Vol 17 No 9, September 2008
1674-1056/2008/17(09)/3241-06

Chinese Physics B

© 2008 Chin. Phys. Soc.
and IOP Publishing Ltd

# GEKF, GUKF and GGPF based prediction of chaotic time-series with additive and multiplicative noises*

Wu Xue-Dong(伍雪冬)† and Song Zhi-Huan(宋执环)

*State Key Laboratory of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou* 310027*, China*

(Received 5 November 2007; revised manuscript received 1 December 2007)

On the assumption that random interruptions in the observation process are modelled by a sequence of independent Bernoulli random variables, this paper generalize the extended Kalman filtering (EKF), the unscented Kalman filtering (UKF) and the Gaussian particle filtering (GPF) to the case in which there is a positive probability that the observation in each time consists of noise alone and does not contain the chaotic signal (These generalized novel algorithms are referred to as GEKF, GUKF and GGPF correspondingly in this paper). Using weights and network output of neural networks to constitute state equation and observation equation for chaotic time-series prediction to obtain the linear system state transition equation with continuous update scheme in an online fashion, and the prediction results of chaotic time series represented by the predicted observation value, these proposed novel algorithms are applied to the prediction of Mackey–Glass time-series with additive and multiplicative noises. Simulation results prove that the GGPF provides a relatively better prediction performance in comparison with GEKF and GUKF.

**Keywords:** additive and multiplicative noises, different generalized nonlinear filtering, chaotic time-series prediction, neural network approximation

**PACC:** 0545

## 1. Introduction

Prediction has played a fundamental role in most spheres of scientific activity. In practice, due to noise or measurement error there are limits to the predictability of chaotic systems, and, in particular, it is by now well understood that long-term prediction of their behaviour is simply not possible due to the exponential divergence of trajectories seen in such systems.[1] Nevertheless, accurate short prediction turns out to be perfectly feasible, therefore, analysis and prediction of chaotic time series become a more and more reliable tool for the study of complicated dynamics from measurements,[2] as a result, the last ten years have seen the development of a variety of new forecasting algorithms based upon the theory of dynamical systems, such as radical basis function neural network (NN),[3,4] multilayer perceptron,[5,6] recurrent (NN)[7] and support vector machine,[8−13] which possess the abilities to approximate nonlinear systems. However, the specified solution based on the training of these networks is satisfactory to the given signal but unsatisfactory to the new input signal, and these networks can only approximate the given trajectory based on the training of the sampling points.

In addition, different filtering methods with continuous update scheme have been employed in the prediction of chaotic time series. Ma and Ten[14] and van der Merwe[15] use the extended Kalman filtering (EKF) and unscented Kalman filtering (UKF) for chaotic time series prediction, Zhang *et al*[16,17] proposes the particle filtering (PF) algorithm to deal with the non-Gaussian noise in chaotic states estimation, but these algorithms[14−18] only consider the uncertain observations perturbed by an additive noise. In fact, there are many practical situations in which the chaotic signal appears in the observation in a random manner so that there are intermittent failures in the observation mechanism at certain times. In these situations, at each time, there is a positive probability called false alarm probability that only noise is observed and, hence, that the observation does not con-

*Project supported by the National Natural Science Foundation of China (Grant No 60774067) and the Natural Science Foundation of Fujian Province of China (Grant No 2006J0017).

†E-mail: woolcn@163.com

tain the chaotic signal. To describe these uncertain observations, the observation equation, with the usual additive measurement noise, is reformulated by multiplying, at any sample time, the signal function by a binary random variable taking the values one and zero (Bernoulli random variable); the value one indicates that the chaotic signal is present in the observation whereas the value zero reflects the fact that the observation is only noise. Thus, the observation equation involves both an additive and a multiplicative noise which models the uncertainty about the chaotic signal being present in each observation, this is called the uncertainty of observation in this paper.

