UDC 007.52:517.938

Krunoslav HORVAT
Ines ŠOIĆ
Ognjen KULJAČA

Adaptive Neural Network Controller for Thermogenerator Angular Velocity Stabilization System

Authors' address (Adresa autora):
Brodarski Institute, Zagreb; e-mail: krunoslav.horvat@hrbi.hr, ines.soic@hrbi.hr, ognjen.kuljaca@hrbi.hr

Received (Primljeno): 2013-04-15
Accepted (Prihvaćeno): 2013-05-14
Open for discussion (Otvoreno za raspravu): 2014-06-30

Original scientific paper

The paper presents an analytical and simulation approach for the selection of activation functions for the class of neural network controllers for ship's thermogenerator angular velocity stabilization system. Such systems can be found in many ships. A Lyapunov-like stability analysis is performed in order to obtain a weight update law. A number of simulations were performed to find the best activation function using integral error criteria and statistical T-tests.

Keywords: *activation function, adaptive neural network, tracking problem*

Adaptivni regulator s neuronskim mrežama za stabilizaciju kutne brzine termogeneratora

Izvorni znanstveni rad

U ovom članku opisan je analitičko-simulacijski pristup izboru aktivacijskih funkcija za klasu regulatora s neuronskim mrežama koji služe za praćenje referentne vrijednosti kutne brzine kod turbogeneratora. Kako bi se odredilo pravilo kojim će se obnavljati težine matrica tj. obavljati učenje mreže, potrebno je bilo provesti analizu Ljapunove stabilnosti. Provedeno je mnogo simulacija uz različite aktivacijske funkcije, a najbolja aktivacijska funkcija određena je pomoću integralnih kriterija, te statističkoga T-testa.

Ključne riječi: *adaptivna neuronska mreža, aktivacijska funkcija, problem praćenja*

1 Introduction

The paper presents a novel procedure for designing adaptive neural network (ANN) controller for ship's thermogenerator angular velocity stabilization system. All of thermogenerator system parameters can change in practice, thus the need for adaptive control. If all parameters change it is hard to use conventional adaptive control methods that identify process parameters. Instead, we use ANN that due to universal approximation property identifies the whole process under control rather than process parameters. The use of ANN is illustrated on a simplified angular velocity stabilization system for thermogenerator.

Other non-adaptive methods for thermogenerator angular velocity stabilization system ([1], [2] and many others) are presented. There are also neural network (NN) control algorithms which show satisfactory performance for similar systems, but they require off-line training, which often cannot be done on real systems ([3], [4], [5] and others). The off-line training problem will be removed by adaptive neural network.

Adaptive neural network given here does not require a priori training, but is capable of on-line training [6]. The controller described here represents an advanced and performance-enhanced version of NN control scheme given in [7]. A numbers of simulations were performed to find the best activation function which is chosen using integral error criteria and statistical T-tests as described in [8].

The paper is organized as follows. Mathematical preliminaries are given in Section 2. It is followed by background on neural networks in Section 3. Section 4 presents a simplified model of ship's thermogenerator angular velocity system, and Section 5 describes the neural network control scheme. Section 6 provides an example of simulation-based design. Finally, conclusions are given in Section 7.

2 Mathematical preliminaries

Let R be the real numbers, R^n the real n vectors and R^{mxn} the real mxn matrices. For a matrix $A = \lfloor a_{ij} \rfloor,\ A \in R^{nxm}$, the Frobenius norm is defined as

$$\|A\|_F^2 = tr(A^T A) = \sum_{i,j} a_{ij}^{\ 2} \qquad (1)$$

where *tr()* is trace operation. The associated inner product is $\langle A, B \rangle_F = tr(A^T B)$. The Frobenius norm is denoted by $\|\bullet\|$ throughout this paper, unless otherwise specified. The trace of matrix $A = \lfloor a_{ij} \rfloor$ satisfies $tr(A) = tr(A^T)$. For any *mxn* or *nxm* matrix *C*, we have $tr(BC) = tr(CB)$.

To prove stability we used proposition which states that a system is uniformly ultimately bounded (UUB) if it has a Lyapunov function whose time derivation is negative in an annulus of certain width around the origin [9].

3 Neural network structure

A mathematical model of an NN ([6]) is shown in Figure 1. This NN has two layers of adjustable weights and is known as a two layer NN. The values x_k are the NN inputs and y_m its outputs. Function σ(.) is a nonlinear activation function contained in the hidden layer of the NN. The hidden-layer weights are v_{jk} and the output-layer weights are w_{ij}. The hidden-layer thresholds are θ_{vj} and the output-layer thresholds are θ_{wi}. The number of hidden-layer neurons is L.

