

# Time series analysis using RBF networks with FIR/IIR synapses

Iulian B. Ciocoiu\*

*Faculty of Electronics and Telecommunications, Technical University of Iasi P.O. Box 877, Iasi 6600, Romania*

Received 14 February 1997; accepted 25 March 1998

---

## Abstract

Radial basis functions networks (RBF) with dynamic synapses are introduced. The novelty aspect consists in replacing the standard scalar values of the output weights by discrete-time FIR/IIR filters. LMS-type learning algorithms are derived and simulation results for prediction of chaotic time series are reported. © 1998 Elsevier Science B.V. All rights reserved.

*Keywords:* Radial basis functions; Dynamic synapses; Time series analysis; Chaotic time series prediction

---

## 1. Introduction

In the last few years, many papers have addressed the problem of time series analysis by means of artificial neural networks. Purely feedforward [6,12], and local recurrent globally feedforward architectures [1,21] have been tested and prediction of chaotic signals was basically used as a benchmark. The present contribution is related to Radial Basis Functions networks (RBFs), which have been traditionally used as a multidimensional interpolation technique, implementing general mappings  $f: R^m \rightarrow R$  according to [5]:

$$f(\mathbf{X}) = \lambda_0 + \sum_{i=1}^M \lambda_i \phi(\|\mathbf{X} - \mathbf{C}_i\|), \quad (1)$$

where  $\phi$  is a nonlinear function selected from a set of typical ones,  $\|\cdot\|$  denotes the Euclidean norm,  $\lambda_i$  are the tap weights and  $\mathbf{C}_i \in R^m$  are called RBF centers. It is easy to

---

\* Corresponding author. E-mail: [iciocoiu@etc.tuiasi.ro](mailto:iciocoiu@etc.tuiasi.ro)

see that the formula above is equivalent to a special form of a 2-layer perceptron, which is *linear in the parameters* by fixing all the centers and nonlinearities in the hidden layer [15]. The output layer simply performs a linear combination of the (nonlinearly) transformed inputs and thus the tap weights  $\lambda_i$  can be obtained by using the standard LMS algorithm or its momentum version. This leads to a dramatic reduction of the computing time with the supplementary benefit of avoiding the problem of local minima, usually encountered when simulating standard multilayer perceptrons. Moreover, RBF networks also possess universal approximation capabilities, which are critically dependent on the choice of the centers. Typical approaches to this problem belong to one of the following categories [9]: (a) unsupervised clustering algorithms (e.g.,  $k$ -means clustering or Kohonen's self-organizing maps [11]), which offer computational efficiency and convergence speed; (b) supervised training algorithms (e.g., orthogonal least-squares [6] or Kalman filter [10]), which yield smaller errors.

RBFs have been used with significant success in many signal processing applications. Examples include nonlinear function approximation, channel equalization, system identification, to name just a few.

Recently, several recurrent RBF architectures have been proposed [2,14]. The first one uses lateral connections among the neurons from the hidden layer, whereas the second one includes a global feedback path from the output to the input of the system. Anyway, both circuits still perform a simple linear combination of the hidden units outputs.

In this paper we analyze the capabilities of a novel RBF network which is a generalization of that presented in [14], by replacing the scalar values

