Indistinguishable States and Data Assimilation:

The roles of observational uncertainty and model inadequacy are contrasted in the case of physical simulation models when our best forecast models are chaotic (deterministic models with exponential-on-average sensitivity to initial condition). In the perfect model scenario, the framework of indistinguishable states provides an (?effectively Bayesian?) algorithm for constructing accountable (ideal) probabilistic forecasts. Within the perfect model scenario, the model-class in-hand will admit a model trajectory which shadows the observations: specifically, a trajectory which is consistent with the observations given the observational noise model. Outside the perfect model scenario, it can be proven that the set of indistinguishable states is empty, suggesting that no algorithm exists which can provide accountable probability forecasts. Practical implications differ for weather-like forecast applications and climate modelling. Adaptive observations are considered in this context, and it is noted that state-estimation might be profitably distinguished from forecast initialisation. Open questions of data assimilation in climate modelling are also touched upon.

Local Context

Interesting/Useful problems for applied maths:

Full Solution	Approximation	Implementation in R^m
Probabilistic Updating	KF	filter.c
Growth of uncertainty	Lyapunov Vector	Breeding Vector

From yesterday:

"One has to get an initial condition" My goal is an ensemble, no particular initial condition is of particular interest.

"Reducing Uncertainty" Or maintaining realistic uncertainty/confidence?

"Works only in low dimensional systems":

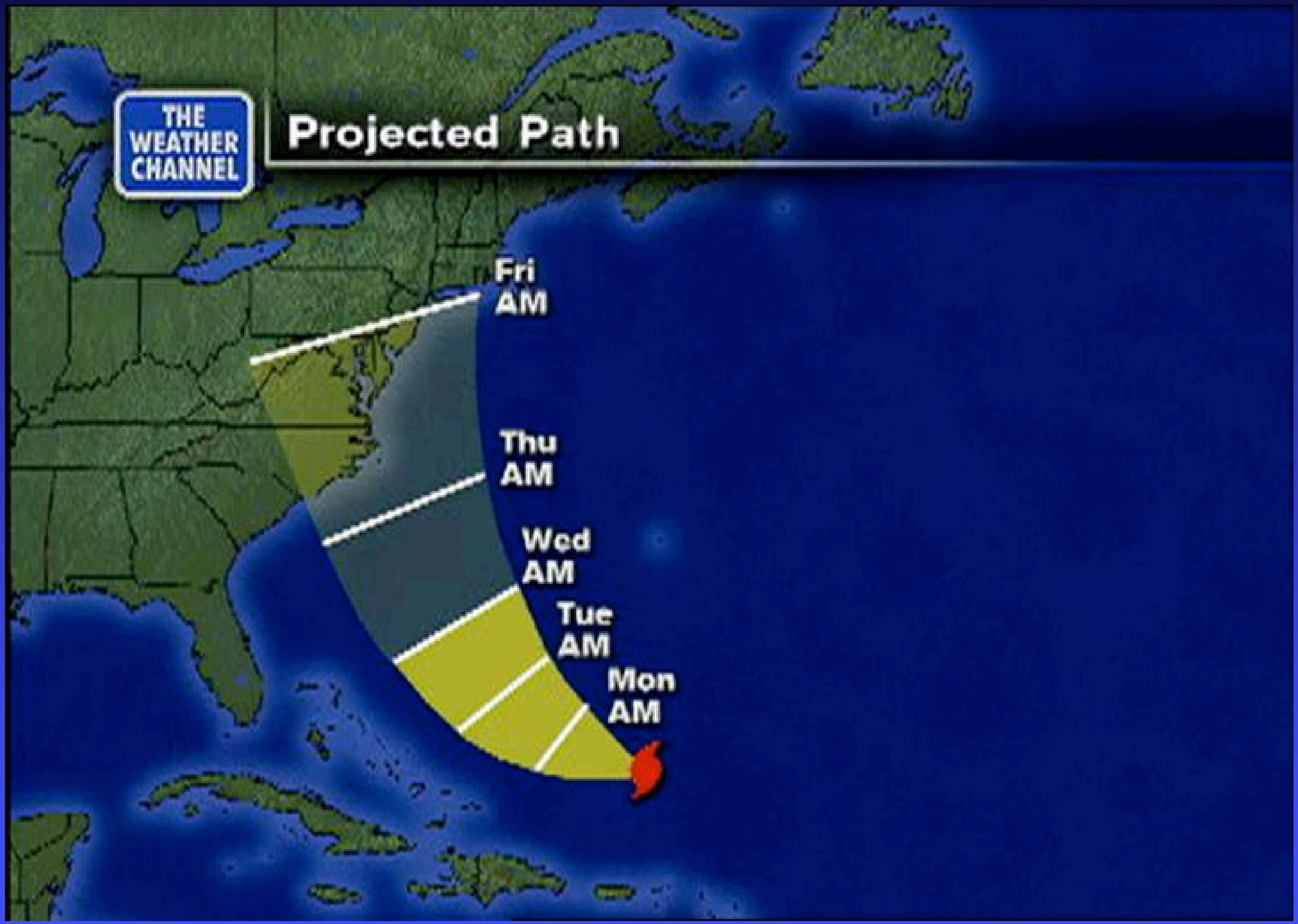
Solutions illustrated in low dimension systems often fail to generalize to hi-D, but difficulties identified in low dimensions systems often fail to magically vanish in hi-D.

Indistinguishable States and Data Assimilation:

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

Mark Roulston, Devin Kilminster, Kevin Judd,
Liam Clarke, Jochen Broecker
lsecats.org

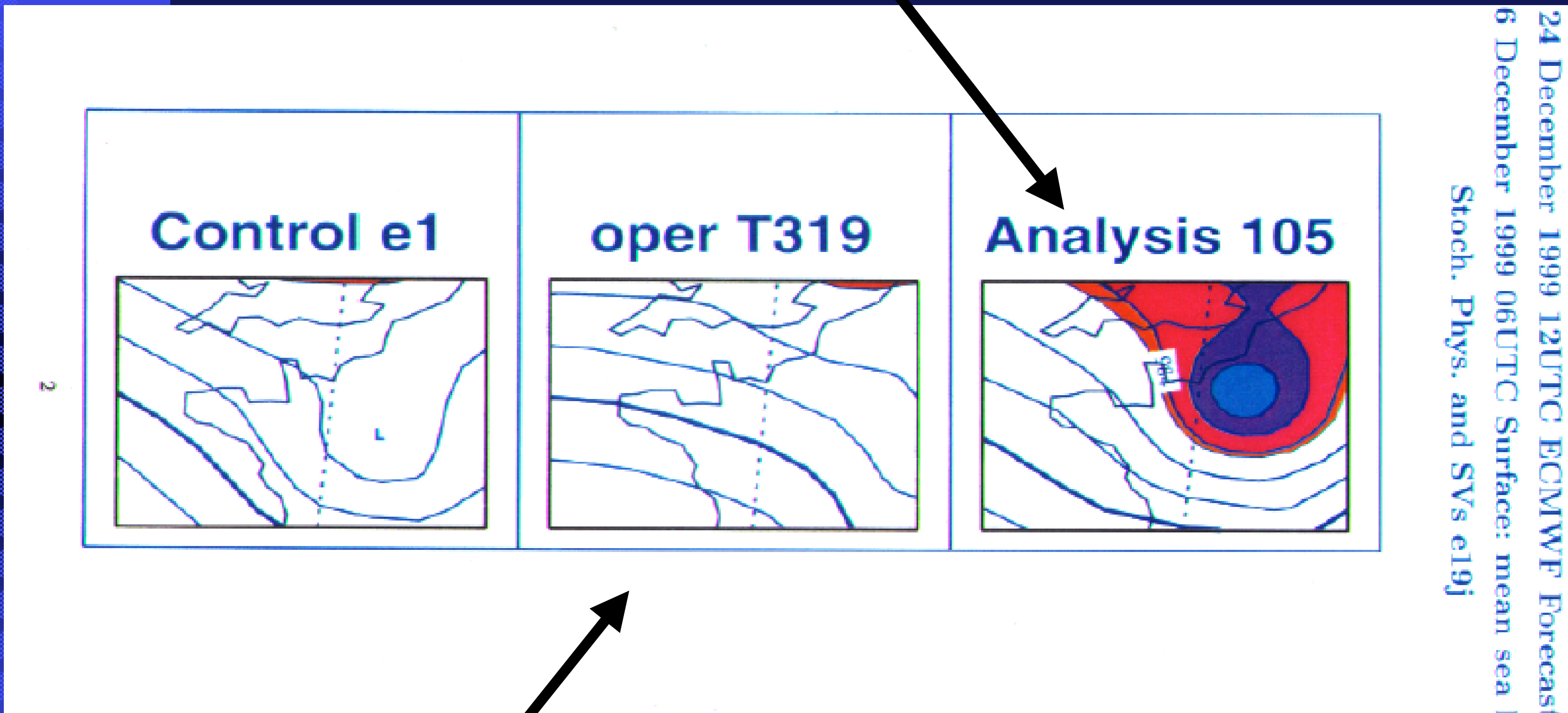
Tropical Storm/Hurricane "Cone of Uncertainty"



The Presentation of Uncertainty (The Weather Channel):

. How should I interpret this distribution?

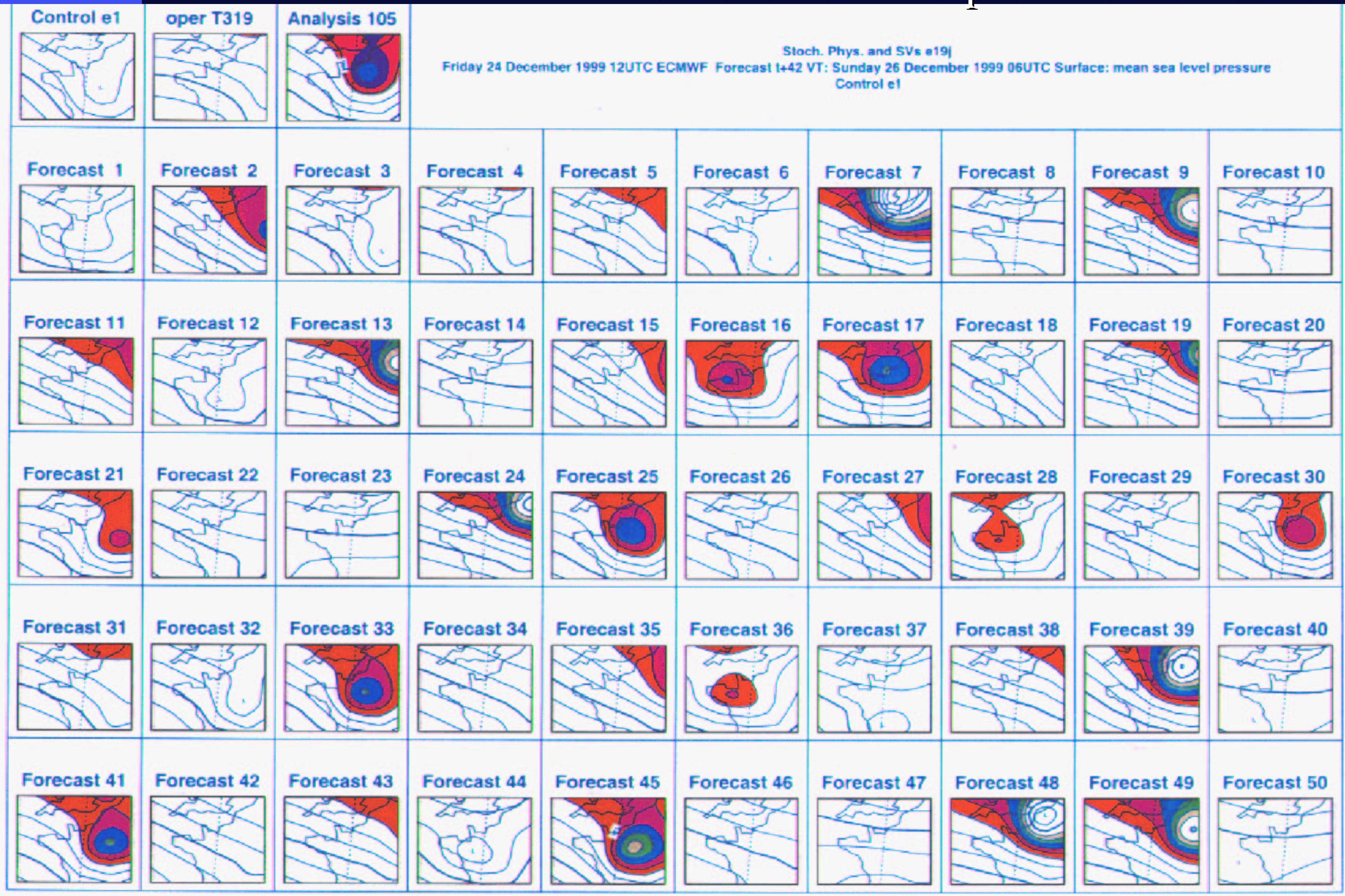
Two 2-day forecast and obs: December 26 1999



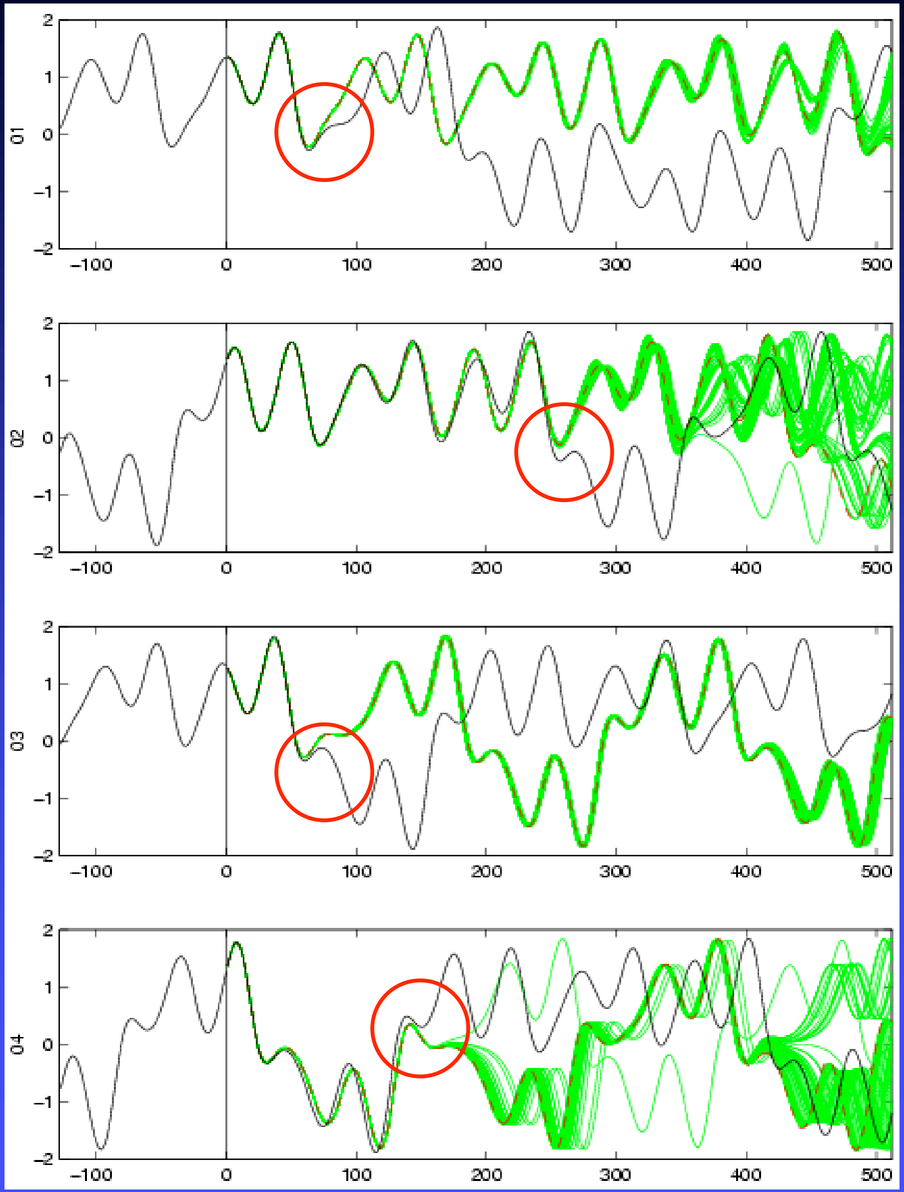
The single forecast from the "best" model

Simulations like these are of great value, even if not a PDF.

How should I interpret this distribution?



Forecasts of a "Chaotic" Circuit



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

How should I interpret these distributions?

Does the model admit a trajectory which is consistent with the targets provided by the obs and the noise model?
(a "shadow")

(And what is noise, really?)

The point of this slide is to stress that there is nothing "wrong" with weather forecasting: these facts hold for dynamic forecasting of *all* physical systems!

Two model states (say, x and y) are *indistinguishable states* (IS) if likely observations of the historical trajectory of x might well have come from y.

A model trajectory *i-shadows* the observations if that trajectory might well have generated the observations, given the noise model.

The distinction between shadowing and IS is that *shadowing* relates a model trajectory to a set of observations (numbers: s_i), while being IS is a relation between two model trajectories given a noise model.

Traditional aims of state estimation:

$$P(x(t_0)\ |\ s_i, F_a(x), a, n)$$

$x(t_0)$	current model state
s_i	observations
$F_a(x)$	dynamical model
a	parameter values
n	obs noise model

Traditional aim of forecasting (in statistics)

$$P(x(t > t_0)\ |\ s_i, F_a(x), a, n)$$

In cases where $F_a(x)$ is imperfect (*i.e.* in practice), these two procedures may have different target different distributions for $P(x(t_0))$.

You will have understood the main point of this talk if you leave it unsure of the target in the second case.

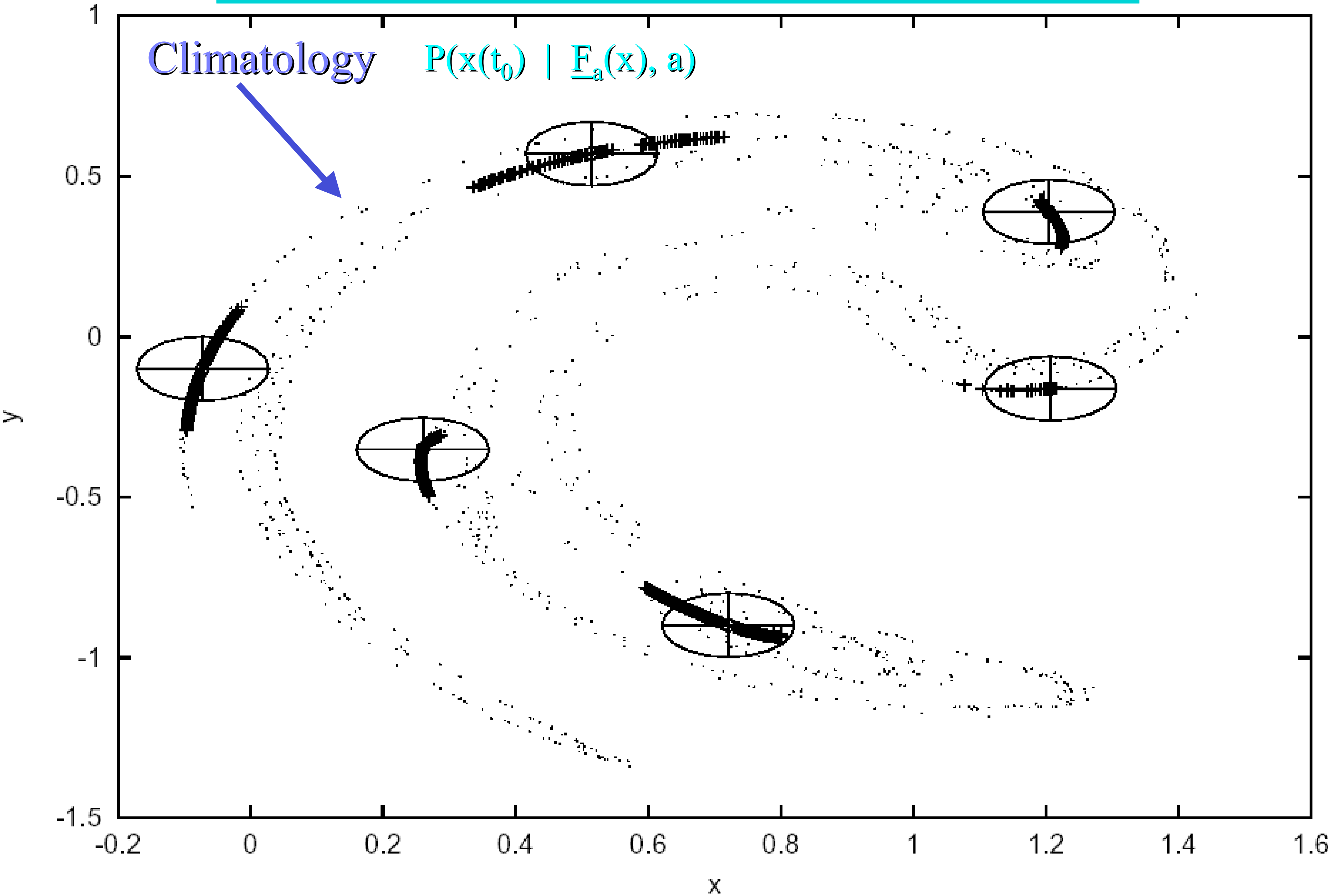
Ikeda System

$$\tilde{x}_{i+1} = 1 + \mu(\tilde{x}_i \cos\theta - \tilde{y}_i \sin\theta)$$

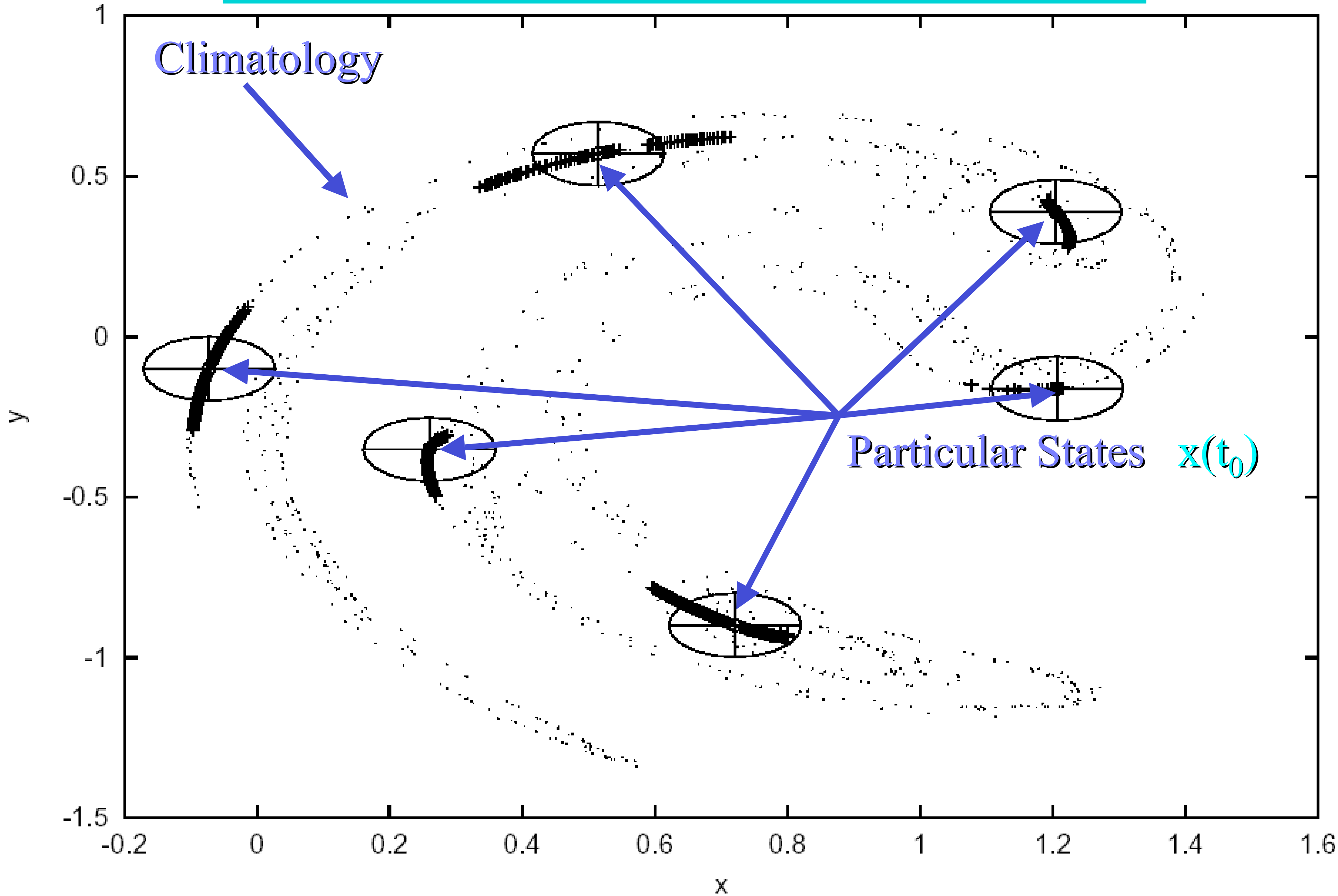
$$\tilde{y}_{i+1} = \mu(\tilde{x}_i \sin\theta - \tilde{y}_i \cos\theta)$$

