#-120 ,

Radial Basis Function Neural Networks in Tracking and Extraction of Stochastic Process in Forestry

PERO J. RADONJA
SE Serbiaforest, Institute of forestry,
Kneza Viseslava 3, 11030 Belgrade
YUGOSLAVIA
E-mail: inszasum@EUnet.yu

ABSTRACT - The performances of classical approximation methods and radial-basis function (RBF) neural networks in ***tracking*** *of the height density data are presented in the first part of the paper. The application of the different classical approximation methods in* ***extraction*** *of unknown biological process from real measured data is considered in the second part of the paper. The advantages of implementation of RBF neural networks in extraction unknown process are analyzed and illustrated with corresponding example.*

Key Words: - Neural networks, radial-basis function networks, curve fitting, tracking and extraction.

1 Introduction

The process of plant growth, typical process in forestry is stochastic long time process. The relation between height of trees and diameter of the trees is very important in the forestry. The curve that represents the mentioned relation is known as the **height curve** and will be considered in this paper. In the elementary cases of the extraction of unknown relation or process is simple a fitting of measured data. However the most frequently the process of extraction includes the many boundaries related with number and position of stationer points, the slopes of asymptotes etc. In other words, the exact shape (form) of the relation (process) is not known but we known its nature. In considered case we known that this function is a nonconstant, bounded and monotone-increasing continuous function.

In classical approach to the approximation of the height curve we have an error of approximation (tracking) even that we have many smoothed data. This is because of the fact that in the classical approach we have a few parameters for optimization.

The application of neural networks, as it will be shown in this paper, gives much better results. The neural networks may be treated as a practical tool for performing a nonlinear input-output mapping of a general nature [1]. Such a viewpoint then permits us to look on generalization as the effect of a good nonlinear interpolation of the input data. In our case the learning process, i.e., training of a neural network may be viewed as a "curve-fitting" problem.

2 Classical Approximation Methods in Tracking of Height Density Data

In this section performances of tracking of the real biological process using the classical approximations methods will be analyzed. Performances of tracking of the height curve will be considered on the height density data. We can say that we consider practically smoothed data. The classical approximation methods introduce the error because the functions for approximations are relatively simple and have only 2-3 parameters for optimization.

In Fig.1 the measured data are denoted by (*). The result of classical Henriks approximation method, [2] is marked by (+) and as we can see the rather great error of tracking (approximation) is introduced.

0-7803-5512-1/99/$10.00 © 2000 IEEE

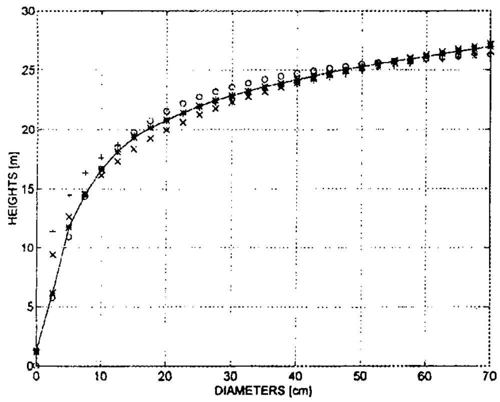

Fig.1 The classical approximation methods in tracking of height density data

The better result gives the Terezaki-Mihajlov method of approximation, [2], data marked by (x), and the best result gives the Noslund method of approximation, [2], data marked by (o). The mean squared tracking (approximation) errors for the mentioned methods are presented in **Table 1.**

Methods of approximation	Mean squared approximation error [m]
Noslund	0.574
Terezaki-Mihajlov	0.831
Henriksen (2)	1.524

Table 1

3 Performances of Neural Networks in Tracking of Height Density Data

It is known that a single hidden layer is sufficient to uniformly approximate any continuous function. It is seen that in our case we can use radial basis networks for successful nonlinear input-output mapping.

