OCEAN APPLICATIONS TECNICAL NOTE NO 21.

Treatment of Systematic Errors in Sequential Data Assimilation[1]

by

M.J Martin[2], N.K.Nichols[2] and M.J.Bell

Met Office
FitzRoy Road, Exeter, Devon. EX1 3PB

[2]Department of Mathematics
P.O. Box 220
University of Reading
Whiteknights
Reading RG6 6AX
United Kingdom

[1]This research was supported in part by the Natural Environment Research Council and by the UK Meteorological Office

Contents

1 Introduction

Data assimilation is a method for estimating the state of a system by combining observational data with a prior estimate of the state, usually obtained from a numerical model. This is an important technique in the prediction of meteorological and oceanographic variables because we need to know how to combine output from models with observations in order to give the most accurate prediction of the state of the system. This is not a trivial task as the model often contains many inaccuracies, as do the observations, and these errors are not always known. Also, the number of variables is as large as the current computer can handle, currently of order 10^7 in the meteorological and oceanographic problems, so there are limitations due to the amount of computing power available for the data assimilation.

There are two main types of data assimilation. Sequential or 3D methods combine a model output and observational data at one time to produce an analysis. The model is then used to forecast forward in time until another group of observations is reached when another analysis is performed and so on. Examples of this type of method are the Successive Corrections Method [1], [7], the Analysis Correction scheme [2], [24], Optimal Interpolation [27], 3D-VAR, the Kalman filter [21] and the ensemble Kalman filter [12]. The second type of method is the 4D approach. Here a model forecast is compared with observations over a long period of time so that some kind of best fit can be obtained over the entire time period to produce an analysis. These methods include 4D-VAR [28], [31], the PSAS method [6] and the Kalman smoother [5]. It can be shown that the Kalman filter and 4D-VAR produce the same solution at the end of a time period for linear unbiased problems.

All of the methods mentioned above make assumptions about the statistics of the observational error and apart from the Kalman filter assume there exists a perfect model of the system. These assumptions include random observational error with zero mean and the Kalman filter assumes the model error is of a random nature. In reality there are often systematic errors in the observations and in the model. These biases might be constant in time but could also change according to seasonal variations for example. If observational biases are known, then there is no problem as they can easily be subtracted out. The model biases are not usually known, however, so they need to be treated in a different way.

In 1969, Friedland [13] described a method for treating a constant bias in the Kalman filter. The method he proposed has since been extended by many authors to include a time-varying bias ,[30], which contains random errors, [20], and also to nonlinear systems, [32]. Dee *et al.* [10] also describes ways of treating bias in the forecast using the Kalman filter. Derber, [11] and Griffith *et al.*, [18], [19], give methods for accounting for bias in the 4D-VAR data assimilation method. This report shows that the type of method given in [13] and [10] can be used in all the sequential data assimilation methods whilst demonstrating how different types of systematic errors can be treated. The problem of noise in the observations leading to a noisy estimate of the bias is noted and a possible remedy put forward.

In this report, four of the sequential methods will be examined, namely the Successive Corrections method, Optimal Interpolation, the Analysis Correction scheme and the Kalman filter. Section 2 reviews the theory behind the methods considered. In Section 3, the models which have been used to test the methods are described and the results shown. Section 4 describes the theory and gives some examples of accounting for bias in data assimilation. Section 5 contains the conclusions reached.

2 Data Assimilation Methods

The sequential data assimilation problem can be described as:

At each time step i, minimise the cost functional

$$\begin{aligned} \mathcal{J} &= (\underline{y}(i) - H(\underline{x}(i))^T R^{-1}(\underline{y}(i) - H(\underline{x}(i))) \\ &+ (\underline{x}(i) - \underline{x}_b(i))^T B^{-1}(\underline{x}(i) - \underline{x}_b(i)) \end{aligned} \tag{1}$$

with respect to the $\underline{x}(i)$.

Here $\underline{x}_b(i)$ is a background state with error covariance matrix B, $\underline{y}(i)$ is a vector of observations containing random white noise with covariance matrix R and H is the observation operator which interpolates from model fields to observation values. The $\underline{x}_a(i)$ is the analysis produced from the solution to this problem and has error covariance matrix A. The analysis produced from the direct minimisation of this cost function is called the 3D-VAR solution and has analysis error covariance matrix $A = (1/2)\mathcal{J}''$ [6], [22]. This method will not be discussed further but is included for completeness.

2.1 Optimal Interpolation

The solution to the minimisation problem in the previous section can be expressed in the linear case as

$$\underline{x}_a(i) = \underline{x}_b(i) + K[\underline{y}(i) - H(\underline{x}_b(i))], \tag{2}$$

$$K = BH^T(HBH^T + R)^{-1}, \tag{3}$$

where K is called the gain matrix, [22]. If H is a nonlinear operator, it should be linearised around a background state. This matrix is then used in equation (3). The solution $\underline{x}_a(i)$ is called the Optimal Interpolation (OI) solution and has analysis error covariance $A = (I - KH)B$ [27],[22]. This method is not strictly optimal over a period of time, however, because the covariance matrices B and R are not evolved with time and are not usually well known. It is therefore sometimes known as Statistical Interpolation.

2.2 The Successive Corrections Method

The Successive Corrections method (SCM) was one of the first data assimilation techniques to be implemented in practical problems. Bergthorsson and Doos [1] were the first to introduce the method in 1955, followed shortly after by Cressman [7]. Corrections are made to a first guess or background state by adding a weighted difference between the observations and the background. This background state can be an output from a model or the climatology of the system. The SCM algorithm can be written as

$$\underline{x}_a^{j+1}(i) = \underline{x}_a^j(i) + W[\underline{y}(i) - H(\underline{x}_a^j(i))], \tag{4}$$

where $\underline{x}_a^0(i) = \underline{x}_b(i)$ is the background state, $\underline{x}_a^k(i)$ is the analysis after k corrections and W is a weighting matrix.

The main point of the method is how we choose the weighting matrix. At first, this matrix was chosen empirically. However, over the years since the method was introduced there have been many suggestions as to the best and most efficient choices. In [7] the weights are chosen to smooth the observations into the analysis so that there are no sharp jumps in the solution.

In [8] Daley shows that if the weighting matrix is chosen so that the SCM converges, then it will always converge to the observations as the number of corrections increases. For this reason, the corrections are usually stopped after only a few iterations to prevent noisy analyses.

2.3 The Analysis Correction Scheme

The Analysis Correction (AC) scheme is a modified SCM. Here, corrections are made to the observations as well as the background state to take into account the error in the observations. In [2] a particular weighting matrix is derived using the method of OI as a reference to obtain the optimal weightings. Lorenc, Bell and MacPherson [24] also give a derivation of the method and show that it converges to the OI solution as the number of iterations increases. The AC algorithm can be written in the linear case as

$$\underline{x}^{j+1}(i) = \underline{x}^{j}(i) + WQ[\underline{y}^{j}(i) - H(\underline{x}^{j}(i))], \tag{5}$$

$$\underline{y}^{j+1}(i) = \underline{y}^{j}(i) - Q[\underline{y}^{j}(i) - H(\underline{x}^{j}(i))], \tag{6}$$

where $\underline{y}^0(i)$ is the vector of observations, $W = BH^TR^{-1}$, $Q = (HW + I)^{-1}$ and $\underline{x}^k(i), \underline{y}^k(i) \longrightarrow \underline{x}_a(i)$ as $k \longrightarrow \infty$. If H is a nonlinear operator then it should be linearised about a background state. This matrix is then used in the formulation of W and Q.

Versions of this scheme are currently in operation at the UKMO in the ocean FOAM model. In the operational scheme, equation (6) is ignored and the observations are used without any correction. Also, an approximation to Q is made to avoid inverting the matrix $HW + I$.

2.4 The Kalman Filter

The Kalman filter is the optimal method over a period of time for linear problems satisfying its assumptions, [16], [17], [21]. The main distinctions between this and the other methods is that the error covariance matrices are evolved with the analysis and random model error is taken into account. In other words, we are solving a minimisation similar to equation (1) but with time as an additional dimension. The solution to this minimisation problem is known as the 4D-VAR analysis. This gives the same solution at the end of a time period as the Kalman Filter for linear problems.

In the Kalman filter, it is assumed that the model error is unbiased, the model error covariance $Q(i)$ is known and the analysis and model errors are mutually uncorrelated. We now change the notation by denoting the background and analysis error covariance matrices B and A by P_f and P_a respectively. With these notations and assumptions, the Kalman filter algorithm can be written as

$$K(i) = P_f(i)H^T(i)[H(i)P_f(i)H^T(i) + R(i)]^{-1}, \tag{7}$$

$$\underline{x}_a(i) = \underline{x}_f(i) + K(i)[\underline{y}(i) - H(i)\underline{x}_f(i)], \tag{8}$$

$$P_a(i) = [I - K(i)H(i)]P_f(i), \tag{9}$$

$$\underline{x}_f(i+1) = M_{i \to i+1}[\underline{x}_a(i)], \tag{10}$$

$$P_f(i+1) = M_{i \to i+1} P_a(i) M^T_{i \to i+1} + Q(i), \tag{11}$$

where $M_{i \to i+1}$ is the model operating between time i and $i+1$ and the analyses are given by the sequence of $\underline{x}_a(i)$.

For a nonlinear problem, the Kalman filter can be extended to give the extended Kalman filter (EKF). Here, the nonlinear model is linearised around a background state so that

$$M_{i \to i+1}[\underline{x}(i)] - M_{i \to i+1}[\underline{x}_b(i)] = \mathbf{M}_{i \to i+1}[\underline{x}(i) - \underline{x}_b(i)] + O(\Delta x^2), \tag{12}$$

where $\mathbf{M}$ is called the tangent linear model. We also have to linearise the observation operator $H(i)$. The only changes that are made to the algorithm (7)-(11) are that the tangent linear model is now used to evolve the error covariance matrices rather than the full model as before and the linearised observation operator must be used. These approximations have the effect of destroying the optimality of the EKF [3].

3 The Models

Two models have been used to investigate the similarities and differences between the four data assimilation techniques described in Section 2. The first is a simple linear system for which we expect all of the assimilation methods to do well. The second is the chaotic nonlinear Lorenz equations, which, it is hoped, will provide an insight into how well these methods are likely to perform when applied to more complicated problems.

3.1 Damped Oscillating System

The damped oscillating system is given by the ordinary differential equation

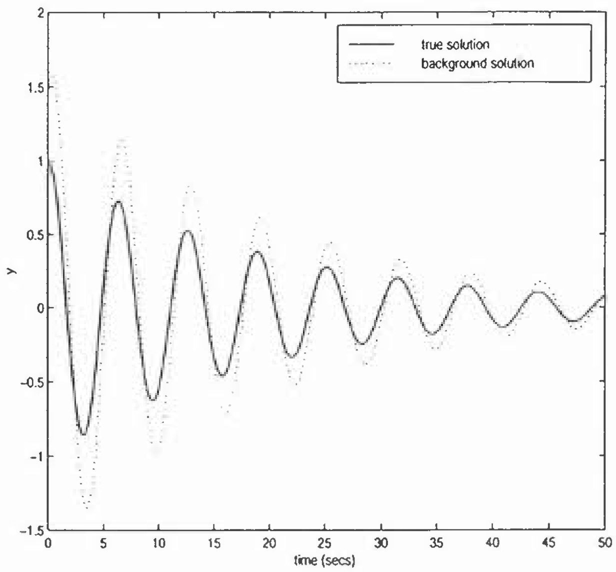
$$\ddot{y} = -k\dot{y} - ny \tag{13}$$

where the damping k and the square of the frequency n are given values of 0.1 and 1 respectively. The differential equation (13) can be expressed as the first order system

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -n & -k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{14}$$

This system has been discretised using a second order Runge-Kutta method with a time step of 0.1. The 'true' solution is given by the discretised equations with initial conditions of $y_1(0) = 1$ and $y_2(0) = 0$. The background or initial guess has initial conditions of $y_1(0) = 1.5$ and $y_2(0) = 0.5$. These two solutions are shown in Figure 1.

Figure 1: The damped oscillating system



Observations of y_1 and y_2 are taken from the true solution at every other time step from t=0 to t=25, with and without random noise added. The random noise is Gaussian with a mean of zero and a variance of 0.01.

Experiment 1 *The analysis is performed on the damped oscillating system using all four assimilation methods from time* 0 *to* 25 *where there are observations every other time step. The analysis at the end of this period is then forecast to time* 50.

3.2 Lorenz Equations

The Lorenz equations are a nonlinear system of three ordinary differential equations. They were originally obtained from the first terms in a Fourier truncation of the flow equations governing thermal convection [26], and are given by

$$\dot{x} = -\sigma(x - y) \tag{15}$$

$$\dot{y} = \rho x - y - xz \tag{16}$$

$$\dot{z} = xy - \beta z. \tag{17}$$

The parameters σ, ρ and β are chosen to have values of 10, 28 and 8/3, respectively, which gives the system chaotic solutions [29]. At these values the system has three equilibrium points, an unstable saddle point at the origin and two unstable spiral points at the coordinates

$$(\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho - 1).$$

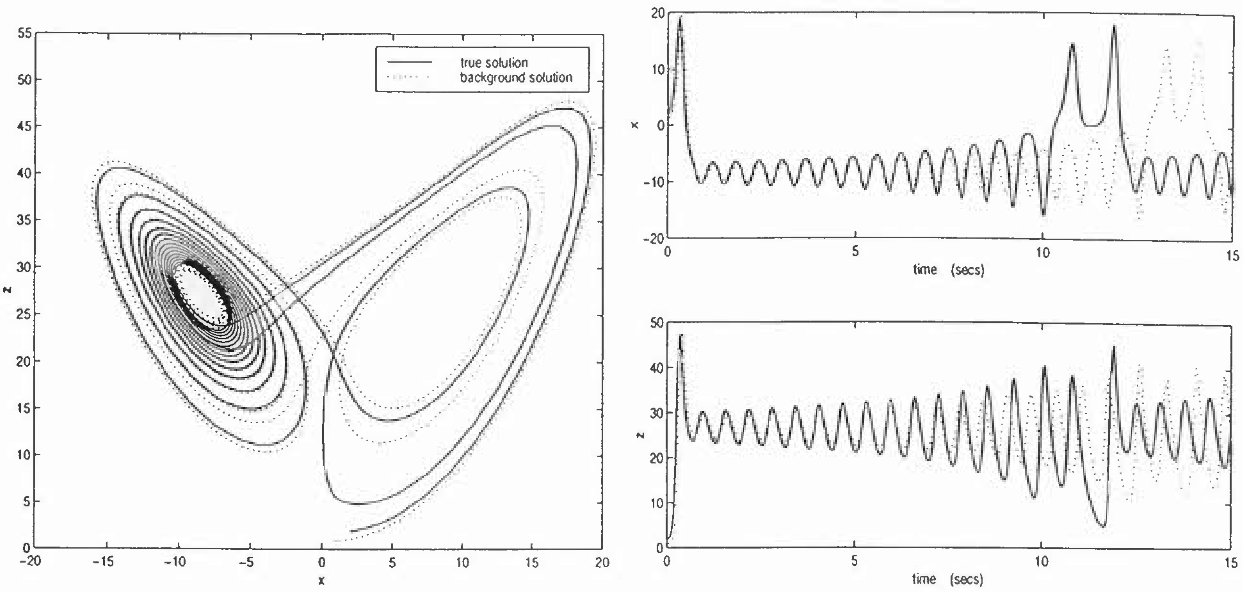
The system has been discretised using a second order Runge-Kutta method with time steps of 0.01. The 'true' solution is given by the discrete system with initial conditions $x_0 = y_0 = z_0 = 2$. The background has initial conditions $x_0 = y_0 = z_0 = 1$. Both solutions are shown in Figure 2.

