# ADAPTIVE CRITIC DESIGN BASED NEUROCONTROLLERS FOR TURBOGENERATOR CONTROL

Ganesh Kumar Venayagamoorthy
Dept. of Electrical and Computer Engineering, University of Missouri-Rolla, USA, email: *gkumar@ieee.org*

## I. Introduction

Nonlinear non-stationary large scale (NNLS) systems such as power system networks, telecommunication networks, financial and transportation systems are difficult to control and manage. Mathematical models of such systems are typically derived based on linear techniques, and wide margins of safety are allowed in order to ensure stable operation. In the era of a deregulated electricity industry, and an emphasis on competitive pricing, it will become necessary to reduce these safety margins as much as possible while still maintaining reliable service. Advanced intelligent identification and control schemes are therefore required in order to maximize the throughput/performance of such systems by reducing the otherwise wide safety margins. Neurocontrollers offer such a possibility.

Power transmission lines are increasingly called upon to transmit more power due to economic considerations but are limited by certain constraints. In many cases, transient stability transfer limits are more constraining than steady-state limits under contingency. On the other hand, operating conditions of large scale power systems are always varying to satisfy different load demands. Automatic control systems are localized around individual generators but are also required to damp the system oscillations that might threaten the system stability as load demands increase or after a major fault, and maintain the system stability under a diversity of operating conditions and different system configurations.

The traditional design of traditional control systems typically use approximately linearized models for the power system, the Automatic Voltage Regulators (AVRs), and the governor. Traditionally the analysis of the power network is simplified to a single machine connected through one transmission line to an infinite bus, referred to as a Single-Machine-Infinite-Bus (SMIB) system. Moreover, in practical systems SMIB subsystems do not exist, and the generators interact against each other. Power system stabilizers are then used to damp these interactive oscillations.

Nevertheless the SMIB is a useful starting point for designers to evaluate the design and dynamic performance of the individual controllers. The nonlinear equations of this SMIB system are then linearized at one operating point and used for the design of traditional controllers. However, when a major fault occurs, the behavior of the power system may be significantly different from that described by the linearized equations. Traditional linear controllers do not guarantee the system stability under such conditions [1].

Various techniques have been developed to improve on the above design techniques by using adaptive controllers [2], but most adaptive control algorithms use linear models, with certain assumptions of types of noise and possible disturbances. Based on these models, traditional techniques of identification, system analysis and synthesis can be applied to achieve the desired performance. However, the turbogenerator system is time varying and nonlinear, with complex dynamic and transient processes, hence it cannot be completely described by such linear models and in addition, certain physical phenomena cannot even be modeled. Furthermore, for the design of adaptive controllers, it typically is assumed that the number of system inputs equals the number of system outputs. Where necessary, this is achieved by using a transformation to reduce the dimensions of the output space, with the drawback that this degrades the description of the system dynamics. While this is possible for SMIB systems, it becomes unwieldy and almost impossible in large multi-machine power system. Consequently, the issues of unmodeled dynamics and robustness arise in practical applications of these adaptive control algorithms and hence supervisory control is required in the form of expert manpower in control rooms, all trying to stabilize local variations.

To allow for all these uncertainties, traditional controllers have been designed with large safety margins. In the era of a deregulated electricity industry, and an emphasis on competitive pricing, it will become necessary to reduce these safety margins as much as possible while still maintaining reliable service. This can be achieved by including some form of intelligent control based, for example, on artificial neural networks (ANNs).

ANNs are good at identifying a nonlinear system and then controlling it [3], and they are suitable for multi-variable applications, where they can identify the interactions between the inputs and outputs. This removes the need for an accurate model of the power system. It has been shown [4] that a multi-layer feedforward neural network using deviation signals as its inputs can identify the complex and nonlinear dynamics of a SMIB system with sufficient accuracy to then be used to design a generic controller, which yields optimal dynamic system response irrespective of the generator load and system configurations. Numerous publications have reported on the designs of ANN controllers for turbogenerators, and have presented both simulation and experimental results [5,6,7] showing that ANNs have the potential to supplement and even replace traditional controllers. Detailed studies were carried out previously on the real time implementation of a Continually Online Trained (COT) ANN identifier/controller for a SMIB system and simulation results were validated by actual measurements [6] on a laboratory system [8].