The prediction problem in nonlinear chaotic systems with such uncertain observations has not received much attention up to now. Therefore, the aim in the present paper is to obtain the generalized EKF (GEKF), generalized UKF (GUKF) and generalized Gaussian PF (GGPF) algorithms for a nonlinear chaotic system with uncertain observations in which the Bernoulli random variables describing the uncertainty are independent in combination with the EKF,[19] UKF[20,21] and GPF[22] at first. Then we use weights and network output of NN as state equation and observation equation for chaotic time-series prediction to obtain the linear state transition equation, the result of chaotic time series prediction is represented by the predicted observation value, which is different from Ma and Ten[14] and van der Merwe[15] who use linear regression equation to approximate the state transition equation, a series novel chaotic time-series prediction algorithms with online fashion is proposed based on the GEKF, GUKF and GGPF.

The remainder of this paper is organized as follows, Section 2 describes the nonlinear system state space model; GEKF, GUKF and GGPF formulations used in our study are derived in Section 3; Section 4 introduces our experimental methodology and presents the experimental results, and Section 5 concludes the paper.

## 2. Dynamical system model of chaotic time-series prediction

In order to solve the prediction of chaotic time-series, firstly clean Mackey–Glass time-series is modelled as a nonlinear system. Considering an NN with one hidden layer, let the vector $\boldsymbol{u}_k$ denote the input vector, the vector $\boldsymbol{x}_k$ denote the output of the hidden layer, and the vector $\boldsymbol{y}_k$ denote the output vector of the network. The dynamic behaviour of the NN may be described by a state-space model as follows:

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{u}_k; \boldsymbol{\omega}(f)), \tag{1}$$

$$\boldsymbol{y}_{k+1} = h(\boldsymbol{x}_k; \boldsymbol{\omega}(h)), \tag{2}$$

where $f(\cdot)$ and $h(\cdot)$ denote nonlinear functions characterized by activation functions of the hidden layer, respectively. These nonlinear functions are implemented with parameter vectors $\boldsymbol{\omega}(f)$ and $\boldsymbol{\omega}(h)$.

The dynamics of the NN given in equations (1) and (2) can be recast by stochastic state-space equations, which are suitable for Kalman filtering formulation

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + \boldsymbol{w}(k), \tag{3}$$

$$\boldsymbol{y}'_{k+1} = \gamma_k h(\boldsymbol{\omega}_k, \boldsymbol{u}_k) + \boldsymbol{v}(k), \tag{4}$$

where the parameter vector of the network, $\boldsymbol{\omega}_k$, at time $k$ is reformulated from two parameter vectors

$$\boldsymbol{\omega}_k = \begin{bmatrix} \boldsymbol{\omega}_k(f) \\ \boldsymbol{\omega}_k(h) \end{bmatrix}. \tag{5}$$

Equation (3) is known as the process equation, which specifies the state of the network when it is characterized as a stationary process corrupted by process noise $\boldsymbol{w}(k)$. The state of the system is given by the weight parameters of the network, $\boldsymbol{\omega}_k$. Equation (4) is the observation equation, which represents the desired output of the state-space model as a nonlinear function of the weight vector $\boldsymbol{\omega}_k$ and the input vector to the network $\boldsymbol{u}_k$, this equation is corrupted by additive noise $\boldsymbol{v}(k)$ and multiplicative noise $\gamma_k$, which describes the uncertainty in the observations. The multiplicative noise $\gamma_k$ is a sequence of independent Bernoulli random variables with $P[\gamma_k = 1] = p_k$, and the probability value $1 - p_k$ represents the probability that the observation at time $k$ does not contain the signal.

The process noise $\boldsymbol{w}(k)$ is typically characterized as zero-mean, white noise with covariance given by $E[\boldsymbol{w}(k)\boldsymbol{w}(k)^{\mathrm{T}}] = \boldsymbol{Q}_k$, where $\boldsymbol{Q}_k$ is a diagonal matrix. Similarly, the measurement noise $\boldsymbol{v}(k)$ is also characterized as zero-mean, white noise with covariance given by $E[\boldsymbol{v}(k)\boldsymbol{v}(k)^{\mathrm{T}}] = \boldsymbol{R}_k$, where $\boldsymbol{R}_k$ is a diagonal matrix.

