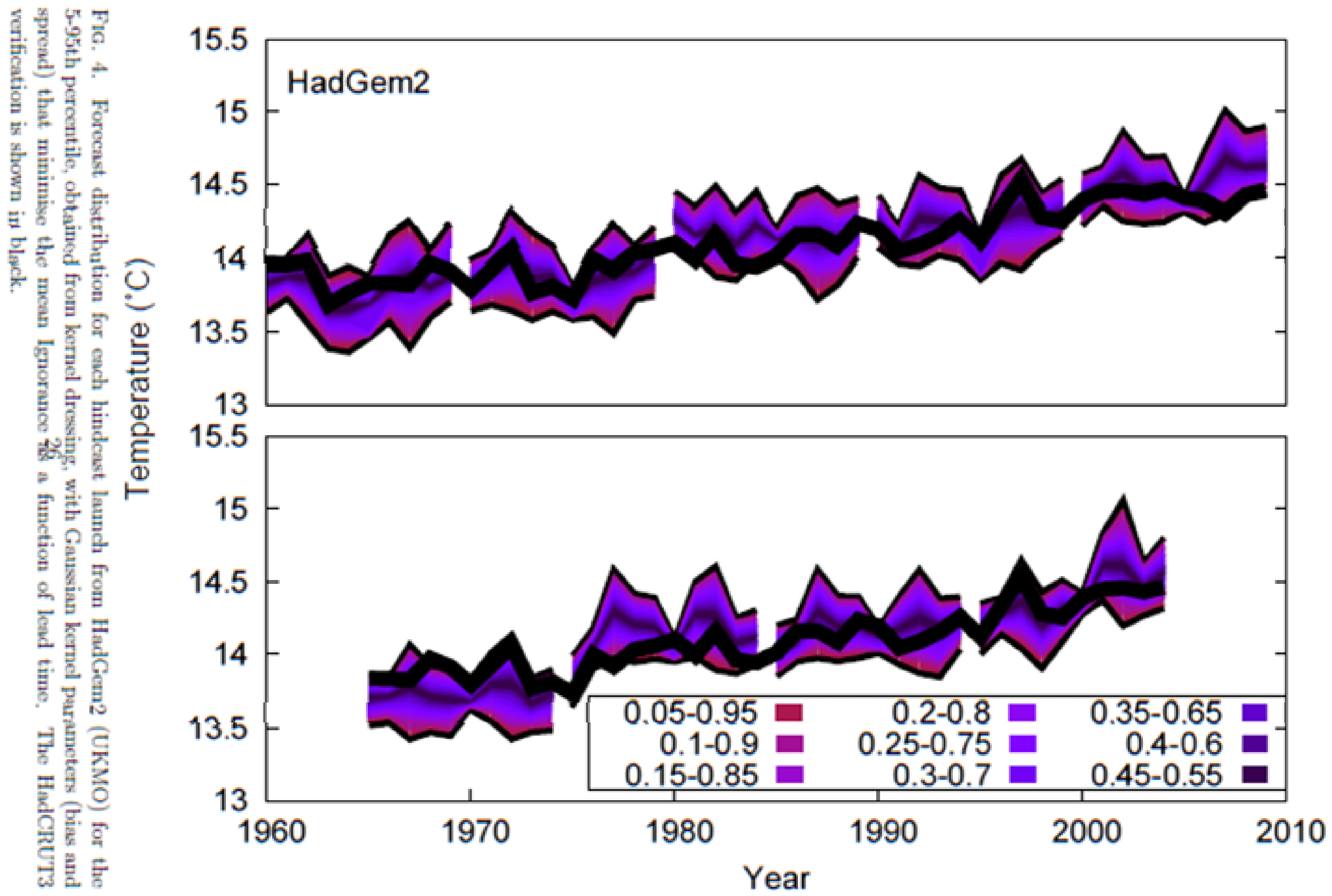
CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

Longer Lead-times require lower resolutions: How is a simulation model to prove its worth?



By adding-value to empirical models, perhaps?

Decadal Predictions



Dressed HadGem2 Decadal Forecasts

E Suckling & L A Smith (2012) Empirical Probability Benchmarks. JoC (in review)

Take eRAP (Dynamic Climatology) as an Empirical Model

LA Smith (1997) The Maintenance of Uncertainty.
Proc International School of Physics "Enrico Fermi"
Course CXXXIII, pg 177-246,

