Generalized unscented Kalman filtering based radial basis function neural network for the prediction of ground radioactivity time series with missing data*

Wu Xue-Dong(伍雪冬)[a)†], Wang Yao-Nan(王耀南)[b)],
Liu Wei-Ting(刘维亭)[a)], and Zhu Zhi-Yu(朱志宇)[a)]

[a)] *School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang* 212003, *China*
[b)] *College of Electrical and Information Engineering, Hunan University, Changsha* 410082, *China*

(Received 5 October 2010; revised manuscript received 19 January 2011)

On the assumption that random interruptions in the observation process are modeled by a sequence of independent Bernoulli random variables, we firstly generalize two kinds of nonlinear filtering methods with random interruption failures in the observation based on the extended Kalman filtering (EKF) and the unscented Kalman filtering (UKF), which were shortened as GEKF and GUKF in this paper, respectively. Then the nonlinear filtering model is established by using the radial basis function neural network (RBFNN) prototypes and the network weights as state equation and the output of RBFNN to present the observation equation. Finally, we take the filtering problem under missing observed data as a special case of nonlinear filtering with random intermittent failures by setting each missing data to be zero without needing to pre-estimate the missing data, and use the GEKF-based RBFNN and the GUKF-based RBFNN to predict the ground radioactivity time series with missing data. Experimental results demonstrate that the prediction results of GUKF-based RBFNN accord well with the real ground radioactivity time series while the prediction results of GEKF-based RBFNN are divergent.

Keywords: prediction of time series with missing data, random interruption failures in the observation, neural network approximation

PACS: 92.20.Td, 05.45.Tp, 05.10.Gg **DOI:** 10.1088/1674-1056/20/6/069201

1. Introduction

The world is naturally radioactive and around 90% of human radiation exposure arises from natural sources such as cosmic radiation, exposure to radon gas and terrestrial radiation.[1] Since radionuclides are not uniformly distributed, knowledge of their distributions in soils and rocks plays an important role in radiation protection and measurement.[2] Some of the exposures are fairly constant and uniform for all individuals everywhere. Other exposures vary widely depending on the location. Therefore, the assessment and the prediction of the radiation dose from natural sources are of particular importance. In this work, we study the time series prediction of radioactivity in the ground at 2-hour intervals over one year (from July 2005 to June 2006).

Time series prediction has played a fundamental role in most spheres of scientific activity. As a result, in the last twenty years a variety of new prediction algorithms based on the theory of dynamical systems have been developed, such as neural networks (NNs)[3−8] and support vector machines (SVMs),[9−11] which possess the ability to approximate nonlinear systems. However, nonlinear time series literature highlights that the good fitting results of nonlinear models do not guarantee an equally good prediction performance.[12] One main reason is that their dynamics and properties are change over time. Moreover, in the modeling of time series using NNs and SVRs, another key problem is the inherent noise of the time series. Using observations without paying attention to noise may lead to fitting unwanted data and may damage the approximation function, and the noise in the data could lead to over-fitting or under-fitting problems.[13] Furthermore, a survey of existing literature reveals that there is a need to

*Project supported by the State Key Program of the National Natural Science of China (Grant No. 60835004), the Natural Science Foundation of Jiangsu Province of China (Grant No. BK2009727), the Natural Science Foundation of Higher Education Institutions of Jiangsu Province of China (Grant No. 10KJB510004), and the National Natural Science Foundation of China (Grant No. 61075028).

†Corresponding author. E-mail: woolcn@163.com

© 2011 Chinese Physical Society and IOP Publishing Ltd

develop efficient forecasting models involving less computational load and a faster forecasting capability.[14] Therefore, how to handle the adaptive ability, the additive noise and the computational load is a important but difficult task for building a prediction model. Keeping these problems in mind, some different filtering methods with a continuous update scheme have been employed in the time series prediction. Ma and Ten[15] and van der Merwe[16] used the EKF and UKF for Mackey–Glass time series prediction, Zhang *et al.*[17,18] proposed the particle filtering (PF) algorithm to deal with the non-Gaussian noise in time series state estimation, but these algorithms[15−18] were studied under the condition without missing data. In fact, irregular observations, missing values and outliers are common in time series data,[19] and many frameworks for pre-estimating the missing data have been developed.[19,20]