The learning problem of the state-space NN model can be thought of as the estimation of the parameters $\boldsymbol{\omega}_k$, and the prediction results of chaotic time series is presented by the predicted measurement value. Learning procedures employing the GEKF, GUKF and GGPF are given in the following section.

# 3. The generalized nonlinear filtering methods

We present a brief description of the GEKF, GUKF and GGPF which are based on the statistical approximations of vector and Refs.[19–22].

## 3.1. Approximations of the statistics of the vector $\boldsymbol{y} = \gamma h(\boldsymbol{x}) + \boldsymbol{v}$

Due to the values of the variable $\gamma$, the density function of $\boldsymbol{y} = \gamma h(\boldsymbol{x}) + \boldsymbol{v}$ has the following mixture form:

$$g(\boldsymbol{y}) = pg(\boldsymbol{y}|\gamma = 1) + (1-p)g(\boldsymbol{y}|\gamma = 0), \quad (6)$$

where $g(\boldsymbol{y}|\gamma = 1)$ is the density of the vector $h(\boldsymbol{x}) + \boldsymbol{v}$ and $g(\boldsymbol{y}|\gamma = 0)$ is that of $\boldsymbol{v}$; hence, from the hypotheses on the variables involved, we have

$$\begin{aligned} E[\boldsymbol{y}] =& pE[h(\boldsymbol{x})], \\ \mathrm{Cov}(\boldsymbol{y}) =& p\mathrm{Cov}(h(\boldsymbol{x})) + p(1-p)E[h(\boldsymbol{x})] \\ & \times E[h^{\mathrm{T}}(\boldsymbol{x})] + \boldsymbol{R}, \\ \mathrm{Cov}(\boldsymbol{x}, \boldsymbol{y}) =& p\mathrm{Cov}(\boldsymbol{x}, h(\boldsymbol{x})). \end{aligned} \quad (7)$$

Therefore, one way to obtain approximations of the above statistics is to approximate $E[h(\boldsymbol{x})]$, $\mathrm{Cov}(h(\boldsymbol{x}))$ and $\mathrm{Cov}(\boldsymbol{x}, h(\boldsymbol{x}))$ and to substitute the approximations in Eq.(7).

## 3.2. Different generalized nonlinear filtering methods

### 3.2.1. The generalized extended Kalman filtering

The EKF is a minimum mean-square-error estimator based on the Taylor series expansion of nonlinear function.[19] Using only the linear expansion terms and the approximations of the statistics of the vector in section 3.1, it is easy to derive the following update equations for the mean and covariance of the Gaussian approximation to the posterior distribution of the states

(1) Initialize with, $\bar{\boldsymbol{x}}_0 = E[\boldsymbol{x}_0]$, $\boldsymbol{P}_0 = E[(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0) \times (\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)^{\mathrm{T}}]$.

(2) For $k = 1, 2, \ldots, M$

(a) Time update

$$\begin{aligned} \bar{\boldsymbol{x}}_{k|k-1} &= f(\hat{\boldsymbol{x}}_{k-1|k-1}), \\ \boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1} &= \boldsymbol{F}_{k-1}\boldsymbol{P}^{\boldsymbol{xx}}_{k-1|k-1}\boldsymbol{F}^{\mathrm{T}}_{k-1} + \boldsymbol{Q}_{k-1}. \end{aligned} \quad (8)$$

(b) Measurement update equations

$$\begin{aligned} \boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1} =& p_k\boldsymbol{H}_k\boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1}\boldsymbol{H}^{\mathrm{T}}_k + p_k(1-p_k) \\ & \times h(\bar{\boldsymbol{x}}_{k|k-1})h^{\mathrm{T}}(\bar{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{R}_k, \\ \boldsymbol{K}_k =& \boldsymbol{P}^{\boldsymbol{xy}}_{k|k-1}[\boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1}]^{-1} = p_k\boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1}H^{\mathrm{T}}_k[\boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1}]^{-1}, \\ \hat{\boldsymbol{x}}_k =& \bar{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - p_k h(\bar{\boldsymbol{x}}_{k|k-1})), \\ \boldsymbol{P}^{\boldsymbol{xx}}_{k|k} =& \boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1} - p_k\boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1}, \end{aligned} \quad (9)$$

where $\boldsymbol{Q}_k$ is the process noise covariance, $\boldsymbol{R}_k$ is the measurement noise covariance, and $\boldsymbol{K}_k$ is the Kalman gain. $\boldsymbol{F}_k$ and $\boldsymbol{H}_k$ are the Jacobians of the observed model.

