Nonlinear Processes
in Geophysics
Discussions

Open Access

Precision variational approximations in statistical data assimilation

J. Ye[1], N. Kadakia[1], P. J. Rozdeba[1], H. D. I. Abarbanel[1,2], and J. C. Quinn[3]

[1]Department of Physics, University of California, San Diego, La Jolla, CA 92093-0374, USA
[2]Marine Physical Laboratory (Scripps Institution of Oceanography), University of California, San Diego, La Jolla, CA 92093-0374, USA
[3]Intellisis Corporation, 10350 Science Center Drive, Suite 140, San Diego, CA 92121, USA

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

NPGD

1, 1603–1620, 2014

Variational approximations in statistical data assimilation

J. Ye et al.

Title Page

Abstract · Introduction

Conclusions · References

Tables · Figures

◄ · ►

◄ · ►

Back · Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

Data assimilation transfers information from observations of a complex system to physically-based system models with state variables $x(t)$. Typically, the observations are noisy, the model has errors, and the initial state of the model is uncertain, so the data assimilation is statistical. One can thus ask questions about expected values of functions $\langle G(X) \rangle$ on the path $X = \{x(t_0), \ldots, x(t_m)\}$ of the model as it moves through an observation window where measurements are made at times $\{t_0, \ldots, t_m\}$. The probability distribution on the path $P(X) = \exp[-A_0(X)]$ determines these expected values. Variational methods seeking extrema of the "action" $A_0(X)$, widely known as 4DVar (Talagrand and Courtier, 1987; Evensen, 2009), are widespread for estimating $\langle G(X) \rangle$ in many fields of science. In a path integral formulation of statistical data assimilation, we consider variational approximations in a standard realization of the action where measurement and model errors are Gaussian. We (a) discuss an annealing method for locating the path X^0 giving a consistent global minimum of the action $A_0(X^0)$, (b) consider the explicit role of the number of measurements at each measurement time in determining $A_0(X^0)$, and (c) identify a parameter regime for the scale of model errors which allows X^0 to give a precise estimate of $\langle G(X^0) \rangle$ with computable, small higher order corrections.

1 Introduction

In a broad spectrum of scientific fields, transferring information in L observed data time series $y_l(t_n)$; $l = 1, \ldots, L$; $n = 0, \ldots, m$ to a physically-based model of the processes producing those observations allows estimation of unknown parameters and unobserved states of the model within an observation window $\{t_0, \ldots, t_m\}$. As a sample of these fields we note applications in meteorology (Talagrand and Courtier, 1987; Evensen, 2009; Lorenc and Payne, 2007), geochemistry (Eibern and Schmidt, 1999),

fluid dynamics (Zadeh, 2008) and plasma physics (Mechhoud et al., 2013), among many others.

The probability distribution $P(X) = \exp[-A_0(X)]$ allows evaluation of expected values of physically interesting functions $G(X)$ along the path. $A_0(X)$ has terms giving a measurement's influence on $P(X)$, terms propagating the model between the measurement times t_n, and a term $-\log[P(x(0))]$ representing the uncertainty in the initial state (Abarbanel, 2013).

We discuss the familiar case where measurement errors are independent at each t_n and Gaussian with covariance $R_m^{-1}(l, l', t)$; $l, l' = 1, \ldots, L$.

The model is a physical differential equation, discretized in space and time and satisfying the D-dimensional stochastic discrete time map $x_a(n + 1) = f_a(x(n)) + R_f^{-1/2}(a, b)\eta_b(n)$; $a, b = 1, \ldots, D$ and Gaussian noise error $\eta_a(n) \sim \mathcal{N}(0, 1)$. We take $R_m(l, l', n) = R_m(n)\delta_{l,l'}$ and $R_f(a, b) = R_f\delta_{a,b}$. $R_m(n)$ is zero except near observation times t_n.

