Using shadowing ratios to evaluate data assimilation techniques

E. Wheatcroft[1] and L. A. Smith[1,2]

[1]Centre for the Analysis of Time Series, London School of Economics
[2]Oxford Centre for Industrial and Applied Mathematics, Oxford University, U.K.
Email: *e.d.wheatcroft@lse.ac.uk*

Abstract

Identifying successful "noise reduction" as such remains a challenge in applications where the "true" values are not know a priori (and linear noise reduction is a game not a true task). We suggest shadowing ratios as a measure of noise reduction when the task at hand involves prediction. Initial condition uncertainty will more or less always limit the lead time of a chaotic model, even when that model reproduces the system dynamics perfectly. Since in reality, observations from such systems tend to be clouded by measurement error, the maximum lead time we can expect to accurately predict using the model will be short. Data assimilation techniques attempt to improve our state estimates, we introduce a new measure which allows us to estimate the quality of these techniques. A model trajectory shadows for as long as it is consistent with the noise model of the observed states. We define the shadowing ratio as the ratio of the length of time the model shadows using the assimilated initial conditions to the length of time the model shadows using some reference data assimilation technique. We use the measure to evaluate the effectiveness of one assimilation technique in particular, Gradient Descent of Indeterminism (GDI). Using the Duffing Map system as an example we first use shadowing ratios to show the effect of using different numbers of observations from the past when applying GDI. We then compare GDI to other assimilation techniques using the measure to compare the effectiveness from a forecasting perspective. Finally, since GDI requires derivative information from the system, we compare the effectiveness of the algorithm when using the exact derivative matrix and when approximating using a forward difference technique.

1 Introduction

Successful prediction of complex process has a huge value to society, particularly in areas like meteorology where it can affect many lives. A modeller of chaotic dynamical systems however, faces big challenges. Even with a perfect model of the dynamics of a system, if past observations are clouded by measurement error, accurate predictions at long lead times will be hard to achieve. Data assimilation is the process of improving on these "noisy" observations by attempting to recover the true trajectory. Shadowing lengths measure how long the model trajectory stays close to reality, for a good model prediction of the future, clearly we desire that these lengths should be as long as possible. We show how we can use ratios of these values to compare models and in particular, assimilation techniques.

2 Shadowing

A shadowing length is time a trajectory stays consistent with the observational noise of future observations. For forecasting purposes it is desirable for these lengths of time to be as long as possible.

definition 1 *A model trajectory shadows for time τ_i if it is consistent with the observational uncertainty at all times t, $0 \le t \le \tau_i$.*

The conditions of a trajectory shadowing observations depend on the observational noise model which we assume to be known. For bounded noise, these condtions are straightforward, for example if the noise model is distributed as $U(-\epsilon, \epsilon)$, the noise is bounded by ϵ. Therefore, a trajectory stops shadowing the observations at the smallest value of i such that $|x_i - S_i| > \epsilon$ where x_i and S_i are points from our model trajectory and the 'noisy' observations respectively.