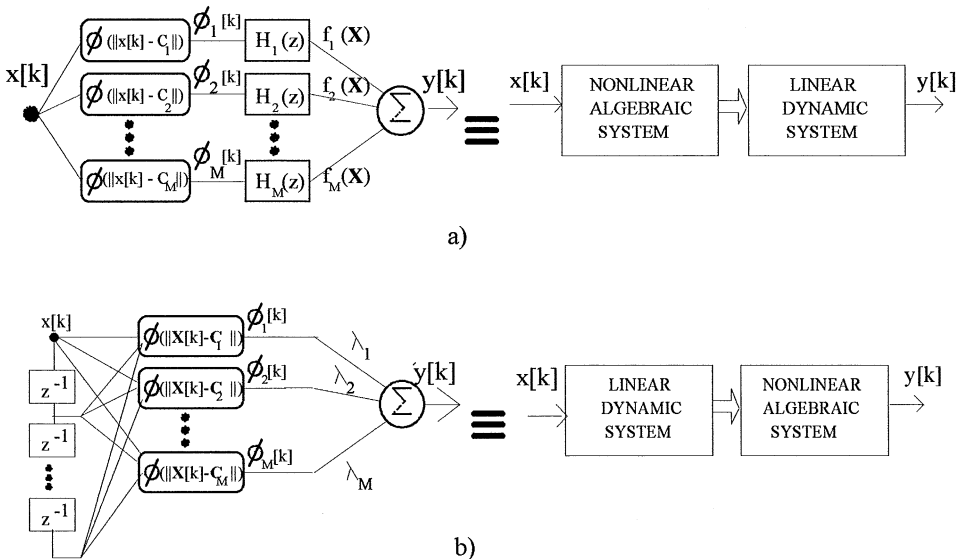


Fig. 1. RBF architectures for time series analysis: (a) the proposed architecture; (b) classical architecture.

of the output weights in the standard architecture by linear discrete time FIR/IIR filters. This idea was inspired by the remarkable performances of multilayer perceptrons with FIR/IIR synapses [1,21], and has been previously studied in the context of system identification [8]. Since RBF learning algorithms have clear advantages in terms of convergence speed and optimality of solution it is expected that such networks would deal with temporal processing problems without excessive computational burden.

In Fig. 1 we present the proposed solution and the standard RBF architecture used for time series analysis. When compared, it is easy to see that the new one performs a “ $z^{-1} - \phi$ ” exchange which enables the input to the hidden layer to become *unidimensional*, and thus greatly simplifying the implementation.

## 2. The learning algorithms

In the sequel we derive LMS-type learning algorithms for the proposed solution. We analyze separately RBF networks with FIR and IIR synapses, respectively. In both cases the system output at time  $k$  is given by:

$$y[k] = \sum_{i=1}^M f_i[k].$$

The instantaneous error at time  $k$  is

$$E[k] = \frac{1}{2} \{e[k]\}^2 = \frac{1}{2} \left\{ d[k] - \sum_{i=1}^M f_i[k] \right\}^2, \quad (2)$$

where  $d[k]$  is the desired output at time  $k$  and  $M$  is the number of centers to be used.

### 2.1. Case of FIR synapses

When FIR filters are used to model the output synapses they are described by different equations of the form

$$f_i[k] = a_{0i}\phi_i[k] + a_{1i}\phi_i[k-1] + \dots + a_{Ni}\phi_i[k-N]; \quad i = 1, \dots, M, \quad (3)$$

where  $a_{ni}$ ,  $n = 0, \dots, N$ ,  $i = 1, \dots, M$ , are real coefficients and  $N$  is the order of the filter (for simplicity all filters are considered of the same order).

The error is minimized by modifying the coefficients  $a_{ni}$  in a direction opposite to the instantaneous gradient of  $E[k]$ :

$$\begin{aligned} \Delta a_{ni}[k] &= -\eta \frac{\partial E[k]}{\partial a_{ni}[k]} = \eta e[k] \frac{\partial y[k]}{\partial a_{ni}[k]} = \eta e[k] \frac{\partial f_i[k]}{\partial a_{ni}[k]} \\ &= \eta e[k] \phi_i[k-n]; \quad n = 0, \dots, N, \quad i = 1, \dots, M, \end{aligned} \quad (4)$$

where  $\eta$  is the learning rate. The relation above is of LMS-type except for the fact that the error factor  $e[k]$  depends on the contribution of all filters and not only the current  $H_i(z)$  filter.

## 2.2. Case of IIR synapses

Adaptive IIR filters suffer from two important drawbacks, namely potential instability problems and the presence of local minima in the error function [18]. Even when the stability conditions are easily met, such as in the case of lattice or biquads, the computational complexity is higher than for FIR adaptive filters. This is why we decided to use a recently introduced IIR structure called *the gamma filter* [19], which imposes both trivial stability conditions and is computationally effective. It may be considered as a generalization of the classical transversal filter, the standard delay operator  $z^{-1}$  being replaced by the so-called *gamma operator*

$$G(z) = \frac{\mu}{z - (1 - \mu)}, \quad (5)$$

where  $\mu$  is a real parameter which controls the memory depth of the filter. As pointed out in [19], gamma filters are superior to standard FIR filters in terms of number of parameters required to model a given dynamics. The filter is stable if  $0 < \mu < 2$ . For  $\mu = 1$ ,  $G(z)$  reduces to the usual delay operator. Gamma neural networks [20] represent a generalization of the Time-Delay Neural Networks (TDNN), which were successfully used in time series prediction [21]. They have the advantage that the memory depth of the system can be adjusted on-line, making it more suitable for capturing the dynamics of the analyzed time series.