With these conditions the action takes a familiar form

$$A_0(X) = \sum_{n=0}^{m} \frac{R_m(n)}{2} \sum_{l=1}^{L} [x_l(n) - y_l(n)]^2$$

$$+ \frac{R_f}{2} \sum_{n=0}^{m-1} \sum_{a=1}^{D} [x_a(n + 1) - f_a(x(n))]^2 - \log[P(x(0))]. \tag{1}$$

The expected value $\langle G(X) \rangle$ of a function $G(X)$ is

$$\langle G(X) \rangle = \frac{\int dX\, G(X) \exp[-A_0(X)]}{\int dX \exp[-A_0(X)]}. \tag{2}$$

One interesting function $G(X)$ is X itself. Its expected value gives us the average path over the measurement window $[t_0, t_m = T]$. Estimates of the parameters and $P(x(T))$ permit prediction for $t > T$.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

To approximate the integral $\langle G(X) \rangle$ we follow the stationary path method of Laplace (Laplace, 1774) and seek extremum paths X^q; labeled by $q = 0, 1 \ldots$ satisfying $\partial A_0(X)/\partial X_\alpha = 0$. The index α is a label for the state and time $\alpha = \{a, n\}$. Paths are sorted by increasing action levels $A_0(X^q)$: $A_0(X^0) \leq A_0(X^1) \leq \cdots$.

2 Evaluating $\langle G(X) \rangle$

We take the the usual data assimilation technique (Talagrand and Courtier, 1987; Evensen, 2009; Lorenc and Payne, 2007) several steps further by

1. showing how to find a consistent global minimum of the action using an annealing method,

2. demonstrating the importance of the number of measurements at each observation time, and

3. explaining how to make systematic perturbative corrections to the value of $\langle G(X^0) \rangle$.

In the neighborhood of X^q, we expand $A_0(X)$ as

$$A_0(X) = A_0(X^q) + (X - X^q)_{\alpha_1} \gamma^2_{\alpha_1 \alpha_2}(X^q)(X - X^q)_{\alpha_2}$$

$$+ \sum_{r=3} \frac{A^{(r)}(X^q)_{\alpha_1 \ldots \alpha_r}}{r!}(X - X^q)_{\alpha_1} \ldots (X - X^q)_{\alpha_r}. \tag{3}$$

Changing variables to $U_\alpha = \gamma_{\alpha\beta}(X^q)(X - X^q)_\beta$ leads to a contribution to the numerator of $\langle G(X) \rangle$ in Eq. (2) from each X^q

$$\frac{\exp\left[-A_0(X^q)\right]}{\det\gamma(X^q)} \int dU \exp\left(-U^2 - V(X^q)\right)$$

$$\left[G(X^q) + W(X^q)\right] \tag{4}$$

where

$$V(X^q) = \sum_{r=3} \frac{A^{(r)}(X^q)_{\alpha_1 \ldots \alpha_r}}{r!}$$
$$(\gamma(X^q)^{-1}U)_{\alpha_1} \cdots (\gamma(X^q)^{-1}U)_{\alpha_r},$$

$$W(X^q) = \sum_{k=1} \frac{G^{(k)}(X^q)_{\alpha_1 \ldots \alpha_k}}{k!}$$
$$(\gamma(X^q)^{-1}U)_{\alpha_1} \cdots (\gamma(X^q)^{-1}U)_{\alpha_k}.$$

In the denominator of Eq. (2), $G(X)$ is replaced by unity. We sum over the contribution of each X^q to evaluate $\langle G(X) \rangle$.

If the lowest action level $A_0(X^0)$ is much smaller than all others, it totally dominates the integral, and

$$\langle G(X^0) \rangle = \frac{\int d\mathbf{U} \exp\left(-\mathbf{U}^2 - V(X^0)\right) [G(X^0) + W(X^0)]}{\int d\mathbf{U} \exp\left(-\mathbf{U}^2 - V(X^0)\right]}, \tag{5}$$

plus exponentially small corrections.

