

PAPER • OPEN ACCESS

On the possibility of correction of the forecasting of the Lorenz attractor dynamic characteristics using experimental data and data assimilation

To cite this article: Timoshenkova Yulia *et al* 2018 *J. Phys.: Conf. Ser.* **1053** 012004

View the [article online](#) for updates and enhancements.

Related content

- [A topological invariant of the Lorenz attractor](#)
N E Klinshpont
- [Data assimilation: Particle filter and artificial neural networks](#)
Helaine Cristina Morais Furtado, Haroldo Fraga de Campos Velho and Elbert Einstein Nehrer Macau
- [Neural networks for emulation variational method for data assimilation in nonlinear dynamics](#)
Helaine Cristina Morais Furtado, Haroldo Fraga de Campos Velho and Elbert E N Macau



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

On the possibility of correction of the forecasting of the Lorenz attractor dynamic characteristics using experimental data and data assimilation

Timoshenkova Yulia^{1,2,3}, Porshnev Sergey¹ and Safiullin Nikolai¹

¹Ural Federal University, Mira. 19, 620002 Ekaterinburg, Russia

²Scientific and Production Association of automatics named after academician N.A.Semikhatov, Mamina-Sibiriyaka 145, 620075 Yekaterinburg, Russia

³Email: juliatimoshenkova@gmail.com

Abstract. In this paper the application of the Data Assimilation method based on Ensemble Kalman Filter to forecasting the Lorenz attractor dynamic characteristics is described. In the article the EnKF algorithm is described as applied to forecasting of the Lorenz attractor coordinates. The assessment of influence of the filter parameters on the quality of correction of the forecast is carried out. The importance of using before obtained forecast results in the filter window is noted. Primary principles of correction of forecast of dynamic characteristics in nonlinear systems based on EnKF are formulated. The obtained results allow the conclusion about the necessity of applying the Data Assimilation method to carrying out forecasts of various dynamic characteristics in nonlinear systems using the EnKF as a tool to be drawn.

1. Introduction

A task of forecasting of values of one-dimensional time series (TS), by which a series of index values w_k ($k = 1 \dots P$) of observed system sorted in time order is meant, is topical. The task arises in different areas of people activity: economics, technology, science.

The task of forecasting is to calculate member $m+1$ of TS basing on m preceding values. While solving the task by means of methods not using mathematical system models (formal forecasting methods) that generated the forecasted TS, it is assumed that:

$$w_t = F(w_{t-1}, w_{t-2}, \dots, w_{t-p}, a_1, \dots, a_m), \quad (1)$$

where F represents a function defined by selected forecast method, w_t is a vector of forecasted values, $w_{t-1}, w_{t-2}, \dots, w_{t-p}$ are preceding values of the TS, a_1, \dots, a_m represent model parameters. For example, when using the ARMA method (Autoregressive moving-average model) the vector of forecasting values is presented as follows:

$$w_t = c + \varepsilon_t + \sum_{i=1}^p \alpha_i w_{t-i} + \sum_{i=1}^q \beta_i \varepsilon_{t-i}, \quad (2)$$

where c – is a constant value, ε represents white noise, α_i, β_i – are real numbers – the autoregressive and the moving average coefficients, p, q – are integers defining the model order. [1]



However, the experience of application of formal forecasting methods shows that the achievable accuracy of the forecast is not high enough. It is indirectly affirmed by the large amount of methods. [2]

An alternative TS forecasting method is Data Assimilation method. It uses mathematical model of a system that generated the forecasting TS. When using the Data Assimilation method, the TS forecast is carried out according to the following system of equations: [1]

$$\begin{aligned} w_f(t+1) &= M(p, w_t(t)), \\ y_o &= H(w_f(t)), \end{aligned} \quad (3)$$

where M – is a nonlinear operator of the state model, w_f represents a vector of forecasted values, p is a vector of the model parameters, H – represents nonlinear operator of the TS observation.

In the paper the DA method is applied to forecasting of the Lorenz attractor coordinates. The Lorenz attractor is a nonlinear dynamic system that reveals chaotic behavior.