Observations of x, y and z are obtained from the true solution with and without noise added. This noise has a Gaussian distribution with zero mean and a variance of 0.01.

Experiment 2 *The analysis is performed on the Lorenz equations using all four assimilation methods from time* 0 *to* 7.5 *where there are observations every three time steps which are (a) perfect (b) noisy. A forecast is then made to time* 15.

Experiment 3 *The analysis is performed on the Lorenz equations using all four assimilation methods from time* 0 *to time* 20 *with* 6 *evenly spaced observations which are (a) perfect (b) noisy.*

Figure 2: The Lorenz equations



3.3 Results

Figures corresponding to the experiments not shown in this section can be seen in the Appendix.

3.3.1 The Oscillating System

Experiment 1

In the SCM, the weighting matrix W was chosen to be $0.5 \times I$ and 2 corrections were performed at each observation time. As has been said, the SCM converges to the observations as the number of corrections increases, which is good if the observations are perfect, but not good if the observations contain noise. This is shown in Figure 3.

In the AC and OI schemes, the background and observation error covariance matrices were chosen to be I and $0.1 \times I$ respectively when the observations were perfect. When noise was present, the background error covariance matrix was equal to I and the observation error covariance matrix was equal to I multiplied by the variance of the noise, 0.01. The AC scheme was iterated twice at each observation time. As can be seen from Figures 4 and 5, the two schemes have produced an identical analysis when there is noise present in the observations. Both methods produce a fairly noisy analysis, but they are an improvement on the SCM which produced sharper jumps. The SCM can be made to perform more smoothly when noise is present by choosing the weighting function in a different way and by performing fewer iterations.

When implementing the Kalman filter, the forecast error covariance matrix at $t = 0$ was chosen to be I, the observation error covariance matrix was chosen to be $0.1 \times I$ when the observations were perfect and I multiplied by the variance of the noise when noise was added. The model error covariance, $Q(i)$ was always set to be I. The Kalman filter produces the most accurate results when noise is present in the observations as shown in Figure 6. The analysis follows the true solution very closely and seems to ignore the observational noise which is what was expected in this simple linear case, given the correct statistical information.

Figure 3: Experiment 1. Successive Corrections assimilation on the damped oscillating system: (i) with perfect observations (ii) with noise on the observations

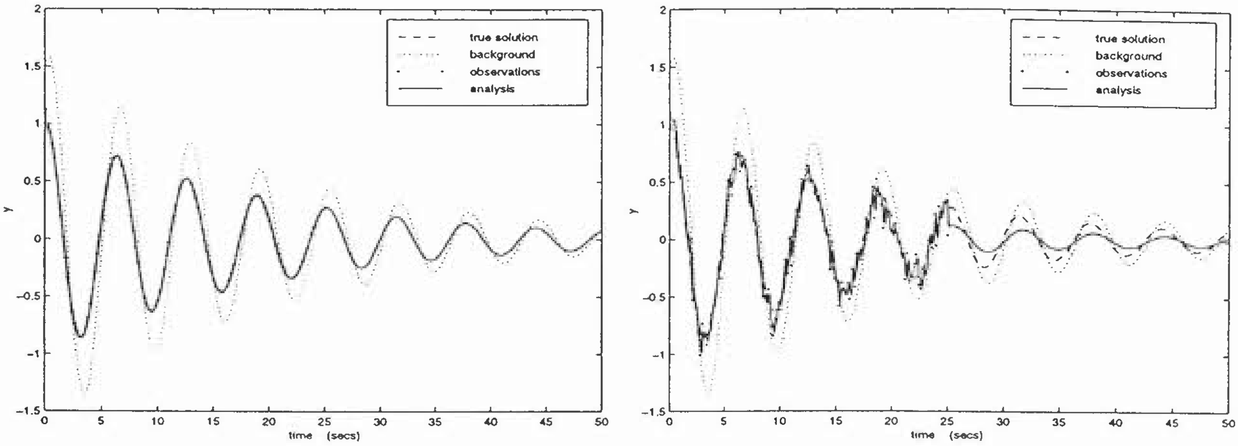


Figure 4: Experiment 1. Analysis Correction assimilation on the damped oscillating system: (i) with perfect observations (ii) with noise on the observations

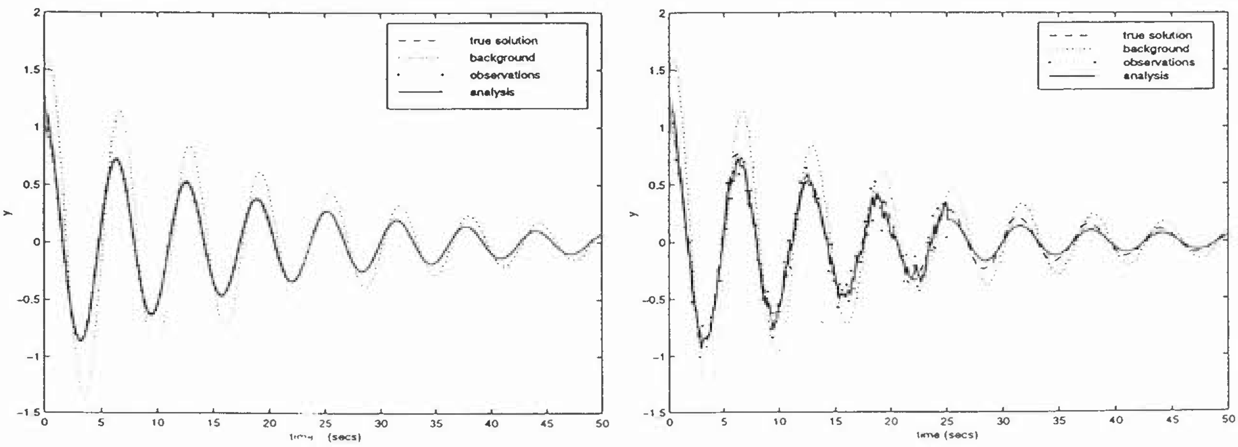


Figure 5: Experiment 1. Optimal Interpolation assimilation on the damped oscillating system: (i) with perfect observations (ii) with noise on the observations

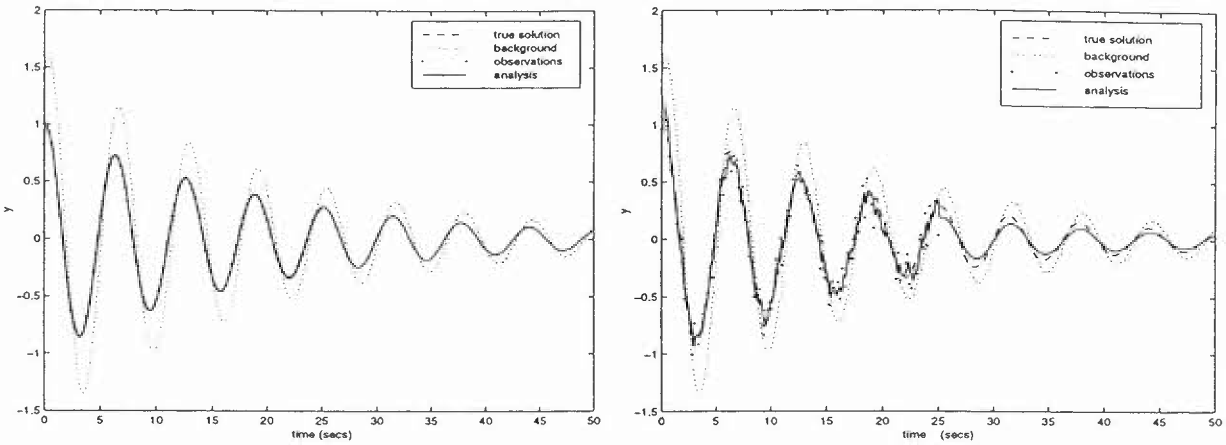
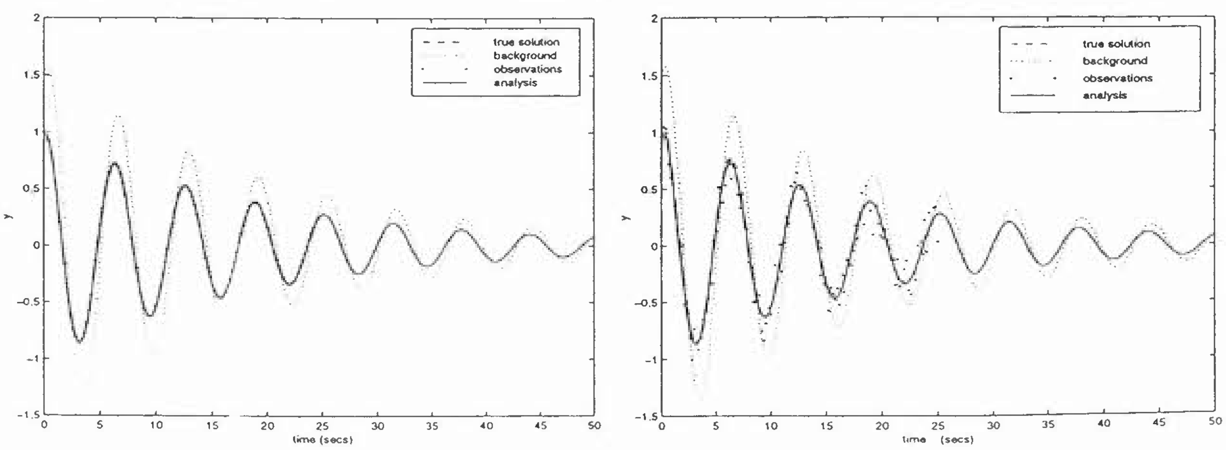


Figure 6: Experiment 1. Kalman filter assimilation on the damped oscillating system: (i) with perfect observations (ii) with noise on the observations



3.3.2 The Lorenz equations

Experiment 2

When the assimilation methods were performed on the Lorenz equations with perfect observations, all of the schemes did very well and managed to follow the true solution exactly. This is shown by the fact that the forecast still agrees with the true solution even at the end of the time period in Figure 7.

Due to the chaotic nature of the Lorenz equations, none of the methods produce a good forecast for very long after the analysis has finished if there is any noise on the observations as shown in Figure 8. This is because a small perturbation in the initial conditions produces a very different solution. Figure 9 shows that the EKF appears to have produced the worst analysis as the forecast seems to move around the wrong attractor after 12 seconds. Although it is understood that the EKF is not optimal for nonlinear systems, this is still surprising.

Experiment 3

With perfect observations, the SCM did not follow the true solution after about 11 seconds of assimilation, as shown in Figure 10. This is probably because at this point the true solution begins to alternate between the two attractors more frequently. Sharp jumps in the analysis can be seen when a new observation is reached. In practice this would be avoided by smoothing in the observation over a time window. The AC scheme also did fairly badly at about the same point in the experiment. If these two methods are iterated more times at each observation point, they produce better solutions. The analysis using the AC scheme after 2 and 5 iterations is shown in Figures 11 and 12. After 5 iterations the AC analysis appears to have converged and is now similar to the OI solution. OI and the Kalman filter produced very good analyses which follow the true solution throughout the time interval.

When the experiments were run with noisy observations, Figure 13 shows that the OI analysis followed the true solution for about 10 seconds but failed to go with the true solution around the correct attractor after that. The EKF produced the best analysis, as shown in Figure 14. This is a surprisingly good analysis given the number of observations. This is probably because the observation at 12 seconds is close to the true solution and brings the analysis back on track after deviating slightly in the preceding second. The other three methods did not follow the true solution after it had started to switch between the two attractors, which is not surprising given the number of observations.

Figure 7: Experiment 2a. Assimilation on the Lorenz equations with perfect observations using the SCM

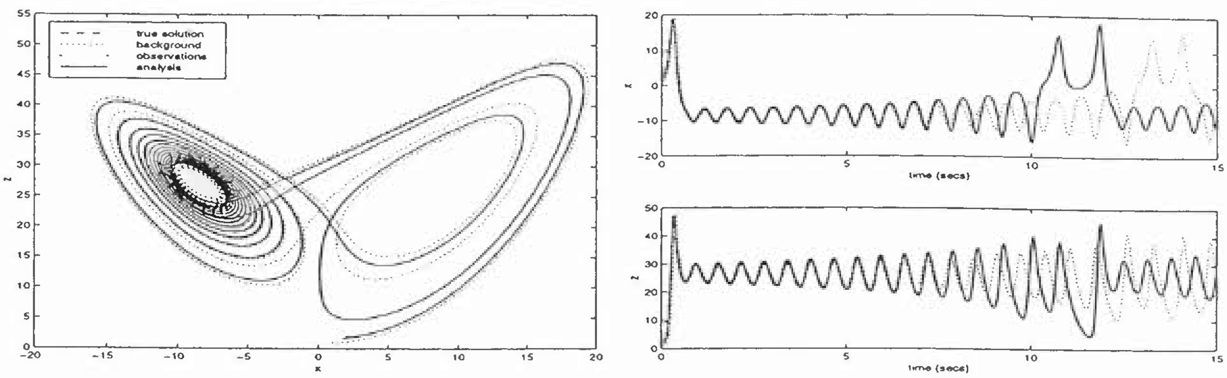


Figure 8: Experiment 2b. Assimilation on the Lorenz equations with noise on the observations using OI

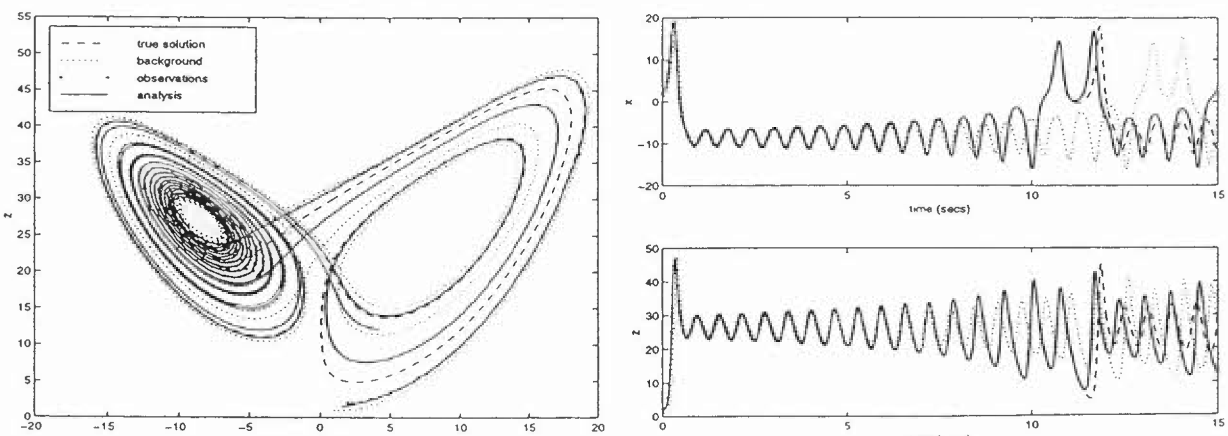


Figure 9: Experiment 2b. Assimilation on the Lorenz equations with noise on the observations using the EKF