3 Annealing to find a consistent global minimum action level $A_0(X^0)$

We now turn to an annealing method to find the path X^0, and within this method examine the importance of the number L of measurements at each observation time t_n. We then present an argument that in the integral Eq. (4) the terms $V(X^q), W(X^q)$ behave as inverse powers of R_f as R_f/R_m becomes large. This would leave us with $\langle G(X) \rangle \approx G(X^0)$ with corrections which are a power series in R_f^{-1}. The implication is that all statistical data assimilation questions, as R_f/R_m becomes large, can be approximated by the contribution of the path X^0 with the lowest action level $A_0(X^0)$ along

with corrections one can evaluate via standard perturbation theory. Variations about $\langle G(\boldsymbol{X^0})\rangle$ would be small and computable.

The term $P(\boldsymbol{x}(0))$ in the action is often written assuming Gaussian variation about some base state $\boldsymbol{x}_{\text{base}}$, so $-\log[P(\boldsymbol{x}(0)] \propto (\boldsymbol{x}(0) - \boldsymbol{x}_{\text{base}})^2 R_{\text{base}}/2$, and this has the form of the measurement term evaluated at $n = 0$. We incorporate this expression into the term with coefficient R_m in the action and no longer display it.

The annealing method is based on the observation (Quinn, 2010) that the equation for the extrema $\boldsymbol{X^q}$ simplifies at $R_f = 0$. The solution is $x_l(n) = y_l(n)$; $l = 1, 2, \ldots, L$ for all observation times. The other $(D - L)$ components of the model state vector are undetermined, and the solution is degenerate. As we increase R_f, the action levels split, and depending on R_m, R_f, L and the precise form of the dynamical vector field $\boldsymbol{f}(\boldsymbol{x})$, there will be $1, 2, \ldots$ minima of $A_0(\boldsymbol{X})$.

As R_f/R_m grows large, the errors in the model diminish relative to the measurement errors and impose high precision $\boldsymbol{x}(n+1) \approx \boldsymbol{f}(\boldsymbol{x}(n))$ on the model dynamics. The noise in the measurement has been taken to be Gaussian $\mathcal{N}(0, \sigma^2)$, so the measurement error term in the action satisfies a χ^2 distribution with mean $R_m \sigma^2 (m+1) L/2$ and standard deviation $R_m \sigma^2 \sqrt{(m+1)L/2}$.

The action level for large R_f approximately approaches a lowest value

$$A_0(\boldsymbol{X^0}) \rightarrow \frac{R_m \sigma^2}{2} L(m+1) \left[1 \pm \frac{1}{\sqrt{(m+1)L/2}} \right]. \qquad (6)$$

σ^2 is the noise level in $y_l(n)$ and $(m+1)$ is the number of measurement times t_n. This provides a consistency condition on the identification of the path $\boldsymbol{X^0}$ by giving a consistent global minimum action value $A_0(\boldsymbol{X^0})$.

If the action levels revealed by our annealing procedure do not give this, it is a sign that the data is inconsistent with the model.

3.1 Annealing procedure

The annealing process proceeds as follows: with very small initial R_f, we call it R_{f0}, solve the $(m+1)D$-dimensional search problem with an optimization algorithm that seeks minima of $A_0(X)$. Start the search with a set of trial paths whose components are selected from a uniform distribution within limits suggested by examining the times series generated by the model $x \to f(x)$ (or any other selection process for the initial guess). This will generate a collection of approximate paths X^q. Increase R_f by a small increment (we choose $R_f = \{R_{f0}2^\beta\}$, where $\beta = 0, 1, \dots$), and using the paths found for the previous R_f as initial guesses, find a new set of approximate X^q.

Continue this process until the lowest action level path X^0 produces an $A_0(X^0)$ near Eq. (6). If the action level $A_0(X^0)$ is substantially less than the action level on the next path $A_0(X^0) \ll A_0(X^1)$, all expected values $\langle G(X) \rangle$ are given by X^0, and corrections due to fluctuations about that path. The contribution to the expected values of the path X^1 with the next action level is exponentially small, of order $\exp[-(A_0(X^1) - A_0(X^0))]$.