The work reported here extends earlier work of the author and presents a technique for designing a turbogenerator neurocontroller, which overcomes the stability issues and the computational load of online training. In this technique, the neurocontroller uses adaptive critic designs based on reinforcement learning and dynamic programming. The main elements of a typical section of a power system are described below. Adaptive critic designs are described in Section III. Simulation and practical results are presented in Section IV and V respectively.

**II. Single-Machine-Infinite-Bus Power System**

A single-machine-infinite-bus power system is shown in Fig. 1. This system consists of a generator connected to the infinite bus through a transmission line. An infinite bus has a fixed voltage and frequency. The generator is equipped with a traditional AVR and turbine governor. The design of each of the traditional controller is usually based on a SMIB linearized model, thereby ignoring the true dynamics of the rest of the generators that exist on the power system. The performance of the traditional AVR and governor controllers in Fig. 2 degrade, once the operating conditions of the system change. Fig. 2 shows a SMIB system in which the AVR and governor have been either replaced or supplemented by an ANN controller (also called a neurocontroller).

Moreover, the controllers of the generators in multi-machine power system respond or react according to local information and may create power flow oscillations between the generators and between areas. These oscillations can be damped out using power system stabilizers on some of the generators, but in a large multi-machine power system the positions for such stabilizers is not a trivial decision. An intelligent nonlinear controller (ANN based) can be used as shown in Fig. 2 to replace the power system stabilizer and to make intelligent decisions to ensure local and global stability. In the multi-machine system, each generator could be equipped with such an intelligent nonlinear controller which will take intelligent decisions based on whatever information is supplied to it. Typically this would be local information of the generator's speed, voltage, and power flowing from it.

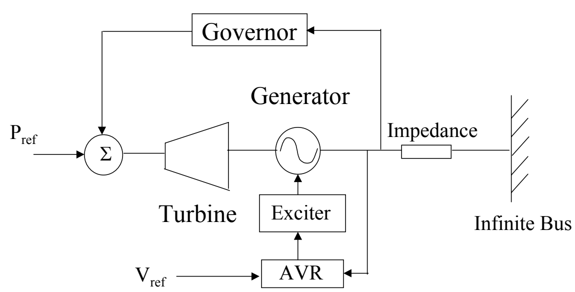


*Fig. 1 A single-machine- infinite-bus power system with a traditional controller*

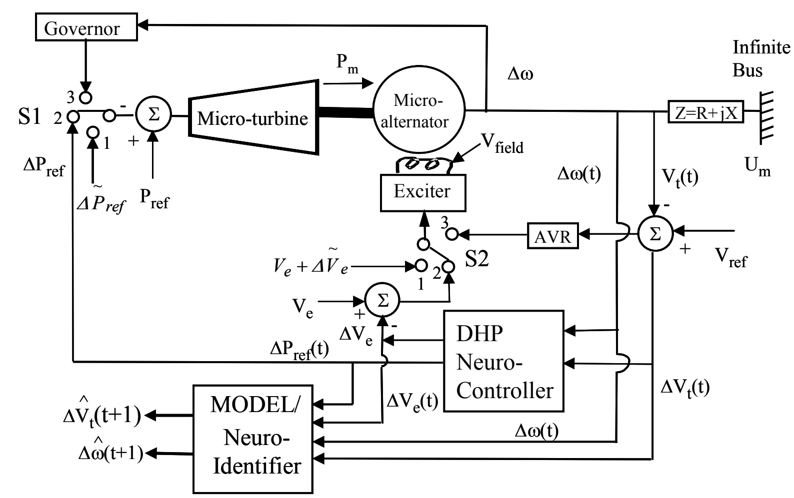


*Fig. 2 A single-machine-infinite-bus power system with an intelligent nonlinear controller (ANN based)*