2. Formulation of the Data Assimilation method for the Lorenz Attractor

The Lorenz attractor appears in the task of describing air convection dynamics, when the air is located over a heating surface. The Lorenz attractor mathematical model is described using a system of three nonlinear ordinary differential equations that represent the convection final amplitude:

$$\begin{aligned} \frac{dx}{dt} &= \sigma \cdot (y - x) \\ \frac{dy}{dt} &= x \cdot (r - z) - y \\ \frac{dz}{dt} &= x \cdot y - b \cdot z, \end{aligned} \quad (4)$$

where $\sigma = \nu/k$ is the Prandtl number, $r = Ra/Ra_c$ is the Rayleigh number (normalized), $b = 4/(1+a^2)$ is geometric factor, x is the convection intensity, y is the difference in temperature between the ascending and the descending flows; z is the vertical thermal profile deviation from the linear one. [2]

An important feature of the system is that the system's behavior becomes random, when the parameters σ , b and r reach values of $\sigma=10$, $b = 8/3$ and $r \geq 24.06$. In figure 1 an example of the Lorenz attractor in its chaotic mode is presented. In phase space the attractor's topology is represented by a tangle of trajectories, in which two areas can be distinguished. The solution is located inside one of the areas at each point in time. And a shift of the state of the system towards one area or another is absolutely unpredictable. [3]

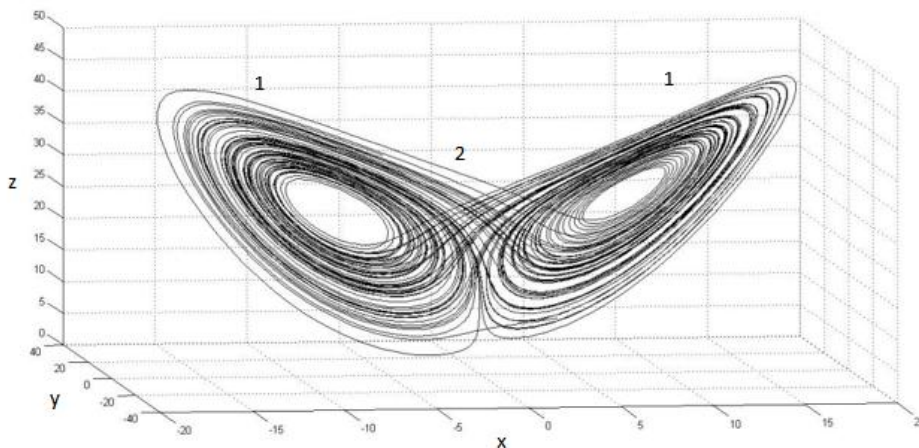


Figure 1. The Lorenz attractor.

3. Ensemble Kalman Filter

Now we shall concretize the equations of the Data Assimilation method as applied to the Lorenz attractor. The Lorenz attractor state at the current point in time is characterized by the following set of coordinates:

$$w = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (5)$$

When applying the Kalman Filter algorithm the estimation of the observed vector (forecasted parameter) X state is carried out at every time step. The filter operation is based on consecutive execution of three steps: analysis, forecasting, correction of the forecasting

The Discrete form of the Lorenz 3 model can be written as:

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} 1 - \sigma\tau & (r - z(t))\tau & y(t)\tau \\ \sigma\tau & 1 - \tau & x(t) \\ 0 & -x(t) & 1 - b\tau \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \quad (6)$$

where τ represents length of a time step used in the model integration.

EnKF allows acquiring more precise descriptions of reality, which can be used in forecasting, due to making use of multiple representations of its parameters. The representations get more accurate as new information incomes. All of the realizations allow us to acquire statistics that describe relationships between the parameters of the models, their responses and the models accordance to the real data.

Considering that the Lorenz system components depend on one another, the forecasting is carried out separately for the x , y , z components; since z component depends on both x and y it is forecasted the last using already forecasted values of x and y .

As an example, let us take 100 realizations of a three-dimensional model. Each realization is calculated using the KF forecast resulting in 100 forecast variations. Having multiple model realizations and multiple forecasts allows us to calculate the covariance between the attractor's parameters and the observed values. Using the covariance, the model parameters can be adjusted based on the difference between the model and the observation. The adjustments differ due to the individual distinctions between the models and the real values. Then the mean value is calculated.

The EnKF algorithm application to forecasting the Lorenz attractor characteristics is described further. Here H represents an identity diagonal 3×3 matrix. The expected values of the model's error and the error of forecast are taken equal to zero. The errors are input with accordance to the covariance matrices.