The annealing procedure discussed above is different from standard simulated annealing (Aguiar e Oliviera et al., 2012). We call this annealing because varying R_f is similar to varying a temperature in a statistical physics problem where R_f is inversely proportional to the temperature. At high temperatures, small R_f, the dynamics among the degrees of freedom $x_a(n)$ is essentially irrelevant, and we have a universal solution where the observed degrees of freedom match the observations. As the temperature is decreased, the dynamics plays a more and more significant role, and allowed paths "freeze" out. Action levels play the role of energy levels in statistical physics, and as the action is positive definite, the paths are directly analogous to instantons in the Euclidian field theory represented by $A_0(X)$ (Zinn-Justin, 2002).

4 Examples from the Lorenz96 Model

We now present the results of a set of calculations on the Lorenz96 (Lorenz, 2006) model used frequently in geophysical data assimilation discussions as a testbed for proposed methods. The model has D degrees of freedom $x_a(t)$ satisfying the differential equation

$$\dot{x}_a(t) = x_{a-1}(t)(x_{a+1}(t) - x_{a-2}(t)) - x_a(t) + f; \tag{7}$$

$a = 1, \ldots, D$; $x_{-1}(t) = x_{D-1}(t)$, $x_0(t) = x_D(t)$, $x_{D+1}(t) = x_1(t)$; $f = 8.17$, for which the $x_a(t)$ are chaotic (Kostuk, 2012). We studied $D = 20$ and added f as an additional degree of freedom satisfying $\dot{f} = 0$.

We performed a twin experiment in which we solved these equations with an arbitrary choice of initial conditions using a fourth order Runga–Kutta solver with $\Delta t = 0.025$ over 160 steps in time. Here, $t_0 = 0$ and $t_m = T = 4$. We then added iid Gaussian noise with zero mean and variance $\sigma^2 = 0.25$ to each time series. $L = 1, 2, \ldots$ of the data series were then represented in the action at each measurement time t_n during our annealing procedure.

In the action we selected $R_m = 4$, the inverse variance of the noise added to the data in our twin experiment, so the minimum action level we expect is $80.5L \pm 8.97\sqrt{L}$. The paths are $(m + 1)(D + 1) = 3381$-dimensional. Our search for minimum paths used a BFGS routine (Press et al., 2012) to which we provided an analytical form of the gradient of $A_0(\boldsymbol{X})$. The search was initialized with 100 times with initial paths from a uniform distribution of values in the interval [-10,10].

In Fig. 1 we display the computed action levels for $L = 5, 7, 8$ and 9. For $L = 5$ there are many close action levels associated with the extremum paths of the action Eq. (1); as L increases, the lowest action level visibly separates from the others. At the bottom of each panel, we indicate the lowest action level value and its standard deviation. The next-lowest action level $A_0(\boldsymbol{X^1})$, for $L = 5, 7, 8$, and 9, is at 403.8, 749.8, 1161.6, and 2256.1, respectively. This means for $L = 5$ we would have to sum over the contributions

of many paths X^q to evaluate the expected value $\langle G(X) \rangle$. At $L = 7$ or higher, X^0 dominates the integral. Our estimate for the forcing parameter, set to 8.17, was 8.22 at large β.

It is important to note that if we begin our search for the extremum states X^q at large values of R_f we are almost sure to miss the actual path X^0 which gives the lowest action level, since the Hessian matrix of A_0 is ill-conditioned when R_f is large. See Fig. 4.6 in (Quinn, 2010).

The real test of an estimation procedure data assimilation is not accuracy in the estimation, but accuracy in prediction beyond the observation window. As this is a twin experiment, we show in Fig. 2 the data, the estimated state variable, and the predicted state variable for an an observed variable $x_3(t)$ and for an unobserved variable $x_{12}(t)$ for $L = 8$. In a real experiment, we could not compare our estimates for the parameters or the unobserved state variables. Although the estimation procedure for the path X^0 with the minimum action value is rather good, in estimating 12 unobserved states and one parameter, there are, of course, errors in our knowledge of the full state $x(t = 4)$. The predictions lose their accuracy in time because of the chaotic nature of the trajectories at $f = 8.17$.