## III. Adaptive Critic Designs

Adaptive critic designs are neural network designs capable of optimization over time under conditions of noise and uncertainty. A family of ACDs was proposed by Werbos [9] as a new optimization technique combining concepts of reinforcement learning and approximate dynamic programming. For a given series of control actions that must be taken sequentially, and not knowing the effect of these actions until the end of the sequence, it is impossible to design an optimal controller using the traditional supervised learning artificial neural network (ANN). The adaptive critic method determines optimal control laws for a system by successively adapting two ANNs, namely an ***action neural network*** (which dispenses the control signals) and a ***critic neural network*** (which 'learns' the desired performance index for some function associated with the performance index). These two neural networks approximate the Hamilton-Jacobi-Bellman equation associated with optimal control theory. The adaptation process starts with a non-optimal, arbitrarily chosen, control by the action network; the critic network then guides the action network towards the optimal solution at each successive adaptation. During the adaptations, neither of the networks need any 'information' of an optimal trajectory, only the desired cost needs to be known. Furthermore, this method determines optimal control policy for the entire range of initial conditions and needs no external training, unlike other neurocontrollers.

Dynamic programming prescribes a search which tracks backward from the final step, retaining in memory all suboptimal paths from any given point to the finish, until the starting point is reached. The result of this is that the procedure is too computationally expensive for most real problems. In supervised learning, an ANN training algorithm utilizes a desired output and, having compared it to the actual output, generates an error term to allow the network to learn. The backpropagation algorithm is typically used to obtain the necessary derivatives of the error term with respect to the training parameters and/or the inputs of the network. However, backpropagation can be linked to reinforcement learning via the critic network which has certain desirable attributes.

The technique of using a critic, removes the learning process one step from the control network (traditionally called the "action network" or "actor" in ACD literature), so the desired trajectory is not necessary. The critic network learns to approximate the ***cost-to-go*** or strategic utility function (the function $J$ of Bellman's equation in dynamic programming) and uses the output of the action network as one of its inputs, directly or indirectly. Different types of critics have been proposed. For example, Watkins [Watkins 1] developed a system known as Q-learning, explicitly based on dynamic programming. Werbos, on the other hand, developed a family of systems for approximating dynamic programming [Werbos 7]; his approach subsumes other designs for continuous domains. For example, Q-learning becomes a special case of Action-Dependent Heuristic Dynamic Programming (ADHDP), which is a critic approximating the $J$ function (see section B below), in Werbos' family of adaptive critics. A critic which approximates only the derivatives of the function $J$ with respect to its states, called the Dual Heuristic Programming (DHP), and a critic approximating both $J$ and its derivatives, called the Globalized Dual Heuristic Programming (GDHP), complete this ACD family. These systems do not require exclusively neural network implementations, since any differentiable structure is suitable as a building block. The interrelationships between members of the ACD family have been generalized and explained in detail by Prokhorov [10].

*Dual Heuristic Programming*

The critic neural network in the DHP scheme in Fig. 6, estimates the derivatives of $J$ with respect to the vector $\Delta Y$, and learns minimization of the following error measure over time:

$$J(t)=\sum_{k=0}^{\infty} \gamma^k U(t+k) \tag{1}$$

$$\|E\|=\sum_t E^T(t)E(t) \tag{2}$$

where

$$E(t)=\frac{\partial J[\Delta Y(t)]}{\partial \Delta Y(t)}-\gamma\frac{\partial J[\Delta Y(t+1)]}{\partial \Delta Y(t)}-\frac{\partial U(t)}{\partial \Delta Y(t)} \tag{3}$$

where $\partial(.)/\partial\Delta Y(t)$ is a vector containing partial derivatives of the scalar (.) with respect to the components of the vector $\Delta Y$. The DHP critic neural network structure has two linear output neurons as shown in Fig. 7. The critic neural network's training is more complicated than in HDP, since there is a need to take into account all relevant pathways of backpropagation as shown in Fig. 6, where the paths of derivatives and adaptation of the critic are depicted by dashed lines.