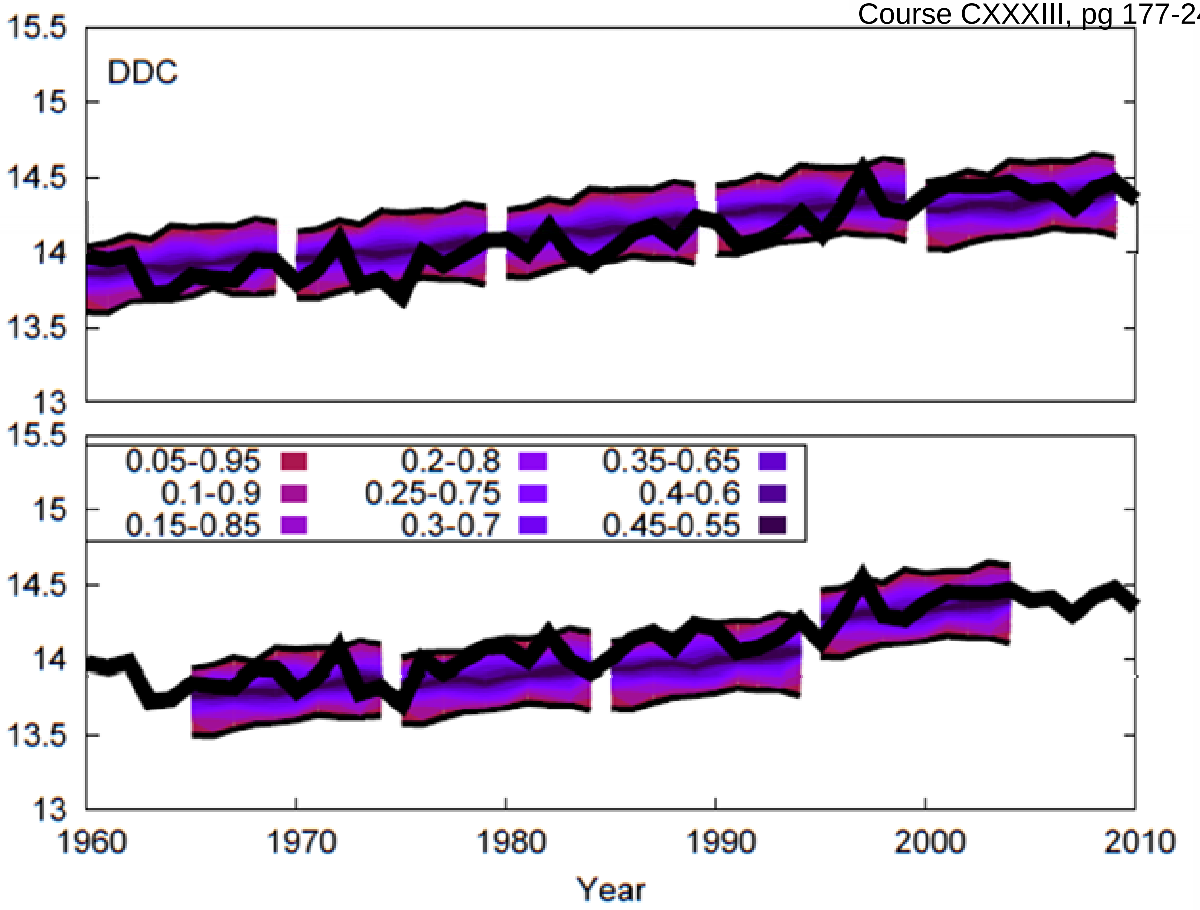


FIG. 5. Forecast distribution for every fifth launch from the direct dynamic climatology (DDC) model. The 5-95th percentiles are shown, obtained from kernel dressing, with Gaussian kernel parameters that minimise the mean Ignorance as a function of lead time. The HadCRUT3 verification is shown in black.

Dressed Dynamic Climatology Decadal Forecasts:
a better option for blending than the static climatology?

E Suckling & L A Smith (2012) Empirical Probability Benchmarks. JoC (in review)