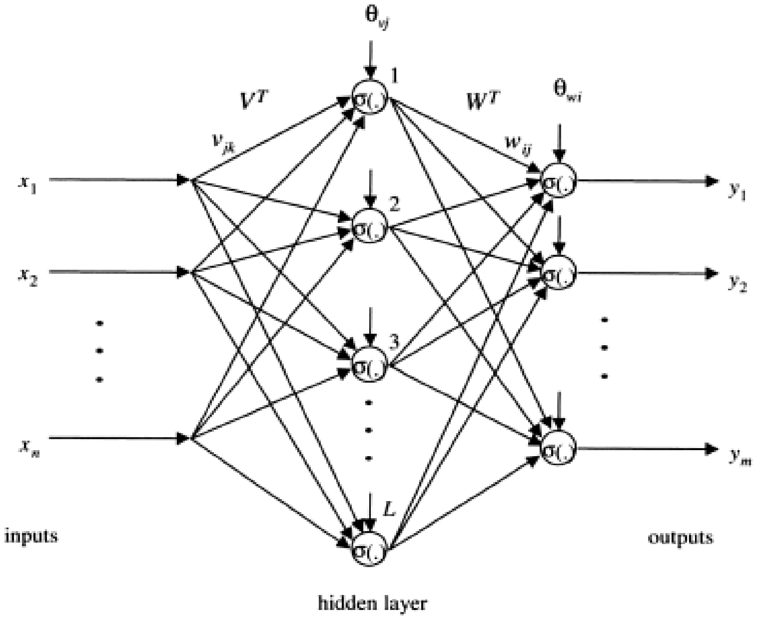

Figure 1 **Two-layer neural network**
Slika 1 **Dvoslojna neuronska mreža**

Net output is given by

$$y = W^T \sigma(V^T x) \qquad (2)$$

with

$$V^T = \begin{bmatrix} \theta_{v1} & v_{11} & \cdots & v_{1n} \\ \theta_{v1} & v_{21} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{v1} & v_{L1} & \cdots & v_{Ln} \end{bmatrix} \qquad (3)$$

and

$$W^T = \begin{bmatrix} \theta_{w1} & w_{11} & \cdots & w_{1L} \\ \theta_{w2} & w_{21} & \cdots & w_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{w1} & w_{m1} & \cdots & w_{mL} \end{bmatrix} \qquad (4)$$

where the output vector is $y = [y_1 \quad y_2 \quad \cdots \quad y_m]^T$. Due to the fact that the thresholds appear as the first columns of the weights matrices, we must define the input vector augmented by „1“ as $x = [1 \quad x_1 \quad x_2 \quad \ldots \quad x_n]^T$. Also, the σ(.) is the augmented hidden-layer function vector, defined for a vector $w = [w_1 \quad w_2 \quad \cdots \quad w_L]^T$ as $\sigma = [1 \quad \sigma(w_1) \quad \cdots \quad \sigma(w_L)]^T$.

The main property of NNs we are concerned with for control is the function approximation property. Let *f(x)* be smooth function from $R^n \rightarrow R^m$. Then it can be shown that if the activation functions are suitably selected, as long as *x* is restricted to a compact set $S \in R^n$, then for some sufficiently large number of hidden-layer neurons *L*, there exist weights and thresholds such as

$$f(x) = W^T \sigma(V^T x) + \varepsilon(x) \qquad (5)$$

The value of $\varepsilon(x)$ is called the neural network functional approximation error and we can always find a number ε_N such that $\varepsilon(x) \leq \varepsilon_N$ for all $x \in S$.

The computing power of the NN comes from the facts that the activation functions are nonlinear and that weight matrices (*W* and *V*) can be tuned through some learning procedure.

It has been shown that if the first-layer weights *V* are fixed, then the approximation property can be satisfied by tuning the output weights *W*. If we use this property, NN output becomes $y = W^T \phi(x)$. For this to occur $\phi(x) = \sigma(V^T x)$ must be a basis [6].

For the activation functions, it is common to use the Gaussian activation function, sigmoidal functions and hyperbolic tangents.

4 Angular velocity system for thermogenerator

A model of angular velocity system is shown in the figure below.

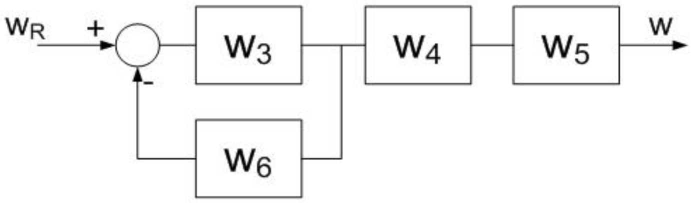

Figure 2 **Model of angular velocity system**
Slika 2 **Matematički model sustava kutne brzine**

The transfer functions are given as:

$$W_3 = \frac{k_3}{s} \tag{6}$$

$$W_4 = \frac{k_4}{T_3 s + 1} \tag{7}$$

$$W_5 = \frac{k_5}{T_4 s + 1} \tag{8}$$

$$W_6 = k_6 \tag{9}$$

where W_3 is valve positioning servomotor, W_4 and W_5 are turbines and W_6 is stabilization feedback of the valve positioning servomotor. Model input is angular velocity reference value.

This system is linear and the need for adaptive control or use of the function approximation property of the neural network is not obvious. Because all the system parameters can and do change during the operation, it is understandable that adaptive control scheme would perform better than non-adaptive control.

The system shown in Figure 2 can be written in state-space form as

$$\begin{bmatrix} \dot{y}_P \\ \dot{y}_T \\ \dot{w} \end{bmatrix} = \begin{bmatrix} -k_3 k_6 & 0 & 0 \\ \frac{k_4}{T_3} & \frac{-1}{T_3} & 0 \\ 0 & \frac{k_5}{T_4} & \frac{-1}{T_4} \end{bmatrix} \cdot \begin{bmatrix} y_P \\ y_T \\ w \end{bmatrix} + \begin{bmatrix} k_3 \\ 0 \\ 0 \end{bmatrix} \cdot x_R \tag{10}$$

$$\dot{x} = Ax + Bu$$

$$y = C \cdot \begin{bmatrix} y_P \\ y_T \\ w \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_P \\ y_T \\ w \end{bmatrix} \qquad (11)$$

where y_p is the output from the valve positioning servomotor and y_t is the output from the first turbine.