The transfer function  $H_i(z)$  may be written as

$$H_i(z) = \sum_{d=0}^D w_{id}[G_i(z)]^d; \quad i = 1, \dots, M, \quad (6)$$

where every  $G_i(z)$  function has its own  $\mu_i$  parameter and  $D$  is the maximum delay (for simplicity,  $D$  is the same for all filters). Equivalently, we may write

$$\begin{aligned} \phi_{id}[k] &= (1 - \mu_i)\phi_{id}[k - 1] + \mu_i\phi_{i(d-1)}[k - 1]; \\ d &= 1, \dots, D; \quad 1 = 1, \dots, M, \end{aligned} \quad (7)$$

where  $\phi_{i0}[k] = \phi_i[k]$ ,  $\phi_{iD}[k] = f_i[k]$  and  $\phi_{id}[k]$  represent (gamma) delayed versions of the incoming signals  $\phi_i[k]$ .

Following the derivation of the learning algorithm presented in [19] the weights  $w_{id}$  and coefficients  $\mu_i$  modify according to

$$\begin{aligned} \Delta w_{id}[k] &= \eta e[k]\phi_{id}[k]; \quad i = 1, \dots, M; \quad d = 0, \dots, D, \\ \Delta \mu_i &= \eta \sum_{d=0}^D e[k]w_{id}[k]\alpha_{id}[k]; \quad i = 1, \dots, M; \quad d = 0, \dots, D, \end{aligned} \quad (8)$$

where  $\alpha_{id}[k] = (\partial_{id}[k]/\partial\mu)$  verifies the recursive equation ( $\alpha_{i0}[k] = 0$ )

$$\alpha_{id}[k] = (1 - \mu)\alpha_{id}[k - 1] + \mu\alpha_{i(d-1)}[k - 1] + \phi_{i(d-1)}[k - 1] - \phi_{id}[k - 1];$$

$$i = 1, \dots, M; d = 1, \dots, D.$$

**Remark.** Since gamma filters are recursive, the RBF network with gamma synapses may be classified as a *local recurrent architecture*.

### 3. Simulation results

We have tested the efficiency of the proposed approach for chaotic time series analysis. A training sequence of 500 points was obtained by integrating the well-known Lorenz equations by Runge–Kutta method with step size 0.1. The data was split into 2 parts: 450 points were used for training and the remaining 50 for assessing the generalization capability of the network by cross-validation. The resulting signal has been used as input to the RBF networks and one-step ahead prediction has been performed. Intensive computer simulations have been carried out using both FIR and gamma synapses and different numbers of centers. In our study we have selected the centers using Kohonen’s self-organizing maps, since they are known to offer a proper image of the distribution density of the data, and because the algorithm is computationally efficient. We used Gaussian-type nonlinear functions of the form:  $\phi(x) = \exp(-x^2/\sigma^2)$ , where the spread parameter  $\sigma$  is typically selected by means of heuristic methods (in our simulations we used the  $k$ -nearest neighbors algorithm [11]). Anyway, as pointed out in [16], universal approximation is still possible by taking constant width.

The results of the simulations are presented in Fig. 2 for FIR synapses of order 5 and 10, and for various number of centers. It is easy to see that the neural network is able to closely approximate the original time series. In Fig. 3 the phase portraits of the true system and the approximating one are presented. The training phase was stopped when the error on the cross-validation set reached the minimum value, and the normalized mean square error (NMSE) for the training set is given in Table 1 (these values are obtained by averaging over 10 separate runs for each experiment).

The results show that the errors decrease when the number of centers and/or the order of the filters increase, but a number of 20 centers and synapses of order 10 are sufficient in order to obtain satisfactory performances. Gamma synapses seem better than FIR synapses, and they both outperform standard RBF networks.

