Using shadowing ratios to evaluate data assimilation techniques

E. Wheatcroft1 and L. A. Smith1,2
1Centre for the Analysis of Time Series, London School of Economics
2Oxford Centre for Industrial and Applied Mathematics, Oxford University, U.K.
Email: e.d.wheatcroft@lse.ac.uk

Abstract

Identifying successful "noise reduction" as well as creating a challenge in applications where the "true" values are not linear or (linear) noise reduction is a quite a true task. We suggest shadowing ratios as a measure of noise reduction when the task at hand involves prediction. Initial condition uncertainty will now or later always limit the long-time of a chaotic model, even when that model reproduces the system dynamics perfectly. Since in reality, observations from each system 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 estimate, we introduce a new measure which allows us to estimate the quality of these techniques. A model trajectory shadows for a long time if it is consistent with the noise model of the observed state. We define the shadowing ratio as the ratio of the lengths of time the model shadowed the assimilated initial conditions to the length of time the model shadowed 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 derivatives and when approximating using a forward difference technique.

1 Introduction

Successful prediction of complex processes has a large 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 prediction of 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 ratios 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 in time is 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 Model trajectory shadows for time \( t_i \) if it is consistent with the observational uncertainty at all times \( t, 0 \leq t \leq t_i \). The conditions of a trajectory shadowing observations depend on the observational noise model which we assume to be known. For bounded noise, these conditions are straightforward; for example if the noise model is distributed as \( N(0, \sigma) \) the noise is bounded by \( \sigma \). Therefore, a trajectory stays shadowing the observations at the smallest value of such that \(|x_t - S_t| < \varepsilon\) where \( S_t \) and \( x_t \) are points from our model trajectory and the "noisy" observations respectively.

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

\[
\gamma = \frac{t_{\text{shadow}}}{t_{\text{ref}}}
\]

where \( t_{\text{shadow}} \) is the shadowing length of an individual set of initial conditions obtained using our "new" assimilation method and \( t_{\text{ref}} \) is the mean or median shadowing length for initial conditions assimilated using some reference method.

When on average, \( \gamma \geq 1 \), 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 significantly. Each with 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 \(|x_{t+\tau} - F(x_t)|\)) of a pseudo-orbit. Iterating repeatedly, this 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 Stanford and L. A. Smith (Gradient Descent for the Point Vertex model).

**Example**

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 for 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 \( \alpha = 2.75 \) and \( b = 0.1 \) while trimmed gaussian noise added with \( \varepsilon = 0.05 \), trimmed at \( 3 \sigma \) 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.

**Conclusions**

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 always choose the one which makes the model trajectory close to reality for as long as possible.

**References**


**Acknowledgements**: The support of NERC is greatly acknowledged.