CATS CENTRE FOR THE ANALYSIS OF TIME SERIES | LSE | NATURAL ENVIRONMENT RESEARCH COUNCIL

HOW TO QUANTIFY THE PREDICTABILITY OF A CHAOTIC SYSTEM?

E.A. Sienkiewicz[1] | L.A. Smith[1,2] | E.L. Thompson[1]

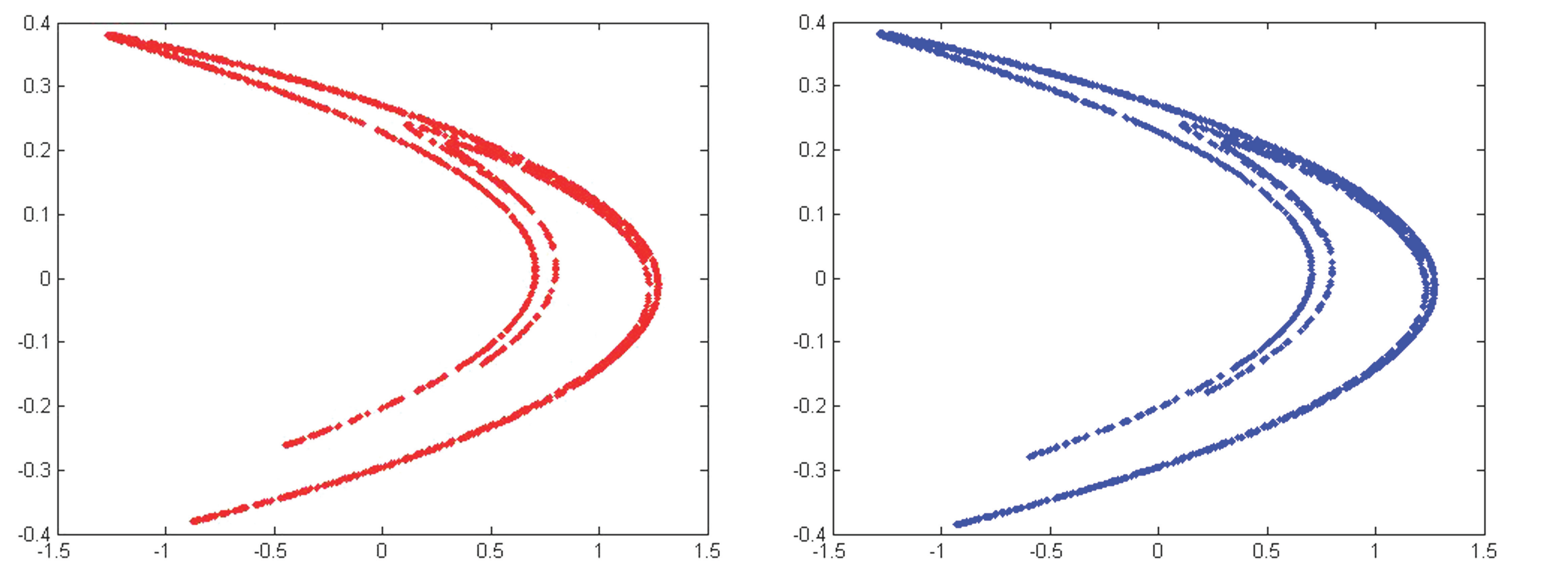
1 Centre for the Analysis of Time Series, London School of Economics | 2 Oxford Centre for Industrial and Applied Mathematics, Oxford University, U.K.

ABSTRACT

Models are tools that describe reality in form of mathematical equations. For example General Circulation Models (GCM) represent actual climate system and are used to investigate major climate processes and help us better understand certain dependencies amongst climate variables. Global forecasts help foresee severe weather anywhere on the planet and save many lives, although meteorology is unreliable in long run. A model is only an approximate representation of nature, which is reflected by model error. In addition, small uncertainties in the initial conditions usually bring up errors in the final forecasts. We can handle initial condition uncertainty but not model error. This study examines how to quantify predictability of complex models with an eye towards experimental design.