In order to validate the model we may chose between two alternatives:

- (a) computing correlation functions between the prediction error (residual sequence) and several linear and non-linear combinations of past inputs and outputs and verify if all of them fall within the 95% confidence interval [3,7], since if the model structure is adequate the residuals should be uncorrelated with any combination of this kind. Another widely-used statistical method is represented by the so-called chi-squared tests [4];

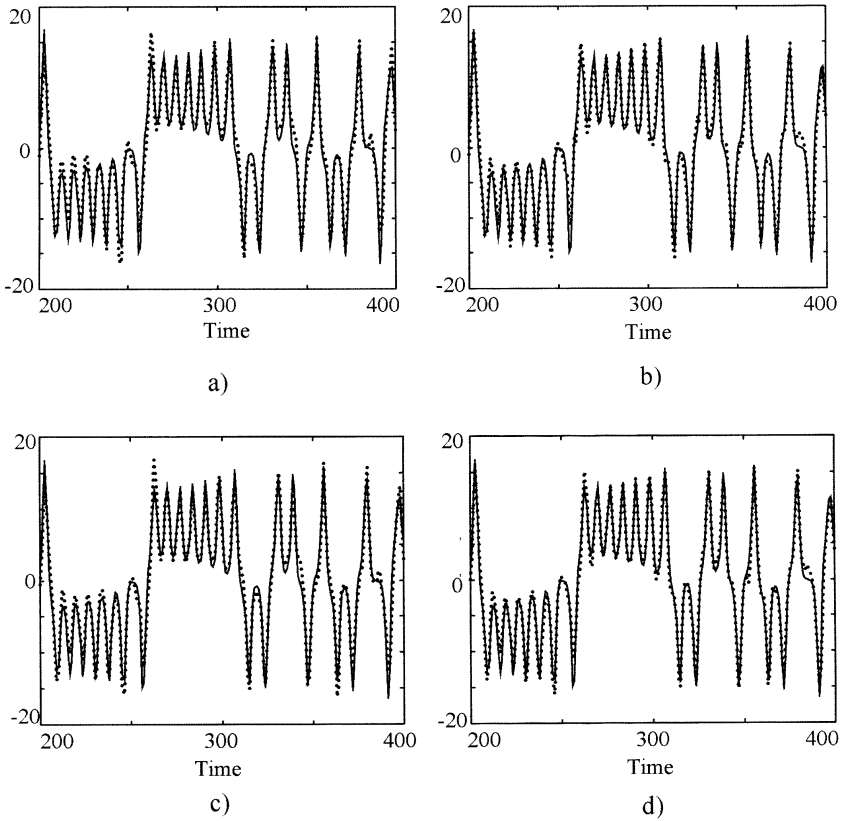


Fig. 2. Simulation results for chaotic time series prediction using FIR synapses: (a) order 5, 20 centres; (b) order 10, 20 centres; (c) order 5, 25 centres; (d) order 10, 25 centres. Solid line – original sequence; dotted line – approximating sequence.

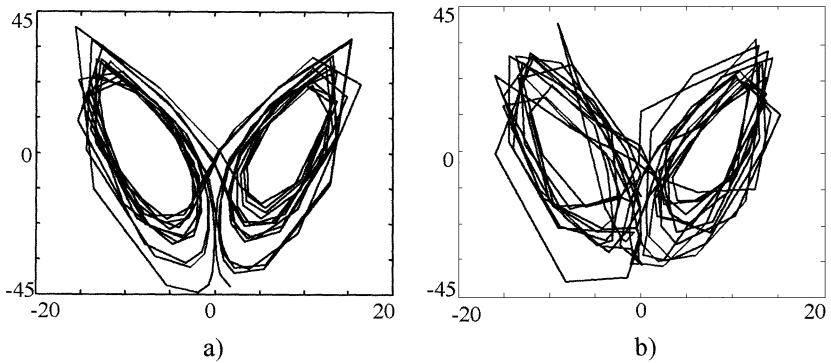


Fig. 3. Phase portraits of the: (a) original system; (b) approximating system.

Table 1

Final values of the normalized mean-squared error (NMSE) on the training set ( $\times 10^{-4}$ )

No. of centres	FIR synapses of order			Gamma synapses of order			RBF standard	
	5	10	15	5	10	15	5	10
5	2.286	1.917	1.991	2.286	1.917	1.991	17.33	19.54
10	0.848	0.66	0.442	0.774	0.663	0.442	11.8	21.02
15	0.591	0.47	0.368	0.591	0.442	0.368	10.69	21.02
20	0.585	0.42	0.331	0.591	0.402	0.365	9.95	20.28
25	0.555	0.4	0.295	0.545	0.368	0.331	9.95	20.28
30	0.516	0.25	0.258	0.442	0.331	0.33	9.95	21.02

(b) split the data into a training sequence and a testing one and performing *iterative prediction*, which means that the *predicted* value at time  $k$  is used as an input to predict the value at moment  $(k + 1)$ . We have used both approaches, although the first one is more suitable in this case, since the number of data points is not very large.