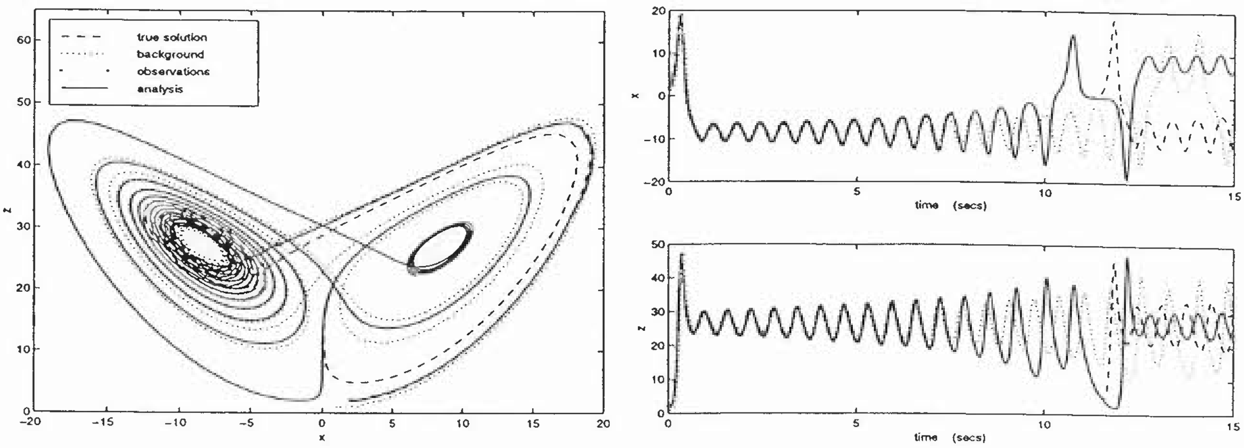


Figure 10: Experiment 3a. Assimilation on the Lorenz equations with perfect observations using the SCM

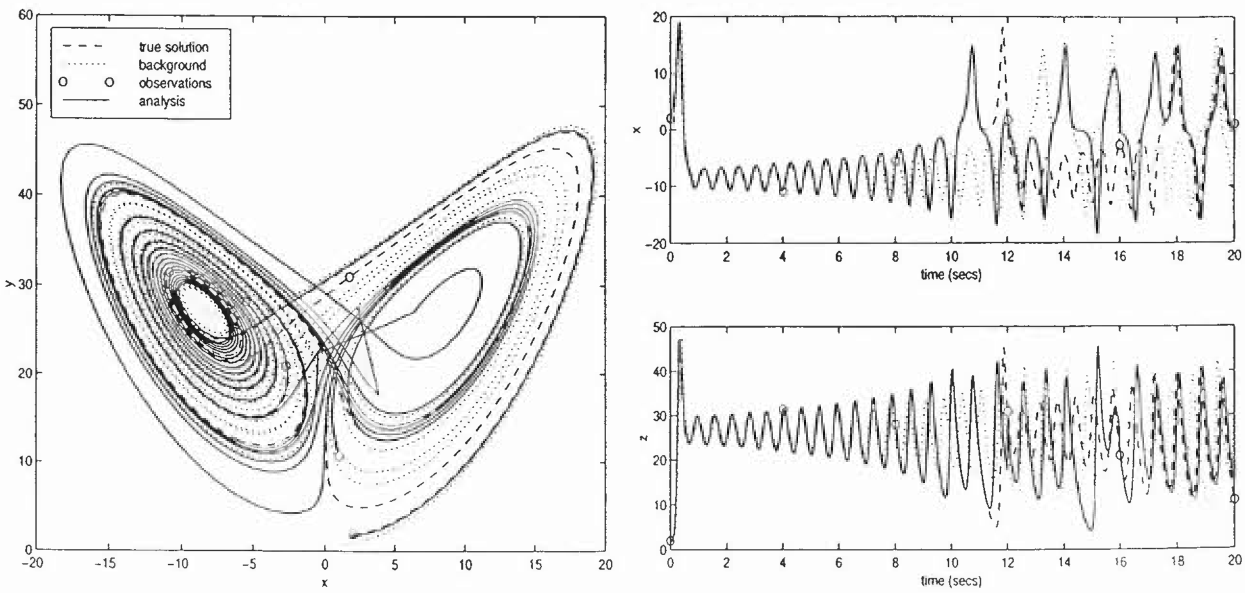


Figure 11: Experiment 3a. Assimilation on the Lorenz equations with perfect observations using the AC scheme after 2 iterations

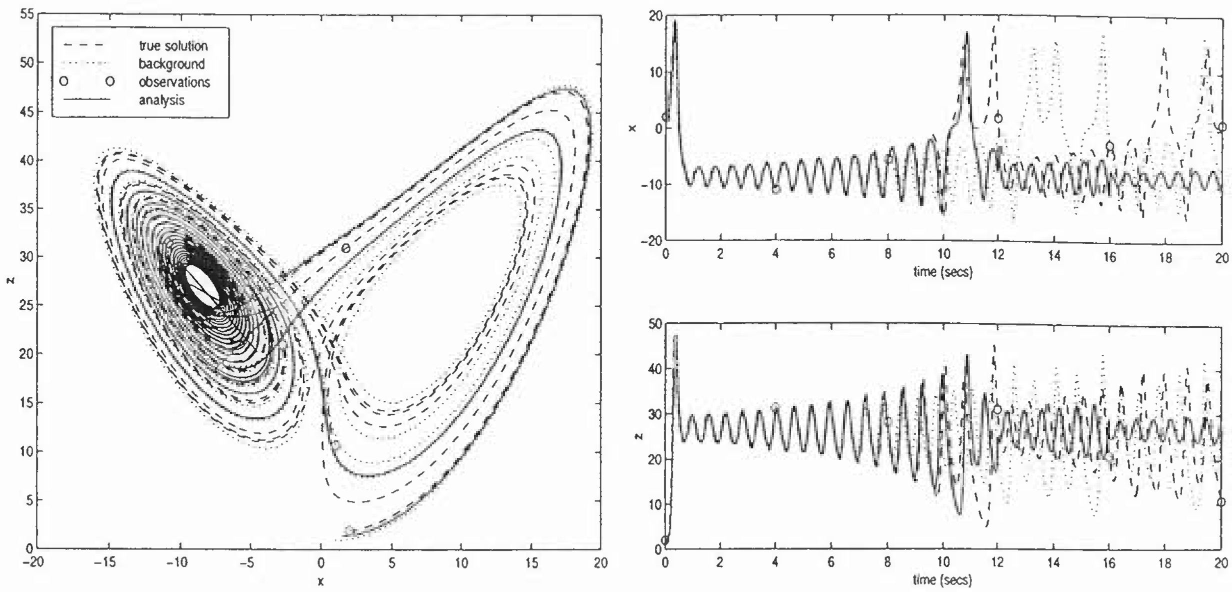


Figure 12: Experiment 3a. Assimilation on the Lorenz equations with perfect observations using the AC scheme after 5 iterations

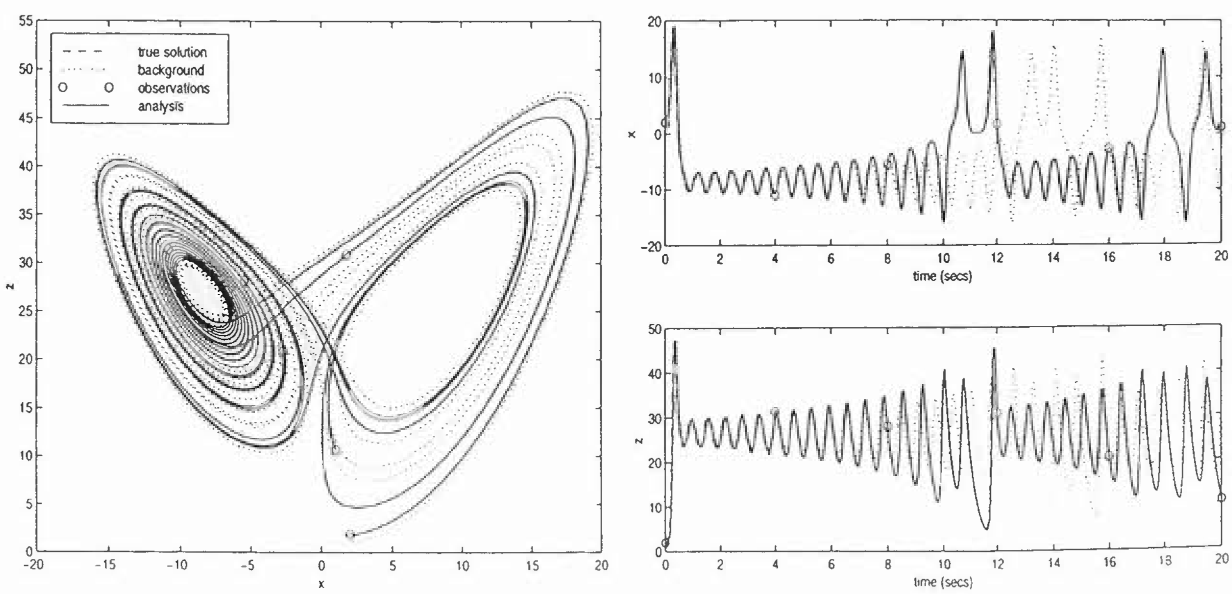


Figure 13: Experiment 3b. Assimilation on the Lorenz equations with noise on the observations using OI

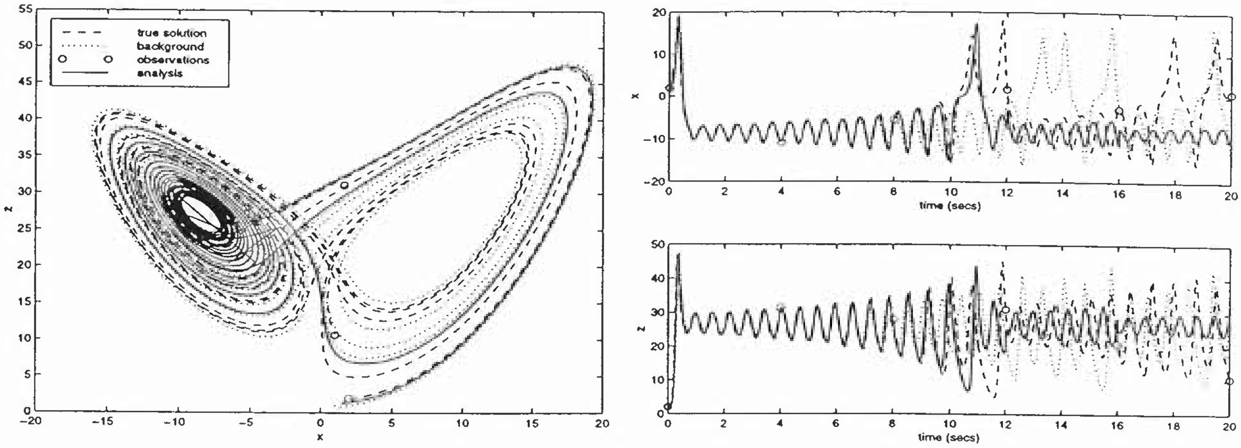
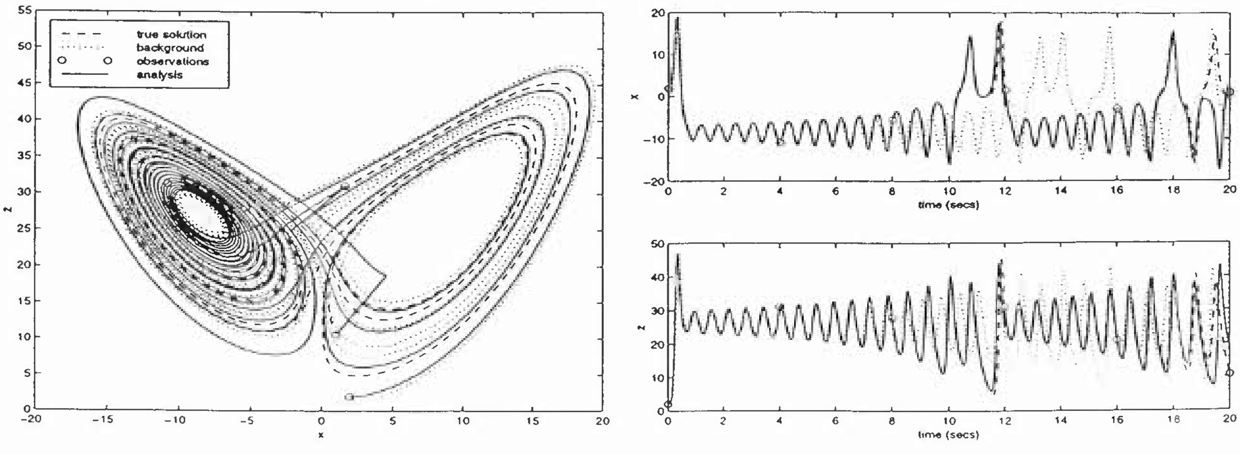


Figure 14: Experiment 3b. Assimilation on the Lorenz equations with noise on the observations using the EKF



4 Accounting for Bias

In reality there are often processes which affect the model and observations in systematic ways. For instance observing instruments may have a bias error which is unknown. Also there are many ways in which the model of a system will be different from the true system. Errors may occur due to the finite approximation of an infinite system, use of the wrong coefficients or incorrect boundary conditions to name a few. These types of errors are not accounted for in the traditional data assimilation methods and therefore need special treatment.

4.1 General Theory

4.1.1 Convergence of the analysed state

Using the notation of previous sections the system of the model and observations is written as

$$\underline{x}(i+1) = M_{i \to i+1}[\underline{x}(i)], \tag{18}$$

$$\underline{y}(i+1) = H_{i+1}[\underline{x}(i+1)], \tag{19}$$

where $\underline{x}(0)$ is unknown. The data assimilation problem for this system has so far been solved by the usual sequential methods. It has been shown that with a perfect model this process works well. However, if we have some unknown bias in the model, the methods described so far do not take this into account (see Section 4.2).

If we assume that the model is biased then we can write the system as

$$\underline{x}(i+1) = M_{i \to i+1}[\underline{x}(i)] + \underline{b}(i), \tag{20}$$

$$\underline{y}(i+1) = H_{i+1}[\underline{x}(i+1)], \tag{21}$$

where $\underline{b}(i)$ is the bias in the model at time i and $\underline{x}(0)$ and $\underline{b}(0)$ are unknown. There is now the problem of solving for the state and the bias at some later time. If we implement the sequential data assimilation process for this system then the equations become

$$\underline{x}_b(i+1) = M_{i \to i+1}[\underline{x}_a(i)] + \underline{b}(i), \tag{22}$$

$$\underline{x}_a(i+1) = \underline{x}_b(i+1) + W_i(\underline{y}(i+1) - H_{i+1}[\underline{x}_b(i+1)]). \tag{23}$$

We now assume that the model for the state and the model for the bias can be linearised to produce the tangent linear models as in Section 2.3 and that the observation operator can also be linearised. We also assume that the matrices M, H and W are constant matrices. Combining equations (22) and (23) under these assumptions gives the evolution of the analysed state as

$$\begin{aligned} \underline{x}_a(i+1) &= (M - WHM)\underline{x}_a(i) + (I - WH)\underline{b}(i) \\ &+ W\underline{y}(i+1). \end{aligned} \tag{24}$$

If we now take the difference between this analysed state and the true state $\underline{x}_t(i)$ obtained from equation (18), assuming that the observations come from the true state, we get the error equation

$$\underline{e}(i+1) = (M - WHM)\underline{e}(i) - (I - WH)\underline{b}(i), \tag{25}$$

where $\underline{e}(i) = \underline{x}_t(i) - \underline{x}_a(i)$. This shows that if our model has a bias then the error between the true and model state will be forced by that bias. One possible way of overcoming this problem is to assume that the bias follows some model, N, which is known [18]. The system now becomes

$$\underline{x}(i+1) = M_{i\to i+1}[\underline{x}(i)] + \underline{b}(i), \tag{26}$$

$$\underline{b}(i+1) = N_{i\to i+1}[\underline{b}(i)], \tag{27}$$

$$\underline{y}(i+1) = H_{i+1}[\underline{x}(i+1)]. \tag{28}$$

We can then augment the state vector to include the vector of biases so that the system can be written in the same form as in previous sections. We write

$$\begin{aligned} \underline{z}_b(i+1) &= \begin{pmatrix} \underline{x}_b(i+1) \\ \underline{b}_b(i+1) \end{pmatrix} = \begin{pmatrix} M_{i\to i+1} & I \\ 0 & N_{i\to i+1} \end{pmatrix} \begin{pmatrix} \underline{x}_a(i) \\ \underline{b}_a(i) \end{pmatrix} \\ &\equiv F_{i\to i+1}\underline{z}_a(i), \end{aligned} \tag{29}$$

$$\underline{y}(i+1) = \begin{pmatrix} H & 0 \end{pmatrix} \begin{pmatrix} \underline{x}_t(i+1) \\ \underline{b}_t(i+1) \end{pmatrix} \equiv G\underline{z}_t(i+1), \tag{30}$$

$$\underline{z}_a(i+1) = \underline{z}_b(i+1) + \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} [\underline{y}(i+1) - G\underline{z}_b(i+1)], \tag{31}$$

where $\underline{z}$ is the augmented state vector.