In the DHP scheme, application of the chain rule for derivatives yields

$$\frac{\partial J[\Delta Y(t+1)]}{\partial \Delta Y_i(t)} = \sum_{i=1}^{n} \lambda_i(t+1)\frac{\partial Y_i(t+1)}{\partial \Delta Y_i(t)} + \sum_{k=1}^{m}\sum_{i=1}^{n} \lambda_i(t+1)\frac{\partial \Delta Y_i(t+1)}{\partial A_k(t)}\frac{\partial A_k(t)}{\partial \Delta Y_i(t)} \tag{4}$$

where $\lambda_i(t+1) = \partial J[\Delta Y(t+1)]/\partial \Delta Y_i(t+1)$, and $n$, $m$ are the numbers of outputs of the model and the action neural networks, respectively. By exploiting eq. (4), each of $n$ components of the vector $E(t)$ from eq. (3) is determined by

$$E(t) = \frac{\partial J[\Delta Y(t)]}{\partial \Delta Y_i(t)} - \gamma\frac{\partial J[\Delta Y(t+1)]}{\partial \Delta Y_i(t)} - \frac{\partial U(t)}{\partial \Delta Y_i(t)} - \sum_{k=1}^{m}\frac{\partial U(t)}{\partial A_k(t)}\frac{\partial A_k(t)}{\partial \Delta Y_i(t)} \tag{5}$$

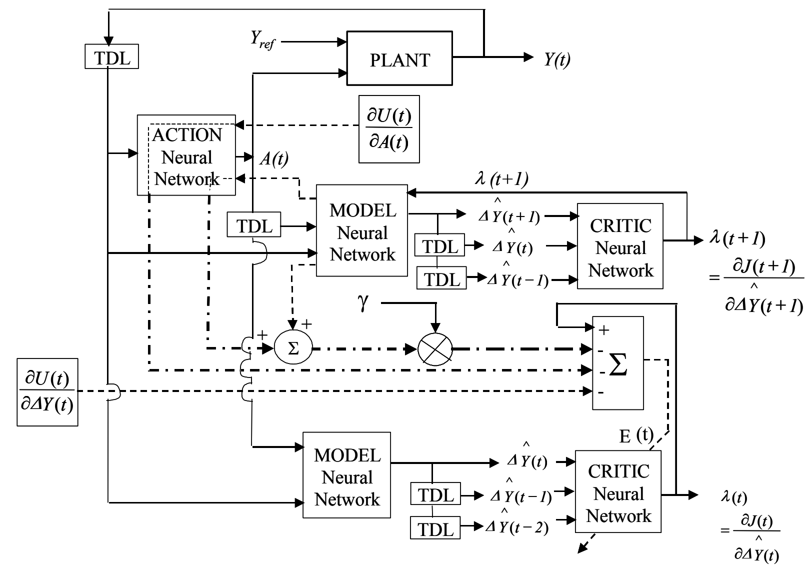


*Fig. 6 DHP critic neural network adaptation. This diagram shows the implementation of (4). The same critic network is shown for two consecutive times, t and t +1. Discount factor $\gamma = 0.5$. BP paths are shown by dashed lines. The output of the critic network $\lambda(t+1)$ is backpropagated through the Model from its outputs to its inputs, yielding the first term of (4) and $\partial J(t+1)/\partial A(t)$. The latter is backpropagated through the Action from its output to its input forming the second term of (4). BP of the vector $\partial U(t)/\partial A(t)$ through the Action results in a vector with components computed as the last term of (6). The summation produces the error vector E(t) for critic training.*

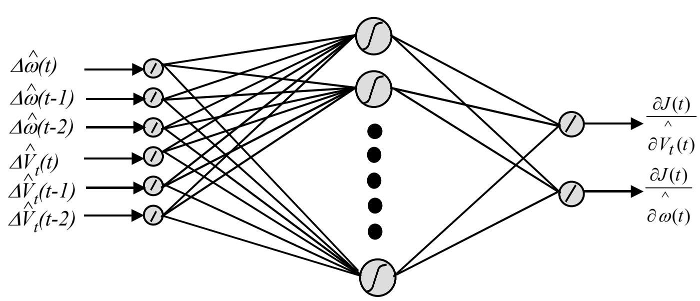


*Fig. 7 DHP critic neural network structure with six inputs, ten sigmoidal hidden layer neurons and two linear output neurons.*

The adaptation of the action neural network in Fig. 6 is illustrated in Fig. 8 which propagates $\lambda(t+1)$ back through the model network to the action network. The goal of such adaptation can be expressed as follows [10]:

$$\frac{\partial U(t)}{\partial A(t)} + \gamma\frac{\partial J(t+1)}{\partial A(t)} = 0 \quad \forall \ t \tag{6}$$