As in [3,7] we have considered the following correlation functions:  $\Phi_{ee}[k]$  ( $k > 0$ ),  $\Phi_{xe}[k]$  ( $\forall k$ ),  $\Phi_{e(ex)}[k]$  ( $k > 0$ ),  $\Phi_{x_2'e}[k]$  ( $\forall k$ ),  $\Phi_{x_2'e_2}[k]$  ( $\forall k$ ), where  $ex[k] = e[k + 1]x[k + 1]$ ,  $x^2[k] = x^2[k] - \text{avg}(x^2[k])$  (time average of  $x^2[k]$ ). The results are plotted in Fig. 4 and indicate that all the correlation functions fall practically within the 95% confidence interval.

Simulation results for iterative prediction are presented in Fig. 5. The network is clearly able to closely predict several values into the future, and the performances improve when the order of the synaptic filters is larger.

In Fig. 6 we give simulation results using gamma filters of order 5 and 10, respectively. We have also tested the case of common  $\mu$  for all synaptic filters, but the results were poorer than in previous cases, and the learning parameter had to be set to smaller values in order to obtain stable convergence.

#### 4. Conclusions

The aim of this contribution was to introduce a novel RBF architecture capable of dealing with temporal processing applications. Although the universal approximation capabilities of such networks need to be rigorously proved, they have advantages over both classical RBF networks as they use unidimensional input vectors and over the multilayer perceptrons with FIR/IIR synapses due to a simpler learning algorithm. The system could be generalized by still considering tapped delay lines at the input and dynamic synapses. Improved performances could be obtained by employing superior techniques for selecting the centers, such as those presented in [6,13,17].