3.1. THE EnKF ALGORITHM

The algorithm input data:

y_{obs} – is the observation vector;

M – is a nonlinear operator of the state model;

H – is a linear operator of the TS observation;

R – is a covariance matrix of the observation errors distribution;

Q – is a covariance matrix of the modeling errors distribution.

For the Lorenz3 model operators of the observation and the forecast are taken in accordance with the following formulas:

$$M = \begin{bmatrix} 1-\sigma & (r-z(t))\tau & y(t)\tau \\ \sigma & 1-\tau & x(t) \\ 0 & -x(t) & 1-b\tau \end{bmatrix} \quad (7)$$

The resulting value of the covariance matrices is approximation of the error covariance matrix. During calculations the mean of a vector of the ensemble real values is taken. During the realization of the Lorenz attractor model, it is assumed that the system components are distributed normally. Consequently, diagonal matrix of the standard deviation values represents the covariance observation error matrix. For the Lorenz3 model covariance matrices of errors of observation and forecast are calculated in accordance with the following formulas:

$$\begin{aligned} R &= E \cdot \xi_f \\ Q &= E \cdot \xi_o \end{aligned} \quad (8)$$

where ξ_f, ξ_o – are the dispersions of errors of forecast and observation, E – is identity diagonal 3×3 matrix. [4, 5]

The EnKF algorithm is realized by carrying out the following sequence of actions:

1. Formation of N -sized ensemble of data using mathematical model at the starting time point. Every realization has its own set of properties. [4]

$$w_a^m = F(w), m = 1, 2, \dots, N \quad (9)$$

where F – is a function describing the system, μ – is a vector of random modeling errors that is normally distributed and has a covariance matrix.

2. Calculation of the forecast values for every model in the ensemble, getting values of the observed parameters at step n for the state vector obtained at the $n-1$ step. After that the fixed-point iteration method in the form of the following equations is used for the model:

$$w_f^m(t + \Delta t) = M(w(t)) + \gamma, m = 1, 2, \dots, N \quad (10)$$

where γ – is a vector of random modeling errors that is normally distributed and has a covariance matrix.

3. The calculation of the covariance error matrix to correct the forecast:

- Formation of a covariance error matrix for the forecast vector:

$$P_f = \frac{1}{N-1} \sum (w_f(\tau_i) - \overline{w_m})(w_f(\tau_i) - \overline{w_m})^T, \quad (11)$$

- Calculation of Kalman weight operator K [5]

$$K = P_f \cdot H \cdot (H \cdot P_f \cdot H^T + R)^{-1}, \quad (12)$$

- Update of the ensemble model values in accordance with the weight operator and the observation vector y_{obs} . The covariance forecast matrix values are updated in accordance with the corrected ensemble value:

$$w_m(\tau_i) = w_m(\tau_i) + K(y_{obs}(\tau_i) - H \cdot w_m(\tau_i)) \quad (13)$$

- Update of the forecast error covariance matrix in accordance with a more precise ensemble w_m . Calculation of the value of analyzed components w_a and of the analysis error covariance matrix. Here I – is an identity diagonal matrix. [5]

$$x_a(\tau_i) = \frac{1}{N-1} \sum w_m(\tau_i) \quad (14)$$

$$P_a = \frac{1}{N-1} \sum (w_m(\tau_i) - \bar{w}_a)(w_m(\tau_i) - \bar{w}_a)^T$$

4. Steps 2 and 3 are executed consecutively for each n moment in time, when new measurements income. The set of measurements is not necessarily stays the same at each step, measurements may income constantly.

4. Numerical experiments

The Lorenz system of DE's with $\sigma=10$, $r = 28$, $b = 8/3$ parameters was taken as a nonlinear model that becomes the Lorenz attractor and enters chaotic mode due to the selected values. Values of time parameter (t) given in seconds are the input of the system. For analyzing and forecasting initial time which is 10 seconds, because prior the attractor is in chaotic mode which makes forecasting harder and complicates statistical processing of results. As a result of modeling the $x(t)$, $y(t)$, $z(t)$ attractor components are obtained from the output. White noise is additively summarized with the obtained TS values resulting in the observation components.

TS with x , y , z parameters, adjusted based on the noise impact on the systems, is obtained as a result of the noise addition. The result of filter's work is three time-series containing x' , y' , z' parameters values adjusted based on forecasting and following correction.