When performing the sequential data assimilation methods on this augmented state, assuming M, N and H to be linear and time invariant, the evolution of the analysed state now becomes

$$\begin{aligned} \underline{z}_a(i+1) &= \begin{pmatrix} M - W_1HM & I - W_1H \\ -W_2HM & N - W_2H \end{pmatrix} \underline{z}_a(i) \\ &+ \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \underline{y}(i+1). \end{aligned} \tag{32}$$

The true state is now given by

$$\underline{z}_t(i+1) = \begin{pmatrix} M & I \\ 0 & N \end{pmatrix} \underline{z}_t(i). \tag{33}$$

The difference between the true and analysed states, assuming that the observations come from the true state, is now given by

$$\underline{e}(i+1) = \begin{pmatrix} M - W_1HM & I - W_1H \\ -W_2HM & N - W_2H \end{pmatrix} \underline{e}(i), \tag{34}$$

where $\underline{e}(i) = \underline{z}_t(i) - \underline{z}_a(i)$. We now see that as long as the spectral radius of the matrix in equation (34) is less than 1, the analysed state will converge to the true state as $i \longrightarrow \infty$.

It is also possible to treat observational bias in a similar way, that is, to augment the state vector to include a vector of biases. The observation operator is now changed so that if the bias is assumed to be constant then

$$\underline{z}(i+1) = \begin{pmatrix} M_{i\rightarrow i+1} & 0 \\ 0 & I \end{pmatrix} \underline{z}(i), \tag{35}$$

$$\underline{y}(i+1) = \begin{pmatrix} H & I \end{pmatrix} \underline{z}(i+1). \tag{36}$$

4.1.2 Applying the Kalman filter to the augmented state

In the previous section, the models for the state and bias evolution were assumed to be perfect, as were the observations. If we now generalise and add random errors to these components, the system becomes

$$\begin{aligned}
\underline{x}(i+1) &= M_{i\to i+1}[\underline{x}(i)] + \underline{b}(i) + \underline{\xi}_x(i), & (37)\\
\underline{b}(i+1) &= N_{i\to i+1}[\underline{b}(i)] + \underline{\xi}_b(i), & (38)\\
\underline{y}(i) &= H_i[\underline{x}(i)] + \underline{\eta}(i), & (39)
\end{aligned}$$

where the Greek letters denote random vectors, $\underline{\xi}_x(i)$ has covariance matrix $Q_x(i)$, $\underline{\xi}_b(i)$ has covariance matrix $Q_b(i)$ and $\underline{\eta}(i)$ has covariance matrix R_i. Now assume that $M_{i\to i+1}$, $N_{i\to i+1}$, and H_i are linear and let $\underline{z} = \begin{pmatrix} \underline{x} \\ \underline{b} \end{pmatrix}$. Then rewriting the above equations gives

$$\begin{aligned}
\underline{z}(i+1) &= \begin{pmatrix} M_{i\to i+1} & I \\ 0 & N_{i\to i+1} \end{pmatrix} \underline{z}(i) + \begin{pmatrix} \underline{\xi}_x(i) \\ \underline{\xi}_b(i) \end{pmatrix} \\
&= F_{i\to i+1}\underline{z}(i) + \underline{\xi}_i & (40)\\
\underline{y}(i) &= \begin{pmatrix} H_i, & 0 \end{pmatrix} \underline{z}(i) + \underline{\eta}_i = G_i \underline{z}(i) + \underline{\eta}_i. & (41)
\end{aligned}$$

We can now apply the Kalman filter algorithm to this augmented system. This formulation is the same as that in Friedland's paper [13] except for the addition of model error to the model of the bias, [20], and inclusion of a general model for the evolution of the bias, [30]. A form of the extended Kalman filter has also been derived for the case of a time-varying bias with random error for nonlinear systems, [32], but is not described here.

4.1.3 Choice of bias evolution model

There are many possible choices for the model N of the bias, and it is useful if something is known about the nature of the bias *a priori*. For instance, if the bias is constant in time it would be sensible to specify $N = I$. If the bias is due to errors in the system parameters then the tangent linear model for the state of the system might be used to represent the bias, i.e. $N_{i\to i+1} = \mathbf{M}_{i\to i+1}$. If a constant bias is assumed, the state equations for the augmented state are given by

$$\underline{z}(i+1) = F_{i\to i+1}\underline{z}(i), \qquad (42)$$

$$\underline{y}(i+1) = G\underline{z}(i+1), \tag{43}$$

where $F_{i \to i+1} = \begin{pmatrix} M_{i \to i+1} & I \\ 0 & I \end{pmatrix}$ and $G = \begin{pmatrix} H & 0 \end{pmatrix}$.

It is now possible to use the same methods for the assimilation as before, applied to the augmented system. The analysis equations for the four types of assimilation take the same form as in Section 2 but with $\underline{x}$ replaced by $\underline{z}$, $M_{i \to i+1}$ replaced by $F_{i \to i+1}$ and H replaced by G.

4.2 Results

Experiments with bias using the damped oscillating system and the Lorenz equations have been carried out and are presented in this section. The numerical methods used for discretising the oscillating system and the Lorenz equations as well as the step sizes used are the same as those in Section 3. In the oscillating system, the observations are taken every other time step and in the Lorenz equations there is an observation every three time steps. When the observations contain noise, the noise has a mean of zero and a variance of 0.01. The error covariance matrices for the SCM and the Kalman filter have the same values as in Section 3 but are extended to be of the correct dimensions for the augmented state. This means that $W_1 = 0.5$ and $W_2 = 0.5$ in the SCM and in the Kalman filter, the full forecast error covariance matrix at $t = 0$ was chosen to be I and the model error covariance was still set to be I. These covariance matrices do not contain the correct error statistics of the problem, which would be difficult to obtain in practice, but should give an idea of how the method will work when applied to a realistic problem.

4.2.1 Oscillating System

Bias was added to the model by including a constant forcing term in the y_1 and y_2 equations of 0.1 and -0.1 respectively giving

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -n & -k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0.1 \\ -0.1 \end{pmatrix}. \tag{44}$$

The effect this change has on the SCM when there is no special treatment of the bias is shown in Figure 15. This does a bad job of estimating the true state and as soon as the forecast is begun the error escalates as there is nothing to constrain the solution to the unbiased true solution. Figure 16 shows the analysis when the bias correction term was added. The bias has been accounted for and we even have an estimate of the bias with which to produce a more accurate forecast. When noise is added to the observations the results of the bias correction are not so promising. This is due to the fact that the analysed bias fluctuates about the actual bias and the point where the forecast is started is in error. This means that the bias in the rest of the forecast is also in error. This is shown in Figure 17. This problem could be overcome by simply averaging the bias over a period of time before the forecast is started. This has been done in Figure 18 which shows an improvement in the analysis although the error is still large. The bias that has been estimated at the end of the assimilation period is, however, very accurate. An alternative is to average the bias over a time window as the assimilation is running and to move the window along with the assimilation. This has been done using a time window of 5 seconds and the results are shown in Figure 19. This averaging seems to produce a very good estimate of the bias and extends the time over which the forecast is accurate.

The same experiments have also been performed using the Kalman filter to see whether this method will produce better results when noise is present in the observations. Figure 20 shows that the Kalman filter will not produce a good analysis when the oscillating system contains a bias. When this bias is accounted for in the same way as before, the Kalman filter does follow the true solution as shown in Figure 21. Noise is still a problem when using the Kalman filter although we do get a less noisy analysis in Figure 22 than in Figure 17. Averaging at the end of the assimilation has similar effects on the forecast as it had when using the SCM which is shown in Figure 23. Averaging over a window produces a similar forecast to the one produced using the SCM as shown in Figure 24.

To see how well this model for the bias would work when a different type of bias was introduced, the coefficients in the oscillating system were altered. Now k is 0.3 and n is 1.2. The results of using the SCM on this system are shown in Figure 25. It is clear that the bias is not constant as assumed, resulting in an estimated bias that is not accurate, even with perfect observations. The same experiment was performed but using the model of the system for the model of the bias i.e. $N_{i \to i+1} = M_{i \to i+1}$. The results are shown in Figure 26 which shows a forecast which agrees with the true solution for longer than with a constant bias. The results are still not perfect however.

Figure 15: Successive Corrections assimilation on the damped oscillating system with a constant forcing added to the model without bias correction

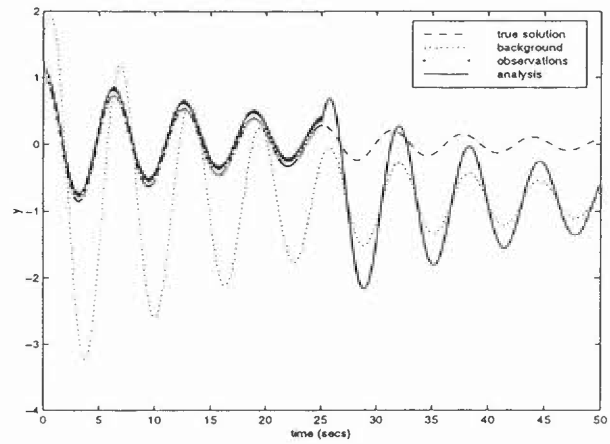
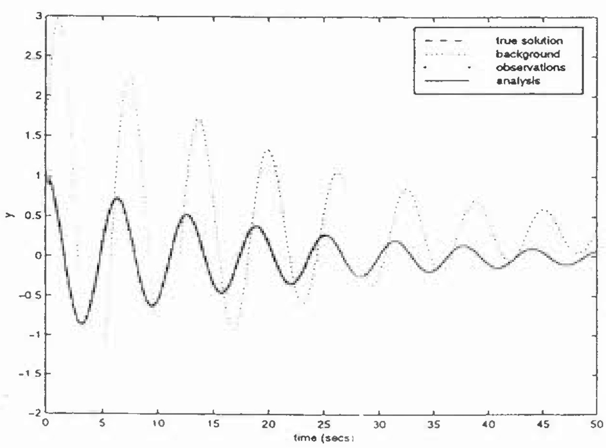


Figure 16: Successive Corrections assimilation on the damped oscillating system with a constant forcing added to the model with bias correction: (i) the state (ii) the analysed bias



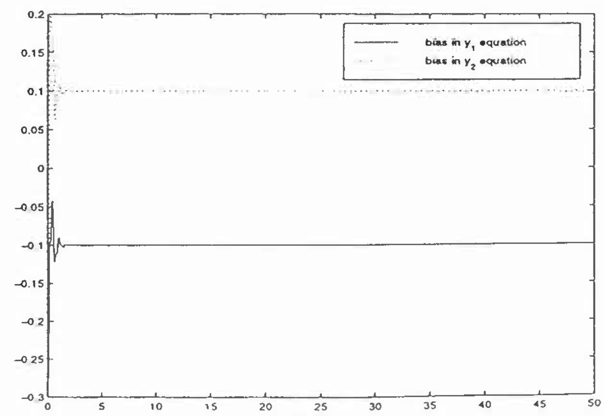


Figure 17: SCM assimilation on the damped oscillating system with bias correction and noise on the observations: (i) the state (ii) the analysed bias

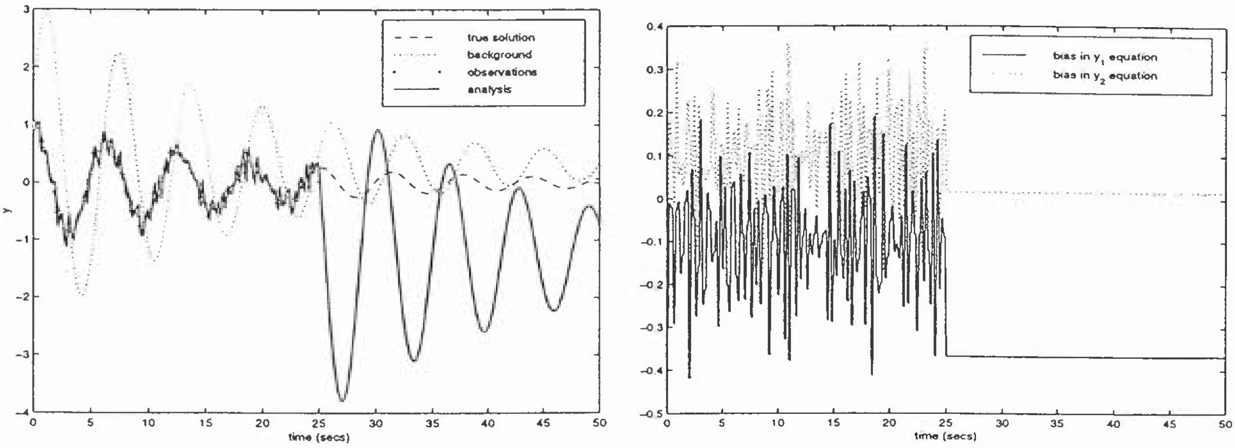


Figure 18: SCM assimilation on the damped oscillating system with bias correction and noise on the observations with averaging of the bias at the end of the assimilation period: (i) the state (ii) the analysed bias

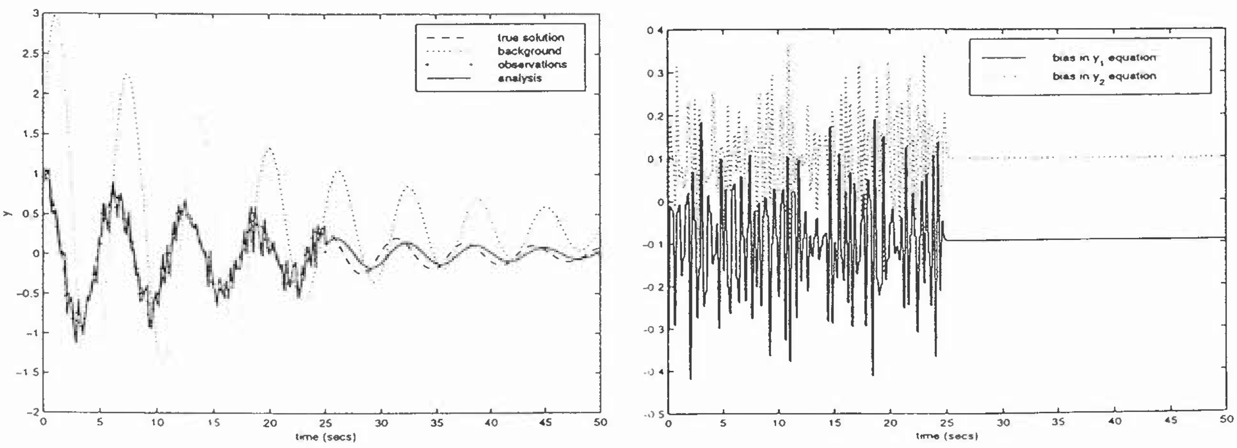


Figure 19: SCM assimilation on the damped oscillating system with bias correction and noise on the observations with averaging of the bias over a moving window of observations: (i) the state (ii) the analysed bias