In order to solve these problems and at the same time include adaptive ability, additive noises, computational load and missing data for building the time series predicted model, in this work we suggest the GEKF-based RBFNN and the GUKF-based RBFNN for the prediction of ground radioactivity time series with missing records. RBFNN can approximate virtually any measurable function up to an arbitrary degree of accuracy and nonlinear filtering approaches with adaptive ability can handle additive noises and reduce the computational load to some degree due to their iterative algorithm structures. On the assumption that random interruptions in the observation process are modeled by a sequence of independent Bernoulli random variables, we generalize two kinds of nonlinear filtering methods with random interruptions in the observation based on the EKF and the UKF, which are shortened as GEKF and GUKF, respectively. When the RBFNN and the GEKF or GUKF are combined, the GEKF-based RBFNN and the GUKF-based RBFNN can be used to predict the ground radioactivity time series with missing observed data by setting each missing observed data to be zero. The GEKF-based RBFNN and the GUKF-based RBFNN used here are formulated in the predictive corrective form which underlies the general unobserved component approach to the state space estimation and prediction, and the prediction results are represented by the predicted observation values of GEKF-based RBFNN and GUKF-based RBFNN.

The rest of this paper is organized as follows. The dynamic model of time series prediction is introduced in Section 2. In Section 3 the GEKF and GUKF principles are presented. Then, the GEKF-based RBFNN and the GUKF-based RBFNN applied in this study are presented in Section 4. Subsequently, some conclusions drawn from the present study are given in Section 5.

2. Dynamic model of time series prediction

2.1. Radial basis function neural network configuration

A radial basis function (RBF) consists of the m-dimensional input $\boldsymbol{u} = [\, I_1 \;\; I_2 \;\cdots\; I_m \,]$ passed directly toward a hidden layer where there are assumed to be c neurons in the hidden layer. Each of the c neurons in the hidden layer uses an activation function, which is a function of the Euclidean distance between the input and an m-dimensional prototype vector. Each hidden neuron contains its own prototype vector as a parameter. The output of each hidden neuron is then weighted and passed toward the output layer. The outputs of the network consist of the sum of the weighted hidden layer neurons. Figure 1 shows a schematic of an RBFNN. The response of an RBFNN in the form of Fig. 1, where the hidden layer function is in the form of $g(\boldsymbol{s}) = \exp[-\boldsymbol{s}/\beta^2]$ (where β is a real constant) can be written as follows:

$$\begin{aligned} y &= [\, w_{11} \;\cdots\; w_{1c} \,][\, g\,\|\boldsymbol{u}-\boldsymbol{v}_1\|^2 \;\cdots\; g\,\|\boldsymbol{u}-\boldsymbol{v}_c\|^2 \,]^{\mathrm{T}} \\ &= \boldsymbol{W}[\, g\,\|\boldsymbol{u}-\boldsymbol{v}_1\|^2 \;\cdots\; g\,\|\boldsymbol{u}-\boldsymbol{v}_c\|^2 \,]^{\mathrm{T}} \\ &= h(\boldsymbol{W}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_c, \boldsymbol{u}). \end{aligned} \tag{1}$$

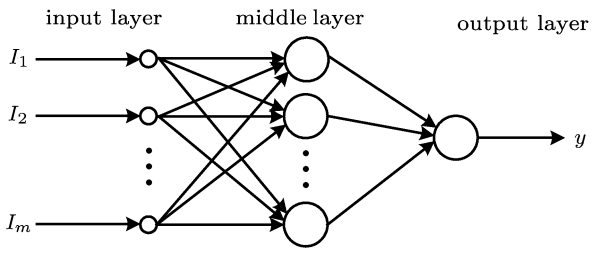


Fig. 1. Radial basis function network architecture.

