the London School of Economics and Political Science

Smith institute
for industrial mathematics and system engineering

Operational Approaches to Managing Weather Risk: From Hours to Decades

22 June 2004

London School of Economics
Centre for the Analysis of Time Series

Supported by

the London School of **Economics** and **Political Science**

Smith *institute* *for industrial mathematics and system engineering*

Operational Approaches to Managing Weather Risk: From Hours to Decades

22 June 2004
London School of Economics

Introduction

The Centre for the Analysis of Time Series (CATS[1]) at LSE is leading one of the Smith Institute[2] Faraday research projects: DIME (Direct and Inverse Methods in End-to-End Environmental Estimation)[3]. The principal investigator on this project and Director of CATS, Dr Leonard Smith, is also involved with the NOAA[4] THORPEX Research Program[5], which is directed towards serving society's needs for weather and water information, the Framework 5 DEMETER[6] project on seasonal forecasting, and the Framework 6 ENSEMBLES[7] project on forecast reliability on all time scales.

The technology emerging from these projects, namely probabilistic forecasting, can place environmentally induced risk in the user domain to aid decision-making. Whilst there exist a number of important impact studies, the existing plethora of environmental data and emerging forecasting tools remain under-utilised by government and commercial organisations. There is a need to 'translate' this technology into society and commerce in general.

The Smith Institute and CATS are scoping applications of this technology and plan to increase awareness of its potential benefits in the wider industrial and commercial communities. This event is the first in a series of awareness raising workshops for industry; future meetings will be held both in the UK and in the US. Follow-up activity will be offered to help those who wish to apply these methods in their businesses.

Aims and purpose

Mitigation of weather risk through direct action is an operational reality, even if hedging and insurance options have been put in place. This meeting will survey the application of current and future sources of weather information towards the active management of weather risk in practice. The aim is to provide a realistic evaluation of what is possible with current forecast information, and what is likely to be available in the near future. This includes combining observations, multiple forecasts, and historical information in the context of specific user applications. The value of weather forecast information depends on the particular user application. We will use case studies as spring boards for discussing problems of interest to the participants.

Format

The event is organized to provide, in the morning, an overview of new options for dealing with weather risk; the afternoon will provide an opportunity both for detailed illustrations (through case studies) and the presentation by industry of particular challenges. Further information is available on the workshop web page www.smithinst.ac.uk/Events/WeatherRisk/Programme .

---

[1] CATS: http://www.lse.ac.uk/collections/cats/
[2] Smith Institute: http://www.smithinst.co.uk/
[3] DIME: www.smithinst.ac.uk/Projects/RA-RMS/index_html
[4] National Oceanic and Atmospheric Administration, www.noaa.gov/
[5] The Hemispheric Observing System Research and Predictability Experiment, www.wmo.int/thorpex/
[6] http://www.ecmwf.int/research/demeter/
[7] http://www.smhi.se/sgn0106/if/rc/projects/ENSEMBLES.html

Faraday Partnership

the London School of Economics and Political Science

Smith institute for industrial mathematics and system engineering

Programme

10:00-10:30 Coffee and Registration
10:30-11:30 Introduction and Overview
Beyond Hedging: Examples of Active Management of Weather Risk
11:30-12:00 Information Providers
Weather forecast information in the 21st Century
12:00-12:30 New Challenges, New Needs
12:30-13:30 Lunch
13:30-15:00 Illustrations through Applications (with Discussion)
Using Ensemble Forecasts
(Forecasting Wind Energy Generation)
(Wave Height Forecasts and Off-Shore Safety)
Using Multi-Model Forecasts
(Demand Forecasting: Moving Beyond the Consensus Forecast)
(Likely Precip in the Medium Range)
Early Warning of High-Impact Events
(Extended hot or cold spells)
(Potential Storms)
Options on More Complex Decision Making
(Ship Routing)
(Contingency Chains)
Correlations, Climate, and the Use of Simulation Models
(Its always windy somewhere: Repackaging Risk)
(Seasonal forecast information & DEMETER)
(How reliable are current climate forecasts?)
15:00-15:10 Tea/Coffee
15:10-15:45 Forecast Evaluation & DIME
15:45-16:00 Summary and Follow-up

Expected Outcomes

- Increased industry awareness of the methods emerging from DIME and of the potential impact of applications.
- Higher profile for sources of state of the art meteorological information.
- Methods for comparative analysis of commercially available forecasts.
- Increased number of application domains.

Further information

For further information or to explore potential applications please contact: Melvin Brown, Melvin@smithinst.co.uk Tel: (+44) (0)7980 580556 or Lenny Smith, L.Smith@lse.ac.uk.

Faraday Partnership

Operational Approaches to Managing Weather Risk

Leonard Smith, Jochen Broecker, Liam Clarke, Devin Kilminster
Centre for the Analysis of Time Series,
London School of Economics

lsecats.org

Questions we aim to address today

How to best exploit modern weather forecast information in managing operational weather risk?

How can you make better use of weather forecasts?

Which (if any) are the most useful forecasts?

Why might you buy several?

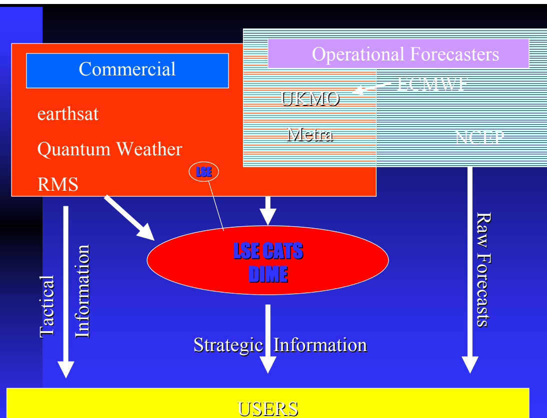
Can we determine if the forecasts pay for themselves?

What CATS wants from today

Potential future proof of value studies.

Information on how to make the DIME site more useful.

Expand the CATS community of numerate users.



Overview

- Weather Forecasts and Uncertainty
- The Role of CATS and DIME
  - Neutral Forecast Evaluation
  - Proof of Value Studies
  - Design of Operational Strategies
- Examples of Operational Weather Risk
  - Wind Farm Production
  - Significant Wave Height
  - Electricity Demand
- Beyond Probability Forecasting
  - Challenges of imperfect models

5th May

weather.com

10-Day Forecast				
Printable Forecast			Daytime High / Overnight Low (°F)	Precip.
Today May 05		Sunny	N/A°/59°	0 %
Thu May 06		Sunny	86°/62°	0 %
Fri May 07		Sunny	88°/63°	0 %
Sat May 08		Sunny	88°/63°	10 %
Sun May 09		Mostly Sunny	88°/65°	20 %
Mon May 10		Partly Cloudy	86°/64°	20 %
Tue May 11		Partly Cloudy	84°/65°	20 %
Wed May 12		Partly Cloudy	83°/63°	30 %
Thu May 13		Mostly Sunny	85°/63°	10 %
Fri May 14		T-Storms	81°/62°	60 %

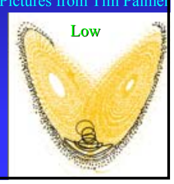
Last Updated Thursday, May 6, 2004, at 1:30 AM Eastern Daylight Time

We know these are not equally reliable!
But sometimes they are more reliable than others.

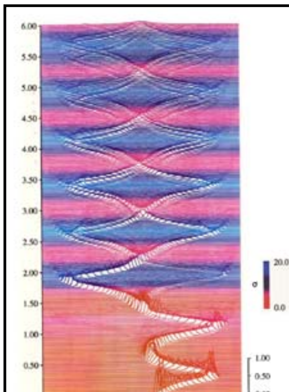
What about the showers forecast for Friday?
But *how big a grain of salt* should we assign to each forecast?

Predictability

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*

This is a probability forecast.

But the value of a probability forecast must be computed in terms of *your* cost function. (meteorological skill is neither necessary nor sufficient!).

The key aim is to find the time scales on which the forecast has information relevant to *your* operations.

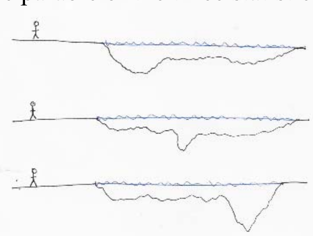
And since the models are imperfect, how can we judge that they are worth their cost to *you?*

Proof of Value Studies

The most useful demonstrations are:
- End-to-End
- Expressed in the your language
  - MW not mb
  - Ideally quantified by £ in (and £ out)
- Immediately deployable (data streams exist)
- Illustrated on a known historical period/location

But why not simply use a consensus forecast?

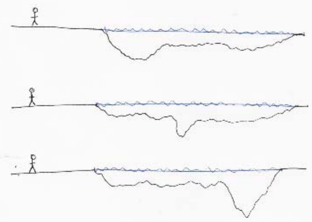
The parable of the three statisticians



Three statisticians wish to cross a river, each has a forecast of depth which indicates death; but they take the consensus forecast (the ensemble mean), and find that the forecast depth is acceptable.

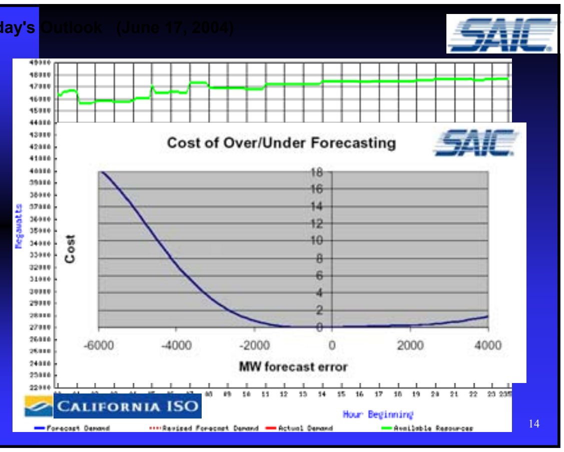
And so they attempt to cross...

The parable of the three statisticians



Key points:
Rather asymmetric cost function
  (overly deep by 2 inches << shallow by 2 inches)
RMS error of the forecast is irrelevant
Unclear how to best combine (imperfect) models.