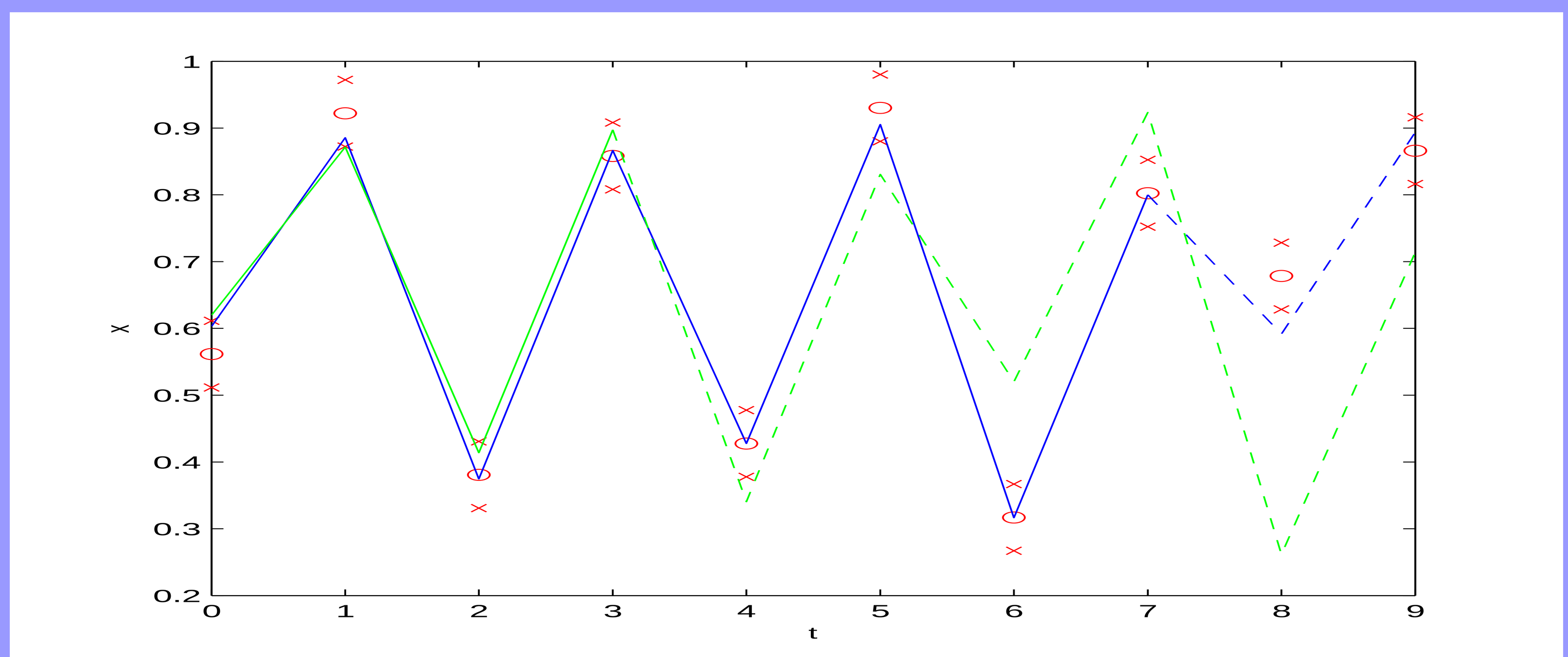

Figure 1: Two shadowing trajectories of the logistic map, the red lines are 'noisy' observations with noise model $U(-0.05, 0.05)$, the distance between the crosses is where a trajectory is said to shadow the point. The blue line shadows for 7 points and the green line shadows for 3 points. Where the lines are solid they shadow the observations and when they are dashed, they don't.

3 Shadowing Ratios

We introduce *shadowing ratios* as a way of comparing assimilation techniques. When choosing an assimilation method, it is desirable that on average at least, the resulting trajectory found from feeding initial conditions from our 'new' method, should be able to shadow longer than the trajectory from some reference method. Shadowing ratios give us an idea of the magnitude of these differences.

definition 2 *We define a shadowing ratio as*

$$\gamma = \frac{\tau_{new}}{\bar{\tau}_{ref}} \qquad (1)$$

where τ_{new} is the shadowing length of an individual set of initial conditions obtained using our 'new' assimilation method and τ_{ref} is the mean or median shadowing length for initial conditions assimilated using some reference method.

When on average, $\gamma > 1$, we have an improvement on our reference method. We define each shadowing ratio to correspond to the shadowing length of one set of assimilated noise realisations divided by the mean or median shadowing lengths of the reference assimilation method, that way we can look at distributions of shadowing ratios. Since shadowing lengths are likely to be skewed in distribution, using the mean or the median shadowing length of the reference method as the denominator changes the results fairly significantly. Each tells a slightly different story, so it is often useful to look at both.

4 Gradient Descent of Indeterminism

One particular method of data assimilation is Gradient Descent of Indeterminism (GDI). The algorithm uses the well known optimisation method, 'gradient descent' to reduce the indeterminism (the distance $||x_{i+1} - F(x_i)||$) of a pseudo-orbit. Iterating repeatedly, reduces the indeterminism and moves the points asymptotically towards a system trajectory. The resulting pseudo-orbit can give us an estimate of past and current conditions. For more information on Gradient Descent of Indeterminism, see E Suckling and L. A. Smith (Gradient Descent for the Point Vortex model).