2.2. RBFNN based nonlinear filtering model with random intermittent failures in the observation

Inspired by the successful use of Kalman filtering for training NNs, we present the predictive nonlinear

model with random intermittent failures in the observation in this section. In general, we can view the optimization of the weight matrix $\boldsymbol{W}$ and the prototypes $\boldsymbol{v}_j$ as a weighted least-squares minimization problem, where the error vector is the difference between the RBFNN outputs and the target values for those outputs. Consider the RBFNN of Fig. 1 with m inputs, c prototypes, and n outputs. In order to cast the optimization problem in a form suited for the nonlinear filtering methods, we let the elements of the weight matrix $\boldsymbol{W}$ and the elements of the prototypes $\boldsymbol{v}_j$ constitute the state of a nonlinear system and we let the outputs of the RBFNN constitute the output of the nonlinear system to which the nonlinear filter is applied. Then the state of the nonlinear system can be represented as

$$\boldsymbol{x} = [\, \boldsymbol{W}^{\mathrm{T}} \ \boldsymbol{v}_1^{\mathrm{T}} \ \cdots \ \boldsymbol{v}_c^{\mathrm{T}} \,]^{\mathrm{T}}. \tag{2}$$

The vector $\boldsymbol{x}$ thus consists of all $(n(c+1)+mc)$ of the RBFNN parameters arranged in a linear array. In order to execute a stable nonlinear filtering algorithm, we need to add some artificial process noise $\boldsymbol{v}_k$ and observation noise $\boldsymbol{w}_k$ to the system model. In addition, we use a sequence of independent Bernoulli random variables γ_k (in the binary switching sequence taken are the values 0 or 1) to denote whether the observation can be observed. According to Eqs. (1) and (2), the nonlinear system model to which the nonlinear filter with random intermittent failures in the observation can be applied is

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{v}_k, \\ \boldsymbol{y}_k = \gamma_k h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k, \end{cases} \tag{3}$$

where $\boldsymbol{x}_k$ is the system state vector at time k, $\boldsymbol{v}_k$ is the process noise, $\boldsymbol{y}_k$ is the observation vector, $\boldsymbol{w}_k$ is the observation noise, $h(\bullet)$ is the nonlinear vector function of the state vector and input vector. Variable γ_k, which describes the random intermittent failures in the observation, is a sequence of independent Bernoulli random variables.

After obtaining the nonlinear filtering approach with random intermittent failures in the observation, we can take the filtering problem under missing observed data as a special case of nonlinear filtering with random intermittent failures by setting $\gamma_k = 0$ or $\gamma_k = 1$ when the k-th datum is missed or provided and then use these filtering approaches to predict the ground radioactivity time series.

3. Generalized nonlinear filtering approaches

Due to the values of the variable γ_k, the density function of $\boldsymbol{y}_k = \gamma_k h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k$ has the following mixture form:

$$g(\boldsymbol{y}_k) = p_k g(\boldsymbol{y}_k|\gamma_k = 1) + (1 - p_k) g(\boldsymbol{y}_k|\gamma_k = 0), \tag{4}$$

where $g(\boldsymbol{y}_k|\gamma_k = 1)$ is the density of the vector $h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{w}_k$ and $g(\boldsymbol{y}_k|\gamma_k = 0)$ is the density of $\boldsymbol{w}_k$. Hence,

$$\begin{cases} E[\boldsymbol{y}_k] = p_k \cdot E[h(\boldsymbol{x}_k, \boldsymbol{u}_k)], \\ \mathrm{Cov}[\boldsymbol{y}_k] = p_k \cdot \mathrm{Cov}[h(\boldsymbol{x}_k, \boldsymbol{u}_k)] + p_k(1 - p_k) \\ \qquad\qquad \cdot E[h(\boldsymbol{x}_k, \boldsymbol{u}_k)] \cdot E[h^{\mathrm{T}}(\boldsymbol{x}_k, \boldsymbol{u}_k)] + \boldsymbol{R}_k, \\ \mathrm{Cov}[\boldsymbol{x}_k, y_k] = p_k \cdot \mathrm{Cov}[\boldsymbol{x}_k, h(\boldsymbol{x}_k, \boldsymbol{u}_k)]. \end{cases} \tag{5}$$

Therefore, one way to obtain approximations of the above statistics is to approximate $E[h(\boldsymbol{x}_k, \boldsymbol{u}_k)]$, $\mathrm{Cov}[h(\boldsymbol{x}_k, \boldsymbol{u}_k)]$, $\mathrm{Cov}[\boldsymbol{x}_k, h(\boldsymbol{x}_k, \boldsymbol{u}_k)]$, and to substitute the approximations in Eq. (5). So it is easy to derive the GEKF and GUKF based on Refs. [21] and [22].