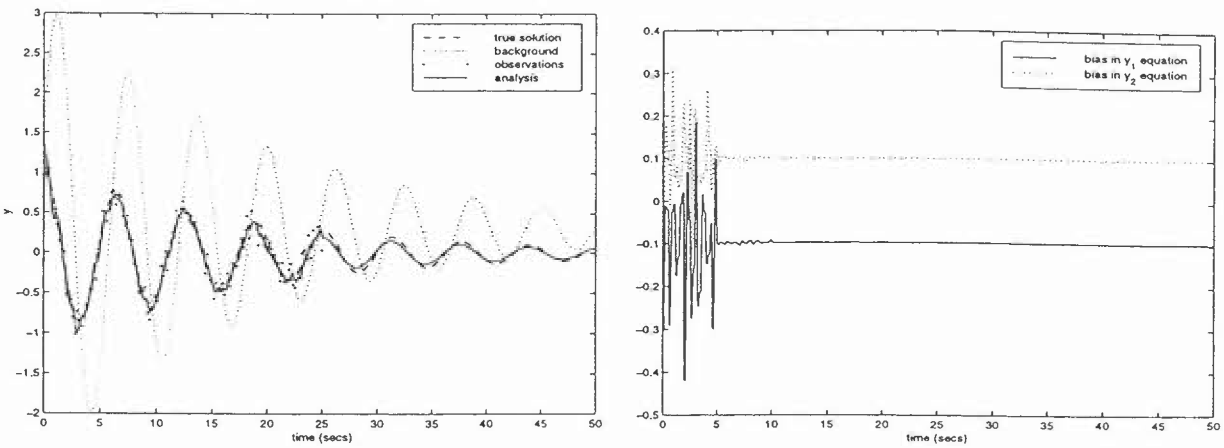


Figure 20: KF assimilation on the damped oscillating system with a constant forcing added to the model without bias correction

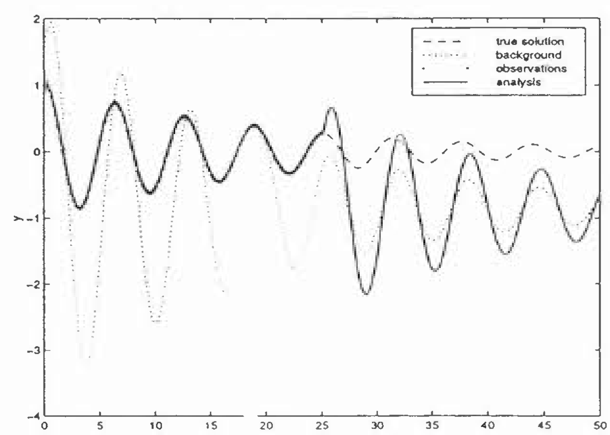


Figure 21: KF assimilation on the damped oscillating system with a constant forcing added to the model with bias correction: (i) the state (ii) the analysed bias

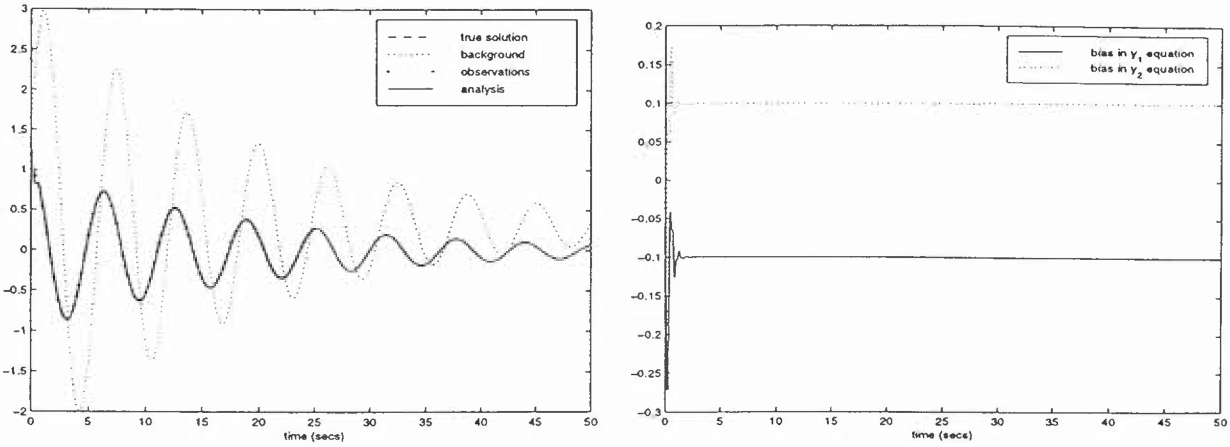


Figure 22: KF assimilation on the damped oscillating system with bias correction and noise on the observations: (i) the state (ii) the analysed bias

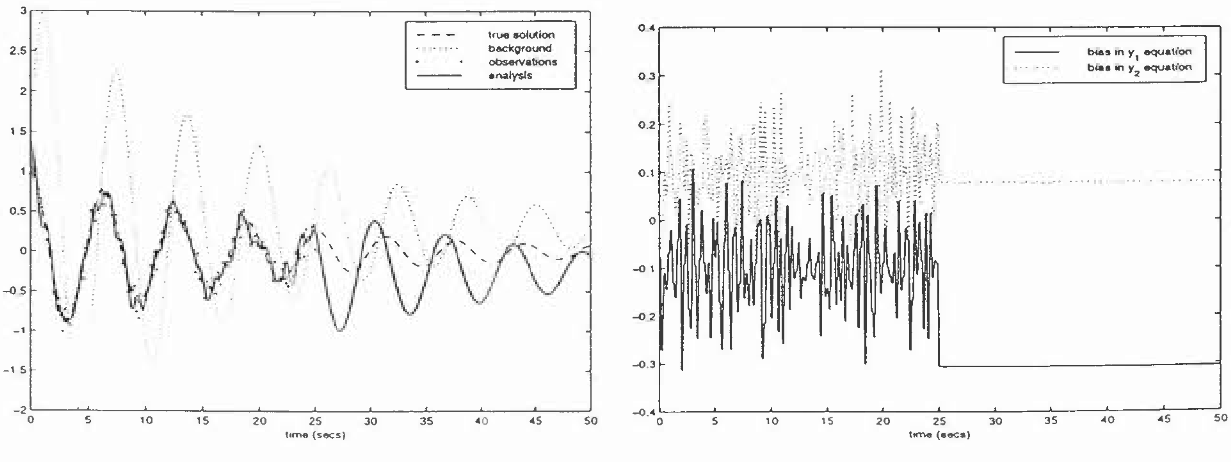


Figure 23: KF assimilation on the damped oscillating system with bias correction and noise on the observations with averaging at the end of the time period: (i) the state (ii) the analysed bias

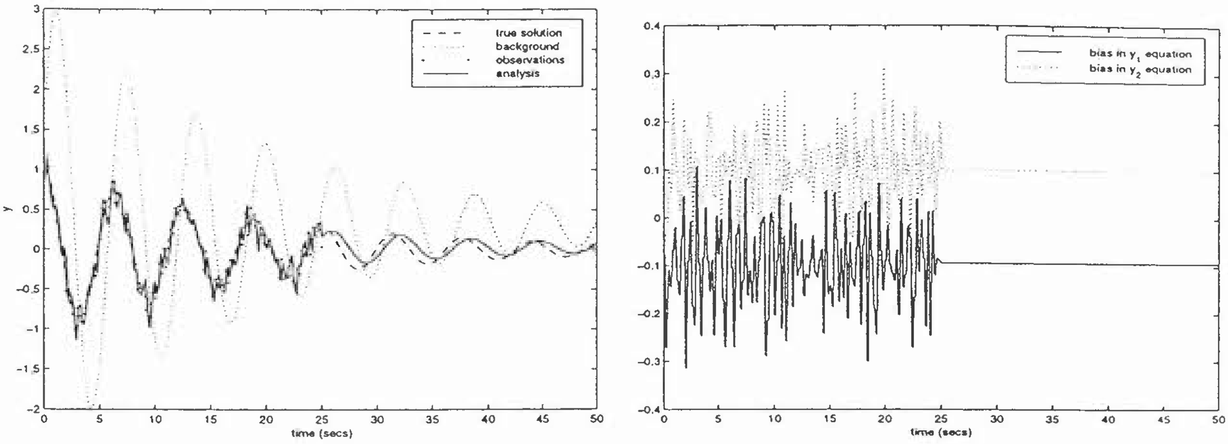


Figure 24: KF assimilation on the damped oscillating system with bias correction and noise on the observations with averaging over a moving window of observations: (i) the state (ii) the analysed bias

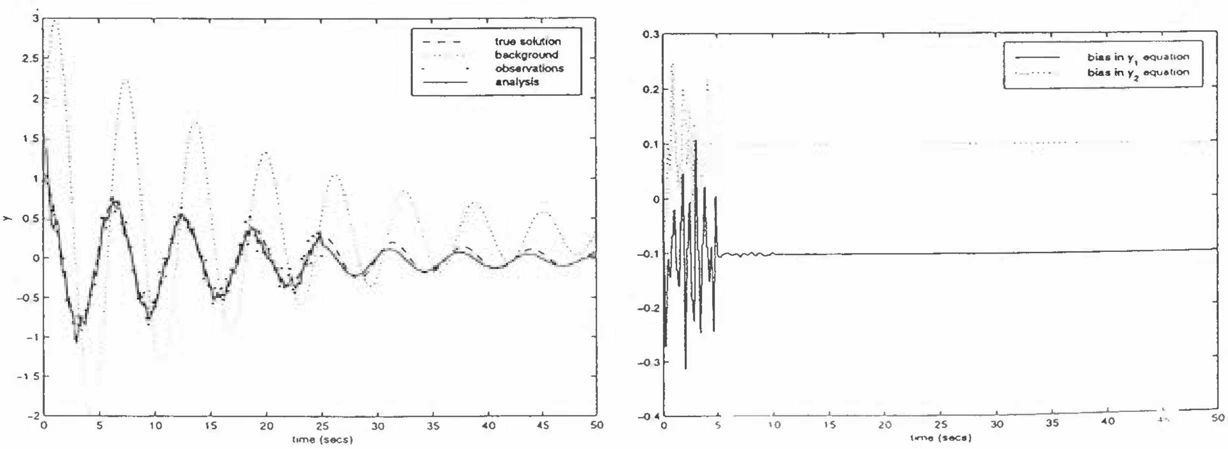


Figure 25: SCM assimilation on the damped oscillating system with incorrect coefficients with constant bias correction: (i) the state (ii) the analysed bias

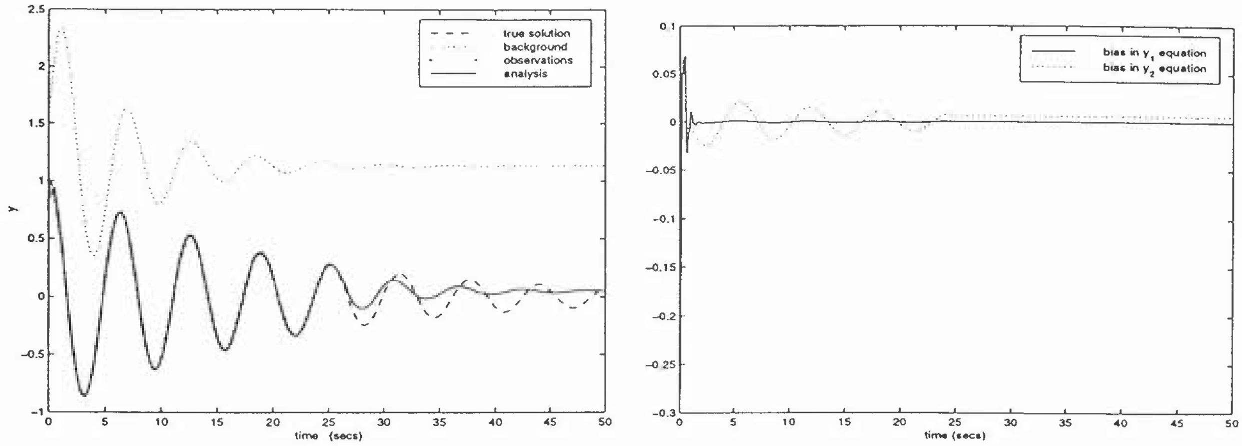
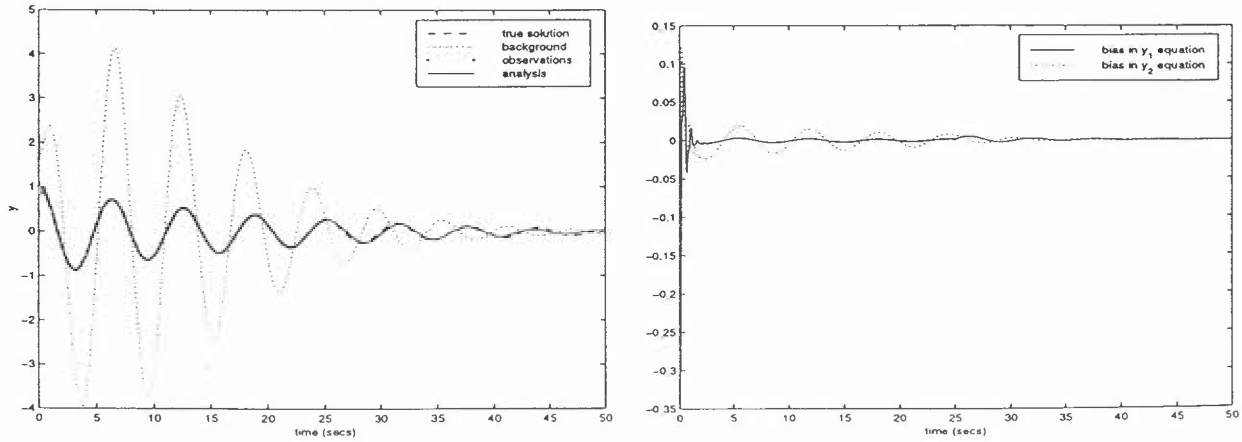


Figure 26: SCM assimilation on the damped oscillating system with incorrect coefficients with oscillating bias correction: (i) the state (ii) the analysed bias



4.2.2 The Lorenz equations

The method of bias correction has also been applied to the Lorenz system. Here a constant bias has been added to each equation, giving

$$
\begin{aligned}
x_{i+1} &= f(x_i, y_i, z_i) + 0.05 && (45)\\
y_{i+1} &= g(x_i, y_i, z_i) + 0.1 && (46)\\
z_{i+1} &= h(x_i, y_i, z_i) + 0.15 && (47)
\end{aligned}
$$

The results of the normal SCM when applied to this biased model are shown in Figure 27. Here, as in the oscillating system, the forecast deviates from the true solution after only a short time after the assimilation is stopped. The results of performing the bias correction using the SCM are shown in Figure 28. This shows a very good agreement between the analysis and the true solution with a good representation of the biases. Figure 29 shows the same experiment but with noise added to the observations. Here, as in the previous section, the bias estimate is greatly affected by the noise and so the forecast is not very good. Using the average of the estimated bias for the forecast produces slightly more accurate results as shown in Figure 30 where the forecast agrees with the true solution for about two seconds longer than without the averaging. When averaging the bias over a moving time window of 50 time steps, the analysed bias is very accurate after initial oscillations as shown on Figure 31. The forecast is also good and agrees with the true solution until about 11 seconds.

The EKF analysis produced similar results to the SCM when there was a bias present in the Lorenz equations as shown in Figure 32. Figure 33 shows that the bias correction term works well with perfect observations in the EKF. When there was noise on the observations, the EKF produced a reasonable analysis but the forecast deviated from the true solution after only a short period of time. This result is shown in Figure 34 where the analysed bias can be seen to be nearer the true bias than when using the SCM. Averaging at the end of the time period produced an accurate analysed bias although the analysed solution only agreed with the true solution up to about 11 seconds as shown in Figure 35. When the bias was averaged over a moving window during the assimilation, the analysed bias was very accurate as shown in Figure 36. The forecast was also slightly improved as it agreed with the true solution for over 11 seconds.