To see how well our procedure works for several unknown parameters, we introduced 10 different forcing parameters f_a into the Lorenz96 model at $D = 10$: $\dot{x}_a(t) = x_{a-1}(t)(x_{a+1}(t) - x_{a-2}(t)) - x_a(t) + f_a$. For $D = 10$, the lowest action level stands out from the rest at $L = 4$. In Table 1 we show our estimates for the ten forcing parameters for $L = 4$, 5, and 6, as well as the actual value used in the calculations of the data. In these estimates, and for the single forcing parameter reported above for $D = 20$, there is a known source of bias (Kostuk et al., 2012). As one can see in the examples it is small here.

To give some sense of what one might expect if the model were `totally wrong`, we presented data from a collection of 1963 Lorenz model (Lorenz, 1963) oscillators oscillators to a Lorenz96 $D = 10$ model.

Twelve time series data are generated by four individual Lorenz63 (Lorenz, 1963) systems with different initial conditions. Gaussian white noise with zero mean and standard deviation $\sigma = 0.5$ are added to each time series. All these "data" $y_l(t)$ are rescaled to lie within $[-10, 10]$.

We then place these signals as "data" in the action with the model taken as Lorenz96 $D = 10$, the single forcing parameter is treated as a time-dependent state variable obeying $\dot{f} = 0$. We use $L = 5$ and 6 as measurements using the data time series taken in the order $y_1(t), y_3(t), y_5(t), y_7(t), y_9(t), y_2(t), y_4(t)$ for the two cases. In Fig. 3 we display the action levels associated with this for $L = 5$ and 6. Results for other values of L are consistent with these. This example provides a graphic illustration of the inconsistency of the data and the model and how this makes its appearance in the annealing procedure.

We now have seen that a consistent global minimum action level can be identified via an annealing process and the dependence of the action levels on the number of measurements has been demonstrated. We use "consistent" as we have no formal proof it is a global minimum, only the consistency condition Eq. (6).

5 Corrections to the contribution of the extremum path to $\langle G(X) \rangle$

We turn back to the evaluation of the path integral for $\langle G(X) \rangle$. In that integral for our action we have

$$\gamma^2_{\alpha\beta}(X) = R_m(n)\delta_{a,l}\delta_{b,l'} + R_f h_{\alpha\beta}(X)$$

$$A_0^{(r)}(X) = R_f g^{(r-2)}(X),$$

for $r \geq 3$. The functions $h(X)$ and $g(X)$ are derivable from the form of $A_0(X)$. These are to be evaluated at $X = X^q$ for the qth extremum path.

In the form of the integral Eq. (4) we see that the term in the exponential with order r in U has order of magnitude about $R_f/(R_m + R_f h(X^q))^{r/2}$ for $r \geq 3$. Similarly in the expansion of $G(X)$ the term of order k has a denominator of order $1/(R_m + R_f h(X^q))^{k/2}$ for

Full Screen / Esc

$k \geq 1$. Taking into account that only terms with even powers of **U** are nonzero in the integral because of symmetry, we have a collection of terms which, as R_f becomes large with respect to R_m, decrease as $1/R_f$ or faster. As a rule of thumb in the calculations we presented for Lorenz96 $D = 20$, we see that for $\beta = 15$ or larger, or $R_f/R_m \approx 100$ or larger, the terms in the path integral beyond the quadratic term in the expansion of the action become small. A stronger estimate would come from evaluating the leading term in $1/R_f$.

In summary, we have examined the path integral formulation of data assimilation and asked how well a variational approximation, which looks for Laplace like estimates of the path integral by seeking extremum paths of the action, works to evaluate the expected values of functions on the path. This approximation is widely used in meteorology and other fields where it is known as weak 4DVar (Talagrand and Courtier, 1987; Evensen, 2009; Lorenc and Payne, 2007). The action is the cost function which is minimized.