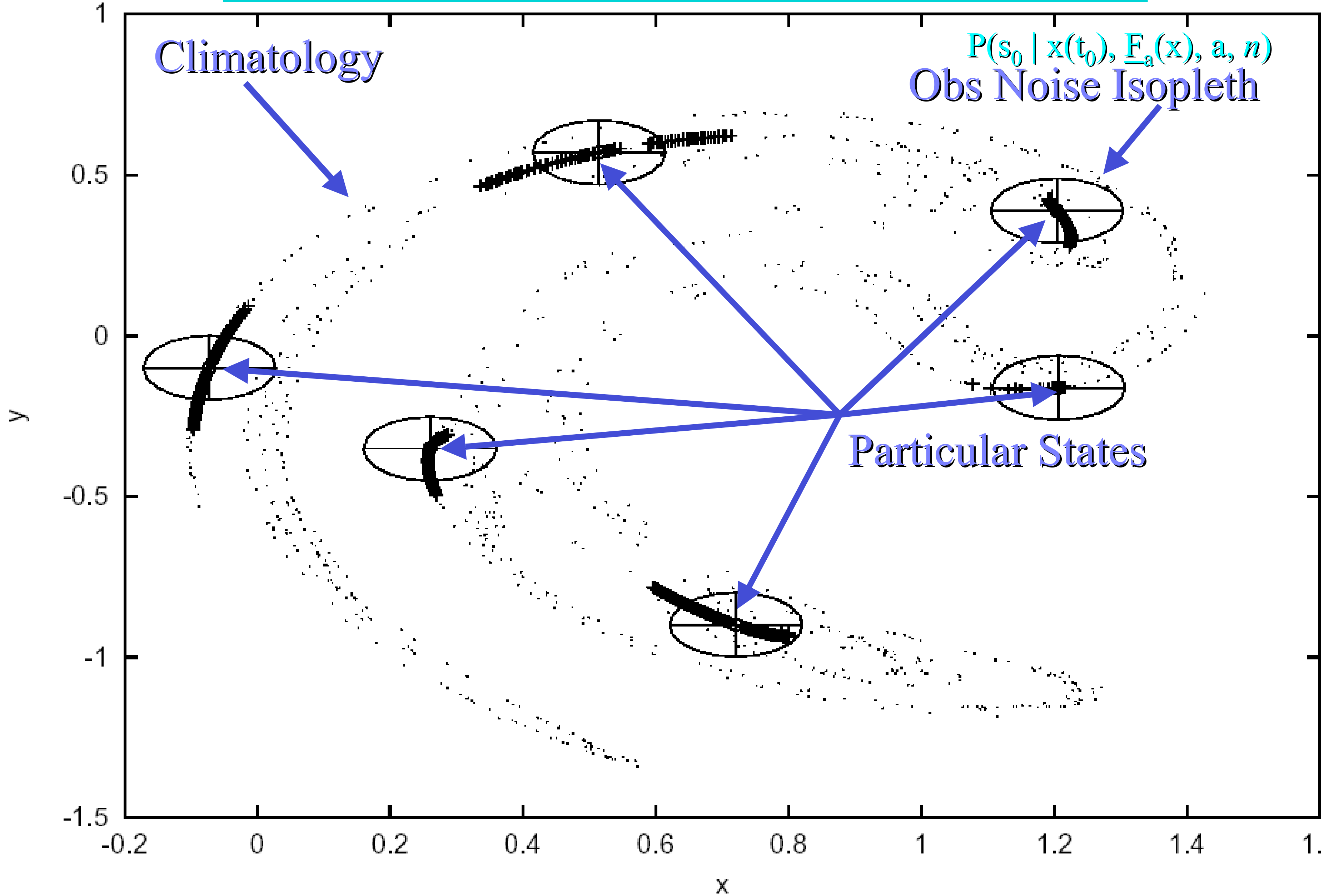
$$\theta = a - b/(1 + \tilde{x}_i^2 + \tilde{y}_i^2)$$

$$a = 0.4, b = 0.6, \mu = 0.83$$

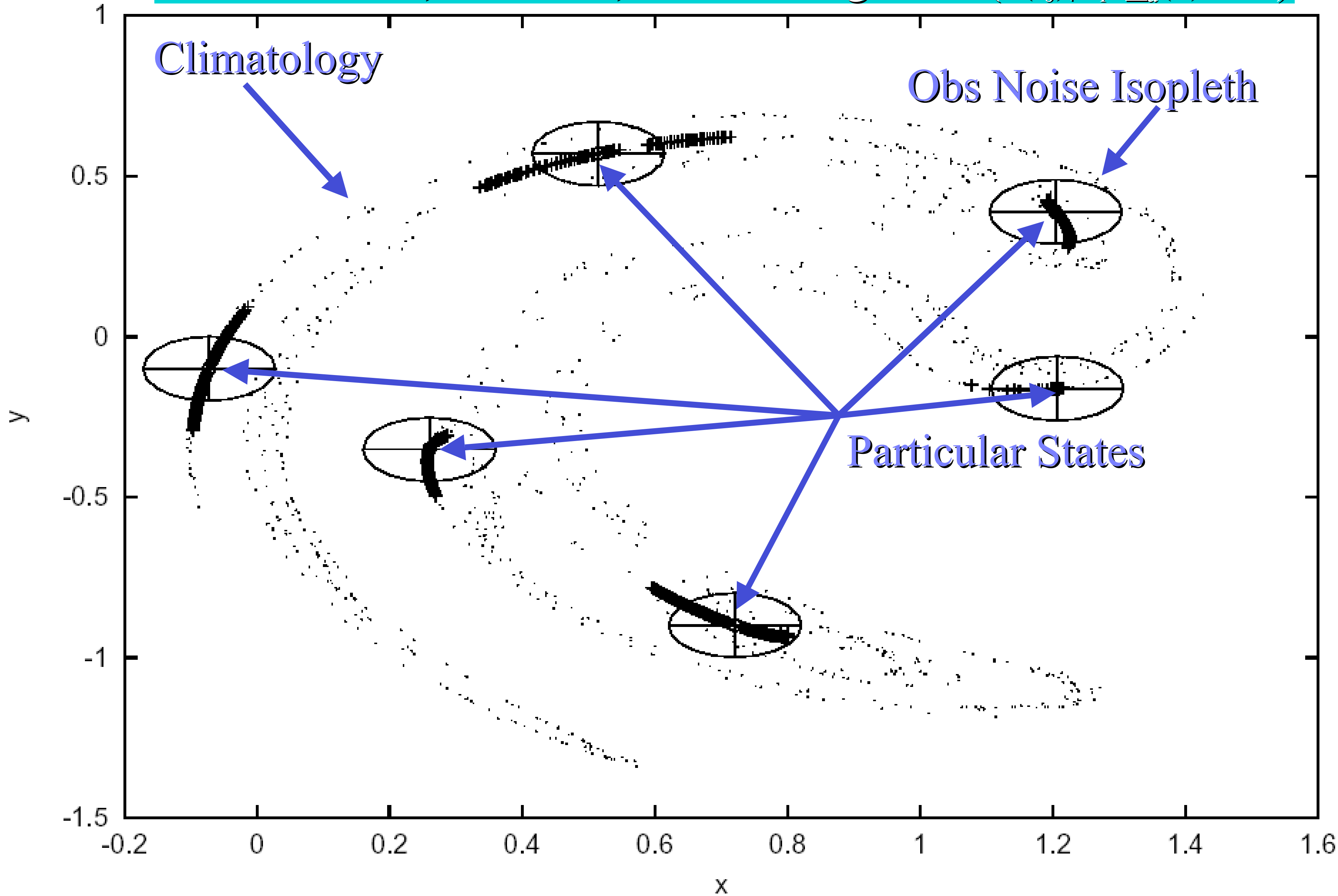


Climatology
P(x(t0) | F_a(x), a)
1
0.5
0
-0.5
-1
-1.5
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
x
y


Climatology
Particular States $x(t_0)$
1
0.5
0
-0.5
-1
-1.5
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
x
y


Climatology
$P(s_0 \mid x(t_0), \underline{F}_a(x), a, n)$
Obs Noise Isopleth
Particular States
1
0.5
0
-0.5
-1
-1.5
y
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
x

What we want, of course, is something like $P(x(t_0)|\, s_i, \underline{F}_a(x), a, n)$



Jargon Normalization

Climate Modelling: How will the invariant measure of the system respond to a change in parameter value?

2xCO_2 yields what change in the global mean temperature

Weather Forecasting: The future PDF of the atmosphere.

What are the chances of "rain" next Thursday?

Parameter: some physics based (usually uncertain) "constant".

boiling point of water at 1 bar (~100); speed of light (=1).

State Variable: potential observable expected to evolve

Temperature in the Joe's office.

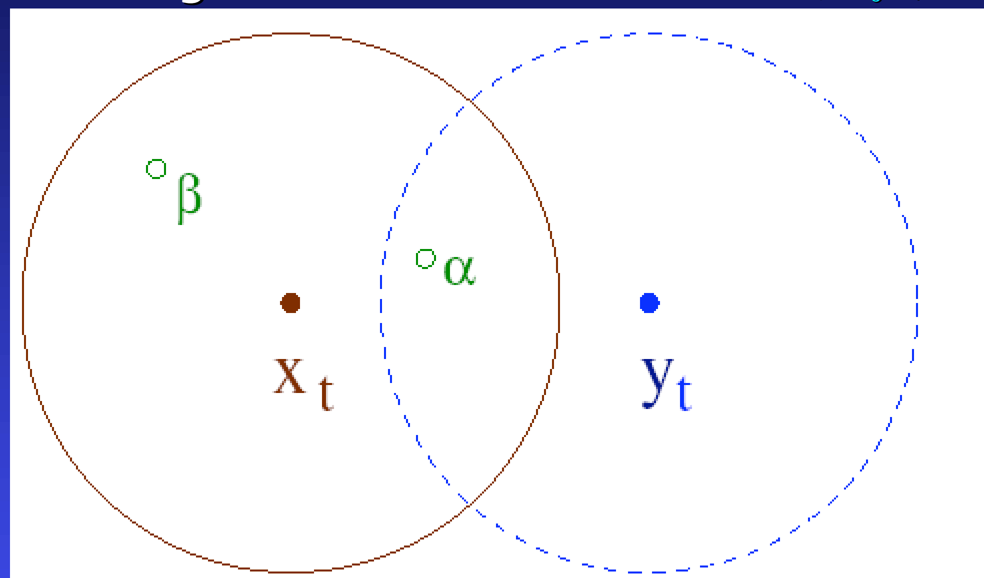
Invariant Measure :: Climatological Distribution :: "Attractor"

Shadow: Relation between a model trajectory and a set of obs.

Indistinguishable States: Relate Model trajectories to each other.

Indistinguishable States

Given a noise model, what is the probability that two states x and y are indistinguishable given one observation? $P(y(t_0) \mid x(t_0), \underline{F}_a(x), a, n)$



NOT: $P(y(t_0) \mid s_i, x(t_0), \underline{F}_a(x), a, n)$

If the observational noise is bounded, a single observation can distinguish x & y.

Given a series of noisy observations of the trajectory of x, what is The probability that the observations came from the trajectory of y?

In the perfect model scenario (PMS) if the true trajectory of the system is x_t, $t = 0, -1, -2, \ldots$, then the final state x_0 is distinguishable with probability one from another state y_0, which is the final state of a trajectory y_t, if $Q_\rho(y_0|x_0) = 0$, where

$$Q_\rho(y_0|x_0) = \prod_{t \leq 0} q_\rho(y_t - x_t),$$

$$q_\rho(b) = g_\rho(b)/g_\rho(0), \tag{1}$$

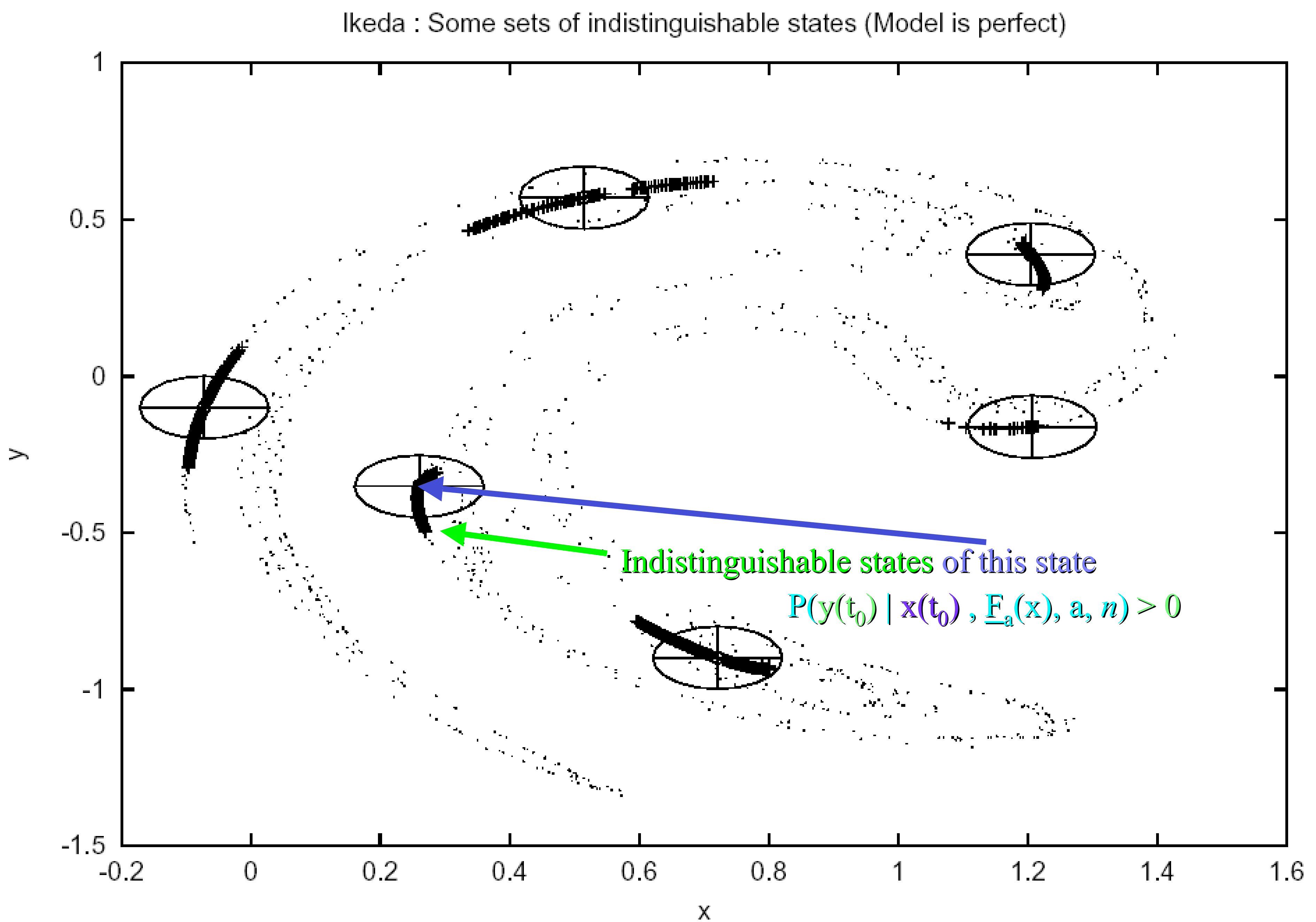
$$g_\rho(b) = \int \rho(z)\rho(z-b)\,dz,$$

and ρ is the probability density of the additive observation error.

Let H(x) be the set of states for which Q > 0.
Within PMS:

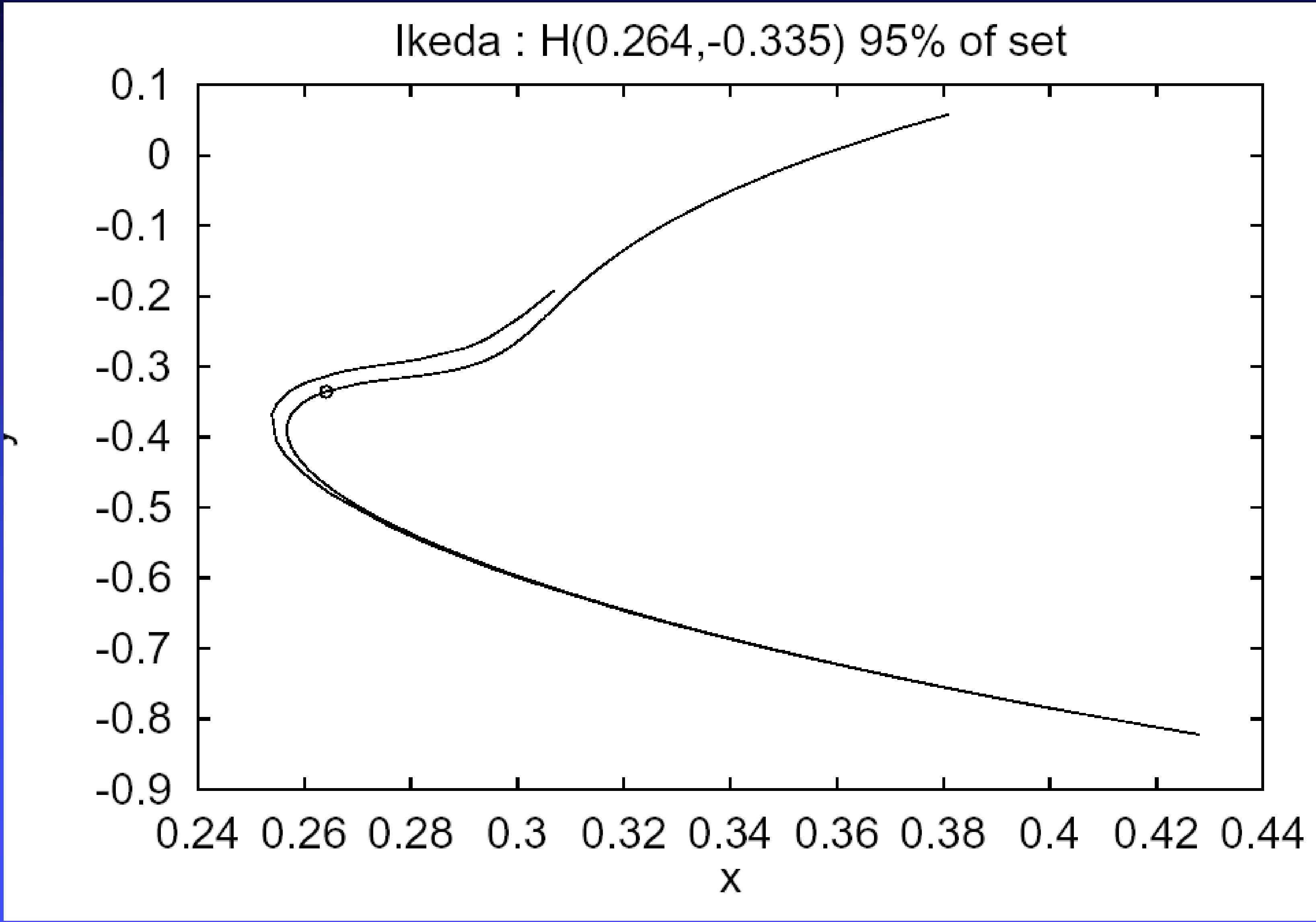
1) H(x) always contains x
2) H(x) also contains a (weighted) segment of the unstable set of x

Judd and Smith (2001) Physica D

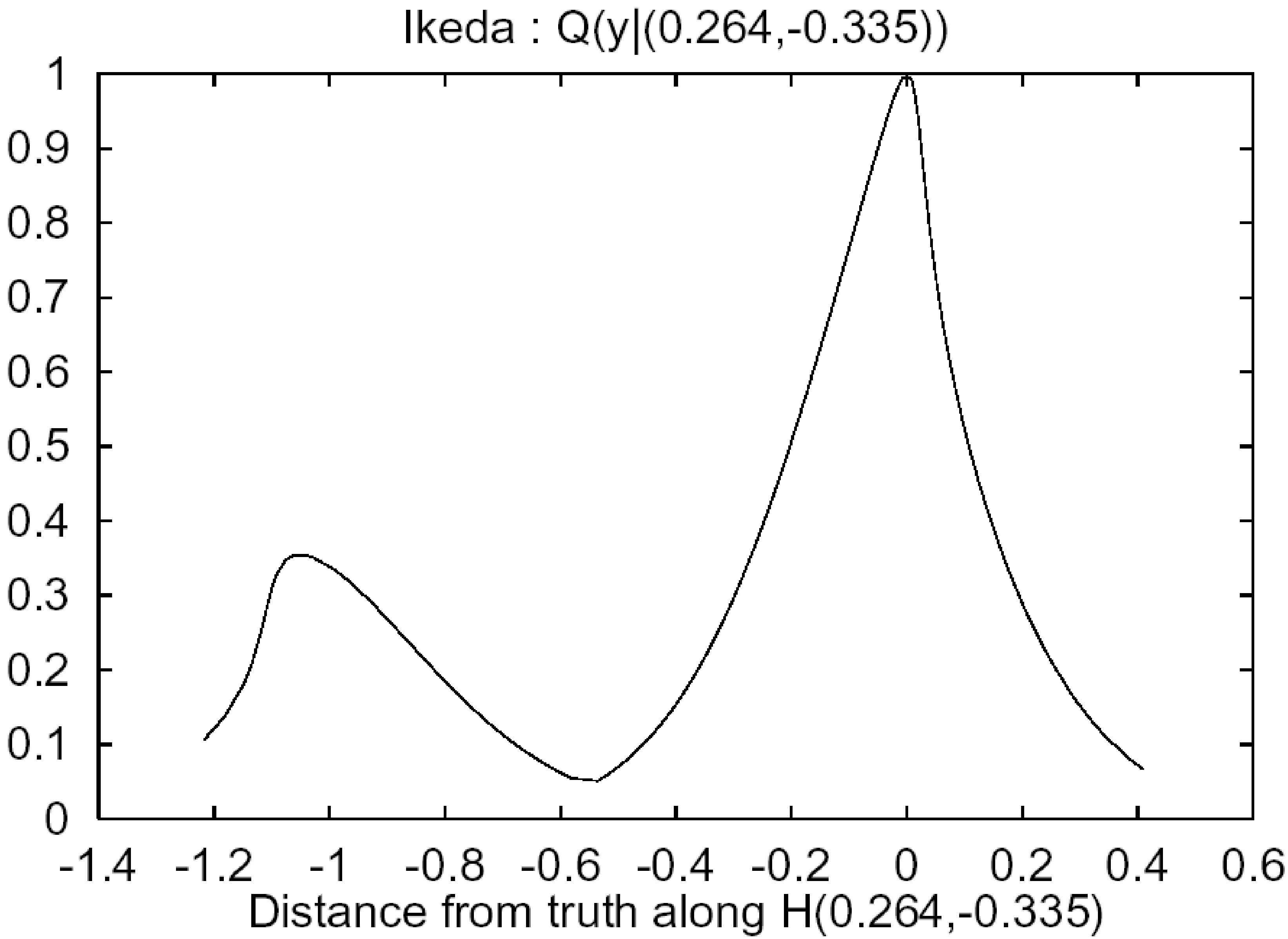


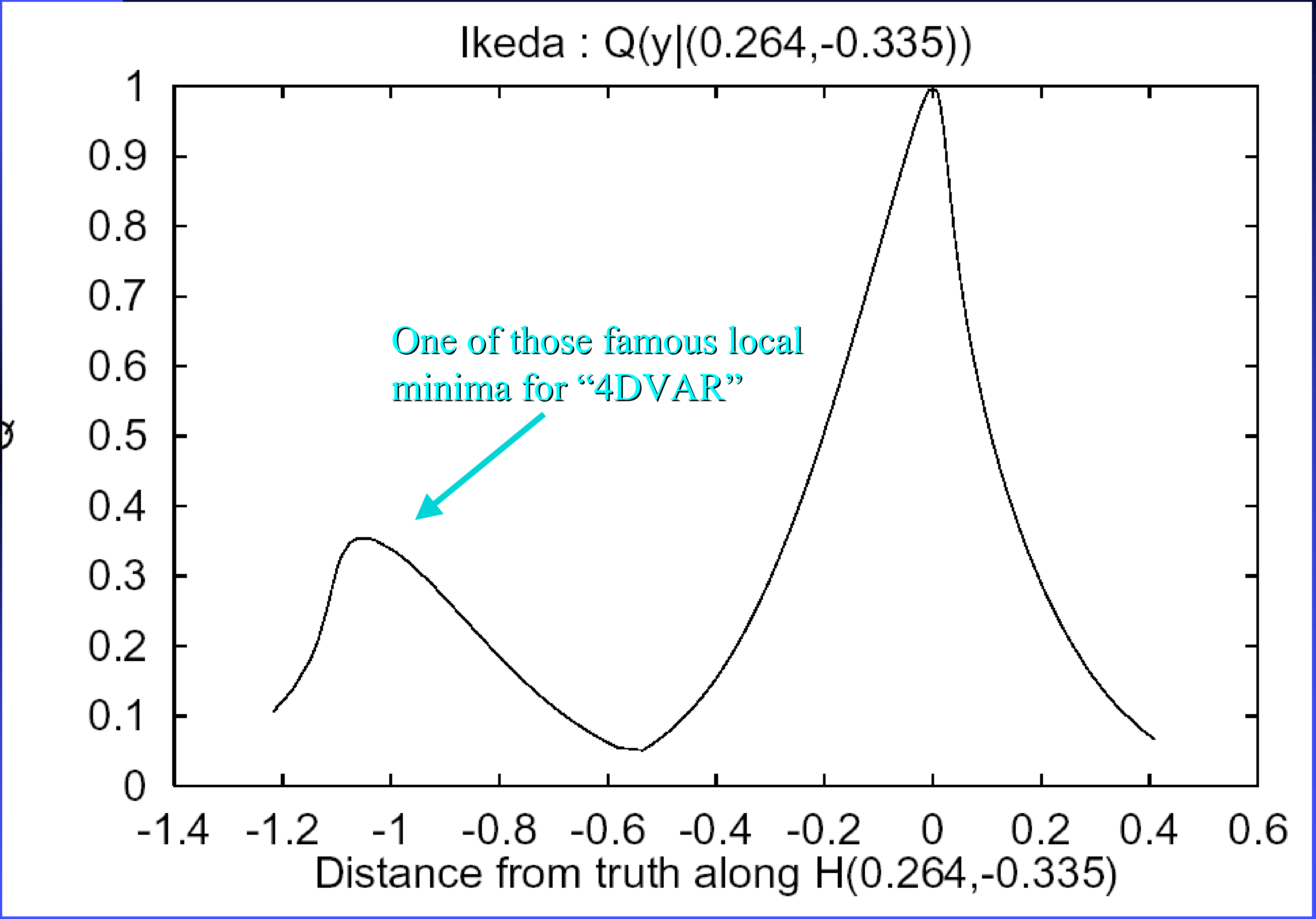
Ikeda : Some sets of indistinguishable states (Model is perfect)
1
0.5
0
-0.5
-1
-1.5
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
x
y
Indistinguishable states of this state
$P(y(t_0) \mid x(t_0)\ , \underline{F}_a(x), a, n) > 0$