5 Adaptive neural network control

We use the neural network shown in Figure 1. If the first layer weights are initialized randomly and then fixed to form a basis $\rho(x)$, the NN output (2) becomes

$$y = W^T \rho(x) \qquad (12)$$

In the system we have a reference value so here we have to solve a tracking problem. It is assumed that a reference input is bounded and that this will be true as long as this system is in normal mode of operation. In Figure 3 an NN control scheme is presented.

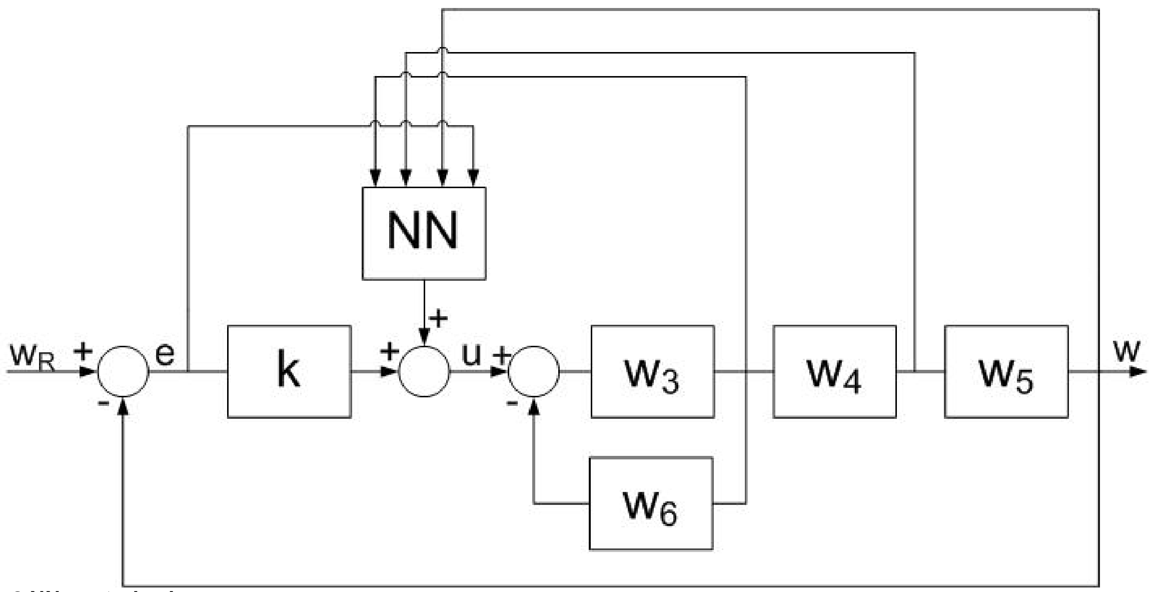

Figure 3 **NN control scheme**
Slika 3 **NN upravljački model**

Control signal is given by

$$u = ke + W^T \rho(x) \qquad (13)$$

where $e = w_R - w$, and the weights updates are provided by

$$\dot{W} = F\rho(x)w - k_w \|x\| FW \qquad (14)$$

with F any symmetric and positive definite matrix and k_w positive design parameter. Weight matrix W and vector x are ultimately uniformly bounded (UUB) and the system will be stable in Lyapunov sense as long as

$$\|W\| > \frac{D}{2k_w} + \sqrt{\frac{D^2}{4k_w^{\ 2}} + \frac{\sigma_{\max}(P)}{k_w}} \qquad (15)$$

or

$$\|x\| > \frac{\frac{D^2}{4k_w} + \sigma_{\max}(P)d_R}{\frac{1}{2}\sigma_{\min}(Q)} \tag{16}$$

where $x = [e \quad y_P \quad y_T \quad w]^T$. Terms D, P and Q are defined in subsequent stability proof.

Proof follows:

First we can rewrite state-space form by introducing (13) into (10)

$$\dot{x} = Ax - BCkx + Bkw_R + BW^T\rho(x) \tag{17}$$

Then the following substitutions are introduced: $d_R = Bkw_R$, $d_K = BCk$ and $A' = A - d_K$. Equation (17) now becomes:

$$\dot{x} = A'x + d_R + BW^T\rho(x) \tag{18}$$

Now we have to define the Lyapunov candidate:

$$L = \frac{1}{2}x^T Px + \frac{1}{2}W^T F^{-1}W \tag{19}$$

with P a diagonal and positive definite matrix. In this design W is a vector because we have only one output. The Lyapunov derivate is:

$$\dot{L} = \frac{1}{2}\dot{x}^T Px + \frac{1}{2}x^T P\dot{x} + W^T F^{-1}\dot{W} \tag{20}$$

By introducing (18) into (20) we obtain

$$\dot{L} = \frac{1}{2}\left[\rho(x)^T WB^T + d_R{}^T + x^T A'^T\right]Px + \frac{1}{2}x^T P\left[A'x + d_R + BW^T\rho(x)\right] + W^T F^{-1}\dot{W} \tag{21}$$

$$\dot{L} = \frac{1}{2}\rho(x)^T WB^T Px + \frac{1}{2}d_R{}^T Px + \frac{1}{2}x^T A'^T Px + \frac{1}{2}x^T PA'x + \frac{1}{2}x^T Pd_R + \frac{1}{2}x^T PBW^T\rho(x) + W^T F^{-1}\dot{W}$$

$$\dot{L} = x^T PBW^T\rho(x) + x^T Pd_R + \frac{1}{2}x^T\left(A'P + PA'\right)x + W^T F^{-1}\dot{W} \tag{22}$$

