JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. C4, PAGES 8015–8031, APRIL 15, 1994

Interactive Kalman filtering

Gerd Bürger and Mark A. Cane
Lamont-Doherty Earth Observatory, Palisades, New York

Abstract. Data assimilation via the extended Kalman filter can become problematic when the assimilating model is strongly nonlinear, primarily in connection with sharp, "switchlike" changes between different regimes of the system. The filter seems too inert to follow those switches quickly enough, a fact that can lead to a complete failure when the switches occur often enough. In this paper we replace the key feature of the filter, the use of local linearity for the error model update, with a principle that uses a more global approach through the utilization of a set of preselected regimes. The method uses all regime error models simultaneously. Being mutually incompatible, a compromise between the different error models is found through the use of a weighting function that reflects the 'closeness' of the error model to the correct model. To test the interactive Kalman filter a series of numerical experiments is performed using the double-well system and the well-known Lorenz system, and the results are compared to the extended Kalman filter. It turns out that, depending on the set of preselected regimes, the performance is worse than, comparable to, or better than that of the extended Kalman filter.

1. Introduction

The main feature of the Kalman filtering technique in data assimilation problems is the statistically optimal incorporation of dynamical model information into the assimilation procedure. This is done via a weighting between observation and model information which reflects the error structure of both of them. If the error is large the weighting is small and vice versa. Since, by definition, the error itself is unknown, one can only rely on some of its statistical properties. An optimal description of the error structure and its evolution would require a dynamical model of the complete probability distribution of the error. But this is, theoretically and practically, only treatable in very trivial cases. However, under certain circumstances the error structure becomes "relatively" simple, and it is possible to formulate a correct error model that is also feasible.

The classical paper by *Kalman* [1960] deals with linear processes. In this case it is used to represent the error structure through the error covariance matrix. Since the evolution of the error is the same as that for the original process, this error matrix evolves along with the square of the system matrix. This, however, is computationally very expensive, especially for applications in oceanography or meteorology, where the model size is some orders of magnitude larger than the classical Kalman filter applications in engineering or ballistic fields. To reduce the computational complexity involved in the Kalman filter, a number of methods have been introduced, many of which are reviewed by *Ghil and Malanotte-Rizzoli* [1991]. The extended Kalman filter (EKF) applies to nonlinear systems, [see *Bucy and Joseph*, 1968; *Gelb*, 1974]. The error model still describes the evolution of a covariance, but now the linear model equations are updated each time step according to the current linearization of the original model. Depending on the inherent time scale of the system the updating frequency can be reduced significantly in order to save computing time. For instance, for atmospheric models it is assumed that an update frequency of 12–24 h is sufficient, whereas for ocean models this period can be much longer. First applications of the EKF to nonlinear models look promising. *Budgell* [1986] in oceanography and *Lacarra and Talagrand* [1988] in meteorology have demonstrated that working with a variable error model can produce good results if certain conditions are satisfied. These conditions mainly imply that the system must not undergo sharp changes, a behavior which is typical for regimelike systems, where changes between regimes tend to occur in a switchlike manner. Among this kind are the well-studied systems of *Duffing* [1918], *Lorenz* [1963], and *Charney and DeVore* [1979]. For the Lorenz system and (a simplified version of) *Duffing*'s [1918] system, we shall demonstrate in our paper that certain circumstances can lead to a complete failure of the EKF (see also *Budgell* [1986] and *Miller et al.* [1993]).

By going from systems with smooth changes in the local structure to those with sharp changes one can make the following simple observation: A system that switches sharply from one local structure to another should remain relatively constant, with respect to the local structure, in the intermediate time when it occupies a regime. This leads to two complementary questions for the assimilation problem. First, is it possible for an assimilation to follow the abrupt changes of the regimes? Second, is it possible that the intermediate regime states are sufficiently described by a single linear model? To assimilate regimelike systems, we introduce what we call interactive Kalman filtering (IKF). The IKF makes explicit use of the regimes, in the sense that each regime has its own fixed error model. The interaction between the regimes, which governs the transitions, is achieved by a weighting function which reflects the "distance" of the system from that regime.

To demonstrate the IKF we perform Monte Carlo experiments with the double-well system and the Lorenz system.

Paper number 94JC00148.
0148-0227/94/94JC-00148$05.00

In a set of numerical simulations we assimilate these systems by using various groups of regimes and compare the results to the full EKF assimilation. As it turns out, if the regimes are chosen appropriately, the IKF shows a very good performance in all cases. This is not true for the EKF which in certain cases completely fails to trace the regime switches correctly.

2. Extended Kalman Filtering

One enters the realm of the extended Kalman filter [*Bucy and Joseph*, 1968; *Jazwinski*, 1970] when the error model changes with the current state of the system. In the general setting of Kalman filtering there is a physical variable x, which has to be estimated at some time t. Two sources of information are used for the procedure. One is the more or less steady influx of observational data, and the other is the evolution of a model. Normally, neither source is perfect; both are disturbed by some error noise.

For convenience we assume that the model describes the whole physical space, and observations are available at all times and locations. Denoting the true value of x at t by $x(t)$, the observation by $\hat{x}(t)$, and the model function by f, the two available sources read (as vectors)

$$\hat{x}(t) = x(t) + r(t) \tag{1a}$$

$$f(x(t-1)) = x(t) + q(t) \tag{1b}$$

with observational noise $r(t)$ and system noise $q(t)$. The assimilated value $\bar{x}(t)$ will be determined in two steps via

$$\hat{x}(t) = f(\bar{x}(t-1)) \tag{2a}$$

$$\bar{x}(t) = (\mathbf{I} - \mathbf{W})\hat{x}(t) + \mathbf{W}\hat{x}(t) \tag{2b}$$

that is, a composition of observation and model output, weighted by the matrix $\mathbf{W}$ which still has to be determined ($\mathbf{I}$ denotes the unit matrix). The assimilation error, which is defined to be

$$\varepsilon(t) = \bar{x}(t) - x(t) \tag{3}$$

is the crucial quantity in the assimilation procedure. If we expand f about x,

$$f(x+\varepsilon) - f(x) = \frac{\partial f}{\partial x}(x)\varepsilon + n(\varepsilon) \tag{4}$$

where $n(t)$ comprises the nonlinear parts, we can combine (1)–(3) to find that the error evolution obeys the following law:

$$\hat{\varepsilon}(t) = \mathbf{L}_{t-1}\varepsilon(t-1) + n(\varepsilon(t-1)) + q(t) \tag{5a}$$

$$\varepsilon(t) = (\mathbf{1} - \mathbf{W})\hat{\varepsilon}(t) + \mathbf{W}r(t) \tag{5b}$$

with $\mathbf{L}_{t-1} = \partial f/\partial x\,[x(t-1)]$. Since, by definition, the exact error $\varepsilon(t)$ is unknown one should, rather than deal with (5), consider the evolution of the probability distribution of $\varepsilon(t)$. If $\mathbf{W}$ were known for each time step, (5) could be reformulated into the Fokker-Planck equation for the error density function. The weights $\mathbf{W}$ should then be determined in such a way that the expected (squared) error (or whatever error measure one chooses) is minimized.

As we already mentioned, this problem is theoretically as well as practically infeasible, especially in a context such as an atmospheric or ocean model where the number of degrees of freedom is so large. The main problem is, therefore, to find methods that reduce the number of degrees of freedom of the posed problem.

It is often reasonable to assume that the probability density is Gaussian. In this case the error is condensed into a covariance structure that, although still huge, contains a considerably smaller number of parameters. Similarly, one can assume that the process in question is governed, at least locally, by linear dynamics. For the EKF one generally assumes the following: (1) The nonlinearities, $n[\varepsilon(t-1)]$, can be neglected. (2) Observational error $r(t)$ and system error $q(t)$ are white noise processes with zero mean which are mutually uncorrelated. (3) Their respective covariances, $\mathbf{R}_t$ and $\mathbf{Q}_t$, are known a priori.

Under these circumstances the expression for the expected value of the outer product of (5) with itself becomes

$$\hat{\mathbf{P}}_t = \mathbf{L}_{t-1}\mathbf{P}_{t-1}\mathbf{L}_{t-1} + \mathbf{Q}_t \tag{6a}$$

$$\mathbf{P}_t = (\mathbf{1} - \mathbf{W})\hat{\mathbf{P}}_t(\mathbf{1} - \mathbf{W})^T + \mathbf{W}\mathbf{R}_t\mathbf{W}^T \tag{6b}$$

If one knew $\mathbf{W}$ at each time, the evolution of the error, i.e. $\mathbf{P}_t$, would be known. $\mathbf{W}$ is determined by requiring that the expected error, measured by the cost function

$$J = \langle \varepsilon(t)^T \varepsilon(t) \rangle = tr(\mathbf{P}_t)$$

be minimal. The solution matrix $\mathbf{K}_t$, i.e., the solution of $\partial J/\partial \mathbf{W}\,(\mathbf{K}_t) = 0$, is called the Kalman gain matrix:

$$\mathbf{K}_t = \hat{\mathbf{P}}_t(\hat{\mathbf{P}}_t + \mathbf{R}_t)^{-1} \tag{7}$$

(6) and (8) together determine the error model of the extended Kalman filter. It is a coupled recursive system in which $\mathbf{K}_t$ and $\mathbf{P}_t$ are calculated in each time step. In the linear case, that is when $f(x) = \mathbf{L}x$ for a matrix $\mathbf{L}$ independent of the system state x, the error model is much simpler (in that $n(\varepsilon) \equiv 0$) and the EKF reduces to the original Kalman filter (KF).

As a measure of the observational influence on $\bar{x}_i$ at time t we use the function

$$s_i(t) = \left(\sum_{j \le n} (\mathbf{K}_t)_{ij}^2 \right)^{1/2} \tag{8}$$

(compare (2b)) with $\mathbf{K}$ instead of $\mathbf{W}$, and call s_i the sensitivity function of $\bar{x}_i$. Its values range between 0 (no observational impact) and 1 (no model impact).

If all three conditions are met rigorously, the Kalman filter is both linear (as $n(\varepsilon) \equiv 0$) and optimal; moreover, by the derivation above, in this case the error model is true. Neither optimal nor true can be said of the extended Kalman filter. Assumption 1 is only a first-order approximation, and in highly nonlinear systems it is neither guaranteed that the assimilation is optimal nor that the error model is true. On the other hand, the EKF performance is very good for weakly nonlinear systems, that is, if assumption 1 is not violated too harshly, [cf. *Budgell*, 1987]. The violation of assumption 1 is equivalent to the fact that along certain trajectories the system's behavior changes too sharply. In this case the error model is no longer true, and, as we will see below, this can lead to a complete failure of the EKF.

If one is to model the nonlinear second-order effects, one might, as a first step, widen assumption 1 to the following: (1′) The nonlinearities, $n[\varepsilon(t-1)]$, are uncorrelated with the error itself, $\varepsilon(t-1)$, and with both the system error, $q(t)$, and the observation error, $r(t)$.

The main advantage of assumption 1′ is that it allows one to run the EKF with only one minor modification. Instead of using $\mathbf{Q}_t$ in (6a), one uses the new matrix $\mathbf{N}_t + \mathbf{Q}_t$, where $\mathbf{N}_t$ is the (second-order) covariance of the nonlinear terms $n[\varepsilon(t-1)]$. If the error model thus modified is correct, $\mathbf{N}_t$ should be one order of magnitude less than $\mathbf{P}_t$. Of course, the determination of $\mathbf{N}_t$ is highly nontrivial and in many cases not possible at all. In our applications we will use a covariance which is essentially a correlation together with a scaling factor.