1. EXPERIMENT TO DISTINGUISH MODEL ERROR AND CHAOS.

We carry out the experiment, which aims to distinguish model error and chaos. The research is based on a classical Henon Map (model) and simple nonlinear two-dimensional chaotic map, including a seasonal cycle - Henon Map described in "What might we learn from climate forecasts" (L.A. Smith, 2002), which constitutes system.



(a) Classical Henon Map - model *(b) Henon Map, including seasonal cycle - system*

Figure 1: Model and System

In the experiment we derive Gaussian noise ball of 2^{12} points around the point (-0.51, 0.32) (ensemble S) and standard deviation: 10^{-2} and 10^{-3}. We divide the ensemble onto two sets: "1" and "2", each consisting of 2^{11} randomly chosen points. Further, we use a model and a system to evolve each set of initial points (S1 and S2) from the specific ensemble and store their position at the first 64 steps. The scheme for the experiment is illustrated in (Fig.2).

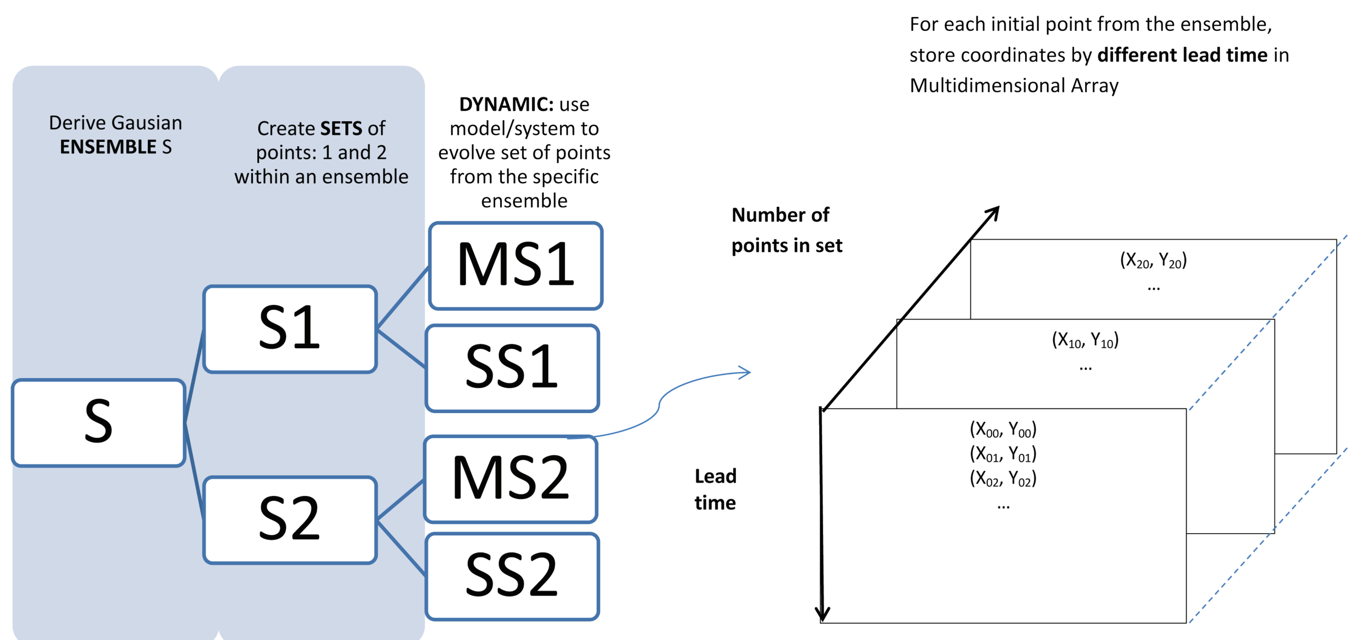


Figure 2: Design of the experiment

To capture the model error we analyse how one set of points behaves under a model and a system at different time steps. To capture chaos we analyse how two sets of points within an ensemble act under the same system.