The third term in (22) is a well known Lyapunov function for linear system ($-Q = A'^T P + PA'$) and after introducing (14) into (22) equation becomes

$$\dot{L} = x^T PBW^T\rho(x) + x^T Pd_R - \frac{1}{2}x^T Qx + W^T\rho(x)w - W^T k_w\|x\|W \tag{23}$$

$$\dot{L} = W^T\rho(x)\left[x^T PB + w\right] + x^T Pd_R - \frac{1}{2}x^T Qx - W^T k_w\|x\|W \tag{24}$$

Let us define $D = \max(\|\rho(x)\|(\|PB\| + 1))$. Activation functions are bounded so we can replace $\rho(x)$ with $\|\rho(x)\|$. Also, we can define $\sigma_{\min}(Q)$ as the minimum singular value of matrix Q and $\sigma_{\max}(P)$ as the maximum singular value of matrix P. After introducing norms we obtain:

$$\dot{L} = \|x\|\|W\|D + \|x\|\sigma_{\max}(P)d_R - \frac{1}{2}\|x\|^2\sigma_{\min}(Q) - \|W\|^2 k_w\|x\| \tag{25}$$

We can rewrite that equation as

$$\dot{L} = -\|x\|\left\{-\|W\|D - \sigma_{\max}(P)d_R + \frac{1}{2}\|x\|\sigma_{\min}(Q) + \|W\|^2 k_w\right\} \tag{26}$$

The Lyapunov derivative is negative as long as the term in parentheses in (26) is positive. The term in parentheses is positive as long as (15) and (16) hold, thus the system is UUB.

6 Simulation based design example

Simulations to determine the best activation function were performed with the following parameters and signals:
$k_3 = 2, k_4 = 0.8, k_5 = 1, k_6 = 0.4, T_3 = 5s, T_4 = 0.3s, F = diag(0.008), k_w = 9.2$,
$w_R = 0.1\sin(0.1t)pu$. Weight matrix V is initialized as random numbers between -0.6 and 0.6 divided by 4 and weight matrix W is initialized as random numbers between -0.6 and 0.6. Number of hidden layers is 6 [8]. Now we can tell that matrix $\rho(x)$ has dimensions 7x1. Simulations were performed for given (original) plant parameters as well as for the plant parameters increased and decreased by 10%. Simulation time was 1000s. Recorded performance indices were Integral of Absolute Error (IAE)

$$J_{IAE} = \int_0^\infty |e(t)| dt \tag{27}$$

and Integral of Squared Error (ISE)

$$J_{IAE} = \int_0^\infty e(t)^2 dt \tag{28}$$

Because the weight matrices were initialized randomly, it was necessary to perform a number of simulations in order to obtain samples with different IAE and ISE values that allow us to compare the mean values and choose the best candidate. We performed fifty simulations for every activation function and for every set of parameters. For activation functions we used Gaussian, tangent hyperbolic and sigmoid activation functions.
These activation functions are defined as follows, respectively

$$f(x) = e^{-a^2(x-b)^2} \tag{29}$$

$$f(x) = \frac{e^{2a(x-b)} - 1}{e^{2a(x-b)} + 1} \tag{30}$$

$$f(x) = \frac{1}{e^{-a(x-b)} + 1} \tag{31}$$

Resulting ISE and IAE mean values for different activation functions are given in Table 1, Table 2 and Table 3.

Table 1 **Mean values of ISE and IAE with original parameters**
Tablica 1 **Srednje vrijednosti IAE i ISE kriterija za originalne parametre**

	ISE	IAE
Gauss	0.10828	7.07312
Sigm	0.02505	3.68759
Tanh	0.01839	3.20787

Table 2 **Mean values of ISE and IAE with decreased parameters**
Tablica 2 **Srednje vrijednosti IAE i ISE kriterija za smanjene parametre**

	ISE	IAE
Gauss	0.08254	6.26309
Sigm	0.02618	3.80704
Tanh	0.02047	3.39294

Table 3 **Mean values of ISE and IAE with increased parameters**
Tablica 3 **Srednje vrijednosti IAE i ISE kriterija za povećane parametre**

	ISE	IAE
Gauss	0.19128	9.65296
Sigm	0.02413	3.58270
Tanh	0.01696	3.05489

From tables above we can see that NN with tangent hyperbolic activation functions shows the best results, but we have to see if differences are statistically significant.

First we will check the distribution of recorded IAE and ISE simulation results for all activation functions. Normal probability plots are given in Figure 4 to Figure 9.

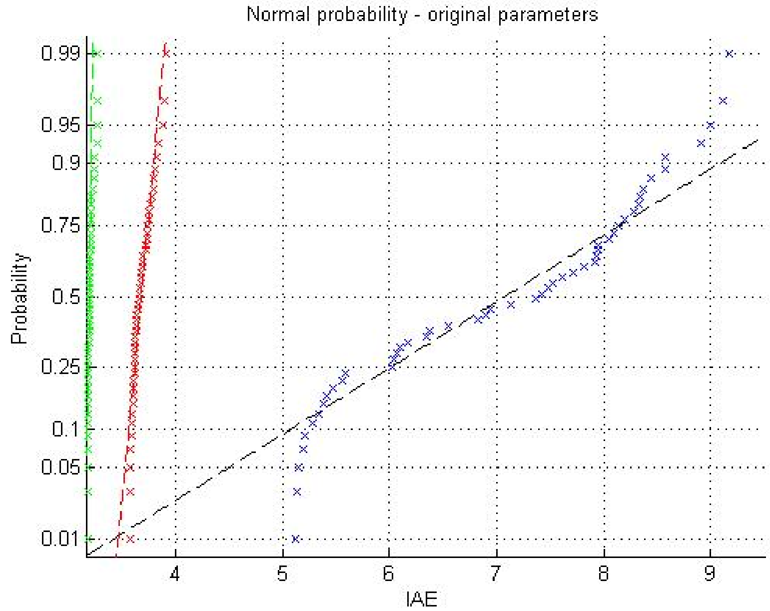

Figure 4 **Normal probability plot for IAE and original parameters**
Slika 4 **Normalna razdioba IAE kriterija uz originalne parametre**