### 3.2.2. The generalized unscented Kalman filtering

Another way to obtain approximations of the statistics of $\boldsymbol{y} = \gamma h(\boldsymbol{x}) + \boldsymbol{v}$ is to use a scaled unscented transformation, the GUKF algorithm that updates the mean and covariance of the Gaussian approximation to the posterior distribution of the states is given as follows based on Refs.[20, 21] and section 3.1.

(1) Initialize with, $\bar{\boldsymbol{x}}_0 = E[\boldsymbol{x}_0]$, $\boldsymbol{P}_0 = E[(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0) \times (\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)^{\mathrm{T}}]$

(2) For $k = 1, 2, \ldots, M$

(a) Calculate the sigma points and the weights

$$\begin{aligned} &\boldsymbol{\chi}_{0,k-1} = \bar{\boldsymbol{x}}, \\ &\boldsymbol{\chi}_{i,k-1} = \bar{\boldsymbol{x}} + (\sqrt{(n_x + \lambda)\boldsymbol{P}_k})_i \quad i = 1, \ldots, n_x, \\ &\boldsymbol{\chi}_{i,k-1} = \bar{\boldsymbol{x}} - (\sqrt{(n_x + \lambda)\boldsymbol{P}_k})_i \quad i = n_x + 1, \ldots, 2n_x, \\ &W^{(m)}_0 = \lambda/(n_x + \lambda), \\ &W^{(c)}_0 = \lambda/(n_x + \lambda) + (1 - \alpha^2 + \beta), \\ &W^{(m)}_i = W^{(c)}_i = 1/\{2(n_x + \lambda)\} \quad i = 1, \ldots, 2n_x, \end{aligned} \quad (10)$$

where $n_x$ is the dimension of the state vector, $\alpha$ determines the spread of the sigma points around $\bar{x}$ and is usually set to a small positive value. $\kappa$ is a secondary scaling parameter which is usually set to 0, and $\beta$ is used to incorporate prior knowledge of the distribution of $x$ (for Gaussian distributions, $\beta = 2$ is optimal). $(\sqrt{(n_x + \lambda)\boldsymbol{P}_k})_i$ is the $i$th row of the matrix square root. $\lambda$ is defined by $\lambda = \alpha^2(n_x + \kappa) - n_x$.

(b) Time update

$$\begin{aligned}
\boldsymbol{\chi}_{i,k|k-1} =& f(\boldsymbol{\chi}_{i,k-1}),\\
\bar{\boldsymbol{x}}_{k|k-1} =& \sum_{i=0}^{2n_x} W_i^{(m)} \boldsymbol{\chi}_{i,k|k-1},\\
\boldsymbol{P}_{k|k-1} =& \sum_{i=0}^{2n_a} W_i^{(c)} [\boldsymbol{\chi}^{\boldsymbol{x}}_{i,k|k-1} - \bar{\boldsymbol{x}}_{k|k-1}]\\
& \times [\boldsymbol{\chi}^{\boldsymbol{x}}_{i,t|t-1} - \bar{\boldsymbol{x}}_{k|k-1}]^{\mathrm{T}} + \boldsymbol{Q}_k,\\
\boldsymbol{\gamma}_{i,k|k-1} =& h(\boldsymbol{\chi}^{\boldsymbol{x}}_{i,k|k-1}),\\
\bar{\boldsymbol{z}}_{k|k-1} =& \sum_{i=0}^{2n_x} W_i^{(m)} \boldsymbol{\gamma}_{i,k|k-1}.
\end{aligned} \tag{11}$$