Figure 27: SCM assimilation on the Lorenz equations with a constant forcing added to the model without bias correction

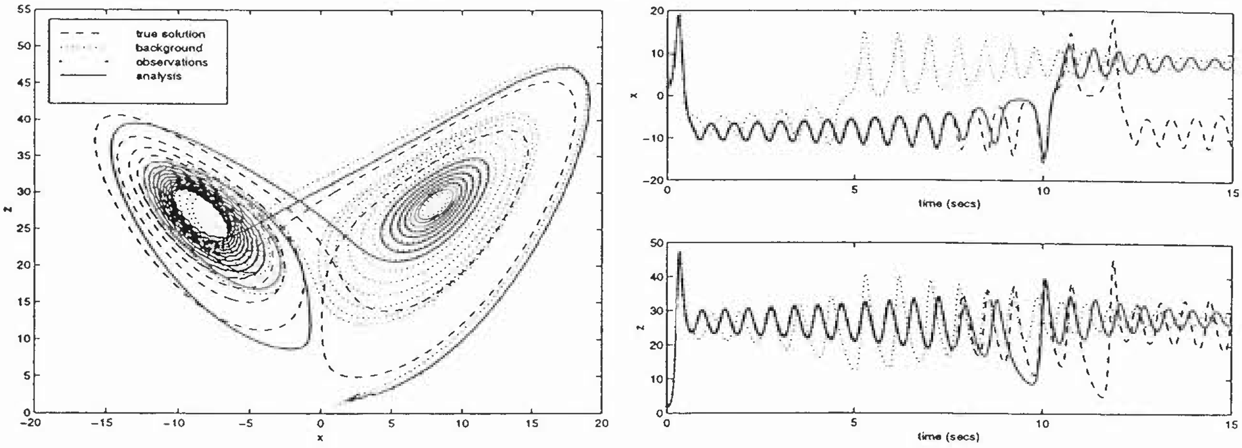


Figure 28: SCM assimilation on the Lorenz equations with a constant forcing added to the model with bias correction: (i) the state (ii) x-component (iii) the analysed bias

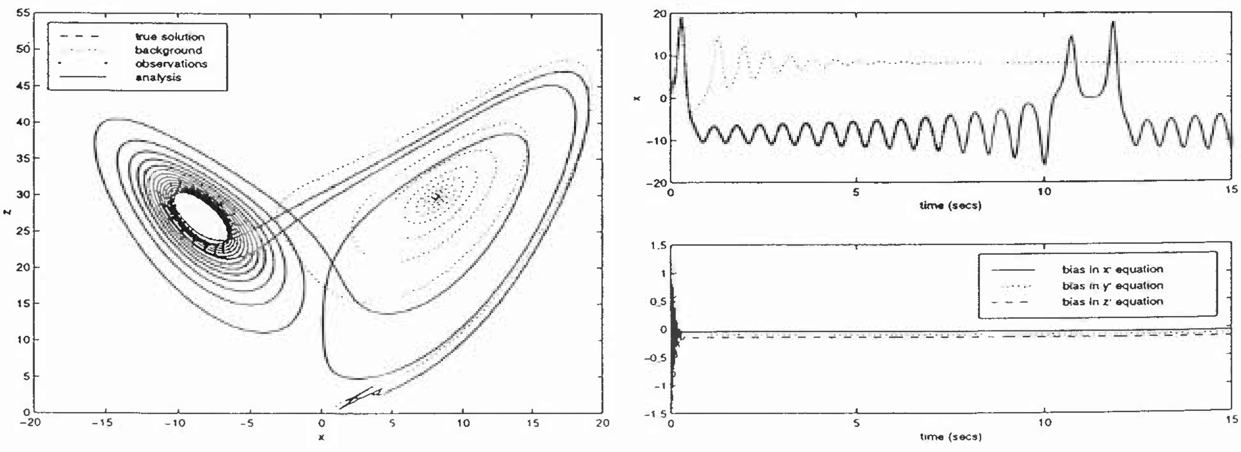


Figure 29: SCM assimilation on the Lorenz equations with noise on the observations using bias correction: (i) the state (ii) x-component (iii) the analysed bias

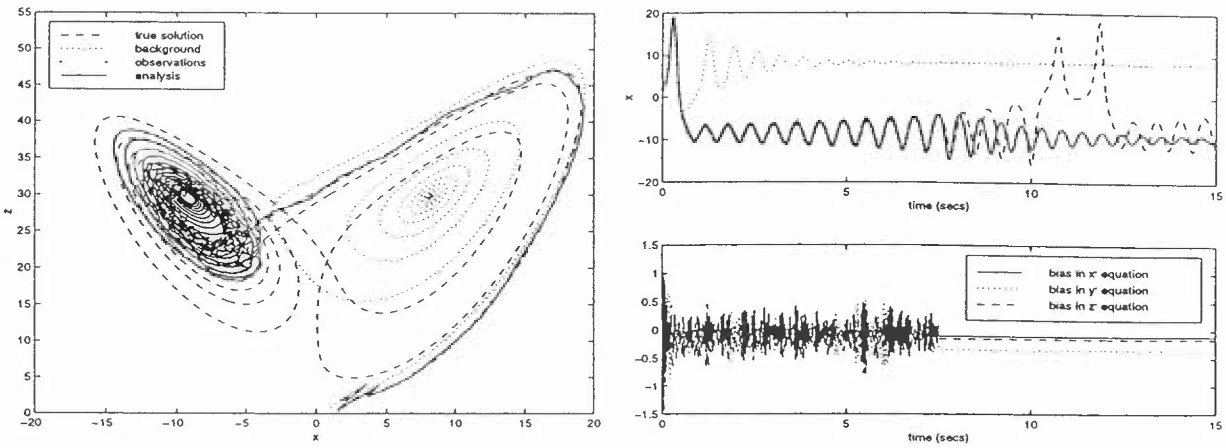


Figure 30: SCM assimilation on the Lorenz equations with noise on the observations using bias correction with averaging at the end of the assimilation period: (i) the state (ii) x-component (iii) the analysed bias

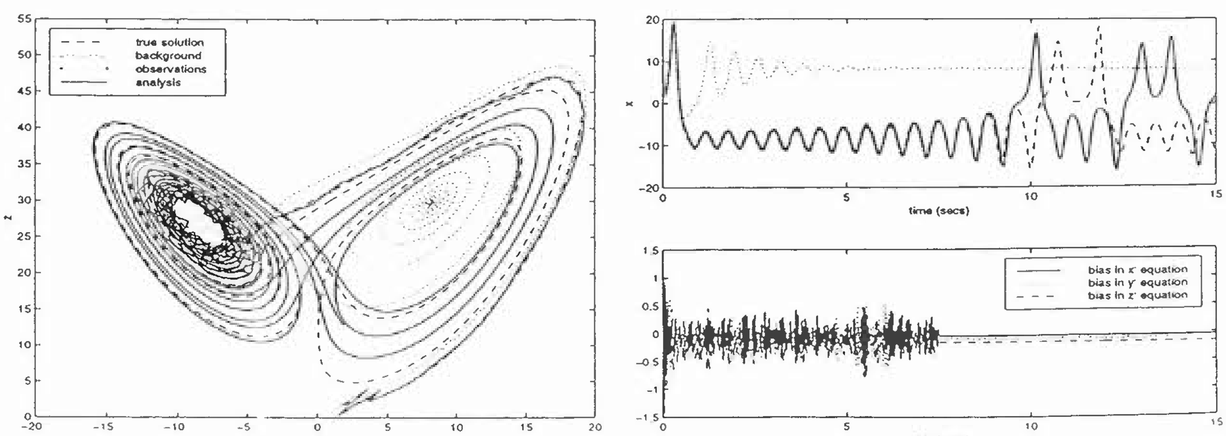


Figure 31: SCM assimilation on the Lorenz equations with noise on the observations using bias correction with averaging over a moving window of observations: (i) the state (ii) x-component (iii) the analysed bias

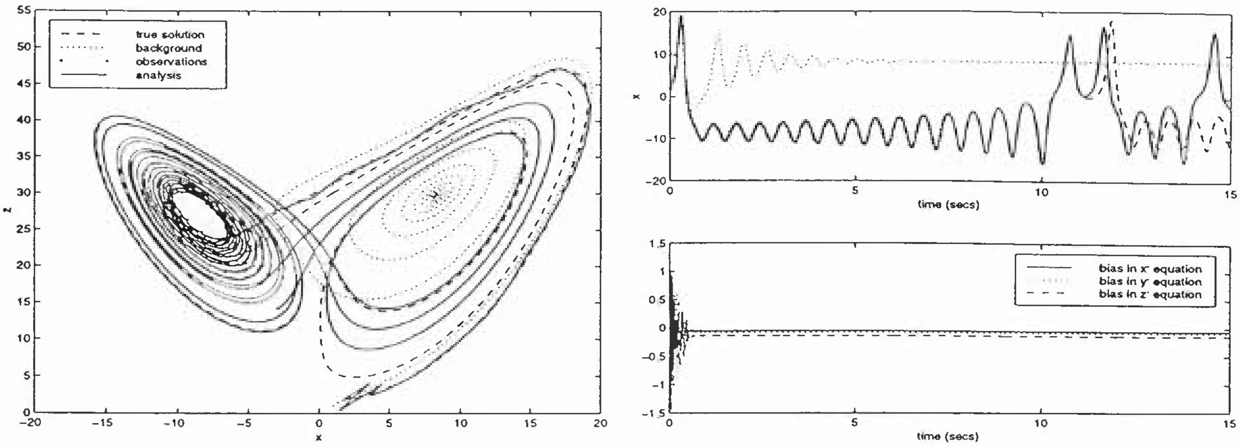


Figure 32: EKF assimilation on the Lorenz equations with a constant forcing added to the model without bias correction

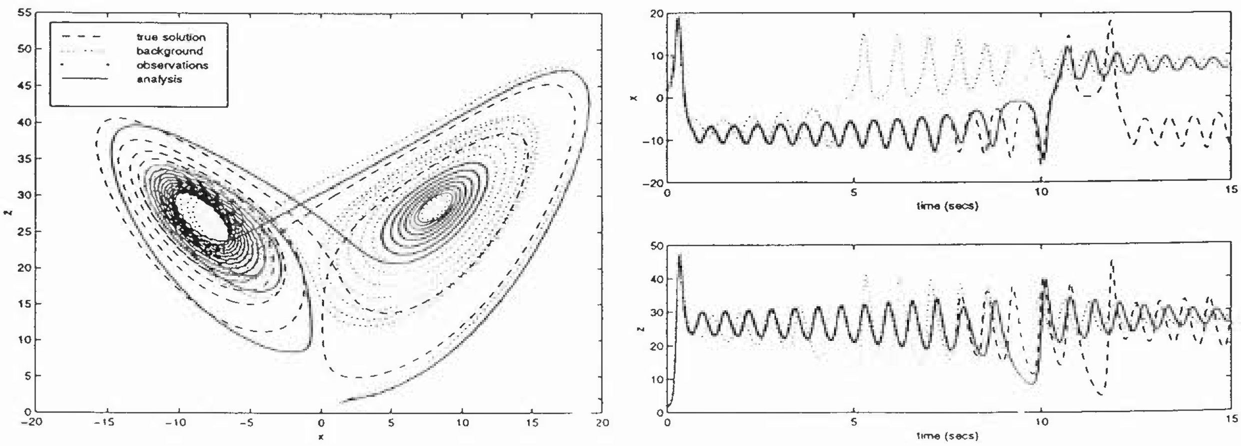


Figure 33: EKF assimilation on the Lorenz equations with perfect observations using bias correction: (i) the state (ii) x-component (iii) the analysed bias

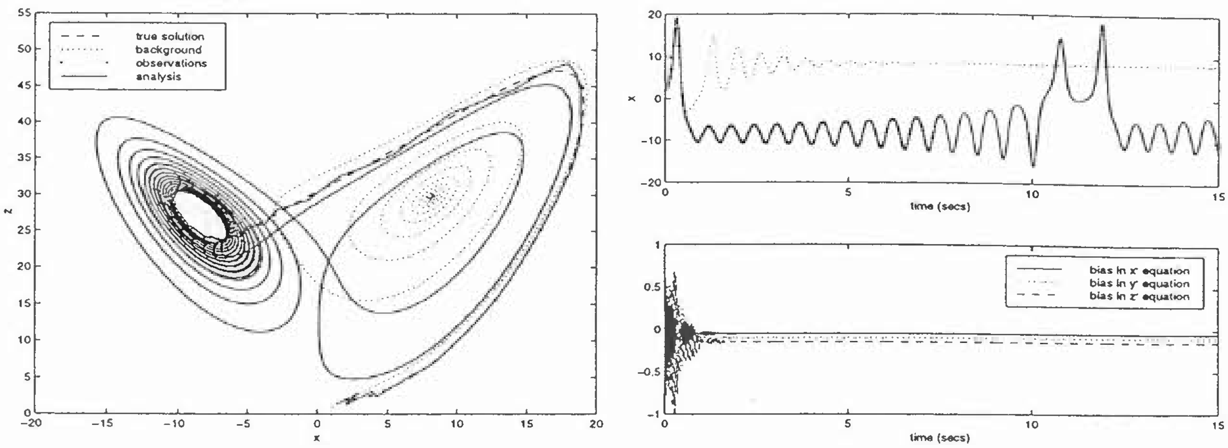


Figure 34: EKF assimilation on the Lorenz equations with noise on the observations using bias correction: (i) the state (ii) x-component (iii) the analysed bias

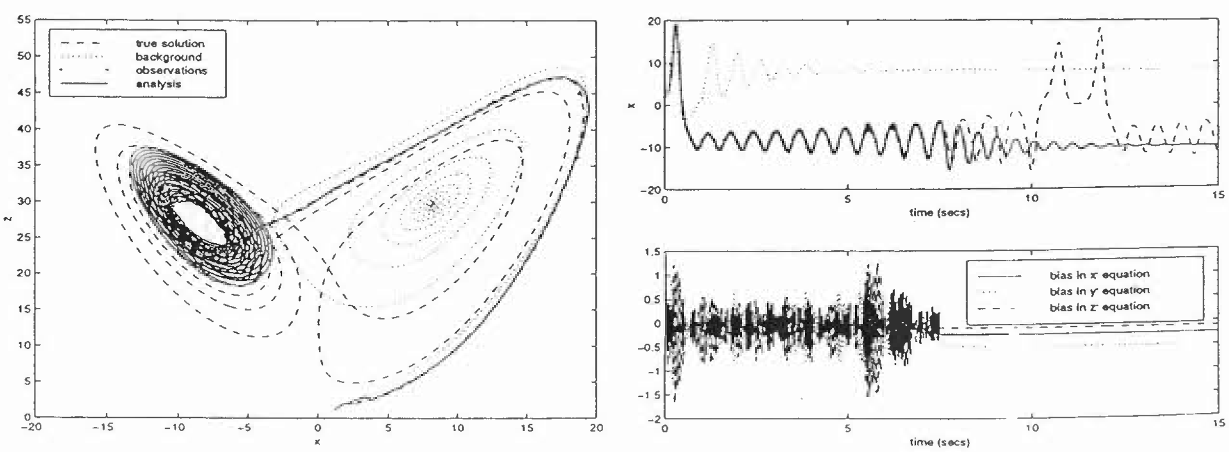


Figure 35: EKF assimilation on the Lorenz equations with noise on the observations using bias correction with averaging at the end of the time period: (i) the state (ii) x-component (iii) the analysed bias