2. ANALYSIS OF THE EXPERIMENT. RELATIVE ENTROPY.

To check the properties of the distributions we use relative entropy. It is a measure of the difference between the two probability distributions P and Q expressed by the equation below.

$$D_{KL}(P \parallel Q) = \sum_i \log_2 \left(\frac{P(i)}{Q(i)}\right) P(i).$$

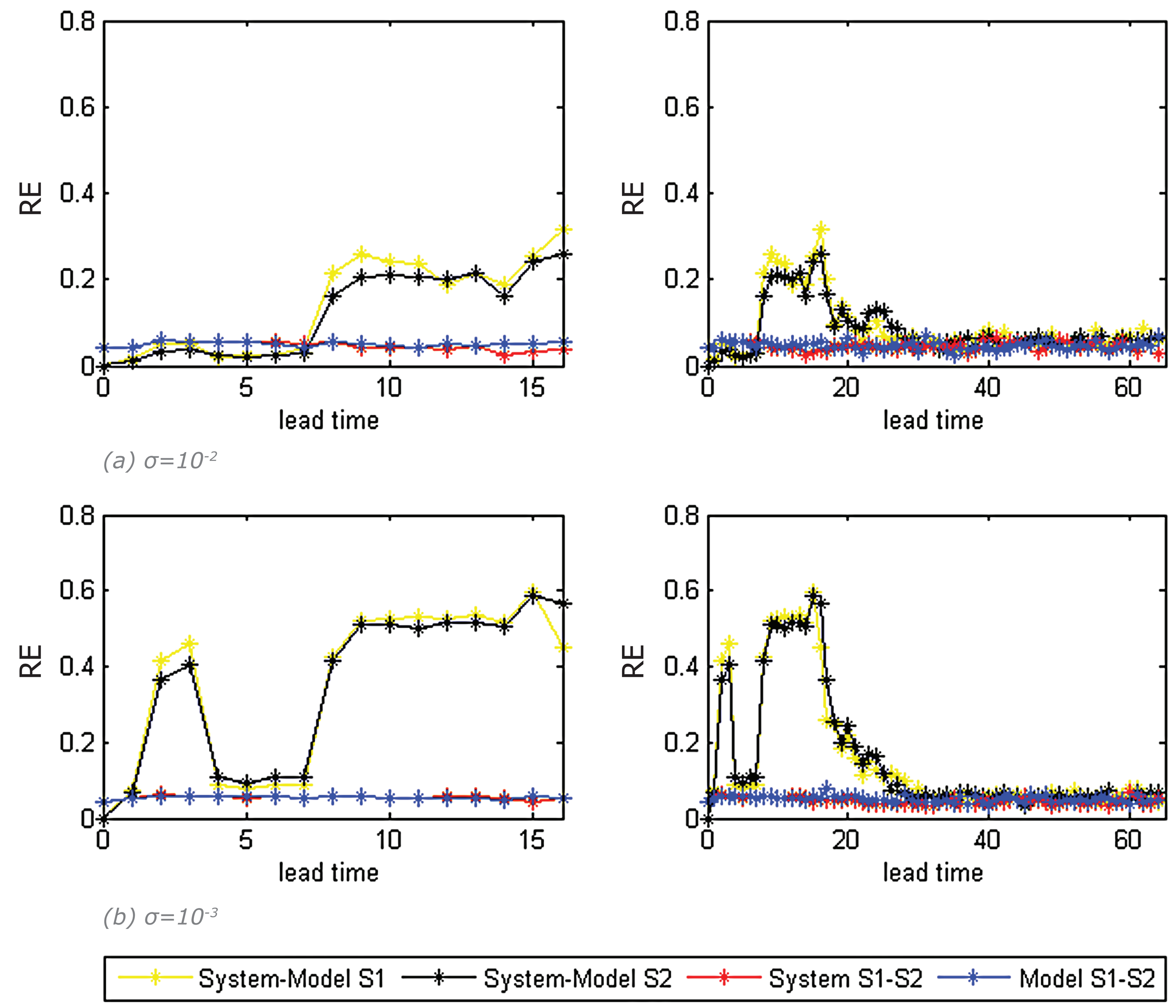
Where: i - represents number of bins. In our case there are 2^6 bins P -is a probability distribution of the model (MS1, MS2) Q - is a probability distribution of the system (SS1, SS2).