(c) Measurement update

$$\begin{aligned}
\boldsymbol{P}^{\boldsymbol{zz}}_{k|k-1} =& \sum_{i=0}^{2n_x} W_i^{(c)} [h(\boldsymbol{\chi}^{\boldsymbol{x}}_{i,k|k-1}) - \bar{\boldsymbol{z}}_{k|k-1}]\\
& \times [h(\boldsymbol{\chi}^{\boldsymbol{x}}_{i,k|k-1}) - \bar{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}},\\
\boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1} =& p_k \boldsymbol{P}^{\boldsymbol{zz}}_{k|k-1} + p_k(1-p_k)\bar{\boldsymbol{z}}_{k|k-1}\bar{\boldsymbol{z}}^{\mathrm{T}}_{k|k-1} + \boldsymbol{R}_k,\\
\boldsymbol{P}^{\boldsymbol{xz}}_{k|k-1} =& \sum_{i=0}^{2n_x} W_i^{(c)} [\boldsymbol{\chi}_{i,k|k-1} - \bar{\boldsymbol{x}}_{k|k-1}]\\
& \times [h(\boldsymbol{\chi}_{i,k|k-1}) - \bar{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}},\\
\boldsymbol{K}_k =& P^{\boldsymbol{xy}}_{k|k-1} [\boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1}]^{-1} = p_k \boldsymbol{P}^{\boldsymbol{xz}}_{k|k-1} [\boldsymbol{P}^{\boldsymbol{yy}}_{k|k-1}]^{-1},\\
\hat{\boldsymbol{x}}_{k|k} =& \bar{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - p_k \bar{\boldsymbol{z}}_{k|k-1}),\\
\boldsymbol{P}^{\boldsymbol{xx}}_{k|k} =& \boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1} - p_k \boldsymbol{K}_k (\boldsymbol{P}^{\boldsymbol{xz}}_{k|k-1})^{\mathrm{T}},
\end{aligned} \tag{12}$$

where $\boldsymbol{Q}_k$, $\boldsymbol{R}_k$ and $\boldsymbol{K}_k$ are the same as GEKF.

### 3.2.3. The generalized Gaussian particle filtering

From Corollary 1 of Ref.[22], we may deduce that the GPF provides better approximations to the integrations involved in the update processes than EKF and UKF in the presence of severe nonlinearities. The complete GGPF algorithm that updates the mean and covariance of the Gaussian approximation to the posterior distribution of the states is given as follows (more details can be found in Ref.[22]):

(1) Initialize with $\bar{\boldsymbol{x}}_0 = E[\boldsymbol{x}_0]$, $\boldsymbol{P}_0 = E[(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)^{\mathrm{T}}]$

(2) GGPF measurement update algorithm

(a) Draw samples from the important function $q(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ and denote them as $\{\boldsymbol{x}_k^{(i)}\}_{j=1}^N$ (where $N$ is the number of sample particles).

(b) Obtain the respective weights by

$$w_k^{(i)} = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}^{(i)}_{k|k}) N(\boldsymbol{x}_k^{(i)}; \hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1})}{q(\boldsymbol{x}^{(i)}_{k|k}; \boldsymbol{x}_k|\boldsymbol{y}_{1:k})}. \tag{13}$$

The choice of the importance sampling function $q(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ depends on the problem (see Refs.[23, 24] for details). A simple choice for $q(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ is $p(\boldsymbol{x}_k|\boldsymbol{y}_{0:k-1}) = N(\boldsymbol{x}_k^{(i)};\ \hat{\boldsymbol{x}}_{k|k-1},\ \boldsymbol{P}^{\boldsymbol{xx}}_{k|k-1})$ since samples from this density can be easily obtained. In this case, we can obtain

$$\begin{aligned}
w_k^{(i)} =& p(\boldsymbol{y}_k|\boldsymbol{x}^{(i)}_{k|k}) = (2\pi)^{-n_x/2} |\boldsymbol{R}_k|^{0.5}\\
& \times \exp\Big[ -\frac{1}{2} (\boldsymbol{y}_k - p_k h(\boldsymbol{x}_{k|k-1})^{(i)})^{\mathrm{T}} \boldsymbol{R}_k^{-1}\\
& \times (\boldsymbol{y}_k - p_k h(\boldsymbol{x}_{k|k-1})^{(i)}) \Big].
\end{aligned} \tag{14}$$

