Real Time Series analysis and modelling

Aid machine learning with dynamical insight

Hailiang Du
Department of Mathematical Sciences, Durham University
hailiang.du@durham.ac.uk
Data Science Institute, London School of Economics and Political Science
h.l.du@lse.ac.uk

All theorems are true, All models are wrong. All data are inaccurate. What are we to do?

The aim of this course is to teach you how to deal with real data, to increase your scepticism regarding reliable modelling in practice, and to expand the tool box you carry to include nonlinear techniques, both deterministic and stochastic with the aid of dynamical insight.
In short: to get you to think before you compute (and perhaps afterwards too.)

Lecture 5

Uncertainty quantification and model discrepancy recognition

Uncertainties

- Initial condition uncertainty
- Parameter uncertainty
- Model structure uncertainty

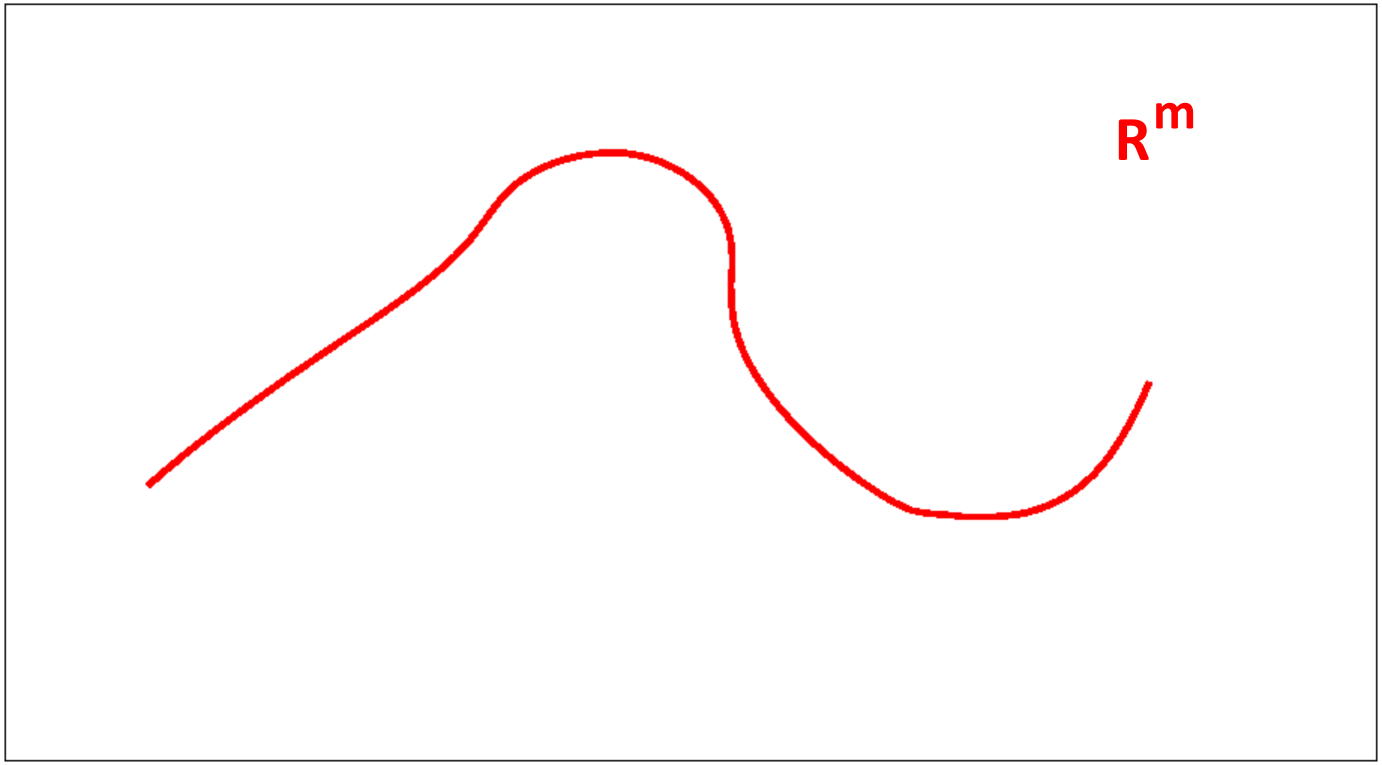
Ensemble predictions

- Initial condition ensemble
- Perturbed parameter ensemble
- Multi-model ensemble…….

why not use pdf directly?

Initial condition ensemble

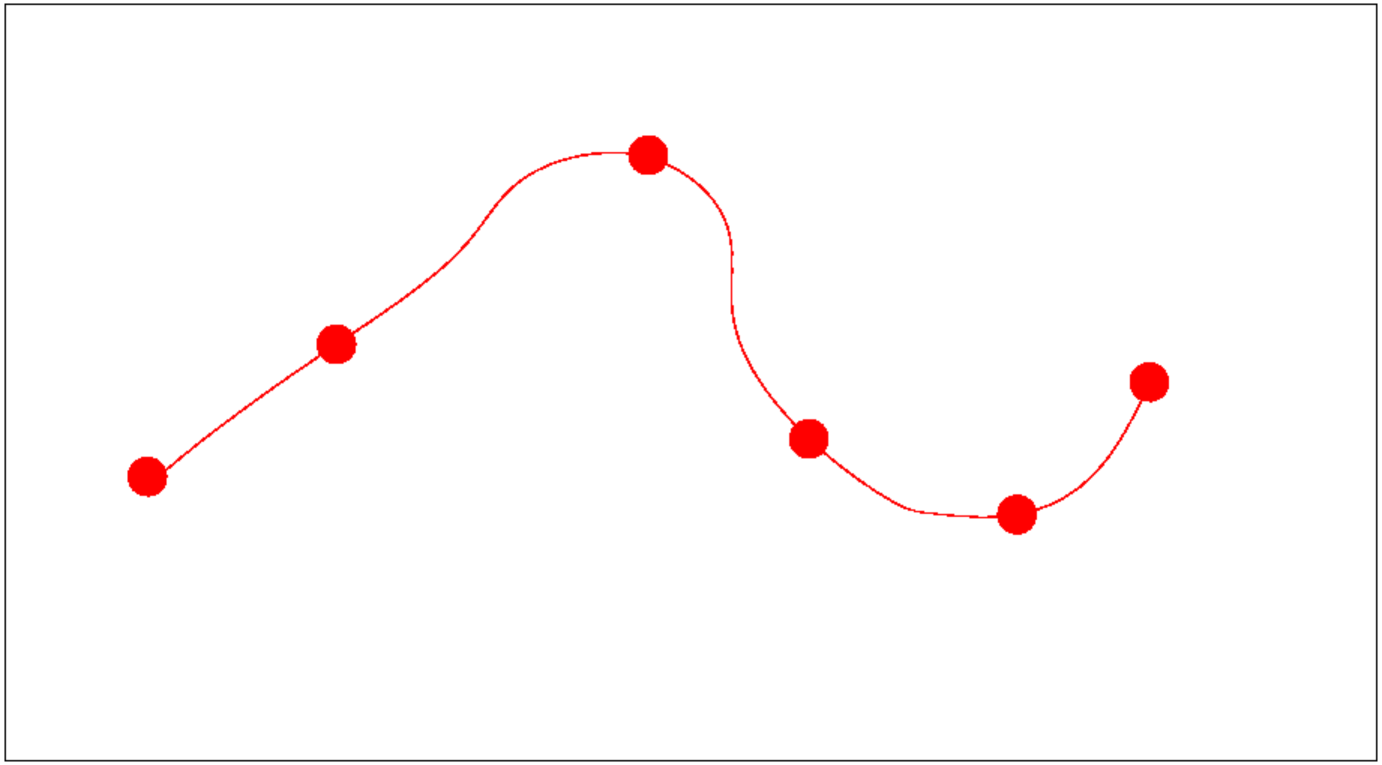
Let's start over, and visualise modelling a dynamical system.



… starting with a segment of the true trajectory.

Initial condition ensemble

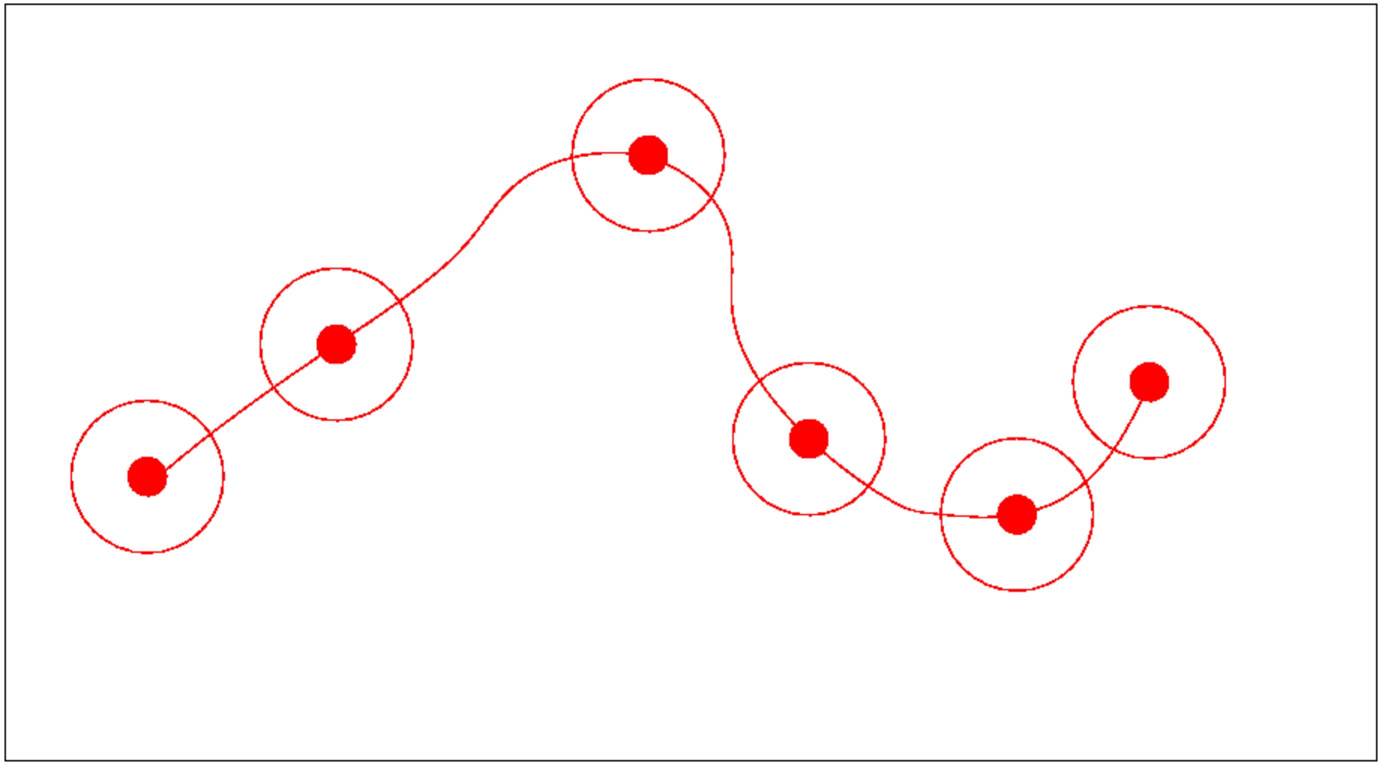
In perfect model case we needs only one state in time.



A segment of the true trajectory, sampled in time.

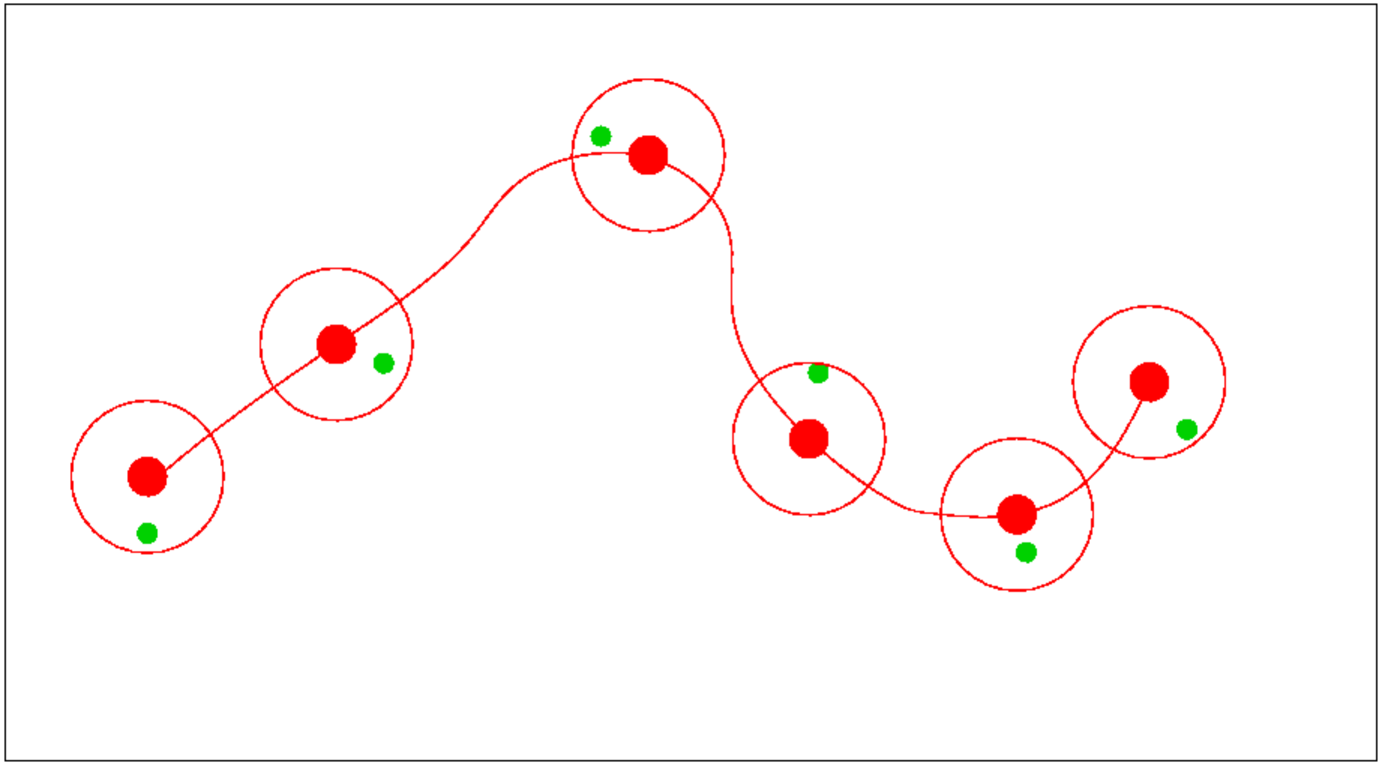
Initial condition ensemble

Unfortunately we have observational noise.



The observational noise yields a PDF for the observations.

Initial condition ensemble

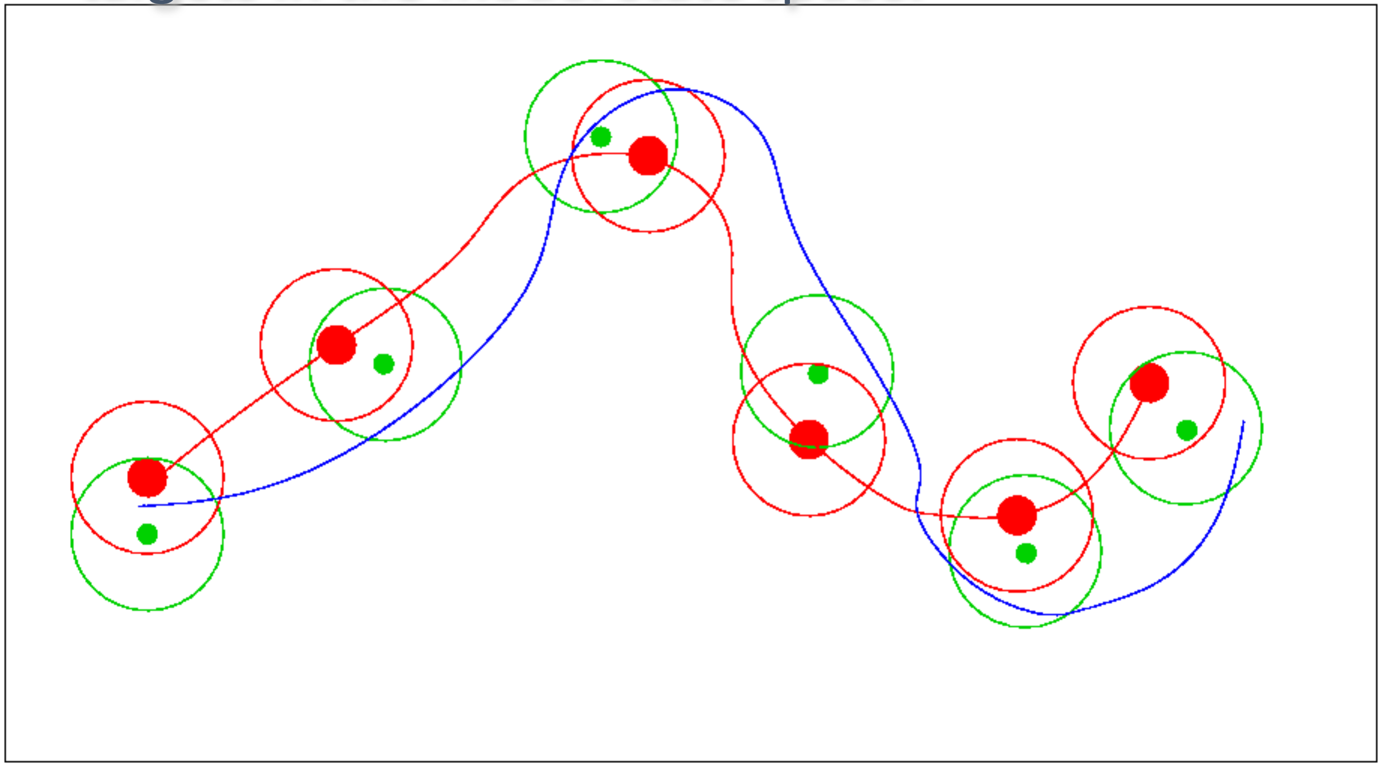


A particular realization of the observational noise.