Another way to remedy the problem of strong nonlinearity is to introduce higher-order moments for the statistical estimates or to introduce variational methods as those by *Miller et al.* [1993]. But as they already point out, this would increase the computational costs considerably and would be infeasible for higher dimensional systems.

Although by our above assumptions one vastly reduces the complexity of the assimilation problem, the computational burden for the determination of $\mathbf{K}_t$ is still immense. The central operation in (6), the calculation of the forecast error covariance $\hat{\mathbf{P}}_t = \mathbf{L}_{t-1}\mathbf{P}_{t-1}\mathbf{L}_{t-1}^T$, requires (even for the linear filter) as much as $O(n^2)$ multiplications for a system of dimension n (considering the sparseness of the system matrix). This is tolerable for systems which have $n < 1000$, as in typical engineering problems. The dimensionality of an atmosphere or ocean model, however, is much larger, so that the filter is hardly feasible in these cases. To overcome this obstacle various techniques have been developed to replace the optimality of the filter with other, less expensive procedures [see *Parrish and Cohn*, 1985; *Todling and Ghil*, 1990]. For smaller models, like the El Niño/Southern Oscillation (ENSO) model of *Zebiak and Cane* [1987], the Kalman filter seems appropriate (M. A. Cane et al., unpublished manuscript, 1993).

3. Regimes

As we already pointed out, strongly nonlinear and unstable systems tend to occupy regimes. To formally define the properties which make a system regimelike, let us assume that the system is given by the (discrete) equation

$$x(t) = f(x(t-1)) \tag{9}$$

Now we assume that there are N states ω_i, $i \le N$, in the neighborhood of which the system function f behaves almost linearly. That means one $x(t-1)$ is close to ω_i we can decompose $x(t-1)$ via

$$x(t-1) = \omega_i + \xi_i(t-1) \tag{10}$$

and expand f about ω_i, such that the linear system

$$x(t) = f(\omega_i) + \frac{\partial f}{\partial x}(\omega_i)\xi_i(t-1) = f(\omega_i) + \mathbf{L}_i\xi_i(t-1) \tag{11}$$

is a good approximation of f at time t. The linear system $(\omega_i, \mathbf{L}_i)$ is called a regime of f. Now, whatever the quality of the approximation, we may consider the linear system $(\omega_i, \mathbf{L}_i)$ to be acting on the whole state space, including the states where f and $(\omega_i, \mathbf{L}_i)$ behave very differently. A decomposition of f into regimes would be achieved and the system itself be regimelike, if for each state (of the attractor) there is at least one regime which resembles f at that instant.

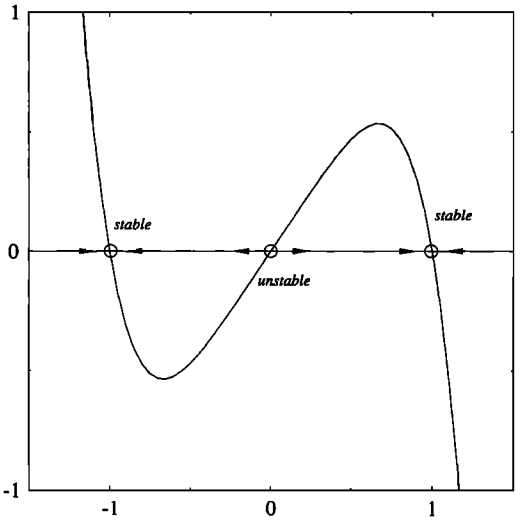

Figure 1. The double-well system. A typical realization of a dynamical system with three equilibrium points, where one unstable equilibrium, the origin, is surrounded by two (quasi)stable equilibria (the wells).

Very often one chooses equilibrium states, i.e., those states x for which $f(x) = x$, as basic states for a linearization like (11). From our present standpoint, however, every other state might equally well serve as a regime. One example would be the mean state of a solution. As is the second case below, the mean state need not be an equilibrium, especially in systems which are predominantly nonlinear.

4. Two Examples

Stochastic Noise: The Double Well System

Probably the simplest example of a nonlinear system which can exhibit two different regimes is the following one-dimensional equation

$$\dot{x} = x - x^3 + n(t) \tag{12}$$

where the external forcing term $n(t)$ shall represent a white noise process. The system is a simpler form of the one by *Duffing* [1918]. Without this forcing, (12) has three different equilibria $x = 0 \pm 1$ (compare Figure 1). The two outer points are stable, whereas the origin is unstable. That means, starting from some initial point which is not the origin itself, any solution will approach one of the two stable points and stay there forever. The stochastic forcing now makes the system switch randomly from one basin to the other; the frequency of the switching depends directly on the amplitude of the forcing.

Deterministic Chaos: The Lorenz System

The equations

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy \end{aligned} \tag{13}$$

describe the nonlinear interaction of the three dominant modes of the Oberbeck-Boussinesq equations for fluid convection in a two-dimensional layer heated from below. The system has been introduced by *Lorenz* [1963] and is the first example of a system which can exhibit what since then has been called deterministic chaos. Depending on its three parameters, σ (the Prandtl number), ρ (the Rayleigh number) and β (an aspect ratio), the system bifurcates through a large variety of substantially different behavior, from simple steady states to limit cycles and finally to chaos. In our study we analyze the system with the most commonly used parameter values of $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$. In this case (13) has three different fixed points, $\omega_0 = (0, 0, 0)$ and $\omega_\pm = (\pm 8.5, \pm 8.5, 27)$, all of which are unstable saddle points. The first fixed point describes a pure conductive heat transfer. Attached to this point are two stable and one unstable manifold, where one of the stabilities as well as the instability is very strong (large negative and positive eigenvalues, respectively). Their interplay seems crucial for the results which are described below. The other two equilibria are almost symmetric to each other. They are quasi-stable in the sense that a stable, one-dimensional manifold is accompanied by a two-dimensional weakly unstable manifold.

This constellation leads to the system's characteristic chaotic behavior: to switch irregularly between the two regimes of ω_- and ω_+. Once close to one of the quasi-stable equilibria, the system has the tendency to slowly spiral away from it. The switching occurs when the system enters the attracting region of the origin: The strong stability there highly accelerates the system so that it can "escape" from its current regime. The strong instability in turn provides the impetus for driving the system away to the other regime, and the same process starts all over again. It is the combination of these strong positive and negative feedback processes which is responsible for the fast error growth and the unpredictability of the Lorenz system. However, contrary to the double-well system the evolution itself is completely deterministic since there is no stochastic forcing involved at all. For details concerning the Lorenz system we refer to the work by *Guckenheimer and Holmes* [1983].

5. Interactive Kalman Filtering

As an alternative to the EKF we propose an assimilation procedure which makes use of the fact that the system that is to be assimilated decomposes into regimes $(\omega_i, \mathbf{L}_i)$, $i \le N$, for some number N. Roughly speaking, the EKF's updating of the error model via the current local linearization is replaced by a procedure that uses only the linear regime models. Being mutually incompatible, the regime error models advance the error covariance differently. Among all these 'suggestions' a compromise is made by the use of a weighting. The weighting reflects, for a single regime and arbitrary state of the system, the 'closeness' of the state to that regime relative to all other regimes. Because in each single assimilation step all regimes $(\omega_i, \mathbf{L}_i)$ are thus involved, the IKF is more globally inspired than the EKF. In detail the error model is determined in the following way.

Suppose we already assimilated data up to the time $t - 1$. We therefore have calculated an assimilation value $\bar{x}(t-1)$ and the error covariance $\mathbf{P}_{t-1}$; furthermore, we are able to determine the forecast value $\hat{x}(t) = f[\bar{x}(t-1)]$. Now we calculate N (temporary) guesses of the forecast error covariance $\hat{\mathbf{P}}_t$, by using (6a) with the respective regime models $\mathbf{L}_i$.

$$\hat{\mathbf{P}}^{(i)} = \mathbf{L}_i \mathbf{P}_{t-1} \mathbf{L}_i^T + \mathbf{Q}_t \tag{14}$$

Once we have assigned weights $\beta_i(t)$ to each single regime, we can define the forecast error covariance by

$$\hat{\mathbf{P}}_t = \sum_{i \le N} \beta_i(t) \hat{\mathbf{P}}^{(i)} \tag{15}$$

The weight $\beta_i(t)$ should reflect the "distance" of the regime $(\omega_i, \mathbf{L}_i)$ to the perfect model at the state $\bar{x}(t-1)$. If we expand according to (10),

$$\bar{x}(t-1) = \omega_i + \xi_i(t-1) \tag{16}$$

this distance can be defined by

$$\delta_i(t) = \| f(\omega_i) + \mathbf{L}_i \xi_i(t-1) - \hat{x}(t) \| \tag{17}$$

where $\| \ \|$ denotes the Euclidean norm in the phase space. There is a lot of freedom now to choose the actual weighting. We decided to take the value

$$\beta_i(t) = \frac{1/\delta_i^2(t)}{\sum_{j \le N} 1/\delta_j^2(t)} \tag{18}$$

This completes the definition of the interactive Kalman filtering (IKF). Note that in (14) we could have chosen an individual system noise for each single regime. The weighting in (18) relative to 1 only expresses the fact or, better, the assumption that $\mathbf{P}$ is changing only close to the regimes, but this, of course, depends on the choice of the regimes.

The number of regimes N that participate in the IKF crucially affects the computational costs of the filter. The expensive covariance updating (14) has to be calculated N times for each time step. On the other hand, one does not have to determine the local error model each time step. Hence depending on N, the IKF can be less or more expensive than the EKF, although an exact estimate of the respective costs is hard to give.

6. Testing the EKF and IKF

We test the various assimilation methods through a series of simulations with the two nonlinear models introduced above, the double-well system and the Lorenz system. (This choice coincides with that of *Miller et al.* [1993].) To initialize the error model (6) and (8) we always started with the final value $\mathbf{P}$ of some former assimilation run.

Both systems have three different equilibria; the unstable origin and two accompanying stable or quasi-stable equilibria. For the double-well system the mean state is the origin and hence is an equilibrium. This is not the case for the Lorenz system, and we chose the mean state as a regime of its own. We dealt with three different choices of regimes for the IKF. They are IKF1, the mean state as the only regime; IKF2, the two (quasi)stable equilibria; and IKF3, all three equilibria.

The first method corresponds to a normal linear Kalman filter approach. Note that for the Lorenz system the first choice deals with a regime that is not an equilibrium.