In previous work with variational principles for data assimilation, we are unaware of any procedure such as our annealing method using R_f to identify a consistent global minimum action. Nor are we aware of a systematic exploration of the dependence of the action levels on the number of measurements at each observation time. (Of course, there is the discussion in Quinn, 2010, which generated this work.) Finally, we do not know of a discussion of the corrections to the variational approximation, which here is shown to consist of small perturbations when the resolution of the model error term in the action is increased, namely R_f becomes large.

6 Conclusions

We have worked within a framework where the measurement errors and the model errors in the data assimilation are Gaussian, with the inverse covariance of the model errors taken to be of order R_f. We have shown that the path which gives a consistent global minimum action can be traced by an annealing procedure starting with a setting

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

where the dynamical model essentially plays no role, $R_f \approx 0$, then systematically increasing the influence of the model dynamics. As the scale of the model error reaches about 100 times the scale of the measurement error, the path X^0 associated with a consistent global minimum action level dominates the path integral with corrections of order $1/R_f$. The important role of the number of measurements L at each measurement time is also demonstrated.

In this paper we do not address the typical situation where the number of measurements actually available is less than that needed to allow the ground level of the action to lie well below the next level. For a solution to that problem we have used information from the waveform of the measurements as shown in some detail in (Rey et al., 2014).

The results here justify the use of the variational approximation in data assimilation, focus on the role of the number of measurements one requires for accuracy in that approximation, and permits the evaluation of systematic corrections to the approximation. We anticipate that our use of Gaussian model error and measurement error is a convenience and that other distributions of these errors will permit the same set of statements about the value of variational approximations to statistical data assimilation.

References

Abarbanel, H. D.: Predicting the Future: Completing Models of Observed Complex Systems, Springer, New York, 2013. 1605

Aguiar e Oliviera, H., Ingber, L., Petraglia, A., Petraglia, M. R., and Machado, M. A. S.: Stochastic Global Optimization and Its Applications with Fuzzy Adaptive Simulated Annealing, Vol. 35, Springer, New York, 2012. 1609

Eibern, H. and Schmidt, H.: A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, J. Geophys. Res.-Atmos., 104, 18583–18598, 1999. 1604

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer, New York, 2009. 1604, 1606, 1613

Kostuk, M.: Synchronization and statistical methods for the data assimilation of HVc neuron models, Ph.D. thesis in Physics, University of California, San Diego, available at: http://escholarship.org/uc/item/2fh4d086 (last access: 2 October 2014), 2012. 1610

Kostuk, M., Toth, B., Meliza, C., Margoliash, D., and Abarbanel, H.: Dynamical estimation of neuron and network properties II: path integral Monte Carlo methods, Biol. Cybern., 106, 155–167, 2012. 1611

Laplace, P. S.: Memoir on the probability of causes of events, Mémoires de Mathématique et de Physique, Tome Sixième, 1774. 1606

Lorenc, A. C. and Payne, T.: 4D-Var and the butterfly effect: statistical four-dimensional data assimilation for a wide range of scales, Q. J. Roy. Meteor. Soc., 133, 607–614, 2007. 1604, 1606, 1613

Lorenz, E. N.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141, 1963. 1611, 1612, 1620

Lorenz, E. N.: Predictability – a problem partly solved, in: Predictability of Weather and Climate, edited by: Palmer, T. and Hagedorn, R., Cambridge University Press, 40–58, 2006. 1610, 1619

Mechhoud, S., Witrant, E., Dugard, L., and Moreau, D.: Combined distributed parameters and source estimation in tokamak plasma heat transport, in: 2013 European Control Conference (ECC), Zurich, 17–19 July, 47–52, 2013. 1605

Press, W. H., Teokulsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes in C: the Art of Scientific Computing, Cambridge University Press, 2012. 1610