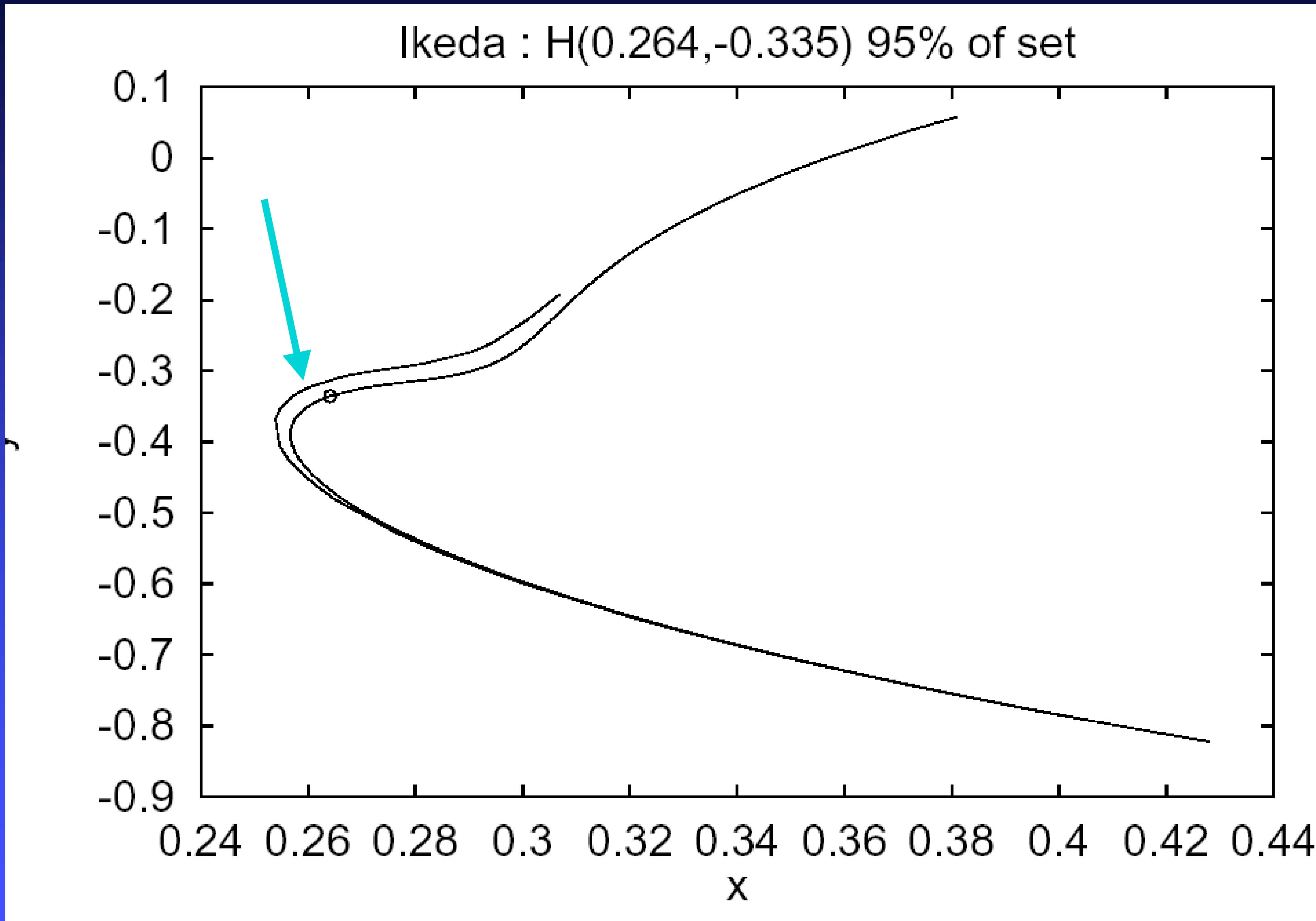
In practice, we might first estimate x*, the maximum likelihood state given (s(t), t= 0, -1, -2, …), and then determine H(x*).



Of course the RMS skill of x* is irrelevant: with Prob 1, x* will not yield the "best" forecast.


Ikeda : Q(y|(0.264,-0.335))
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-1.4
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
Distance from truth along H(0.264,-0.335)

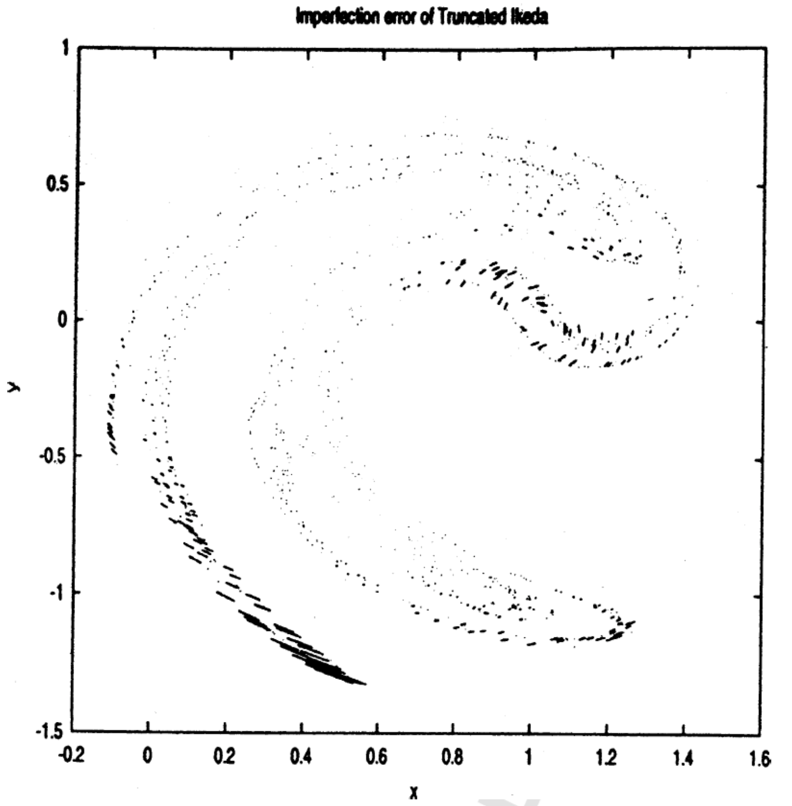

Ikeda : Q(y|(0.264,-0.335))
One of those famous local minima for "4DVAR"
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-1.4
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
Distance from truth along H(0.264,-0.335)



Ikeda : H(0.264,-0.335) 95% of set
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
0.24
0.26
0.28
0.3
0.32
0.34
0.36
0.38
0.4
0.42
0.44
x

A (good) Imperfect Model for the Ikeda System

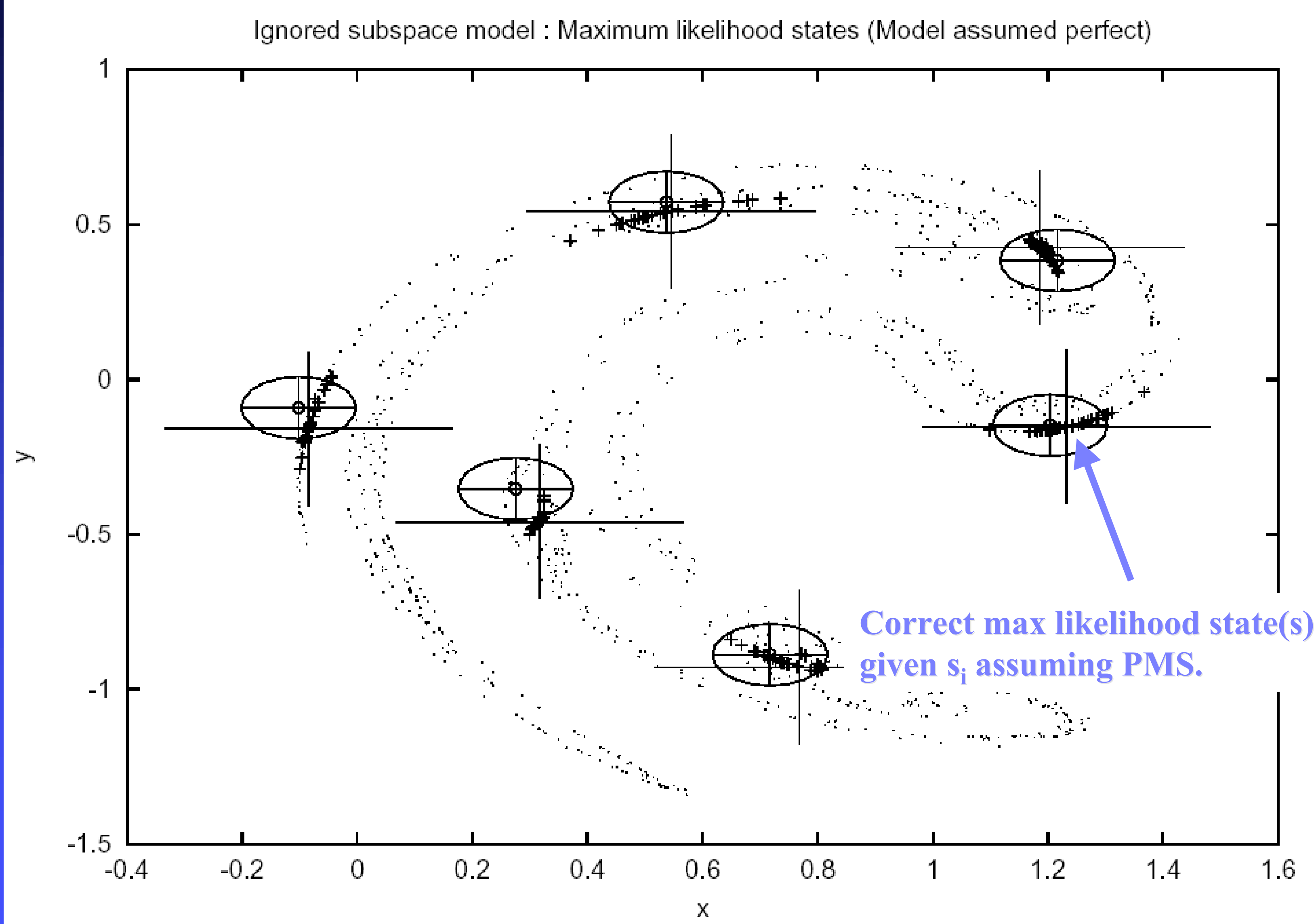
$$\cos\theta = \cos(w+\pi) \mapsto -w + w^3/6 - w^5/120,$$
$$\sin\theta = \sin(w+\pi) \mapsto -1 + w^2/2 - w^4/24,$$



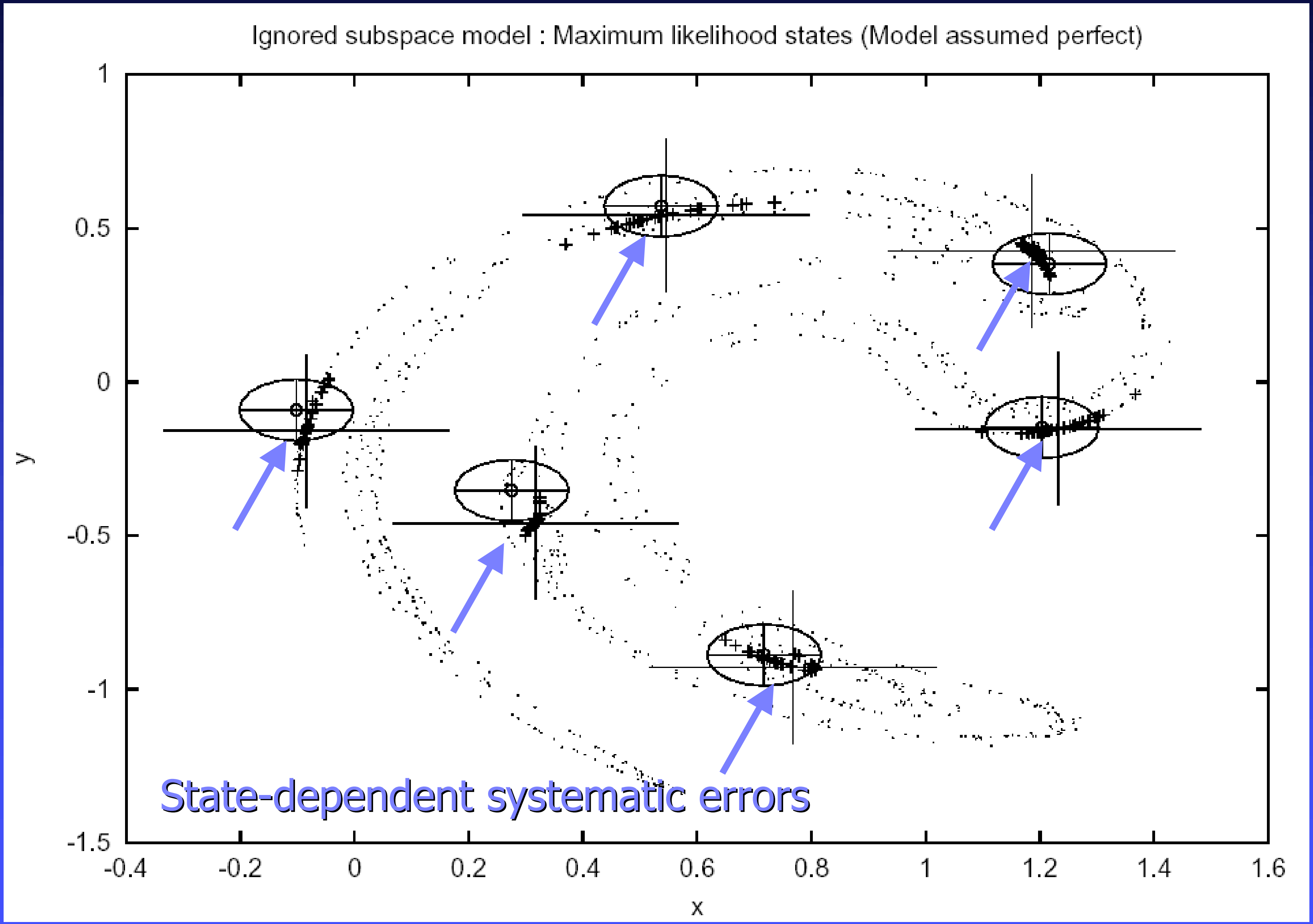
A good but imperfect model may be constructed using a finite truncation of the trigonometric expansions.

Aside: Which parameter values *should* be used in that case?

Assuming PMS when the model is imperfect introduces state-dependent systematic errors:



Assuming PMS when the model is imperfect introduces state-dependent systematic errors:



And what about the set of indistinguishable states?

In short: H(x) is empty.

As t goes to minus infinity, Q(x) goes to 0 for all x (including the trajectory that ends at x_0 = the "true x").

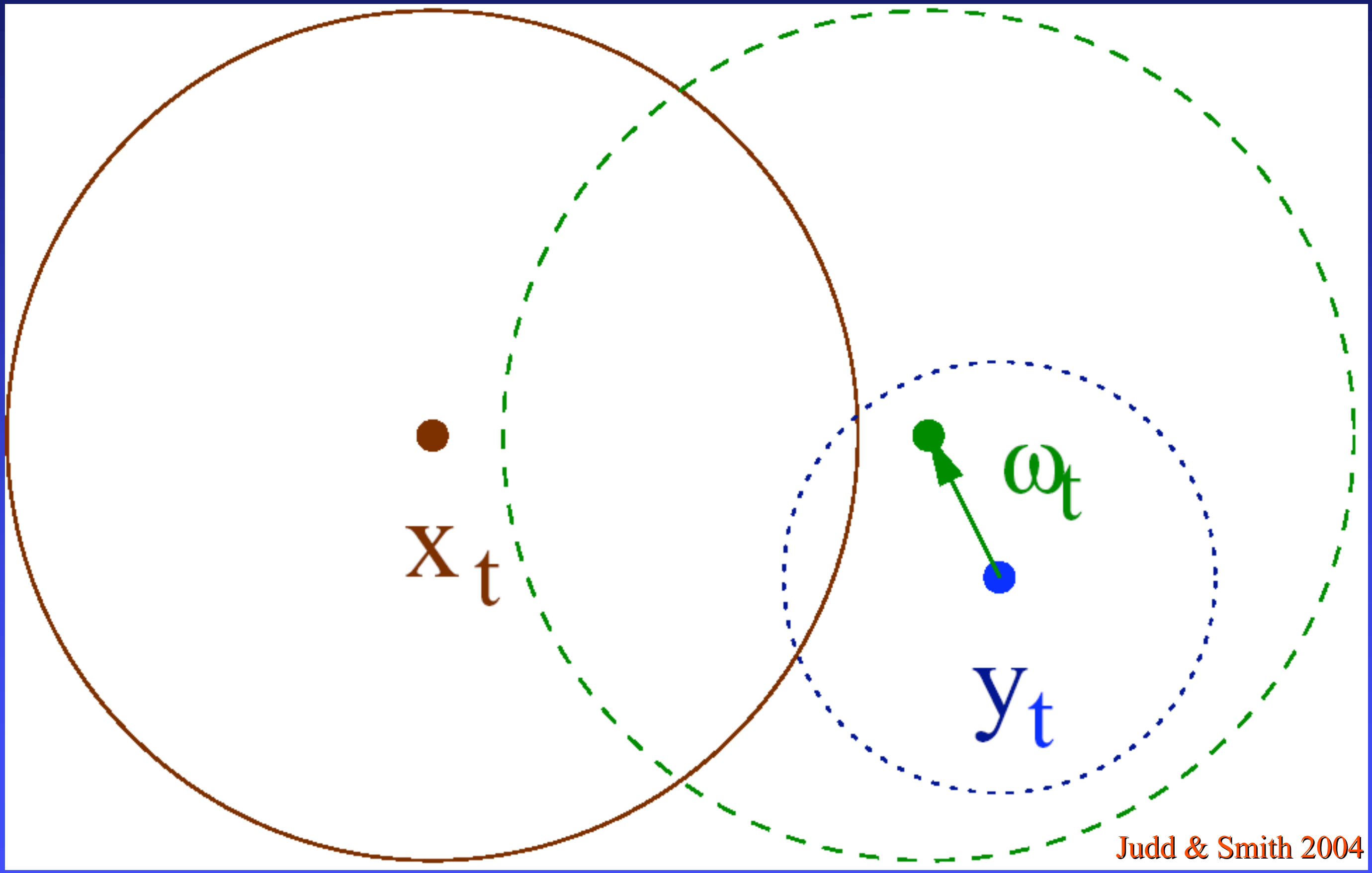
Of course we can ignore this (by renormalizing)…

Or we can make "better" DA/forecasts by considering pseudo-orbits of the model:

> explicitly adding dynamical noise and then computing the set of indistinguishable states:

Given that H(x) is empty:

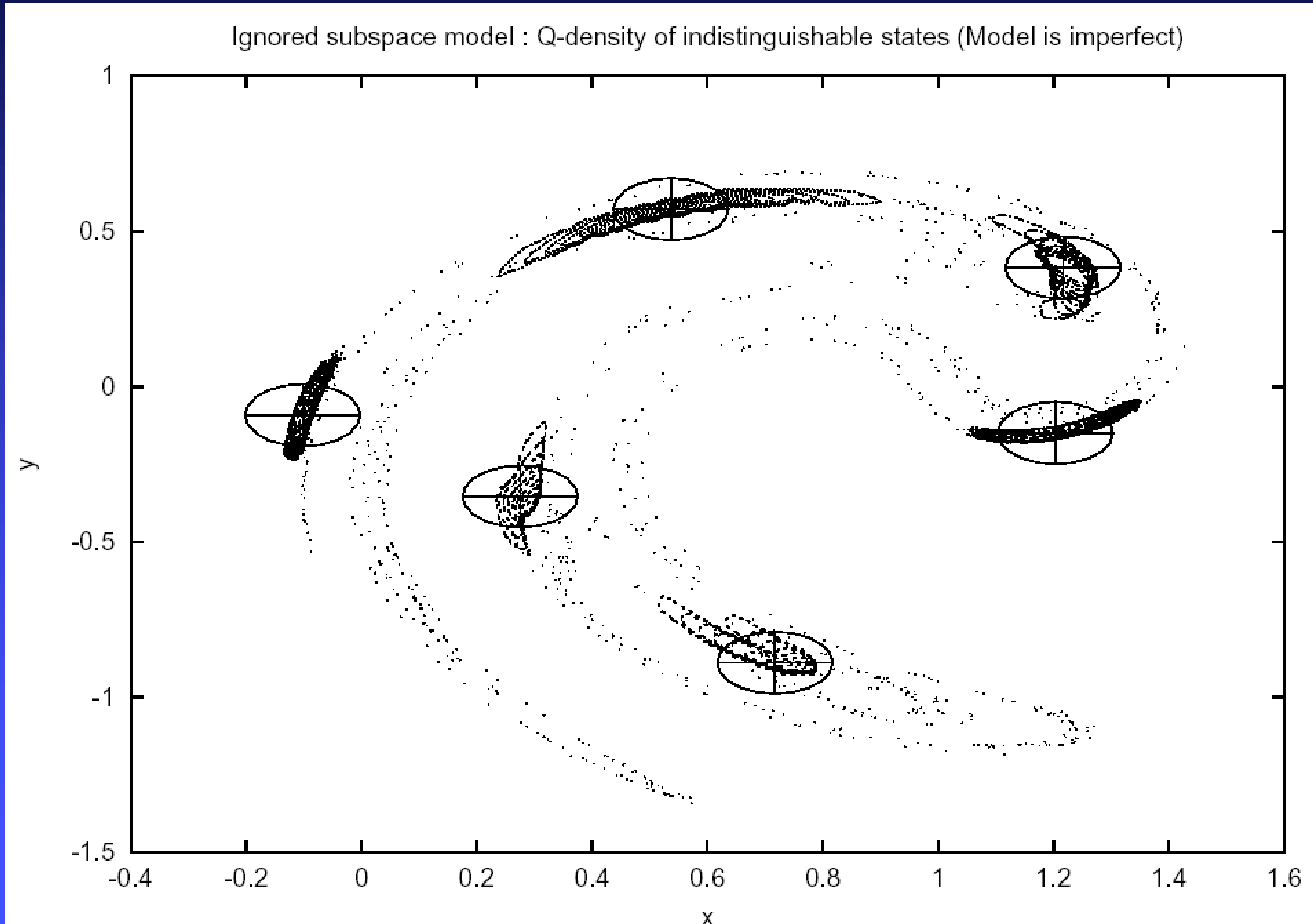
We can make “better” DA/forecasts if we change our model class and then compute the set of indistinguishable states of the new model. For example: use a stochastic model (add dynamical noise to the old model):



But all hope of an accountable PDF forecast is lost!

State estimation using pseudo-orbits out-performs those that assume PMS...

But what is the point? What is the goal?