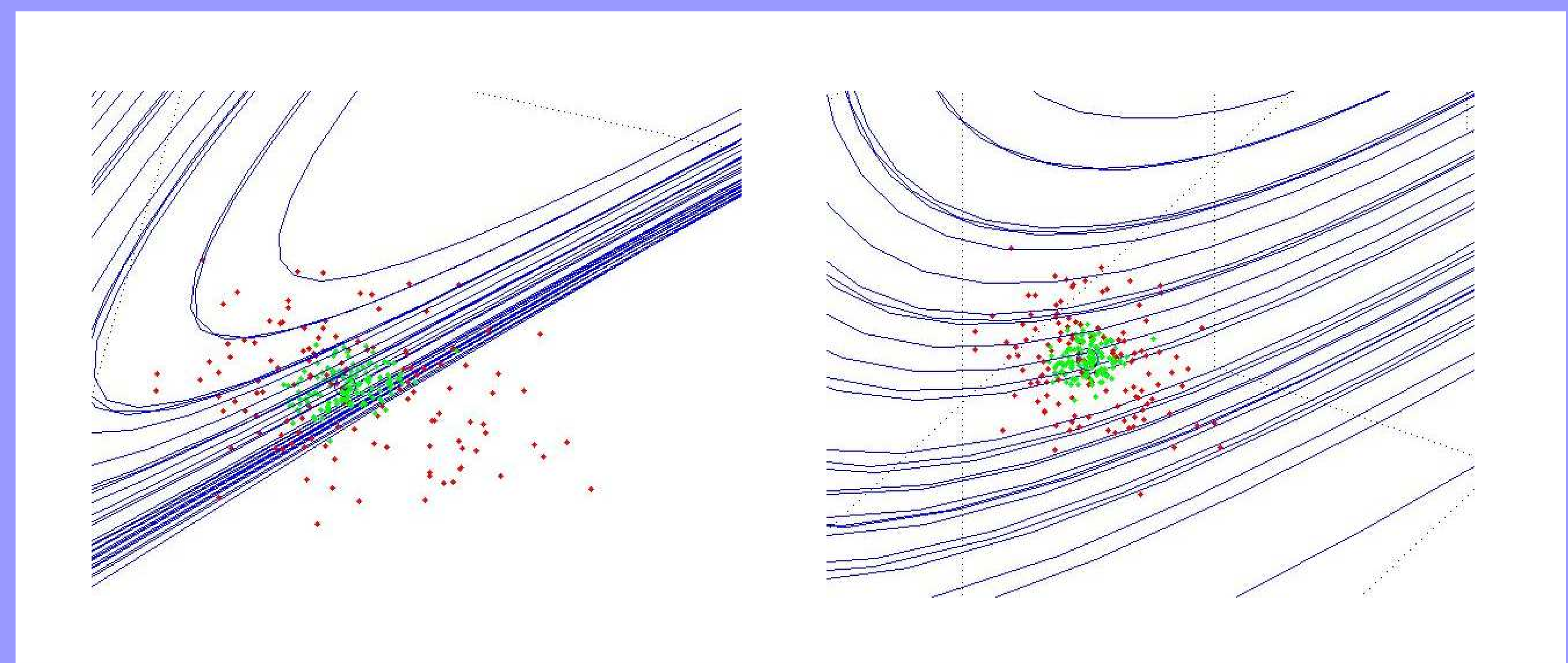

Figure 2: The midpoint of the assimilation window for the Lorenz '63 system of equations showing noisy observations(red) and results of assimilation (green), the 'truth' is the centre of the black circle. Both plots show the same points but from a different perspective.

5 Examples

The gradient descent algorithm uses derivative information to attempt 'noise reduction'. This is not always possible to do analytically, in these cases we can use some approximate method to estimate the jacobian matrix, in this case a forward difference method. We can use shadowing ratios to assess the impact of this substitution. Figure 3 shows the shadowing ratios for GDI applied to the Duffing Map with a=2.75 and b=0.2 and white trimmed gaussian noise added with $\sigma = 0.05$, trimmed at 3 standard deviations. Our reference method here is using no assimilation and running the model from our observations. It is hard to distinguish between the two sets of shadowing ratios and we can conclude that a lack of gradient information makes little difference to Gradient Descent of Indeterminism.