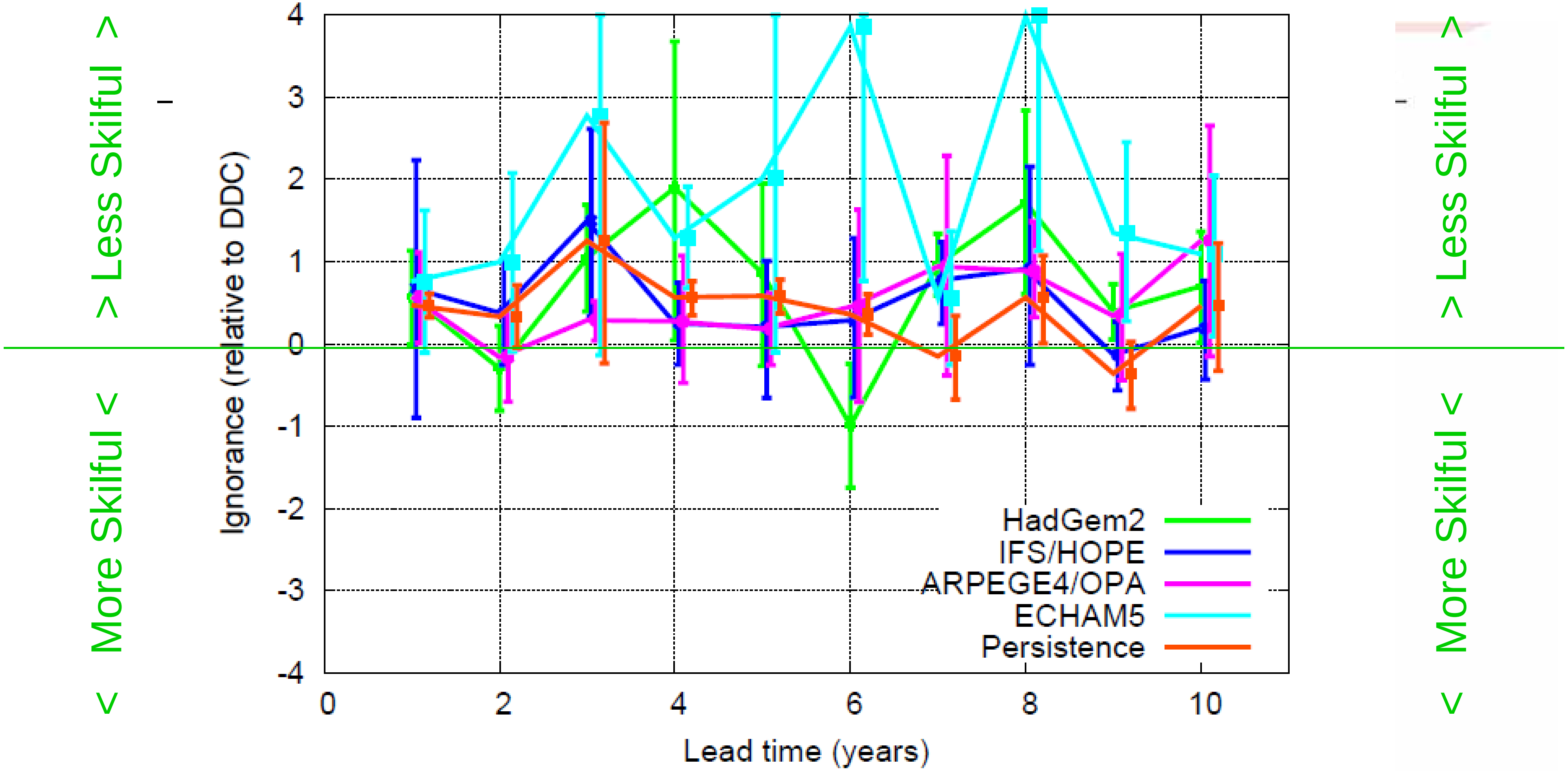


FIG. 11. Ignorance of the ENSEMBLES models relative to DDC as a function of lead time.

GCMs do not outperform this empirical model (even in GMT).

Does this trouble you?

E Suckling & L A Smith (2012) Empirical Probability Benchmarks. JoC (in review)

Questions of Perspective

Empirically based probability forecasts systematically outperform GCMs.

Does this bother you? Or cheer you?
Operationally, why not welcome this source of added skill?

One can argue that ultimately simulation GCMs are our only hope for long range climate prediction; but one cannot argue that today's GCMs they are appropriate for climate services *a priori*.

There are interesting ethical, legal, and rational questions surrounding the provision of information known likely to mislead, but perhaps our central concern should be maintaining the credibility of science-based policy and science-informed action.

It would be interesting to revisit Thompson (1957) and ask where investment is likely to have the largest immediate reward in climate services (purely in terms of deliverables on the table in 2014).

Thank you

J Bröcker, LA Smith (2007) Scoring Probabilistic Forecasts: The Importance of Being Proper *Weather &Fore*, **22,** 2

H Du and L A Smith (2012) Parameter estimation using ignorance *Physical Review E 86, 016213*

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

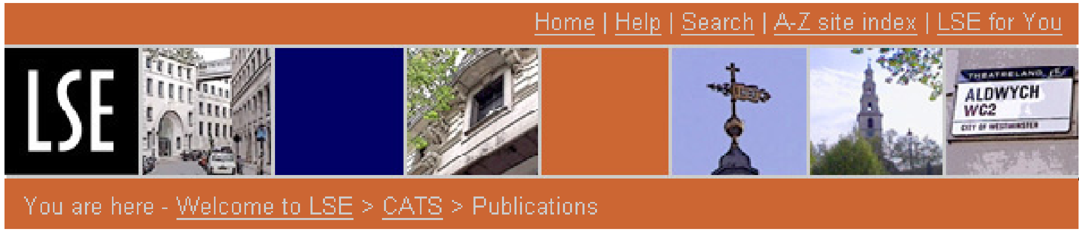
Is it plausible to provide a PDF of hottest or stormiest summer day in 2080's Oxford???