We shall consider the design of neural network as an approximation (curve fitting) problem in a high dimensional space. In fact we are trying to find a surface in a multidimensional space that provides a best fit to the training data. In other to show the performances of neural networks in process of tracking we shall consider the same data which we have analyzed in the previous section

The first layer, i.e., the input layer is made up of sensory units. The second layer involves the hidden units that provide a set of radial-basis functions. The transformation from the input space to the hidden-unit space is nonlinear. The transformation from the hidden-unit space to the output space is linear.

Evidently that application of two layers neural networks, [3], is sufficient in our case. The most important design parameters are a sum-squared error goal and a spread constant of radial basis functions. In our case, for the considered data the experiments show that value of the spread constant cane changes in great boundaries. For example if we adopt for the error goal value of 0.02 and for the spread constant of 30.0 we shall get the **mean** squared approximation error, **MSAE**, of 1.58cm, Fig.2. Similarly if value of the spread constant is only 8.0, the **MSAE** is 1.37cm.

We have got a well-behaved, i.e., relatively smooth error surface for the considered problem. Indeed the final mean squared tracking (approximation) error of each approximation procedure (procedures for spread constant of 30.0 and 8.0) does not vary significantly.

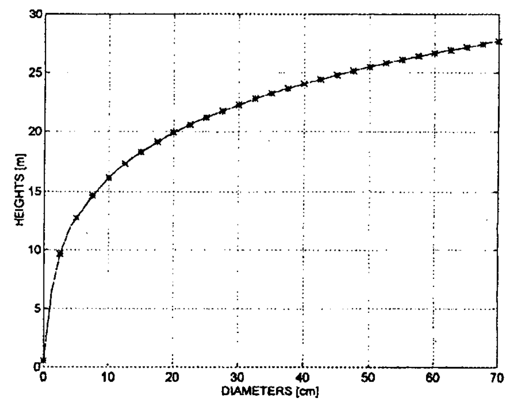

Fig.2. Application of neural networks in tracking of height density data

As we can see the approximation error is more of 30 times less than the best result of the classical approximation method. Noslund approximation has error of 57.4cm, **Table 1.** We can say that we have got the ideal tracking or smoothing of the real data. This is obviously consequence of the more complex process of approximation by radial basis neural networks.

The record of training errors i.e. a process of adaptation (training) is presented in Fig.3. We can see that the adaptation of network is performed for some more than 20 epochs. The process of the adaptation is convergent without using any techniques like these based on additive noise or random weights, [4] and [5].

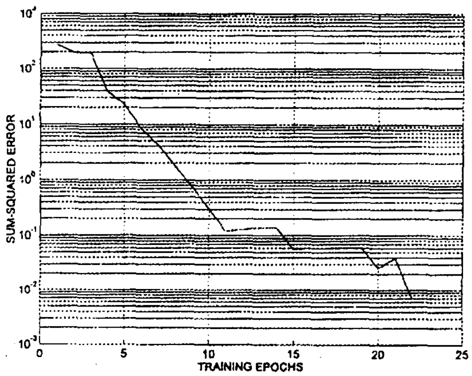

Fig.3. The adaptation of network

The best manner to demonstrate the performances of tracking is the value of MSAE. However, we can consider the all errors of tracking separately, also. These errors are very small as we can see on Fig.4. The biggest error is about 4cm.

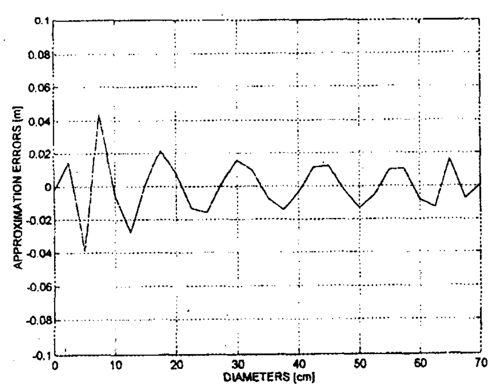

Fig.4. The errors of tracking

In order to see the effect of any parameter in obtained approximation we have to adopt the big error goal. Assume that the error goal is 1000 times bigger than the error in the case chown in Fig.2.