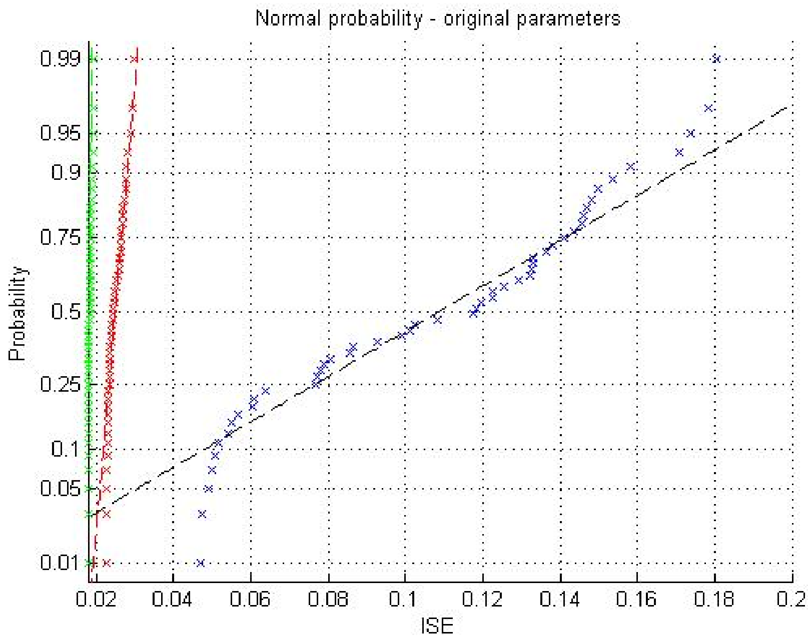

Figure 5 **Normal probability plot for ISE and original parameters**
Slika 5 **Normalna razdioba ISE kriterija uz originalne parametre**

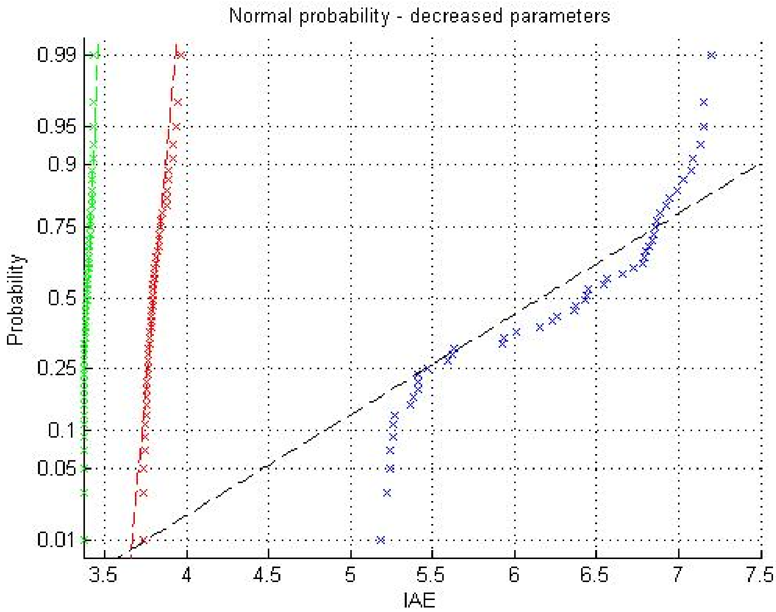

Figure 6 **Normal probability plot for IAE and decreased parameters**
Slika 6 **Normalna razdioba IAE kriterija uz smanjene parametre**

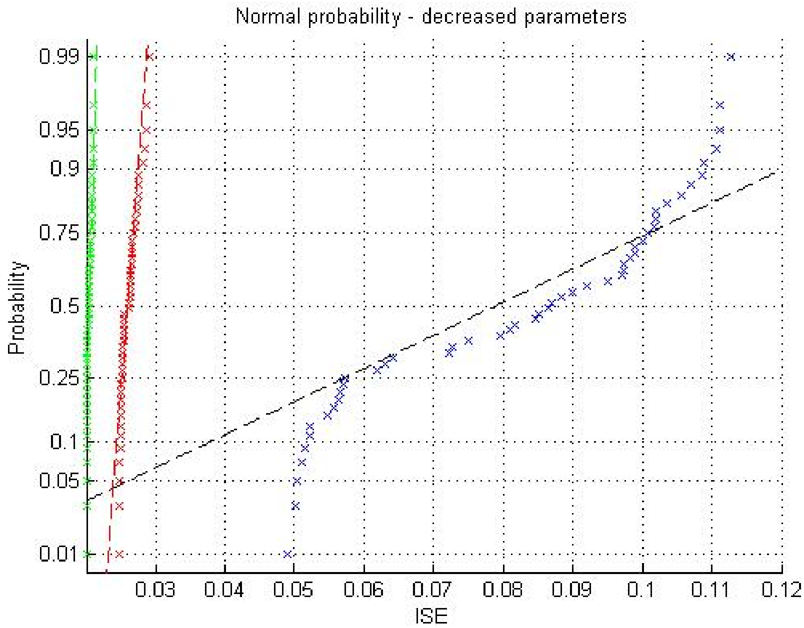

Figure 7 **Normal probability plot for IAE and decreased parameters**
Slika 7 **Normalna razdioba ISE kriterija uz smanjene parametre**