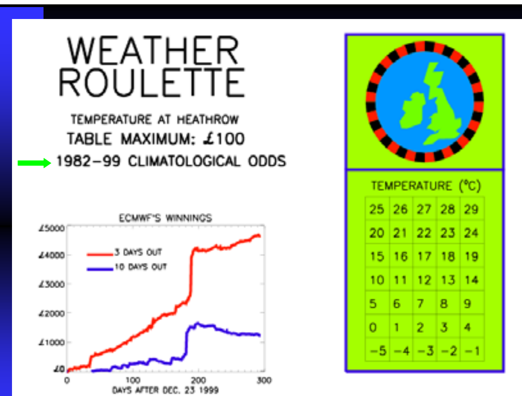
Today's Outlook (June 17, 2004)

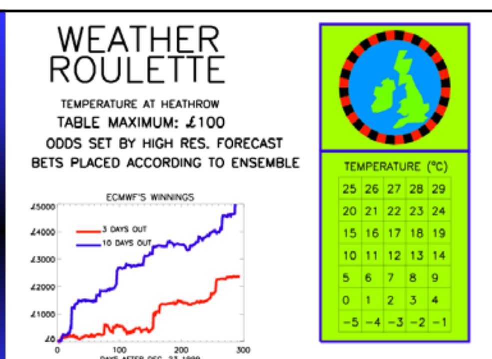


An illustrative example

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

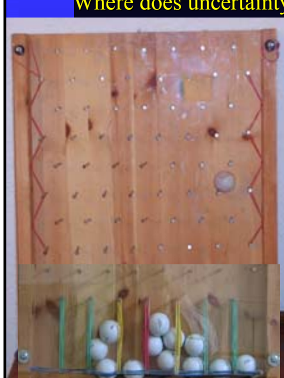
How would a probability forecast based on the ECMWF ensemble forecast fare against a house that set its odds using climatology?





How can we measure this kind of skill?
How do we allow a fair comparison of different forecasts?

Where does uncertainty come from?



Think of forecasting past each row of nails as another day ahead: we can do well for a day or two, but should we appear confident at day 8?

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

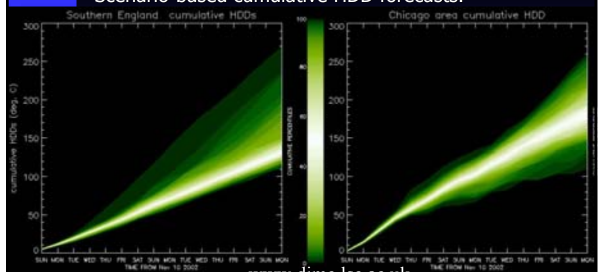
Operational Weather Ensembles:
US and European Services: 1992
Canada: Now
Japan: 'Soon'

The NAG Board (Not a Galton Board)

Tropical Storm/Hurricane “Cone of Uncertainty”

The Presentation of Uncertainty (The Weather Channel)

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 traditional forecast.

Managing operational weather risk:

Relative Income in Wind-energy Experiment

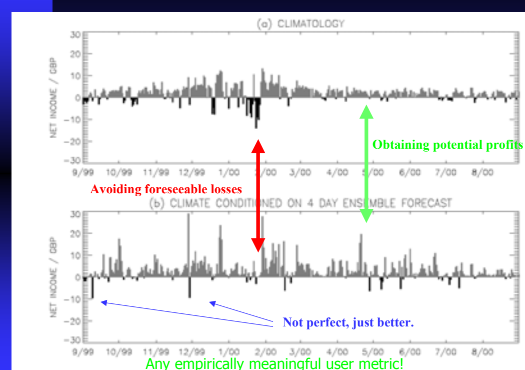
Real electricity prices (30 min)

Real wind data (1 min)

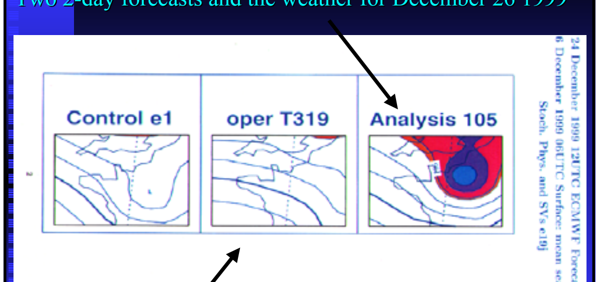
Real ECMWF Forecasts

Fictional Wind Farm

Wind farm profits

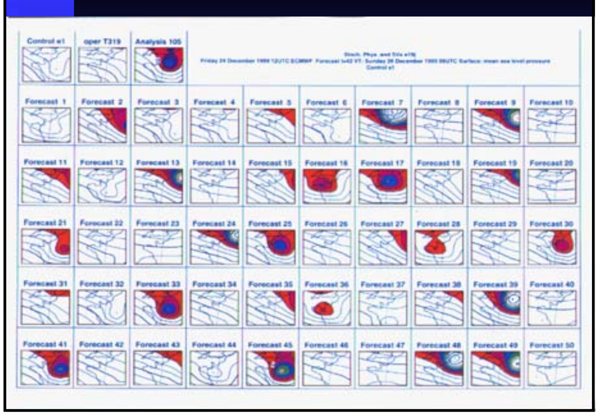


Two 2-day forecasts and the weather for December 26 1999

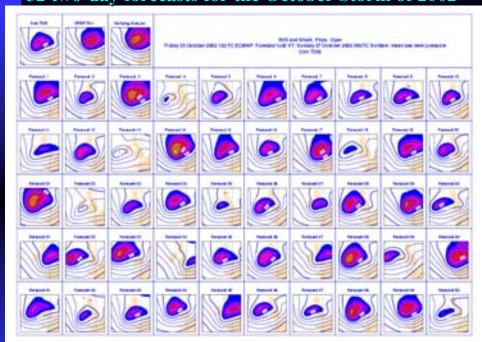


The single forecast from the “best” model

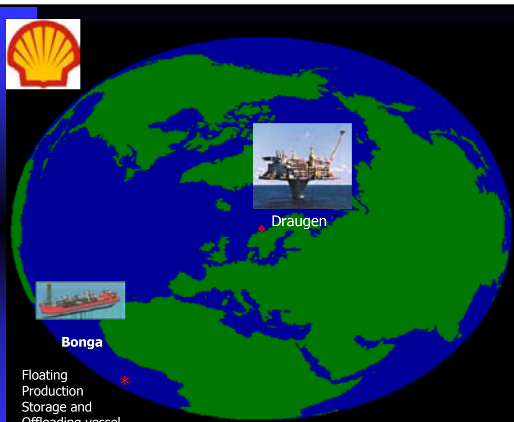
Early warning of the French storm of Christmas 1999.



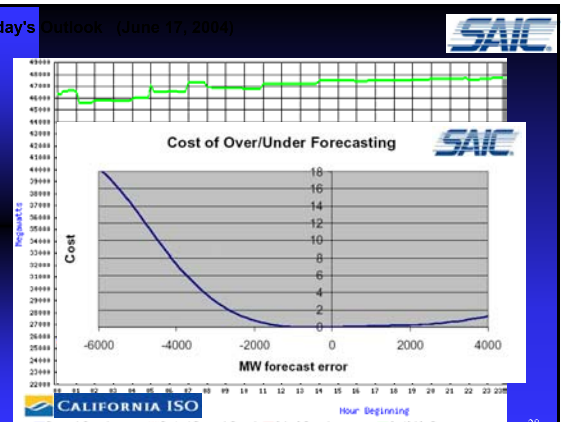
52 two-day forecasts for the October Storm of 2002



Each “postage stamp” represents a golf ball, and most have storms…



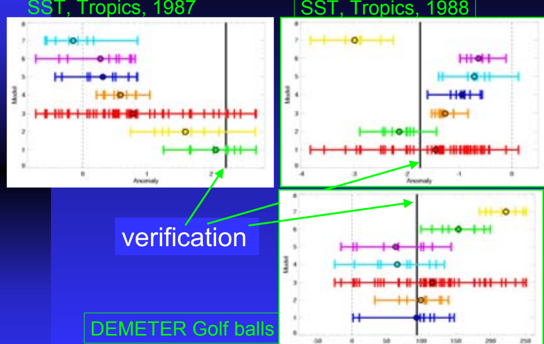
Today's Outlook (June 17, 2004)



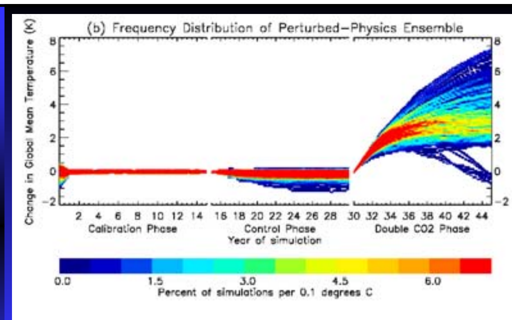
Seasonal Forecasts and DEMETER

SST, Tropics, 1987

SST, Tropics, 1988



European Centre for Medium-Range Weather Forecasts (ECMWF)



This distribution represents the evolution of several thousand full GCM climate model runs; but what should we call it?
It is not a PDF…

www.climateprediction.net

Sources of information:

Multiple initial conditions (golf balls)

Different models (different boards of nails)

Different start dates

But the value of the information is determined by the decisions made; these decisions may be complex.

The remains of the day:

Weather information
Operational Weather Risk at NG Transco

Examples, Illustrations, and Discussion

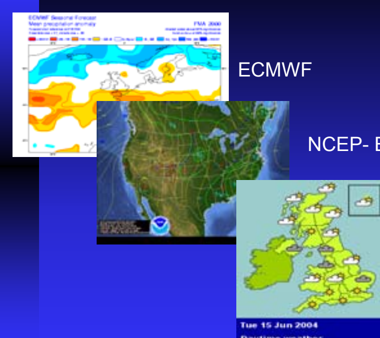
Weather Forecast Information

Overview

- Sources of weather forecast information
- Determining forecast value
- Proof of Value
- Verification Data

Forecast Characteristics

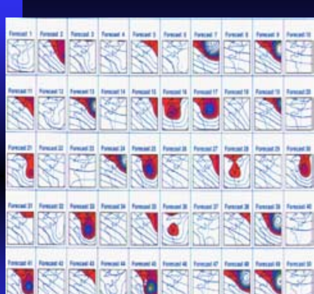
- Many sources of weather forecast information
- What are the distinguishing and relevant features for the user
  - geographical coverage
  - time scales
  - format
  - cost



ECMWF

NCEP- ETA

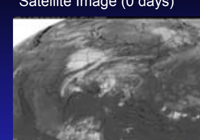
UKMO



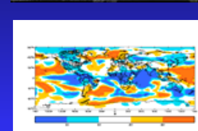
ECMWF Ensemble

ECMWF High Resolution Run

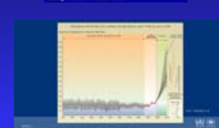
Satellite Image (0 days)



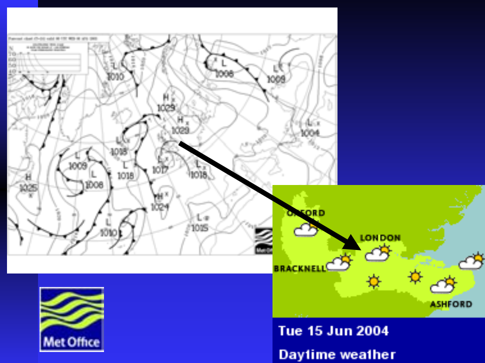
UKMO 2 day forecast



ECMWF seasonal temperature



IPCC



CATS Weather Information Sources

Institution	Description	Type
ECMWF	European Centre for Medium-Range Weather Forecasts	International
NCEP	National Centres for Environmental Prediction	National Non-commercial
Met Office	UK Meteorological Centre	National
Metra	New Zealand Meteorological Centre	National
EarthSat	Private company	Commercial
Weather Channel	Private company	Commercial
Quantum Weather	Weather Risk Forecasting Company	Commercial
Demeter	Seasonal Forecasting project	Academic
ClimatePrediction	Climate change project	Academic

The question must be

Do these forecasts have value for the user?