So we have retreated from prediction to mere state estimation.

Outside PMS we can "solve" this problem within the indistinguishable states framework; but note we have already cheated a bit:

1) The state space and the model-state space are different spaces! (Lorenz's "subtractable")
2) We should be allowed a projection operator (here taken to be the identity).

And how do we know if something works?

1) It is better under data denial.
2) It yields "better" (more useful) forecasts.
   (IGN or ?Rel Entropy? via dressing and users cost functions…)
3) Internal consistency and beauty.

Traditional aims of state estimation:

$$P(x(t_0)\ |\ s_i, F_a(x), a, n)$$

Traditional aim of forecasting (in statistics)

$$P(x(t > t_0)\ |\ s_i, F_a(x), a, n)$$

In cases where $F_a(x)$ is imperfect (*i.e.* in practice), these two procedures may have different target different distributions for $P(x(t_0))$.

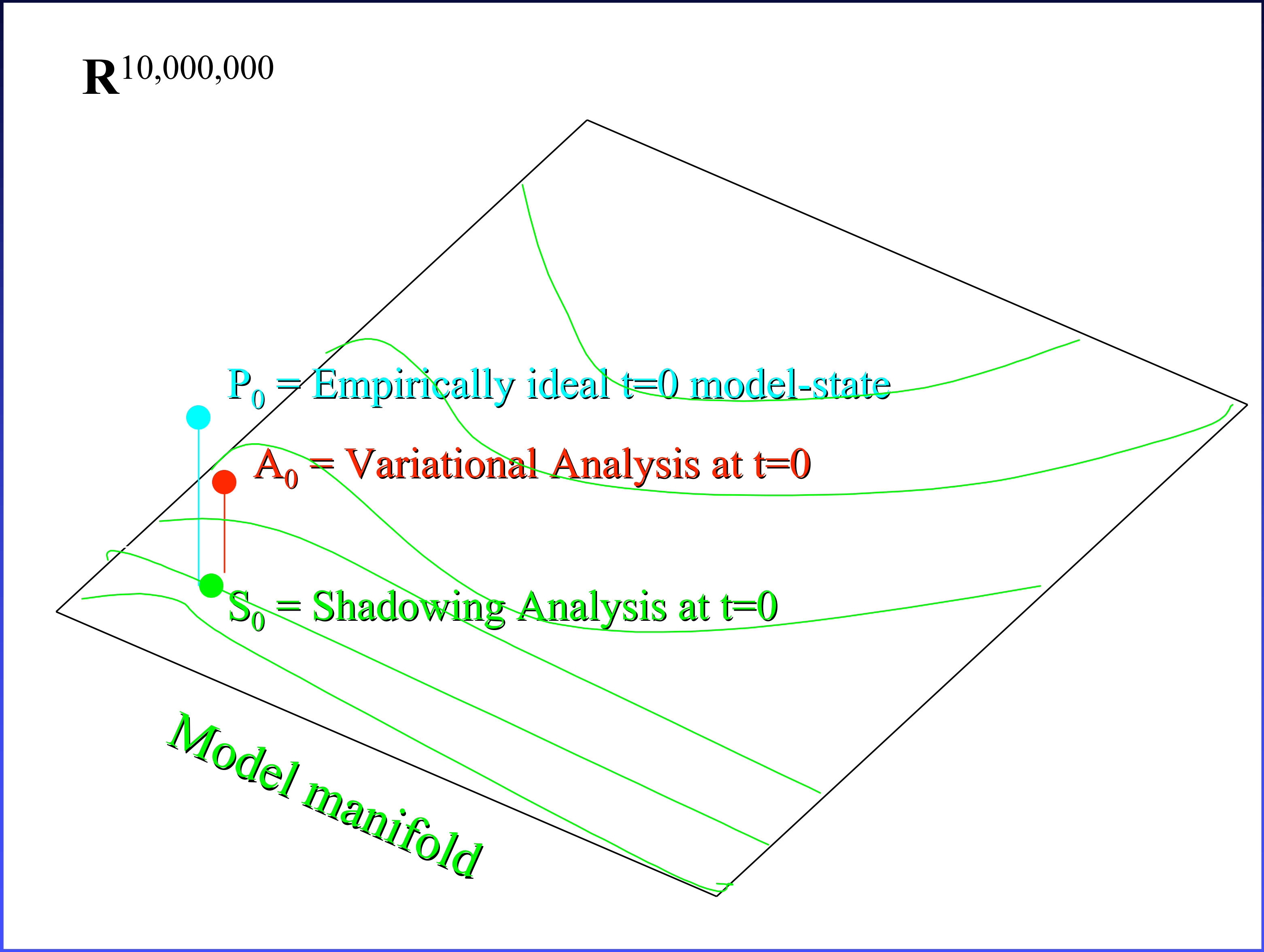
Evaluation of $P(x(t_0))$ via data denial is not expected to yield the same ranking as forecast evaluation of $P(x(t>t_0))$.

Future directions:

1) Accounting for lower dimensional dynamics

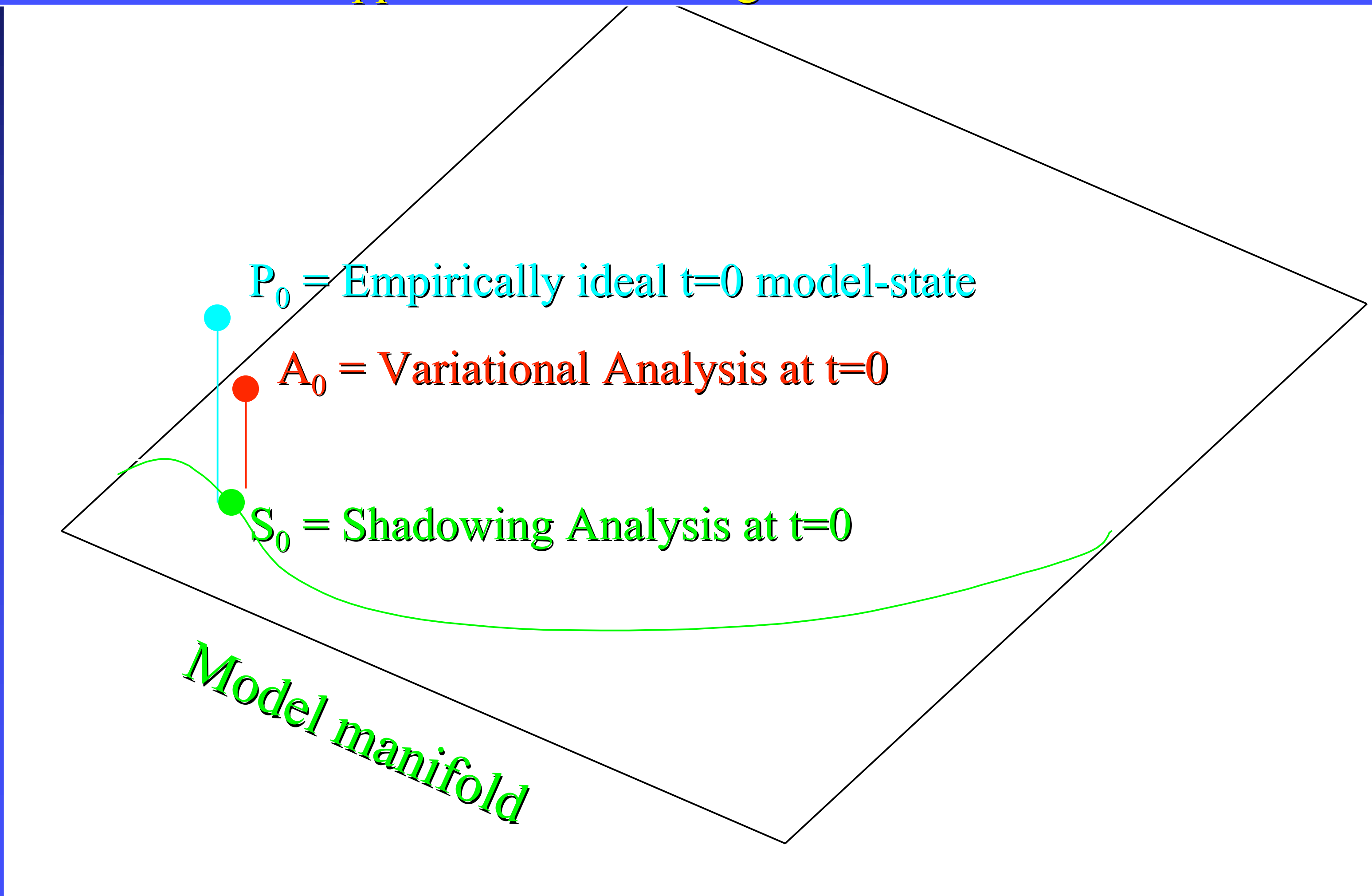
2) Adaptive obs and indistinguishable states

3) Data assimilation given unrealistic state-of-the-art models


$\mathbf{R}^{10,000,000}$
P_0 = Empirically ideal t=0 model-state
A_0 = Variational Analysis at t=0
S_0 = Shadowing Analysis at t=0
Model manifold

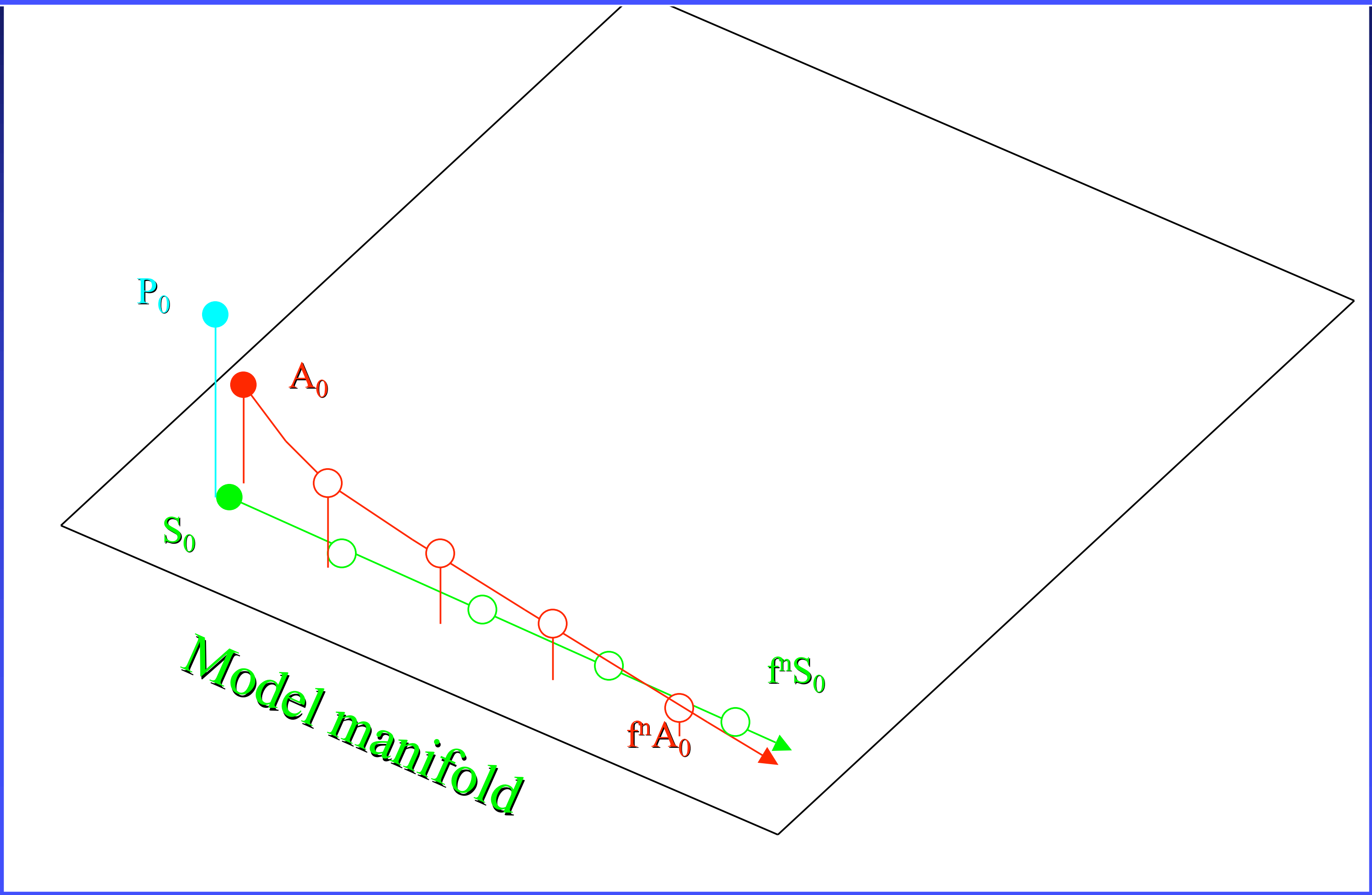
We might keep P_n as a target/verification,
but P_0 is unlikely to provide model-initial condition(s).

Variational Assimilation pulls the initial conditions away from the manifold.
What happens when we "let go" and forecast…

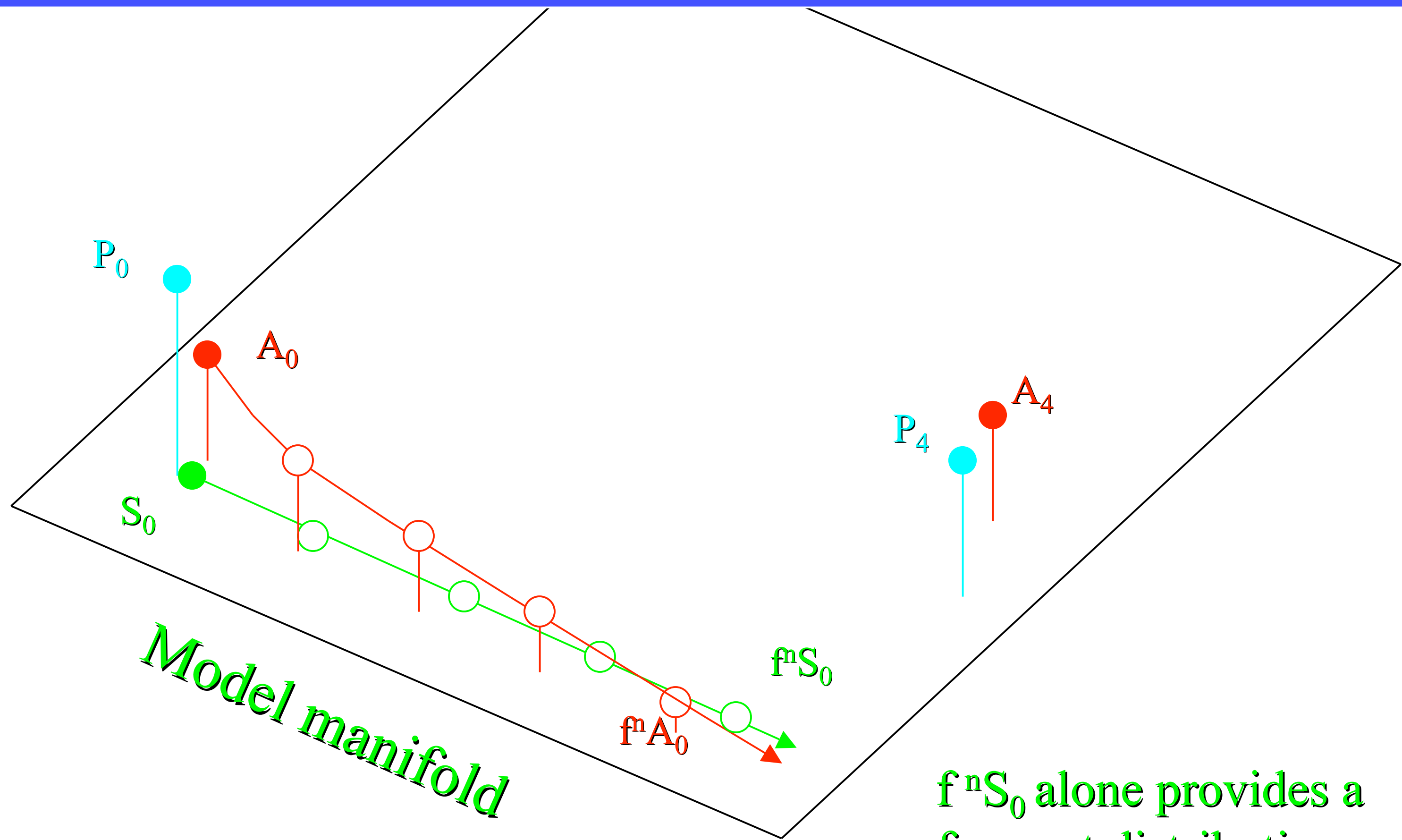


What happens when we "let go" and forecast…

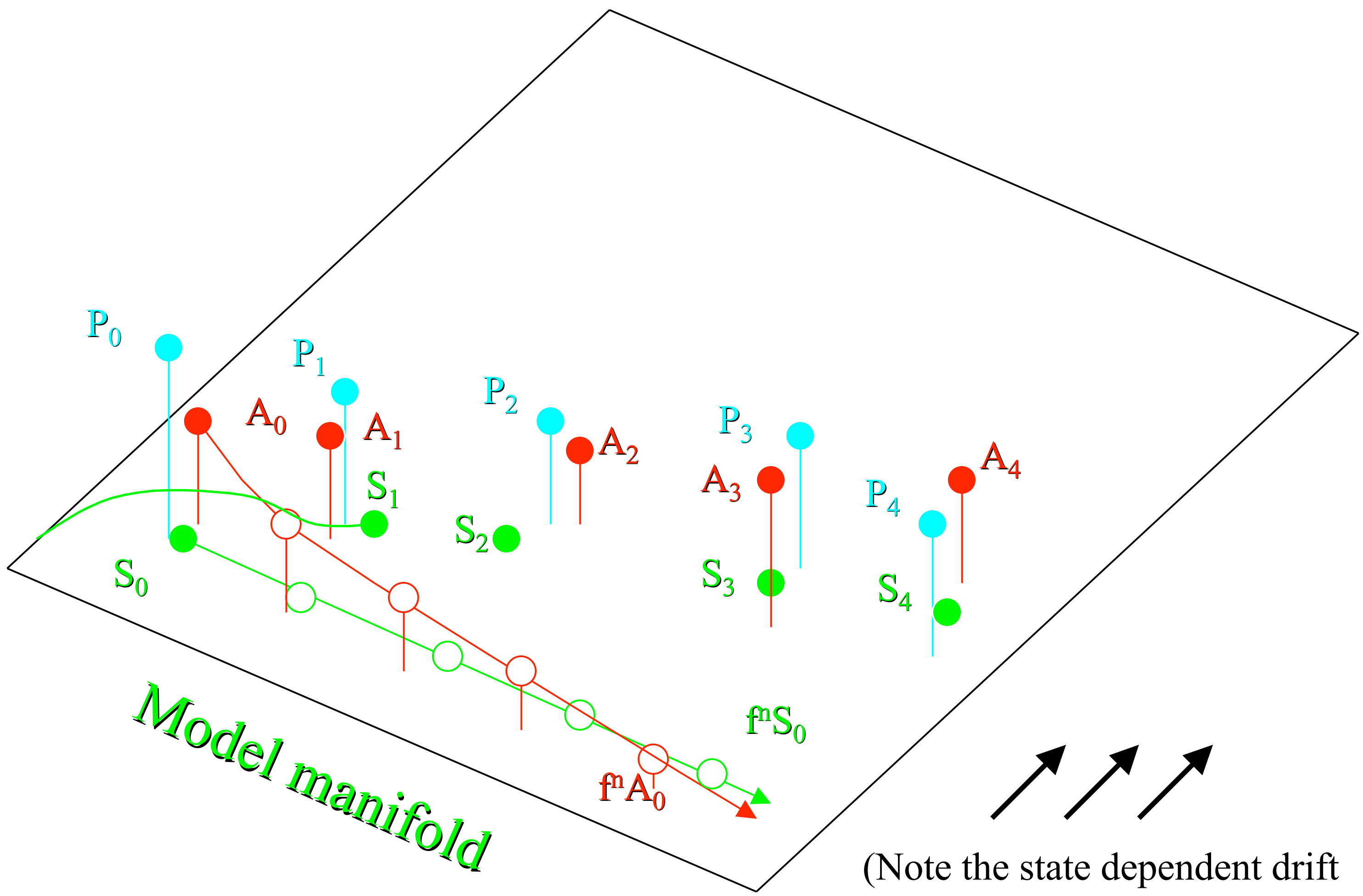
A_0 immediately falls toward *somewhere* on the manifold



We are allowed a projection operator to map $f^{n}S_0$ into a distribution; we take this freedom even if we verify against P!

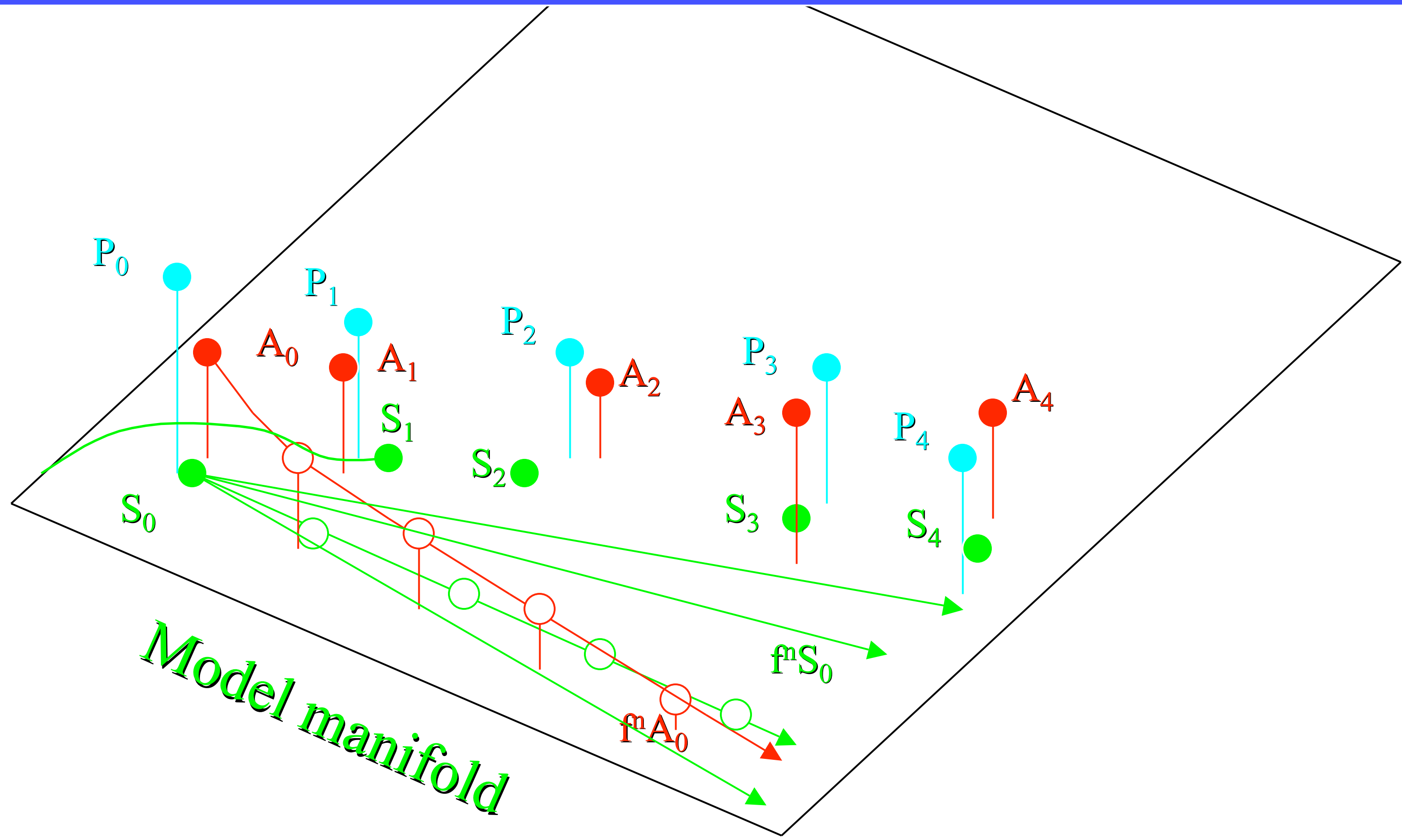


$f^{n}S_0$ alone provides a forecast distribution

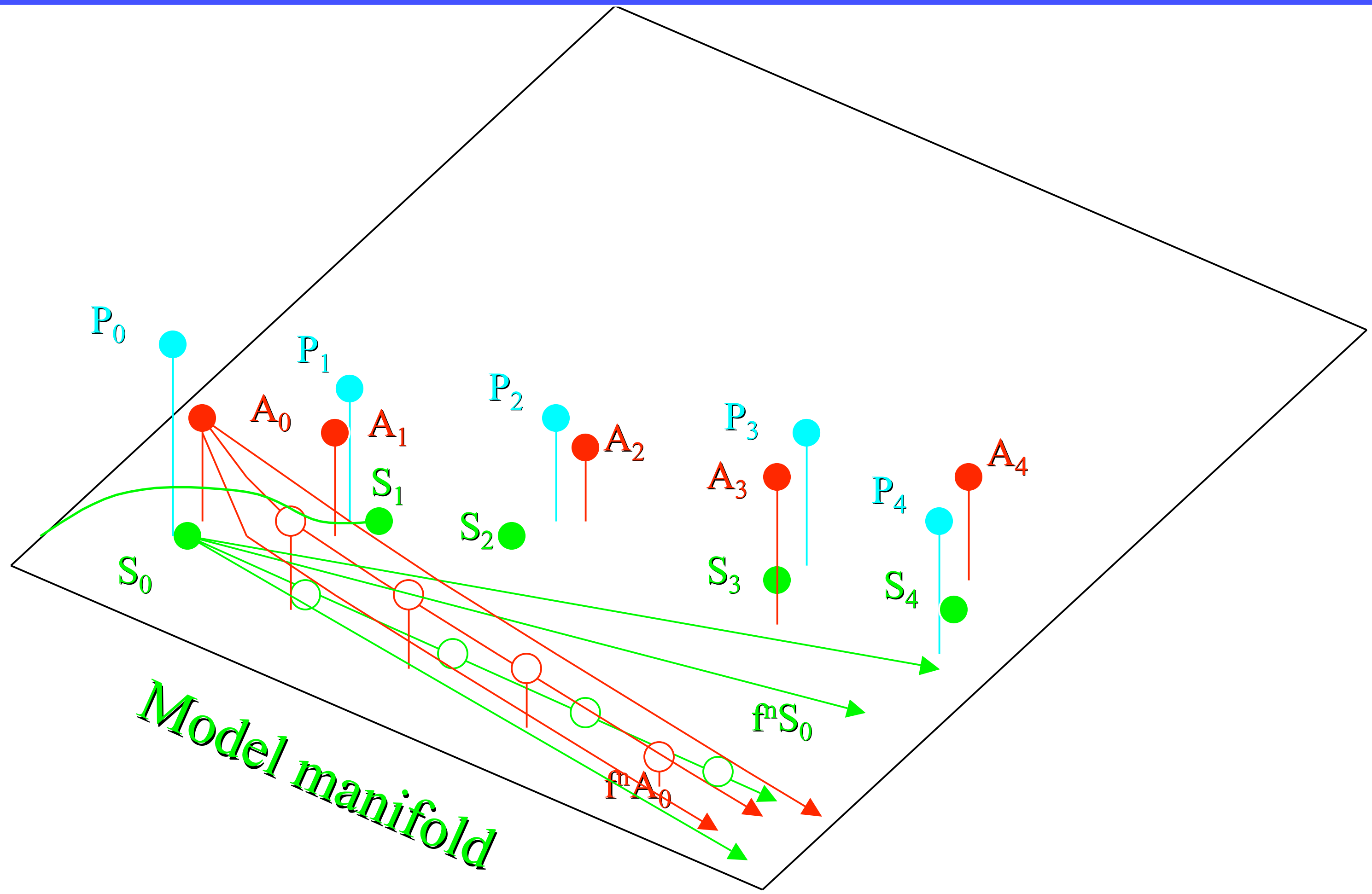


(Note the state dependent drift
due to model inadequacy…)