(c) Normalize the weights as

$$\tilde{w}_k^{(i)} = w_k^{(i)} \Big/ \sum_{i=1}^N w_k^{(i)}. \tag{15}$$

(d) Estimate the mean and covariance by

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k|k} =& \sum_{i=1}^N \tilde{w}_k^{(i)} \boldsymbol{x}^{(i)}_{k|k},\\
\boldsymbol{P}^{\boldsymbol{xx}}_{k|k} =& \sum_{i=1}^N \tilde{w}_k^{(i)} (\boldsymbol{x}^{(i)}_{k|k} - \hat{\boldsymbol{x}}_{k|k})(\boldsymbol{x}^{(i)}_{k|k} - \hat{\boldsymbol{x}}_{k|k})^{\mathrm{T}}.
\end{aligned} \tag{16}$$

(3) GGPF time update algorithm

(a) Retain weighted samples of $p(\boldsymbol{x}_k|\boldsymbol{y}_{0:k})$ obtained in the measurement update and denote them as $\{\boldsymbol{x}_k^{(i)}, w_k^{(i)}\}_{i=1}^N$.

(b) For $j = 1, \ldots, N$, sample from $p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k = \boldsymbol{x}_k^{(i)})$ to obtain $\{\boldsymbol{x}_{k+1}^{(i)}\}_{i=1}^N$.

(c) Compute the mean and covariance as

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k+1|k} =& \sum_{i=1}^N w_k^{(i)} \boldsymbol{x}^{(i)}_{k+1|k},\\
\boldsymbol{P}^{\boldsymbol{xx}}_{k|k} =& \sum_{i=1}^N \tilde{w}_k^{(i)} (\boldsymbol{x}^{(i)}_{k+1|k} - \hat{\boldsymbol{x}}_{k+1|k})(\boldsymbol{x}^{(i)}_{k+1|k} - \hat{\boldsymbol{x}}_{k+1|k})^{\mathrm{T}}.
\end{aligned} \tag{17}$$

## 4. Experiments

In all the experiments carried out in this paper, we have made use of the Mackey–Glass delay-differential equation.[25]

$$\frac{\mathrm{d}x(t)}{\mathrm{d}(t)} = \frac{ax(t-\tau)}{1 + x^r(t-\tau)} - bx(t). \tag{18}$$

This equation was derived as a model of blood production and the Mackey–Glass chaotic signal can be obtained from its integration. $x(t)$ represents the concentration of blood at time $t$, that is, when the blood is produced and $x(t-\tau)$ is the concentration of blood when there is a demand for more blood. In patients

with different pathologies, such as leukemia, the time $t$ may become very large, and the concentration of blood will oscillate, becoming chaotic when $\tau \geq 17$. We are thus going to choose values for $\tau$ in this range so as to test our strategy in the hardest conditions. In order to be as standard as possible in the series we employ, the parameter are chosen as $a = 0.2$, $b = 0.1$, $r = 10$, $\tau = 17$ and $x(0) = 0.1$ individually in chaotic time-series equation (18). As for the noise of clean chaotic time-series, an additive white Gaussian noise is added to the clean chaotic time-series with 3dB signal noise ratio, and the probability value that the observation in each time consists of noise alone and does not contain the chaotic signal is a sequence of independent Bernoulli random variables with known probability which varies from 0.1 to 0.9.