Quinn, J. C.: A path integral approach to data assimilation in stochastic nonlinear systems, Ph.D. thesis in Physics, University of California, San Diego, available at: http://escholarship.org/uc/item/0bm253qk (last access: 2 October 2014), 2010. 1608, 1611, 1613

Rey, D., Eldridge, M., Kostuk, M., Abarbanel, H. D., Schumann-Bischoff, J., and Parlitz, U.: Accurate state and parameter estimation in nonlinear systems with sparse observations, Phys. Lett. A, 378, 869–873, 2014. 1614

Talagrand, O. and Courtier, P.: Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. I: Theory, Q. J. Roy. Meteor. Soc., 113, 1311–1328, 1987. 1604, 1606, 1613

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Zadeh, K. S.: Parameter estimation in flow through partially saturated porous materials, J. Comput. Phys., 227, 10243–10262, 2008. 1605

Zinn-Justin, J.: Quantum Field Theory and Critical Phenomena, Oxford University Press, 2002. 1609

Table 1. Known and estimated forcing parameters for the Lorenz96 Model at $D = 10$, $L = 4, 5,$ and 6.

Known f_a	$L = 4$	$L = 5$	$L = 6$
5.7	5.742	5.737	5.768
7.1	7.096	7.080	7.094
9.6	9.696	9.686	9.654
6.2	6.156	6.174	6.131
7.5	7.605	7.592	7.604
8.4	8.353	8.330	8.349
5.3	5.310	5.278	5.214
9.7	9.679	9.703	9.643
8.5	8.632	8.629	8.626
6.3	6.334	6.336	6.308
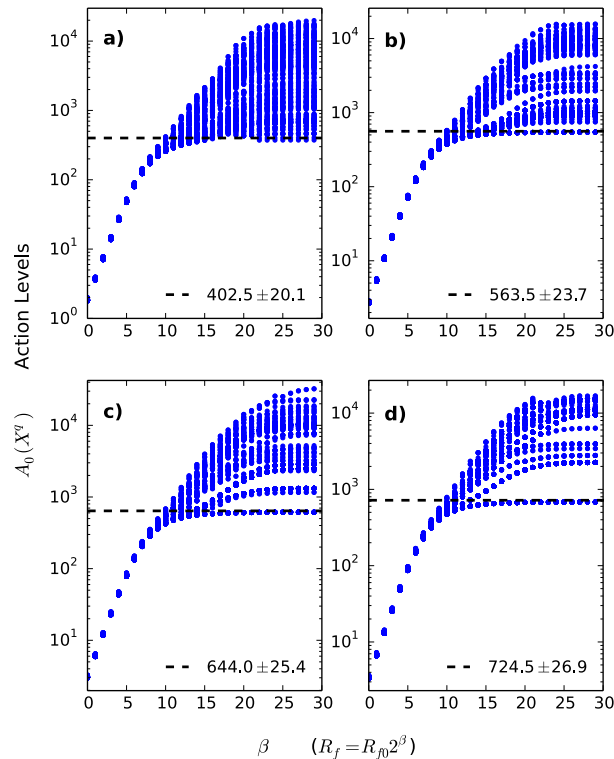
Figure 1. Action levels as a function of R_f for Lorenz96 model, $D = 20$, $R_{f0} = 0.01$. **(a)** At $L = 5$ we used $y_1(t)$, $y_3(t)$, $y_5(t)$, $y_7(t)$, $y_9(t)$ in the action; **(b)** at $L = 7$, $y_{11}(t)$, $y_{13}(t)$ are added; **(c)** at $L = 8$, $y_{15}(t)$ is added; **(d)** at $L = 9$, $y_{17}(t)$ is added. The expected values of the lowest action level are denoted by black dashed lines.