The weights' update expression [10], when applying backpropagation, is as follows:

$$\Delta W_{A2} = -\eta_4 \left[ \frac{\partial U(t)}{\partial A(t)} + \frac{\partial J(t+1)}{\partial A(t)} \right]^T \frac{\partial A(t)}{\partial W_{A2}} \tag{7}$$

where $\eta_4$ is a positive learning rate and $W_{A2}$ is the weights of the action neural network in the DHP scheme. The structure of the action neural network is identical to that of the action network in the HDP scheme. The general derivation of the equations in this section are shown in [10] in detail. The word "Dual" is used to describe the fact that the target outputs for the DHP critic training are calculated using backpropagation in a generalized sense; more precisely, it does use dual subroutines (states and co-states) to backpropagate derivatives through the model and action neural networks, as shown in Fig. 6. The dual subroutines and more explanations are found in [9, 11].

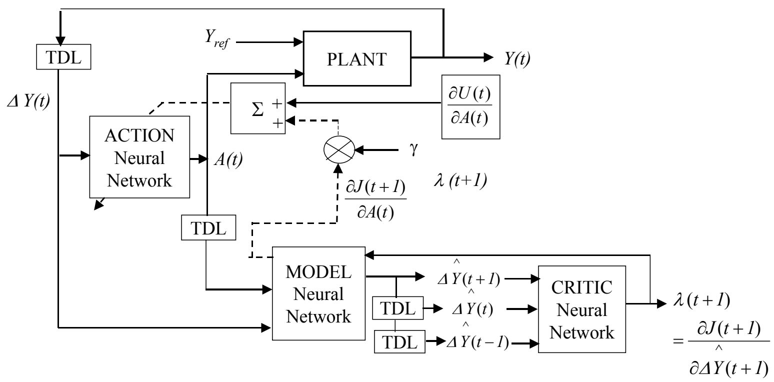


*Fig. 8 DHP action network adaptation. BP paths are shown with dashed lines. The output of the critic $\lambda(t+1)$ at time $(t+1)$ is backpropagated through the Model from its outputs to its inputs, and the resulting vector is multiplied by the discount factor $\gamma$ and added to $\partial U(t)/\partial A(t)$. Then an incremental adaptation of the action network is carried in accordance with (7).*

## IV. Simulation Results

For two different tests, the transient performance of the DHP neurocontroller is compared in simulation, with that of the conventional controllers (AVR and turbine governor) [6].

*Step changes in the terminal voltage reference $V_{ref}$ or $V_e$ (Fig. 2)*

Figs. 9 and 10 show the terminal voltage and the rotor angle of the micro-alternator for ± 5% step changes in the terminal voltage with the micro-alternator operating at 1 pu power and 0.85 lagging power factor, and line impedance $Z_1$ = 0.02 + j 0.4 pu. The DHP neurocontroller clearly outperform the conventional controllers. The DHP neurocontroller has a better damping than the conventional controller. The damping factor and rise time is influenced by the local utility function and the discount factor in eq. (2), and the inherent characteristics of the DHP scheme.

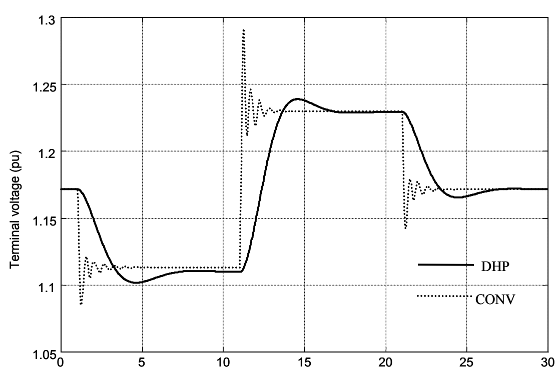


*Fig. 9 Terminal voltage of the micro-alternator for ± 5% step changes in the terminal voltage reference (transmission line impedance $Z_1$)*