4.1. EXPERIMENT – Evaluation of Ensemble Size in the EnKF

The research is aimed at determining dependencies of the result of correction of forecasting on various filter parameters and the attractor's (stable or unstable) state (result for 50 ensembles of filter in figure 2). Corrected forecast accuracy evaluation is carried out by calculating the means square error (MSE). The MSE computes mean values of squares of errors or deviations; it is a difference between the estimation and estimated value. MSE is a measure of quality of estimation, and it is always a non-negative value, the values closest to zero are considered the best. For Lorenz 3 model need use ensemble filters, because this method uses several sets of conditions.

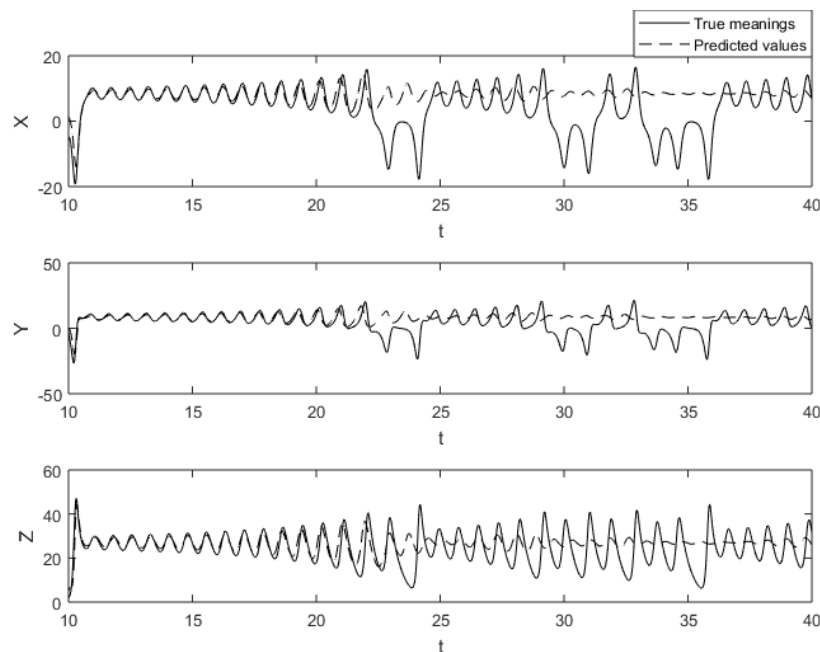


Figure 2. The results of an experiment using the filter size of 50 ensembles and dispersions of the errors equal to 10 and 5

Further the results of the filter's work on 150 ensembles with error dispersions of 10 and 5 correspondingly are reviewed. The obtained results are presented in figure 3. Applying EnKF using given parameters is impractical. The MSE value roughly equals to 15 for both x and y coordinates. The MSE values obtained as results of previous experiments are equal to approximately 0, 2.