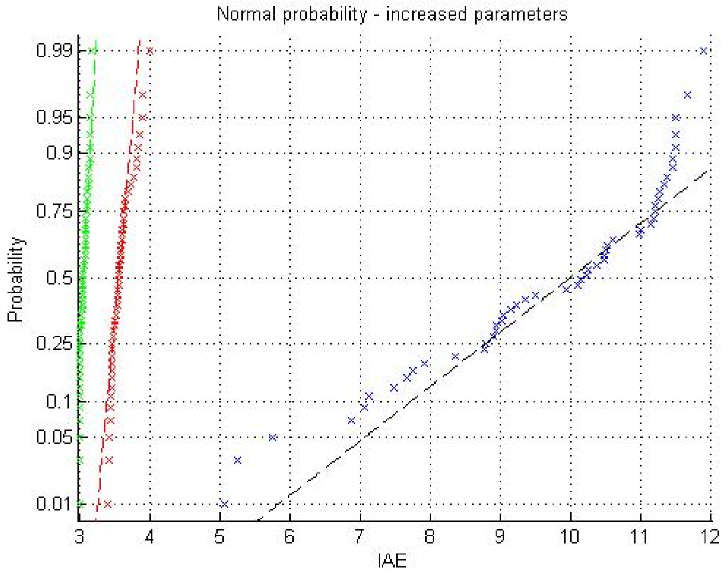

Figure 8 **Normal probability plot for IAE and increased parameters**
Slika 8 **Normalna razdioba ISE kriterija uz povećane parametre**

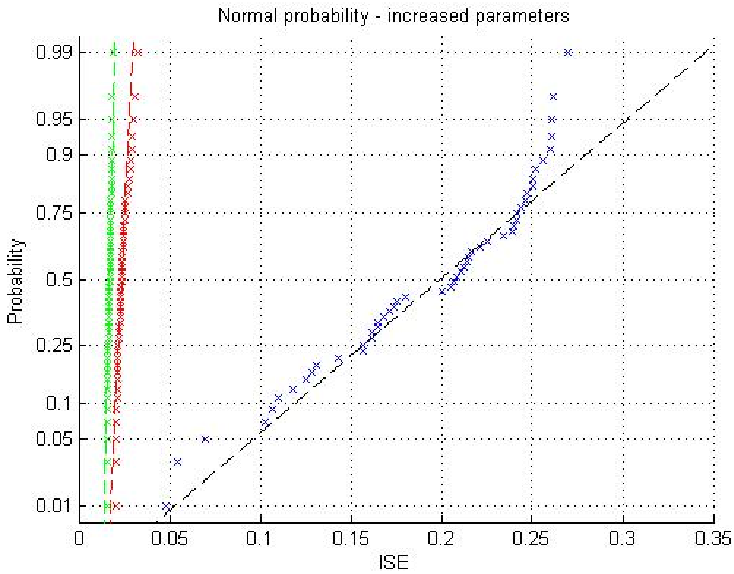

Figure 9 **Normal probability plot for IAE and increased parameters**
Slika 9 **Normalna razdioba ISE kriterija uz povećane parameter**

From Figure 4 to Figure **9** we can see that distributions are normal, so now we can compare samples by standard T test for hypothesis testing, with H_0 hypothesis stating that the means are the same and H_1 hypothesis stating that one mean is smaller than the other. This test is performed with 0.025 level of significance.

T test results are given in Table 4, Table 5 and Table 6. Subscripts assigned to the criteria name show activation function for obtained mean value of integral error.

Table 4 **T test results for IAE and ISE and original parameters**
Tablica 4 **T test rezultati za ISE I IAE kriterij uz originalne parametre**

	Test	P-value	T-value
$IAE_{tanh}<IAE_{sgm}$	yes	2.20498e-059	-36.71284
$IAE_{sgm}<IAE_{gss}$	yes	3.62346e-034	-18.57106
$ISE_{tanh}<ISE_{sgm}$	yes	2.71357e-042	-23.45751
$ISE_{sgm}<ISE_{gss}$	yes	1.63668e-026	-14.54235

Table 5 **T test results for IAE and ISE and decreased parameters**
Tablica 5 **T test rezultati za ISE I IAE kriterij uz smanjene parametre**

	Test	P-value	T-value
$IAE_{tanh}<IAE_{sgm}$	yes	1.03131e-068	-46.91115
$IAE_{sgm}<IAE_{gss}$	yes	1.29194e-044	-24.99742
$ISE_{tanh}<ISE_{sgm}$	yes	6.26923e-055	-32.78719
$ISE_{sgm}<ISE_{gss}$	yes	3.14151e-034	-18.60571

Table 6 **T test results for IAE and ISE and increased parameters**
Tablica 6 **T test rezultati za ISE I IAE kriterij uz povećane parametre**

	Test	P-value	T-value
$IAE_{tanh}<IAE_{sgm}$	yes	2.93560e-043	-24.08945
$IAE_{sgm}<IAE_{gss}$	yes	2.09944e-043	-24.18575
$ISE_{tanh}<ISE_{sgm}$	yes	9.03157e-030	-16.20051
$ISE_{sgm}<ISE_{gss}$	yes	6.89979e-037	-20.12731

If we look at Table 4 toTable **6** we can see that all results are statistically significant and now we can say that NN control with the tangent hyperbolic activation function gives the best results in control since it gives the smallest integral errors.