When spread constant, sc, has value of 330 the obtained result is shown in Fig.5

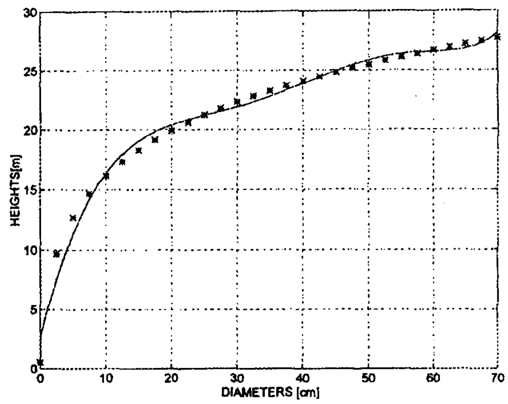

Fig.5 The efficiency of tracking for sc=330 and e_g=20

Similarly, for the same value of the error goal and very small value of spread constant, sc=3, the efficiency of tracking is illustrated by Fig.6.

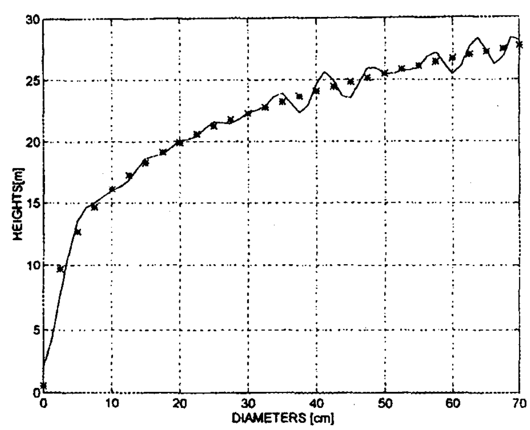

Fig.6 The efficiency of tracking for sc=3 and e_g=20

4 The Classical Approximation Methods in Process of Extraction

Generally the process of extraction of an unknown process is more seriously than the problem of tracking from preceding sections. In this section we shall consider the one realistic problem of extraction. The available data, data denoted by (*), are presented in Fig. 7. As we can see these data are with great error and much rarely than data from the preceding section.

As was mentioned in section 1 a number of approximation methods (techniques) have been developed and employed for this and similar purpose. We shall analyse the application of Noslund, Terezaki-Mihajlov and Henriksen approximation methods. In Fig.7 only the results of the best and the worst classical approximation methods are presented.

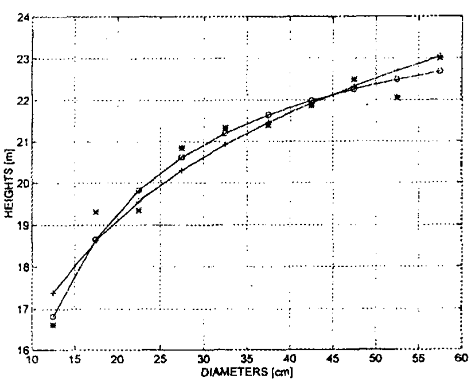

Fig.7 The classical approximations in extraction of unknown process from real measured data

In Fig.7 application of Noslund and Henriksen approximation methods are denoted by (o) and (+), respectively. The values of **MSAE** for 3 classical approximation methods are presented in **Table 2.**

Methods of approximation	MSAE [cm]
Noslund	38.54
Terezaki-Mihajlov	40.75
Henriksen (2)	50.45

Table 2

5 Extraction of Unknown Process by RBF Neural Networks

It is known that radial neurons with narrow transfer function, that is, with the small spread constant, are good for steep part of curve. On the other hand the neurons with the large spread constant are good for relatively a flat part of curve, **Appendix A**. Therefore in our case the best results are obtained if we use two neural networks with corresponding general weighting coefficients GWC_1 and GWC_2. The values of general weighting coefficients show the influence of each neural network on final result separately.