◉ 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



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.

Consider an Empirical Forecast Using e-RAP

For simplicity:
Consider the last 50 observed annual first differences in GMT
Add each to the current year to get an ensemble for next year.
Use two-year differences to get an ensemble for two years hence.
And n-year difference to get an ensemble for n years hence.

Dress as with the simulations models.

(never use the observation of a year within the forecast range at any point, of course)

Select a Skill Score for Probability Forecasts

We will examine IGN = $-\log_2(P(X))$ I.J. Good (1952)

J Bröcker, LA Smith (2007) Scoring Probabilistic Forecasts: The Importance of Being Proper *Weather and Forecasting*, **22** (2), 382-388

Note that:
(a) Smaller is better;
(b) IGN = 0 implies no skill beyond the reference forecast.

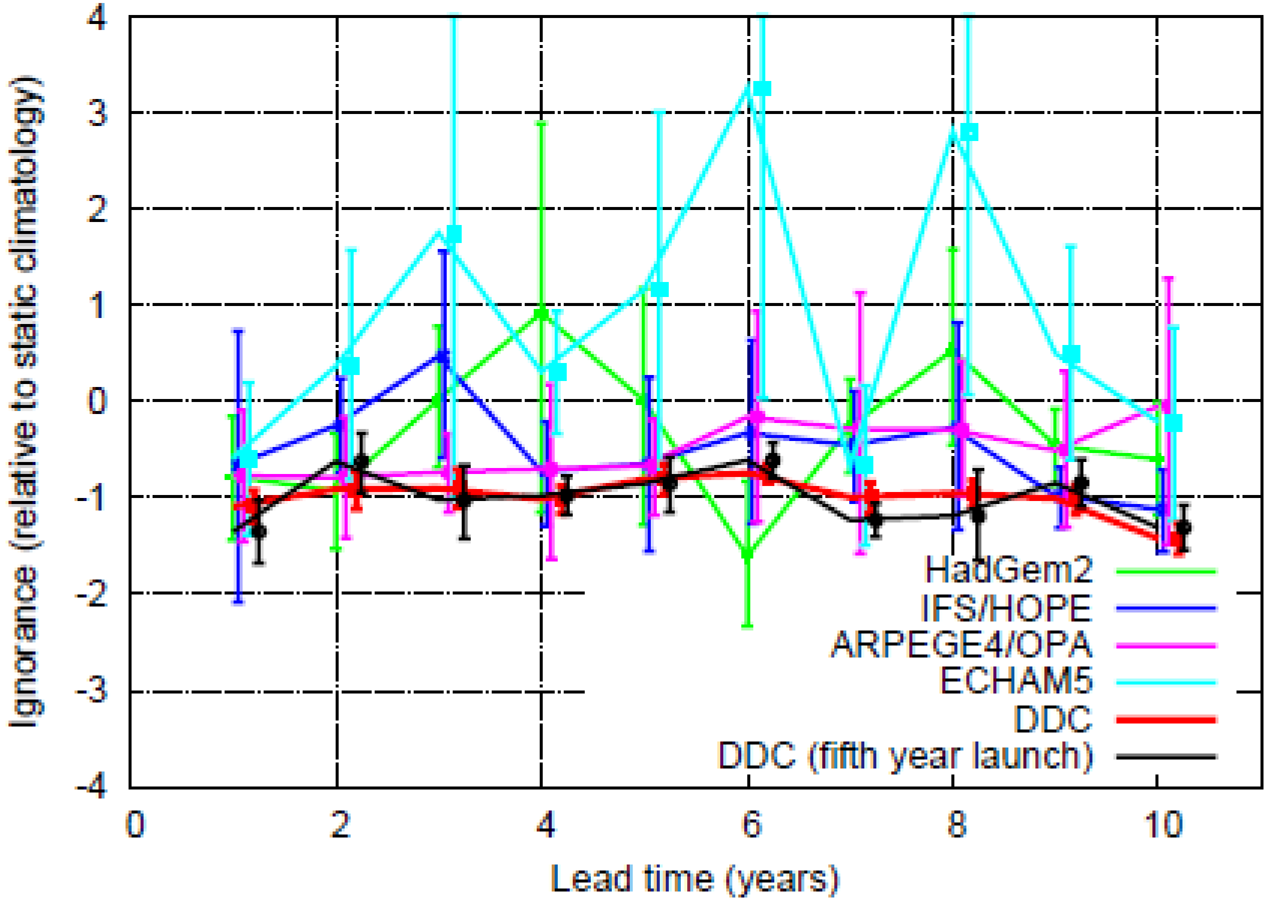


FIG. 8. Ignorance as a function of lead time relative to the static climatology for each of the four ENSEMBLES hindcast systems and the DDC model.

This empirical model outperform GCMs even in GMT.