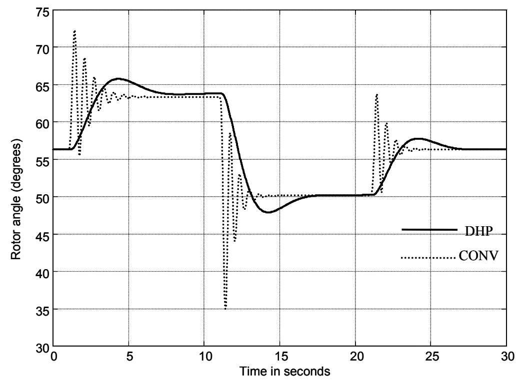


*Fig. 10 Rotor angle of the micro-alternator for ± 5% step changes in the terminal voltage reference (transmission line impedance $Z_1$)*

*Short circuit test with increased transmission line impedance (Fig. 2)*

In power systems faults such as three phase short circuits occur from time to time, and because they prevent energy from the generator reaching the infinite bus, it means that most of the turbine shaft power goes into accelerating the generator during the fault. This represents a very severe transient test for the controller performance. Figs. 11 and 12 show the terminal voltage and the rotor angle of the micro-alternator in Fig. 2 operating under the same conditions as in Figs. 9 and 10, and with increased line impedance $Z_2 = 0.025 + j\ 0.8$ pu, but with a temporary three phase short circuit applied at the infinite bus for 50 ms at t = 1 s. Increasing the line impedance represents the case of one of two parallel transmission lines, or part of a ring connected power system, being switched out. Figs. 11 and 12 show that the DHP controller clearly beat the conventional controllers in terms of offering the greatest oscillation damping, especially in the rotor angle. The DHP controller proves its robustness to changes in the system configurations.

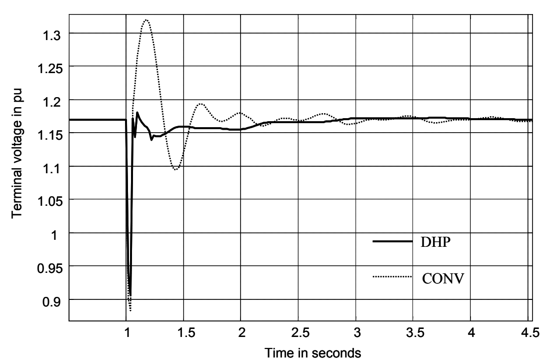


*Fig. 11 Terminal voltage of the micro-alternator for a temporary 50 ms three phase short circuit (transmission line impedance $Z_2$)*

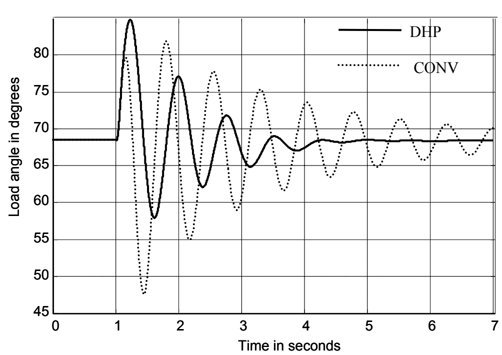


*Fig. 12 Load/Rotor angle of the micro-alternator for a temporary 50 ms three phase short circuit (transmission line impedance $Z_2$)*

## V. Practical Results

At two different operating conditions, the transient performance of the DHP neurocontroller is compared experimental in real time, with that of the conventional controllers (AVR and turbine governor) [6], as well as with that of the AVR equipped with a PSS (whose parameters are carefully tuned [12] for the first set of operating conditions).

*Removal of a series transmission line at the first operating condition (P = 0.2 pu, Q = 0 pu)*
At the *first* operating condition, the series transmission line impedance is decreased at time $t$ = 10 seconds from Z = 0.0044 + j1.50 pu to Z = 0.0022 + j0.75 pu by closing switch S2 (Fig. 13). Figs. 14 and 15 show the terminal voltage and the rotor angle response for this test. Clearly the DHP neurocontroller again exhibits superior damping over the performance of the conventional controllers even when equipped with a PSS.