The two-stage approximation process is similar to the spline technique for curve fitting, [6], in the sense that the effects of neurons are isolated and the approximations in different regions of the input space may be individually adjusted.

Assume that the sum-squared error goals for two neural networks are 2.0 and 3.0 respectively. The spread constants of radial basis functions are 44.0 and 66.0 respectively, also.

The obtained approximation, i.e. result of extraction of unknown biological process, is shown in Fig.8.

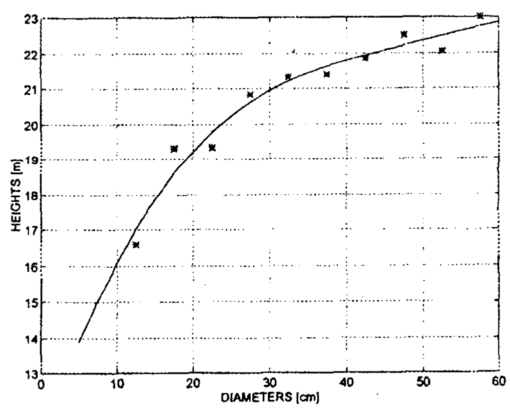

Fig.8 The two neural networks in extraction of unknown process from real measured data

We can see that the obtained curve has very acceptable shape. As it is known the height curve is a nonconstant, bounded and monotone-increasing continuous function.

The process of adaptation is very convergent; the adaptation of network is performed for only two epochs, **Appendix B**. It can be seen that using the radial basis network we get the superfast function approximation.

The errors of extraction (approximation) versus the diameters are presented in Fig.9. The comparison with the results from section 3, Fig.4, shows that the maximal value of the extraction errors is nearly 20 times greater than the maximal value of errors of tracking.

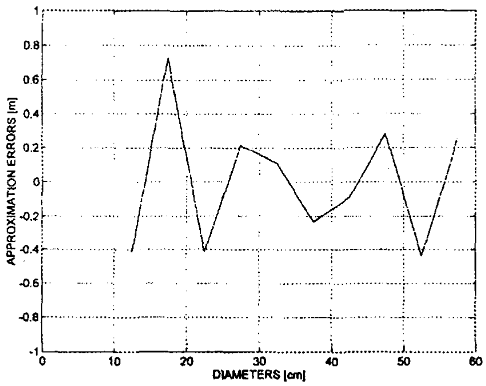

Fig.9 The error of extraction

It can be shown that the best result, i.e., the minimal value of MSAE, is obtained when general weighting coefficients GWC_1 and GWC_2, have value of 0.75 and 0.25, respectively. The obtained result, 35.83cm, is better than the result that we have got using the best classical method. Noslund approximation method gives error of 38.54cm, **Table2.**

6 Conclusion

The implementation of radial-basis function neural networks, in problem of **tracking** of the height density measured data, ensures the significant advantages. The mean squared error of tracking is more of 30 times less than the best result of the classical approximation methods. This is obviously consequence of the more complex process of approximation by neural networks. In the other hand, in the classical approach, the relatively big error of tracking is introduced always because the functions for approximation are relatively simple and have only 2-3 parameters for optimization.

The **extraction** of an unknown process from a noisy measured data is more realistic problem and the very frequently a requirement in many engineering disciplines. The available data are usually with big error and the process of extraction includes the many boundaries related with number and position of stationer points, the slope of asymptotes etc.

The advantages of application the neural networks in process of extraction of the unknown biological process are shown by illustrative example. The two radial basis networks are used in the two-stage approximation process. This approach permit to us to use the radial basis functions with the different spread constants. In this way the effects of neurons are isolated and the (approximation) extraction in different regions of the input space is individually adjusted. The obtained result is significantly better than the best result of classical approximation method, as it was shown in the considered case.

Appendix A

The obtained approximation of the height curve when the sum-squared error goal has value of 2.0 and the spread constant of radial-basis function network has value of 44, is shown in Fig.10.