Since in our study the model output plays the role that

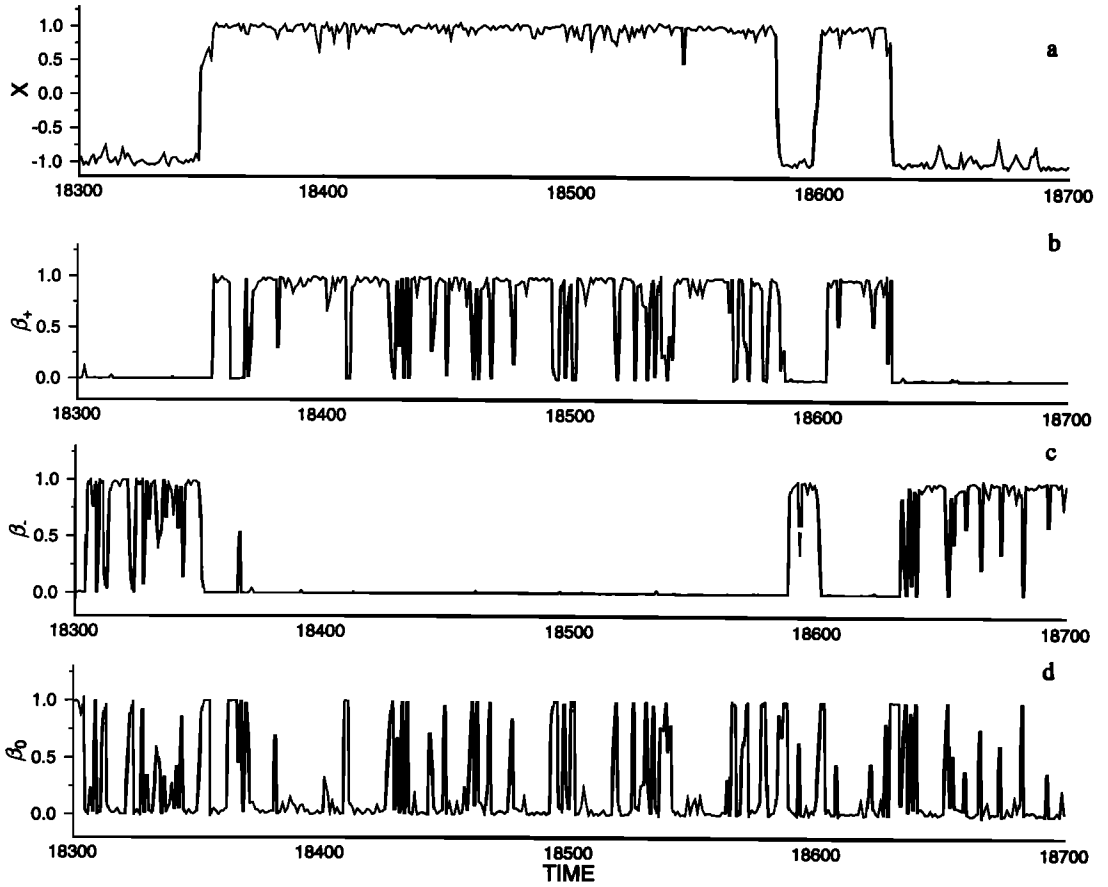

Figure 2. (a) A small section of a typical realization of the double-well system together with the three regimes (b) +1, (c) −1, and (d) 0 of the IKF3 assimilation (for noise levels see text). The regime weights closely follow the true values, with sharp changes during the transitions. The short delays are caused by the delays in the assimilation. Note the larger spikes of the 0 regime weight during the switching.

reality normally plays, and the system noise should therefore be zero, we impose certain perturbations on the model through an additional white noise forcing term (which exists for the double-well system anyway) with prescribed covariance **Q** constant with time. We assume that **Q** is exactly known for the assimilation and has the form $\mathbf{Q} = \mu\mathbf{I}$. A perfectly known system noise is, theoretically, the best case possible for the EKF assimilation. However, since the systems under consideration are strongly nonlinear, a better approach is to guess the system noise via some enlarged covariance $\mathbf{Q} + \mathbf{N}$, where **N** should comprise all nonlinear contributions to the error.

Since we do not know of a straightforward way to model **N**, we use the following approach. Starting from a fixed correlation matrix **C**, the form of which is given in the appendix, we perform a series of assimilations that use $\mathbf{N} = \nu\mathbf{C}$, with ν varying in a certain range. We not only vary the system noise but also the true forcing **Q**. That is, for a series of forcing covariances, $\mu\mathbf{I}$, μ varying, we are assimilating with system noise covariances $\mu\mathbf{I} + \nu\mathbf{C}$, ν varying. As a performance index for a single assimilation, identified by the pair (μ, ν), we take the scalar

$$\Delta(\mu, \nu) = \frac{tr(\mathbf{P})}{tr(\mathbf{R})} \tag{19}$$

which is the percentage of the measured (not modeled) analysis error variance (**P**) relative to the imposed observational noise variance (**R**).

In most experiments the observations were inserted each time step. We assumed that the observational error was white with a covariance structure of the form $\psi\mathbf{I}$, with a scalar ψ that reflects the overall variability of the process as described below.

Double-Well System

To solve system (12) we used the International Mathematics and Statistics Libraries routine DGEAR with a step size of 0.9. The single runs were over 100,000 time steps. This large number is necessary in order to stabilize the statistics, especially when the forcing noise is very small. The forcing noise parameter μ was chosen from the series 0.01, 0.02, $\cdots$, 0.1, and the ν values were 0, 0.01, 0.02, $\cdots$, 0.1; the observational noise variance was held constant at a value of $\psi = 1$.

In Figure 2 we see a section of a typical realization together with the three regime weights of the IKF3 assimilation, using $(\mu, \nu) = (0.1, 0)$. The switching between regimes −1 and +1 can clearly be seen in the x values themselves, but the weights β also reflect these switches convincingly, especially β_- and β_+. Note the asymmetry in

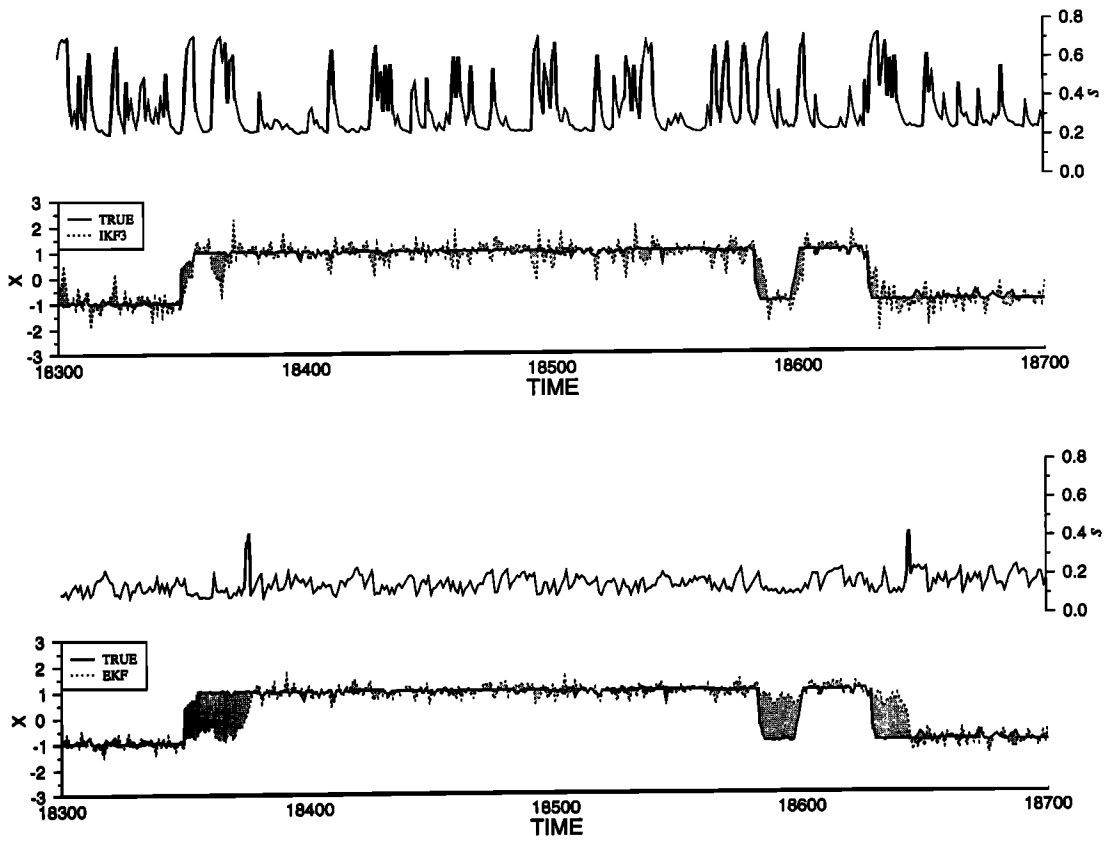

Figure 3a. The same as Figure 2 except now with assimilations. (bottom) The EKF assimilation; the delay in regime switching (lower curve) is too long so that the short double switch cannot be traced. Inside the regimes, if they are correct, the observational noise impact is small. Both phenomena are mainly due to the character of the sensitivity function (upper curve). The sensitivity function has rather low values and only small variability, with maximum spikes not reaching beyond 0.5. Note that each transition is accompanied by a spike event. (top) The IKF3 assimilation; we see a very good reproduction of the original data, with all regime switches simultaneously traced (lower curve). The observational noise impact during the regimes is slightly larger than for the EKF. There seems to be a lower bound in sensitivity (upper curve) of about 0.25, which is much larger than any lower bound of the EKF sensitivity.

the weighting: A weight of 1 is much noisier than a weight of 0. The spike-like behavior of β_0 reflects the instability of the origin. These spike events become significant during transitions (namely, the first and fourth transition) when they have a much longer duration than normal. We will return to this point later.

Figures 3a and 3b depict, in the same time section, the true and assimilated values of x for the EKF and IKF3; we show the case $(\mu, \nu) = (0.1, 0.05)$. The main observation is the complementary behavior of the filters during the regime time and the transitions. Once a regime is established, the EKF seems to be less disturbed by observational noise and closer to the true values than the IKF3. However, it is just because of this that it needs a longer time to react when a (true) regime transition occurs. The double transition $+1 \rightarrow -1 \rightarrow +1$ near time step 18,600 is not traced at all by the EKF. This general feature is reflected in the sensitivity function for x, as defined by (8): It shows much larger values, and a much larger variability, for the IKF3. Small peaks during an EKF switch contrast with large and broad episodes during IKF3 switches. Moreover, there seems to be a threshold in the IKF3 sensitivity function below which its values never fall, and if there is such a threshold for the EKF, it is certainly lower. We have $\Delta = 32\%$ for the EKF and $\Delta = 16\%$ for the IKF3. The complete Δ matrices for both filters are given in Tables 1 and 2. We see that the overall error is much smaller for the IKF3. Note that the introduction of second-order errors (i.e., $\nu > 0$) for the nonlinearities only worsens the performance for both filters.

What has been said about the relation between the EKF and IKF3 can also be said for the IKF2 and IKF1 (compare Figure 3b). The IKF2's behavior inside regimes can be compared to the EKF; it is more undisturbed by noise and closer to the true values. However, the transitions are traced more reliably by the IKF1. Note that unlike the EKF, the IKF2 performs the double regime switch near step 18,600. The different behavior is again reflected in the sensitivity function, with very small, constant values for the IKF2 and large (and constant) values for the IKF1. The constancy of the IKF2 sensitivity is caused by the symmetry of the single localizations at -1 and $+1$; both have the same linear error model. For the case of Figure 3b (that is, for the pair $(\mu, \nu) = (0.1, 0.05)$) we have $\Delta = 11\%$ and $\Delta = 58\%$ for the IKF2 and IKF1, respectively. So the IKF2 performs even better than the IKF3. Table 2 shows that this is not the case for many values of (μ, ν).