A 6-6-1 NN, which constitutes six inputs, six hidden neurons and one output neuron, is trained with GEKF, GUKF and GGPF algorithms in an online fashion. To enable a fair result of the chaotic time-series prediction proposed in this paper, the prediction is averaged across a Monte Carlo simulation consisting of 50 runs. Each run is carried out with a different noise sample, and the simulation is run for a time of 1000 steps. After the training phase for 500 time steps, the chaotic time-series prediction results and the prediction absolute error for the remaining 500 time steps with $N$=200 and $p$=0.8 are shown in Fig.1 (where "Idea" is the clean Mackey–Glass chaotic time-series) and Fig.2. Figure 3 plots the mean square error (MSE) for 50 random realizations with particles $N$=200 and $p$=0.8. A comparison of prediction average MSEs with different uncertain probability which varies from $p$=0.1 to $p$=0.9 are also shown in Fig.4.

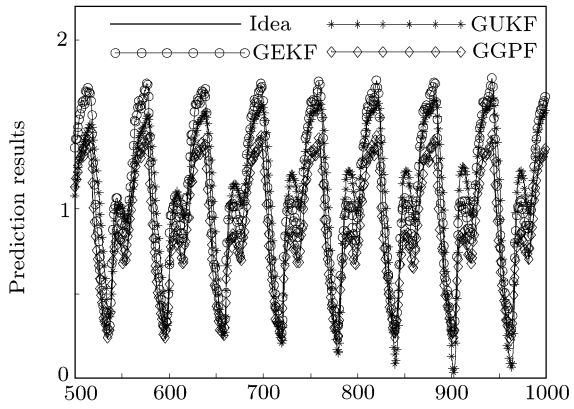



**Fig.1.** Mackey–Glass time-series prediction results with probability $p$=0.8, $N$=200 for GPF (50 runs average).

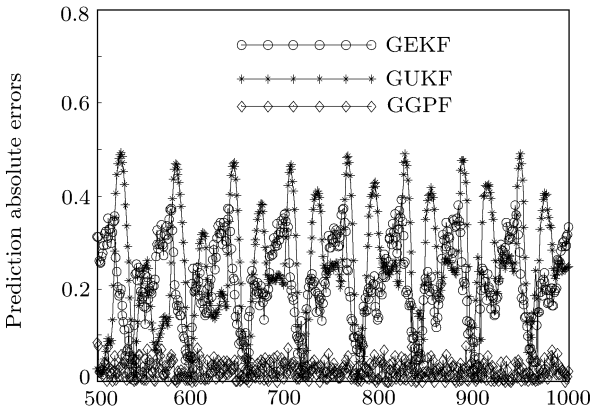



**Fig.2.** Mackey–Glass time-series prediction absolute errors with probability $p$=0.8, $N$=200 for GPF (50 runs average).

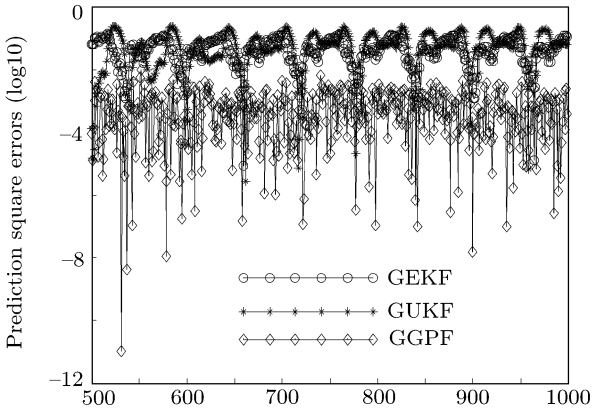



**Fig.3.** Mackey–Glass time-series prediction square errors with probability $p$=0.8, $N$=200 for GPF (50 runs average).