Demonstrating Value

- A forecast's proof of value if achieved by showing that including the forecast information significantly benefits the user
- How might we show this?

CATS' Proof of Value

Weather Information

User Information

Cost Function

Weather Verification

Reality

Profit (?)

User Verification

Wind Generation (High Res)

High Resolution Forecast

Electricity Market

Amount to bid

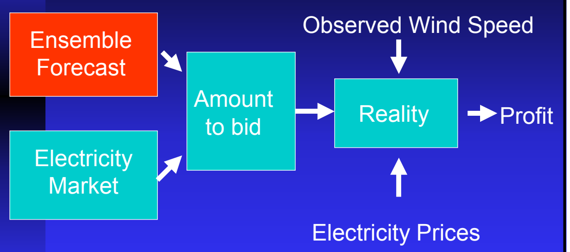
Observed Wind Speed

Reality

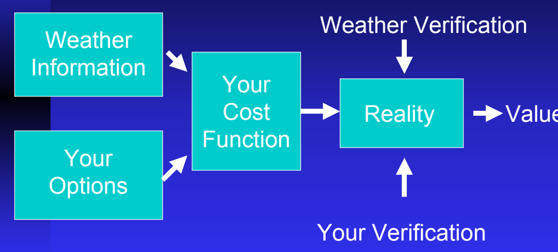
Profit

Electricity Prices

Wind Generation (Ensemble)



Your Example



Additional Options

- Inputs
  - multiple models
  - climatology
- Processing
  - different kernel methods
  - bias correction
- Action
  - different cost functions

Determining Value

- Weather Forecasts
- User Information
- Cost Function
- Verification
  - meteorological
  - economic

Weather Forecasts

- Size of forecast archive limits the extent of proof of value study
  - statistical significance
  - performance (better dressing)

User Information

A user’s cost function usually requires additional information, for example

- local meteorological observations
- plant efficiency
- true demand

Cost Function

- Cost function maps input variables into user action
- CATS can assist users in developing their cost functions

Verification

- Availability of meteorological data
- Availability of user verification
- What is the target
  - model analysis
  - WMO station data
  - economic variables / user targets

DIME

- The DIME website exists to contrast and value numerical weather forecast models
- DIME produces general measures of skill for many different users

CATS

- Specific proof of value studies can be carried out by CATS
  - Wind generation
  - Wave height forecast
  - Electricity generation
  - Weather Roulette
- Just need user problems

Summary

- Forecast information is available
- Expertise exists to process it
- Looking for industry problems to apply the methods to

Using Ensemble Products

Overview

Ensembles and Applied Probability Forecasting
- Wind Farm Production
- Significant Wave Height

• Beyond Probability Forecasting
- Distribution Forecasts and Electricity Demand

• Conclusions

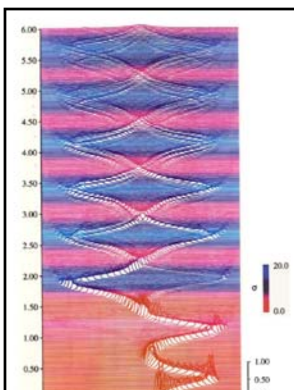
What is a Probability Forecast?

The options:
• empirically accountable forecast
• subjective (but well defined) image of our uncertainty/belief
• any curve which has been normalised to unit area and mathematically manipulated in a certain manner.

In no case can a single probability forecast be judged 'right' or 'wrong'.

We need to look at a series of forecasts and contrast predicted probability with observed relative frequency.

You should feel free to do this if someone gives/sells you a "probability forecasts".



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

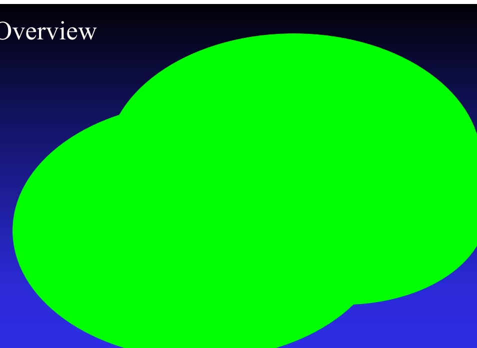
Given a probability forecast, a cost function, and a set of options we can either

• maximise expected utility
• find the best historical option

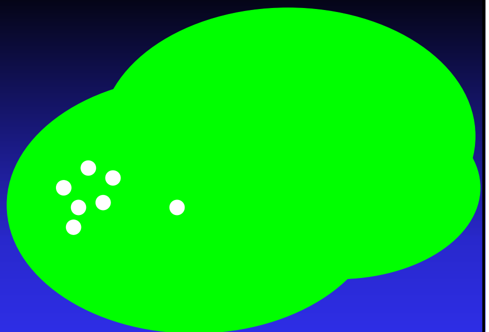
We can then look at the savings a given strategy implies, and see if it justifies its costs.

Initially, think of each ensemble as a scenario…

Overview



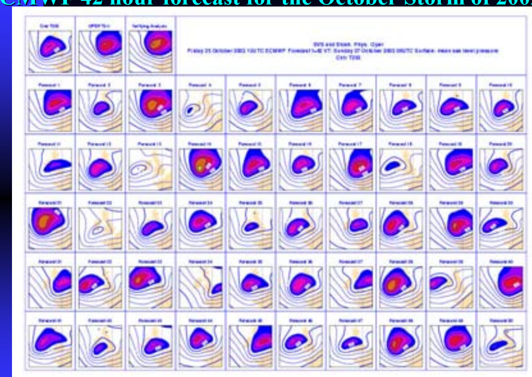
Traditionally, a using Monte Carlo ensemble has been likened to employing a shotgun



Perhaps a better analogy is that of an ensemble of golf balls, placed on the green by a player who cannot see the hole.

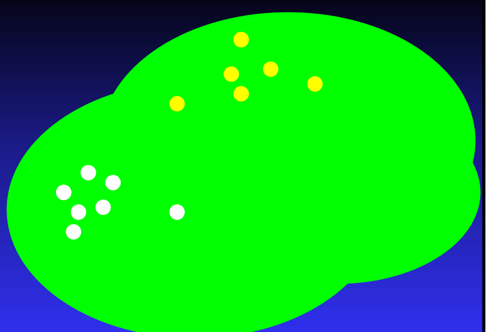
7

ECMWF 42 hour forecast for the October Storm of 2002



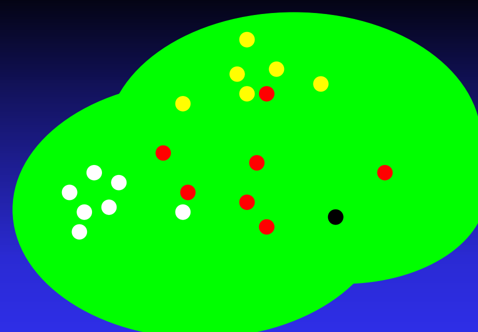
Here each simulation looks physically reasonable; most have storms: I do not want my car sitting under a tree.

8



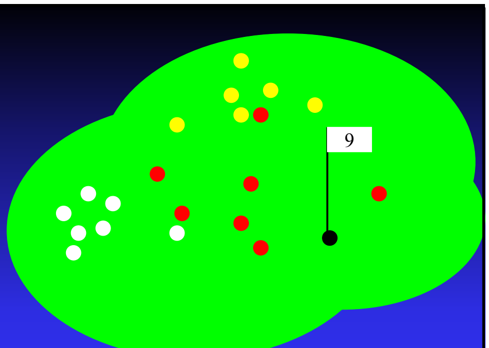
Adding a second player (model) can help.
Note that these points are sampling different distributions,
And even if the input noise is Gaussian, the output(s) is not!

9



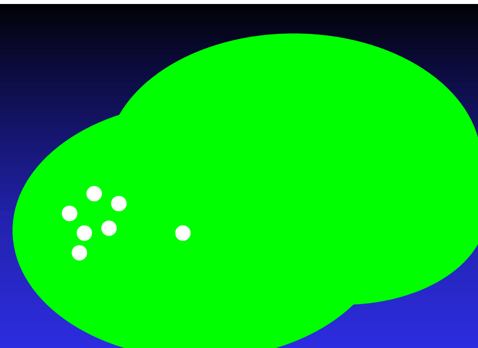
How can we best use these distribution in practice (rather than in improving our theory)?
(Extract their information content, not their component values)

10



Historical data on the target (and forecasts) is the key.

11



First off, we should not treat the distribution as delta functions!

12

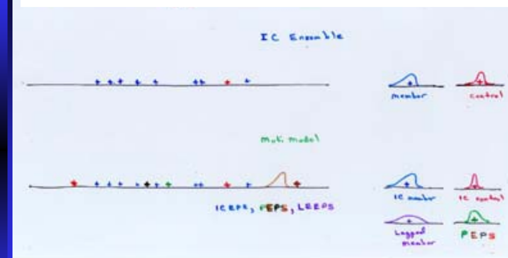
"Dressing" a Simulation

(from simulations to forecasts)

- Kernels based on historical forecast errors or forecast goals can be used to "dress" multi-model runs:
  - For a single forecast (one member ensemble) historical errors can be added directly to current forecast to produce a "statistical ensemble"
  - For an ensemble forecast minimum ignorance kernels (or historical "best member" errors).

13

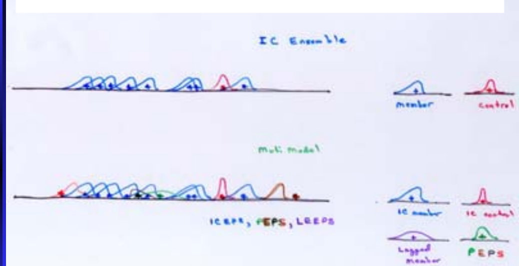
Dressing an Ensemble as a Collection of Scenarios:



Each class of ensemble member is dressed with its own kernel.
Members of an EPS, DEMETER, a PEPS or LEEPS are easily included.

14

Dressing an Ensemble as a Collection of Scenarios:



Although this yields a smooth distribution function, it is unlikely that a physical scientist would want to call it a probability forecast, as the models are (each) known to be imperfect *a priori*.