From Table 3 we see that there is an instability of the IKF2 for various noise levels. Especially for the pair $(\mu, \nu) =$

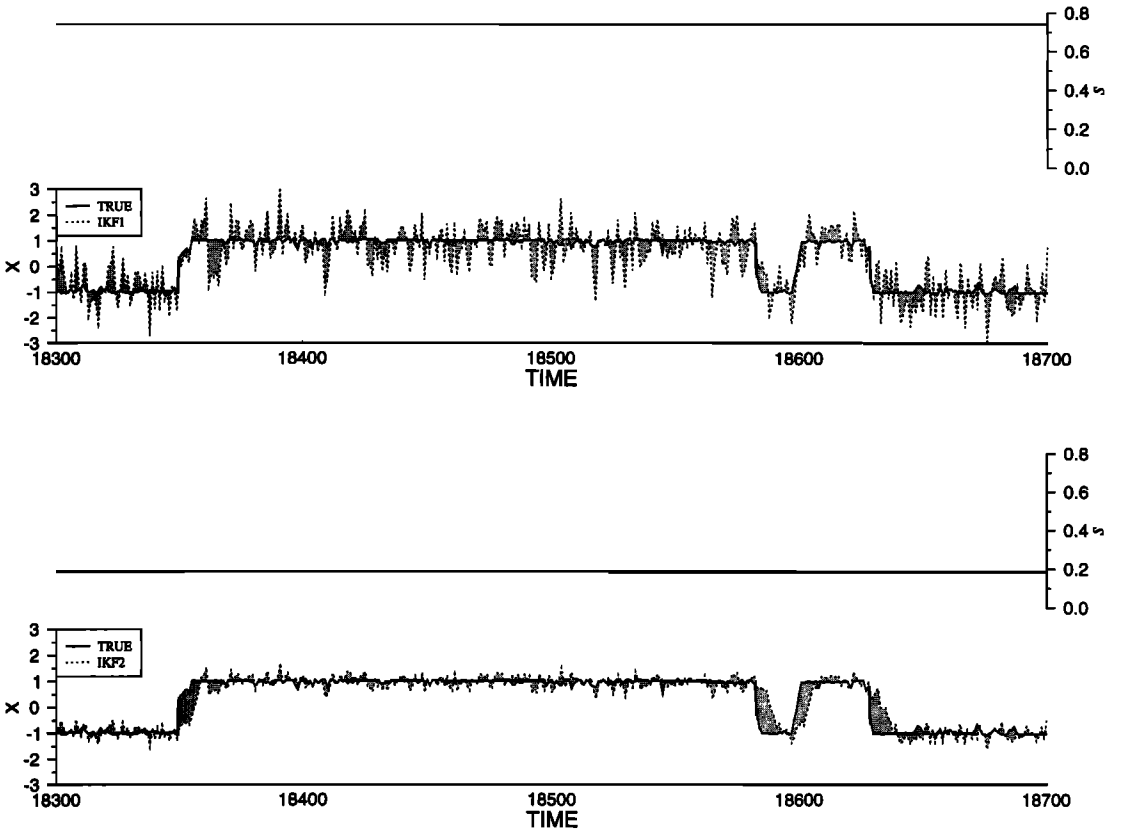

Figure 3b. The same as Figure 3a, only for the IKF2 and IKF1 assimilations. (bottom) The IKF2 case. The regime jumps can still be traced but their delay is increased compared to the IKF3. This is mainly due to the fact that the sensitivity function is now constant at a value of about 0.2. Note that this value is larger than the limit of the EKF sensitivity. (top) The IKF1 assimilation follows the jumps simultaneously, which is caused by the high level of observational noise influence ($s \approx 0.7$). This is paid for by an increased perturbation of the assimilation by observational noise during the regimes.

(0.06, 0), the IKF2 shows an error of $\Delta = 104\%$. That means the filter is actually worsening the assimilation, compared to pure observation. But through an increase of the nonlinear noise terms the filter can be improved considerably, lowering errors to 5–6%. The IKF1 error (see Table 4) seems to be rather independent from the various noise levels, thus reflecting the fact that as long as the sensitivity to observations is large enough, the simple linear filter cuts the observational noise roughly in half.

How a change in the observation frequency affects the assimilation is exemplified by the following: using the parameters $(\mu, \nu) = (0.1, 0)$, we performed five EKF and IKF3 experiments, with observations taken every 1st, 2nd, 3rd, 4th, and 5th time step. A typical section of the various

Table 1. The Δ Matrix of the EKF for the Double-Well System, Using an Observational Noise Scale of $\psi = 1$

	ν										
μ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	4	5	5	6	6	7	7	8	8	9	9
0.02	10	11	11	12	12	13	13	13	14	14	14
0.03	15	15	16	16	16	16	17	17	17	18	18
0.04	18	19	19	19	19	19	20	20	20	21	21
0.05	21	21	22	22	22	22	22	23	23	23	24
0.06	22	23	23	23	23	24	24	24	24	25	25
0.07	25	26	26	26	26	27	27	27	27	28	28
0.08	27	27	28	28	28	28	29	29	29	30	30
0.09	29	29	30	30	30	30	31	31	31	32	32
0.1	31	31	32	32	32	32	33	33	33	33	34

Rows represent various white noise forcing scales through different μ's; columns represent different ν's, i.e., nonlinear contributions to the system noise. The increase of ν does not diminish the error.

Table 2. The Δ Matrix of the IKF3 for the Double-Well System Using an Observational Noise Scale of $\psi = 1$

	ν										
μ	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	0	0	1	1	2	3	4	5	7	8	9
0.02	0	1	1	2	3	4	5	7	8	9	10
0.03	1	1	2	3	4	5	7	8	9	10	11
0.04	1	2	3	4	5	7	8	9	10	12	13
0.05	2	3	4	6	7	8	9	10	12	13	14
0.06	3	5	6	7	8	10	11	12	13	14	15
0.07	5	6	8	9	10	11	12	13	14	15	16
0.08	7	8	9	11	12	13	14	15	16	17	17
0.09	9	10	11	12	13	14	15	16	17	18	19
0.1	11	12	13	14	15	16	17	18	19	19	20

Rows and columns are the same as Table 1. The errors throughout are smaller than for the EKF and increase slightly with larger ν.

Table 3. The Δ Matrix of the IKF2 for the Double-Well System Using an Observational Noise Scale of $\psi = 1$

	ν										
μ	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	0	0	1	1	12	11	2	2	3	3	4
0.02	0	1	1	12	11	2	2	3	3	4	4
0.03	1	1	12	11	2	2	3	3	4	4	5
0.04	1	12	11	2	2	3	3	4	4	5	5
0.05	59	60	6	3	3	4	4	4	5	5	6
0.06	104	26	7	5	5	5	5	5	6	6	6
0.07	44	11	8	6	6	6	6	6	7	7	7
0.08	25	15	10	8	7	7	7	8	8	8	9
0.09	20	14	10	9	9	9	9	9	9	10	10
0.1	19	15	13	12	11	11	11	11	11	11	12

Rows and columns are the same as Table 1. With the exception of the band of very large errors, the performance is better than all other methods. A value of Δ = 104% indicates a persistent failure of regime tracing.

Table 4. The Δ Matrix for the IKF1 for the Double-Well System Using an Observational Noise Scale of $\psi = 1$

	ν										
μ	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
0.01	55	55	55	56	56	56	56	56	56	56	56
0.02	55	55	56	56	56	56	56	56	56	56	57
0.03	55	56	56	56	56	56	56	56	56	57	57
0.04	56	56	56	56	56	56	56	57	57	57	57
0.05	56	56	56	56	56	57	57	57	57	57	57
0.06	56	56	57	57	57	57	57	57	57	57	58
0.07	57	57	57	57	57	57	57	57	58	58	58
0.08	57	57	57	57	57	57	58	58	58	58	58
0.09	57	57	57	57	57	58	58	58	58	58	58
0.1	57	57	57	58	58	58	58	58	58	58	58

Rows and columns are the same as Table 1. There is no sensitivity of the IKF1 for the noise parameters. Roughly, the observational error is cut in half.

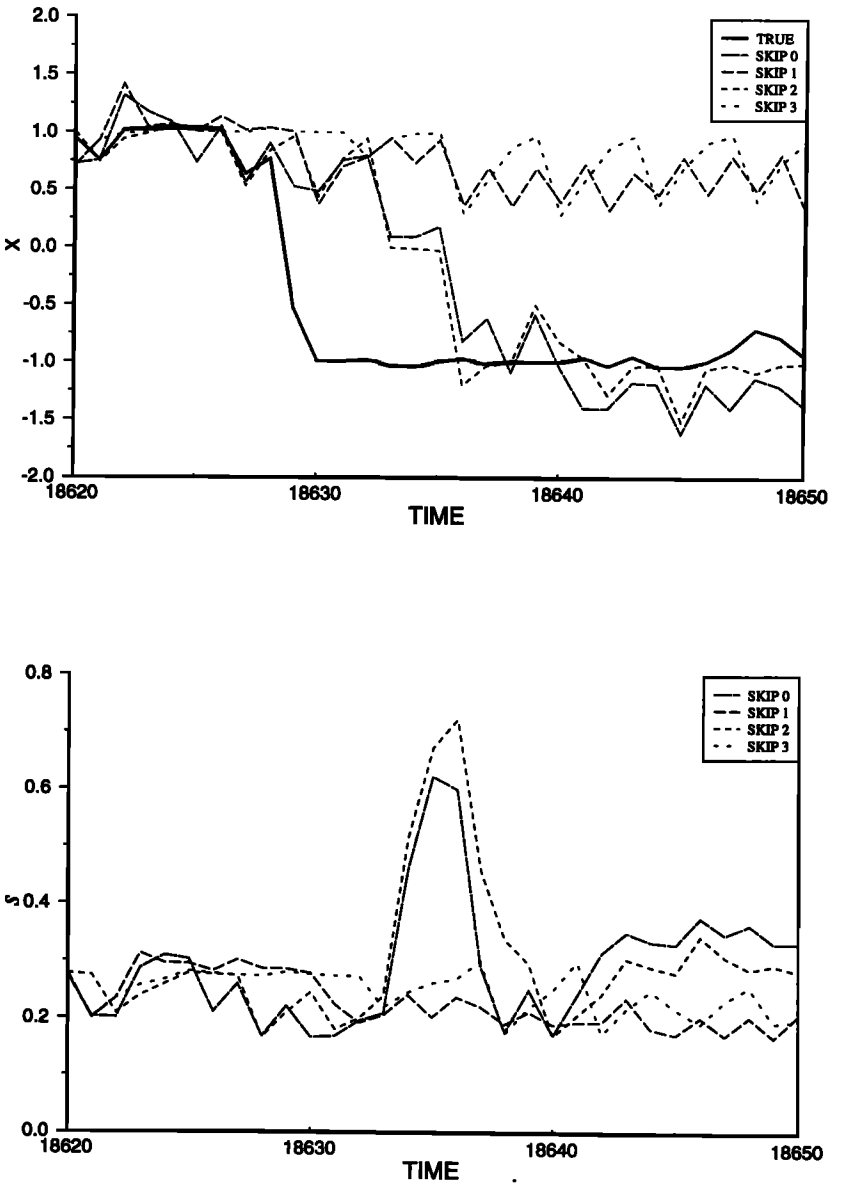

Figure 4a. (top) A section of the double-well system undergoing a switch and for EKF assimilations with decreasing observation frequency. Generally, fewer observations cause a delay in the switching of the assimilation. The exception for the SKIP-1 case is caused by a number of bad observations at the time. (bottom) The various sensitivity functions of the assimilations. The spikes occur when the assimilation switches.