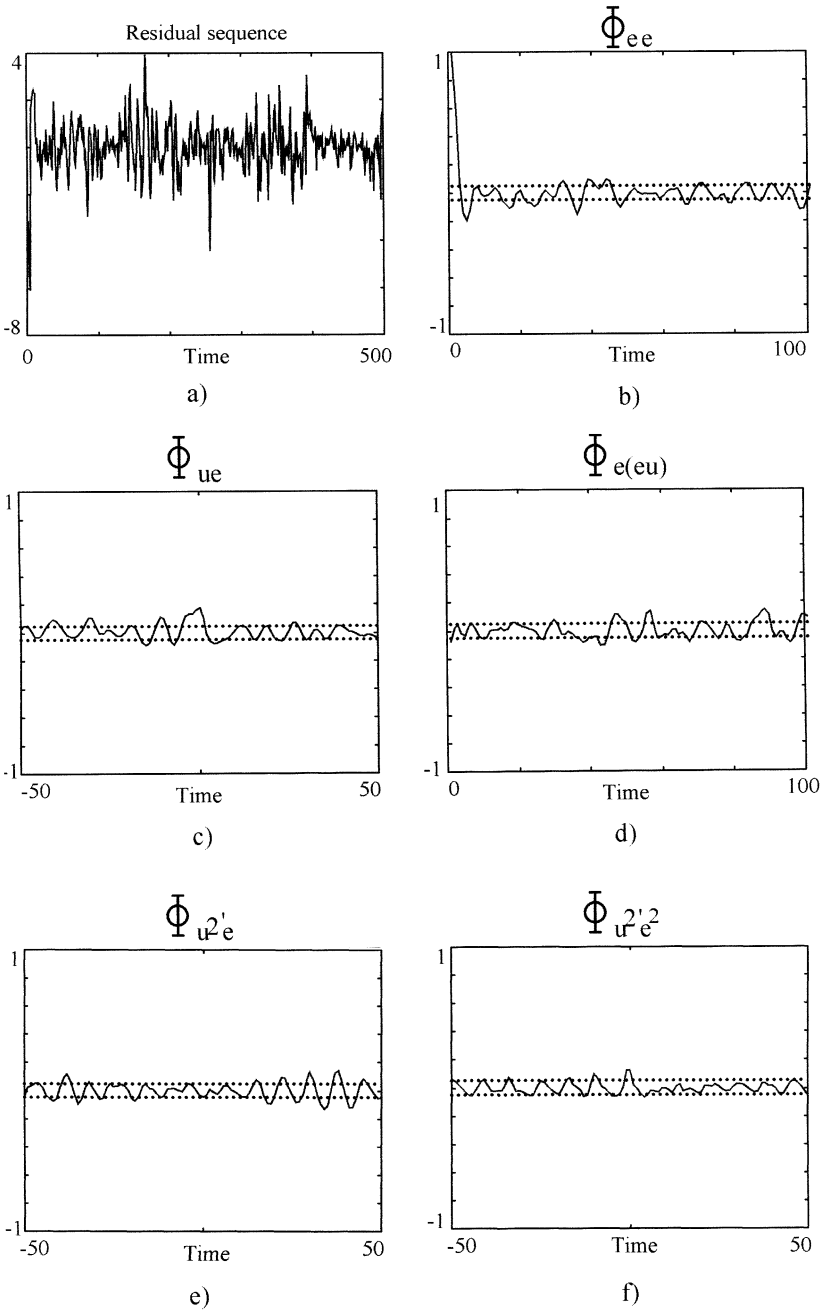


Fig. 4. Model validation statistical tests: (a) residual sequence  $e[k]$ ; (b)  $\Phi_{ee}[k]$ ; (c)  $\Phi_{ue}[k]$ ; (d)  $\Phi_{e(eu)}[k]$ ; (e)  $\Phi_{x_2^2'e}[k]$ ; (f)  $\Phi_{x_2^2'e_2}[k]$ . Dashed line – 95% confidence interval.



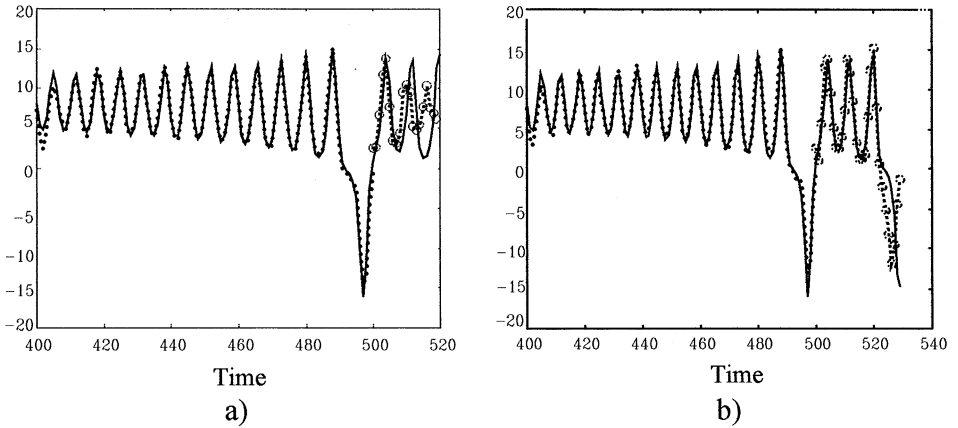


Fig. 5. Iterative prediction using FIR synapses: (a) order 10, 20 centres; (b) order 10, 25 centres.

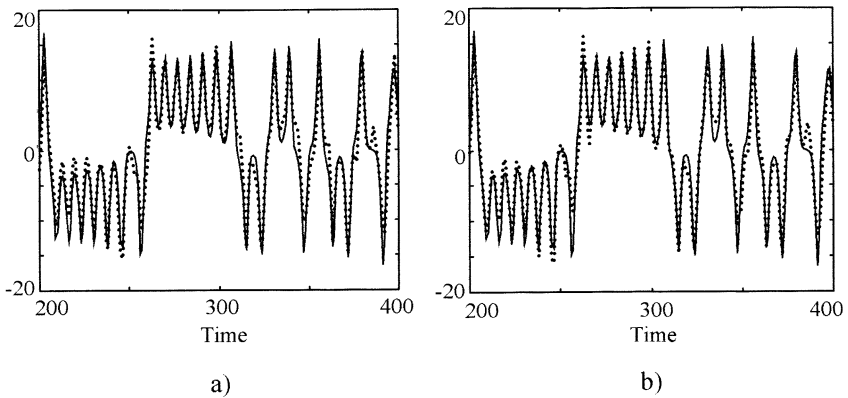


Fig. 6. Simulation results using gamma synapses: (a) order 5, 20 centres; (b) order 10, 20 centres. Solid line – original sequence; dotted line – approximating sequence.