We expect the RE for sets developed on model and system to be at zero at t=0 since the distributions are identical, and to converge close to zero after a large number of time steps, because both attractors are similar. The interesting part should be in the middle, where the model error results in very large relative entropy, which means an overconfident forecast. For sets developed under the same dynamics we anticipate that the RE should be close to zero since they move around the same attractor. The small difference between distributions is a result of initial uncertainty.

Conclusions:

1. Relative entropy (RE) for a set evolved on a system and a model (yellow, black lines) reaches its maximum at around 20th time steps and then falls down to near 0 as time increases.
2. RE between two sets of points derived from the same distribution and evolved using system (red line) or model (blue line) is slightly above zero, but never falls down to zero.
3. RE for the Gaussian ball with standard deviation 10^{-2} (case a) stays small for few steps before growing.
4. The time when RE starts to fall down increases as Gaussian ball of initial conditions shrinks.
5. Value of RE is negatively correlated with the size of ensemble and value of its standard deviation.

Figure 3: Charts on the right: RE for first 16 time steps between set of points:S1 (yellow line) or S2 (black line) developed on model and system -to capture model error; RE between two sets of points S1 and S2 within an ensemble developed under the system (red line) or model (blue line). Charts on the right side show RE as function of lead time for the first 64 time steps. Top charts concern results of ensemble with standard deviation (a) 10^{-2} and (b) 10^{-3}.



3. HOW DO THESE RESULTS RELATE TO WEATHER MODELS?

In SHORT TERM (step 1-7) model is **useful** and it gives a **good prediction**. It tells where the points are going to be and the relative entropy shows that. It is a good prediction, because it is constrained. It is analogical to a weather forecast of "between 3-6 degrees for tomorrow's temperature, when the actual value is 5 degrees".

In MEDIUM TERM (step 8-20) the degree by which system distribution differs from the model distribution is the biggest. One could compare it to a forecast of "between 7 degrees and 10 degrees for tomorrow's temperature", when the actual value is 5 degrees. In this case it would be actually **wrong prediction** (not just inaccurate), and the relative entropy, gives a very bad score for that.

In the LONG RUN (step 20+) model looks like climatology and it is **not useful**. It is giving a good prediction but it would not be any more useful than looking at old statistics. It's like a forecast of "between minus 20 and plus 15 degrees" for tomorrow's temperature.

4. BIGGER PICTURE.

In the experiment we noticed that, as we decrease the number of members in the ensemble or cut down the noise level, the forecast changes and RE reflects this (it has higher value). We also expect longer time step till we lose predictability. Do we always observe such improvement? The problem is that, curves presented in Figure 3 change a lot depending on where on chaotic system we build the ensemble of initial conditions. This issue is illustrated in Figure 4.

5. CONCLUSIONS

Predictions in meteorology are limited in time because of (i) **initial uncertainty** are highly sensitive to initial conditions and (ii) the presence of **model error** reflecting the fact that model is only an approximate representation of nature.

An interesting feature of chaotic systems is that as we go to the different places on attractor, the **time till we lose the predictability varies by location** of the ensemble of initial conditions. We can apply different techniques to change the way we make the forecast on the system i.e. by increasing the size of the ensemble or by cutting down the noise level, but this does not guarantee any improvement.