EKF assimilations is shown in Figure 4a (compare Figure 2). The true signal performs a regime switch $+1 \rightarrow -1$. The SKIP-0 (i.e., the original) assimilation traces this switch with a delay of about six time steps. If we take observations only at every second time step, the assimilation cannot trace the switch anymore, at least not during the time interval considered. Because there was a good observation in time, the SKIP-2 assimilation was able to trace the switch, unlike the SKIP-4 assimilation which missed it. There are a number of bad observations, and this can be seen in the SKIP-1 assimilation which is worse than the SKIP-2. The behavior of the IKF3 assimilation is different, as we see in Figure 4b. The switch is recognized almost immediately when observations were available at all times. With data insertions every other time step, the tracing of the switch is delayed by about eight time steps. As in the EKF assimilation, the performance improves when we take observations only every third time step, now switching the regime three steps later. In the SKIP-3 case the switching still occurs, with a delay of some 15 time steps.

All these results are in accordance with the investigations of *Miller et al.* [1993], who studied the EKF's sensitivity to missing observations in greater detail.

Lorenz System

To numerically realize system (13) we again used DGEAR, with a step size of 0.025. This gives an average oscillation period around the unstable convective equilibria of ~40 time steps. Because the Lorenz system will undergo regime switches without any additional noise, the total

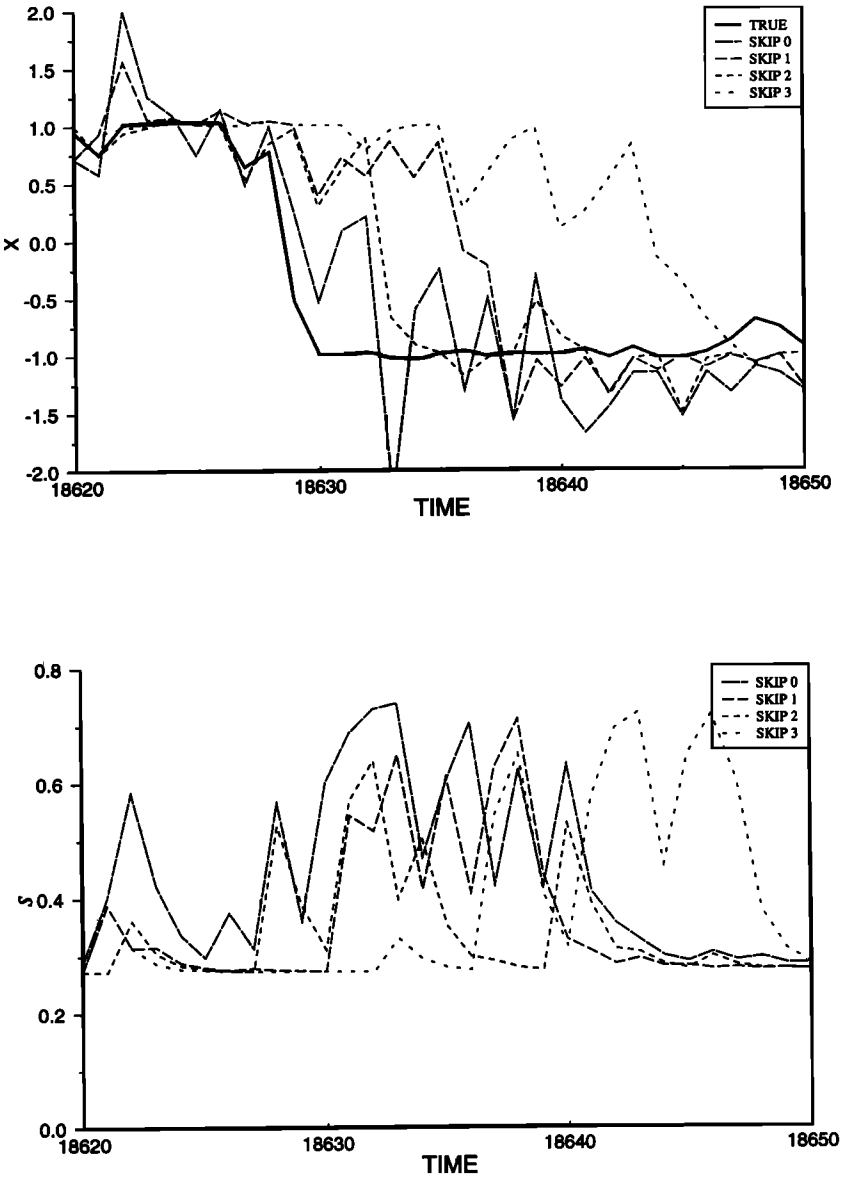

Figure 4b. The same as Figure 4a, now with the IKF3 assimilation. Note that the delay time is much shorter now. Even the SKIP-3 assimilation traces the switch after about 20 time steps. A main difference from the EKF is the width of the sensitivity spikes; they are much wider for the IKF3.

integration time can be much smaller than that for the double-well system. However, to achieve stability in the statistics we integrated (13) over 20,000 time steps. Figure 5 indicates that the mean duration time in a single regime is about four cycles, such that there are about $20{,}000/(4 \times 40) \approx 100$ regime changes. If we write (13) in the form $\dot{x}(t) = \mathcal{L}[x(t)]$, then the system we actually deal with is $\dot{x}(t) = \mathcal{L}[x(t)] + n(t)$. Here $n(t)$ is a random process with a prescribed covariance of the form $\mu\mathbf{I}$, with μ being taken from a specific set as described below. As we already mentioned, we use a system noise covariance of the form $\mu\mathbf{I} + \nu\mathbf{C}$. Because now there is "internal" noise in the process which comes from its chaotic properties we let μ and ν vary in a much broader range; we used the range 10^{-7}, $10^{-6}, \cdots, 10^{1}$. As in the former example, the observational noise covariance was taken to be $\psi\mathbf{I}$, with ψ assuming various values, as described below.

We see the x coordinate of the process, calculated for a noise level of $(\mu, \nu) = (10^{-5}, 10^{-5})$ (in a typical section in Figure 5) together with the three IKF3 regime weights. The observational noise scale was $\psi = 49$ which is about the average variability of the system confined to one of the quasi-stable regimes. Again, as in the double-well system, the regimes are traced very well, with sharp changes during the transitions. Moreover, the spikelike behavior of the ω_0 regime is unchanged, although not every spike leads to a transition. We conclude that the weighting β gives a very natural definition of "regime".

Figure 6a shows the EKF assimilation (x values) in a smaller section, this time using an observational noise level of $\psi = 100$. From the assimilated curve we clearly see that the regimes are not traced confidently (the shaded areas). The assimilation tends to lose contact with the true process, although once this contact is established the assimilation is

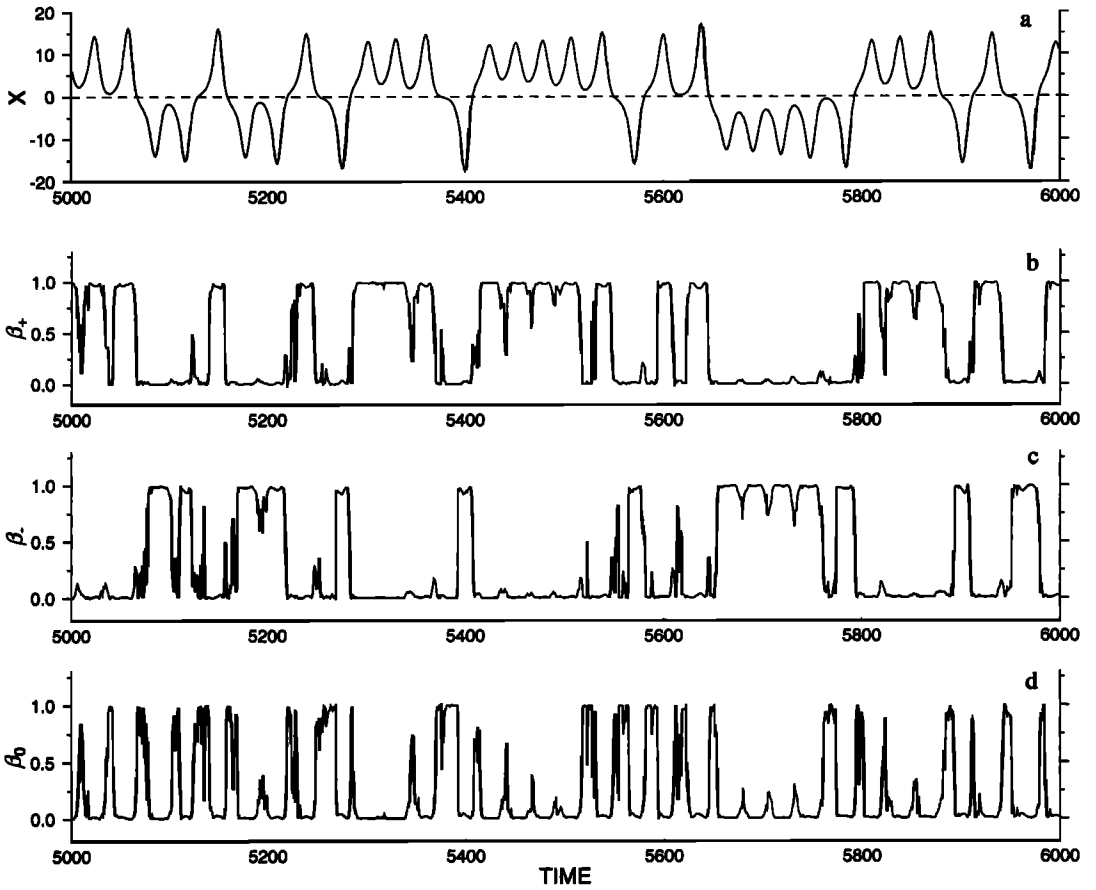

Figure 5. (a) A typical realization of the Lorenz system (the x coordinate). One clearly sees the two quasi-stable regimes (negative and positive x values). The three lower panels show the regimes as detected through the IKF3 assimilation weights (b) $\omega+$, (c) $\omega-$, and (d) ω_0 (noise parameters are $(\mu, \nu) = (10^{-3}, 10^{-3})$ and $\rho = 100$). As in the double-well system, there is a sharp change in the weights during the transitions; these appear as spikes in the ω_0 weight.

actually quite good. The general pattern described by the sensitivity function is, starting from almost zero values after a transition, a gradual increase to maximum values of about 0.07 when the regime is dying. After this, in a time span of about five steps, there is a sharp drop to zero values and a return to the old values. This is probably caused by the successive transition of ω_0 through the strongly stable and unstable manifolds. During this period, observational noise has practically no chance to enter the assimilation. The regimes are traced confidently by the IKF3, see Figure 6b, with only minor delays during the transitions. The filtered values show larger high-frequency disturbances by observational noise than by the EKF. This is reflected in the sensitivity function, which never approaches zero at any time or changes as sharply between extreme values; its values never fall below 0.02. This is probably one of the main reasons for the different behavior of the EKF and IKF3. This is especially clear at time steps 5500–5650. Whereas the IKF3 sensitivity assumes lower values only after ~5660, that is, when the process is stable in ω_-, the EKF shows zero values at ~5620 and ends up in the wrong regime. We are not in a position to explain the different behavior of the sensitivity functions.

The IKF2 assimilation (Figure 6c) occupies an intermediate position between EKF and IKF3. This applies to the small-scale observational noise seen in the assimilated values (which is smaller under the EKF and larger under the IKF3) as well as the delays needed to follow the transitions (larger under EKF and smaller under IKF3). Compared to the other filters, the IKF2 sensitivity function completely lacks any greater variability, departing only slightly from the mean value that marks the minimum sensitivity of the IKF3. Similar to the double-well system, this is easily explained by the similarity of the two quasi-stable regimes. Their higher stability lets the sensitivity settle at a value too low for observational noise to become effective. A constant value of about 0.05 seems large enough to let the IKF1 assimilation always stay in the correct regime (Figure 6d). The enlarged IKF1 sensitivity can also be seen during the period in the ω_- regime where the errors are larger compared to the other assimilations.