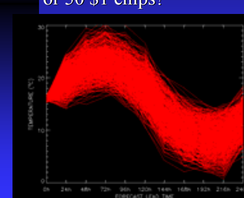
Bayesian model averaging fails to be internally consistent for the same reason.

15

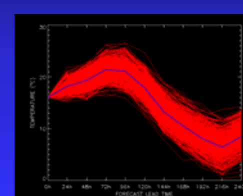
Reasons for Scenario Dressing:

- "Probability" Forecasts:
  - Accounting for finite ensemble size
  - Accounting for typical model inadequacy
- Equal Counting Statistics:
  - A fair comparison between EPS of different sizes/compositions
  - A fair comparison between an EPS and a single hi res BFG simulation

Would you rather play informed roulette with one $50 chip or 50 $1 chips?



100 Element Dressed EPS Forecast



100 Element Dressed BFG Forecast

16

In general, a skill scores should reflect the information content of our forecasts. (Ideally the relevant information.)

RMS error and correlations with a mean are at best irrelevant!

Gambling skill scores are relevant:

Brier score: $\text{Prob(verification)}^2$

Ignorance: -log [Prob(verification)] (see R&S MWR 2002)

Both are "proper" skill scores (you bet according to your forecast)

Differences in ignorance reflect *expected* wealth doubling times

Your cost function provides the most relevant metric for the score.

17

Relative Income in Wind-energy Experiment

Real electricity prices (30 min)

Real wind data (1 min)

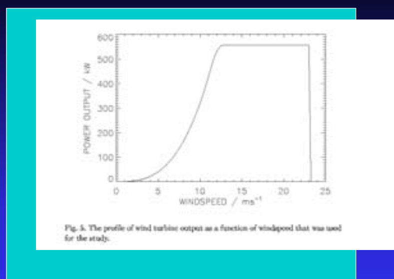
Real ECMWF Forecasts

Fictional Wind Farm

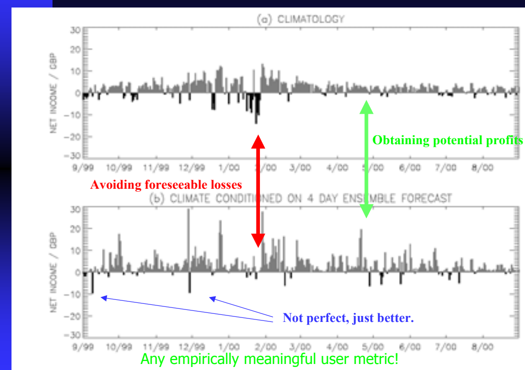
Forecast In:	Targets (historical data):
ECMWF Ensemble	Observed Wind
	Observed Electricity Price

18

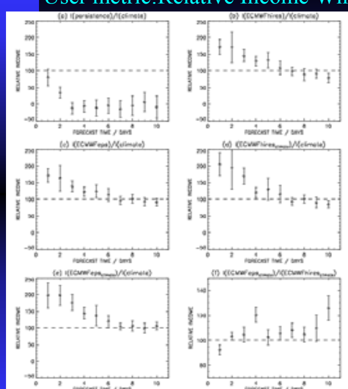
Relative Income in Wind-energy Experiment



Wind farm profits



User metric:Relative Income Wind-energy Experiment



Real electricity prices (30 min)
Real wind data (1 min)
Real ECMWF Forecasts
Fictional Wind Farm

From day 3 to day 10 the ECMWF EPS has the largest expected weekly income.

Relative Income in Wind-energy Experiment

Real electricity prices (30 min)

Real wind data (1 min)

Real ECMWF Forecasts

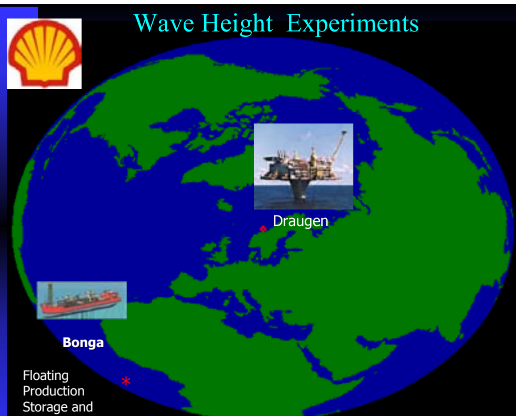
Fictional Wind Farm

Forecast In:	Targets:
ECMWF Ensemble	Observed Wind
	Observed Electricity Price

Result: Ensemble Products yield 20% increase in profits at three day lead time

Questions?

Wave Height Experiments



Wave Height Experiment

Real ECMWF Forecasts

Observed Wave Data

Forecast In:	Targets:
ECMWF Ensemble	Significant Wave Height
ECMWF hi-res	Ship Heave

ECMWF Wave Model

- The global ECMWF atmospheric model is coupled to a global wave model.
- The wave model is driven by each of the 51 members of the ensemble forecast to produce an ensemble of 51 wave forecasts.

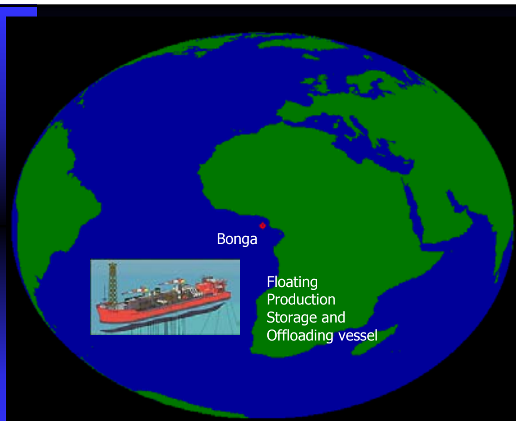
25

Wave Variables Forecast at ECWMF

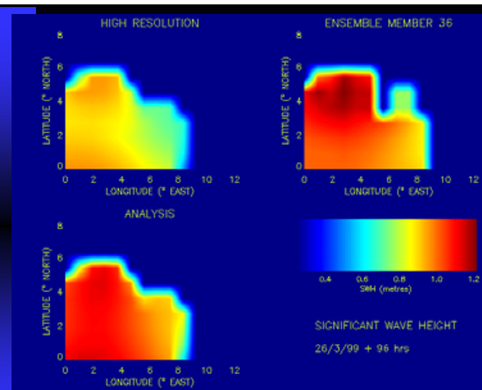
- x significant wave height
- peak period of 1D spectra
- mean wave period
- x swell height
- mean swell period
- mind wave height
- mean wind wave period
- wave slope
- mean wave direction

- …and more
- 2D wave spectra forecast but not archived

26

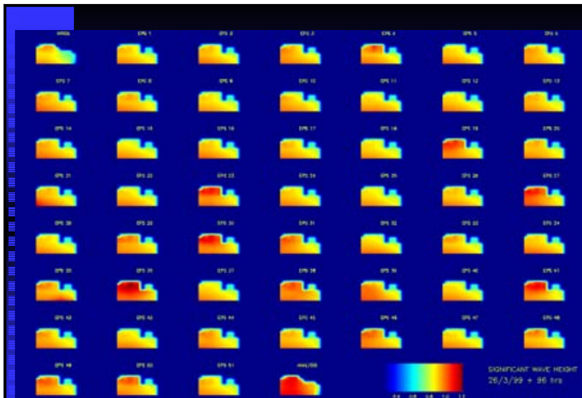


27



To verify against the analysis or your observations?

28

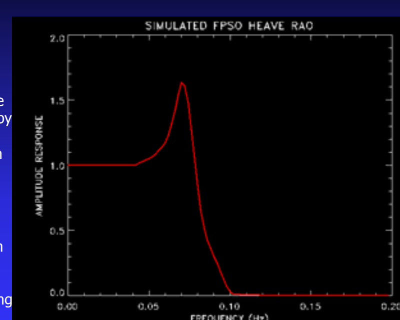


29

Target: FPSO Response Amplitude Operator

Heave data for the FPSO is obtained by passing (30 min) buoy data through this operator.

Source: WAMIT-MOSES study from Hull University (simulated using numerical modelling Package)



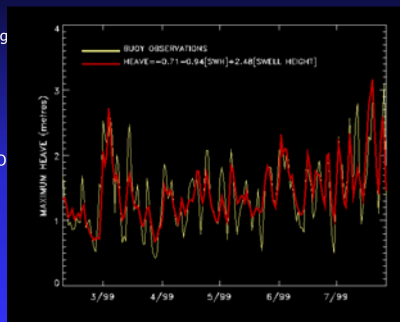
30

MOS Model of FPSO 24 hr Max Heave

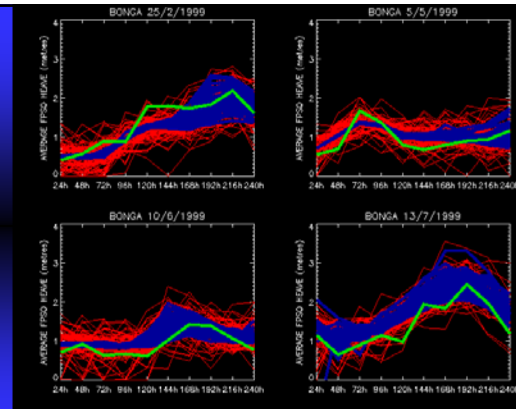
Fitting performed using ensemble mean at a lead time of 24 hours.

"Buoy observations" calculated using wave spectra measured by the buoy and the FPSO RAO.

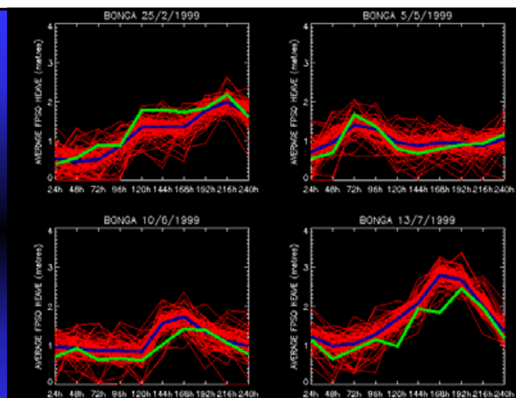
ECMWF Heave is a linear combination of model SSH and SWH.