(This is all we *ever* get.)

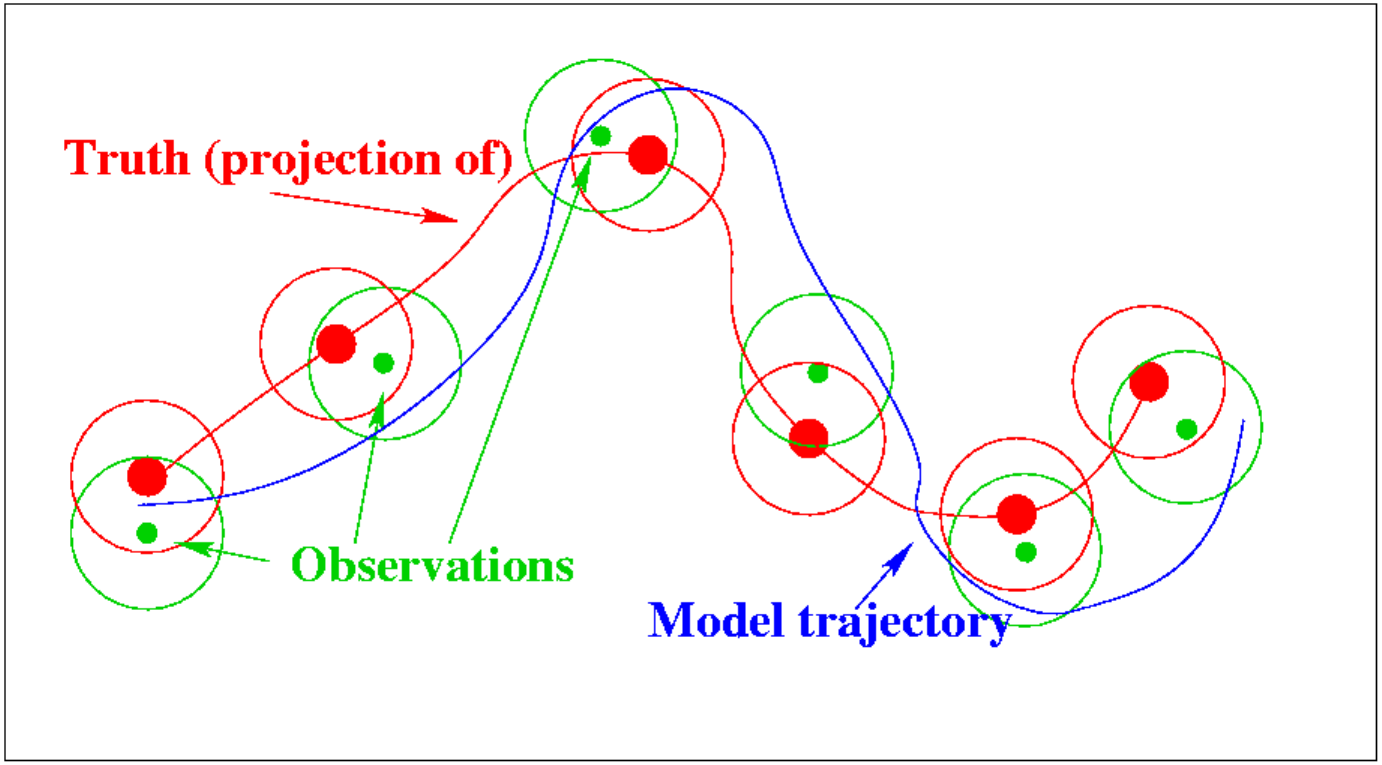
Initial condition ensemble

The observations are inexact, yet they provide targets in the model state space.



And finally, a model trajectory that "shadows" the obs.

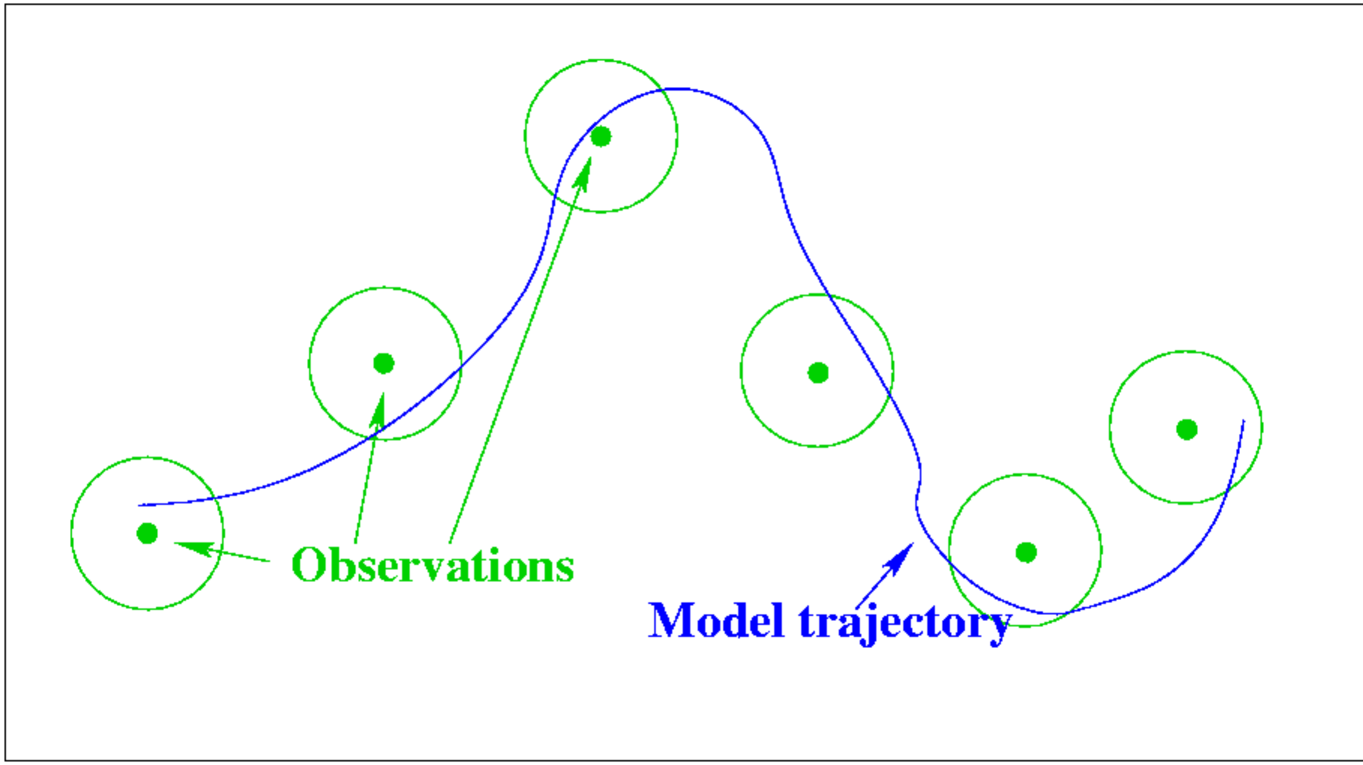
Initial condition ensemble



And truth mere hypothesis.

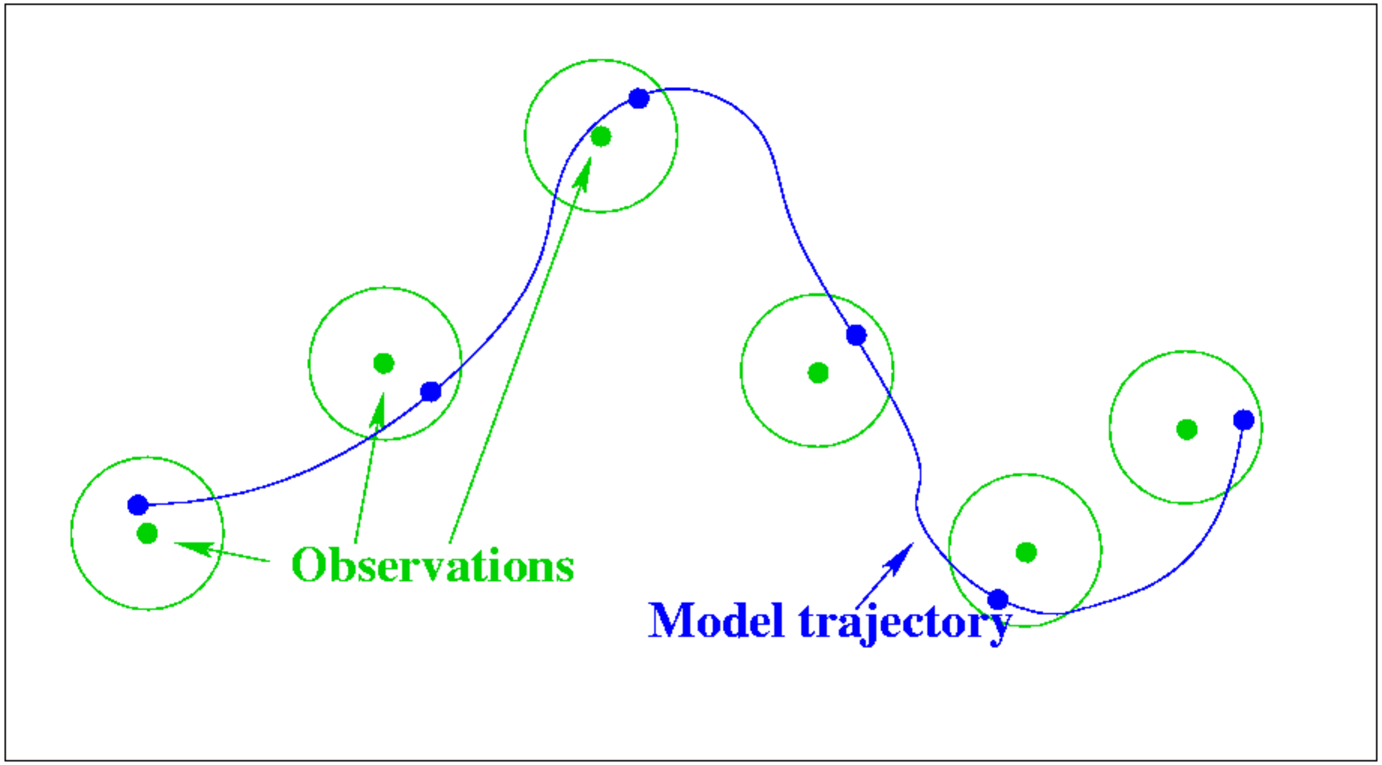
Initial condition ensemble

We have no need of/access too this hypothesis.



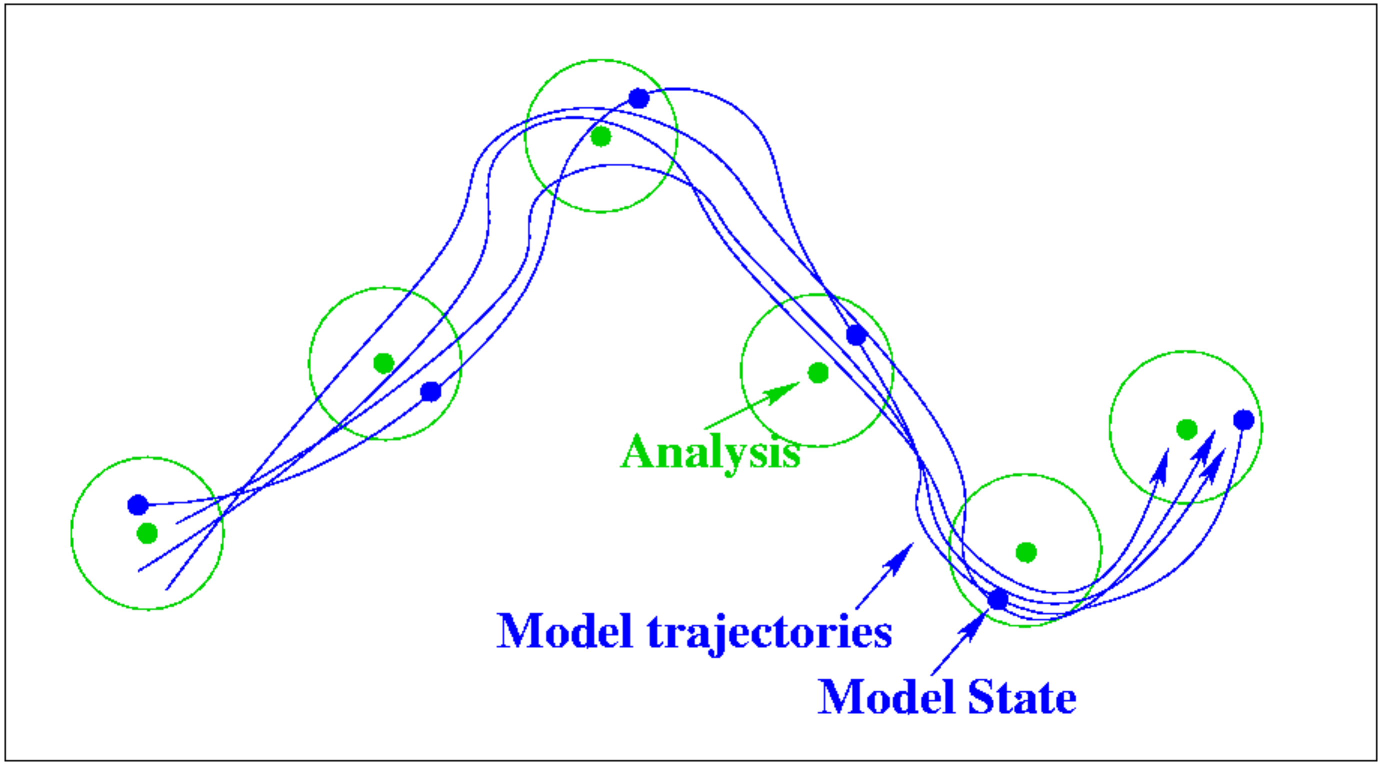
The only dynamical system we have is our model(s); the truth is not out there.