This can be verified by looking at the Δ values. We have $\Delta = 32\%$ for the EKF, and $\Delta = 8\%$, 35%, and 22% for the IKF3, IKF2, and IKF1 assimilations, respectively. Tables 5–8 give a thorough overview of what happens for the case of the smaller observational noise with $\psi = 49$.

For the EKF (Table 5) we see that there are large errors when the system noise is small. Only when the nonlinear contributions ν go beyond 10^{-3} does the EKF trace the process confidently. This is in accordance with *Miller et al.*

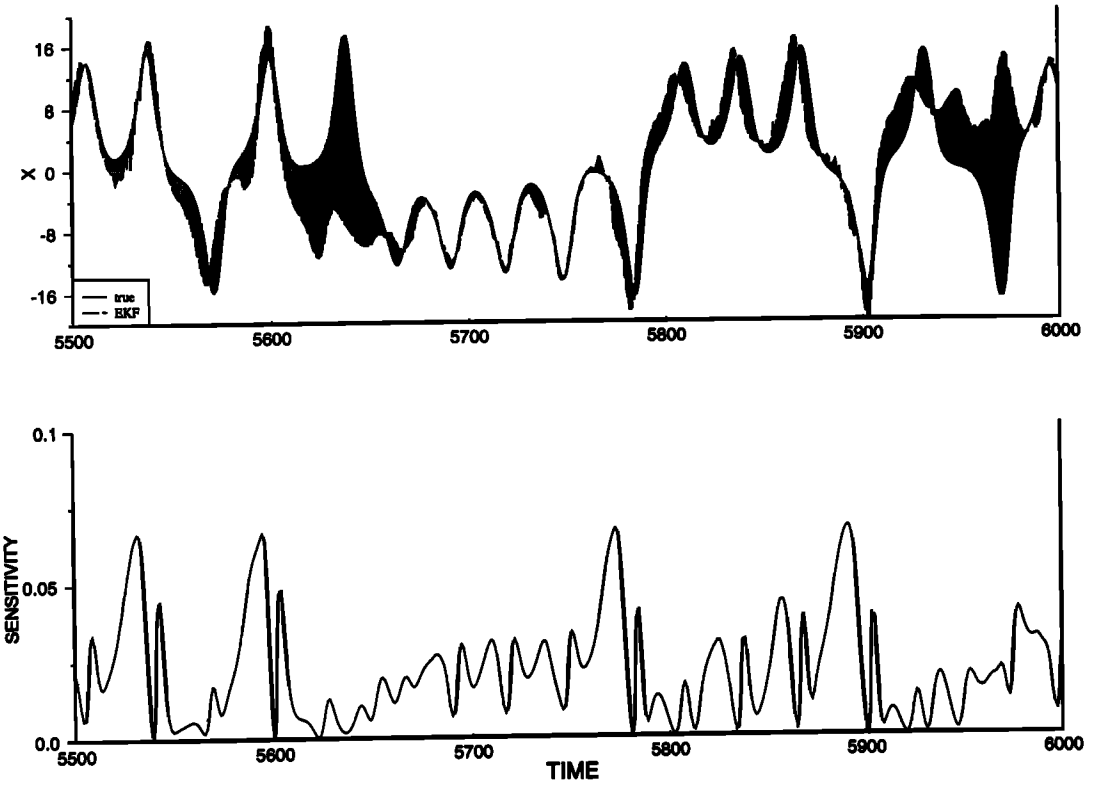

Figure 6a. (top) The EKF assimilation for the same situation as in Figure 5. There are periods when the EKF is in the wrong regime (shaded areas). Note that once the assimilation is in the correct regime, its errors are quite low. (bottom) The x sensitivity function on the assimilation of x. Note the sharp increase of the functions during the regime switches. Note also that at certain times the sensitivity is almost zero.

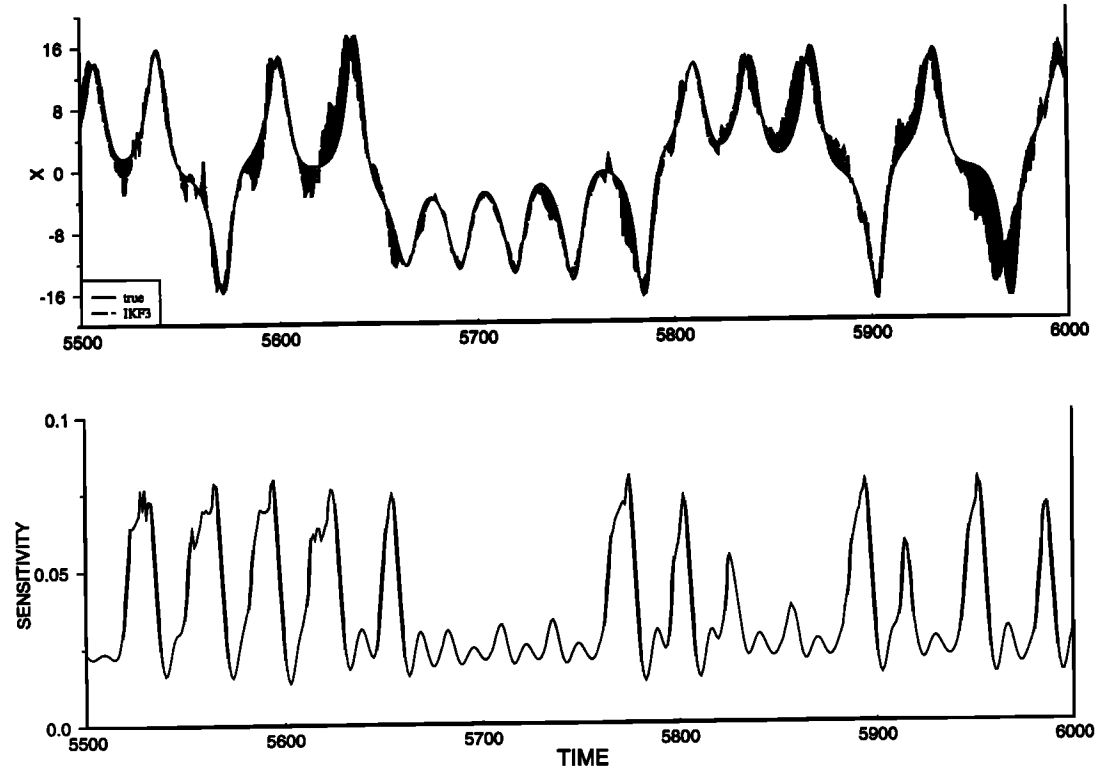

Figure 6b. Like Figure 6a except now for the IKF3. (top) Unlike the EKF, the IKF3 assimilation is able to trace all regime switches. This is accompanied by slightly enlarged perturbations caused by observational noise, as, for example, at time step 5560. The sensitivity function, (bottom) which is responsible for this shows larger average values than the EKF sensitivity. Moreover, the number of sensitivity spikes is significantly larger than in the EKF. Note that there are IKF spikes during those cycles when the EKF fails to trace the transition.

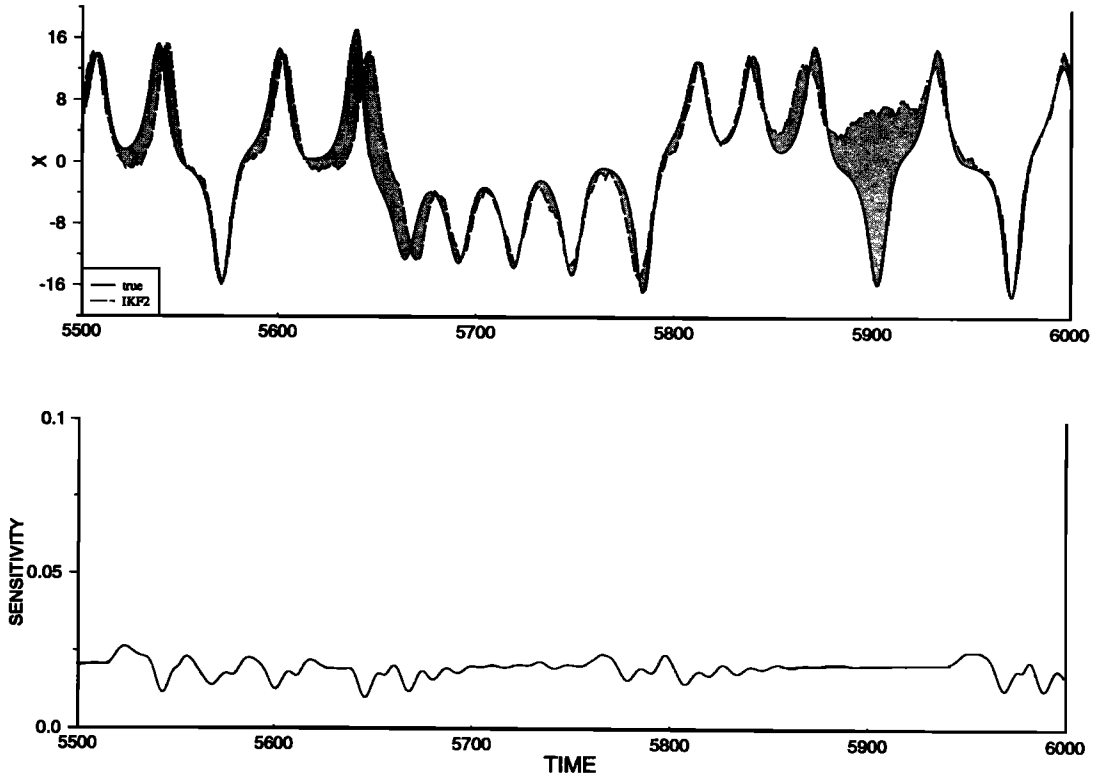

Figure 6c. The IKF2 assimilation. (top) The assimilation itself is very similar to the EKF case, although (bottom) the sensitivity shows much less variability compared to the EKF sensitivity. It varies slightly around a mean value of 0.025, a value that is too low for fast transitions to be captured. The low variability comes from the lack of an instability which would let the model error grow.