## References

- [1] A.D. Back, A.C. Tsoi, FIR and IIR synapses, a new neural network architecture for time series modelling, *Neural Comput.* 3 (3) (1991) 375–385.
- [2] S.A. Billings, C.F. Fung, Recurrent radial basis function networks for adaptive noise cancellation, *Neural Networks* 8 (2) (1995) 273–290.
- [3] S.A. Billings, S. Chen, Extended model set, global data and threshold model identification of severely non-linear systems, *Internat. J. Control* 50 (5) (1989) 1897–1923.
- [4] T. Bohlin, Maximum-power validation of models without higher-order fitting, *Automatica* 14 (1978) 137–146.
- [5] D.S. Broomhead, D. Lowe, Multivariable functional interpolation and adaptive networks, *Complex Systems* 2 (1988) 321–355.

- [6] S. Chen, C.F.N. Cowan, P.M. Grant, Orthogonal least squares learning algorithm for radial basis function networks, *IEEE Trans. Neural Networks* 2 (2) (1991) 302–309.
- [7] S. Chen, S.A. Billings, C.F.N. Cowan, P.M. Grant, Non-linear systems identification using radial basis functions, *Internet J. Systems Sci.* 21 (12) (1990) 2513–2539.
- [8] I.B. Ciocoiu, RBF networks with FIR/IIR synapses, *Neural Process. Lett.* 3 (1) (1996) 17–22.
- [9] S. Haykin, A. Ukrainec, Neural networks for adaptive signal processing, in: N. Kalouptsidis, S. Theodoridi (Eds.), *Adaptive System Identification and Signal Processing Algorithms*, Prentice-Hall, New York, 1993.
- [10] S. Haykin, *Neural Networks – A Comprehensive Foundation*, IEEE Press, New York, 1994.
- [11] T. Kohonen, *Self-Organization and Associative memory*, Springer, New York, 1988.
- [12] A. Lapedes, R. Farber, *Nonlinear signal processing using neural networks: Prediction and system modelling*, LA-VR-87-2662, Los Alamos National Laboratory, 1987.
- [13] D. Lowe, Adaptive radial function nonlinearities, and the problem of generalization, *First IEE Int. Conf. on ANN*, 1989, pp. 171–175.
- [14] M.W. Mak, A learning algorithm for recurrent radial basis functions networks, *Neural Process. Lett.* 2 (1) (1995) 27–31.
- [15] J. Moody, C.J. Darken, Fast learning in networks of locally-tuned processing units, *Neural Comput.* 1 (1989) 281–294.
- [16] J. Park, I.W. Sandberg, Universal approximation using radial basis function networks, *Neural Comput.* 3 (1991) 246–257.
- [17] A. Saha, J. Keeler, Algorithms for better representation and faster learning in radial basis function networks, Technical Report ACT-NN-028-90, Austin, TX, 1990.
- [18] J.J. Shynk, Adaptive IIR filtering, *IEEE ASSP Mag.* (1989) 4–21.
- [19] J.C. de Vries, P.G. de Oliveira, The gamma-filter-a new class of adaptive IIR filters with restricted feedback, *IEEE Trans. Signal Process.* 41 (2) (1993) 649–656.
- [20] B. de Vries, J.C. Principe, The gamma model – a new neural model for temporal processing, *Neural Networks* 5 (4) (1992) 565–576.
- [21] E.A. Wan, Time series prediction by using a connectionist network with internal delay lines, in: A.S. Weigend, N.A. Gershenfeld (Eds.), *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading, MA, 1994, pp. 195–217.



**Iulian B. Ciocoiu** received a BS degree in Electronics in 1988 and a Ph.D. in Electrical Engineering in 1996, both from the Technical University of Iasi, Romania. He is an Associate Professor at the Faculty of Electronics and Telecommunications department of Signals, Circuits, and Systems at the Technical University of Iasi. His research interests include neural networks learning algorithms, complex dynamics in electronic systems, and time series prediction.