3.1. Generalized extended Kalman filtering

The EKF is a minimum mean squared error (MSE) estimator based on the Taylor series expansion of nonlinear function. 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.

(i) Initialize with $\bar{\boldsymbol{x}}_{1|0} = E[\boldsymbol{x}_0]$, $\boldsymbol{P}_{1|0}^{xx} = E[(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_{1|0})(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_{1|0})^{\mathrm{T}}]$.

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

(a) Measurement update equations,

$$\begin{cases} \boldsymbol{P}_{k|k-1}^{yy} = p_k \boldsymbol{H}_k \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^{\mathrm{T}} + p_k(1 - p_k) \\ \qquad\qquad \times h(\bar{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_k) h^{\mathrm{T}}(\bar{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_k) + \boldsymbol{R}_k, \\ \boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}^{xy} [\boldsymbol{P}_{k|k-1}^{yy}]^{-1} \\ \qquad = p_k \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^{\mathrm{T}} [\boldsymbol{P}_{k|k-1}^{yy}]^{-1}, \\ \hat{\boldsymbol{x}}_k = \bar{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k [\boldsymbol{y}_k - p_k h(\bar{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_k)], \\ \boldsymbol{P}_{k|k}^{xx} = \boldsymbol{P}_{k|k-1}^{xx} - p_k \boldsymbol{K}_k \boldsymbol{H}_k \boldsymbol{P}_{k|k-1}^{xx}. \end{cases} \tag{6}$$

(b) Time update,

$$\begin{cases} \bar{\boldsymbol{x}}_{k+1|k} = f(\hat{\boldsymbol{x}}_{k|k}), \\ \boldsymbol{P}_{k+1|k}^{xx} = \boldsymbol{F}_k \boldsymbol{P}_{k|k}^{xx} \boldsymbol{F}_k^{\mathrm{T}} + \boldsymbol{Q}_k, \end{cases} \tag{7}$$

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

3.2. Generalized unscented Kalman filtering

The GUKF algorithm that updates the mean and covariance of the Gaussian approximation to the posterior distribution of the states is given as follows (See Ref. [22] by Julier and Uhlmann for an in-depth discussion and some generalizations on UKF).

i) Initialize with $\bar{\boldsymbol{x}}_{1|0} = E[\boldsymbol{x}_0]$, $\boldsymbol{P}^{xx}_{1|0} = E[(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_{1|0})(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_{1|0})^{\mathrm{T}}]$.

ii) For $k = 1, 2, \ldots, M$.

(a) Calculate the sigma points and the weights

$$\begin{cases} \boldsymbol{\chi}_{0,k|k-1} = \bar{\boldsymbol{x}}_{k|k-1}, \\ \boldsymbol{\chi}_{i,k|k-1} = \bar{\boldsymbol{x}}_{k|k-1} + \left(\sqrt{(n_{\boldsymbol{x}}+\lambda)\boldsymbol{P}_{k|k-1}}\right)_i, \ i = 1, \ldots, n_{\boldsymbol{x}}, \\ \boldsymbol{\chi}_{i,k|k-1} = \bar{\boldsymbol{x}}_{k|k-1} - \left(\sqrt{(n_{\boldsymbol{x}}+\lambda)\boldsymbol{P}_{k|k-1}}\right)_i, \ i = n_x + 1, \ldots, 2n_{\boldsymbol{x}}, \\ W_0^{(m)} = \lambda/(n_{\boldsymbol{x}} + \lambda), \\ W_0^{(c)} = \lambda/(n_{\boldsymbol{x}} + \lambda) + (1 - \alpha^2 + \beta), \\ W_i^{(m)} = W_i^{(c)} = 1/\left\{2(n_{\boldsymbol{x}} + \lambda)\right\}, \qquad i = 1, \ldots, 2n_{\boldsymbol{x}}, \end{cases} \tag{8}$$