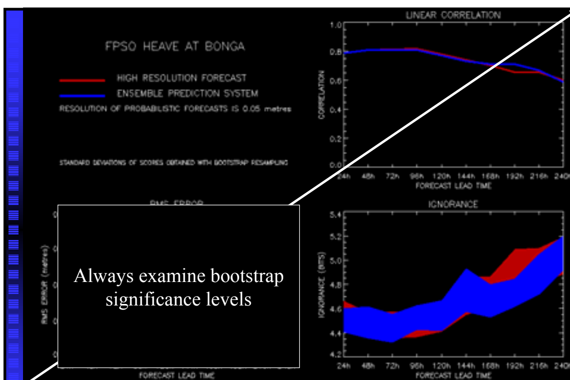
This fit is then used to forecast heave from each ensemble member



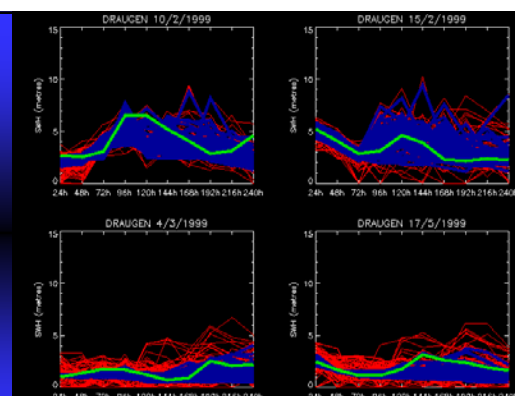
dynamical ensemble is blue, dressed ensemble is red, verification (buoy) is green



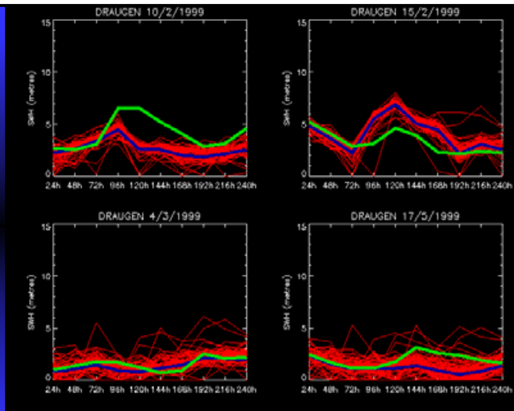
dynamical ensemble is blue, dressed hi-res forecast is red, verification (buoy) is green



In this case, the dressed hi res does well.
Of course with a bigger archive or tropical SVs…



dynamical ensemble is blue, dressed ensemble is red, verification (buoy) is green



dynamical ensemble is blue, dressed hi-res forecast is red, verification (buoy) 37

Wave Height Experiment

Real ECMWF Forecasts

Observed Wave Data

Forecast In:
ECMWF Ensemble
ECMWF hi-res

Targets:
Significant Wave Height
Ship Heave

Result: Ensemble Products significantly more effective in the North Sea, Hi-res Products of similar value in Bonga

38

At Draugen the ensemble has significantly more value than the hi-res BFG.

But there is no reason to expect these dressed forecasts will yield reliable probability forecasts (that is, that events which are forecast at 10% will occur with a relative frequency ~ 0.10)

Do we need to treat these as "good probability forecasts"?

The odds on achieving a "good=accountable" probability forecast look rather long; while we already have "good=useful" in-hand!

39

What is a Probability Forecast?

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

Then we can construct a pretty good (empirically) PDF. 40

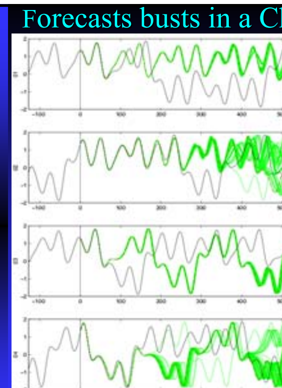
What is a Probability Forecast?

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 sell) only the former!

41

Forecasts busts in a Chaotic Circuit



512 member ensembles
Best known 1-step model
512 step free running forecasts

Short term (weather) forecasts are very skilful.

"Seasonal" forecasts suffer from model drift (and eventually from model irrelevance!)

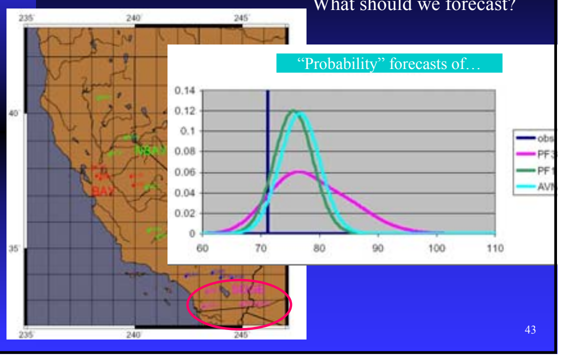
A model can add value as long as it adds information, it need not have traditional "skill."

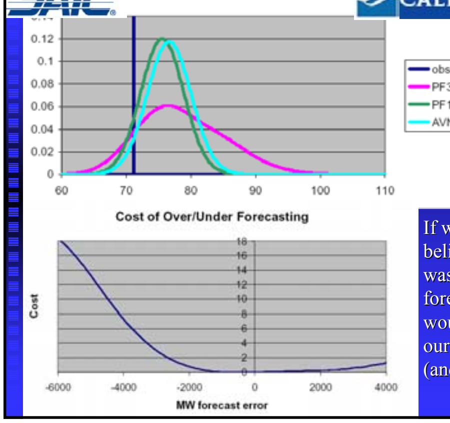
In seasonal forecasts, as in golf, there is at least some level of recurrence...

whereas in climate forecasting...

42

4 regions, 20 stations, one ultimate demand forecast: generation. What should we forecast?





If we really believed that PF3 was a probability forecast, then we would maximise our expected utility (and lose money!).

Expected utility yields an unhelpful results for several reasons:

First it tends to heavily wait the "low probability tails" where even the users cost function may be in error.

Second, we know (empirically) that our best weather forecasts do **not** score well as accountable probability forecasts!

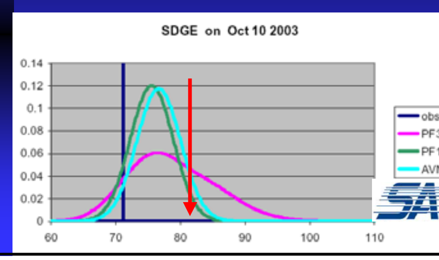
An alternative approach is simply to base decisions on the isopleth of the forecast distribution that is observed (historically) to yield maximum utility.

For Cal ISO, this yields a significant savings over the demand forecasting using the 'consensus forecast'.

Playing Isopleths

An alternative is simply to base decisions on the isopleth of the DF that is observed (historically) to yield maximum profit.

For Cal ISO, this yields a *significant* savings over the demand forecasting using the 'consensus forecast'.



Challenges to probability forecasts from imperfect models

Each of the models is imperfect.

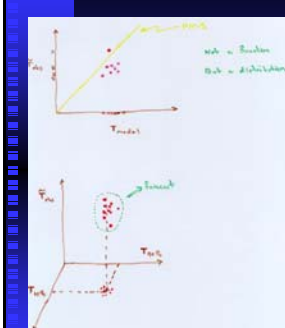
The *ad hoc* assumption that their distribution can be mapped into uncertainty in the verification is unsupported.

We might aim to:
- extract information, not scenarios.
- condition on their joint distribution, rather than some averaging over an *ad hoc* model class…

Or might we ask for less than physically meaningful probability forecast?

Outside the Perfect Model Scenario, the question is the key.

A Forecasting in a Product Space

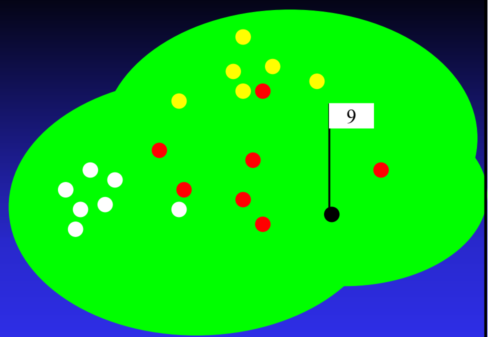


Rather than looking for a function that relates the model variables to an observable, we can consider the distribution of obs directly.

This analogue approach remains useful when model error cannot be ignored.

Treating the Ensemble as a Distribution:

Dressing interprets the individual ensembles members as scenarios; an alternative approach is to draw analogues based on the statistics of the ensemble.



How can we best use these distribution in practice?
Historical target data is the key.

Seasonal Information

Development of a
European Multi-Model Ensemble System
for
Seasonal to Interannual Prediction

A subset of the models used in this study will *soon* be operational.

Multi-model ensemble system

- DEMETER system: 7 coupled global circulation models

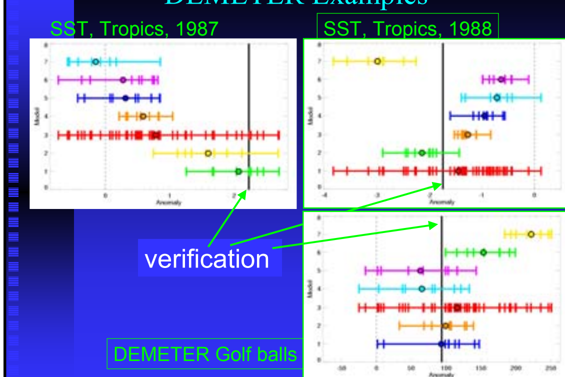
www.ecmwf.int

Partner	Atmosphere	Ocean
ECMWF	IFS	HOPE
LODYC	IFS	OPA 8.3
CNRM	ARPEGE	OPA 8.1
CERFACS	ARPEGE	OPA 8.3
INGV	ECHAM-4	OPA 8.2
MPI	ECHAM-5	MPI-OM1
UKMO	HadCM3	HadCM3

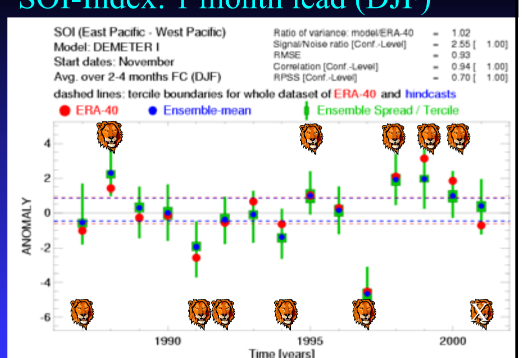
9 member ensembles
ERA-40 initial conditions
SST and wind perturbations
4 start dates per year
6 months hindcasts

- Hindcast production for: 1987-1999 (1958-2001)

DEMETER Examples

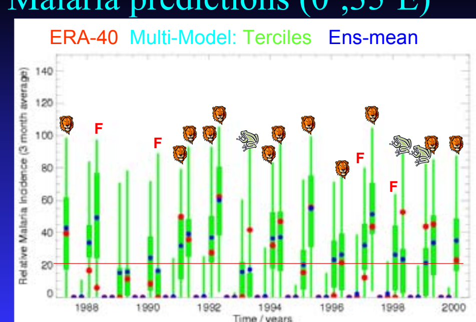


SOI-Index: 1 month lead (DJF)



The sign of every major anomaly (except the last) is clearly indicated

Malaria predictions (0º,35ºE)



We need empirical verification against malaria cases, rather than ERA-40. F

Conclusions

Ensemble forecasts contain a wealth of information.

Users are interested in exploiting this information, once we demonstrate that it provides (net) value.