Simulation responses for $w_R = 0.1\sin(0.001t)pu$ are given in Figures 10 to 12. It can be seen that the system is stable in all cases.

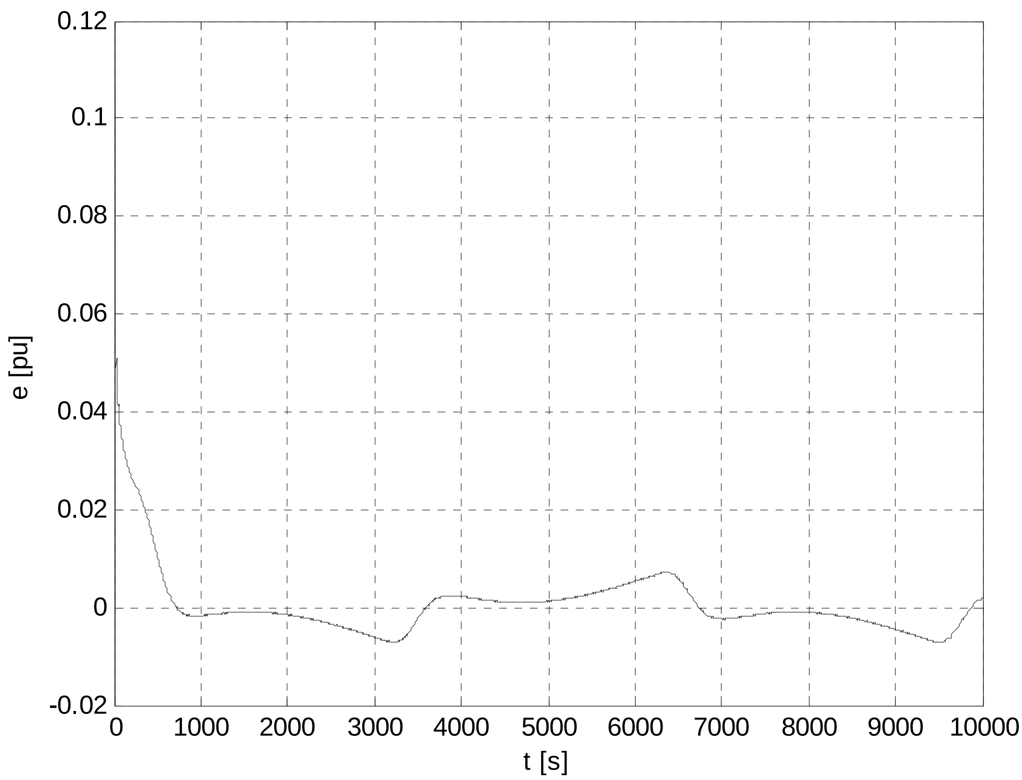

Figure 10 **Error plot for nominal parameters**
Slika 10 **Greška za nazivne parametre**

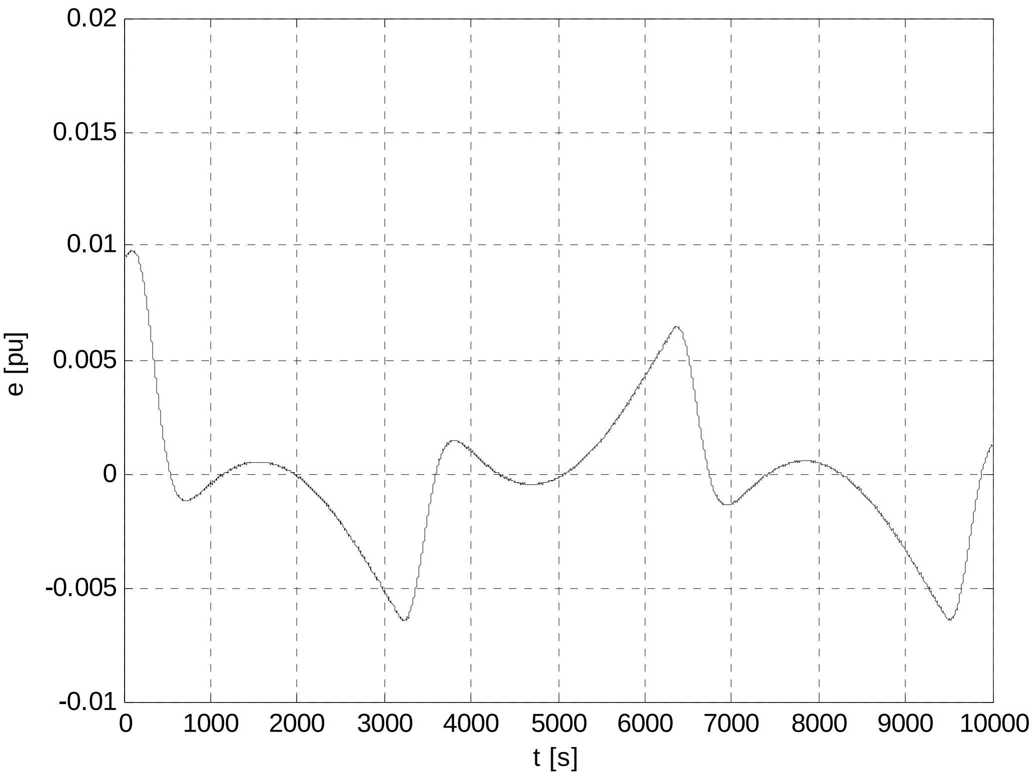

Figure 11 **Error plot for 10% decreased parameters**
Slika 11 **Greška za 10% smanjene parametre**