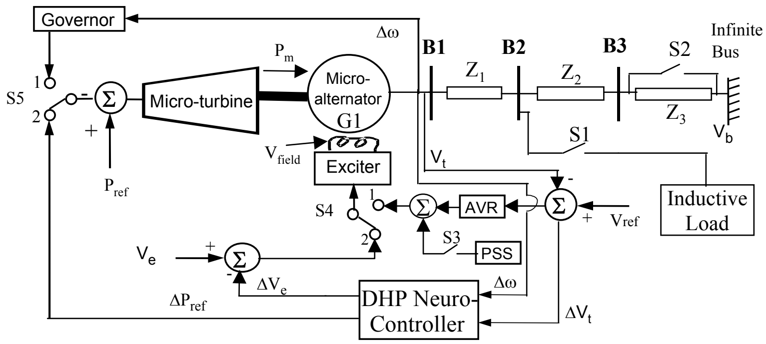


*Fig. 13 Practical laboratory power system model (transmission line impedance $Z_1$ = 0.01 +j0.25, $Z_2$ = 0.012 +j0.50 and $Z_3$ = 0.022 +j0.75).*

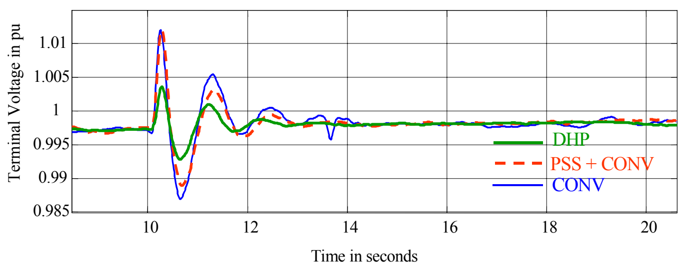


*Fig. 14. Terminal voltage response for series transmission line impedance decrease by closing switch S2 (Fig. 1) for P = 0.2 pu & Q = 0 pu.*

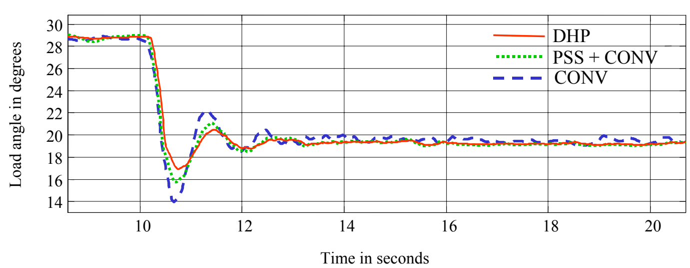


*Fig. 15. Load angle response for series transmission line impedance decrease by closing switch S2 (Fig. 1) for P = 0.2 pu & Q = 0 pu.*

*Removal of a series transmission line at the second operating condition (P = 0.3 pu, Q = 0 pu)*
At the *second* operating condition, the series transmission line impedance is increased at time $t$ = 10 seconds from Z = 0.0044 + j1.50 pu to Z = 0.0022 + j0.75 pu by closing switch S2 (Fig. 13). Figs. 16 and 17 show the terminal voltage and the rotor angle response for this test. The DHP neurocontroller provides the best damping despite the change in the operating condition, unlike with the conventional controllers (AVR, governor and PSS). This proves that the DHP neurocontroller has learned the new operating condition and has the ability to perform robustly at all operating conditions.

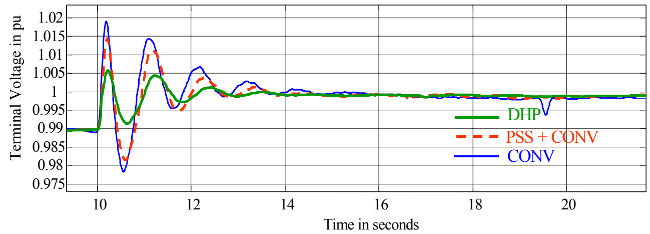


*Fig. 16. Terminal voltage response for series transmission line impedance decrease by closing switch S2 (Fig. 1) for P = 0.3 pu & Q = 0 pu.*