The value of an EPS (or a hi-res model run adds to an existing ensemble) will depend not only on its cost and meteorological quality, but also on the size and accessibility of a forecast archive.

For numerate users, we should take care to offer only what we might deliver.

M Roulston *et al* (2003) Renewable Energy **28:** 585-602

M Roulston, J Ellepola and LA Smith (2004) *J of Coastal Engineering* (in review)

LA Smith & M. Altalo (2004) Public Utilities Monthly (in preparation)

LA Smith (2003) Predictability Past Predictability Present. ECMWF.

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

K Judd and LA Smith (2001) *Indistinguishable States I* , Physica D **151**: 125-151 *Indistinguishable States II*, Physica D ?2004?

A Weisheimer, L.A.S. and K Judd (2004) A New Look at DEMETER forecasts via Bounding Boxes (in review)

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

Using More than One Model

Introduction

- Ensembles can be used to account for uncertainty in the initial condition.
- Also, ensembles across separate models could be used to account for “model uncertainty” – *Different models are different even for the same initial conditions.*
- How can information from different models be integrated in order to make better decisions?

Temperature Forecasting Game

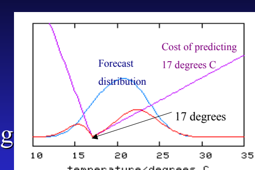
- Problem: Forecast temperature at Schleswig.

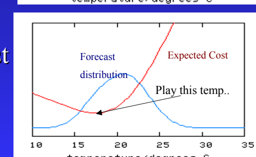
- 3 day lead time
- Two sources of information:
  - dressed ECMWF high resolution forecast
  - dressed ECMWF ensemble forecast
- Lose
  - 1 point for every degree underestimated
  - 5 points for every degree overestimated

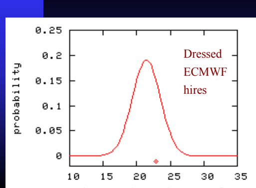
- Given a forecast distribution, how might this game be played?

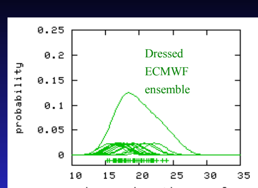
Decision Procedure 1: Expected Cost

- Given a probability forecast and a cost function, calculate expected cost of playing a given temperature. (say 17 degrees)
- Calculate expected cost for every play.
- Choose the smallest.









?

Combined forecast distribution

How can we combine two forecast distributions?

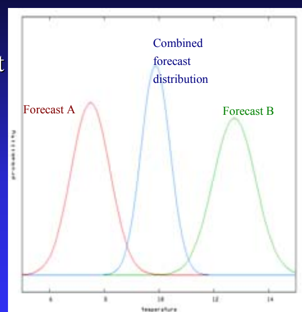
Especially if they “disagree”?

- Look at three methods of combining forecast distributions:
  1. Traditional Combination
  2. Compromise Combination
  3. Mixture
- Compare these with respect to the decision procedure.

Traditional Combination:

Assume each model provides independent information.

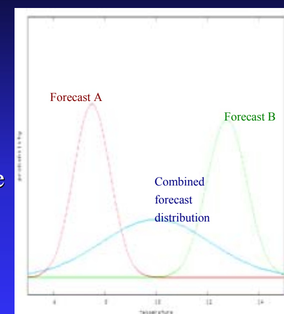
This can assign high probability where neither forecast did.



Compromise Combination:

Take disagreement as an indication of model failure.

This will increase the variance to reflect increased uncertainty.

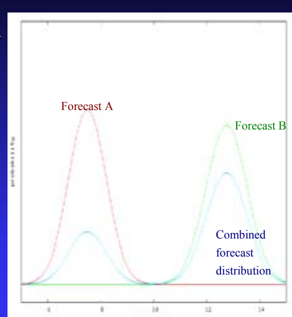


Mixture:

But, failure of one forecast doesn't imply failure of both.

Bimodal forecast reflects possibility that one or other *hasn't* failed.

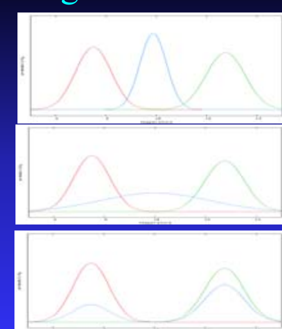
Can weight to reflect relative reliability.

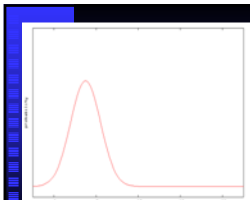


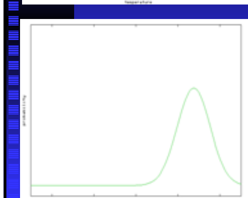
Back to the game:

Where would each combination play?

1. Traditional : ~9
2. Compromise : ~8
3. Mixture : ~6.5







- If Forecast A *didn't* fail:
1. Traditional: Loses ~10 pts
2. Compromise: Loses ~ 5 pts
3. Mixture: small loss

- If Forecast B *didn't* fail:
1. Traditional: Loses ~ 4 pts
2. Compromise: Loses ~5 pts
3. Mixture: Loses ~8 pts

Assuming that disagreement corresponds to failure of only one component forecast:

- The information needed to play well is contained in the mixture.
- That information is *destroyed* by the traditional or compromise combinations.

The Lesson

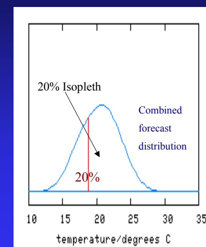
Like the three statisticians:

We must be careful about how we choose to combine information from several sources.

A poor choice can actually result in the destruction of vital information.

Decision Procedure 2: Play "Isopleths"

- Forecast temperature below which a certain percentage of probability falls – an "isopleth".
- Can choose isopleth which minimises cost on historical data.



Results

- Played game for year May 1998 to April 1999:

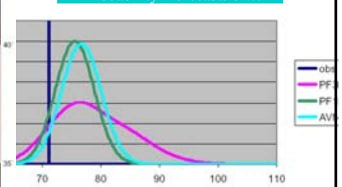
	Loss from Expected Cost	Loss from Isopleths
High Resolution	1076±68	1053±60
Ensemble	1003±49	967±38
Mixture	972±42	948±36

- Mixture is better than either source alone
- Isopleths beat Expected Utility

4 regions, 20 stations, one ultimate demand forecast: generation.

5 models, which target?

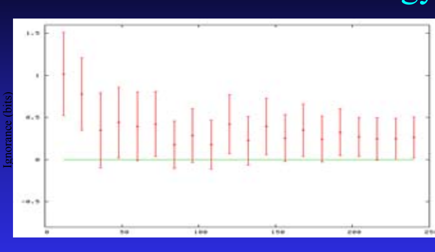
"Probability" forecasts of…



Probability of Precipitation at Schleswig

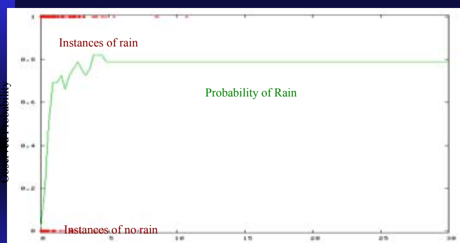
- Case in which relevant information exists but is not obvious.
- Simplest method for predicting probability:
  - Look at ECMWF 50 member ensemble
  - Estimate probability of rain as proportion of ensemble members in which it rains
- Can use "Ignorance" to evaluate performance against climatology.
  - Represents expected outcome of a series of bets.

Raw Ensembles vs Climatology



- Does worse even than climatology.
- Paradoxically, skill *decreases* for small lead times!

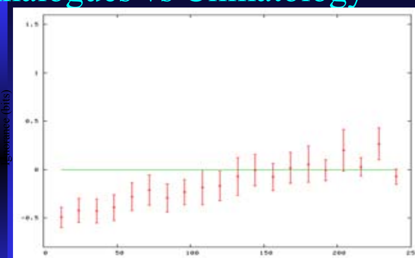
Method of Analogues



ECMWF median of ensemble rainfall prediction (mm)

- Reinterpret output of model. Given median ensemble rainfall, how often *did* it rain?

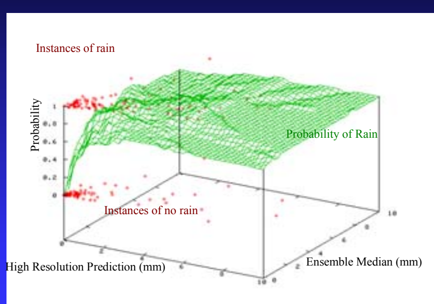
Analogues vs Climatology



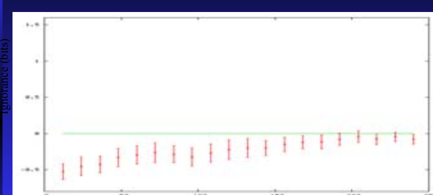
Lead time (Hours)

- Much improved: Skill to 5 days.
- Can do even better!

Idea: Combined Analogues



Combined Analogues vs Climatology



Lead time (Hours)

- Combining the high resolution and the 10%, 50% and 90% isopleths of the ensemble, we achieve skill even to 10 days.

Conclusions

- Information must be combined with care.
  - Lest relevant information is destroyed.
- Models are imperfect.
  - Minimising expected costs would be optimal otherwise, but isopleths win.
- Information can sometimes be found in surprising places.
  - A large verification archive is invaluable.

High Impact Weather Forecasting

High Impact Vs Severe

- High Impact weather need not be severe
- Normal weather can have major effects
- Users determine high impact events
  - rain, good for plants, bad for drying clothes outdoors

Snowfall January 2003

"Snowfall in January is hardly unexpected" – Alistair Darling.
Despite warnings, snowfall in January 200 caused widespread disruption on the UK road network.

- The snowfall was not severe or unexpected.
- The response was correct
- Failed to account for the response of motorists to the same high impact forecast
- Complex optimisation process

Hogmanay 2003

Late cancellation of festivities on New Year Eve due to high winds.

Organisers did not seek forecast information from the Met Office.