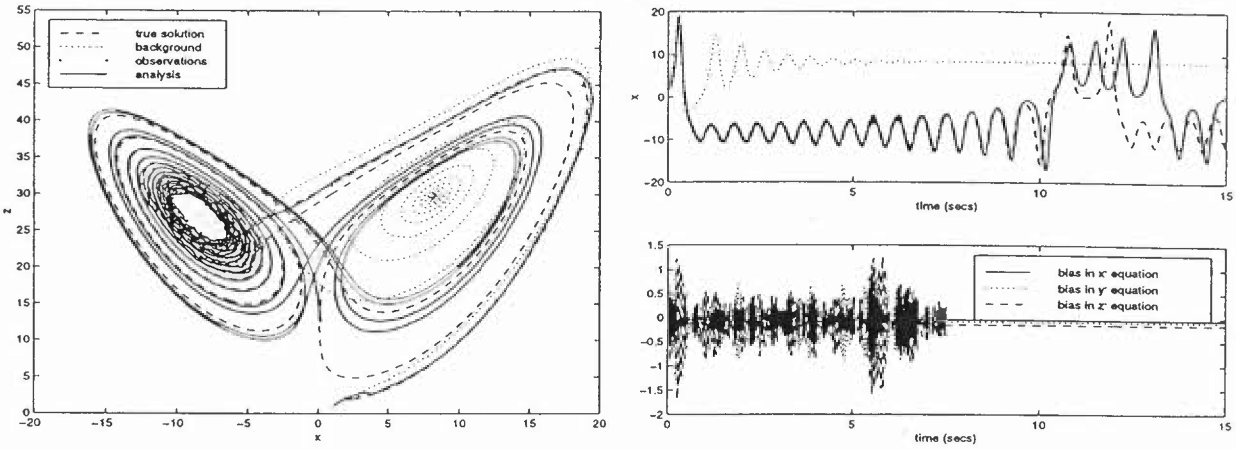
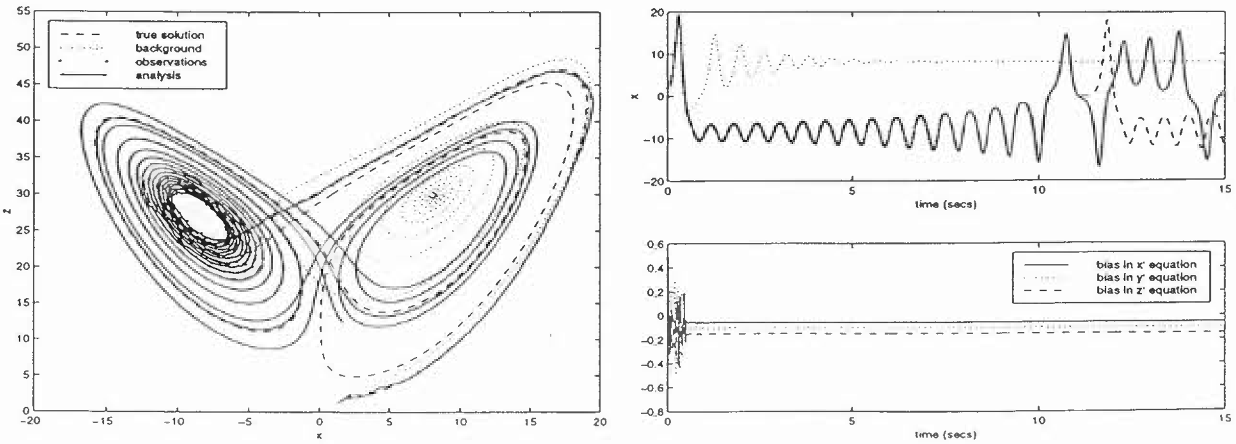


Figure 36: EKF assimilation on the Lorenz equations with noise on the observations using bias correction with averaging over a moving window of observations: (i) the state (ii) x-component (iii) the analysed bias



4.2.3 Summary of Results

The addition of the bias correction term in the data assimilation appears to account well for a constant bias with perfect observations in both the oscillating system and in the Lorenz equations. When noise is introduced, however, the fluctuations in the estimated bias are large compared to the bias itself when using the SCM. The Kalman filter bias estimate contains less noise but is still far from smooth. By averaging the bias prior to a forecast, the effects of the noise were reduced at the end of the assimilation and there were slight improvements in the subsequent forecast in both the SCM and the Kalman filter. Smoothing the analysed bias over a moving time window when there was noise in the observations proved to be effective in the oscillating system and the Lorenz equations and the resulting forecasts were improved.

When the coefficients in the oscillating system were changed, the constant bias correction did not work as well as was hoped. Altering the model to take this different type of bias into account shows improved results. One way in which it might be possible to make an educated guess at the type of model needed for the bias would be to use the scheme with a constant bias and observe how the bias evolves in time. The assimilation could then be done again using this model for the bias.

The experiments done with the SCM and the Kalman filter have not included the correct information about the statistics of the system. In practice, the statistics of the system will not be known very accurately so these experiments give an insight into how the bias correction method will work when applied to real problems. With the correct information incorporated into the error covariance matrices, the bias correction method produces smoother and more accurate analyses.

5 Conclusions

In Section 3.3, the data assimilation methods described in Section 2 were tested on an oscillating system and on the Lorenz equations. The Kalman filter appeared to be the method which produced the best analysis on the linear oscillating system when there was noise on the observations. This was expected given the theory in Section 2. When tested on the nonlinear Lorenz equations, all of the methods produced good analyses when assimilating for 7.5 seconds. The EKF gave the worst forecast although this could be due to the chaotic nature of the Lorenz equations rather than a bad analysis. Assimilating over a longer time interval with fewer noisy observations showed that the EKF actually stayed closer to the true solution than the other methods. The amount of computer power needed to compute the Kalman filter analysis is much larger than for the other methods due to the fact that the error covariance matrices have to be evolved at each time step. This becomes a problem when dealing with systems such as the atmosphere or ocean as the matrices are too large for even modern supercomputers to handle. Simplifications of the Kalman filter do exist however where only a small part of the error covariance matrices are evolved [4], [12].

The method of accounting for bias described in Section 4.1 was tested using the SCM and the Kalman filter. The models were modified by adding a constant bias. Both methods gave promising results when there were perfect observations. However, adding noise to the observations made the analysed bias fluctuate a great deal about the true bias which gave a poor analysis of the state. Using the Kalman filter to account for this gave better results, but the bias was not as smooth as expected. The method of averaging the analysed bias over a moving time window appeared to smooth the bias to the correct value, which gave improved forecasts. When the coefficients were altered in the model for the oscillating system, assuming a constant bias gave a good analysis of the state - it is the forecast which is the problem. Using an oscillating model for the bias gave a much better forecast. In practice, the way the bias evolves would be difficult to ascertain. One possible way round this would be to observe how the bias evolves using a constant bias correction and modify the model for the bias using this knowledge.

References

[1] Bergthorsson, P., Doos, B.R. Numerical weather map analysis. *Tellus*, **7**, 1995, pp 329-340.

[2] Bratseth, A.M. Statistical interpolation by means of successive corrections. *Tellus*, **38A**, 1986, pp 439-447.

[3] Burger, G., Cane, M.A. Interactive kalman filtering. *J. of Geoph. Res.*, **99**, 1994, pp 8015-8031.

[4] Cohn, S.E., Todling, R. Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteor. Soc. Jpn.*, **74**, 1996, pp 63-75.

[5] Cohn, S.E., Sivakumaran, N.S., Todling, R. A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.*, **122**, 1994, pp 2838-2867.

[6] Courtier, P. Dual formulation of four-dimensional variational assimilation. *Quart. J. R. Met. Soc.*, **123**, 1997, pp 2449-2461.

[7] Cressman, G. An optimal objective analysis system. *Mon. Weather Rev.*, **87**, 1959, pp 367-374.

[8] Daley, R., 1991. Atmospheric data analysis. Cambridge University Press.

[9] Dee, D.P. On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1995, pp 1128-1145.

[10] Dee, D.P. Data assimilation in the presence of forecast bias. *Quart. J. R. Met. Soc.*, **124**, 1998, pp 269-295.

[11] Derber, J.C. A variational continuous data assimilation technique. *Mon. Wea. Rev.*, **117**, 1989, pp 2437-2446.

[12] Evensen, G. Advanced data assimilation for strongly nonlinear dynamics. *Mon. Wea. Rev.*, **125**, 1997, pp 1342-1354.

[13] Friedland, B. Treatment of bias in recursive filtering. *IEEE Trans. Auto. Contr.*, **AC-14**, 1969, pp 359-367.

[14] Gauthier, P. Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus*, **44A**, 1992, pp 2-17.

[15] Gelb, A. (ed.), 1974. Applied optimal estimation. MIT Press.

[16] Ghil, M. Meteorological data assimilation for oceanographers. Part I: description and theoretical framework. *Dyn. Atmos. Oceans*, **13**, 1989, pp 171-218.

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

[18] Griffith, A.K., Nichols, N.K. Accounting for model error in data assimilation using adjoint methods. *Compututational Differentiation: Techniques, Applications and Tools*, (Eds. M. Berz, C. Bischof, G.Corliss and A. Greiwank), SIAM, Philadelphia, 1996, pp 195-204.

[19] Griffith, A.K., Nichols, N.K. Adjoint methods for treating model error in data assimilation. *Numerical Methods for Fluid Dynamics*, (Ed. M.J. Baines), ICFD, Oxford University Computing Laboratory, 1998, pp 335-344.

[20] Ignagni, M.B. Separate-bias Kalman estimator with bias state noise. *IEEE Trans. Auto. Contr.*, **AC-35**, 1990, pp 338-341.

[21] Kalman, R.E. A new approach to linear filtering and prediction problems. *Trans. of the ASME ser. D*, **82**, 1960, pp 35-44.

[22] Lorenc, A.C. Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.*, **112**, 1986, pp 1177-1194.

[23] Lorenc, A.C. Optimal nonlinear analysis. *Quart. J. R. Met. Soc.*, **114**, 1988, pp 205-240.

[24] Lorenc,A.C., Bell, R.S., MacPherson, B. The met. office analysis correction data assimilation scheme. *Quart. J. R. Met. Soc.*, **117**, 1991, pp 59-89.

[25] Lorenc, A.C. Iterative analysis using covariance functions and filters. *Quart. J. R. Met. Soc.*, **118**, 1992, pp 569-591.

[26] Miller, R.N., Ghil, M., Gauthiez, F. Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sc.*, **51**, 1994, pp 1037-1056.

[27] Rutherford, I.D. Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sc.*, **29**, 1972, pp 809-815.

[28] Sasaki, Y. Some basic formulisms on numerical variational analysis. *Mon. Wea. Rev.*, **98**, 1970, pp 875-883.

[29] Sparrow, C., 1982. The Lorenz equations: bifurcations, chaos, and strange attractors. Springer-Verlag.

[30] Tacker, E.C., Lee, C.C. Linear filtering in the presence of time-varying bias. *IEEE Trans. Auto. Contr.*, **AC-17**, 1972, pp 828-829.

[31] Talagrand, O. A study on the dynamics of four-dimensional data assimilation. *Tellus*, **33**, 1981, pp 43-60.

[32] Zhou, D.H., Sun, Y.X., Xi, Y.G., Zhang, Z.J. Extension of Friedland's separate-bias estimation to randomly time-varying bias of nonlinear systems. *IEEE Trans. Auto. Contr.*, **AC-38**, 1993, pp 1270-1273.

A Appendix

This appendix contains figures corresponding to the experiments described in Section 3 which are not shown in the main text.

List of Figures

Figure 1: Experiment 2a. Assimilation on the Lorenz equations with perfect observations using the AC scheme

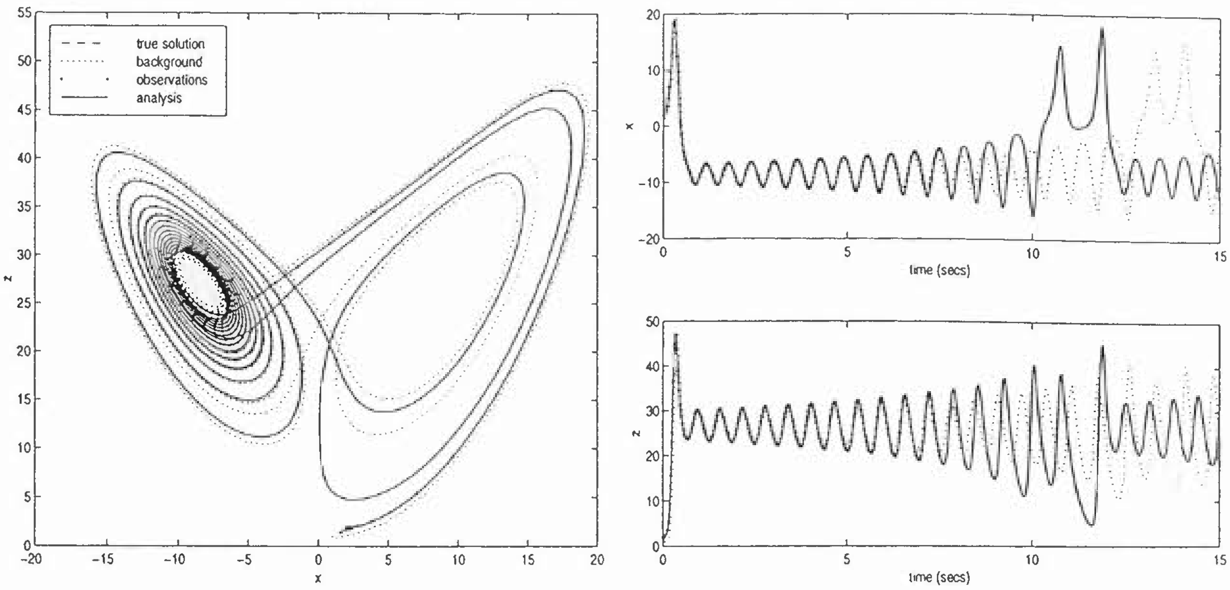


Figure 2: Experiment 2a. Assimilation on the Lorenz equations with perfect observations using the OI scheme

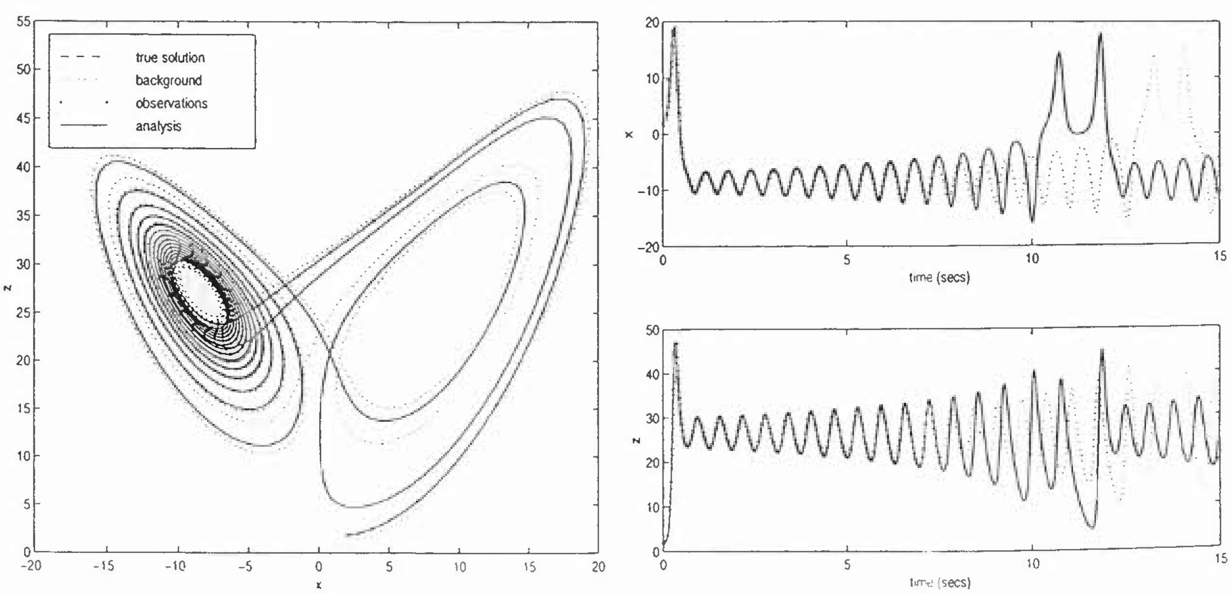


Figure 3: Experiment 2a. Assimilation on the Lorenz equations with perfect observations using the EKF