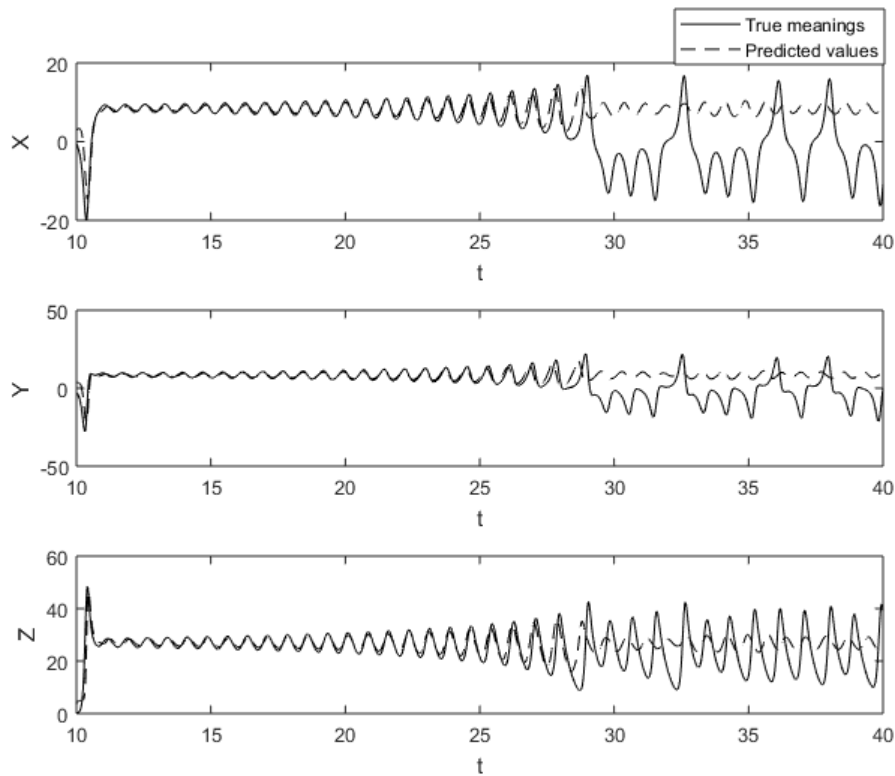


Figure 3. The results of an experiment using the filter size of 150 ensembles and dispersions of the errors equal to 10 and 5

The results of the experiments are compared, the dependency of the forecast accuracy on true values is determined. The estimation is carried out to determine the dependency of the forecast result accuracy on the EnKF characteristics.

During the research the estimation of the Kalman Filter parameter's impact has been carried out, namely, the dependency of the accuracy of the model characteristics forecast result on the number of points in ensemble is determined. The dependency of the MSE value on the filter ensemble size for three components is presented on the plot in the picture.

The dependency is linear (figure 4) which proves the assumption that increase in number of the filter ensemble points leads to decrease in the forecast error.

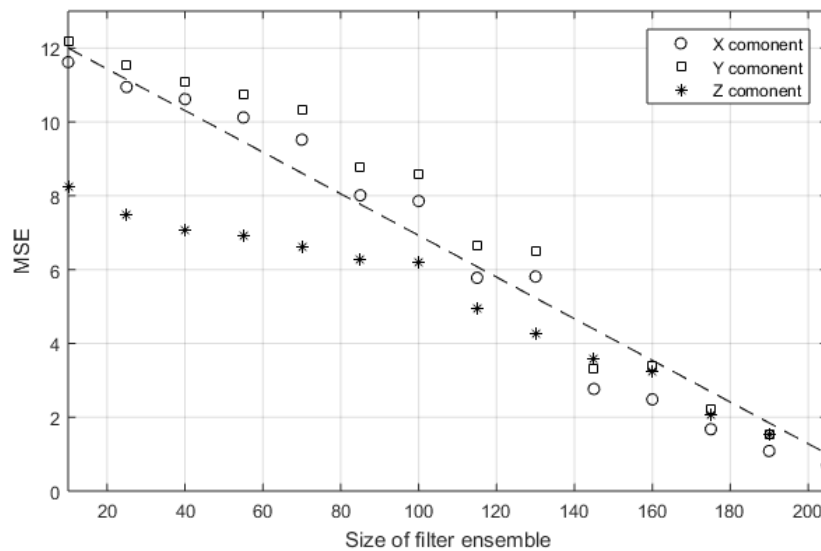


Figure 4. The dependency of the forecast accuracy on the filter ensemble size.

5. Conclusions

During the research, EnKF has been applied to the Lorenz attractor in chaotic mode as a data assimilation method. The methodology has been tested with values of ensemble size ranging from 10 to 205 points.

The conclusion about 10 ensemble being the best option is made basing on the results, since increase in the ensemble size to 200 ensembles causes overfitting. The result is fair for small values of the added noise dispersion.

The results show that the true and estimated values differ (especially strongly at the end of the integration period) due to the chaotic character of the system. When applying EnKF it is necessary to make observations more often in order to obtain accurate forecast.

In some cases, the forecasting fails more than once in unequal periods of time. The most accurate observations of forecast are obtained in a period of time when the model is located in the most stable attractor's area (i.e. the model is spinning around one of the leaves instead of moving from one to another).

References

- [1] Bocquet M Bocquet 2015 Introduction to the principles and methods of data assimilation in geosciences. Publ. by Ecole des Ponts Paris Tech. p 88
- [2] Houtekamer P L and Mitchell H L 1998 Data assimilation using an ensemble Kalman filter technique *Monthly Weather Review* **126(3)** 796-811
- [3] Van den Bossche W 2013 Data assimilation toolbox for Matlab.
- [4] Evensen G 2009 Data assimilation: the ensemble Kalman filter-Springer Science & Business Media
- [5] Van Leeuwen P J et al. SANGOMA: Stochastic Assimilation for the Next Generation Ocean Model Applications EU FP7 SPACE-2011-1 project 283580