Initial condition ensemble



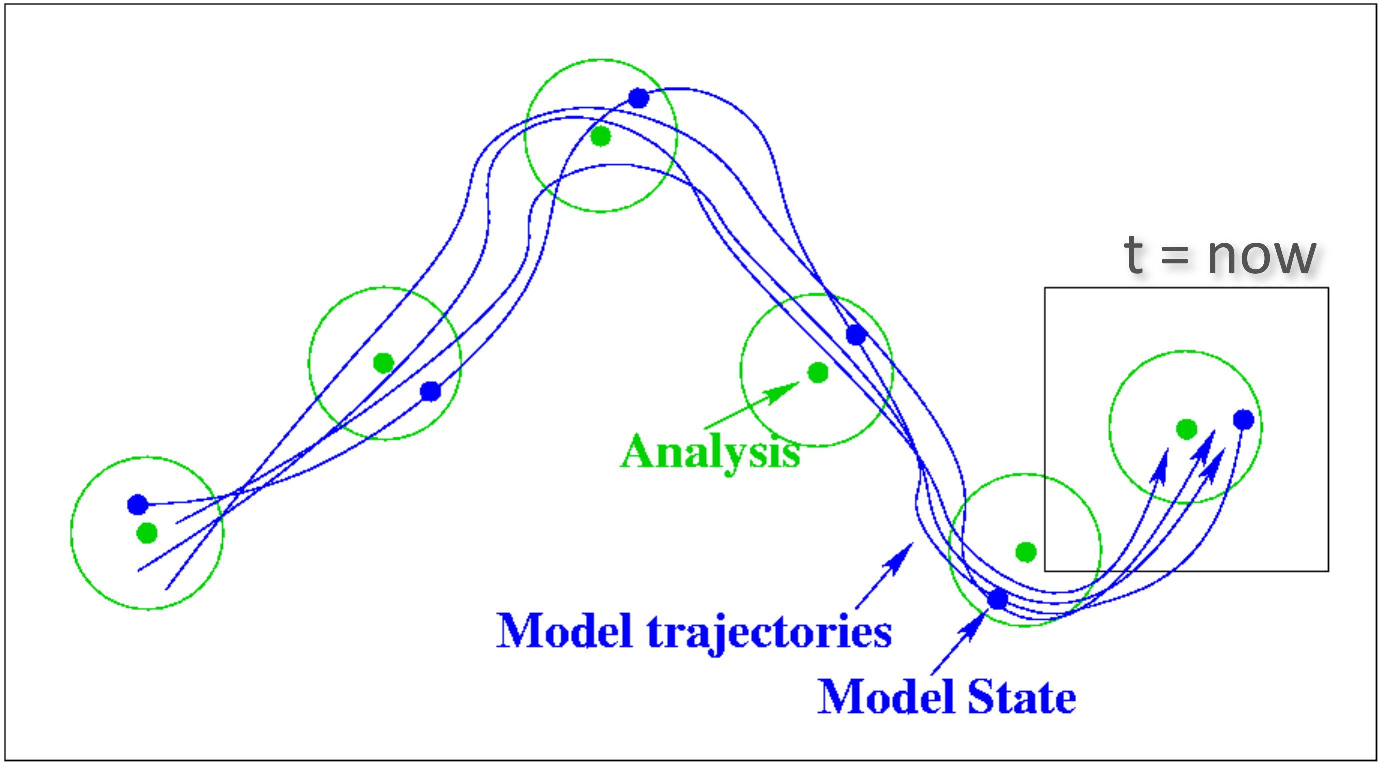
Given the statistics of the "observational noise", the analysis, and the model states, we can still compute the probability of these observations given the "shadowing" trajectory.

Initial condition ensemble



And if one shadowing trajectory exists, there will be many. Of course, if the model is imperfect there may be none at all.

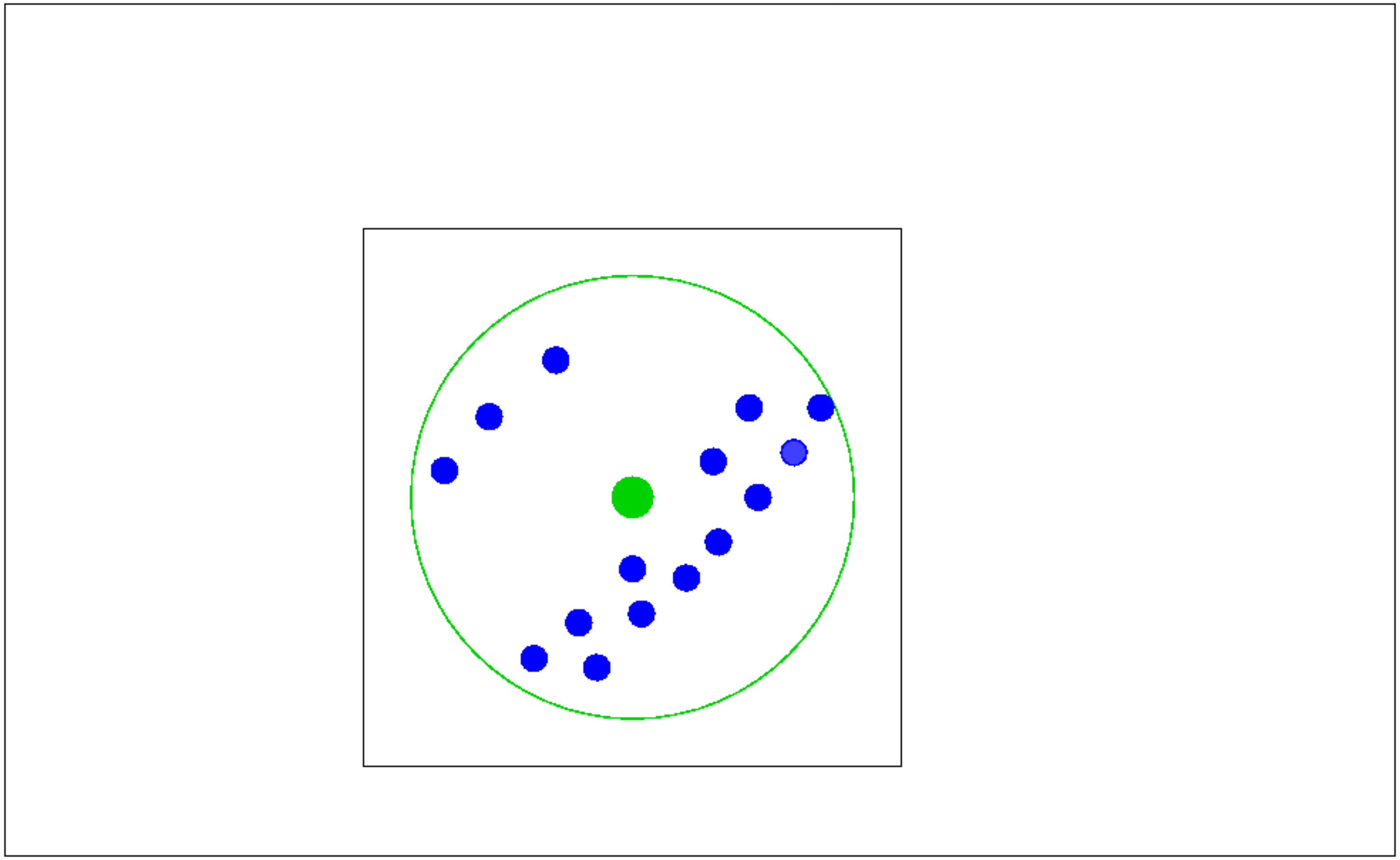
Initial condition ensemble



We can do no better than find a collection of trajectories.

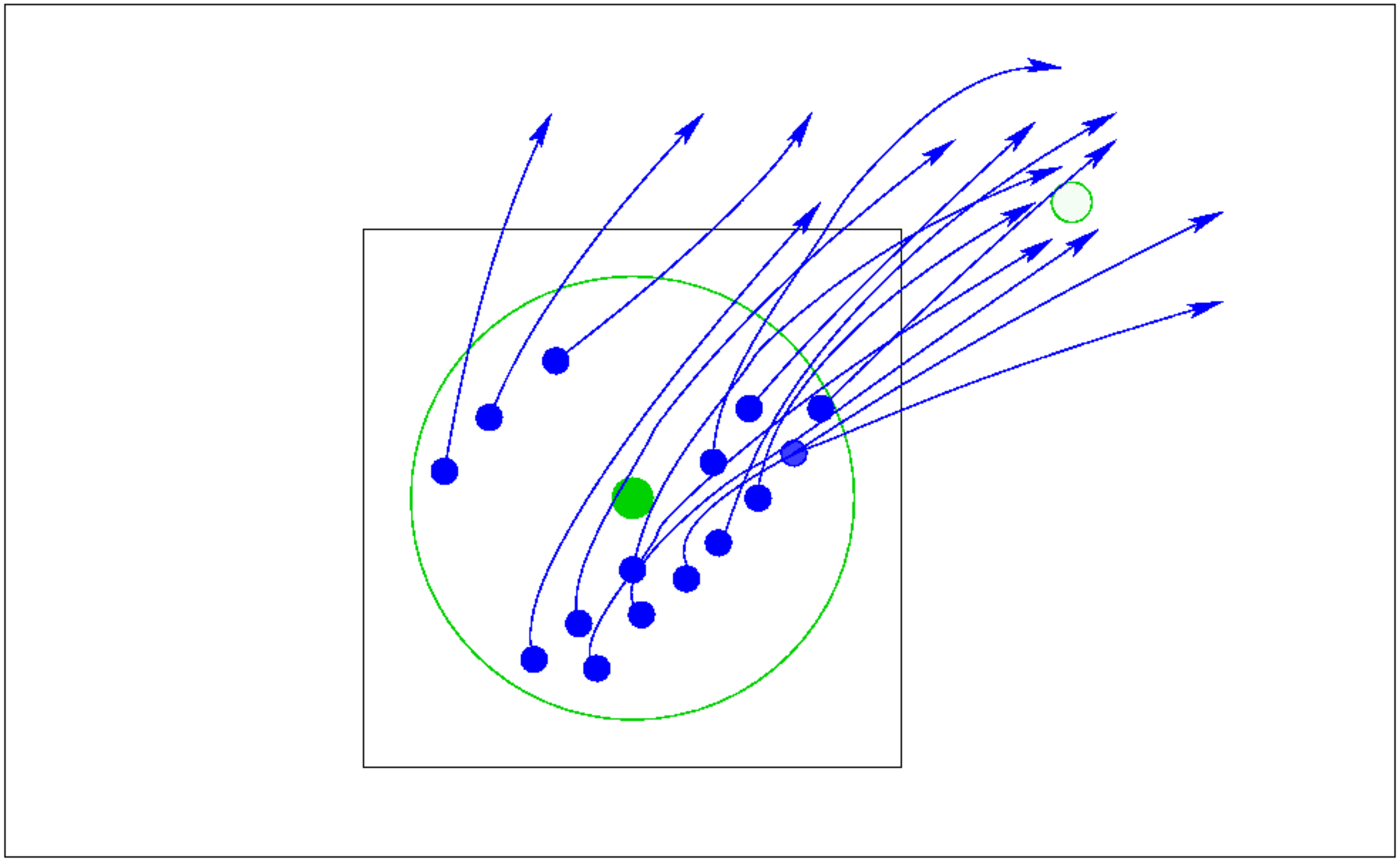
Initial condition ensemble

So an ensemble is **the** coherent goal of state estimation.



Initial condition ensemble

And ensemble forecasts arise naturally.

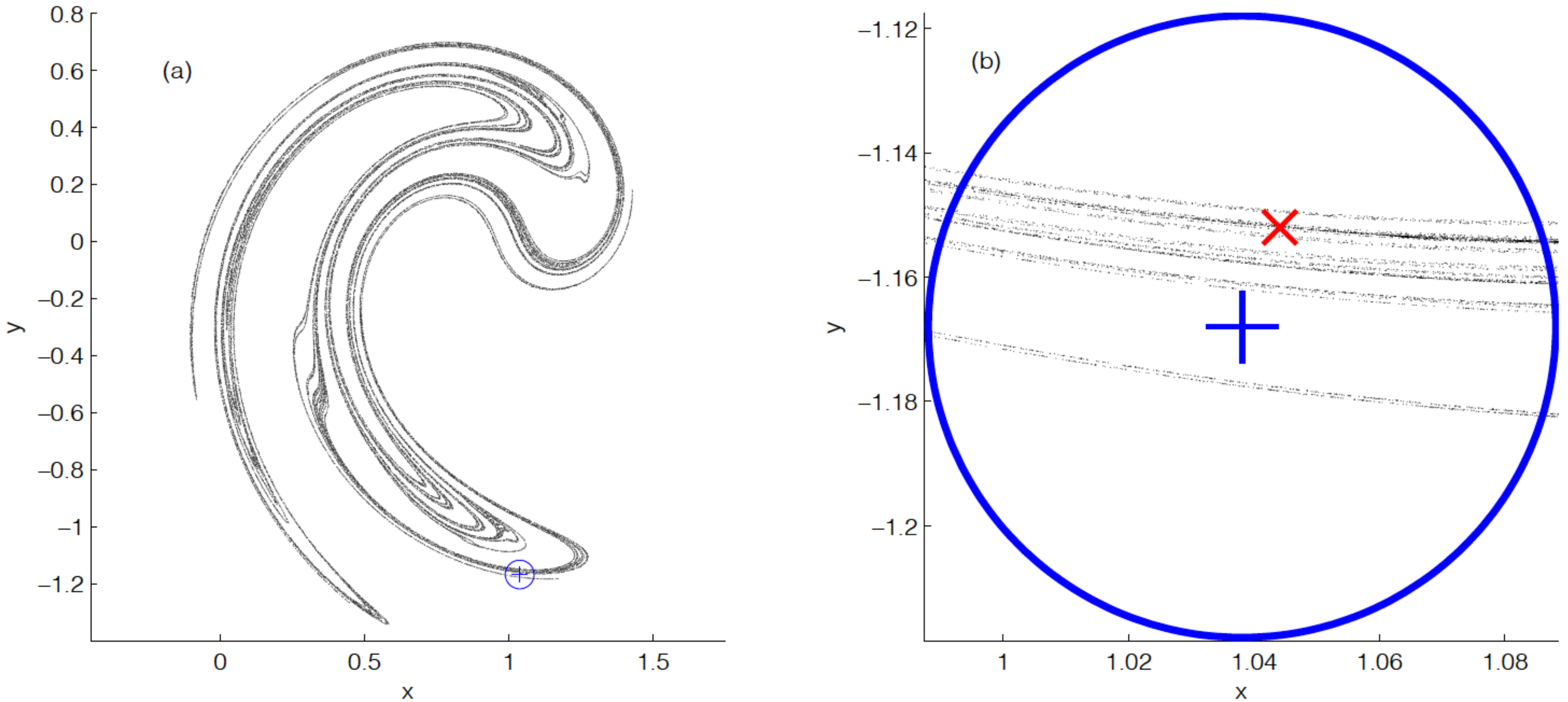


Perfect ensemble

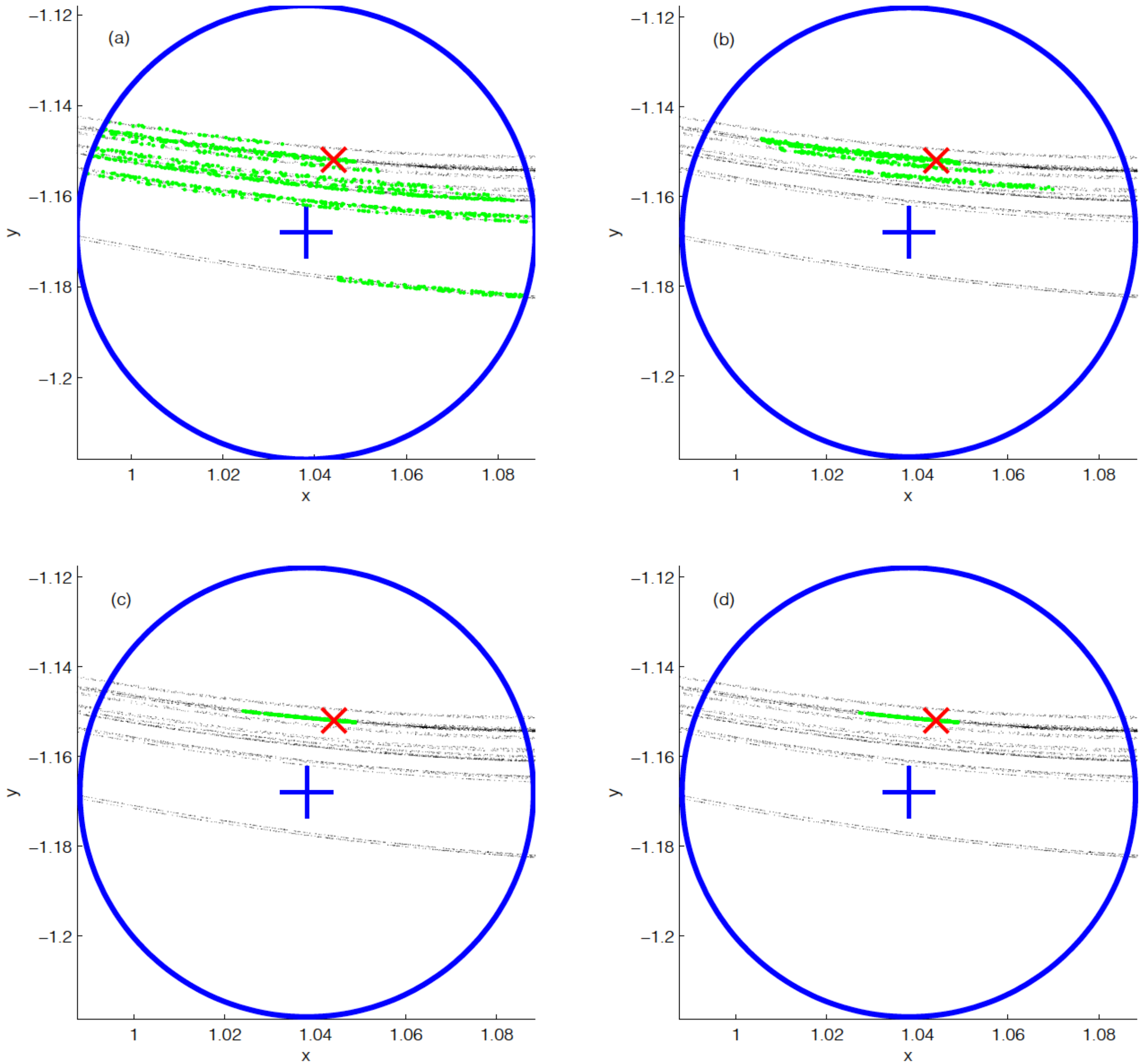
- If we form an ensemble of points both consistent with the observational noise and the model dynamics, we will have a **"perfect ensemble"**

- Each ensemble initial condition must lie on the system's attractor for it to be consistent with both the dynamics and the observational noise.