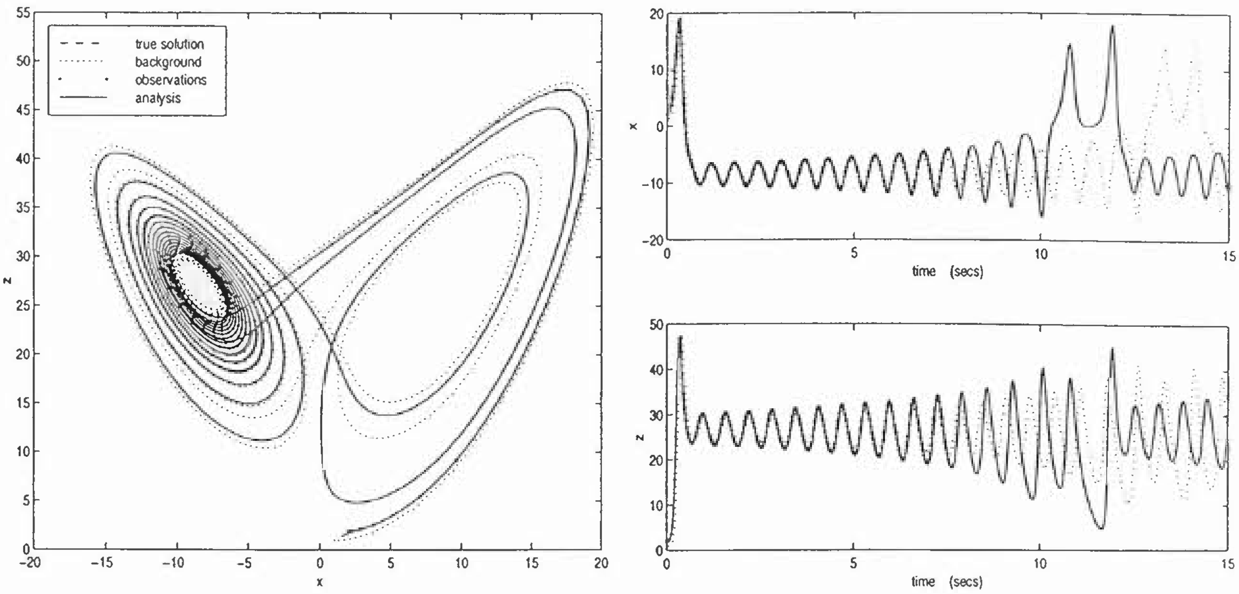


Figure 4: Experiment 2b. Assimilation on the Lorenz equations with noise on the observations using the SCM

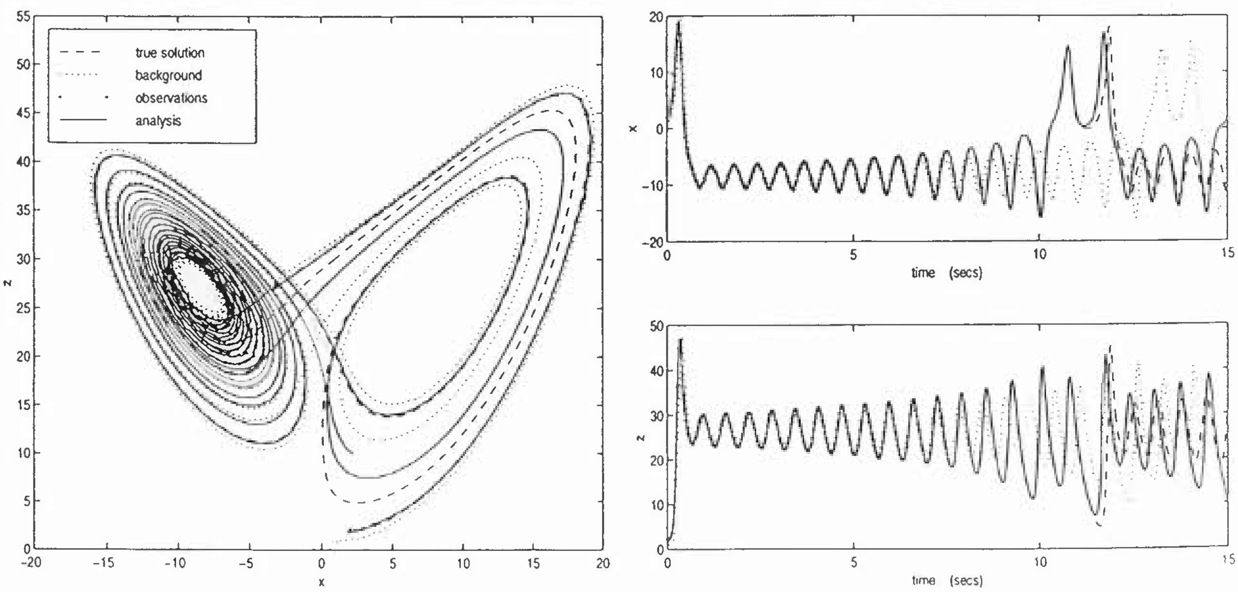


Figure 5: Experiment 2b. Assimilation on the Lorenz equations with noise on the observations using the AC scheme

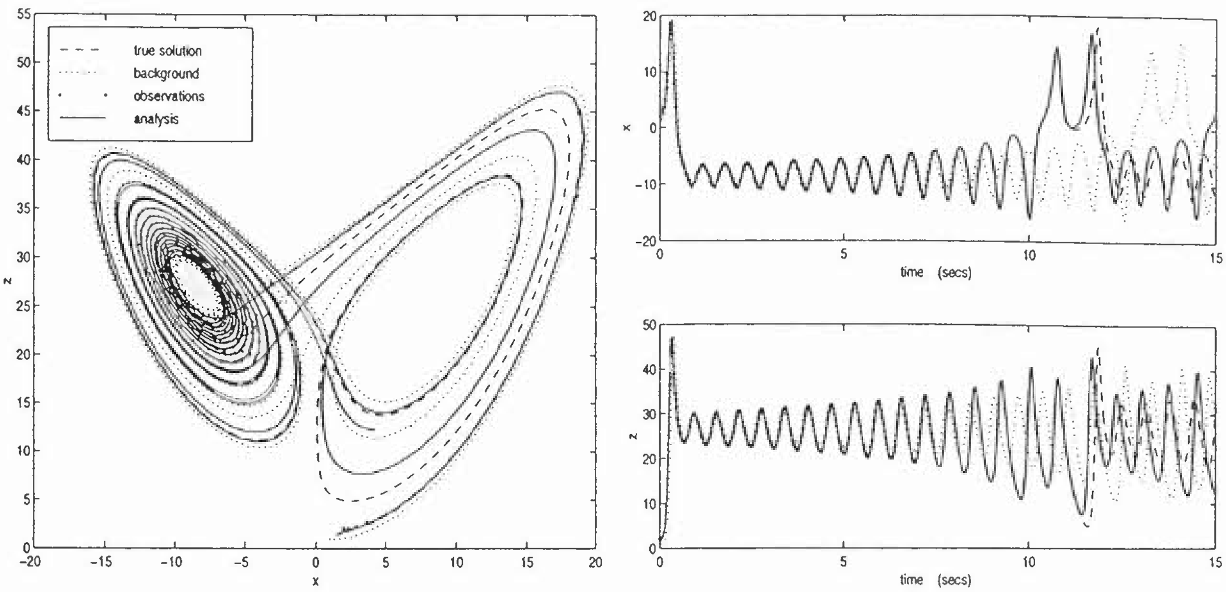


Figure 6: Experiment 3a. Assimilation on the Lorenz equations with perfect observations using OI

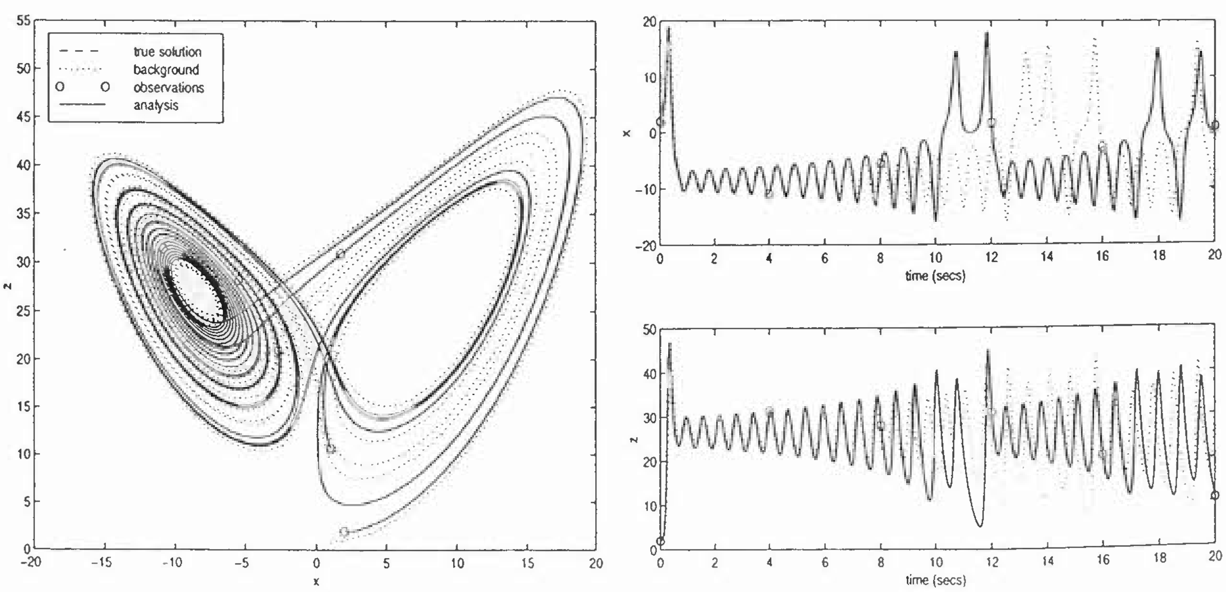


Figure 7: Experiment 3a. Assimilation on the Lorenz equations with perfect observations using the EKF

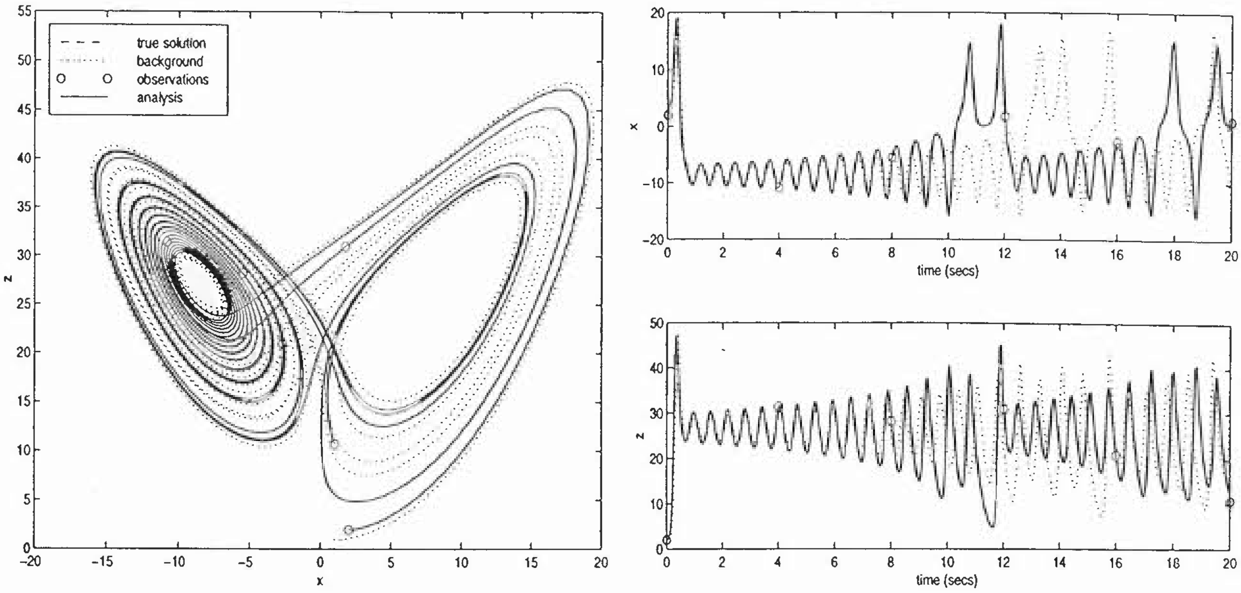


Figure 8: Experiment 3b. Assimilation on the Lorenz equations with noise on the observations using the SCM

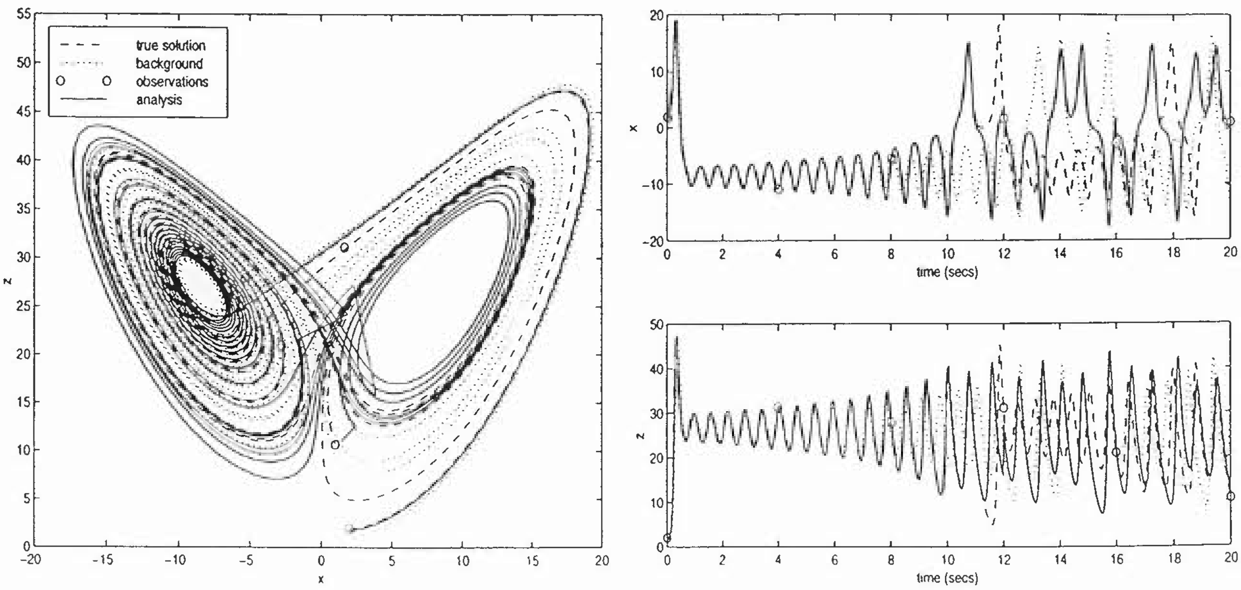


Figure 9: Experiment 3b. Assimilation on the Lorenz equations with noise on the observations using the AC scheme