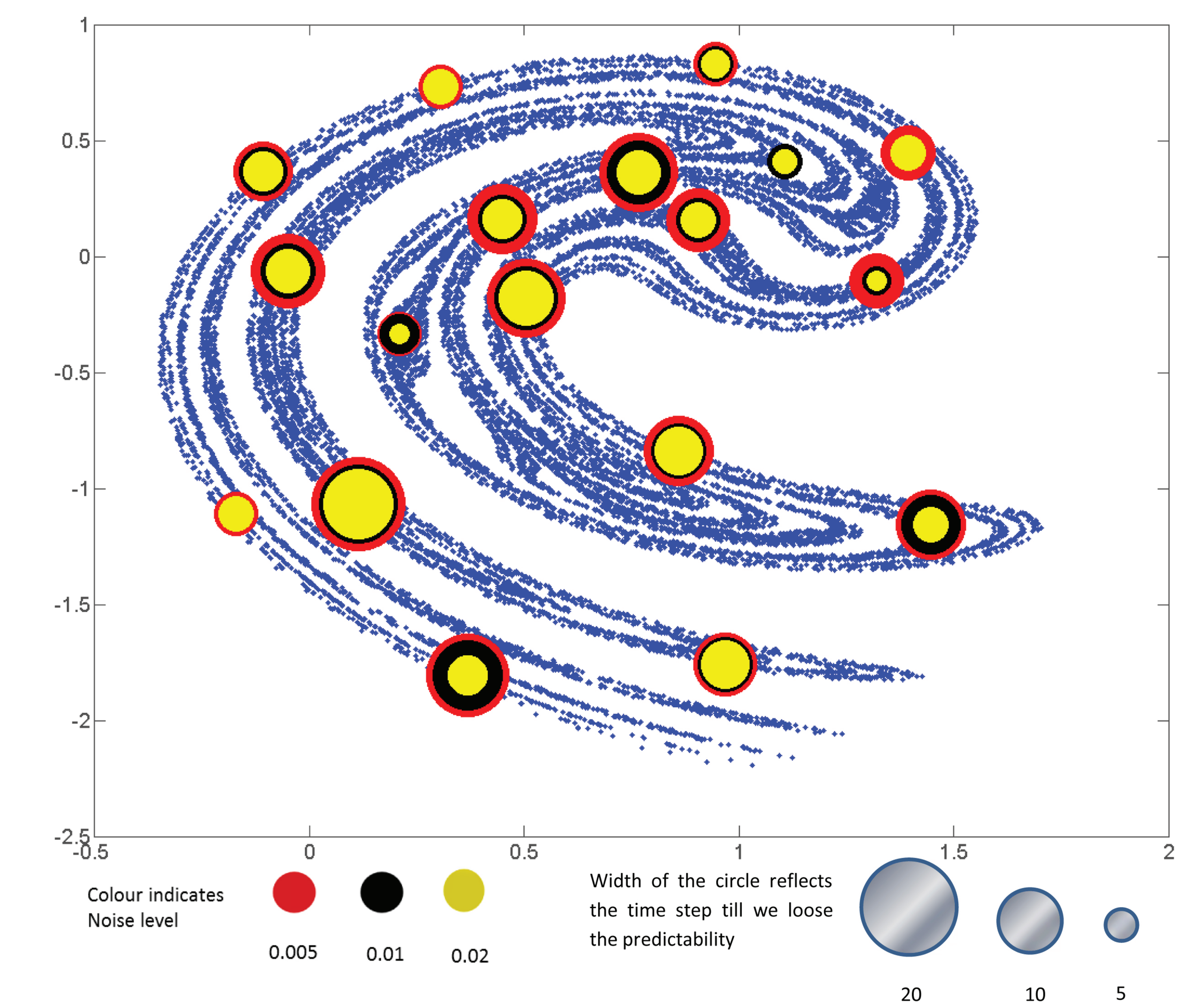


Figure 4: For this different initial conditions where the circles are located on Ikeda Map (another example of two dimensional chaotic systems), the width of the circles tells us how long it takes before we lose the predictability, before we are better than climatology. Colour indicates the noise level. The size of the circle differs as we move around the attractor and we see many different structures in terms of how predictability improved as the noise level decreased.

References:

[1] L. A. Smith. What Might We Learn from Climate Forecasts? Proc. National Acad. Sci. USA 4 (99): 2487-2492.

[2] W. H. Press, S.A. Teukolsky, W. T. Vetterling, B. P. Flannery (2007), Numerical Recipes, The Art of Scientific Computing, 14.7 Information - Theoretic Properties of Distributions. Cambridge University Press, 756-758.

Acknowledgements:

The support of NERC is gratefully acknowledged.