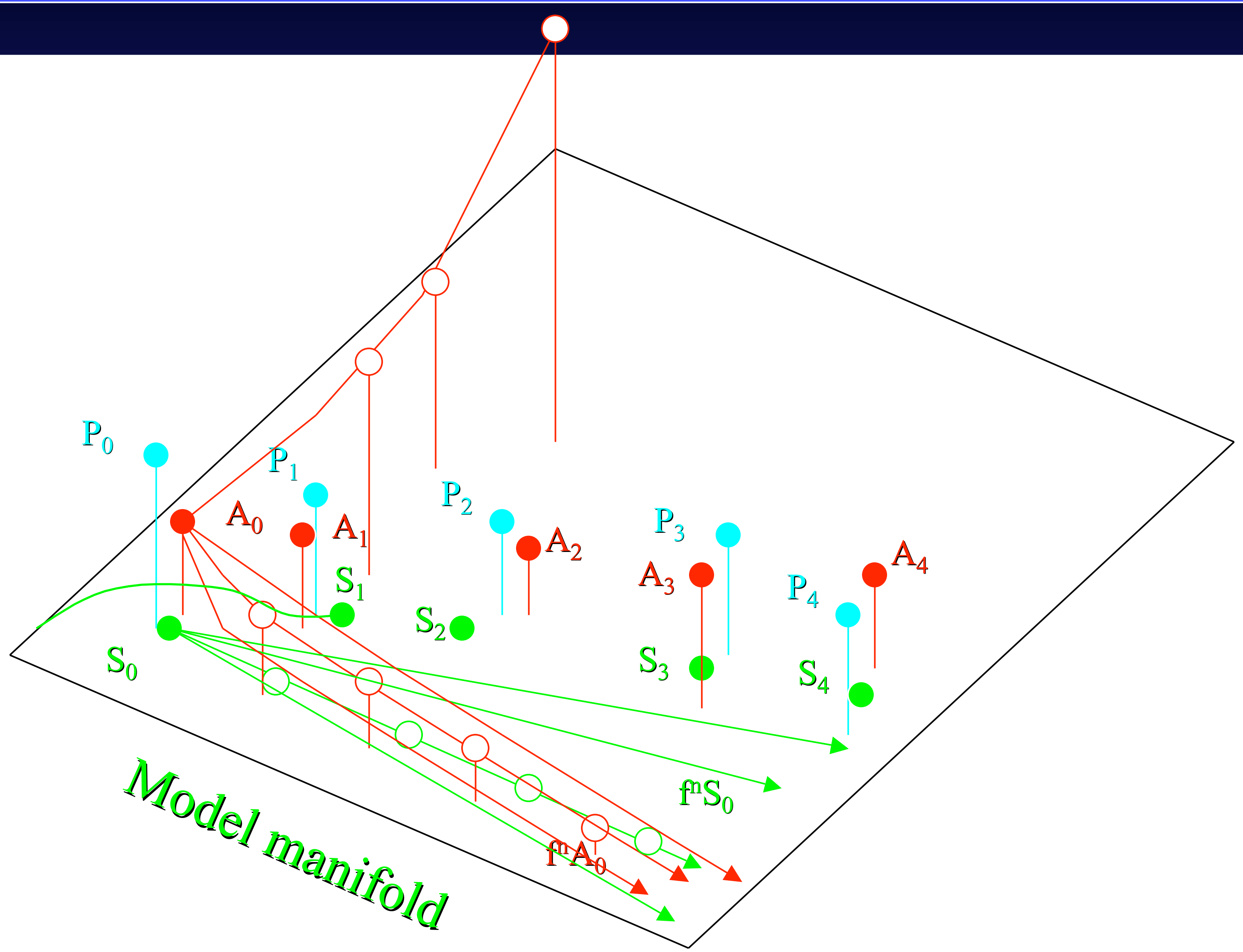
But if we have taken ensembles seriously then we have an ensemble of simulations from near S.



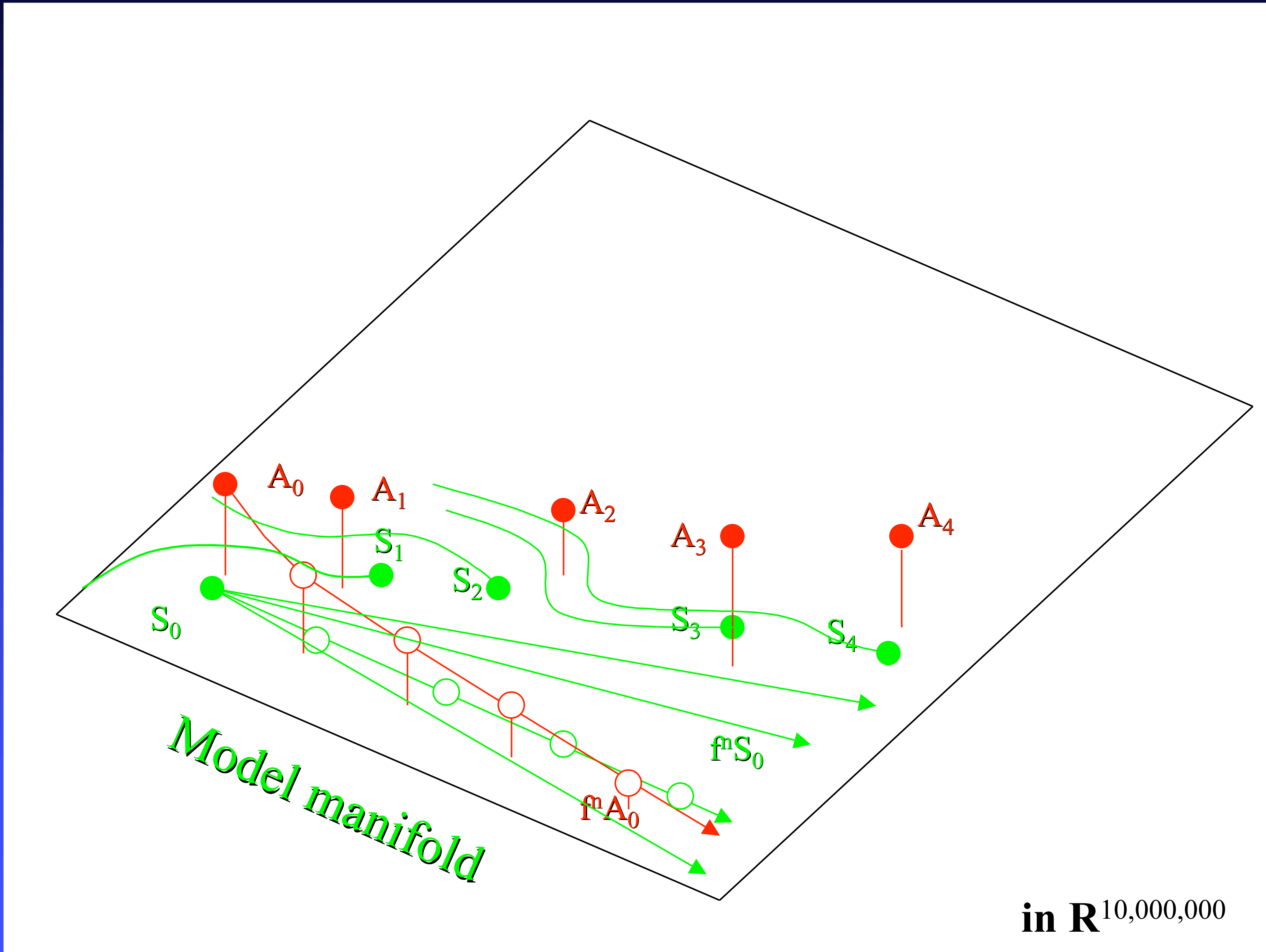
And an ensemble of model simulations from near A.



Of course, points near A can fall onto other bits of the manifold.



What can we know operationally?

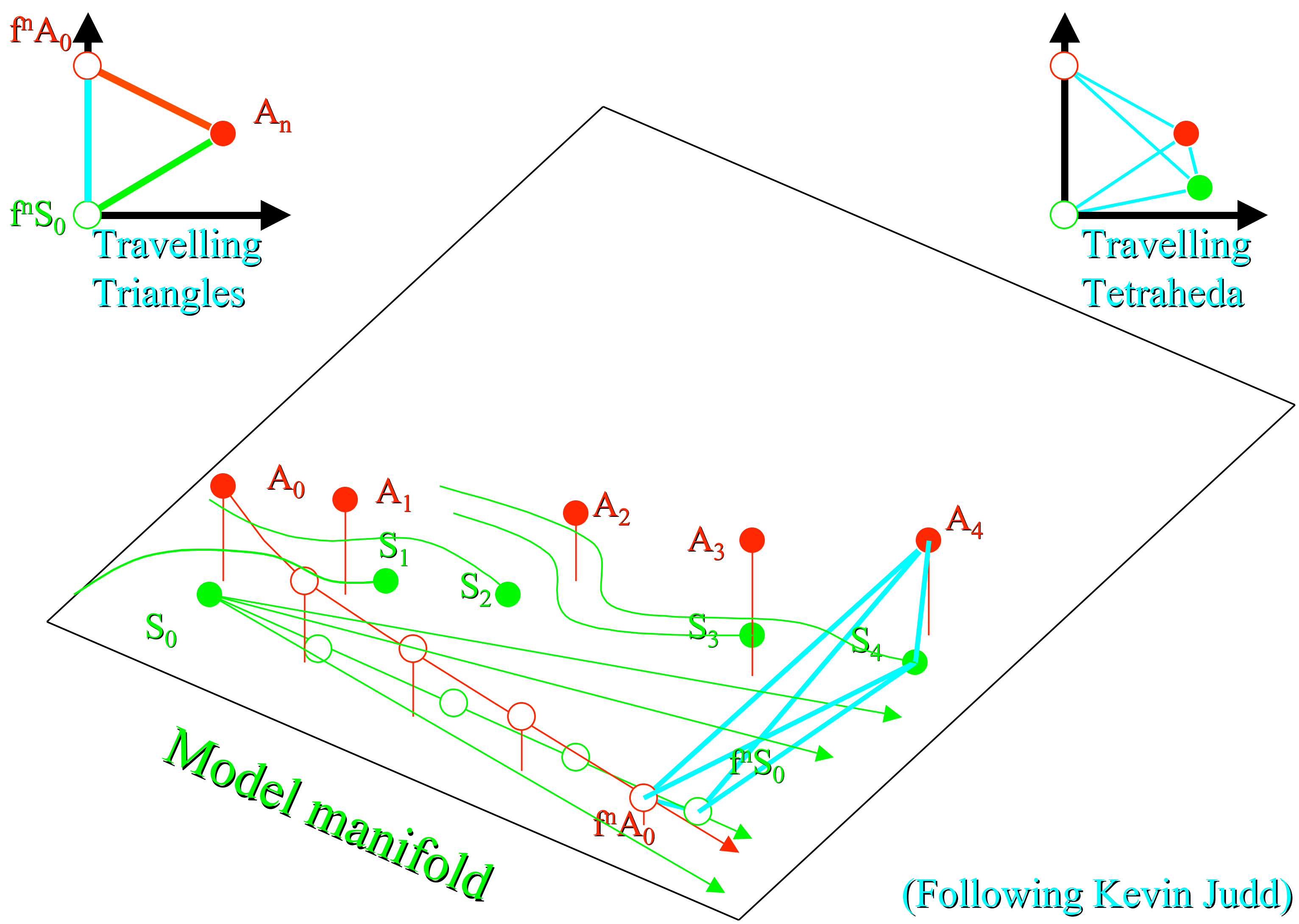


The empirically relevant questions are (include):

Is it better to pull A_0 away from the model manifold or to project S_n back into the obs space?
How would ensembles on/near the model manifold compare to ensembles from perturbed variational analyses?

For other physical systems, taking initial conditions on the model manifold seems to be better; for NOGAPS, Judd et al have suggested that this yields no RMS penalty for a single model run after day two:

But could we ever interpret such diagrams operationally?

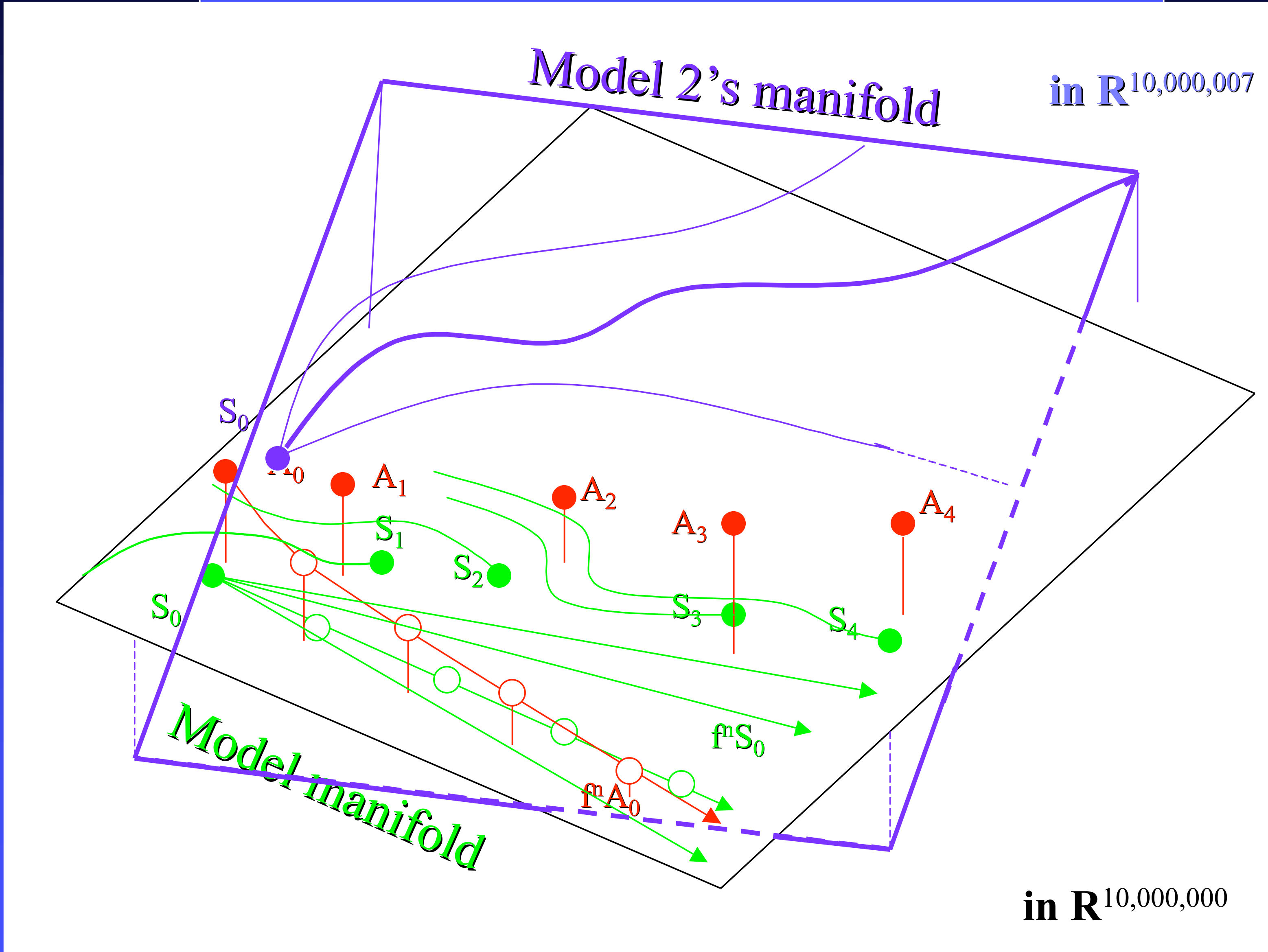


(Following Kevin Judd)

in $\mathbf{R}^{10,000,000}$

What is the aim of DA given two models?

$P(x(t > t_0) \mid s_i, F_a(x), a, n_a\ G_b(x'), b, n_b)$



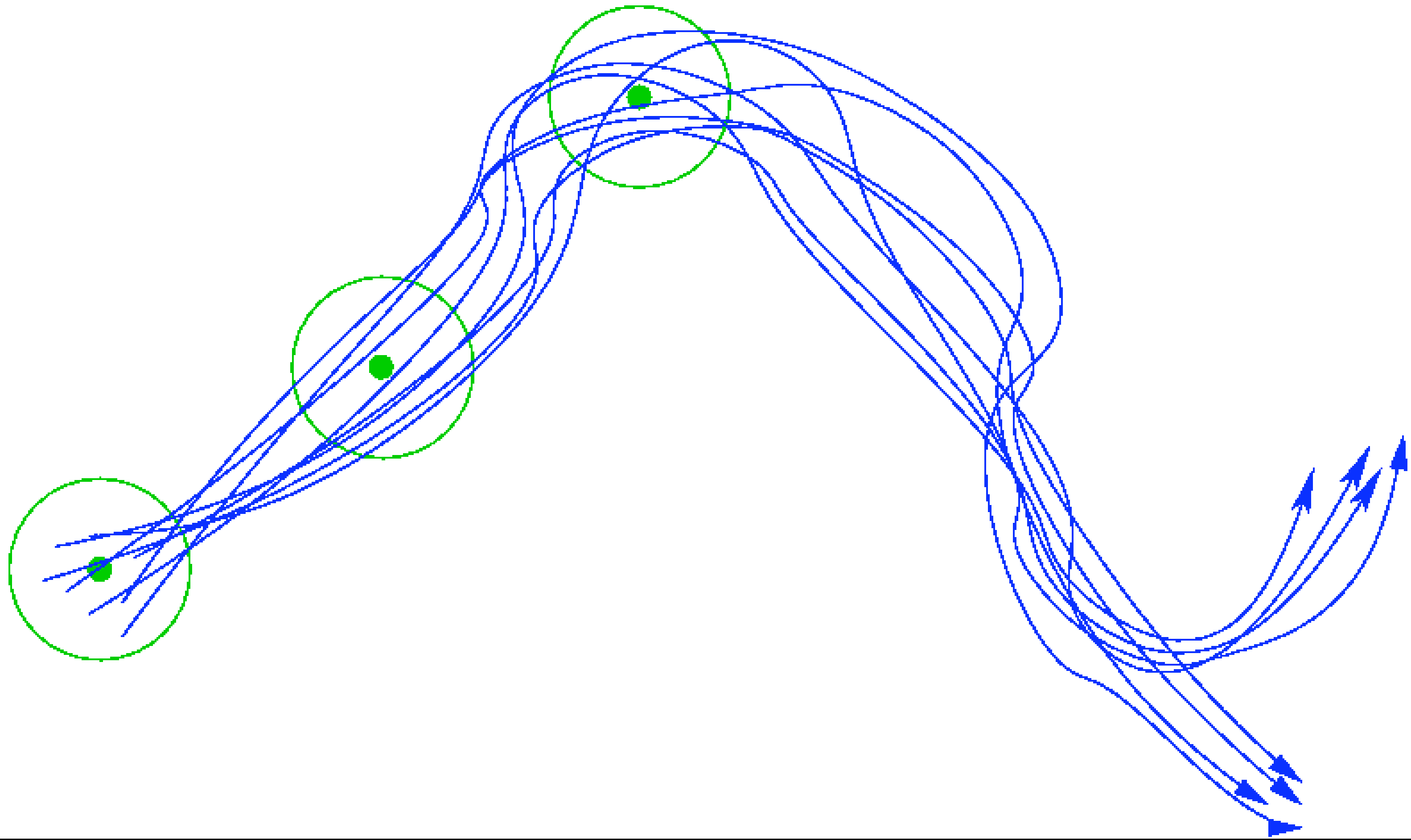
What are the implications of lower dimensional dynamics for Data Assimilation?

Is it better to pull A_0 away from the model manifold or to project S_n back into the obs space?

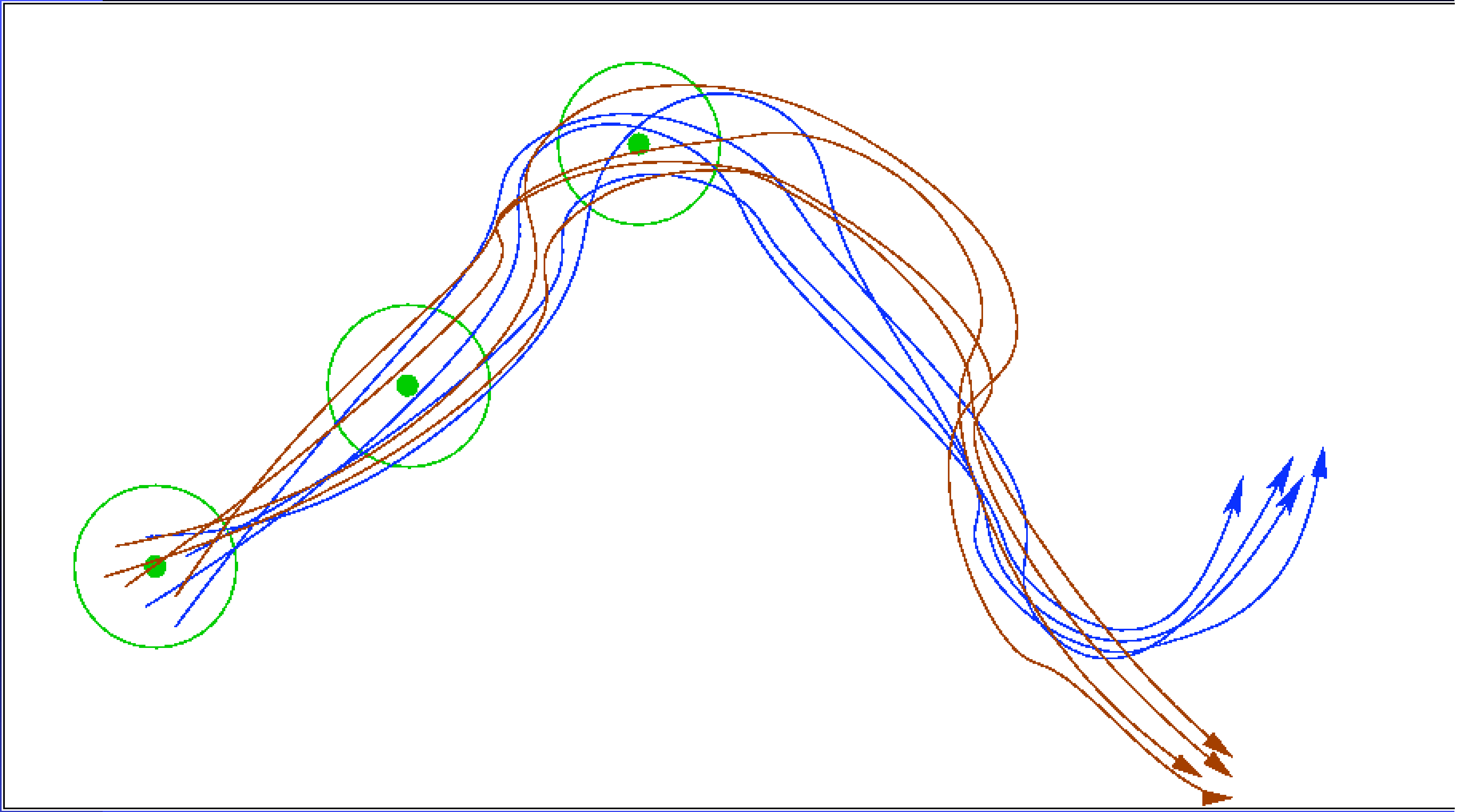
How would ensembles on/near the model manifold compare to ensembles from perturbed variational analyses?