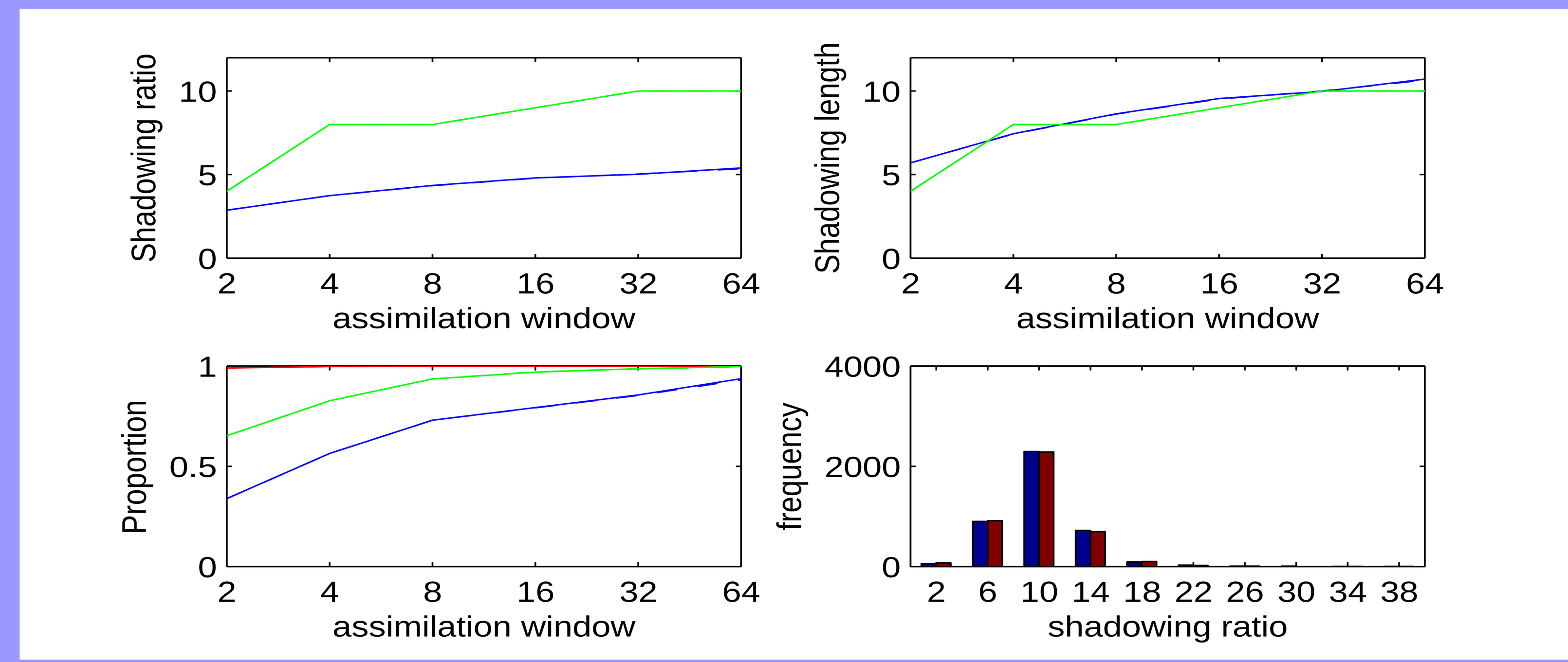

Figure 3: Top left: mean (blue) and median (green) shadowing ratios after assimilating using GDI with exact (solid line) and approximate (dashed line) jacobian. Top right: The same as top left but showing shadowing lengths instead of shadowing ratios. Bottom left: Proportion of median shadowing ratios greater than 2 (red), 4 (green) and 8 (blue). Bottom right: Histogram of median ratios using exact(blue) and approximate (red) jacobian.

Figure 4 shows shadowing ratios for a perfect and an imperfect model of the Duffing map with the same parameter values as figure 4. Again, for our reference method we have used no assimilation. The difference here is much more obvious and we can conclude that Gradient descent has a much larger effect on shadowing lengths for the perfect model than the imperfect model.