Perfect ensemble



Example of perfect ensemble for the Ikeda Map when only one observation is considered. The observational noise is uniformly bounded. In panel a, the black dots indicate samples from the Ikeda Map attractor, the blue circle denotes the bounded noise region where the single observation is the centre of the circle. Panel b is the zoom-in plot of the bounded noise region. The red cross denotes the true state of the system.



Examples of perfect ensemble are shown for the Ikeda Map when more than one observation is considered. The perfect ensemble of different number observations are considered are plotted separately. Two observations are considered in panel (a), 4 in panel (b), 6 in panel (c) and 8 in panel (d). In all the panels, the green dots are indicates the members of perfect ensemble.

Ensemble predictions

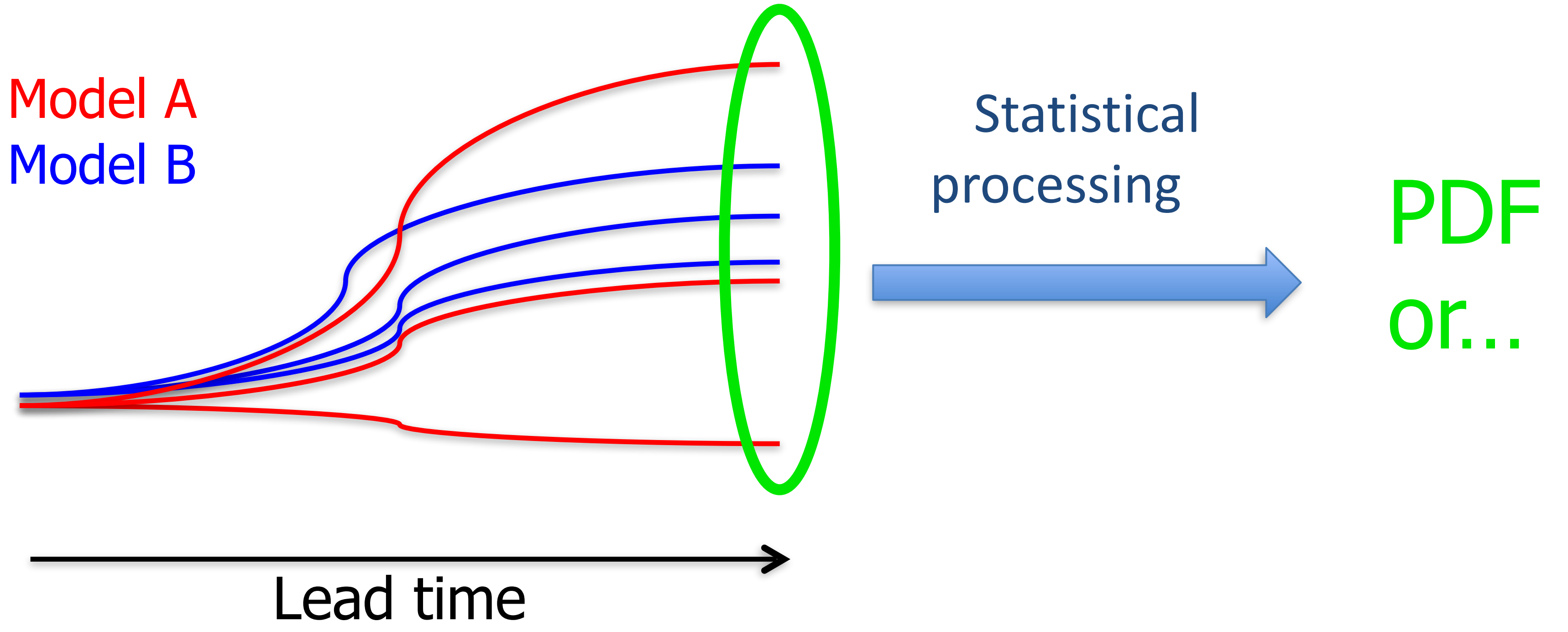
- Initial condition ensemble
- Perturbed parameter ensemble
- Multi-model ensemble…….

Multi-model Ensemble

Statistical approaches are often used to combine ensembles of individual model runs, for example produce PDF based on weighted model output.

Statistical processing

This is a very poor interpretation of models living in high-D



There is no communication between the models; their outputs are only combined at the lead time of interest in the observation space.

Multi-model Cross Pollination异花授粉 in Time (CPT)

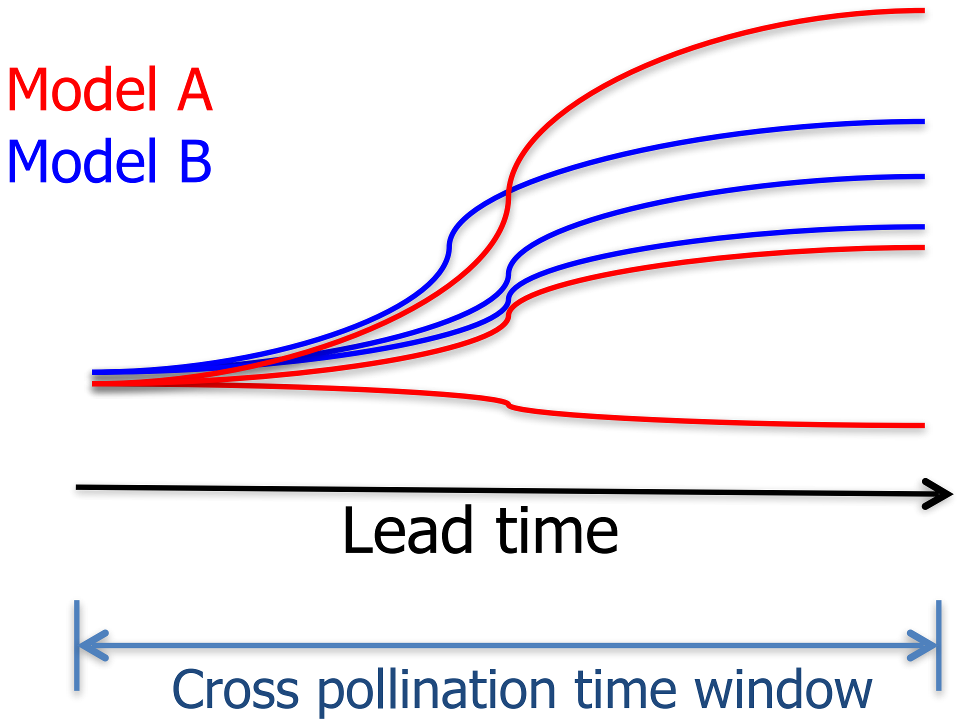
- Statistical approaches are often used to combine ensembles of individual model runs, for example produce PDF based on weighted model output.

- Statistical processing alone can not exploit the richness of multi-model dynamics

- Our aim is to create communication between models in time (between launch and target lead time).

Du, H. and Smith, L.A. (2017) 'Multimodel cross pollination in time', Physica D: Nonlinear Phenomena, Vol. 353-4, pp.31-38.

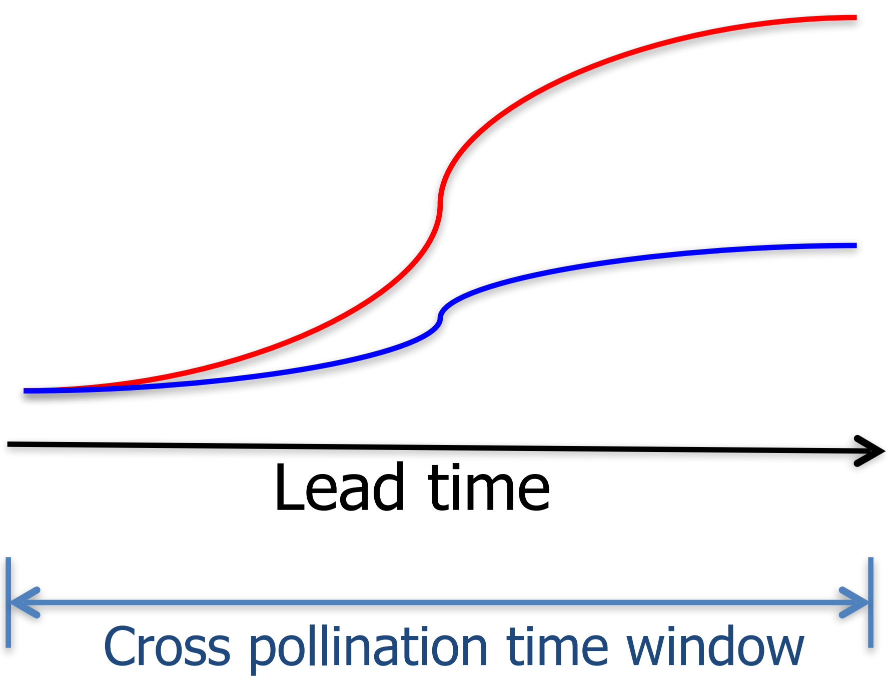
Multi-model Cross Pollination in Time



Note that Model A and Model B usually do not share model space at all, but they could share observation space.

Therefore the communication can be made through the shared observation space.

Multi-model Cross Pollination in Time

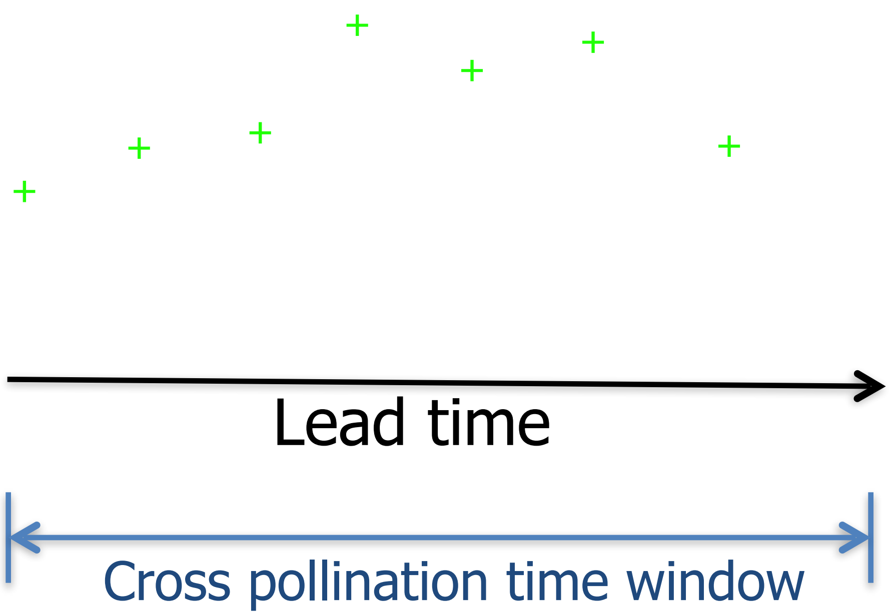


For each model model forecast trajectory, we can sample it (in observation time) to make pseudo observations in the observation space.

We then can mix the pseudo observation from Model A with the pseudo observation from model B create a sequence of pseudo observations.

Multi-model Cross Pollination in Time

This is a very poor interpretation of observations living in high-D

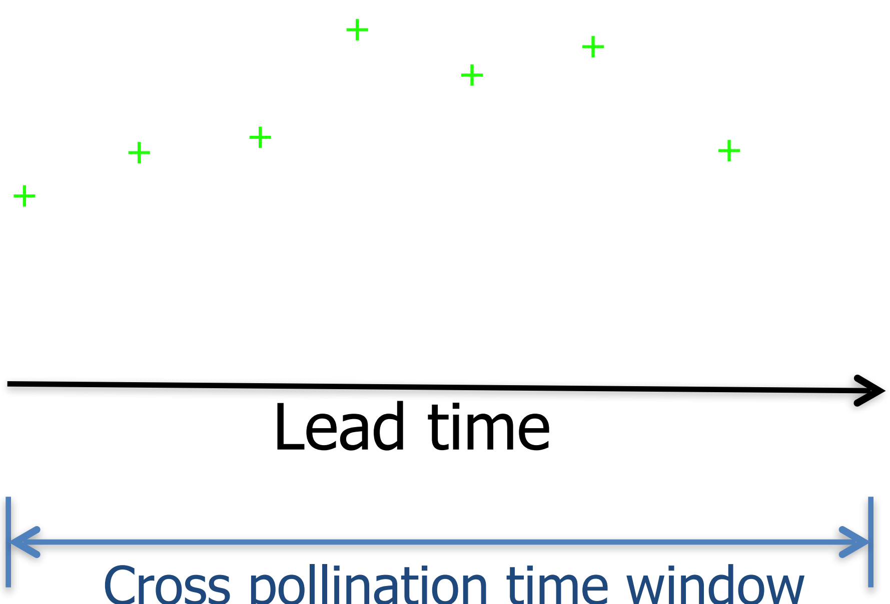


For each model model forecast trajectory, we can sample it (in observation time) to make pseudo observations in the observation space.

We can mix the pseudo observation from Model A with the pseudo observation from model B create a sequence of pseudo observations.

Multi-model Cross Pollination in Time

This is a very poor interpretation of observations living in high-D



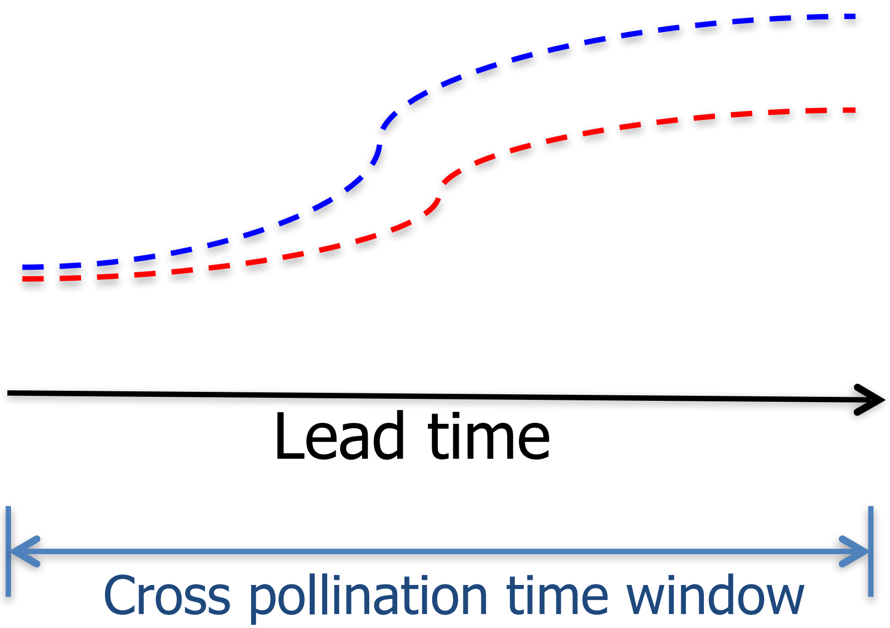
Apply Data Assimilation to the pseudo observations using Model A and Model B respectively.

Of course, Model A and Model B can have very different data assimilation schemes.