- Nothing could be done to save the festivities
- More faith in the forecast and organisers could have forewarned attendees
- Minimise disappointment and frustration

Rocket Launch

- On March 26th 1987 an Atlas Centaur rocket costing $181 million was destroyed shortly after lift off
- The rocket was hit by lightning 49 seconds after launch
- Ideally all available information is included in the decision making process

Extended Cold Spells & Construction

- The number of freezing days over a working week affects the construction industry
- Freezing conditions prevent tasks being performed, e.g. concrete setting
- Information about specific weather conditions can assist project management

Extended Cold Spells & Gas Delivery

- Natural gas used for domestic heating and electricity generation
- Over an extended cold spell demand increases
- Industry should maximise reserves to service possible change in priorities

What makes a High Impact forecast

- User defined
- Economic impact
- Need not be severe
- Mitigable
- Exploitable

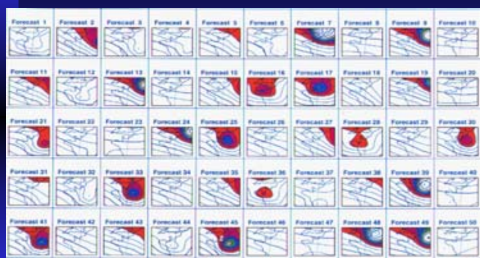
Timing and Duration

- Two distinct aspects of high impact weather forecasts
  - Timing, when an event will occur
  - Duration, how long will the event last for
- CATS aims to identify and extract this relevant information

Early warning of high impact weather events

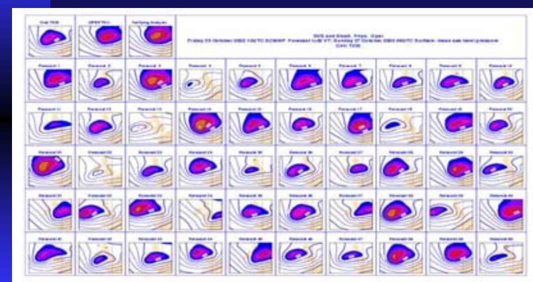
- Ensembles can highlight the possibility of future weather events
- Each ensemble member can be interpreted as a plausible scenario

December 1999



ECMWF Ensemble Forecast

October 2002



ECMWF Ensemble Forecast

Condensing Information

- We do not need all the information in the weather forecast – just the relevant information
- Cannot say definitively that a storm will happen
- But not surprised if it does

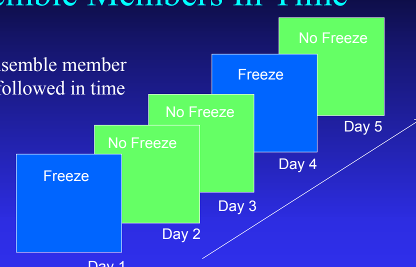
- Nothing special about storms (outside the Weather Channel)
- We can look for any particular weather in an ensemble
  - freezing
  - warm
  - cloudy

Duration Of Weather Events

- A sequence of weather forecasts can provide information on the duration or number of weather events
- Each ensemble and its evolution can be interpreted as a scenario over the forecast time

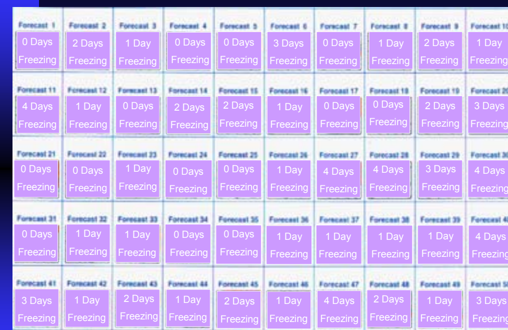
Ensemble Members In Time

Each ensemble member can be followed in time

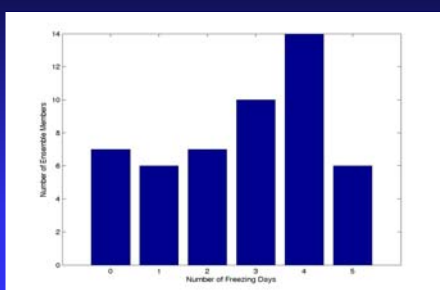


Relevant Information

- We can examine the evolution of each scenario and extract the relevant information



Risk of Freezing



Multiple Models

- We can take scenario information from different weather forecasters
- Can include forecasts that validate at the same time but are issued at different times
- Joint model distribution

Evaluating

- What high impact weather can we usefully forecast?
- DIME aims to determine this
- CATS can help design tailor made forecasts and show proof of value

Summary

- High impact not necessarily severe
- The performance of high impact forecasts can be evaluated
- Information content (economic value) not always initially obvious

More Complex Decision Making

Introduction

Sometimes a series of decisions must be made over a relatively extended period of time. Both the direct consequences and the effect of current decisions on future decisions are important.

An interesting aspect of many of these problems is that, as time progresses, more information in the form of later forecasts can become available. The value of these is strongly dependent on the decision process.

Weather dependent examples:

- When to spin up/spin down generating units.
  - Demand is weather dependent.
  - Units can take significant time to bring online. Flexibility can be impaired by past decisions.
- Routing in transportation.
  - Desire to avoid bad weather.
  - Position is the total of the series of decisions “where next?”. Where next depends on the current position.
- Others?

Ship Routing

Simple ship routing problem:

Hoffschildt, M., J-R. Bidlot, B. Hansen and P.A.E.M. Janssen, 1999: Potential benefits of ensemble forecasts of ship routing. ECMWF Research Department Tech. Memo. No. 287. 25pp

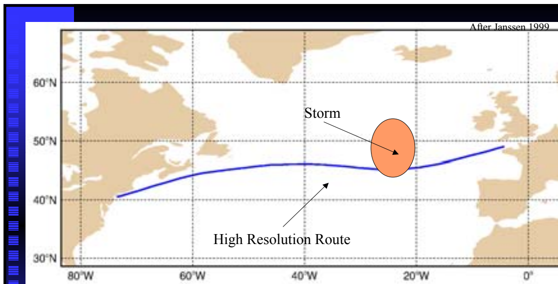
Janssen, P. Potential benefits of ensemble prediction of waves. in ECMWF Newsletter No. 86 – Winter 1999/2000.

Ship model relates cost of running to weather parameters such as wind, waves and swell, taking into account fuel and grease consumption, and damage due to storms…

Perfect Information Case

Given a trustworthy forecast for the weather. It is (at least conceptually) a straightforward matter to find the optimal route.

On the next slide, we show such an optimal route calculated for a crossing leaving Brest on 28 February 1999 and arriving at New York on 7 March 1999. The ECMWF high resolution forecast was used in this case.



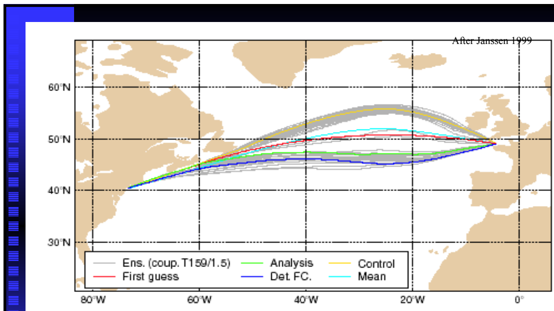
In the absence of inclement weather the optimal route would be near a great circle. Here, though, a storm is forecast and the route bypasses it.

Imperfect Information Case

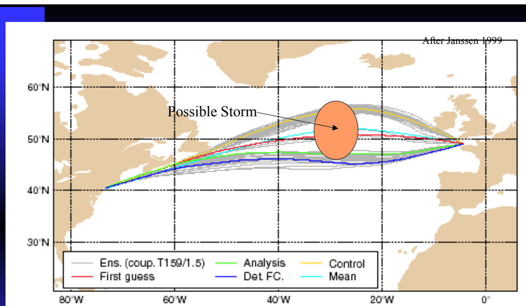
In practice, of course, there is no completely trustworthy forecast. It is prudent to consider the information in an ensemble or forecasts.

An ensemble of routes can be generated by computing an optimal route for each forecast.

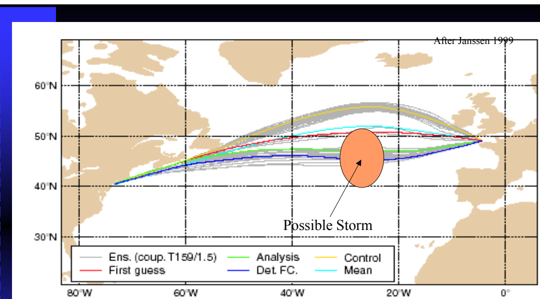
Next we show such an ensemble of route. Also shown are the mean route, the "great circle" route and the optimal route.



Note the "bifurcation" of the ensemble of routes. This can be explained by uncertainty as to the location of the storm.



When the storm forms more northward, the better routes are those to the south.



Similarly, if the storm forms to the south. The few ensemble routes passing through the middle correspond to simulations where the storm never formed.

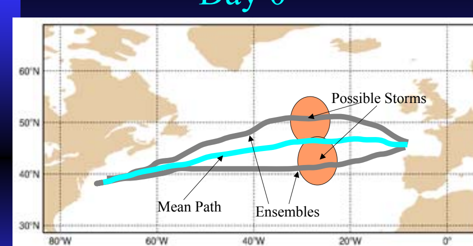
Updating the Route

If most of the ensemble routes clustered, it would be sensible to follow one of them.

Things are less clear if the routes bifurcate.

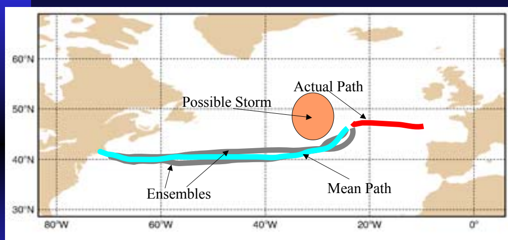
One solution is to follow a compromise route (say the mean) and re-compute the route as new forecasts become available.

Day 0



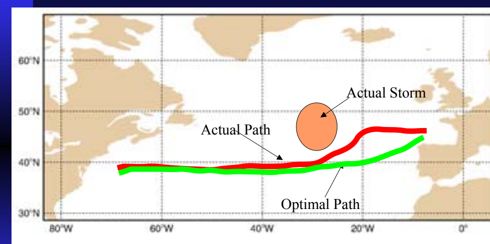
Before setting out, storms are forecast, but position is uncertain. Begin by following the mean path.

Day 1



After a day, the position of the storm has become more certain. Now all ensemble routes pass to the south.

Eventual Route



Actual path lies not too far from the best path possible.

Following the mean ensemble route, but being willing to update the route as newer forecasts become available, can result in good routes.

Saetra has noted that even if the mean ensemble route is followed from the outset with no alterations other than varying speed, then the path taken is, on average, no more than about %0.5 more expensive than optimal.

A possible danger

Note that choosing decision paths by optimising for the conditions in individual ensemble forecasts can lead to riskier paths being over valued:

	Path A	Path B	Path C
Forecast 1	-10	-20	-12
Forecast 2	-20	-10	-12

- Well known weakness in bridge playing programs.
- Would be nice to have some industrial examples.

Conclusions

- Realistic decision processes can be complex.
- Forecast valuation under simple models can be misleading.
- More documented industrial examples would be nice. (Please help!)

Evaluating Weather Forecasts - The DIME Project

Overview – Objectives of DIME

- Evaluating forecast distributions
- Combining forecast distributions
- Disseminating results
- The DIME site

What you can find on the DIME website

- Further perspectives

Why DIME?