E Suckling & L A Smith (2012) Empirical Probability Benchmarks. JoC (in review)

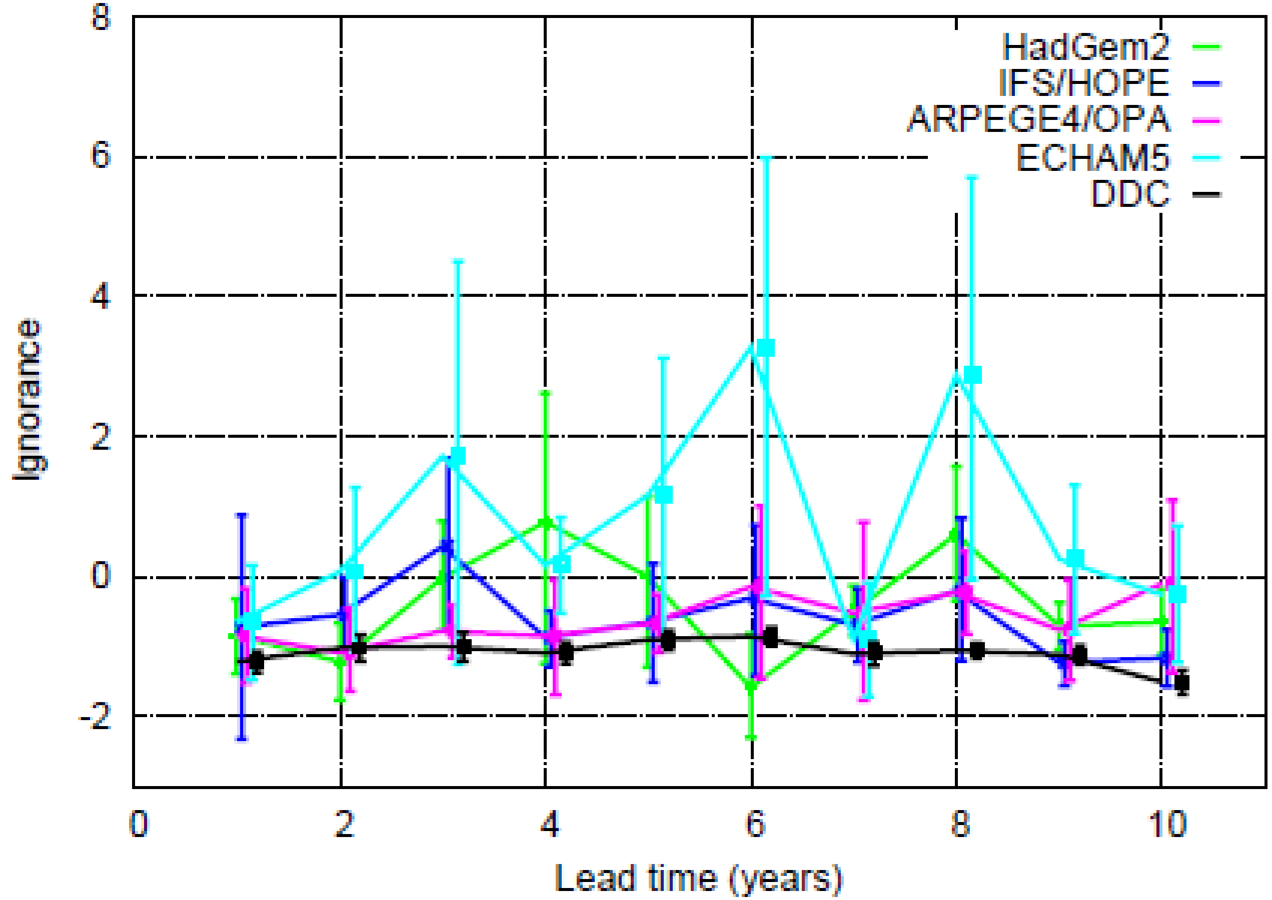


FIG. 6. Ignorance as a function of lead time for each of the four ENSEMBLES multi-model hindcast systems and the DDC model, using a true leave-one-out methodology to fit the kernel parameters. The bootstrap resampling intervals are illustrated at the 10-90th percent level.

This empirical model outperform GCMs even in GMT.