Multi-model Cross Pollination in Time

Data assimilation using Model A

Data assimilation using Model B



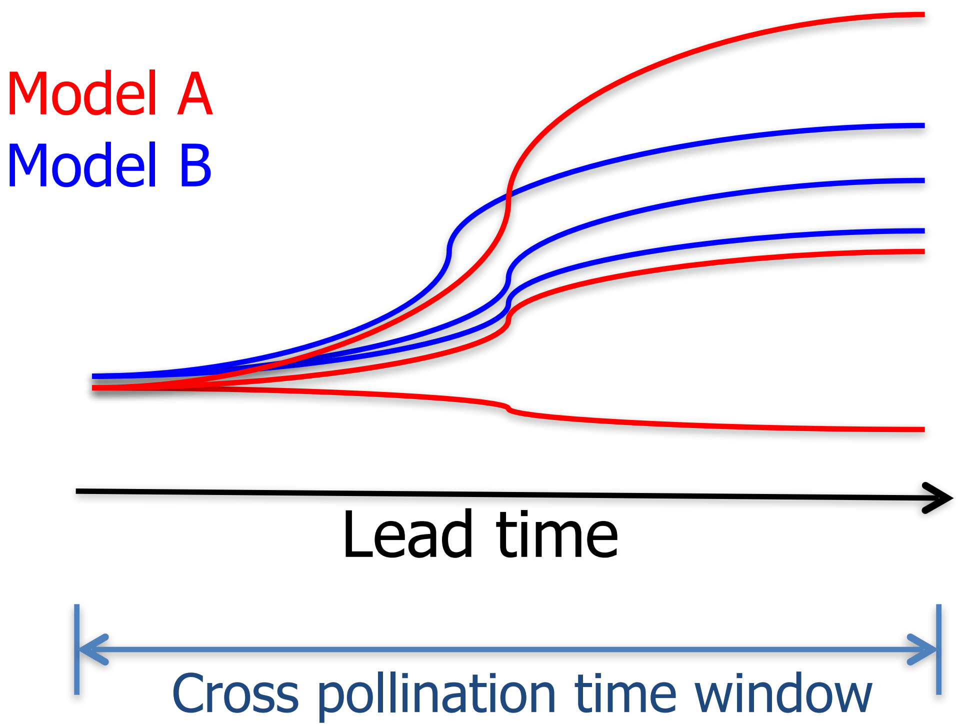
Apply Data Assimilation to the pseudo observations using Model A and Model B respectively.

Of course, Model A and Model B can have very different data assimilation schemes.

Assimilating the future to exploit the richness of multi-model dynamics

Multi-model Cross Pollination in Time

Similarly for Multi-model ensemble



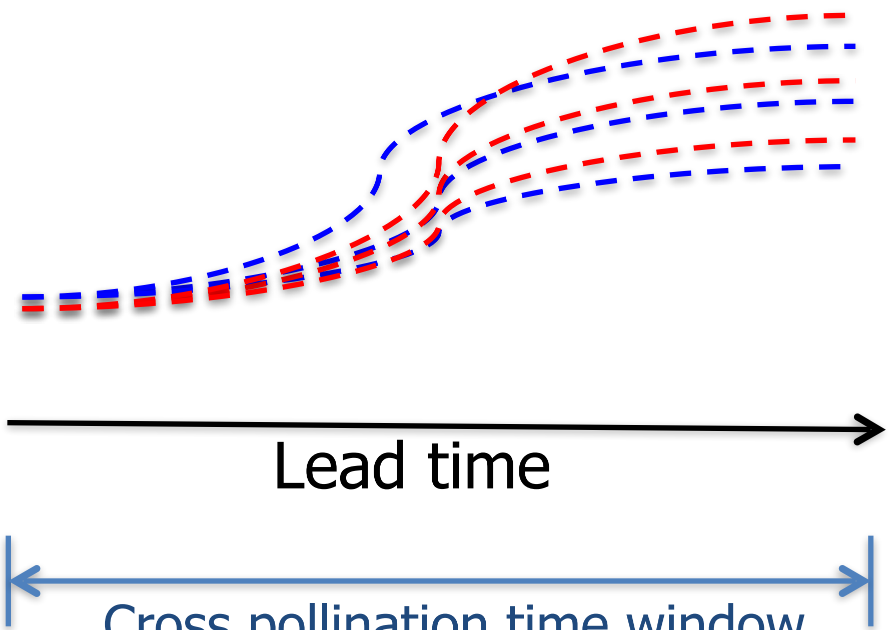
Note that Model A and Model B usually do not share model space at all, but they could share observation space.

Therefore the communication can be made through the shared observation space.

Multi-model Cross Pollination in Time

Data assimilation using Model A

Data assimilation using Model B



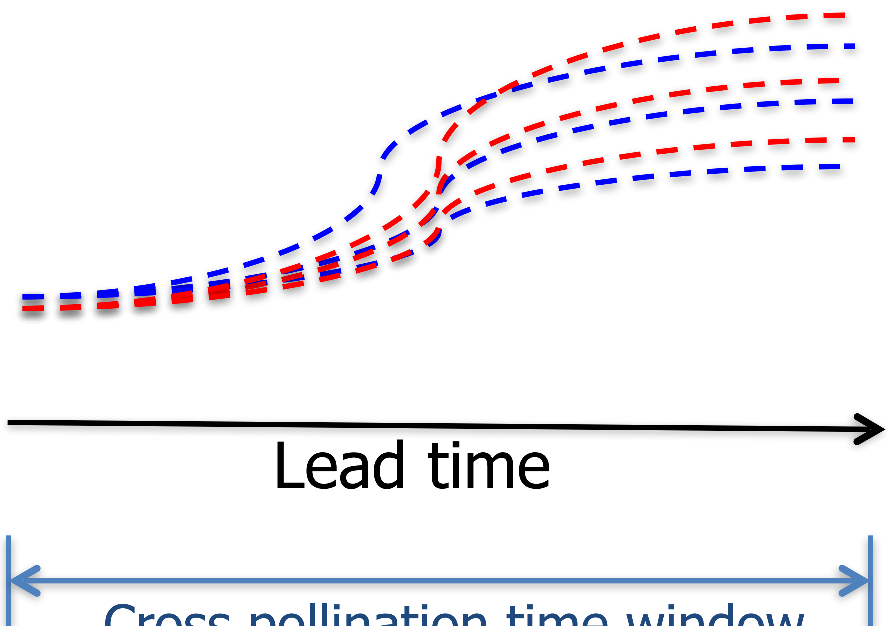
Apply Data Assimilation to the pseudo observations using Model A and Model B respectively.

Of course, Model A and Model B can have very different data assimilation schemes.

Multi-model Cross Pollination in Time

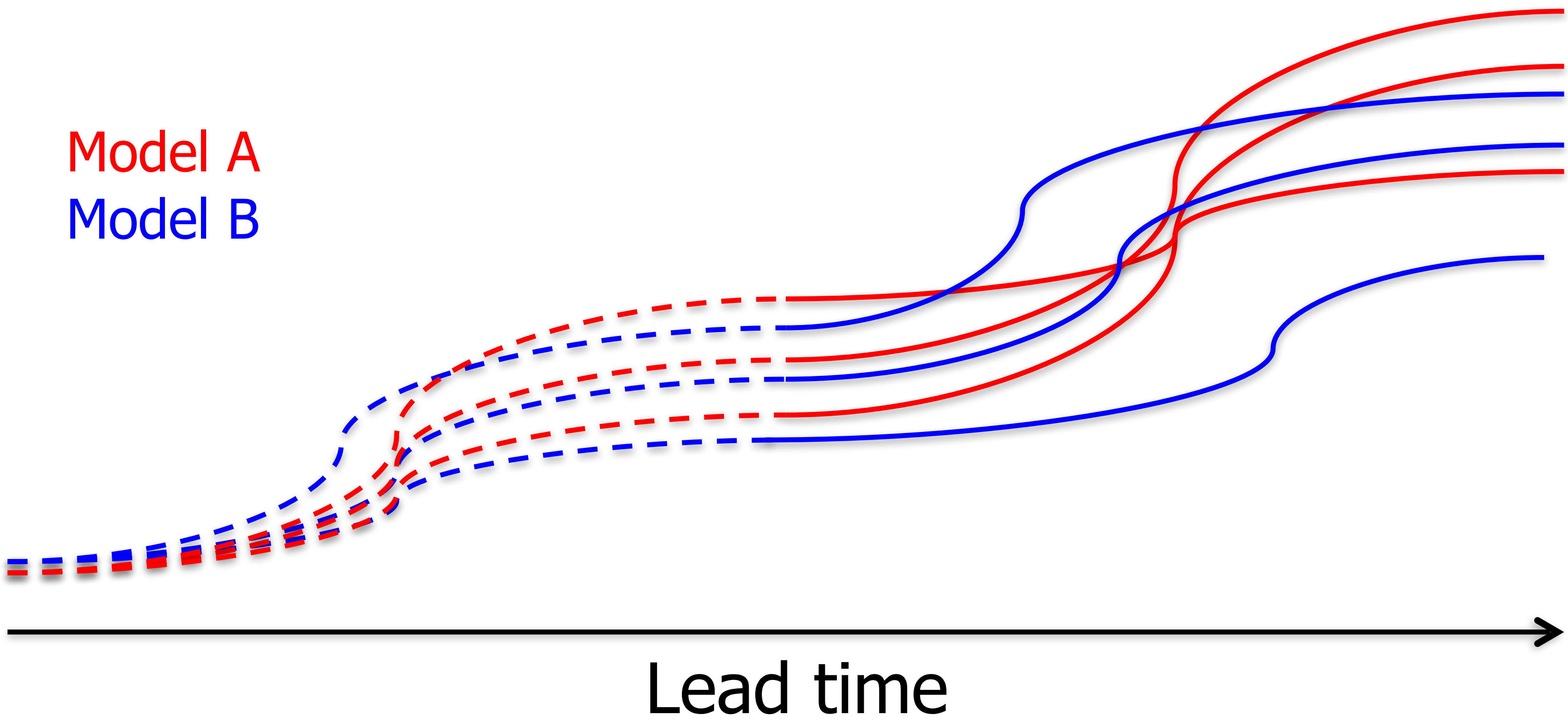
Data assimilation using Model A

Data assimilation using Model B



Iterate the ensemble model states at the end of the cross pollination time window forward under the corresponding model.

Multi-model Cross Pollination in Time



Propagate forward in time and repeat the cross pollination process until the time of interest.

Du, H. and Smith, L.A. (2017) 'Multimodel cross pollination in time', Physica D: Nonlinear Phenomena, Vol. 353-4, pp.31-38.

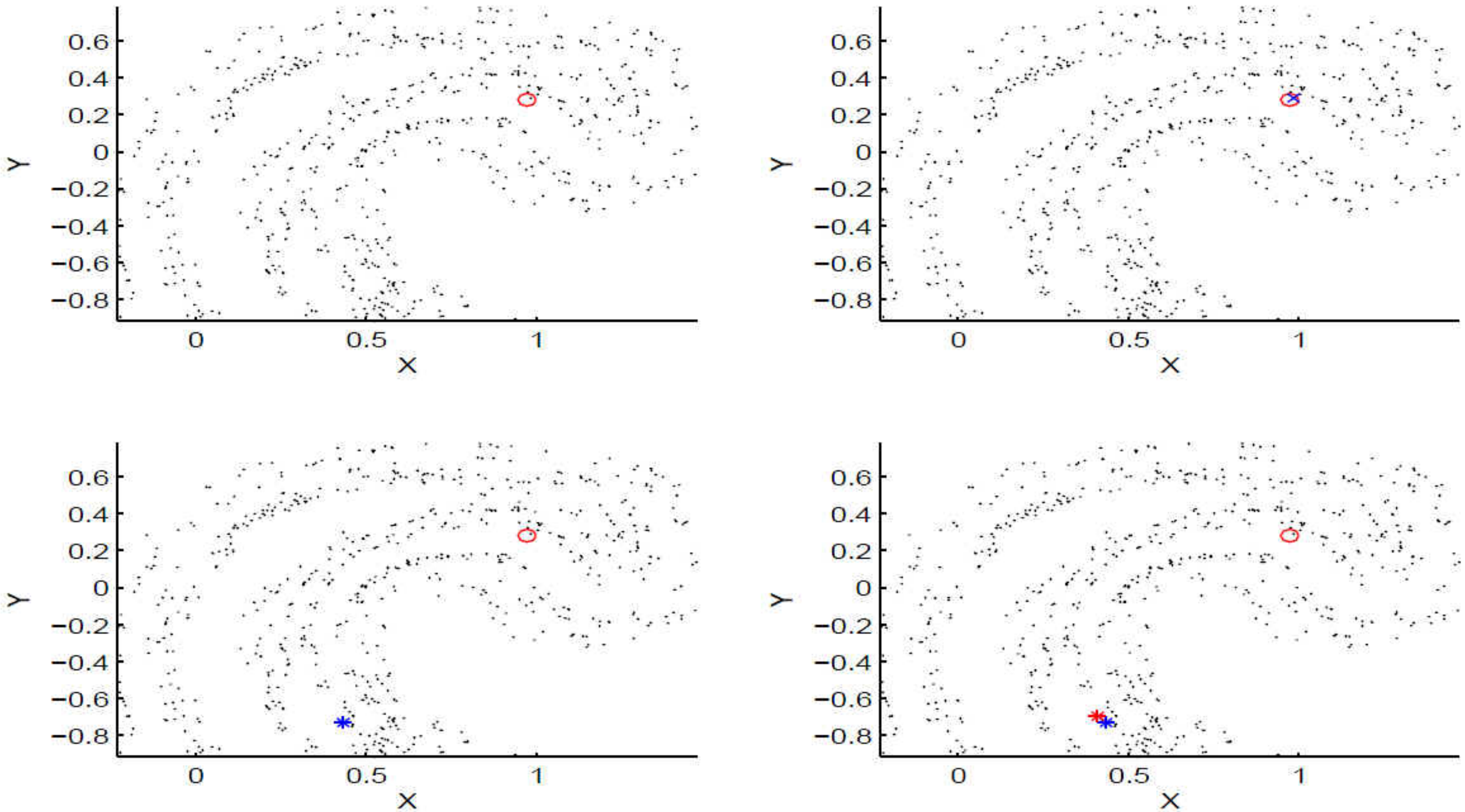
Extract information in delay-space via analogue model

A wide variety of approaches have been explored with the common aim of extracting the information contained in the time ordering of points in delay-space.

To determine the future behaviour assuming that it will be similar to that of the a sample of the "nearby" points.

The most straightforward method is to choose an analogue:
Simply take the point in the reconstruction nearest to the point to be predicted (nearest neighbour) and report for example the nearest neighbour's image or the sum of the nearest neighbour's forward first difference and the current value as the prediction.

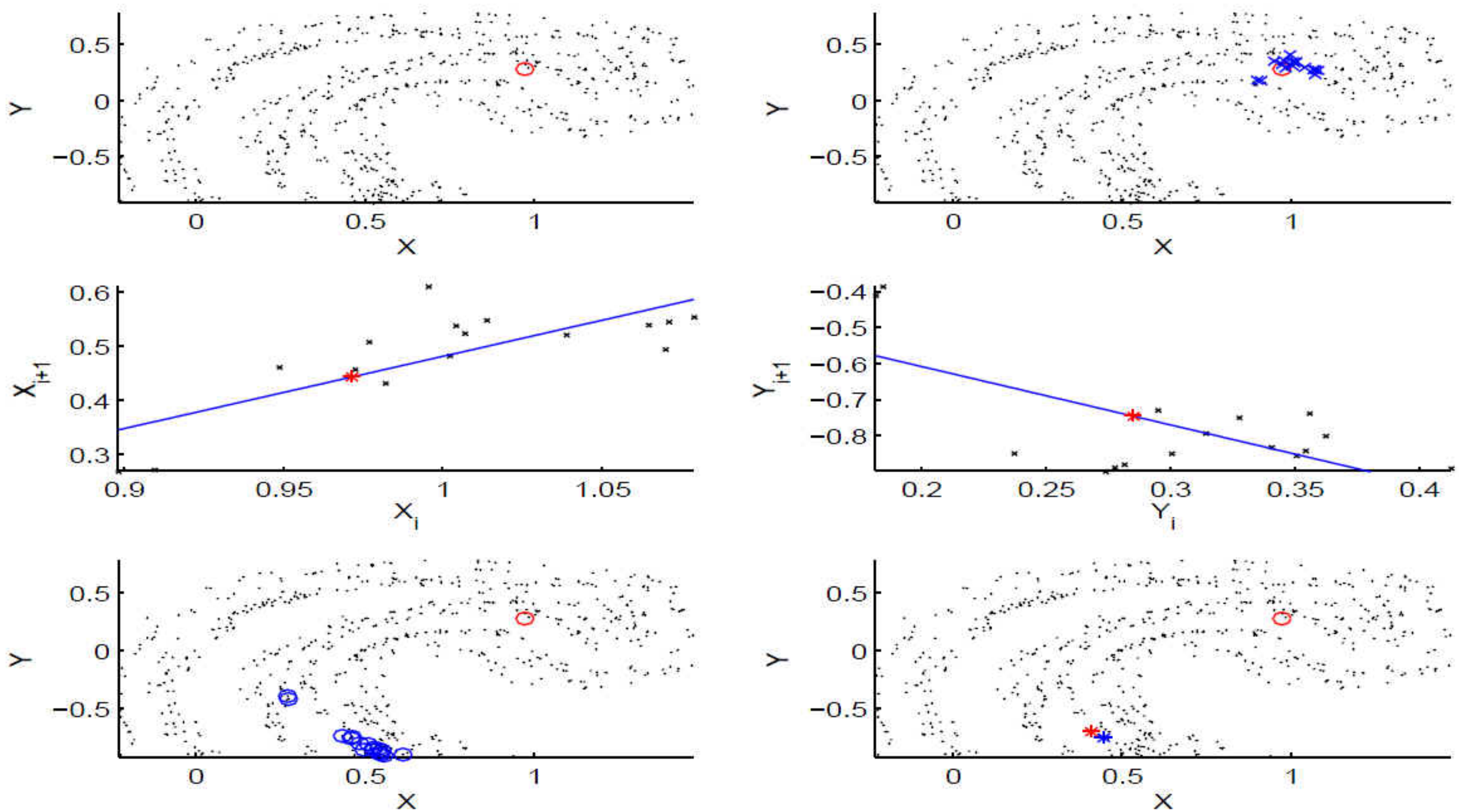
Local Analogue model



Easy to form an ensemble using k nearest points

Local linear model

Given k points within a neighbourhood, a local linear predictor aims at the linear map with the smallest mean square error when interpolating the future observation.