NPGD

1, 1603–1620, 2014

Variational
approximations in
statistical data
assimilation

J. Ye et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

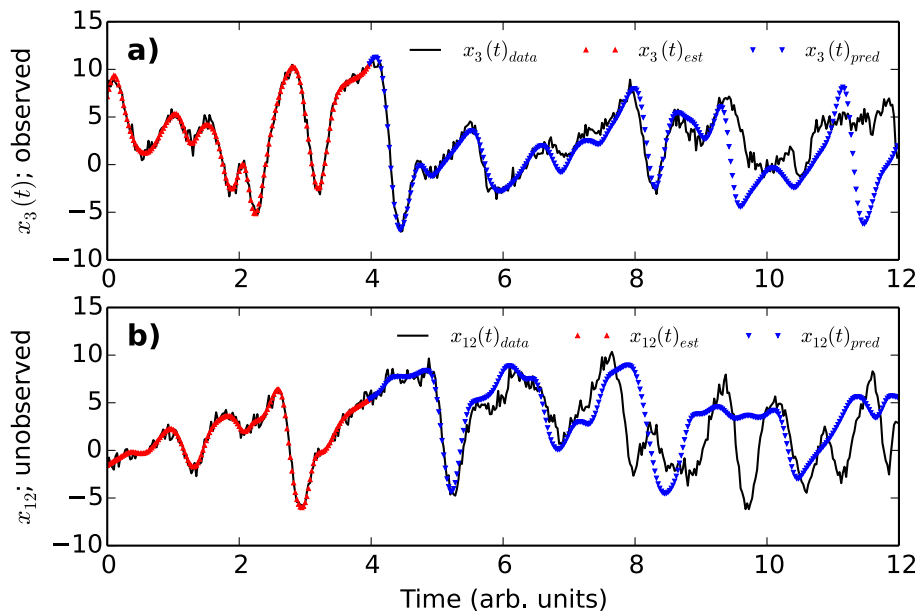Printer-friendly Version

Interactive Discussion

Figure 2. Data, estimated, and predicted time series for the Lorenz96 model (Lorenz, 2006) with $D = 20$, $L = 8$. **(a)** $x_3(t)$ was an observed state variable, and **(b)** $x_{12}(t)$ was unobserved. The data (black) the estimated state variable (red) and the predicted state variable (blue) are shown for each of them.
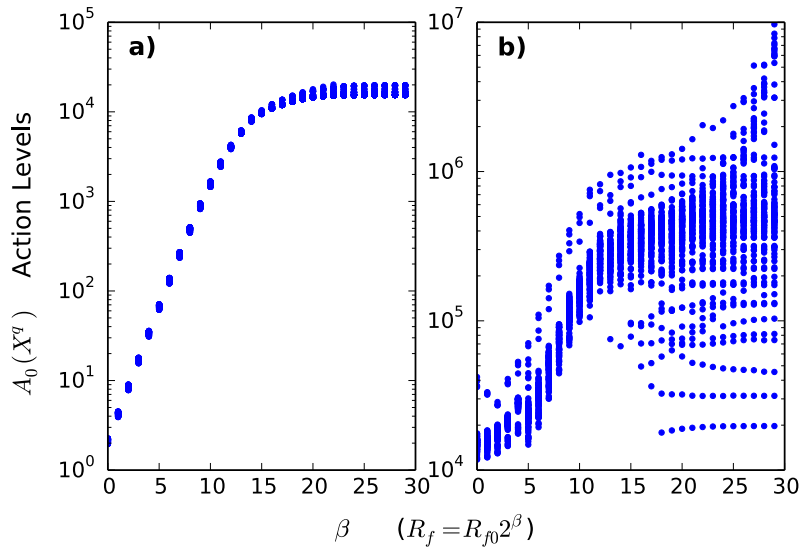
Figure 3. Action levels as a function of R_f for Lorenz96 model, $D = 10$, $R_{f0} = 0.01$. **(a)** $L = 5$ and **(b)** $L = 6$ when the wrong data is used for the Lorenz96 model. We actually used data from four realizations of the Lorenz63 model (Lorenz, 1963). The structure of the action levels vs. R_f shows no trace of the minimum allowed level Eq. (6). This indicates the data and the model are incompatible. The action levels are also quite large, and, for $L = 6$, numerous and not well separated.