E Suckling & L A Smith (2012) Empirical Probability Benchmarks. JoC (in review)

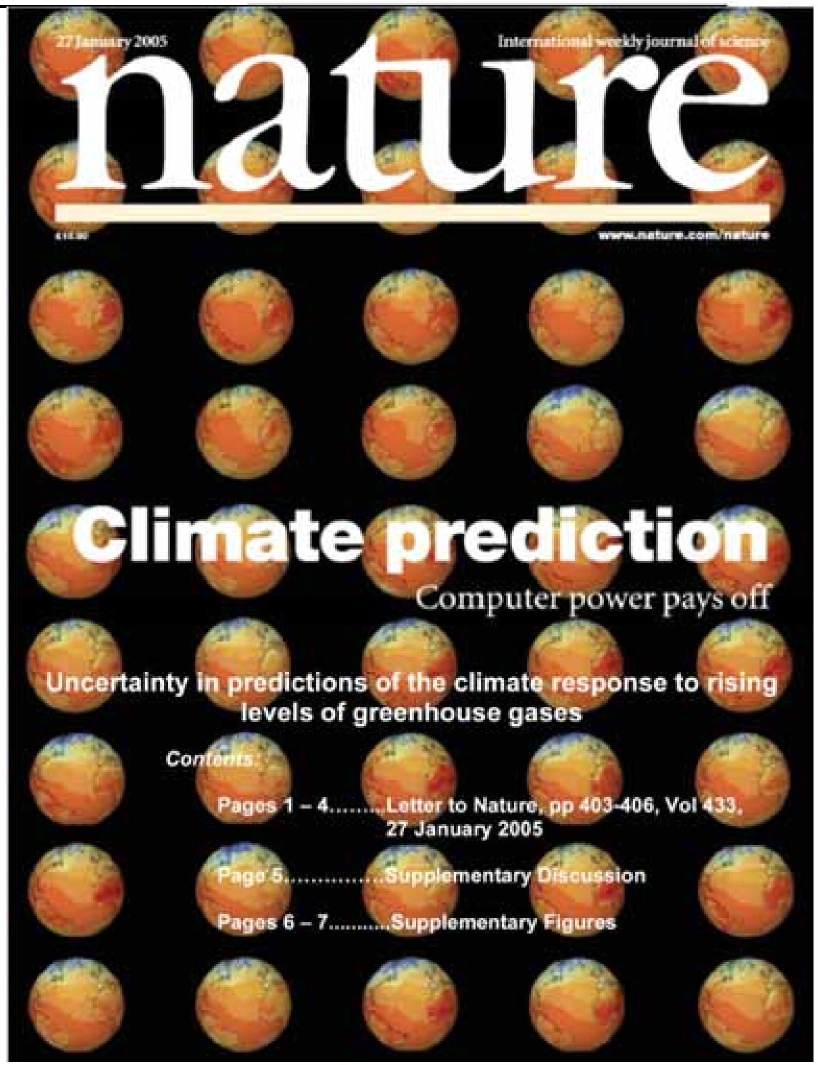
What is the aim of Climate Modelling?

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?