Global Empirical model

Radial Basis Function

Each m dimensional vector, $\mathbf{y}_i$, in the learning set is associated with a (future) value to be predicted, $s(i+\tau_p)$. A predictor is then a map, $\mathcal{J}(\mathbf{y}) : \mathrm{R}^m \to \mathrm{R}^1$ which estimates s for any $\mathbf{y}$. Radial basis function (RBF) predictors consider $\mathcal{J}(\mathbf{y})$ of the form

$$\mathcal{J}(\mathbf{y}) = \sum_{j=1}^{n_c} \lambda_j \phi(||\mathbf{y} - \mathbf{c_j}||)$$

where $\phi(r)$ are radial basis functions

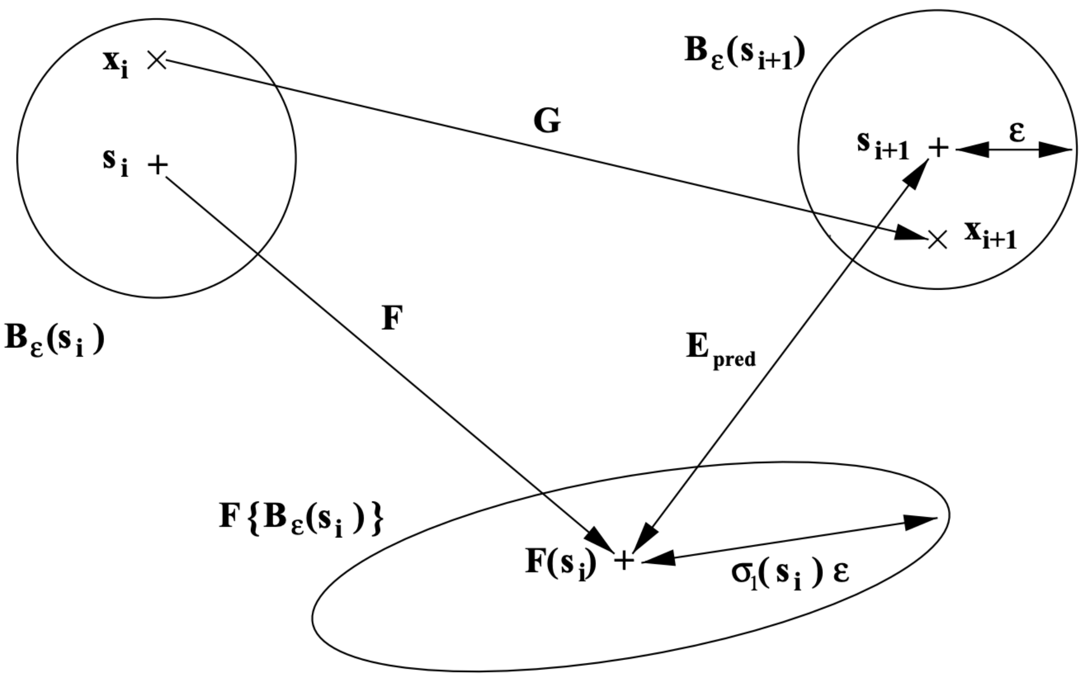
Typical candidates for $\phi(r)$ include $\phi(r) = r^3$ and $\phi(r) = e^{-r^2/\sigma^2}$ where the constant σ reflects the average spacing of the centers $\mathbf{c_j}$. $\mathcal{J}(\mathbf{y})$ is constructed about n_c centers

$$\mathbf{c_j}, \quad j = 1, 2, \ldots, n_c; \quad \mathbf{c_j} \in \mathrm{R}^m$$

chosen to cover the region of state space which the reconstruction explores.

model discrepancy recognition

Consistency test

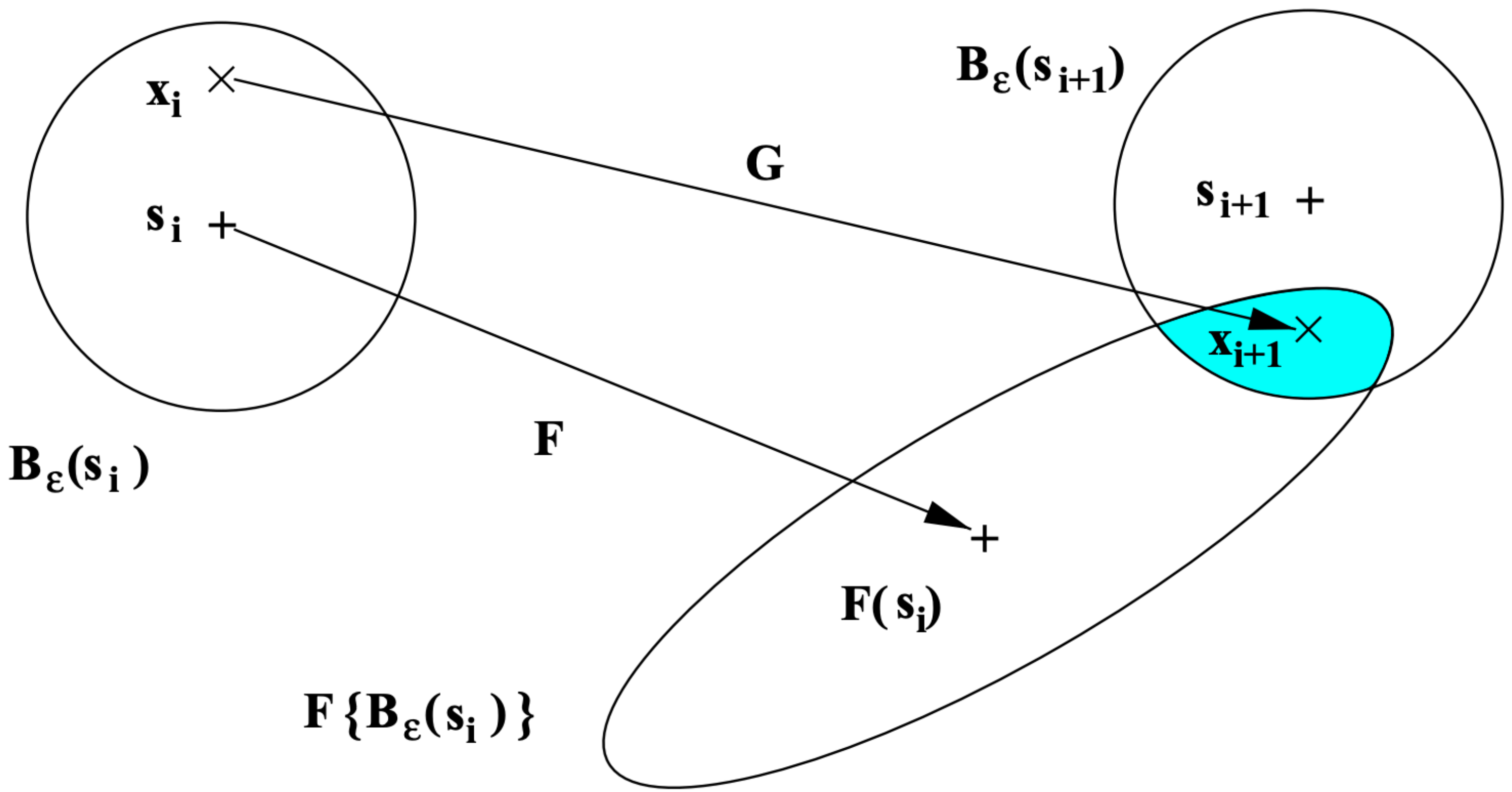


No model trajectory exists which is consistent with the observational uncertainty

McSharry, P.E. and Smith, L.A. (2004) 'Consistent nonlinear dynamics: identifying model inadequacy', Physica D, 192: 1-22.

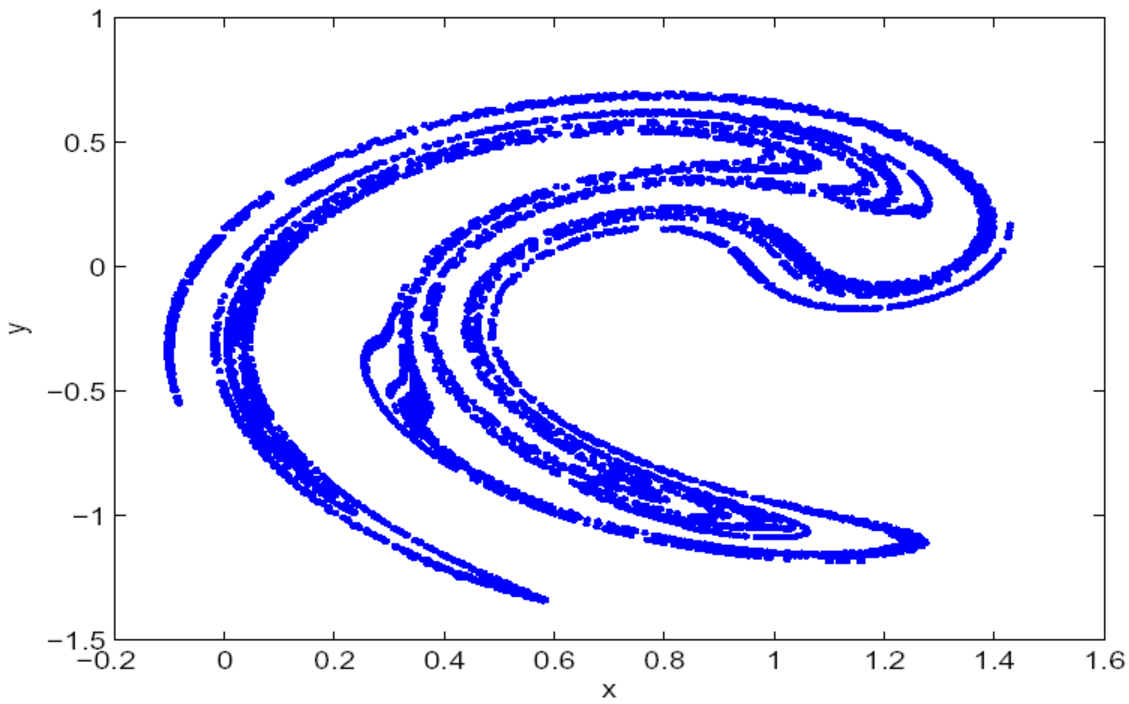
model discrepancy recognition

Consistency test: Consistent nonlinear dynamics



model discrepancy recognition

For Local Linear Models: How large a local region in the state space?



model discrepancy recognition

How large a local region?

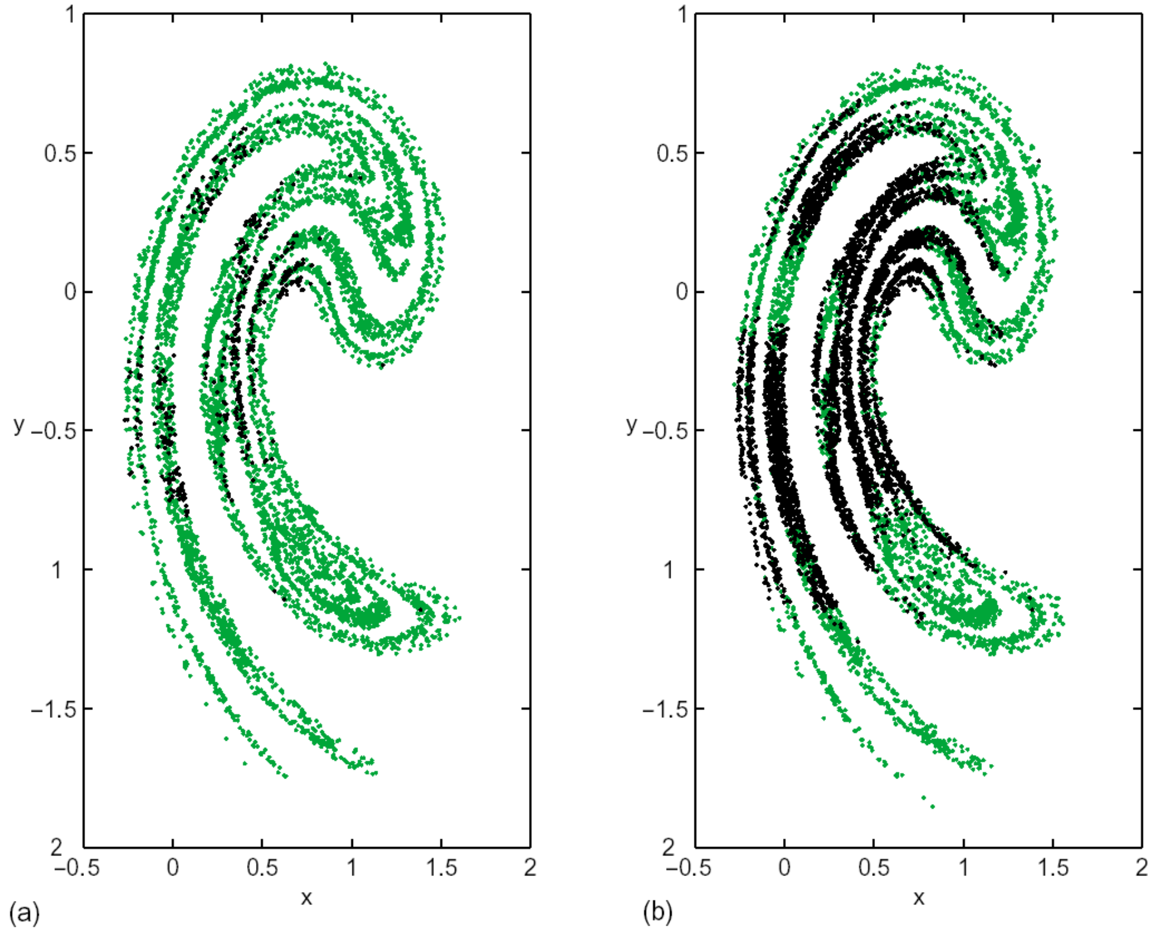
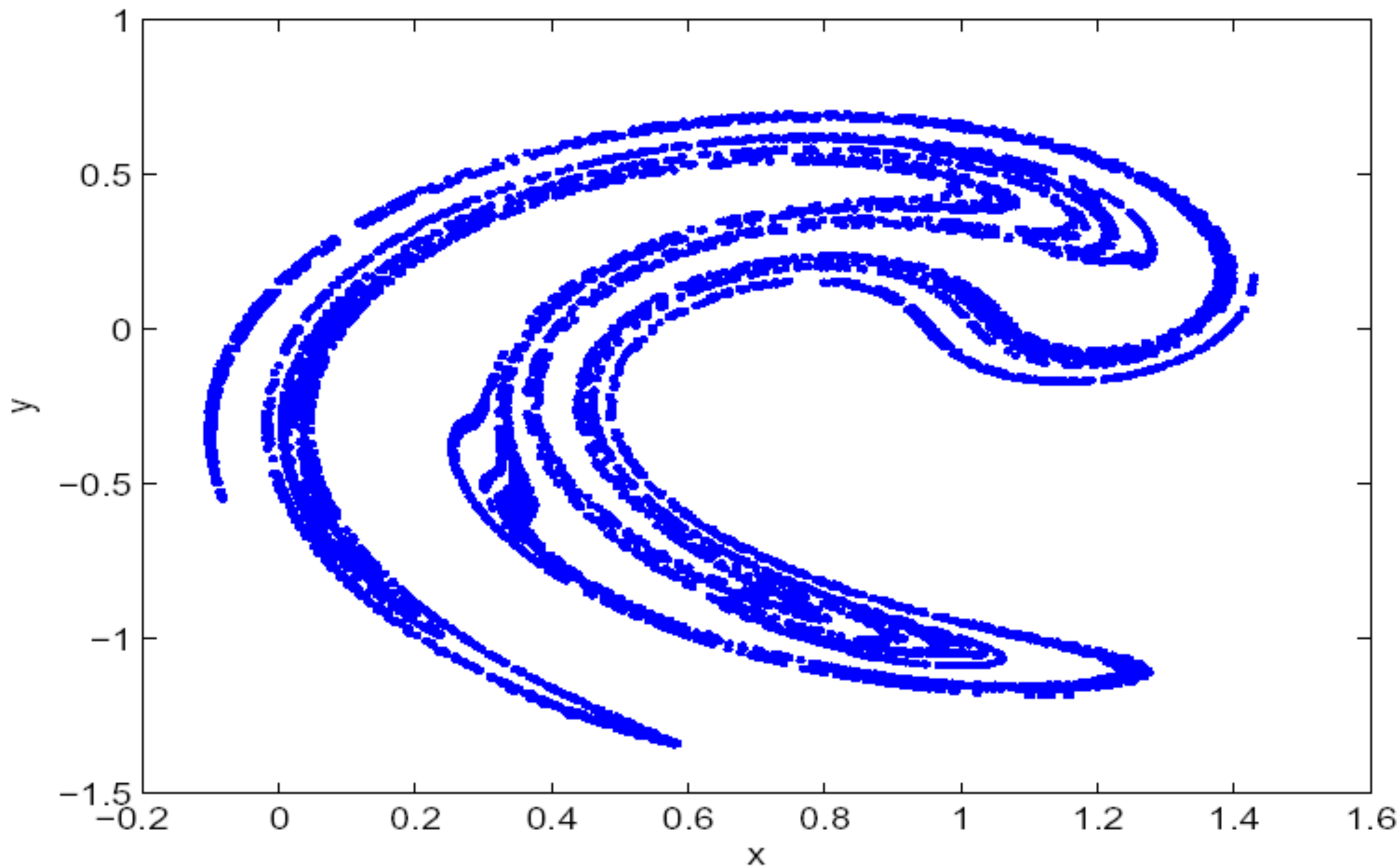
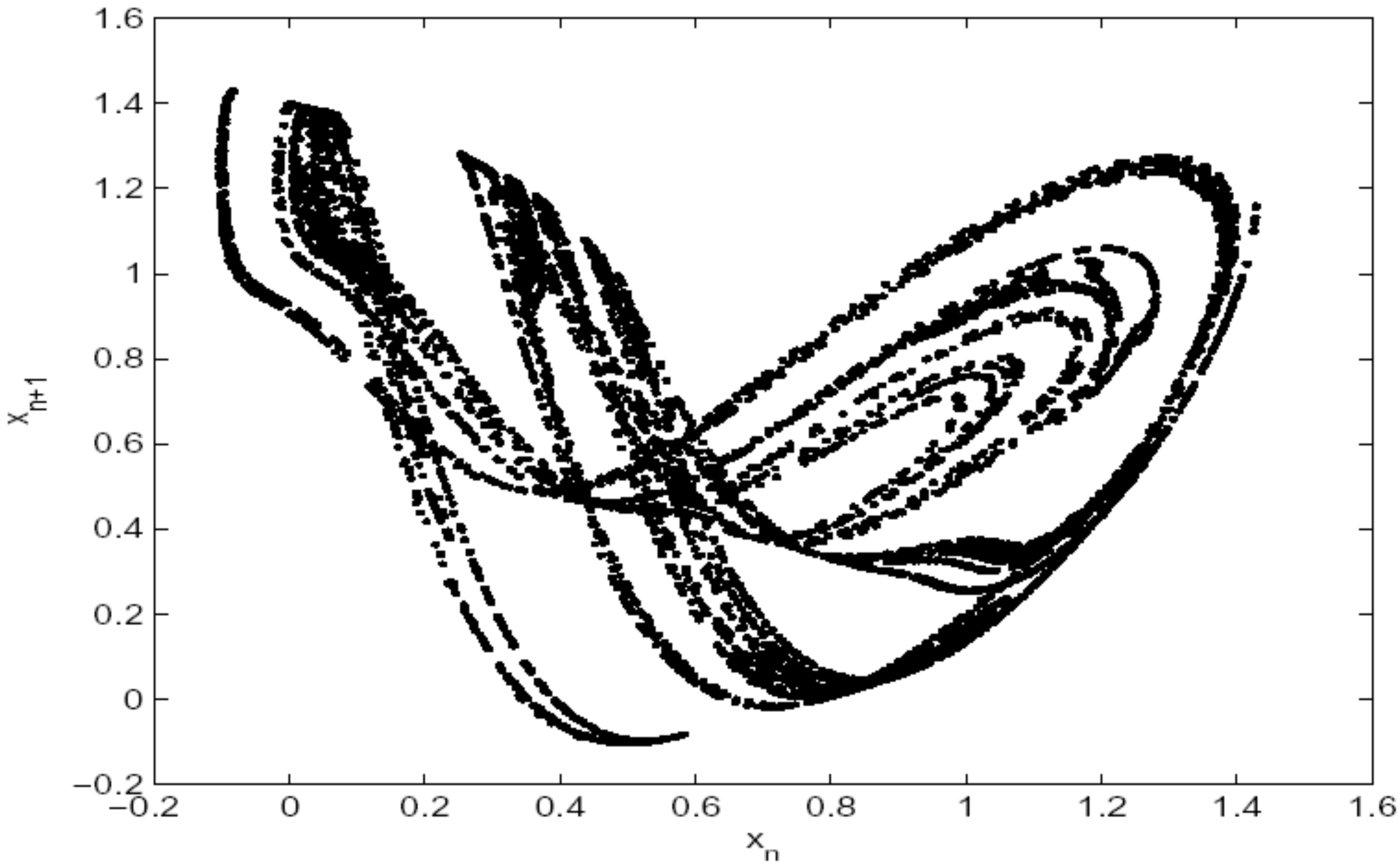