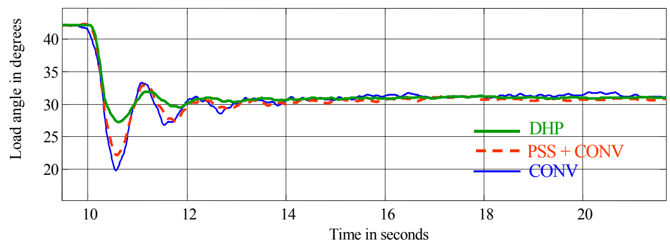


*Fig. 17. Load angle response for series transmission line impedance decrease by closing switch S2 (Fig. 1) for P = 0.3 pu & Q = 0 pu.*

## VI. Conclusions

Adaptive critic based neurocontrollers (DHP) have been successfully designed and evaluated in this project to control turbogenerators in a single-machine-infinite-bus power system. The real-time implementation demonstrates the validity of the concept of using adaptive critic based neurocontrollers and its potential for turbogenerator control. This technique of neurocontroller design does not require continual online training thus overcoming the risks of instability.

## VII. References


[1] Adkins B, Harley RG, "The general theory of alternating current machines", *Chapman and Hall*, London, 1975, ISBN 0-412-15560-5.

[2] Wu QH, Hogg BW, "Adaptive controller for a turbogenerator system", *IEE Proceedings*, Vol 135, Pt D, No 1, 1988, pp 35 – 42.

[3] Hunt KJ, Sbarbaro D, Zbikowski R, Gawthrop PJ, "Neural networks for control systems – a survey", *Automatica*, Vol 28, No 6, 1992, pp 1083 – 1112.

[4] Venayagamoorthy GK, Harley RG, "A continually online trained artificial neural network identifier for a turbogenerator", accepted for publication in the *Proceedings of IEEE International Electric Machines and Drives Conference IEMDC' 99*, Seattle, USA, 9 – 12 May, 1999.

[5] Venayagamoorthy GK, Harley RG, "Simulation studies with a continuously online trained artificial neural network controller for a micro-turbogenerator", *Proceedings of IEE International Conference on Simulation*, University of York, UK, 30 September – 2 October 1998, pp 405 – 412.

[6] Venayagamoorthy GK, Harley RG, "A Continually Online Trained Neurocontroller for Excitation and Turbine Control of a Turbogenerator", *IEEE Transactions on Energy Conversion*, vol. 16, no.3, pp. 261-269.Flynn D, McLoone S, Irwin GW, Brown MD, Swidenbank E, Hogg BW, "Neural control of turbogenerator systems", *Automatica*, Vol 33, No 11, 1997, pp 1961 – 1973.

[7] Flynn D, McLoone S, Irwin GW, Brown MD, Swidenbank E, Hogg BW, "Neural control of turbogenerator systems", *Automatica*, Vol. 33, No. 11, 1997, pp. 1961 – 1973.

[8] Limebeer DJN, Harley RG, Lahoud MA, "A laboratory system for investigating subsynchronous resonance", Paper A80-0190-0, *IEEE PES Winter Power Meeting*, New York, USA, Feb 4 - 8, 1980.

[9] Werbos P, "Approximate dynamic programming for real-time control and neural modeling*", in Handbook of Intelligent Control*, White and Sofge, Eds., Van Nostrand Reinhold, ISBN 0-442-30857-4, 1992, pp 493 – 525.

[10] Prokhorov D and Wunsch D, "Adaptive Critic Designs", *IEEE Transactions. on Neural Networks*, Vol. 8, No. 5, 1997, pp 997-1007.

[11] Werbos PJ, *Roots of backpropagation*, Wiley, USA, 1994, ISBN 0-471-59897-6.

[12] Kundur P, Klein M, Rogers GJ, Zywno MS, "Application of Power System Stabilizers for Enhancement of Overall System Stability*", IEEE Trans. on Power Systems*, vol. 4, no. 2, pp. 614-626, May 1989.

**Acknowledgements**

Support from the National Science Foundation, USA, IEEE Neural Network Society Summer Research Program, the National Science Foundation, South Africa and the University of Natal, South Africa are acknowledged. I am grateful to Professor Ronald Harley from Georgia Institute of Technology, Atlanta, USA and to Professor Donald Wunsch from the University of Missouri-Rolla, USA for their excellent guidance throughout this research project.