What do multiple models look like in this context?

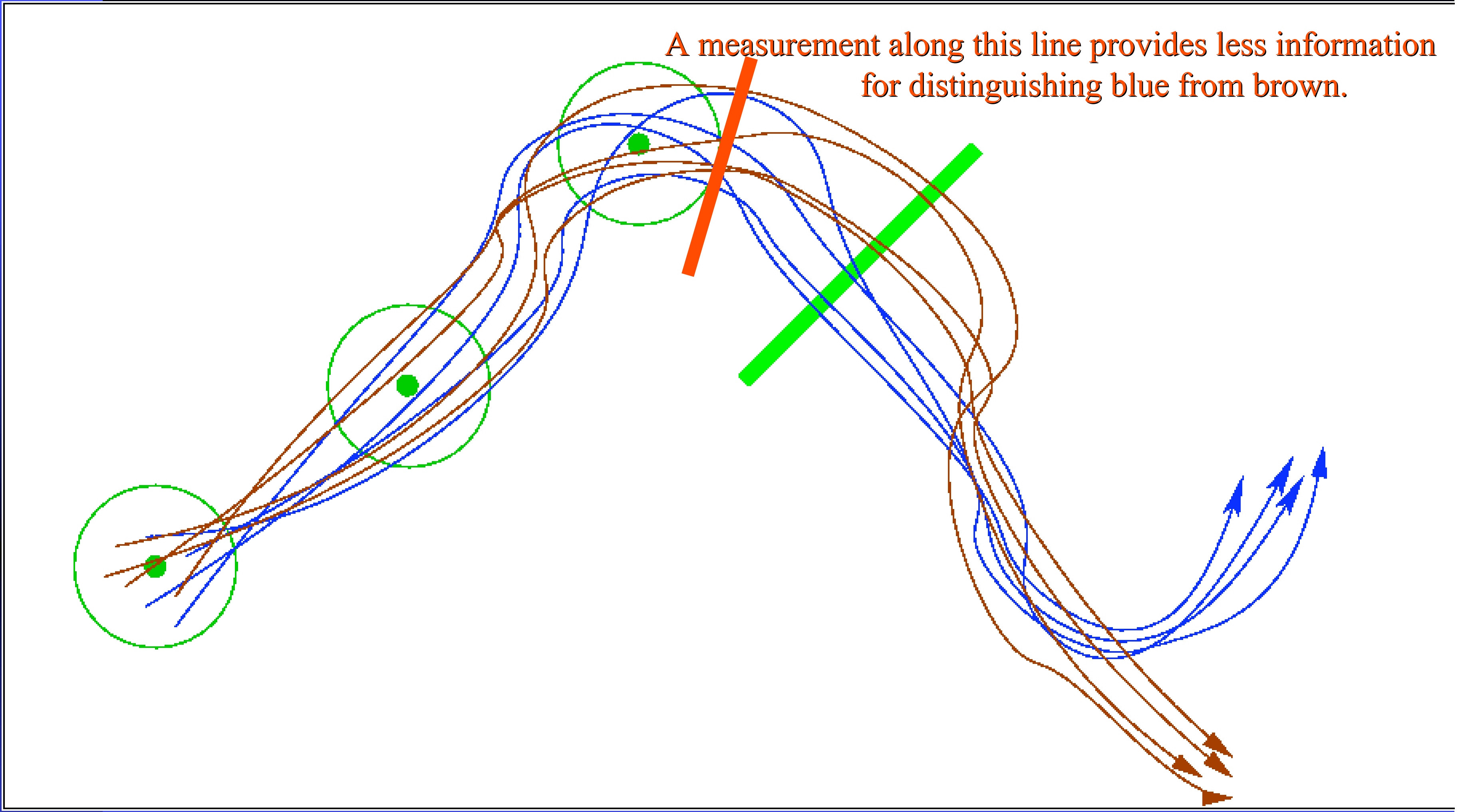
Data Assimilation and Adaptive Observations



Suppose we wish to distinguish two sets of simulations (say, storm/no storm); in terms of indistinguishable states, the AO question is simply "Which observations are most likely to separate these sets?"



To do this, merely color the trajectories in each set, and determine the observation in space and time (post ‘now’) that is likely to yield the most relevant information.



No linearization,
No implicit perfect model assumption,
And the ability to update the AO in light of scheduled obs without rerunning the simulations.

Data Assimilation and the Press:

FREE

METRO

Clubbing: This is the only guide you need

Page 25

Thursday, January 27, 2005 www.metro.co.uk

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.

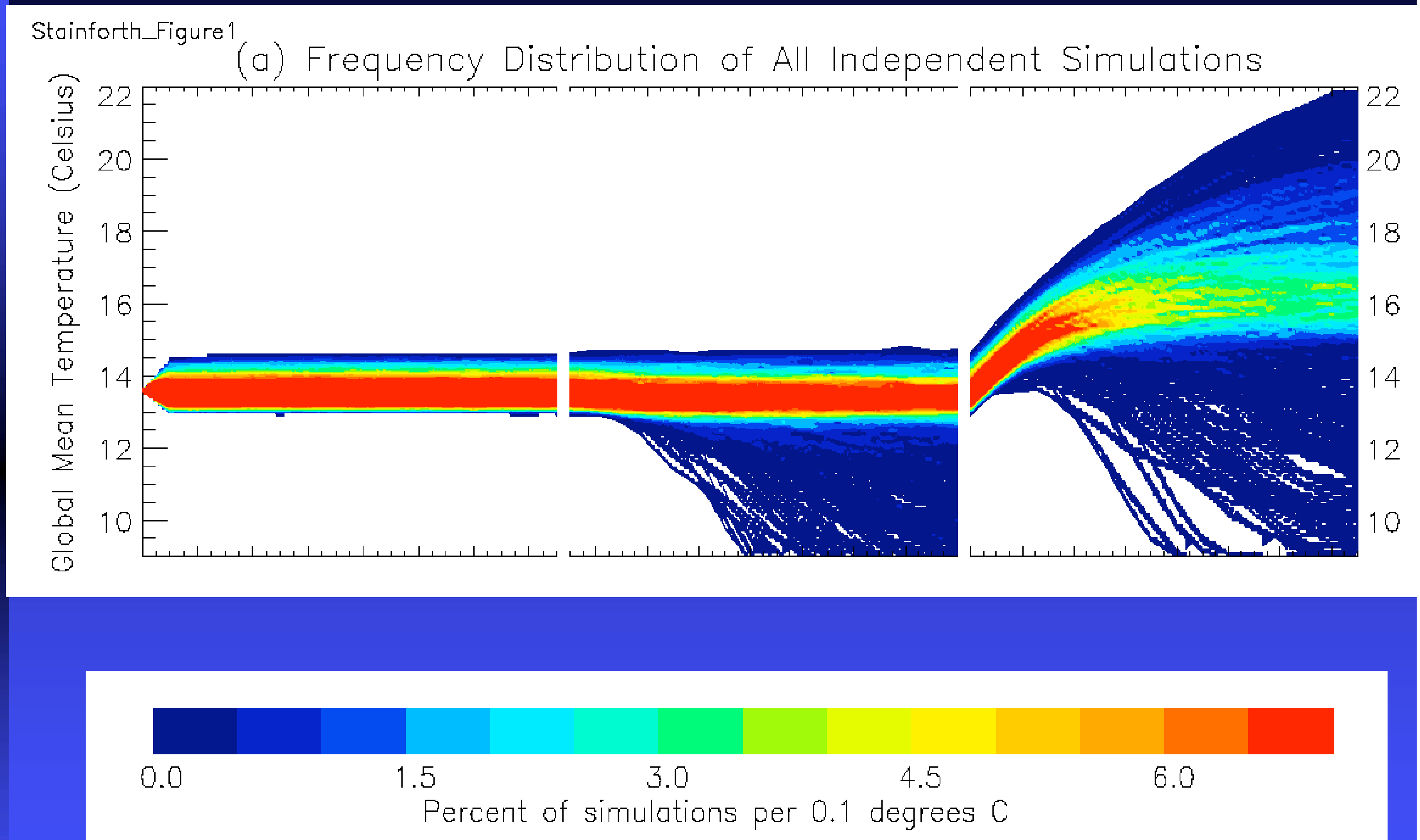
'I think it would probably not be a tropical paradise. The UK would be at the high end of this change, well into the teens as the temperature changes. I don't think we'll be building many snowmen in winter, or going sledging.'

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.

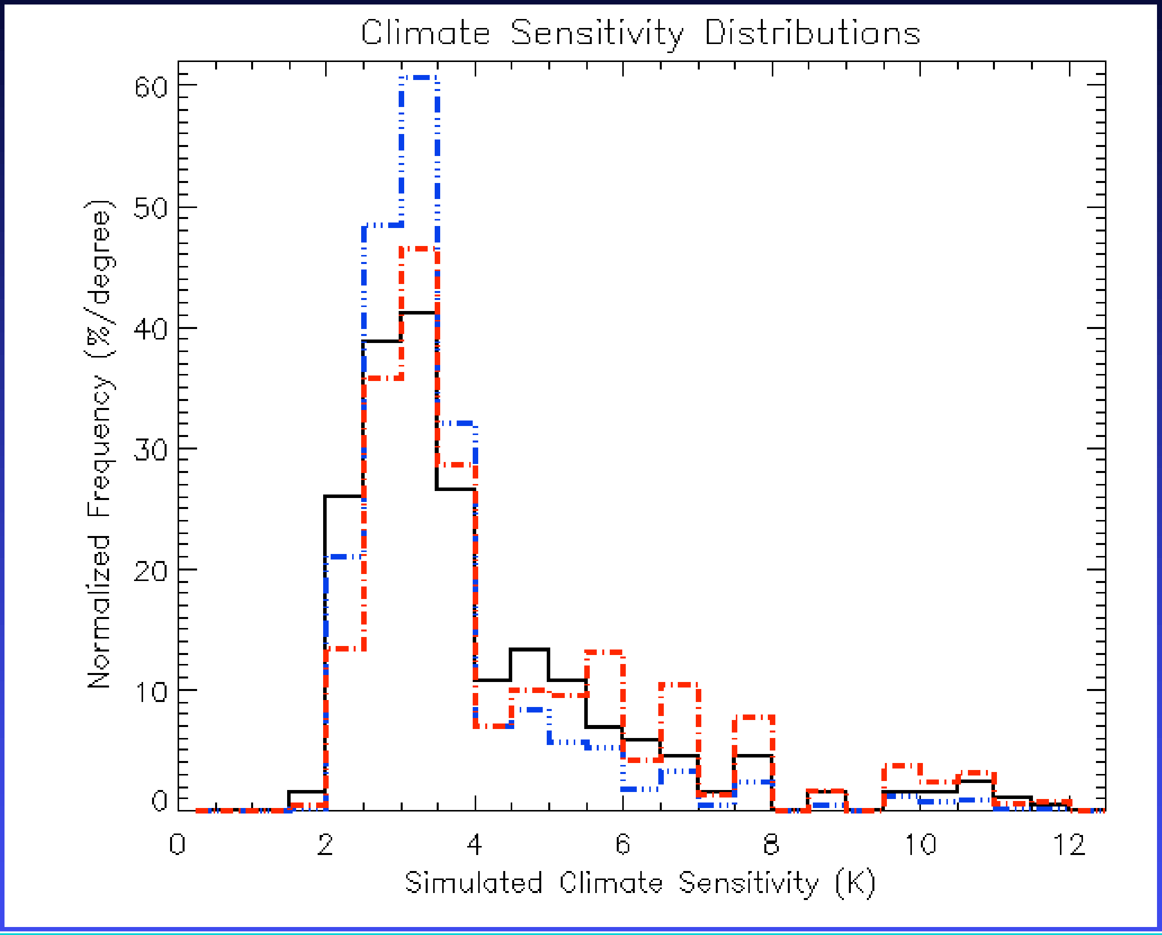
The warning came as Tony Blair used the World Economic Forum in Switzerland to call for action on global warming and to pressure America to sign up to Kyoto.

Blair's call – Page 5

23 Feb 2005


Stainforth_Figure1
(a) Frequency Distribution of All Independent Simulations
Global Mean Temperature (Celsius)
22
20
18
16
14
12
10
0.0
1.5
3.0
4.5
6.0
Percent of simulations per 0.1 degrees C

The model is merely a transfer function; there is no relevant prior.



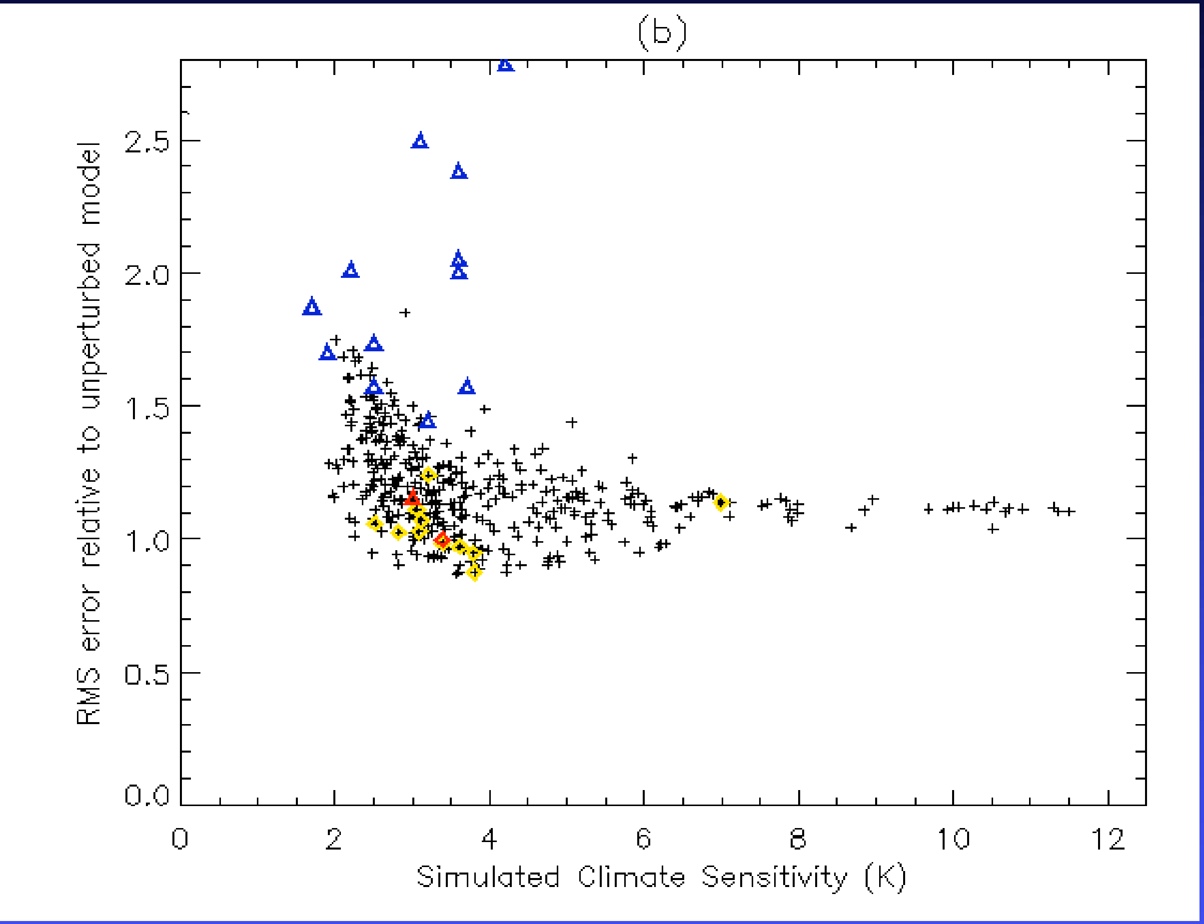
Black: As sampled in cp.n

Blue: Entrainment coef constant

Red: Cloud-to-rain conversion threshold constant

Note that sampling uniform in *a* differs from uniform in *1/a*.

Is the climate*prediction*.net ensemble "state-of-the-art"? (Yes) How can we best use models which are *not* realistic?



Relative RMS Error relative to unperturbed model:

Yellow Diamond: Single Parameter Perturbation

Black Plus: Multiple Parameter Perturbation

Blue Triangle: CMIP II Model

Red Triangle: HadCM3 (same atmosphere with dynamic ocean)

Conclusions and Open IAM Questions:

Forecasts can improve *because* the initial conditions look worse.
How can we best measure forecast quality?

Data assimilation has different aims for now-casting and forecasting.

Data assimilation can produce ensemble(s) of initial conditions, each "on" that model(s) manifold. **Will these yield better probabilistic forecasts?**

Adaptive observations are straight-forward in the IS context (multi-model).

What is the shadowing time of operational NWP models?
Climate models?

Given short shadowing times, how can we best use DA methods for parameter estimation in climate estimation?

Forecasts will be improved by better resolving the projection from the simulation(s) to forecast (Moving further from the identity operator toward conditioning on the joint distribution of all simulations).

Probabilistic forecasts may prove more valuable by not providing probabilities.

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

Dave Stainforth (2005) *et al.* 27 Jan *Nature*

K Judd and LA Smith (2001) *Indistinguishable States I* , Physica D **151**: 125-151; *(2004) Indistinguishable States II,* **196:** 224-242 .

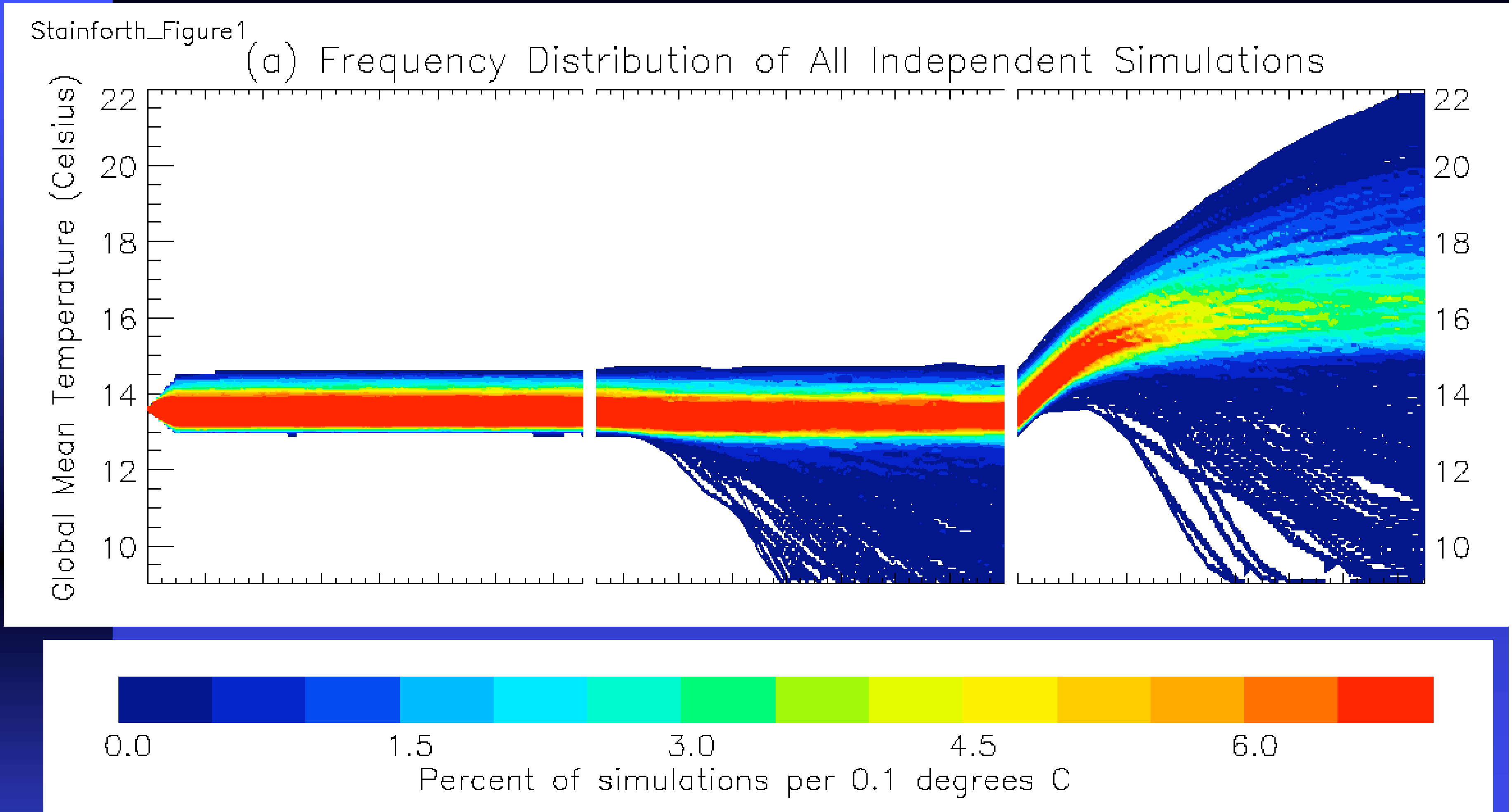
M. Altalo and LA Smith (2004) Environmental Finance **6** (1) 48-49.

M Roulston *and LA Smith* (2003) MWR **130** (6): 1653-1660.

A Weisheimer, L.A.Smith and K Judd (2004) A New Look at DEMETER forecasts via Bounding Boxes Tellus (to appear).

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

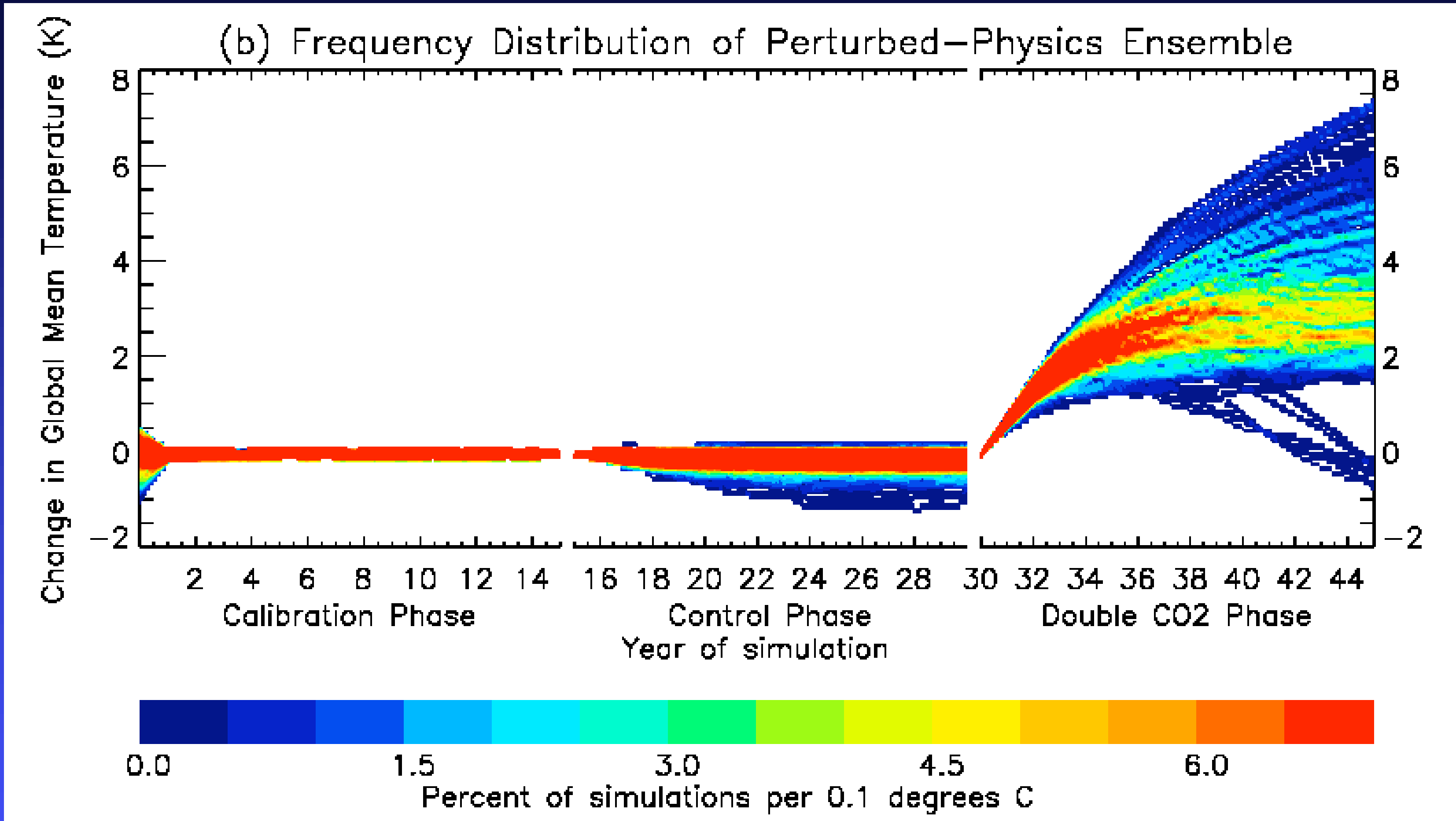
www.lsecats.org lenny@maths.ox.ac.uk



This distribution represents the evolution of several thousand full GCM climate model runs; but what should we call it?