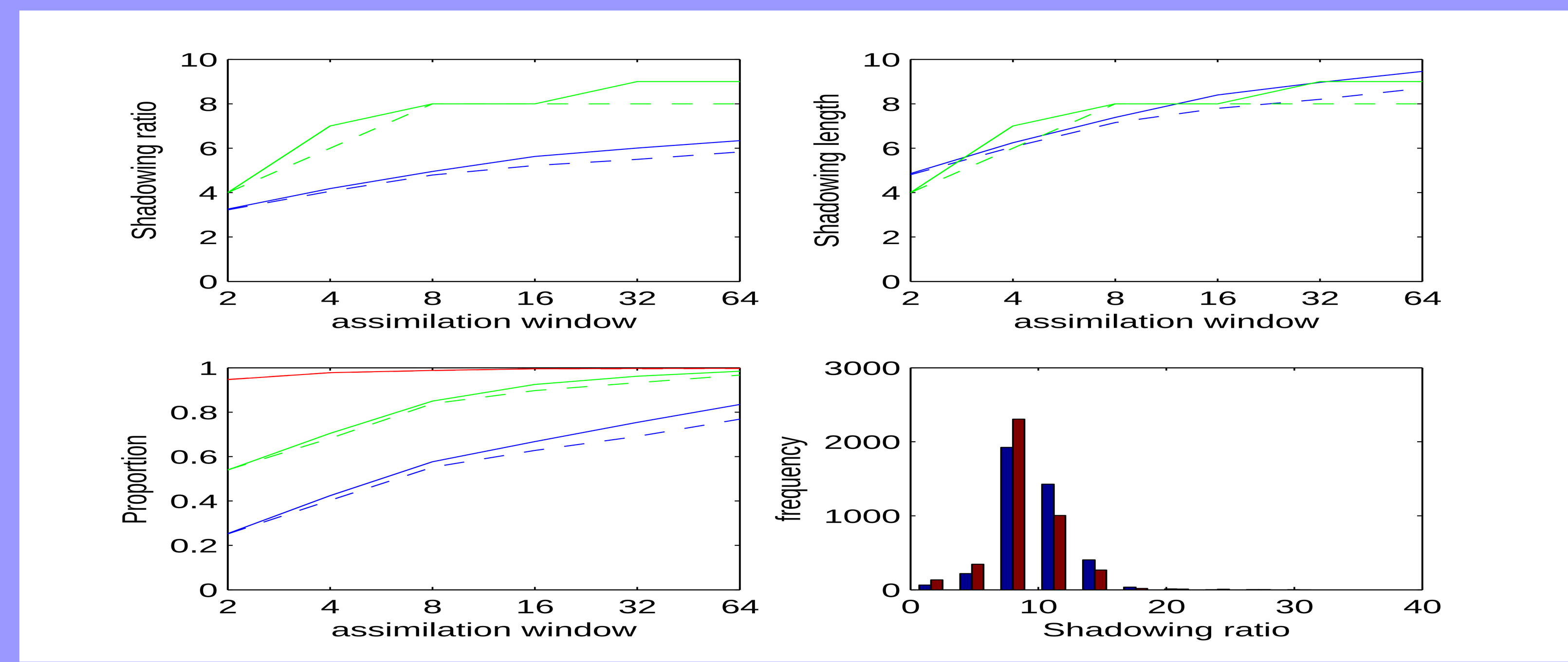

Figure 4: Top left: mean (blue) and median (green) shadowing ratios for a perfect model (solid line) and an imperfect model (dashed) of the Duffing Map. Top right: The same as top left but showing shadowing lengths instead of shadowing ratios. Bottom left: Proportion of median shadowing ratios greater than 2 (red), 4 (green) and 8 (blue). Bottom right: Histogram of median ratios with perfect(blue) and imperfect (red) models.

6 Conclusion

In the context of forecasting, the choice of data assimilation is not a trivial choice, shadowing ratios provide us with a logical way of comparing such techniques in that they will alway choose the one which makes the model trajectory close to reality for as long as possible.

References

[1] L. A. Smith. Disentangling uncertainty and error: on the predictability of nonlinear systems. In: Mees, Alistair I., (ed.) Nonlinear dynamics and statistics. Birkhauser, Boston, pp. (2001) 31-64. ISBN 9780817641634

[2] K. Judd, L.A Smith. Indistinguishable States 1. Perfect Model Scenario. Physica D 151 (2001) 125-141.

Acknowledgements: The support of NERC is gratefully acknowledged.