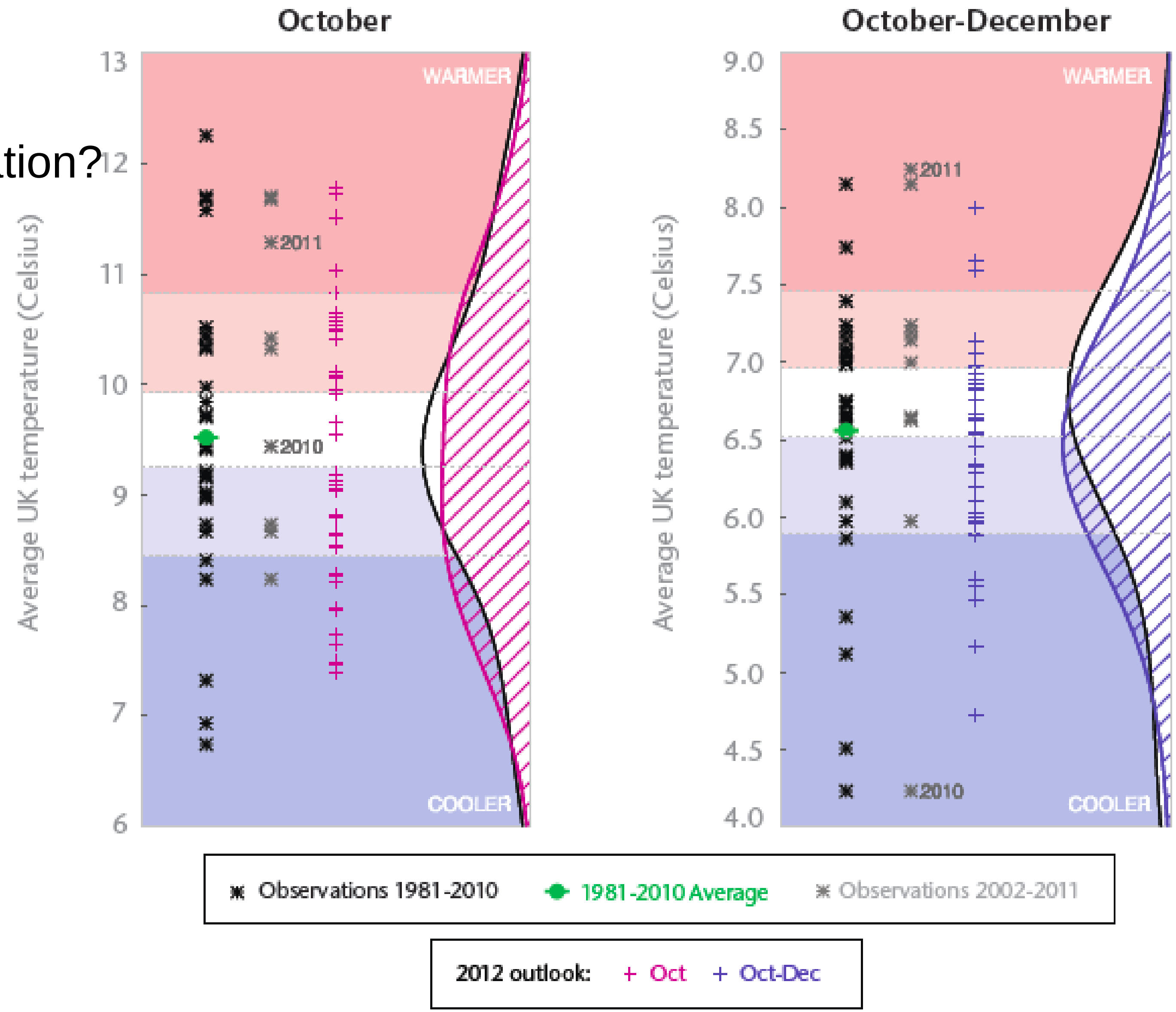


DA Stainforth, T Aina, C Christensen, M Collins, DJ Frame, JA Kettleborough, S Knight, A Martin, J Murphy, C Piani, D Sexton, L Smith, RA Spicer, AJ Thorpe, M.J Webb, MR Allen (2005) Uncertainty in the Predictions of the Climate Response to Rising Levels of Greenhouse Gases *Nature* 433 (7024): 403-406.

The UK MetOffice Queuing the Right U.

Skill?
Value?
Expectation?

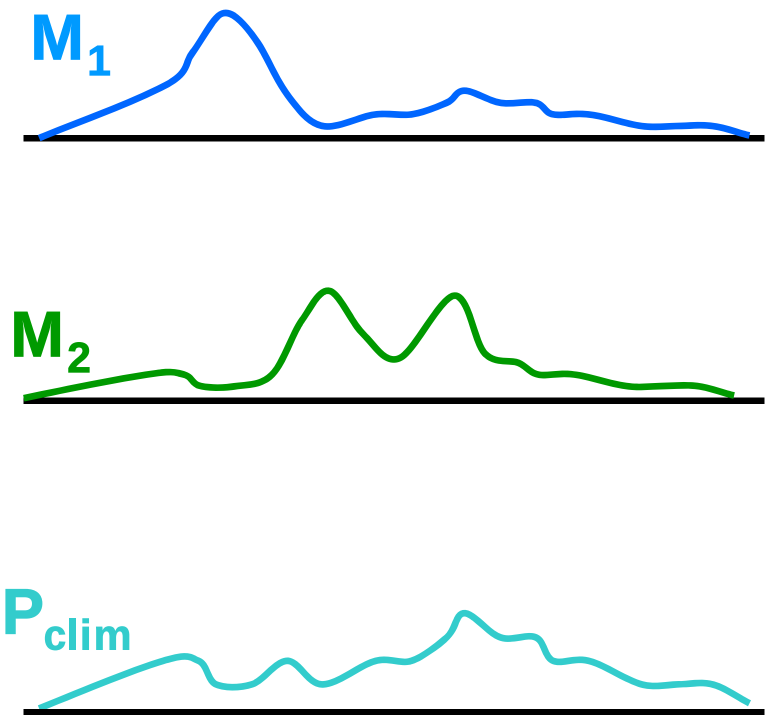
Time to Value?

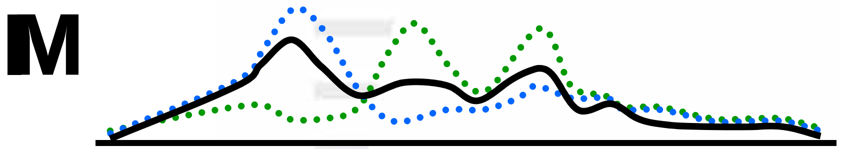


http://www.metoffice.gov.uk/media/pdf/n/3/A3-plots-temp-OND.pdf

Multi-Model Ensembles In - Predictive Distributions Out
(3) Model Distributions to Multi-model PDFs

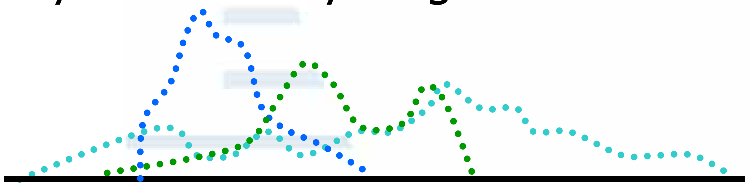
Is this Bayesian if I believe neither "PDF" reflects reality?
And might I then be allowed more flexibility w/o penalty?





$$\mathbb{M} = \omega_1 M_1 + \omega_2 M_2$$

But why not fit everything at once?



$$\mathbb{M} \overset{?}{=} \omega_1 P_1 + \omega_2 P_2 + (1-\omega_1-\omega_2) P_{clim}$$

The answer for seasonal forecasting goes back to the size of the forecast-outcome archive.