How should I interpret this distribution?

Model is a transfer function: what is a “prior”?



Mere frequency distribution or (physically-relevant) probability distribution?

Climate is defined as a distribution of weather states; we must sample initial states in order to describe this distribution and to obtain statistically meaningful results on instabilities.

Model dynamics reduces the impact of the particular initial conditions

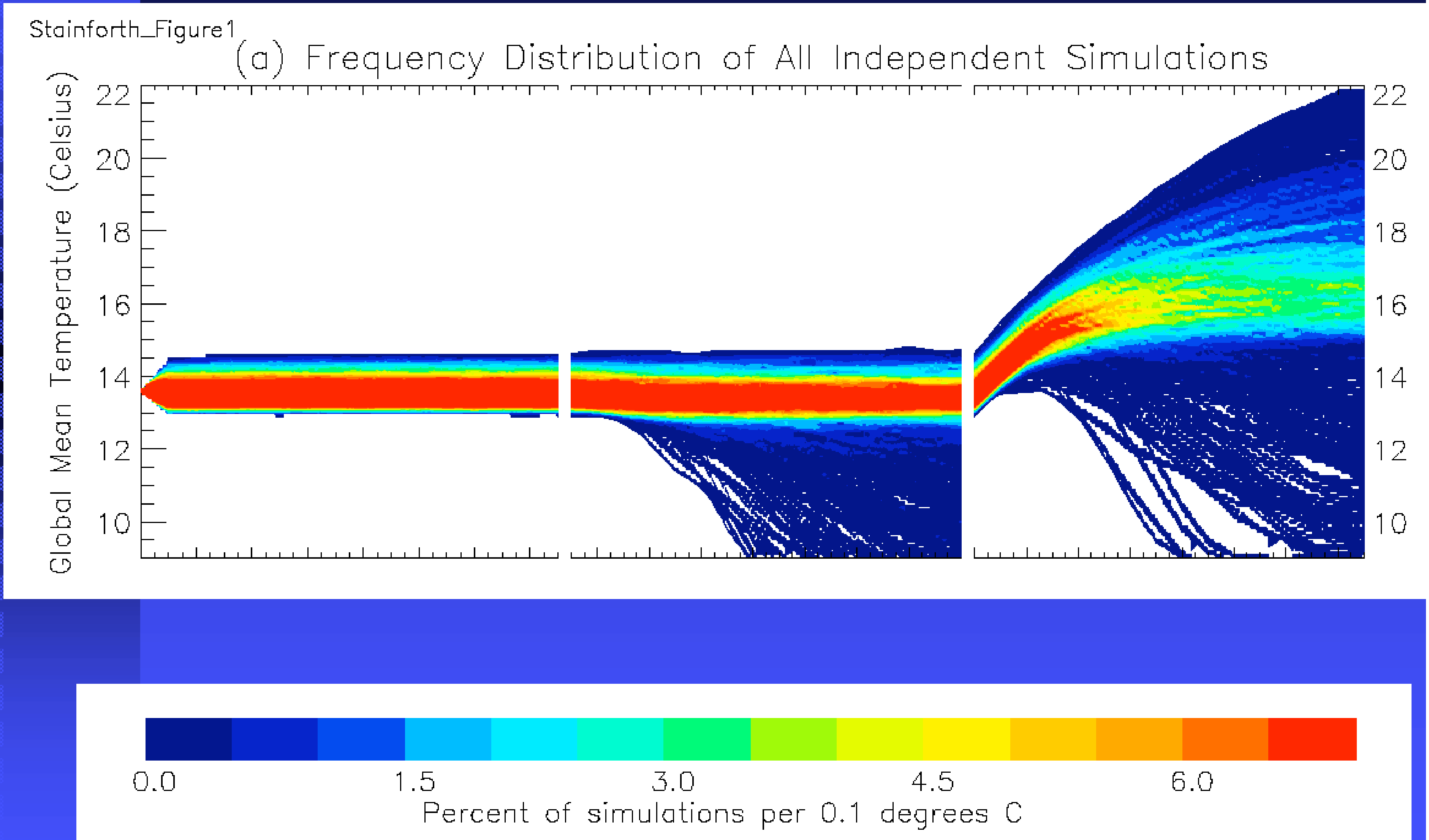
To sample parameter values, however, the input distribution determines the output distribution:

Are all these parameter (and heat flux) values realistic?

Do they yield "state-of-the-art" climates?

Details of the input distribution determine general shape of the sensitivity distribution!

And it is not clear how to begin sampling uncertainty in the model structure!


Stainforth_Figure1
(a) Frequency Distribution of All Independent Simulations
Global Mean Temperature (Celsius)
22
20
18
16
14
12
10
0.0
1.5
3.0
4.5
6.0
Percent of simulations per 0.1 degrees C

Climate is defined as a distribution of weather states; we must sample initial states in order to describe this distribution and to obtain statistically meaningful results on instabilities.

Model dynamics reduces the impact of the particular initial conditions

To sample parameter values, however, the input distribution determines the output distribution:

Are all these parameter (and heat flux) values realistic?

Do they yield "state-of-the-art" climates?

Details of the input distribution determine general shape of the sensitivity distribution!

And it is not clear how to begin sampling uncertainty in the model structure!

Climate is defined as a distribution of weather states; we must sample initial states in order to describe this distribution and to obtain statistically meaningful results on instabilities.

Model dynamics reduces the impact of the particular initial conditions

To sample parameter values, however, the input distribution determines the output distribution:

Are all these parameter (and heat flux) values realistic?

Do they yield "state-of-the-art" climates?

Details of the input distribution determine general shape of the sensitivity distribution!

And it is not clear how to begin sampling uncertainty in the model structure!

Climate is defined as a distribution of weather states; we must sample initial states in order to describe this distribution and to obtain statistically meaningful results on instabilities.

Model dynamics reduces the impact of the particular initial conditions

To sample parameter values, however, the input distribution determines the output distribution:

Are all these parameter (and heat flux) values realistic?

Do they yield “state-of-the-art” climates?

Details of the input distribution determine general shape of the sensitivity distribution!

Questions:

What are we trying to do? Exactly?
What is model inadequacy? What is "uncertainty in the initial condition"?
How to compare/combine simulations? (How do you know an improved model is better?)
How should we judge forecasts? (PDF relevant skill scores? Not RMS!)
- include the verification as on EPS member: ?what is the impact on the skill score?
- beyond a better best member (but without rejecting a perfect model!)
- and go beyond C/L (Questions of 'how much' not 'whether')

When to treat simulations as scenarios (vs product space approach)?
How to forget 'best' ? (and accept/identify a move towards better)
How to let the simulations speak for themselves?
Issues of Statistical Good Practice (the real dark side) [avoid being misled in 10^7 D]
Out-of-sample, bootstrapped significance, fair counting, Imperfect model, nonlinearity ...

Things to distinguish:

Simulations from Forecasts [Forecasts can evolve after the simulations are fixed]
Probability Forecasts from Ensemble Prediction Systems [Deterministic from Unequivocal]
'Useful' spread from 'Bad spread' from 'model-optimal' spread [The path from the goal]
High Impact Forecasts from Severe WeatherForecasts
Model Variables from Physical Variables (Projection Operator P)
Empirical Adequacy vs Internal Model Consistency (Z500)
Improving tomorrow's Forecast from improving 'the' 2020 Simulation
Goal of simulations (shadowing) from that of forecasts (information on an observable)
Simulations as Scenarios from Product Space Approaches
Accurate Forecasts from Useful Forecasts (esp risk adverse users) [both sci and psych]

Data Assimilation

Modern data assimilation techniques force the model to take on aspects of the data which the model cannot support.

Variational approaches *assume* shadowing trajectories exist.

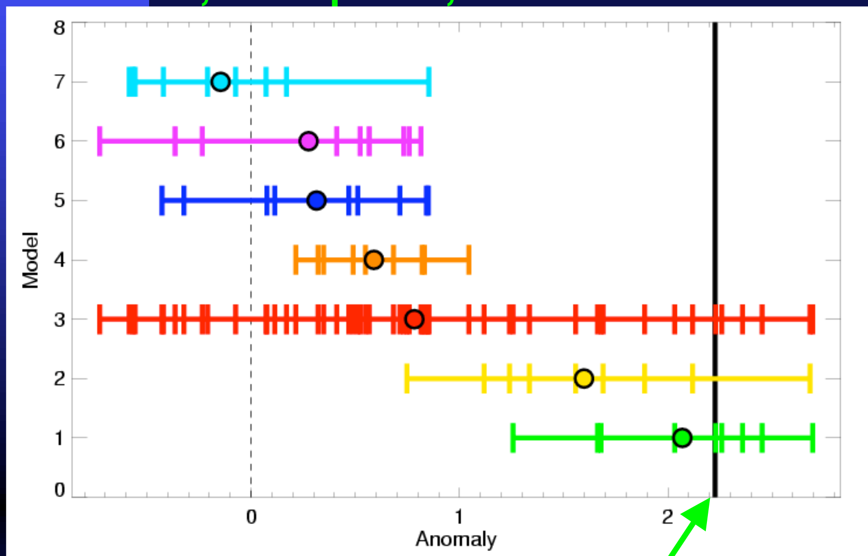
To the extent that our models differ from the system being modelled, they may yield better forecasts if allowed to evolve on their own manifolds…

Forecasting then relies on translating an ensemble of model(s) simulations into forecasts, not treating each as a scenario!

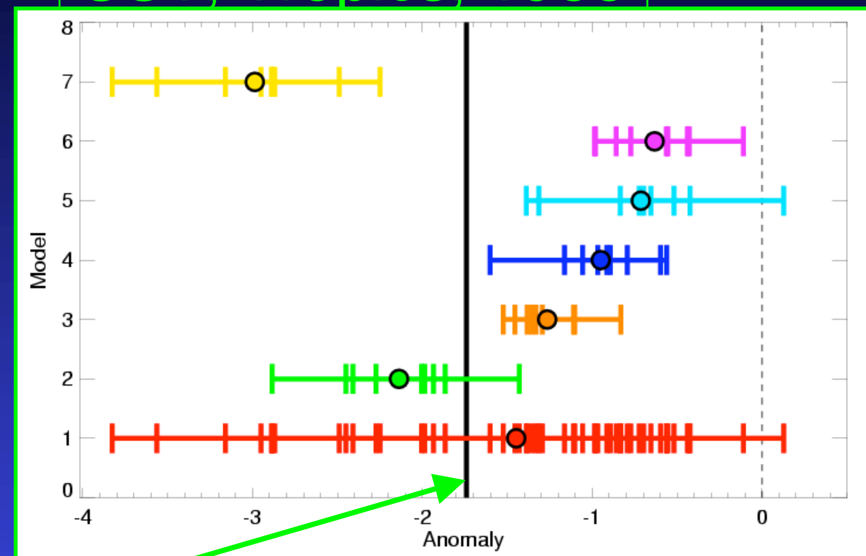
This avoids weaknesses of the Bayesian approach, and Borg-like aspects of 4DVAR.

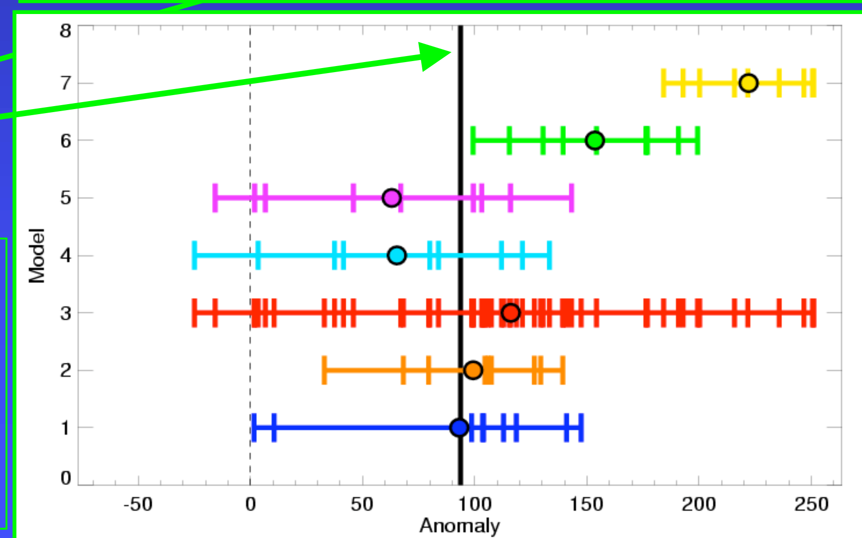
Seasonal Forecasts and DEMETER

SST, Tropics, 1987



SST, Tropics, 1988



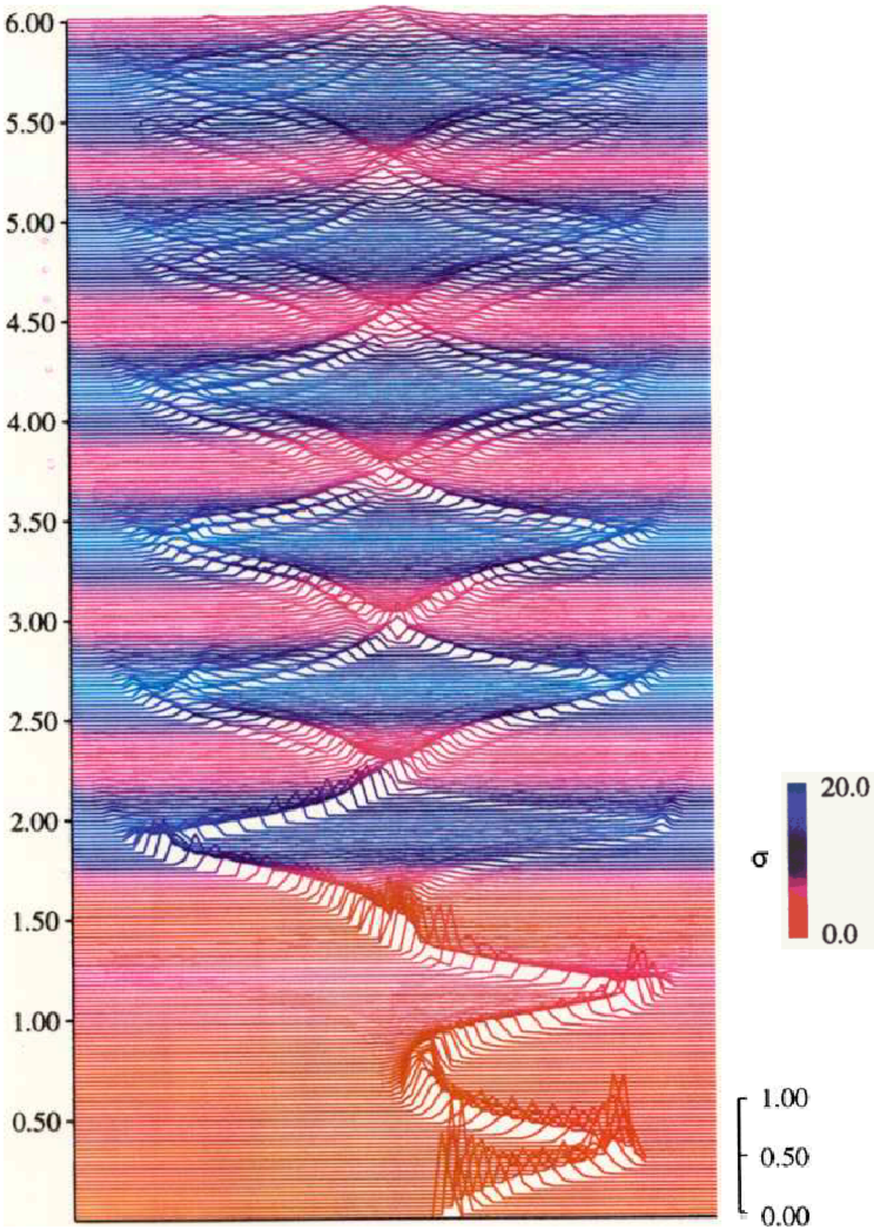


verification

DEMETER Golf balls:

Multi-model (ps-operational)

multi-IC Seasonal forecasts



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

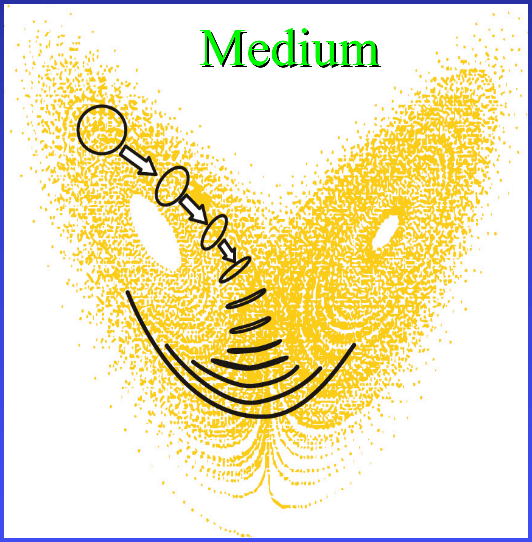
Even with a perfect deterministic model, *the future* is, at best, a probability density function.

And RMS forecast error is at best irrelevant. (McSharry & Smith, PRL, 1999)

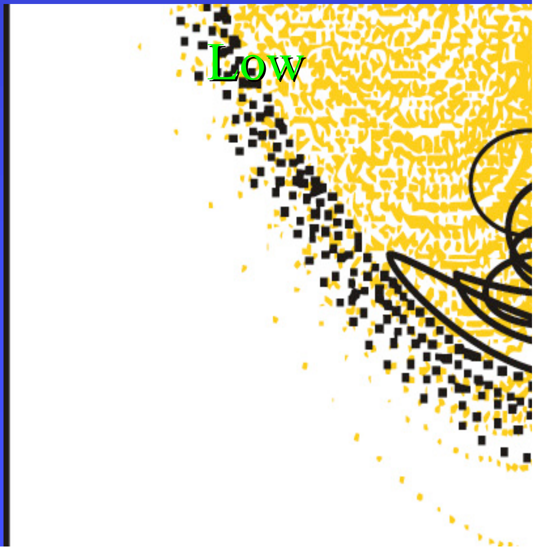
What skill scores should we be using?

Proper? Local?
(Ignorance? Good, 1952; Roulston & S 2003)

Predictability



Pictures from Tim Palmer



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

The empirically relevant questions are (include):

Is it better to pull A_0 away from the model manifold or to project S_n back into the obs space?

For other physical systems, taking initial conditions on the model manifold seems to be better; for NOGAPS, Judd et al have suggest that this yields no RMS penalty after day two: how would ensembles on/near the model manifold compare to perturbed variational analyses?

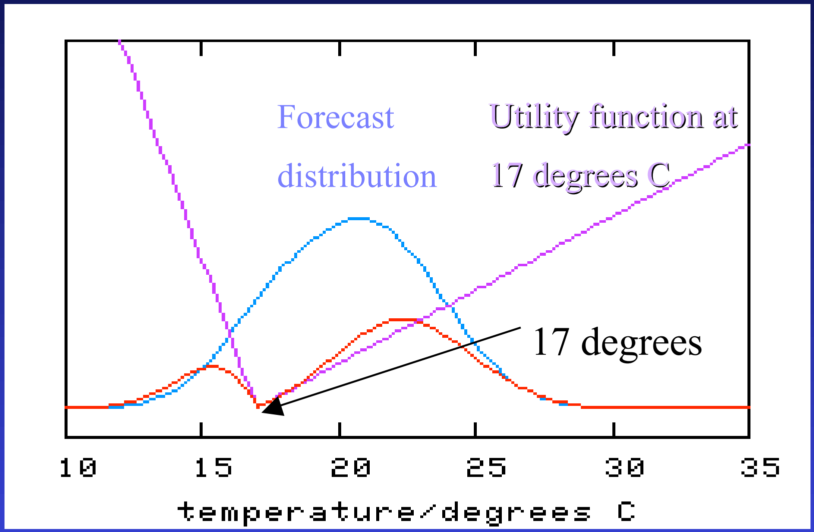
In the multi-model case, we need to look at information content in an empirically meaningful space (obs):
Are better forecasts obtained by interpreting ensemble members as a collection of plausible weather scenarios? or by conditioning on the joint distribution of simulations (multi-IC and multi-model)?

Should we advertise “probability forecasts” or merely “probabilistic forecasts”?

Is it rational to expect probabilities from operational EPS(s)?

Given a forecast and a cost function, we can calculate expected cost of playing every likely temperature. (here, 17 degrees)

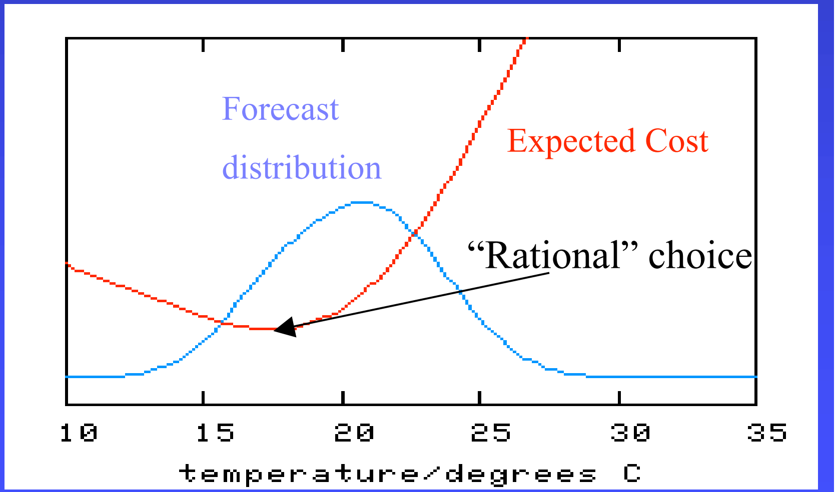
To maximize expected utility, we "should" act on the temperature with smallest expected cost.



For Cal ISO, it proves better to play a empirical quantile: and this is rational, *unless* we insist that the forecast distribution is a probability-DF.

Anyone have a counter-example?

Of course, this can be recast as merely a problem of robust estimation.



The empirically relevant questions are (include):

Is it better to pull A_0 away from the model manifold or to project S_n back into the obs space?

For other physical systems, taking initial conditions on the model manifold seems to be better; for NOGAPS, Judd et al have suggest that this yields no RMS penalty after day two: how would ensembles on/near the model manifold compare to perturbed variational analyses?

In the multi-model case, we need to look at information content in an empirically meaningful space (obs).

Are better forecasts obtained by interpreting ensemble members as a collection of plausible weather scenarios? or by conditioning on the joint distribution of simulations (multi-IC and multi-model)?

How might adaptive obs be taken (based, say, on TIGGE)?

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

K Judd and LA Smith (2001) *Indistinguishable States I* , Physica D **151**: 125-151 *(2004) Indistinguishable States II,* **196:** 224-242 .

Dave Stainforth (2005) *et al.* 27 Jan *Nature*

M. Altalo and LA Smith (2004) Environmental Finance **6** (1) 48-49.

M Roulston *and LA Smith* (2003) MWR **130** (6): 1653-1660.

A Weisheimer, L.A.Smith and K Judd (2004) A New Look at DEMETER forecasts via Bounding Boxes Tellus (to appear).