where $n_{\boldsymbol{x}}$ is the dimension of the state vector, α determines the spread of the sigma points around $\bar{\boldsymbol{x}}_{k|k-1}$ and is usually set to be a small positive value. $\left(\sqrt{(n_{\boldsymbol{x}}+\lambda)\boldsymbol{P}_{k|k-1}}\right)_i$ is the i-th row of the matrix squared root. λ is determined by $\lambda = \alpha^2(n_{\boldsymbol{x}} + \kappa) - n_{\boldsymbol{x}}$ with κ being a secondary scaling parameter which is usually set to be 0, and β is used to incorporate prior knowledge of the distribution of $\boldsymbol{x}$ (for Gaussian distributions, $\beta = 2$ is optimal).

(b) Measurement update

$$\begin{cases} \boldsymbol{P}^{zz}_{k|k-1} = \sum_{i=0}^{2n_{\boldsymbol{x}}} W_i^{(c)} \left[h(\chi_{i,k|k-1}, \boldsymbol{u}_k) - \bar{\boldsymbol{z}}_{k|k-1}\right] \left[h(\chi_{i,k|k-1}, \boldsymbol{u}_k) - \bar{\boldsymbol{z}}_{k|k-1}\right]^{\mathrm{T}}, \\ \boldsymbol{P}^{yy}_{k|k-1} = p_k \boldsymbol{P}^{zz}_{k|k-1} + p_k(1 - p_k)\bar{\boldsymbol{z}}_{k|k-1}\bar{\boldsymbol{z}}^{\mathrm{T}}_{k|k-1} + \boldsymbol{R}_k, \\ \boldsymbol{P}^{xz}_{k|k-1} = \sum_{i=0}^{2n_{\boldsymbol{x}}} W_i^{(c)} \left[\chi_{i,k|k-1} - \bar{\boldsymbol{x}}_{k|k-1}\right] \left[h(\chi_{i,k|k-1}, \boldsymbol{u}_k) - \bar{\boldsymbol{z}}_{k|k-1}\right]^{\mathrm{T}}, \\ \boldsymbol{K}_k = \boldsymbol{P}^{xy}_{k|k-1}[\boldsymbol{P}^{yy}_{k|k-1}]^{-1} = p_k \boldsymbol{P}^{xz}_{k|k-1}[\boldsymbol{P}^{yy}_{k|k-1}]^{-1}, \\ \hat{\boldsymbol{x}}_{k|k} = \bar{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - p_k \bar{\boldsymbol{z}}_{k|k-1}), \\ \boldsymbol{P}^{xx}_{k|k} = \boldsymbol{P}^{xx}_{k|k-1} - p_k \boldsymbol{K}_k(\boldsymbol{P}^{xz}_{k|k-1})^{\mathrm{T}}. \end{cases} \tag{9}$$

(c) Time update

$$\begin{cases} \boldsymbol{\chi}_{i,k+1|k} = f(\boldsymbol{\chi}_{i,k|k}), \quad \bar{\boldsymbol{x}}_{k+1|k} = \sum_{i=0}^{2n_{\boldsymbol{x}}} W_i^{(m)} \boldsymbol{\chi}_{i,k+1|k}, \\ \boldsymbol{P}_{k+1|k} = \sum_{i=0}^{2n_a} W_i^{(c)} \left[\boldsymbol{\chi}_{i,k+1|k} - \bar{\boldsymbol{x}}_{k+1|k}\right] \left[\boldsymbol{\chi}_{i,k+1|k} - \bar{\boldsymbol{x}}_{k+1|k}\right]^{\mathrm{T}} + \boldsymbol{Q}_k, \\ \boldsymbol{\gamma}_{i,k+1|k} = h(\boldsymbol{\chi}_{i,k+1|k}, \boldsymbol{u}_k), \quad \bar{\boldsymbol{z}}_{k+1|k} = \sum_{i=0}^{2n_{\boldsymbol{x}}} W_i^{(m)} \boldsymbol{\gamma}_{i,k+1|k}, \end{cases} \tag{10}$$

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.