?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)

When the best available model is not adequate for purpose

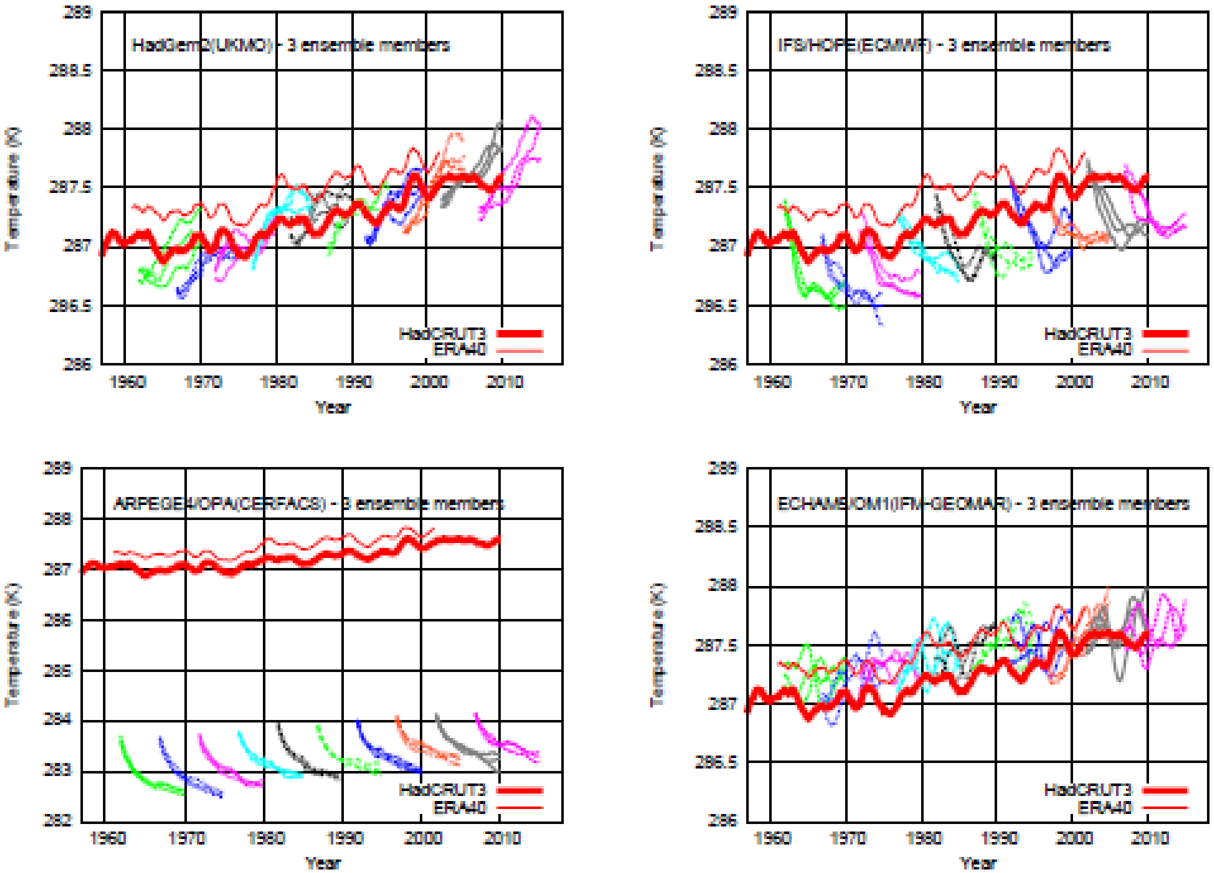


FIG. 1. Global mean temperature (2 year running mean applied) for the four forecast systems - HadGem2 (UKMO), IFS/HOPE (ECMWF), ARPEGE4/OPA (CERFACS) and ECHAM5 (IFM-GEOMAR) - that form Stream 2 of the ENSEMBLES decadal hindcast simulations (Doblas-Reyes et al. 2010). HadCRUT3 observations and ERA40 reanalysis are also shown for comparison. Note that the vertical axis for the ARPEGE4/OPA model is different to the other three panels, reflecting the larger bias in this model.