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

Dressing and Scoring

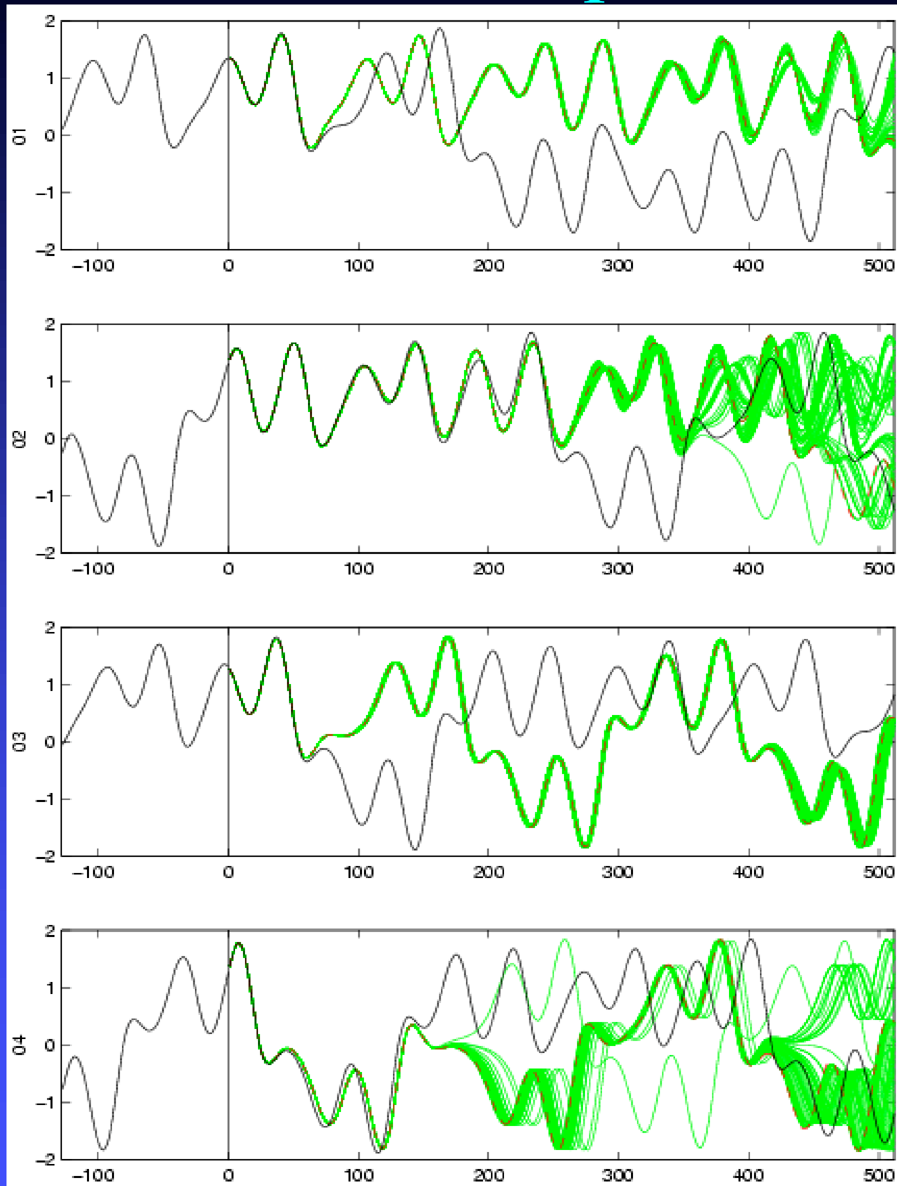
Ensemble forecasts in the model-state space can be *interpreted* as probability forecasts in the target space.

One way to do this is scenario dressing with kernels.

If done: how should these be evaluate? Tuned?

Is conditioning a probability forecast on the joint distribution of a multi-model ensemble feasible?

Recurrent example: "Chaotic" Circuit



Often (in some of the systems I have looked at) model error tends to be episodic and state dependent.

Short term (weather) forecasts are very skilful.

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

Focus on information content, not model-state values…

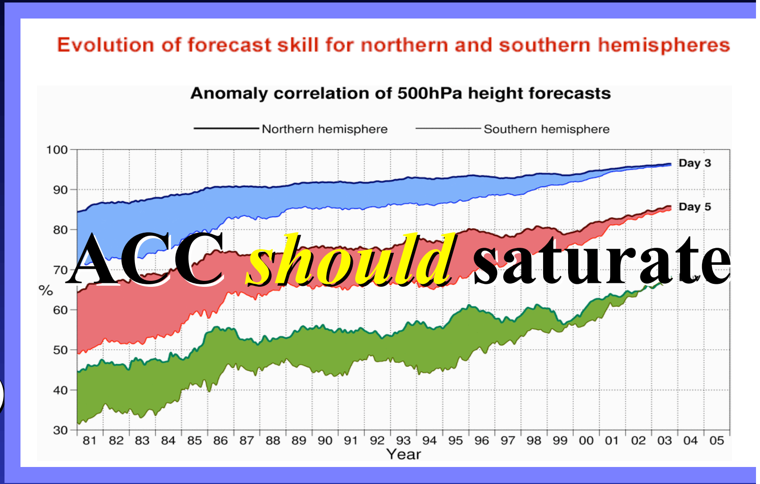
A model can add value as long as it adds information, it need not have traditional "skill."

Should we use kernels?
Or some joint density?

And what about weather/climate?

Nonlinearity couples tuning, DA, dressing… the entire EPS.

Take home point:



From Holingsworth, et al. 2002

What are the implications of:

- Taking Ensembles Seriously?
  TIGGE (mIC,mM,mN, ?mA?):
  - **NO**: RMSE, MAE,…(of what?)
  - "Empirically proper" scores
  - empirical verification (for TIGGE?)

- Taking Model Inadequacy Seriously? *(intrinsic uncertainty)*
  - translating model-simulations into weather-forecasts
    - Information content, not literal interpretation
  - accepting that perfect model studies are misleading
  - preferring initial conditions that "look" worse
- Adaptive observations in the mM, mIC, mN context:
  ? Based on TIGGE?

End-to-End Applications in this Context

Model development

Parameter Estimation

Data Assimilation

(Ensemble) Simulation

Forecasting

Forecast Interpretation

Informed Decision Making

Model(s) Improvement

Nonlinearity Links This End-to-End Chain

Aim merely for internal consistency

Overview of the Options:

Indistinguishable States (Perfect Model)

Probability Forecasts (Perfect Model)

Probability Forecasts (Real World)

Indistinguishable States (Imperfect Model)

Evaluating Probabilistic Forecasts

Lower Dimensional Dynamics (Real World)

Adaptive Obs (Indistinguishable States)

Acting on Probabilistic Forecasts (Real World)

Statistical Good Practice and PoV Studies

Scenarios, Seasonal, Climate and Other Lectures

So (by 2014):

Forecasts may be better *because* the initial conditions look worse.

Forecasts will be improved by better resolving the projection from the simulation(s) to forecast (Moving further from the identity operator toward conditioning on the joint distribution of all simulations).

Probabilistic forecasts may prove more valuable by not providing probabilites.

Data assimilation will produce an ensemble of initial conditions, each “on” the model manifold.

Multi-model adaptive obs will be more straightforward by working in model-state space(s) with large TIGGE-like ensembles.

Eight Current Challenges:

Moving Beyond Scenarios
$P(\boldsymbol{x} \mid \text{obs}; X_1, X_2, X_3; Y_1, Y_2; Z_0)$

Dressing individual ensemble members may be useful, but a better (& more Bayesian) approach would be to condition on the joint distribution of our (imperfect) models.

Projection Operator –or- Ensemble "Bias Removal"

We do not really understand how to map (individual) model states to and from observational space, much less ensembles.

(Coelho et al 10:00 Thursday)

Parameter estimation in nonlinear models

Even with Normal input errors, nonlinearity implies non-normal output errors, complicating not only "state" estimation but also parameter selection.

"Recalibration"

Unlikely in meteorology

von Mises (1928)

Current Challenges:

Limited relevance of the Kalman Filter	"Of course, in general these tasks (prediction, separation, detection) may be done better by nonlinear filters." (Kalman, 1960; first substantial footnote)
Use of 4DVar with imperfect model(s)	The target is no longer a max likelihood state, in fact the model may not support the most "realistic" looking states.
Ensemble "spread" and "bias" correction.	Distinguishing "good spread" and "bad spread" given ~ 100 points in a ~10,000,000-dim space.
Interpreting parametric uncertainty in the "one-off" case (climate).	What are "reasonable" parameter ranges? How climate variables differ from weather? Can a prior distribution and a transfer function yield a policy relevant PDF?

Applications in this Context

Model development	resource distribution for utility (not for naïve realism)
Parameter Estimation	relaxed (to within the physical relevance of then parameterisation)
Data Assimilation	allow each model its manifold, assimilate without re-simulating!
(Ensemble) Simulation	perturb as far in the past as possible: do NOT resample
Forecasting:	true eMOS
Informed Decision Making	a PDF, but not as we know it
Model(s) Improvement	evaluation & forecast archive

Aim first for mere internal consistency?

Aims

- Exploitation and Demonstration of Value
- Ensembles and Applied Probability Forecasting
  - Wind Farm Production
  - Significant Wave Height
  - FPSO Heave
- Better than Probability Forecasting?
  - Beyond the Shotgun
  - Challenges of imperfect models
  - Distribution Forecasts and Electricity Demand
- Conclusions/Transition

(Kevin with the maths…)

Our models are imperfect, we can see through them more clearly if we accept this when initialising, applying, evaluating, selling & improving them.

(Not only does this lead to better science, it is better business.)

METRO Café

the commuters' favourite

Today's The Day
Weird World
Have Your Say
Competitions
60 Sec Interview
Web Watch
Games & Puzzles
Metro Mobile
Metro Events
Metro Courses
Metro Total Golf
METRO TUNES
METRO getaways
LOG IN
REGISTER WITH US

TODAY'S THE DAY

January 27, 2005

back

CRYPTIC CROSSWORD
Today's crossword
Today's print version
collection of crosswords
archive GO

ON THIS DAY IN 1953

A PLEA FROM 200 MPs was rejected and Derek Bentley was condemned to hang. Bentley, 19, and Christopher Craig, 16, were found guilty of murdering a policeman during a burglary in South London. Craig fired the shots but was too young to be hanged. The prosecution claimed Bentley was responsible as he

What are you doing to save environment?

Have your say on our messageboards

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.

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.

Travel archive

MAKE YOUR MARK

YOU DECIDE
Should the UK sign up to the new European Union constitution?

Have your say
Yes
No
Not sure

want more of a say? GO

Metro Message Boards
If you've got something you want to ask or if you just want a chat, the Metro Message Board is the place for you. Drop by and say hello.

DINNER AT A DASH

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.

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

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

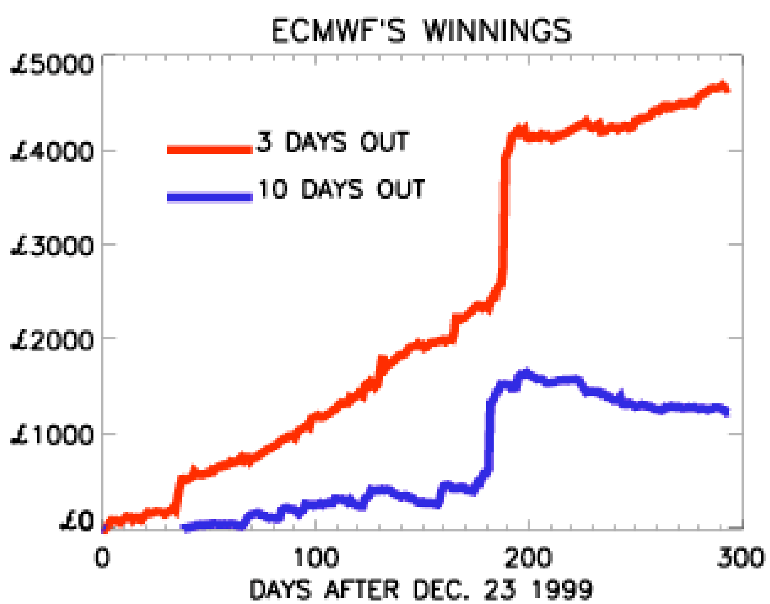
How would a probability forecast based on the ECMWF EPS fare against a house that set its odds using climatology?

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

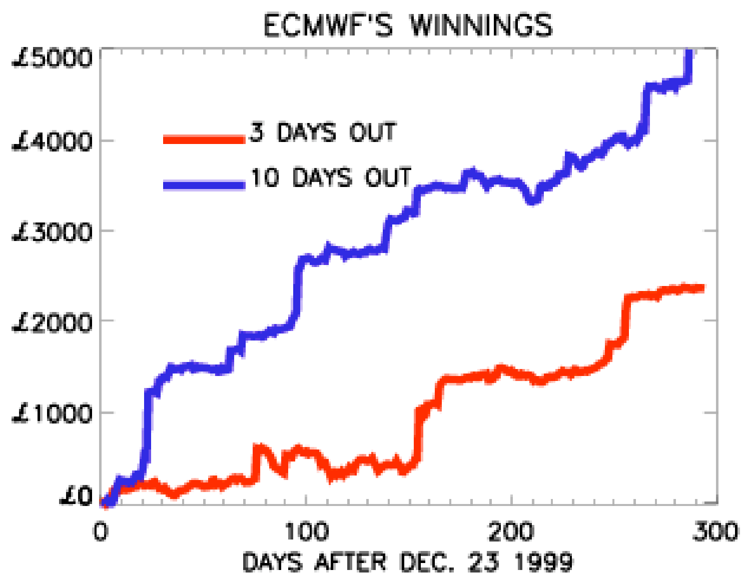
WEATHER ROULETTE

TEMPERATURE AT HEATHROW

TABLE MAXIMUM: £100

ODDS SET BY HIGH RES. FORECAST

BETS PLACED ACCORDING TO ENSEMBLE



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

Dressing allows a fair comparison of EPS and BFG.
How can we measure this kind of skill?

Beyond Scenario Dressing

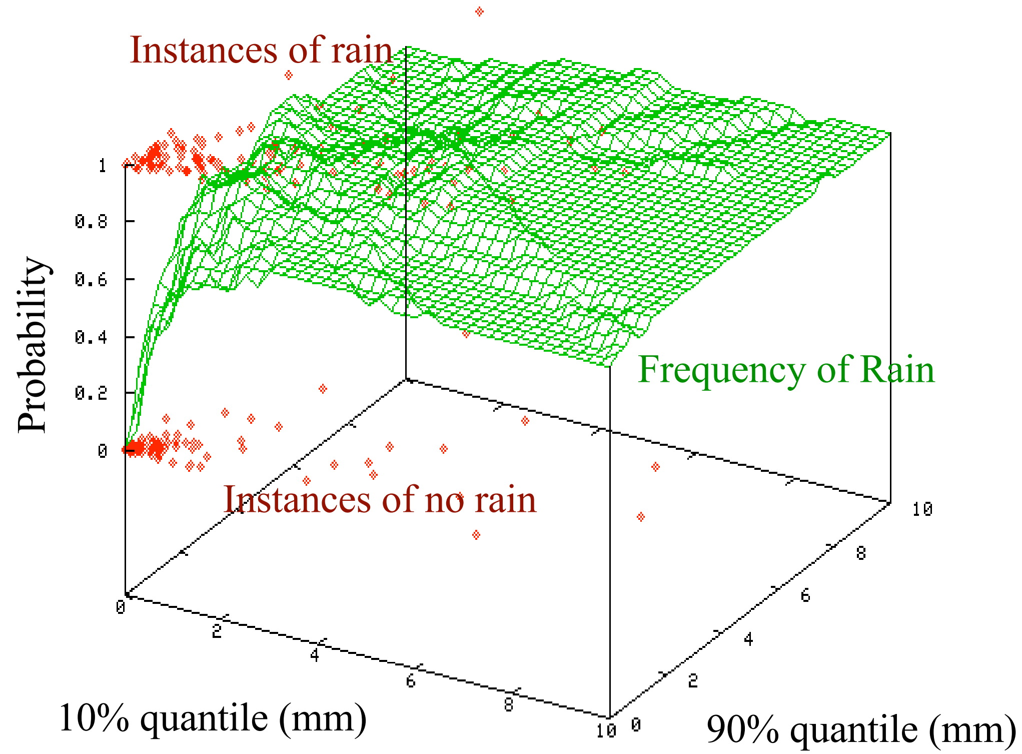
Beyond Weather I: Seasonal   (DEMETER)

Beyond Weather II: Climate   (climateprediciton.net)

Precipitation at Schleswig

We don't have to interpret model-rain *as* a scenario for rain in each ensemble member...

Better to use the joint distribution with the aim of extracting information.



Schematic

i- shadows

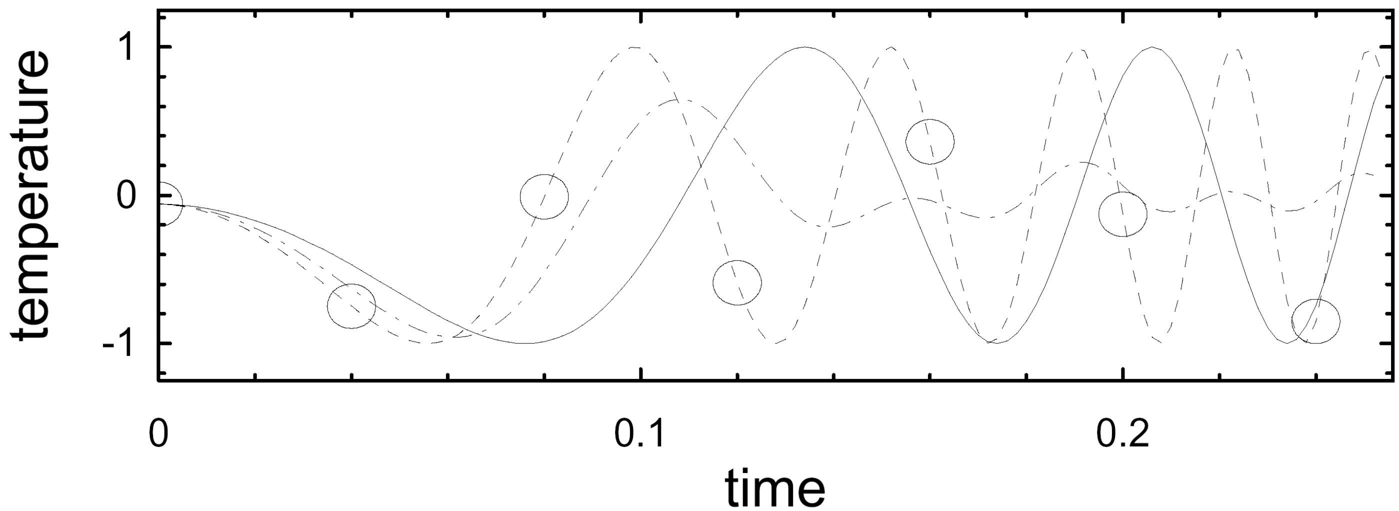


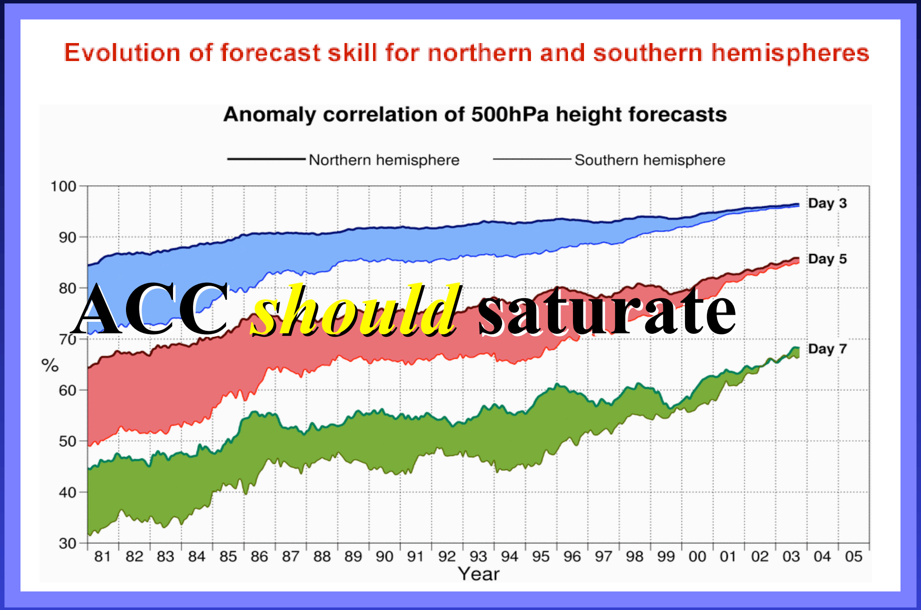
Figure 6. The shadowing dilemma. Given a series of uncertain observations, indicated by the circles, the dot-dashed line indicates the forecast with minimizes the RMS error. It clearly out performs the more realistic model started from the initial observation (the solid line). What we want to determine is whether there exists another initial condition (*e.g.* the dashed line), consistent with the initial observation, for which the trajectory of the realistic model passes within the uncertainty radius of a series of observations.

(Smith 1995)

Note that small changes in the initial condition of a model with realistic trajectories may yield a shadowing trajectories; small changes in the initial condition of the optimal RMS model will not.

If we take ensembles seriously how am I to interpret these distributions (even if our model was perfect):

The selection of skills score is very very important!

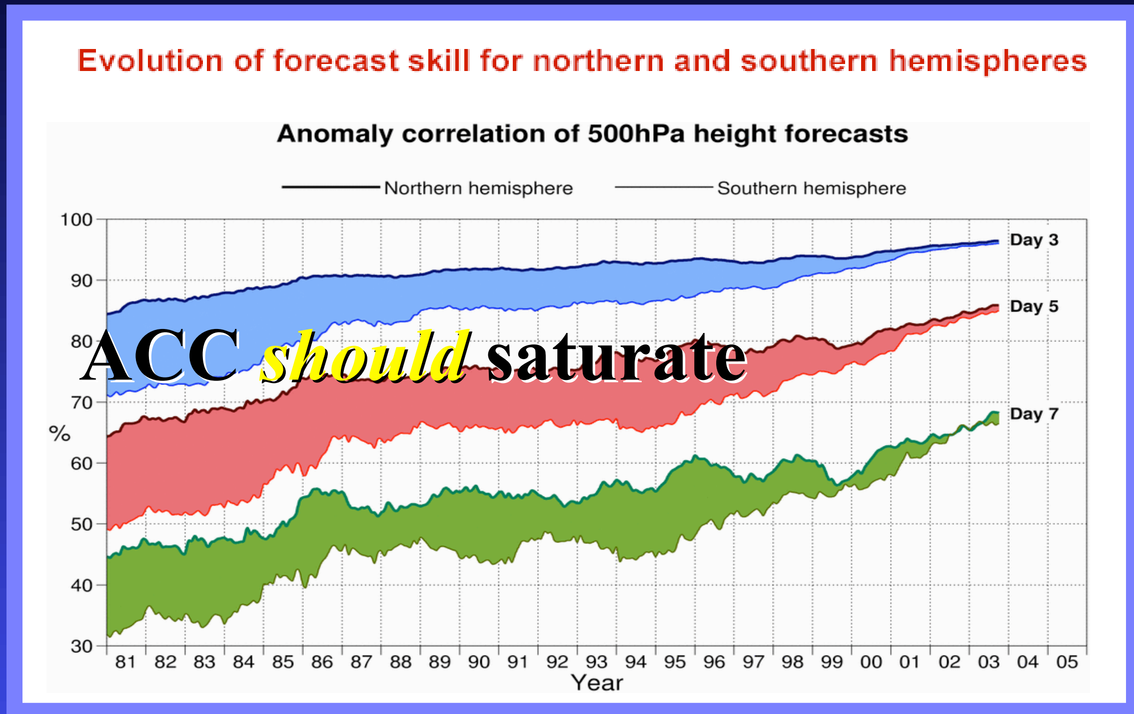


From Holingsworth, et al. 2002

And if I take Model Inadequacy seriously?

- translating model-simulations into weather-forecasts
  - Information content, not literal interpretation
- accepting that perfect model studies are misleading
- preferring initial conditions that "look" worse
- IS and Adaptive observations in the mM, mIC, mN context

If we take ensembles seriously how am I to interpret these distributions (even if our model was perfect):

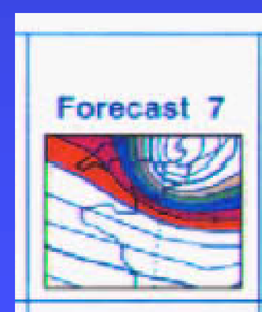


From Holingsworth, et al. 2002

And if I take Model Inadequacy seriously?

- translating model-simulations into weather-forecasts
  - Information content, not literal interpretation
- accepting that perfect model studies are misleading
- preferring initial conditions that "look" worse

- IS and Adaptive observations in the mM, mIC, mN context



Taking Model Inadequacy seriously means:

- translating model-simulations into weather-forecasts
  - Information content, not literal interpretation
- accepting that perfect model studies are misleading
- preferring initial conditions that “look” worse

- IS and Adaptive observations in the mM, mIC, mA context
- And figuring out how to use unrealistic state-of-the-art models

Taking Model Inadequacy seriously means:

- translating model-simulations into weather-forecasts
    - Information content, not literal interpretation
- accepting that perfect model studies are misleading
- preferring initial conditions that "look" worse

- IS and Adaptive observations in the mM, mIC, mA context
- And figuring out how to use unrealistic state-of-the-art models

This is no a shortcoming of meteorology, but of all physical simulation modelling!