- Ensemble weather forecasts appear to be invaluable to weather dependent business
- There remain open questions in handling and valuing ensemble forecasts

DIME aims to be a one-stop weather forecast information site

Point Forecasts vs Ensemble Forecasts

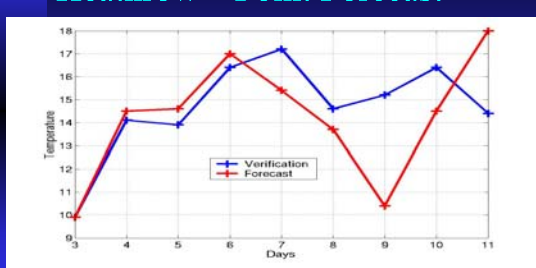
- A point forecast is a single number:

London temperature tomorrow: 25°

- An *ensemble* forecast is a *set* of numbers:

London temperatures tomorrow: 23°, 26°, 27°, 29°, 31°

Temperature at London Heathrow – Point Forecast



Evaluating Point Forecasts

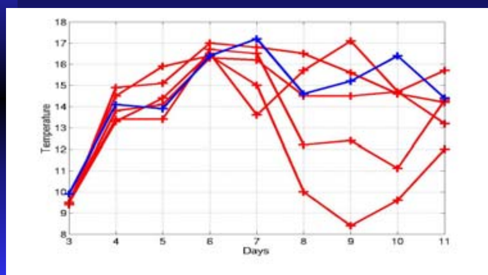
- For a classical point forecasts we can evaluate the error:

London temperature tomorrow: 25°
Reality turns out to be: 28°

Error: 3°

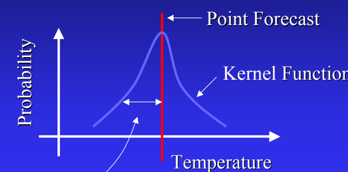
- A point forecast is considered good if the error is small on average

Temperature at London Heathrow – Ensemble Forecast



Enhancing Point Forecasts – Method of *Dressing*

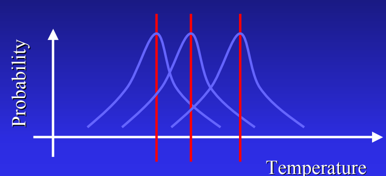
- *Dressing* point forecasts adds uncertainty information



Kernel width needs adjustment –
Many possible ways to do that!

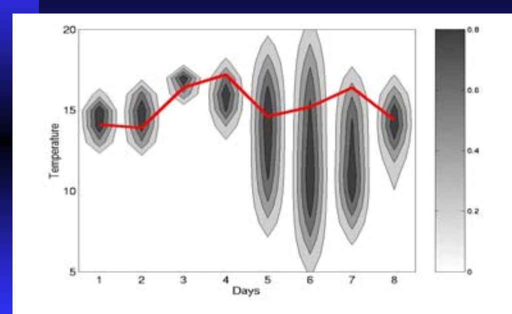
Dressing Ensemble Forecasts

- *Dressing* ensemble forecasts



How do we combine the probability densities?

Dressed Ensemble Forecast



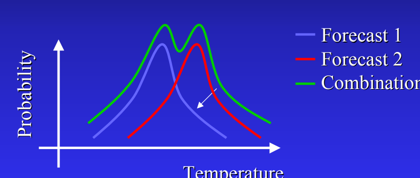
Combining Forecast Distributions

There exist *many different* methods to combine forecast distributions

- Take into account that the different forecasts have different skills
- Equally likely forecasts should be treated equal
- Determine optimal combination by looking at the skill, e.g. Ignorance

Combining Forecast Distributions

- Combining forecasts (recall D. Kilminster)



Evaluating Forecast Distributions

To quantify the potential usefulness of different forecast distributions we need to evaluate their skill

How do we compare reality (a single number) with the forecast (a distribution)?

Different Problems Require Different Skill Scores!

- There exist *many different* methods to measure skill
- DIME aims to investigate the skill of NWP's and dressing techniques using various skill scores
- DIME disseminates background information

What is the best dressing method for your forecasts?

Example of a DIME Result - Assessment of Skill

- Evaluating various skill scores of operational NWP's over a period of time

MODELS	Ignorance Skill Score	Brier Skill Score	…
MODEL 1	4.5	3.2	…

- What do those skill scores mean?

Ignorance And Weather Roulette

- Bet on the outcome of tomorrow's weather

Tomorrow's Temperature

Temperature 4 5 6 7 8 9 10 11 12

Spread wealth for safety

Restrict spread to make money

When Is a Forecast Distribution Good in Weather Roulette?

- A good forecast distribution balances between *spread* and *accuracy*
- Criterion is:

A forecast distribution is the better the more money it yields in weather roulette

- The *ignorance* reflects the expected rate of wealth grow.

Skill Scores for Binary Event Forecasts – The *Brier* Score

- Binary *(yes/no)* events are e.g.: Will it freeze? Will precipitation exceed a threshold? etc.
- A forecast for a binary event:

Forecast:
It freezes with probability p
It doesn't with probability (1-p)

- *Brier* score reflects the quality of binary event forecasts

Comparing Forecasts

- Two different forecasts can be compared by means of their ignorance

NWP schemes	Ignorance Skill Score	Brier Skill Score	…
MODEL 1	4.5	3.2	…
MODEL 2	3.2	3.9	…

Comparing Forecasts

Dressing allows to compare point forecasts and ensemble forecasts

MODELS	Ignorance Skill Score	Brier Skill Score	…
MODEL 1 (Ens.)	4.5	3.2	…
MODEL 2 (Ens.)	3.2	3.9	…
MODEL 3 (Point Forecast)	3.0	1.9	…

Combining Forecasts

- The table grows again…

MODELS	Ignorance Skill Score	Brier Skill Score	…
MODEL 1 (Ens.)	4.5	3.2	…
MODEL 2 (Ens.)	3.2	3.9	…
MODEL 3 (Point Forecast)	3.0	1.9	…
MODEL 1 and MODEL 2	4.8	3.4	…

DIME Products

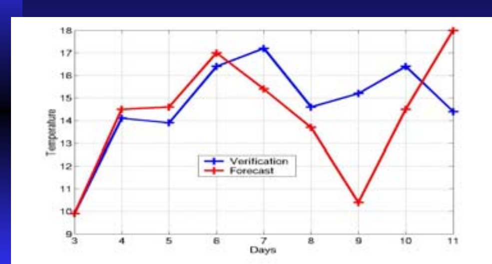
- Objectively compare operational NWP model ensemble forecasts by their skill
- Suggest schemes to combine ensemble forecasts and evaluate the skill of these schemes
- Provide actual weather forecasts for specific locations using a selection of our methods

Dissemination of Products

- The medium to disseminate DIME products is the internet
- The web allows users to interactively request the products they are most interested in

www.dime.lse.ac.uk

What is wrong with this graph?



the London School of Economics and Political Science

Smith institute for industrial mathematics and system engineering

Operational Approaches to Managing Weather Risk: From Hours to Decades

22 June 2004
London School of Economics

The Centre for the Analysis of Time Series (CATS)

The Centre for the Analysis of Time Series (CATS) at the London School of Economics brings together a unique mix of internationally recognised expertise both in deterministic non-linear modelling and stochastic non-linear modelling, as well as a powerhouse of statisticians expert in traditional statistical methods and a number of physical scientists with experience in time series analysis. CATS was established in 2000 and is based within the Department of Statistics at LSE. The School has a long and distinguished history in time series analysis and as part of its strategic plan has invested in developing a world-class centre of excellence in this area.

The Centre for the Analysis of Time Series aims to:

- Address the question of data analysis using both physical insight and the latest statistical methods.
- Focus on non-linear analysis in situation of economic and physical interest, such as weather forecasting.
- Promote awareness of limitations of non-linear analysis and the danger of blindly transferring well-known physics to simulation modelling.
- Focus on end-to-end forecasting, taking account of current uncertainty about the state of the system, model inadequacy and finite computational power.

Suggestions for new areas of interest are always welcome. We interpret analysis rather broadly to include estimation of statistics, prediction and the analysis of forecast systems. We are interested in the development of tools to interpret, value and apply probabilistic forecasts.

Contact details:

Centre for the Analysis of Time Series
Department of Statistics, B717
Columbia House
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Tel: +44 (0) 20 7955 6457
Fax: +44 (0) 20 7955 6273
E-mail: littertray@lse.ac.uk
Website: http://www.lse.ac.uk/collections/cats/

Faraday Partnership

Operational Approaches to Managing Weather Risk: From Hours to Decades

22 June 2004
London School of Economics

The Smith Institute for Industrial Mathematics and System Engineering

The Smith Institute delivers solutions and technical services to companies, through the application of mathematical modelling and analysis. In a knowledge-driven economy, these skills provide cost-effective solutions to operational or design problems, and are also important to the formulation of industrial strategy. The Institute's staff has wide expertise in modelling, data analysis, project management and research coordination. Mathematics is a uniquely transferable discipline, and rapid competitive advantage is often provided by the exploitation of techniques that have found established applications in other sectors of the economy. The Institute is therefore able to provide cost-effective solutions across all sectors. Its technical staff have extensive experience in the energy, telecommunications, food, paper and aerospace industries.

The Smith Institute also manages the Faraday Partnership for Industrial Mathematics, which it launched in 2000 with assistance from the Department of Trade and Industry and the Engineering and Physical Sciences Research Council. Faraday Partnerships promote industrial competitiveness, through improved collaboration between industry and the science base for the purposes of research, development and technology transfer. The Faraday Partnership for Industrial Mathematics currently supports approximately 30 Faraday Associates, who are talented young researchers, engaged on industrial research challenges in some of the strongest university research groups in the UK. The Smith Institute is able to draw on the specialist skills of these research groups as necessary in support of all its activities, putting it in a uniquely strong position to provide companies with cutting-edge solutions.

Our services include: In-house study groups, Knowledge Transfer Partnership supervision, technology translation, scoping studies, collaborative research project management, seminars & training, consultancy.

If you wish to learn more about the variety of mechanisms that we can offer, please contact:

Gillian Hoyle

Smith Institute
Surrey Technology Centre
Guildford
GU2 7YG

Telephone: 01483-579108
Fax: 01483-568710
Email: enquiries@smithinst.co.uk
Website: http://www.smithinst.ac.uk/

Centre for the Analysis of Time Series

Department of Statistics, B717
Columbia House
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Tel: +44 (0) 20 7955 6457
Fax: +44 (0) 20 7955 6273
E-mail: littertray@lse.ac.uk
Web site: http://www.lse.ac.uk/collections/cats/

Smith *institute*
for industrial mathematics and system engineering

Surrey Technology Centre
Guildford
GU2 7YG

Telephone: 01483-579108
Fax: 01483-568710
Email: enquiries@smithinst.co.uk
Web site: http://www.smithinst.ac.uk/