4. Experimental results and analysis

4.1. Experimental data and setup

In this paper we explore the applicability and the reliability of these proposed methods for the time series prediction of radioactivity in the ground in 2-hour intervals over one year (from July 2005 to June 2006). A few data are missing due to a technical problem and the data are available at http://robjhyndman.com/TSDL. In order to compare the different data sets analysed, we have normalized data sets between 0 and 1 and obtained a normalized data y_k as follows:

$$y_k = \frac{y_{ok} - \min(y_{ok})}{\max(y_{ok}) - \min(y_{ok})}, \tag{11}$$

where y_{ok} is the original data, and y_k is the normalized data.

An RBFNN, which constitutes three inputs, fifteen prototypes and one output, is trained with GEKF and GUKF algorithms in an online fashion, respectively. To afford a fair prediction result, the prediction is averaged across a Monte Carlo simulation consisting of 50 runs. In addition, generally the prediction accuracy gets better and better as the iterative steps increase as the proposed filtering model is convergent, so we take the last predicted 95 data as the evaluation criterion of experimental results.

Furthermore, the GEKF and the GUKF are based on the minimum MSE principle, so we take the average mean square error (AMSE) $e_{\rm av}$ to compare the performances of the GEKF-based RBFNN and the GUKF-based RBFNN. The $e_{\rm av}$ is defined as follows:

$$e_{\rm av} = \frac{1}{N}\sum_{k=1}^{N} |y_k - \hat{y}_{{\rm p}k}|^2, \tag{12}$$

where y_k and $\hat{y}_{{\rm p}k}$ are observed and predicted values respectively, and N is the prediction samples.

4.2. Experimental results and discussion

The last 95 predicted data of radioactivity in the ground in 2-hour intervals over one year are shown in Fig. 2 (where "observed" means the real radioactivity time series) and the corresponding absolute error and MSE are shown in Figs. 3 and 4. The normalized AMSEs of GEKF-based RBFNN and GUKF-based RBFNN for the predicted results of radioactivity in the ground are 0.0197 and 6.3301×10^{-5}, respectively. According to the experimental results, it turns out that the GUKF provides much better estimation accuracy than the GEKF, this is our major conclusion and recommendation.

Now we discuss the reasons for the observed results. We know from Eq. (3) that our dynamic system model is nonlinear (although the state transition equation is linear, and the observation equation is nonlinear). The improved performance of the GUKF compared with the GEKF is due to the improved covariance accuracy. The covariance estimations can be quite different for the two filters as shown in Eqs. (6) and (9), and this again makes a difference in Kalman gain in the measurement update equation and hence the efficiency of the measurement update step.

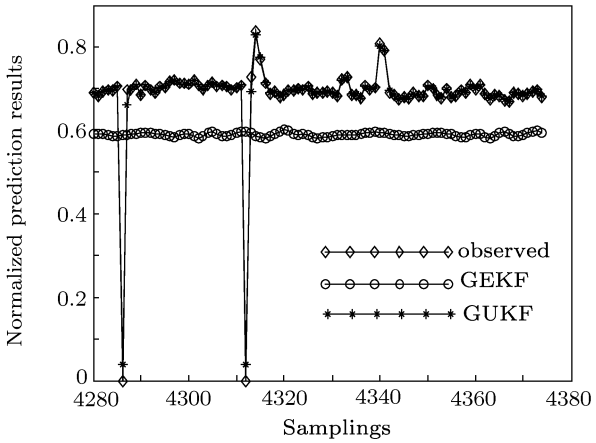


Fig. 2. Prediction results of radioactivity in the ground in 2-hour intervals over one year (from July 2005 to June 2006).

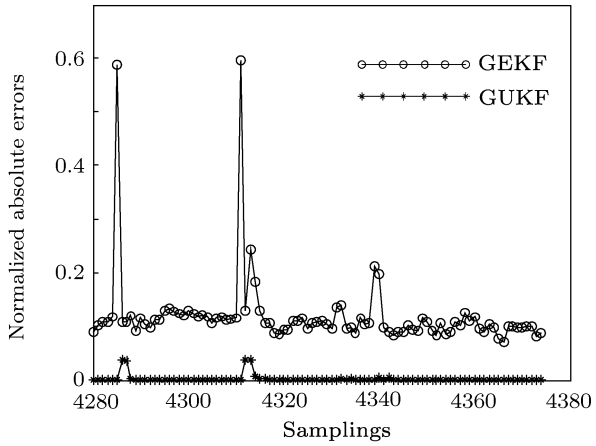


Fig. 3. Absolute errors of radioactivity time series prediction results.