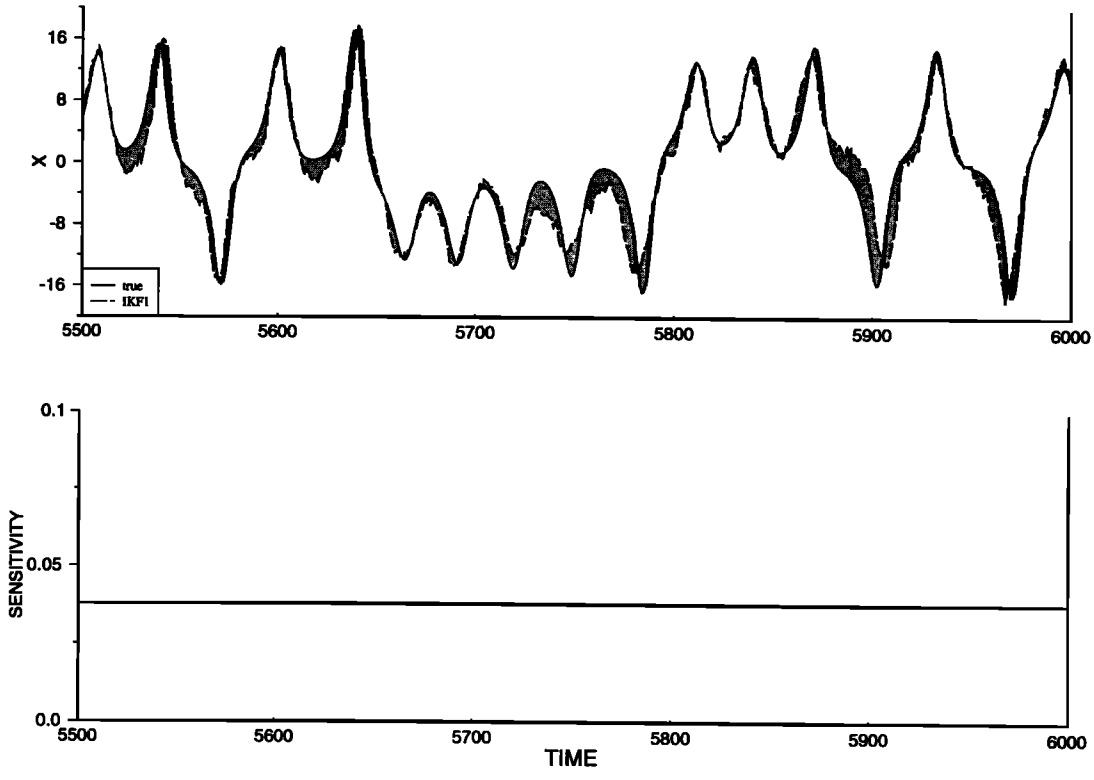

Figure 6d. (top) The IKF1 assimilation (which is equivalent to the linear Kalman filtering). All regime switches (in the specified section) are traced correctly. Compared to the IKF3, there is more observational noise visible in the assimilated values; this is reflected in the larger overall error ($\Delta = 22\%$ compared to $\Delta = 8\%$ for the IKF3). (bottom) The sensitivity function is constant at about 0.05. This value, slightly larger than the one for the IKF2, lies just at the threshold of tracing all regime switches confidently.

Table 5. The Δ Matrix of the EKF for the Lorenz System With an Observational Noise Scale of $\psi = 49$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	34	37	35	33	21	7	6	8	19
−6	32	36	32	37	18	9	6	8	19
−5	48	42	42	37	20	8	6	8	19
−4	27	30	27	24	25	9	6	8	19
−3	16	17	15	15	11	7	6	8	19
−2	10	10	10	10	9	8	6	8	19
−1	7	7	7	7	7	7	7	8	19
0	11	11	11	11	11	11	12	12	21
1	30	30	30	30	30	30	30	30	34

Values of μ and ν are given by their decimal exponents. Note the area of large errors for small noise scales. The performance improves markedly with growing nonlinear contributions ν.

[1993], who find an improvement of the assimilation by increasing the system noise appropriately. The large errors are certainly induced by the EKF's inability to trace regime switches as we have already seen in the example with higher observational noise.

The overall IKF3 performance (Table 6) is much better, completely lacking the window of large errors for small system noise levels; there the errors are constantly about 7%. This is the error level that the EKF reaches when the system noise is increased. The values at very large system noise levels coincide with those for the EKF.

The picture for the IKF2 (Table 7) has changed considerably compared to its behavior in the double-well system. For practically all system noise levels there is now an error of about 40–50% of the observational error. Increasing the nonlinear part of the system noise has no effect at all, probably because there is no mechanism for internal error growth without any instability. Here we see a principal difference from the nonchaotic double-well system.

This is supported by the results of the IKF1 assimilation

Table 6. The Δ Matrix of the IKF3 for the Lorenz System With an Observational Noise Scale of $\psi = 49$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	7	7	7	7	7	7	7	9	19
−6	7	7	7	7	7	7	7	9	19
−5	7	7	7	7	7	7	7	9	19
−4	7	7	7	7	7	7	7	9	19
−3	7	7	7	7	7	7	7	9	19
−2	7	7	7	7	7	7	7	9	20
−1	8	8	8	8	8	8	8	9	20
0	12	12	12	12	12	12	12	13	21
1	31	31	31	31	31	31	31	31	34

Values of μ and ν are given by their decimal exponents. For all noise levels the performance is very good, comparable to the best EKF case with increased ν.

Table 7. The Δ Matrix of the IKF2 for the Lorenz System With an Observational Noise Scale of $\psi = 49$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	41	41	41	40	40	40	25	10	19
−6	50	50	50	50	51	46	30	10	19
−5	52	52	52	52	52	49	31	10	19
−4	46	46	46	46	46	45	29	10	19
−3	40	40	40	40	39	37	25	9	19
−2	45	45	45	45	45	44	30	10	19
−1	30	30	30	30	30	29	22	10	19
0	14	14	14	14	14	14	13	13	21
1	31	31	31	31	31	31	31	31	34

Values of μ and ν are given by their decimal exponents. Note the very large errors for all noise levels, only slightly decreasing with larger ν's.

(Table 8) which uses, as the only regime, the point ω_m = (0.9, 0.9, 23.7) (which is the estimated mean value of the process). This point has, similar to the origin, one unstable and two stable manifolds, the instability being smaller than that of the origin. The errors are significantly smaller than those of the IKF2. For smaller system noise there is a difference of up to 15%, and the performance improves further with larger ν.

To support the view that instabilities are crucial for the error model, we modified the IKF1 in the following way: Instead of using the local linearization as the error model, we fitted a linear (Markov) model to the anomalies relative to ω_m, using the method of least squares [cf. *von Storch et al.*, 1993]. The model thus determined must lack instabilities. The assimilation with this model was entirely useless since the sensitivity was practically zero at all times. This becomes understandable when we consider the vast difference in magnitude between observational and system error. From the definition of the Kalman matrix **K**, compare (7), we see

Table 8. The Δ Matrix of the IKF1 for the Lorenz System With an Observational Noise Scale of $\psi = 49$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	31	31	31	30	30	28	17	9	18
−6	39	39	39	39	38	33	20	10	18
−5	36	36	36	36	36	31	19	10	18
−4	35	35	35	35	36	31	20	10	18
−3	28	28	28	28	28	26	17	9	18
−2	37	37	37	37	37	35	21	10	18
−1	20	20	20	20	20	19	15	10	19
0	14	14	14	14	14	14	14	13	20
1	31	31	31	31	31	31	31	31	34

Values of μ and ν are given by their decimal exponents. As in the double-well system, there is not much sensitivity to the noise parameters. Note that the errors are about one third of the observational noise, compared to one half for the double-well system.

Table 9. The Δ Matrix of the EKF for the Lorenz System With an Observational Noise Scale of $\psi = 0.01$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	3	3	4	6	14	41	84	99	100
−6	4	4	4	7	14	41	84	98	100
−5	5	5	5	7	14	41	84	99	100
−4	9	9	9	10	15	42	84	98	100
−3	22	22	22	22	24	44	84	99	100
−2	57	57	57	57	58	63	86	99	100
−1	91	91	91	19	91	91	93	99	100
0	100	100	100	100	100	100	100	100	100
1	101	101	101	101	101	101	101	101	101

Values of μ and ν are given by their decimal exponents. The results have improved markedly with the smaller observational error. All regime transitions are traced confidently.

Table 11. The Δ Matrix of the IKF2 for the Lorenz System With an Observational Noise Scale of $\psi = 0.01$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	107	101	65	13	14	41	84	99	100
−6	990	844	328	13	14	41	84	99	100
−5	412	407	198	14	14	41	84	99	100
−4	28	27	25	14	15	42	84	98	100
−3	24	24	24	23	25	44	84	99	100
−2	58	58	58	58	58	63	86	99	100
−1	91	91	91	91	91	91	93	99	100
0	100	100	100	100	100	100	100	100	100
1	101	101	101	101	101	101	101	101	101

The values of μ and ν are given by their decimal exponents. Note that extremely large errors occur for small system noise. An error of Δ = 1000%, which is about 10 (=1000 × 0.01) in the dimensionless units of the Lorenz system, indicates that there is no correct regime tracing at all. With larger ν values the errors decrease considerably.

that if the observations are poor enough (large **R**), either the system noise **Q** or the error growth through **L** (compare (6a)) has to be large in order to prevent **K** from becoming small.

The occasional bad performance of the EKF as well as the IKF2 for the Lorenz system is due to a lack of growth in the assimilation error model which could compete with the error growth caused by the chaotic dynamics. The differences between the assimilation methods disappear, and EKF and IKF2 improve, with stronger true white noise forcing. Choosing the noise scale μ from the range $1, 2, \cdots, 10$, with constant $\nu = 0$ and $\psi = 49$, the sensitivity of all assimilations is increased significantly, and the filters are much more governed by the observational input. Moreover, the processes are much less governed by chaotic dynamics than before. It turned out that EKF and IKF3 perform almost identically, with slightly smaller errors of about 1–2% for the EKF, and errors which are about 5% larger for the other methods. All regime transitions are traced almost instantly.

If we keep the variance of the forcing noise small and only reduce the observational error, the chaotic nature of the process is increased. With this setting we made a series of experiments all of which point in the same direction. The EKF and IKF3 still perform well, but IKF2 and IKF1 are much worse now. Using an observational error of $\psi = 0.01$ the EKF and IKF3 results are given in Table 9 and 10. The two are quite comparable; only for very small noise levels does the EKF perform up to 3% better. For increasing noise, both methods approach an assimilation error of 100%, that is, the quality of pure observations.

Extremely large errors of up to 990% occur for the IKF2, compare Table 11, when the system noise is small. This is certainly due to a persistent failure in tracing the regimes correctly, which now weighs much stronger with the reduced observational noise (which is the scale for the Δ measures). The situation is better for the IKF1, Table 12, but still there are many false regimes involved for small system noise.

We finally performed assimilations with missing observa-

Table 10. The Δ Matrix of the IKF3 for the Lorenz System With an Observational Noise Scale of $\psi = 0.01$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	6	6	6	8	14	41	84	99	100
−6	6	6	6	8	14	41	84	98	100
−5	6	6	7	8	14	41	84	99	100
−4	10	10	10	10	16	42	84	98	100
−3	23	23	23	23	25	45	84	99	100
−2	57	57	57	57	58	63	86	99	100
−1	91	91	91	91	91	91	93	99	100
0	100	100	100	100	100	100	100	100	100
1	101	101	101	101	101	101	101	101	101

Values of μ and ν are given by their decimal exponents. There is almost no difference from the EKF errors; only for small system noise the errors are slightly larger.

Table 12. The Δ Matrix of the IKF1 for the Lorenz System With an Observational Noise Scale of $\psi = 0.01$

	ν								
μ	−7	−6	−5	−4	−3	−2	−1	0	1
−7	99	88	45	11	13	41	84	99	100
−6	136	122	53	11	13	41	84	98	100
−5	148	139	61	12	13	41	84	98	100
−4	21	21	20	14	15	41	84	98	100
−3	24	24	24	24	25	44	84	99	100
−2	58	58	58	58	58	63	86	99	100
−1	91	91	91	91	91	91	93	99	100
0	100	100	100	100	100	100	100	100	100
1	101	101	101	101	101	101	101	101	101

Values of μ and ν are given by their decimal exponents. As in the IKF2 case, there are larger errors for small system noise, and they decrease with growing ν.