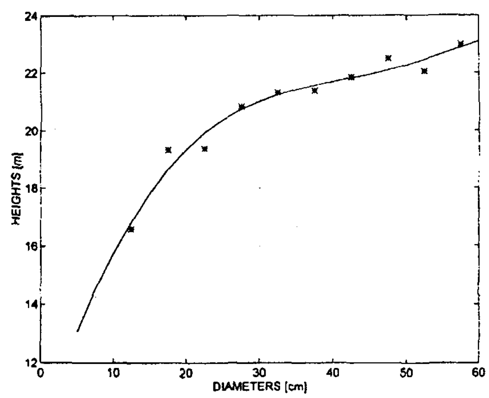

Fig.10 Approximation of the height curve by RBF network with e_g=2.0 and sc=44

It is seen that the obtained shape of the curve is good for small value of diameters. The good slope of the curve is ensured by the relatively small value of spread constant, 44. However the behavior of the curve is not good for the larger values of diameters.

The application of neural network with bigger spread constant, sc= 66, ensures the approximation of the height curve like is presented in Fig.11.

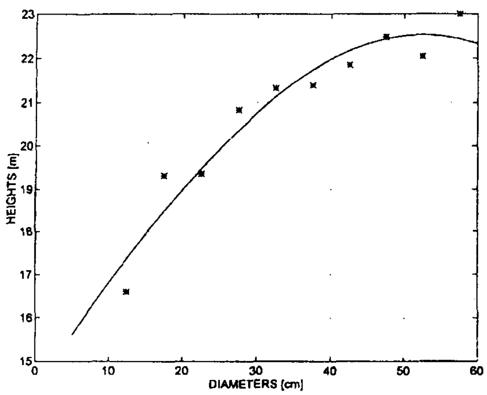

Fig.11 Approximation of the height curve by RBF network with e_g=3.0 and sc=66

The magnitude of the slope of the curve for the small value of diameters is not enough. However, the shape of the curve for the large value of diameters ensures the compensation of the bad behavior of the curve presented in Fig.10 for these values of the diameters.

Appendix B

The training (adaptation) errors versus training epochs are presented in Fig.12. It can bee seen that the training (sum-squared) error goal, e_g=2, is reached for only 2 epochs.

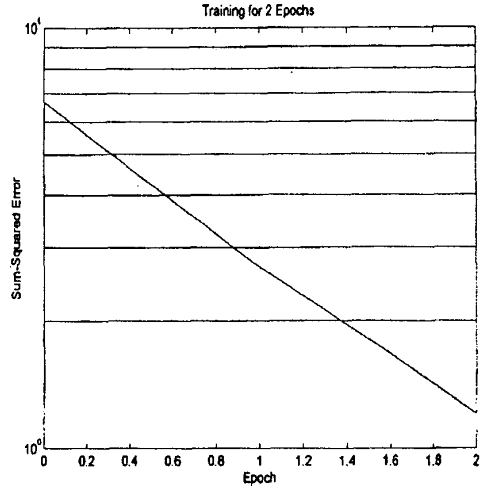

Fig.12 The training (adaptation) errors versus training epochs

References:

[1] Haykin, S. (1994): *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Company, New York.

[2] Vasiljevic, A. (1991): *Program package:* **KOR.**

[3] Neural Network Toolbox, (1994), Version 2.0a, 06-Apr-94, MATLAB for Windows 4.2c.1.

[4] Holmstrom, L. and Koistinem P. (1992): Using Additive Noise in Back-Propagation Training, *IEEE Transaction on Neural Networks,* jan.1992, Vol.3, No.1

[5] Bartlett L. Peter and Tom Downs, (1992): Using Random Weights to Train Multilayer Networks of Hard-Limiting Units, *IEEE Transaction on Neural Networks,* March 1992,Vol.3, No.2.

[6] Spline Function Toolbox, (1994), Version 2.0a, 06-Apr-94, MATLAB for Windows 4.2c.1.