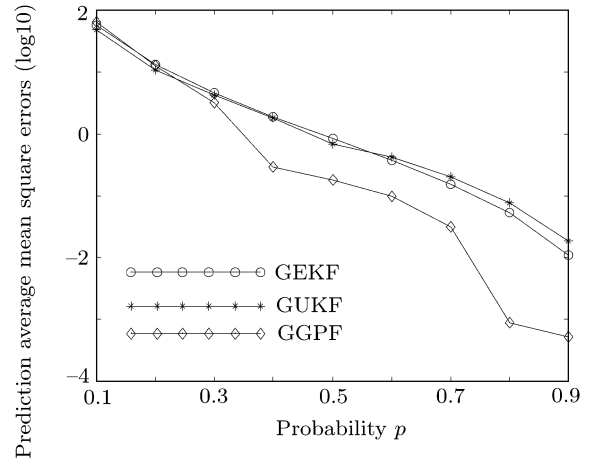



**Fig.4.** The average mean square errors of Mackey–Glass time-series prediction with different uncertain probability $p$, $N$=200 for GPF (50 runs average).

# 5. Conclusions

The prediction problem in nonlinear chaotic systems with uncertain observations perturbed by additive and multiplicative noises at the same time has

not received much attention up to now, this paper obtains the GEKF, GUKF and GGPF algorithms for a nonlinear chaotic system with such uncertain observations. Using weights and network output of NN to constitute state equation and observation equation for chaotic time-series prediction to obtain the linear state transition equation, we apply these proposed novel algorithms to the prediction of Mackey–Glass time-series and present a relatively all-sided comparison. Simulation results of GEKF, GUKF and GGPF based on Mackey–Glass time-series prediction with synthetic data prove that GGPF is a better choice in comparison with GEKF and GUKF.

## References


[1] Stark J 1993 *J. Nonlinear Science* **3** 197

[2] Zhang J, Lam K C, Yan W J, Gao H and Li Y 2004 *Computers and Electrical Engineering* **30** 1

[3] Harpham C and Dawson C W 2006 *Neurocomputing* **69** 2161

[4] Li J and Liu J H 2005 *Acta Phys. Sin* **54** 4569 (in Chinese)

[5] Dudul S V 2005 *Applied Soft Computing* **5** 333

[6] Übeyli E D and Güler I 2004 *Engineering Applications of Artificial Intelligence* **17** 567

[7] Luo J F, Zheng J L and Sun S Y 1999 *Chin. J. Radio Science* **14** 172 (in Chinese)

[8] Yu Z H and Cai Y L 2006 *Acta Phys. Sin.* **55** 1659 (in Chinese)

[9] Ye M Y and Wang X D 2004 *Chin. Phys.* **13** 454

[10] Lau K W and Wu Q H 2007 *Pattern Recognition* in press, corrected proof doi:10.1016/j.patcog.2007.08.013

[11] Yu Z H and Cai Y L 2006 *Acta Phys. Sin.* **55** 1659 (In Chinese)

[12] Zhang J S, Dang J L and Li H C 2007 *Acta Phys. Sin.* **56** 0067 (in Chinese)

[13] Ren R, Xu J and Zhu S H 2006 *Acta Phys. Sin.* **55** 0555 (in Chinese)

[14] Ma J and Ten J F 2004 *Int. Conf. Machine Learning and Cybernetics* (Shanghai, China) **2** 687

[15] van der Merwe R 2004 *Sigma-point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models* (PhD Thesis, Oregon Health & Science University) p1

[16] Zhang B, Chen M Y, Zhou D H and Li Z X 2007 *Chaos, Solitons and Fractals* **32** 1491

[17] Zhang B, Chen M Y and Zhou D H 2006 *Chaos, Solitons and Fractals* **30** 1273

[18] Teck P L and Sadasivan P 2007 *Phys. Lett.* A **365** 309

[19] Chui C K and Chen G 1999 *Kalman Filtering with Real Time Applications* (New York: Springer) p1

[20] Julier S J, Uhlmann J K and Durrant-Whyte H F 1995 *American Control Conf.* **3** 1628

[21] Julier S J and Uhlmann J K 2004 *Proc. IEEE* **92** 401

[22] Kotecha J H and Djuric P M 2003 *IEEE Trans. Signal Processing* **51** 2592

[23] Doucet A, Godsill S J and Andrieu C 2000 *Statist. Comput.* **10** 197

[24] Geweke J 1989 *Econometrica* **33** 1317

[25] Mackey M C and Glass L 1977 *Science* **197** 287