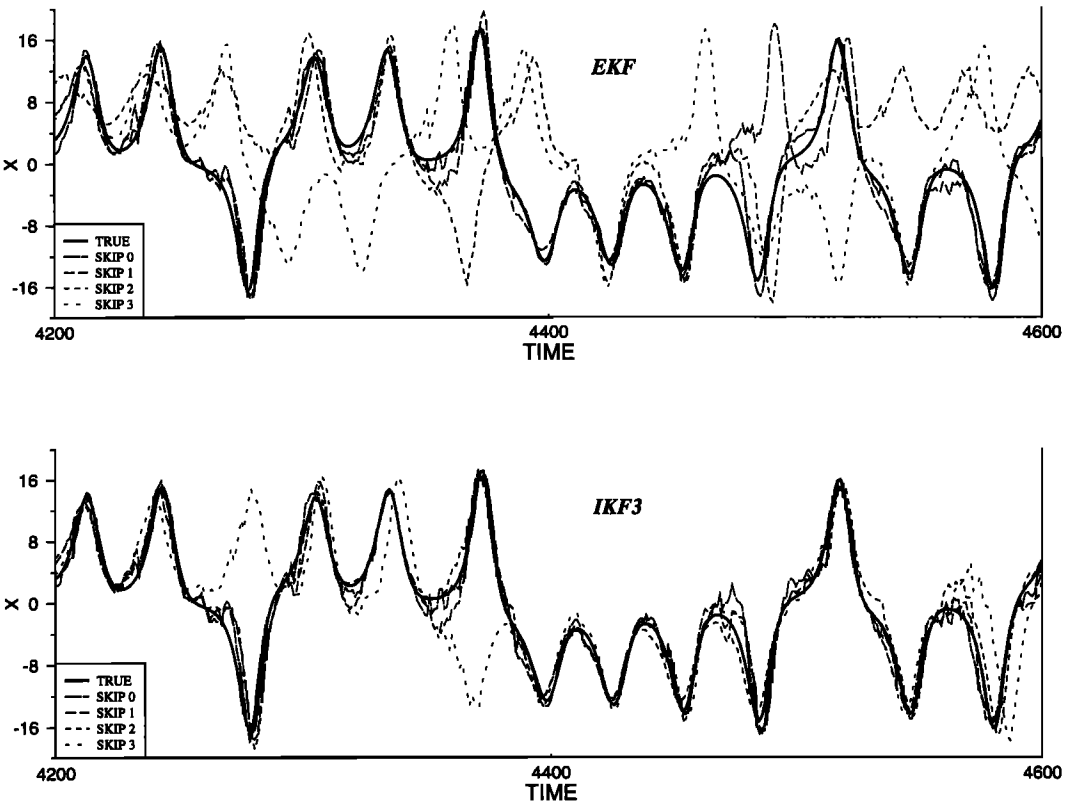

Figure 7. Comparison of EKF and IKF3 with respect to missing observations. (top) The EKF shows a faithful assimilation only for the cases SKIP-0 and SKIP-1. For the other two cases the filter shows a complete failure to trace the systems' evolution. (bottom) The IKF3 performance. Only the case SKIP-3 shows sporadic failures in the tracing of the system. In all other cases the filter stays rather close to the true evolution.

tion. The parameters are those of Figures 6a–6d except that we chose the nonlinear contribution to be $\nu = 0.1$. Figure 7 shows the results for the EKF and IKF3, with observations being withheld for 0, 1, 2, and 3 time steps. Both filters trace the transitions correctly when there is nothing withheld. With observation insertions every second time step, the EKF shows one wrong cycle at time steps 4400–4500, and the IKF3 is without any failure. By further reducing the observation frequency the transition failures become so frequent that the EKF is practically worthless. (This is worse with reduced ν.) The situation is different for the IKF3. The assimilations show a rather close contact with the true values in all but the last experiment where three insertions are withheld.

7. Conclusions

One of the conditions that has to be satisfied for an application of the extended Kalman filter is the smoothness of the assimilating model along possible trajectories. Once this condition is violated the filter quality is degraded drastically. This applies to highly nonlinear systems whose trajectories undergo sharp changes in a relatively short time. These "switches" are characteristic of systems which show regime-like behavior, and the main task for the assimilation is, first, to trace the switches simultaneously and, second, to assimilate correctly inside a regime. As an alternative method we presented the interactive Kalman filter (IKF). The filter makes explicit use of the regimes. By utilizing only the local structure of the regimes the error model of the filter is defined through a procedure that handles the regimes interactively.

To test the IKF and to compare it with the EKF we used two well-known regime-like systems, the double-well system and the Lorenz system, superimposed by a white noise forcing. The EKF failed in tracing regime switches correctly, especially at low noise levels. The incorporation of second-order nonlinear effects into the error model (similar to that of *Miller et al.*, [1993]) could only partly improve the filter results. The assimilation results for the IKF heavily depend on the choice of the participating regimes. When we used all three equilibria as regimes, the IKF performance was much more stable than the EKF. Removing the unstable origin from the set of interacting regimes led, however, to a significant loss in the quality of the assimilation. On the other hand, when using the mean value, which is not an equilibrium in the Lorenz model, as a regime the performance was better than for the IKF2 which used two regimes. We see that choosing which regimes will participate in the IKF is not a trivial task. Moreover, if one was to use more and more regimes, the filter would become as unstable as, and finally equivalent to, the EKF. For the case of equilibrium regimes, however, it seems obvious that the performance will improve with the number of equilibria used.

The lack of crucial equilibria can be fatal, as we have seen for the IKF2. For systems which show very fast error growth, namely chaotic systems, there must be a compara-

ble mechanism of error growth in the corresponding error model of the assimilation. Since there are no unstable equilibria involved, the IKF2 error model's failure is obvious. For the EKF, however, it is not at all easy to sort out the very point where the error model becomes wrong. It appears that the EKF error model reacts too sensitively to the various states during the assimilation, so that the filter finally becomes unstable. Since the IKF3 performs so well in those cases, and since the IKF3 uses "global" information via the equilibria, we argue that it is the local character of the EKF which causes it to fail.

The profit gained by using equilibria (or, more generally, appropriate regimes) has to be paid for by the difficulties that arise in finding them or, to be realistic, finding as many as possible. For simpler systems like the ones we dealt with or, for example, systems that are encountered in the engineering sciences one can often easily write down analytical solutions. However, as soon as the systems become larger and more complex such as those in the geosciences, the equilibria, especially the highly unstable ones, are perfectly hidden in the systems' mathematical formulation as well as in the frizzy output data.

Most likely, one would find equilibria by considering the physics involved in the process. One might even be successful by implementing mathematical methods like, for instance, a conjugate gradient method with the help of a good first guess. This method certainly cannot work without an appropriate prefiltering of the process in order to decrease the number of degrees of freedom. The regime itself, i.e., the local linear model, could then be estimated by the use of Monte Carlo methods.

This method actually worked for the ENSO model of *Zebiak and Cane* [1987]. We found a group of equilibria which look promising in their physical structure as well as in their stability properties. Certainly, the equilibria of that model are of physical interest in their own right. We will use them in a forthcoming study as another test case for the IKF.

Appendix: Nonlinear System Noise

In order to estimate the nonlinear contribution to the assimilation error, we performed Monte Carlo experiments using one long run of 20,000 time steps with no extra forcing. We then created a series of white noise processes $w(t)$ with covariance $\mathbf{C} = \psi\mathbf{I}$, where ψ was one of 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, and superimposed each of these noise processes onto the original process. Given one ρ we determined, from both the undisturbed and the disturbed processes, one-step forecasts for both the full model and its local linearization and calculated their differences; formally this is

$$n(t) = f[x(t) + w(t)] - f[x(t)] - \frac{\partial f}{\partial x}[x(t)]w(t)$$

The nonlinear error covariance, $\mathbf{N} = \langle n(t)n^T(t)\rangle$, is certainly dependent on the different scales of the noise $w(t)$, and modeling of these dependencies is not feasible for our purposes.

As a simplification we applied the following approach. Starting from the assumption that the correlation structure of the nonlinear error is constant only if the error is not too large, we estimated this correlation for each of the processes above. We found that, in fact, the correlation converged to the following matrix:

$$\mathbf{C} = \begin{pmatrix} 1.0 & 0.2 & 0.0 \\ 0.2 & 1.0 & -0.3 \\ 0.0 & -0.3 & 1.0 \end{pmatrix}$$

Here "converged" is used in an informal sense. It means that with $w(t)$ approaching zero, the correlation matrices approached C. The size of the noise is accounted for by a scaling factor ν: $\mathbf{N} = \nu\mathbf{C}$. The best ν has to be found a posteriori.

Acknowledgments. Thanks to Bob Miller for fruitful discussions, especially in the context of missing observations and extensions of the EKF. This work has been supported by NOAA grant NA16-RC-0432-03. This is contribution number 5164 of Lamont-Doherty Earth Observatory of Columbia University.

References

Bucy, R., and P. Joseph, *Filtering for Stochastic Processes With Applications to Guidance*, 2nd ed., Chelsea, New York, 1968.

Budgell, N., Nonlinear data assimilation for shallow water equations in branched channels, *J. Geophys. Res.*, *91*, 10,633–10,644, 1986.

Budgell, N., Stochastic filtering of linear shallow water wave processes, *SIAM J. Sci. Stat. Comput.*, *8*, 152–170, 1987.

Charney, J., and J. DeVore, Multiple flow equilibria in the atmosphere and blocking, *J. Atmos. Sci.*, *36*, 1205–1216, 1979.

Duffing, G., *Erzwungene Schwingungen bei veränderlicher Eigenfrequenz*, F. Vieweg u. Sohn, Braunschweig, Germany, 1918.

Gelb, A., *Applied Optimal Estimation*, MIT Press, Cambridge, Mass., 1974.

Ghil, M., and P. Malanotte-Rizzoli, Data assimilation in meteorology and oceanography, *Adv. Geophys.*, *33*, 141–266, 1991.

Guckenheimer, J., and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.

Jazwinski, A., *Stochastic Processes and Filtering Theory*, Academic, San Diego, Calif., 1970.

Kalman, R. E., A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82D*, 35–45, 1960.

Lacarra, J.-F., and O. Talagrand, Short-range evolution of small perturbations in a barotropic model, *Tellus*, *40A*, 81–95, 1988.

Lorenz, E., Deterministic nonperiodic flow, *J. Atmos. Sci.*, *20*, 130–141, 1963.

Miller, R., M. Ghil, and F. Gauthiez, Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, in press, 1993.

Parrish, D., and S. Cohn, A Kalman filter for a two-dimensional shallow water model, in *Proceedings of the 7th Conference on Numerical Weather Prediction*, pp. 1–8, American Meteorological Society, Boston, Mass., 1985.

Todling, R., and M. Ghil, Kalman filtering for a two-layer, two-dimensional shallow-water model, Paper presented at International Symposium Assimilation of Observations in Meteorology and Oceanography, Am. Meteorol. Soc., Paris, France, July, 1990.

von Storch, H., G. Bürger, R. Schnur, and J. X. von Storch, Principal oscillation patterns: Review Paper, *Rep. 113*, Max Planck Inst. für Meteorol., Hamburg, Germany, 1993.

Zebiak, S., and M. Cane, A model El Nino-Southern Oscillation, *Mon. Weather Rev.*, *115*, 2262–2278, 1987.

G. B. Bürger and M. A. Cane, Lamont-Doherty Earth Observatory, Route 9W, Palisades, NY 10964.

(Received July 13, 1993; accepted November 22, 1993; revised December 28, 1993.)