Fig. 4. Consistent (grey) and inconsistent (black) predictions of the full state space Ikeda map using local linear prediction with neighbourhood radii of (a) 0.15 and (b) 0.20.

Delay reconstruction

Ikeda Map




1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
x_{n+1}
−0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
x_n

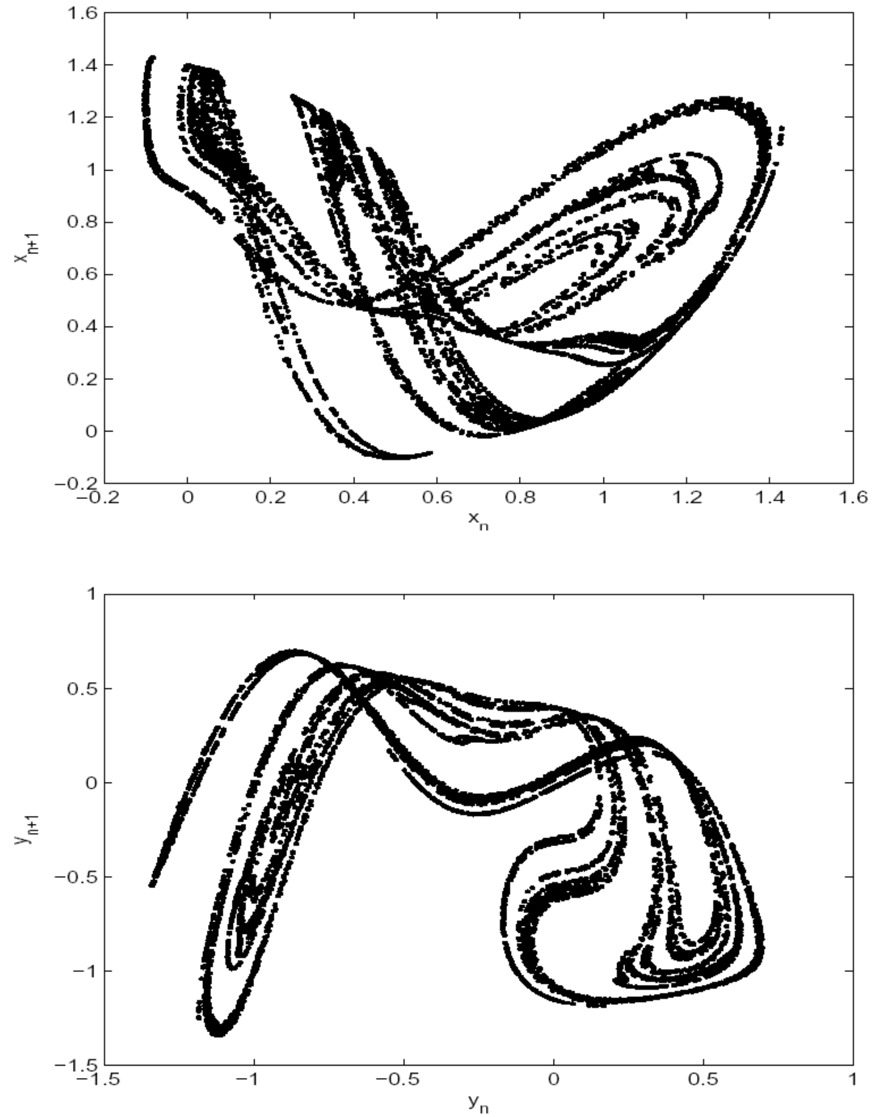


Figure 2: Delay plot of Ikeda map for both x and y

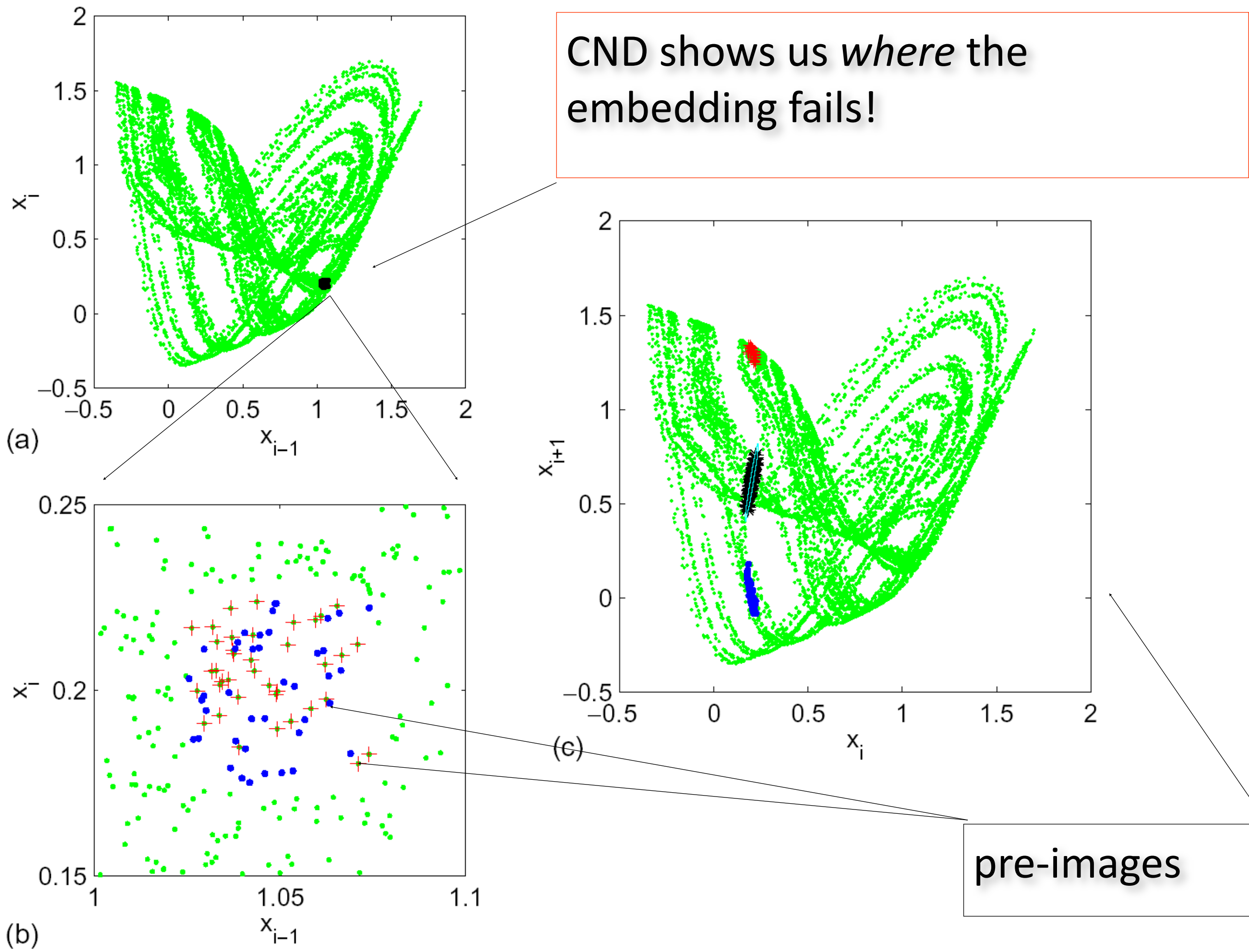


Fig. 5. An illustration of self-intersection in an $m = 2$ reconstruction of the Ikeda map: (a) the reconstructed state space (x_{i-1}, x_i) showing observed points within a square centred at the base-point (1.05, 0.2) being consistent with measurement errors of magnitude $\varepsilon = 0.025$, (b) a zoom-in showing pre-images of points which have large x_{i+1} (pluses) and points which have small values of x_{i+1} (dots) and (c) the reconstructed state space (x_i, x_{i+1}) with images from (b), corresponding local linear predictions (black) and consistency ellipse (grey).

model discrepancy recognition

4.3. Rulkov circuit equations

A second mathematical system will be used to illustrate how CND can identify state space-dependent model error, in this case due to incorrect parameter values. The system is the set of equations defining Rulkov's circuit [33,34] are

$$\dot{x} = y, \qquad \dot{y} = -x - \delta y + z, \qquad \dot{z} = \gamma[\alpha f(x(t)) - z] - \sigma y, \tag{12}$$

where $x(t)$ is the voltage across the capacitor C, $y(t) = \sqrt{(L/C)}i(t)$ with $i(t)$ the current through the inductor L, and $z(t)$ is the voltage across the capacitor C'. Time has been scaled by $1/\sqrt{LC}$. The parameters of this system have the following dependence on the physical values of the circuit elements

$$\gamma = \frac{\sqrt{LC}}{RC'}, \qquad \delta = r\sqrt{\frac{C}{L}}, \qquad \sigma = \frac{C}{C'}. \tag{13}$$

The function $f(x)$ is

$$f(x) = \begin{cases} 0.528, & x \leq -x_a, \\ x(1 - x^2), & -x_a < x < x_a, \\ -0.528, & x \geq x_a, \end{cases} \tag{14}$$

and the control parameter α characterises the gain of the nonlinear amplifier around $x = 0$. The parameters of the circuit correspond to the following values for the coefficients in the differential equations (12): $\gamma = 0.294$, $\sigma = 1.52$, $\delta = 0.534$, $\alpha = 15.6$ and $x_a = 1.2$.

The model is the same set of equations but with $x_a = 1.4$, thus the model is structurally correct, has one parameter in error; all other parameters are exactly the same as the system.

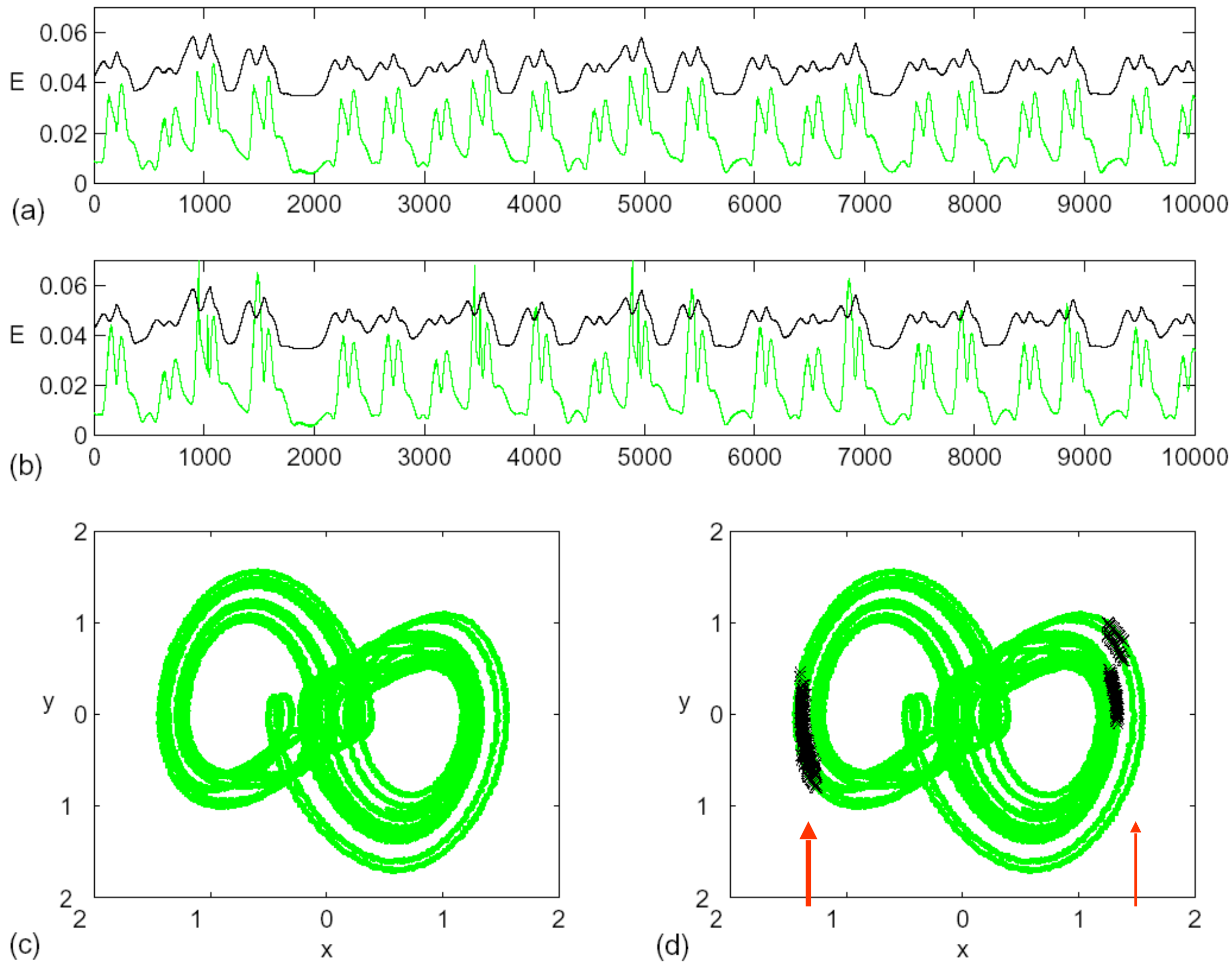


Fig. 7. Consistency analysis of the Rulkov circuit. The upper panels illustrate the prediction errors for the perfect model (a) imperfect model (b). The grey line is the prediction error and the black line is the consistency envelope. Points above the envelope are inconsistent. The lower panels are 2D projections of the delay reconstruction showing both consistent points (grey dots) and inconsistent points (×) for perfect (c) and imperfect model (d).