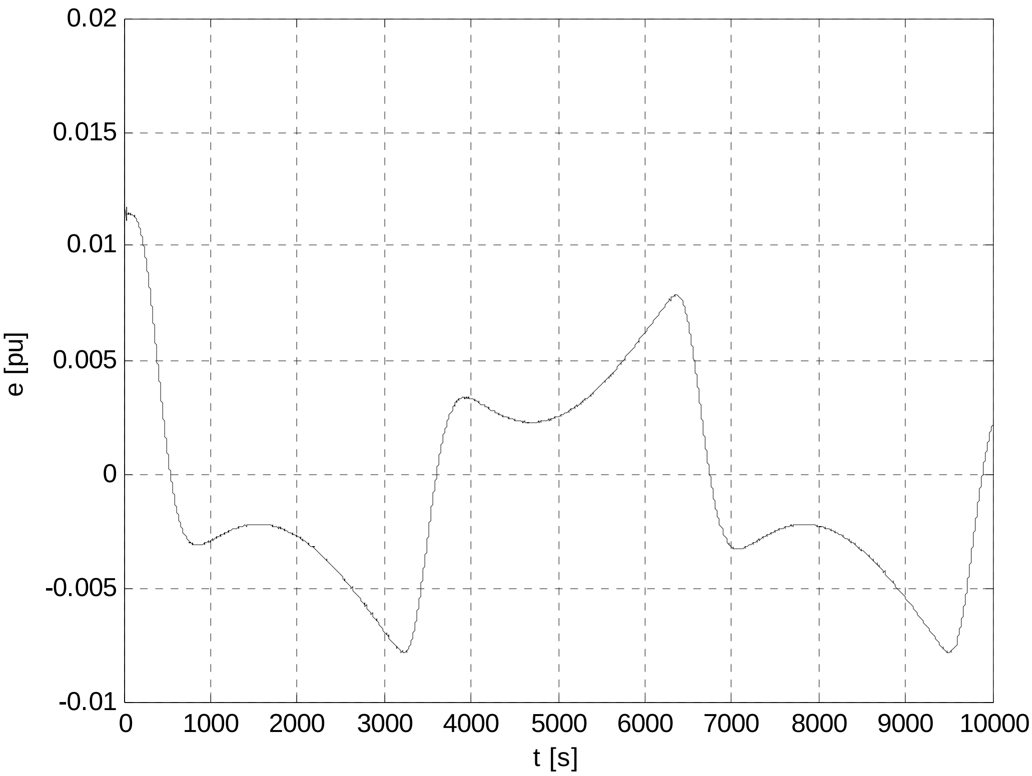

Figure 12 **Error plot for 10% increased parameters**
Slika 12 **Greška za 10% povećane parametre**

7 Conclusion

The paper presents a practical approach to neural network design for angular velocity system for thermogenerator. First, neural network with fixed randomly chosen weights of the first layer is defined. Then a Lyapunov-like stability analysis was performed in order to find the weights update law. The simulation procedure for the selection of the number of nodes and other activation functions is given.

Described procedure can be easily carried out using a digital computer and can be applied for any kind of system with similar dynamics. Resulting control law is always stable and ensures good reference tracking.

References

[1] SUEVALOV, L.F.: "Handbook on Ship Automated Systems", Sudostroenie, Leningrad, 1997 (in Russian).

[2] TOVARNIK, B., KULJAČA, O.: "Fuzzy Control of Ship's Control Systems", Proceedings of the 37th ELMAR, p. 213-216, Pula 1995.

[3] DEMIROREN, A., ZEYNEIGLY, H.I., SENGOR, N.S.: "The application of ANN technique to load-frequency control for three-area power system", IEEE Porto Power Tech Conference, Porto, Portugal, 2001.

[4] BEVRANI, H.: "A novel approach for power system load frequency controller design", Asia Pacific. IEEE/PES Transmission and Distribution Conference and Exhibition, Oct. 2002, p. 184-189.

[5] BIRCH, A.P., SAPELUK, A.T., OZVEREN, C.T.: "An enhanced neural network load frequency control technique", International Conference on Control, 1994, Coventry, UK, Mar. 1994, p. 409-415.

[6] LEWIS, F.L., JAGANNATHAN, S., YESILDIREK, A.: "Neural network control of robot manipulators and nonlinear systems", Taylor and Francis, London UK, 1998.

[7] SADEGH, N.: "A perceptron network for functional identification and control of nonlinear systems", IEEE, Trans. of Neural Networks, 4(6), 1993, p. 982-988.

[8] KULJAČA, O., HORVAT, K., BOROVIĆ, B.: "Design of adaptive neural network controller for thermal power system frequency control", ATKAFF 52(4), 2011, p. 319-328.

[9] BERGHUIS, H., NIJMEIJER, H.: "Robust control of robots via Linear Estimated State Feedback", IEEE Transactions on Automatic Control, Vol. 39, No. 10.