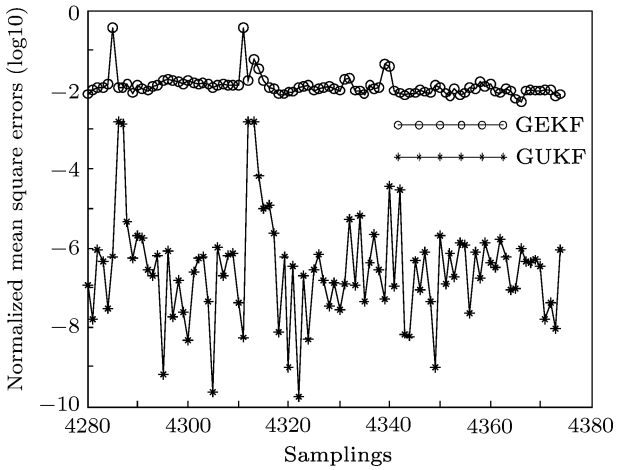


Fig. 4. Mean square errors (MSEs) of radioactivity time series prediction results.

5. Conclusions

Integrating the GEKF and GUKF with RBFNN correspondingly and without needing to pre-estimate

missing data, we directly use the GEKF-based RBFNN and GUKF-based RBFNN to predict the ground radioactivity time series with missing observed data by setting the independent Bernoulli random variables $\gamma_k = 0$ when the k-th datum is missed. From experimental results of our study, it follows that the prediction results of GUKF-based RBFNN accord well with the real ground radioactivity data while the prediction results of GEKF-based RBFNN deviate from those.

References

[1] Rani A and Singh S 2005 *Atmos. Environ.* **39** 6306

[2] Khan H M, Khan K, Atta M A and Jan F 1994 *J. Chem. Soc. Pakistan* **16** 183

[3] Lee C M and Ko C N 2009 *Neurocomputing* **73** 449

[4] Fábio A G and Leandro S C 2008 *Chaos, Solitons and Fractals* **35** 967

[5] Xiao Z, Ye S J, Zhong B and Sun C X 2009 *Expert Syst. Appl.* **36** 273

[6] Cai X D, Zhang N, Venayagamoorthy G K and Wunsch D C 2007 *Neurocomputing* **70** 2342

[7] Ma Q L, Peng H, Qin J W, Zheng Q L and Zhong T W 2008 *Chin. Phys.* B **17** 536

[8] Ding G, Li Y and Zhong S S 2008 *Chin. Phys.* B **17** 1998

[9] Quan T W, Liu X M and Liu Q 2010 *Appl. Soft Computing* **10** 562

[10] Wu Q 2010 *Expert Syst. Appl.* **37** 1776

[11] Wang W J, Men C Q and Lu W Z 2008 *Neurocomputing* **71** 550

[12] Chatfield C 2001 *Prediction Intervals* ed. Armstrong J *Principles of Forecasting: A Handbook for Researchers and Practitioners* (New York: Springer)

[13] Cao L J 2003 *Neurocomputing* **51** 321

[14] Majhi R, Panda G and Sahoo G 2009 *Expert Syst. Appl.* **36** 181

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

[16] van der Merwe R 2004 *Sigma-point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models* Ph. D. Thesis, Oregon Health & Science University p. 1

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

[18] Zhang B, Chen M Y and Zhou D H 2007 *Chaos, Solitons and Fractals* **32** 1491

[19] Kasahara Y, Pourahmadib M and Inoue A 2009 *Stat. Probabil. Lett.* **79** 1637

[20] Bermúdez J D, Vallet A C and Vercher E 2009 *J. Stat. Plan. Infer.* **139** 2791

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

[22] Julier S J and Uhlmann J K 2004 *P. IEEE* **92** 401