Examine 1.2 < |x| < 1.4

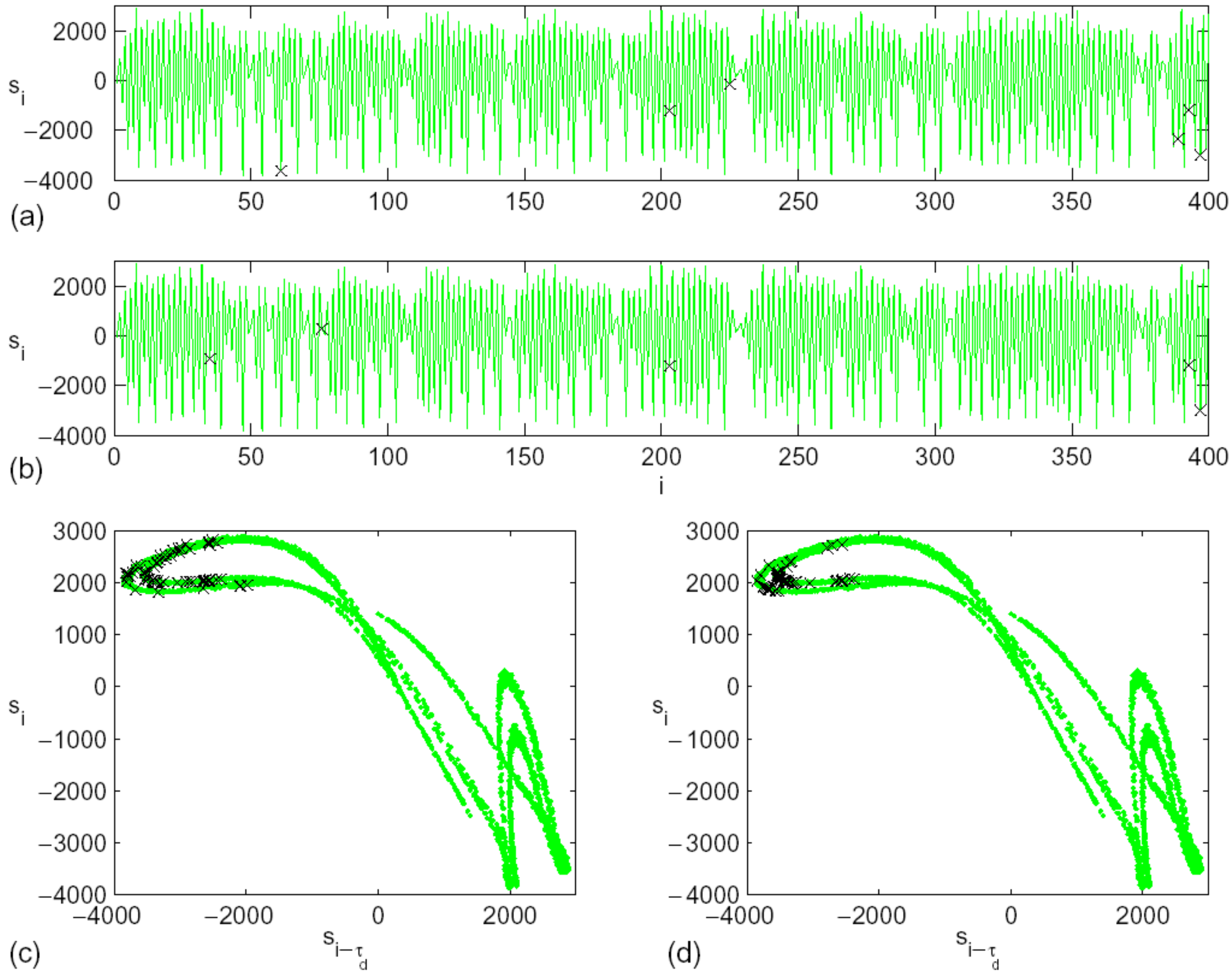
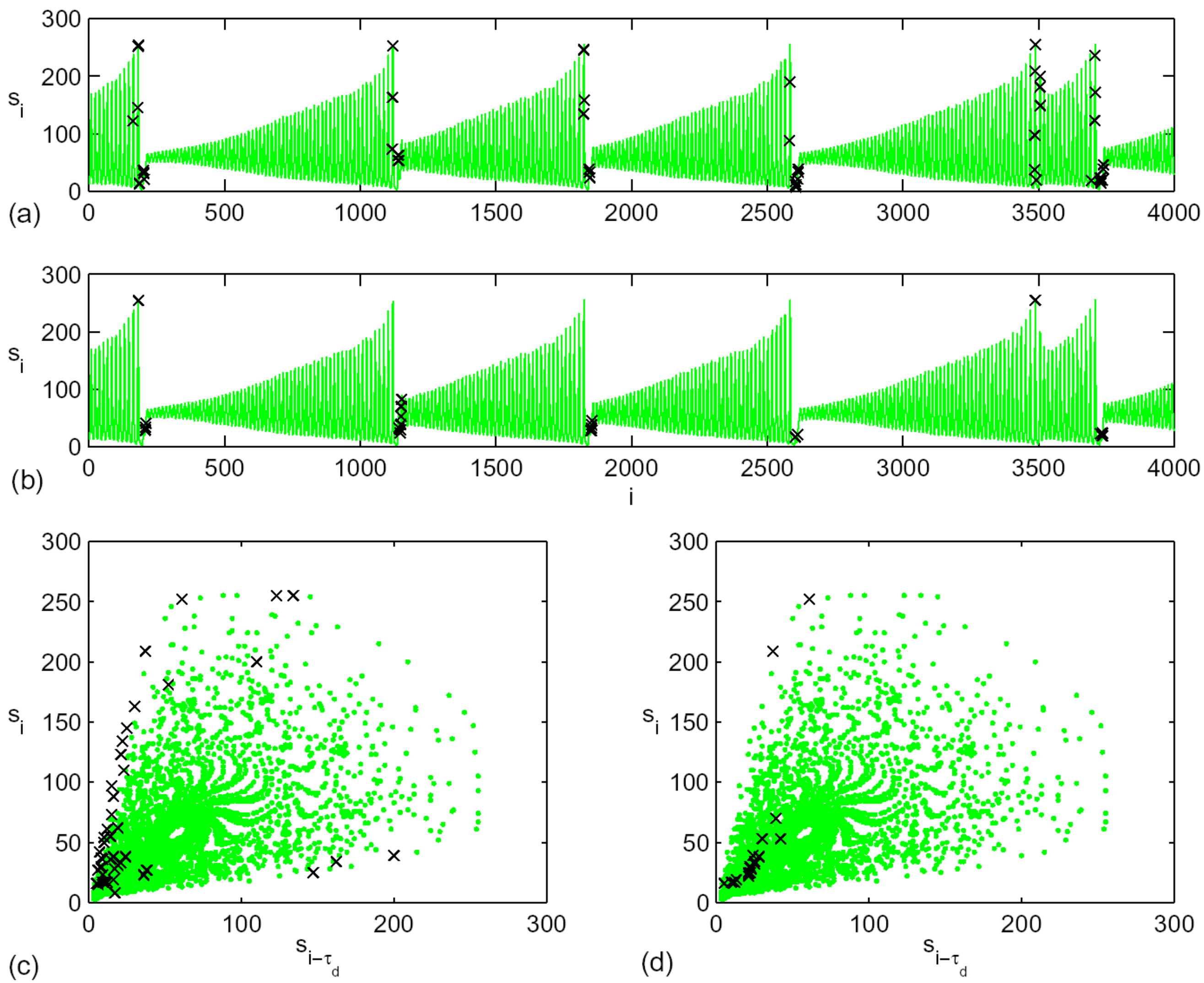


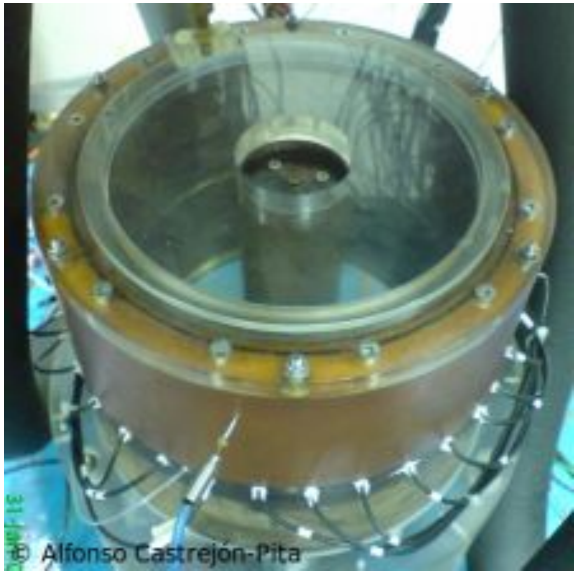
Fig. 10. Predictions of the NMR laser data contrasting models RBF_2 and LL_{2a}. The upper panels illustrate the inconsistent predictions ($\times$) for models RBF_2 (a) and LL_{2a} (b). The lower panels are 2D projections of the delay reconstruction showing both consistent points (grey dots) and inconsistent points ($\times$) for RBF_2 (c) and LL_{2a} (d).

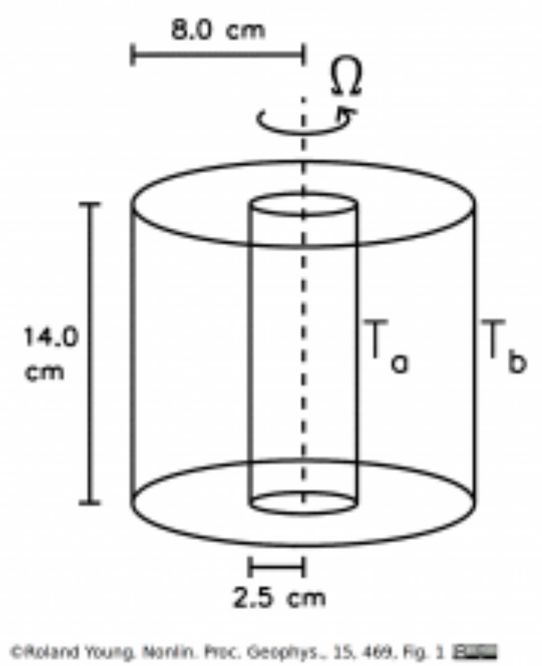


Note that the RBF fails just *before* the collapse, while the LL fails just *after*!

Fig. 8. Predictions of the NH_3 laser data. The upper panels illustrate the inconsistent predictions (×) for models RBF_1 (a) and LL_{1a} (b). The lower panels are 2D projections of the delay reconstruction showing both consistent points (grey dots) and inconsistent points (×) for RBF_1 (c) and LL_{1a} (d).

Annulus





https://www.physics.ox.ac.uk/research/group/geophysical-and-astrophysical-fluid-dynamics

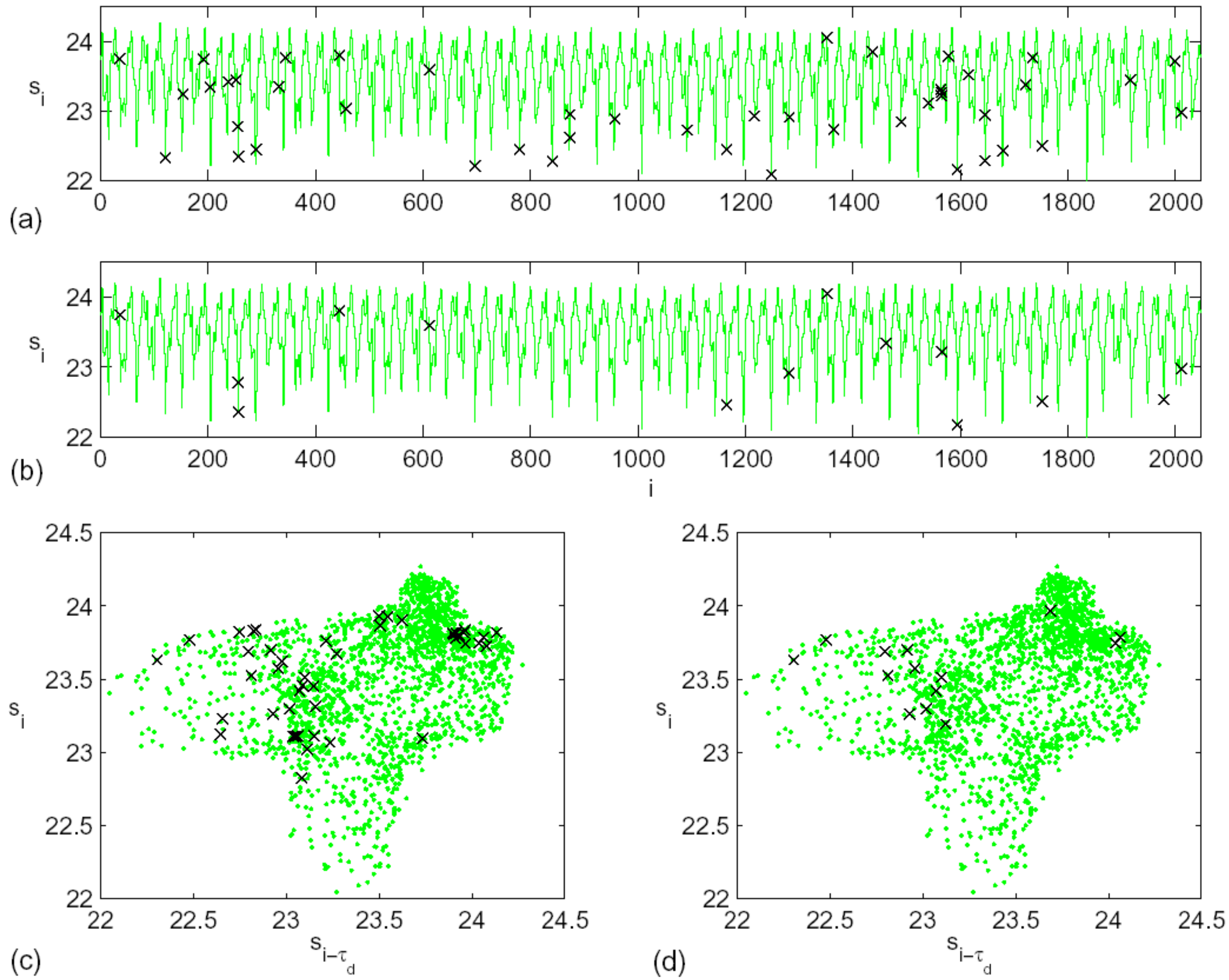


Fig. 13. Predictions of the Annulus data contrasting models RBF_3 and LL_3. The upper panels illustrate the inconsistent predictions ($\times$) for models RBF_3 (a) and LL_3 (b). The lower panels are 2D projections of the delay reconstruction showing both consistent points (grey dots) and inconsistent points ($\times$) for RBF_3 (c) and LL_3 (d).