Improving Power System Neural Network Construction Using Modal Analysis

Joshua Johnson, Shamina Hossain-McKenzie, Uyen Bui, Sriharsha Etigowni*, Katherine Davis, Saman Zonouz*

***Abstract*—Historically, the structure of an Artificial Neural Network (ANN) has been defined through trial-and-error or excessive computation leading to reduced accuracy and increased training time, respectively. For many disciplines, especially power systems, models must both be accurate and support fast computations in order to be viable for large-scale use. These requirements often render poorly structured ANNs useless. However, using power system behavioral knowledge to create an ANN structure could provide a near best case estimate for a model that maximizes accuracy and minimizes computational run-time. This paper considers the relationship between the dominant modes of a power system and the hidden neurons (units) in an ANN. In this study, several ANNs were created with varying number of neurons. These ANNs were used to predict rotor angle response to faults at generator buses that were cleared at varying times and compared with actual responses, as obtained through simulation. The number of neurons used include the hypothesized dominant mode number and five known heuristic estimates. The resultant method is a domain-dependent algorithm to structure an ANN without relying on trial-and-error or additional unnecessary computation time for power system models.**

***Index Terms*–artificial neural network, modal analysis, generator model**

I. Introduction

From load forecasting to dynamic security assessment, artificial neural networks (ANN) have prominent and widespread applications in the power system domain [1]–[3]. Machine learning techniques such as ANN, support vector machine (SVM), clustering, and others have seen an increase in use over recent years, driven by the growing availability of data [3], [4]. Measurement sources include distributed devices such as phasor measurement units (PMUs) as well as smart meters and home appliances. By leveraging this access with powerful machine learning tools, power system decision making, situational awareness, and response are greatly improved.

In particular, ANNs are extensively used in formulating power system solutions. This is due to attractive features such as the ability to learn complex nonlinear relationships and modular structures that can allow parallel processing [5]. An ANN is a biologically inspired programming paradigm that is able to learn from observational data [6]. ANNs are also computing systems; each is composed of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs [7]. The large scale and nonlinearity of power systems are factors that contribute to their complexity, and ANNs hold promise for tackling these challenges.

Short-term load forecasting, defined as predicting future load series minutes, hours, or days ahead, has been achieved with ANN in experiments and practical tests. However, Hippert et al. [1] conducted a review and evaluation of these ANN-based forecasting systems to address skepticism that the ANN use has been systematically proven. Specifically, the study found that the neural network architectures chosen for the data samples were not suitable and perhaps too large with many parameters to be estimated, often resulting in overfitting and poor out-of-sample testing results [1]. The authors also noticed that the models were not systematically tested and that more rigorous results are needed to fully validate.

Besides load forecasting, ANNs have also been utilized to replace complex power system models to aid computation of different power system analyses. Xu et al. [8] considered the challenge of modeling nonlinear three-phase photovoltaic generators (PVGs) for power flow analysis. Transient stability assessment has been paired with ANN by Bahbah et al. [2] with generator angle and angular velocity prediction for multi-machine power systems. Additionally, Qian et al. [9] performed transient stability studies using ANN models for generators, excitation systems, and governors individually and subsequently linked them.

These ANN power system applications, both for load forecasting and replacing complex machine models, indicate experimental success. Yet, as noted by Hippert et al., no systematic approach is apparent, specifically with regard to selecting the number of ANN parameters (e.g., number of layers, number of neurons, activation function) and testing approaches. This paper focuses on developing a data-dependent and *power system-dependent* procedure for the selection of ANN parameters. The approaches used presently rely on trial-and-error by assessing resultant accuracy iteratively, existing model setups that may not translate for a different application, and/or outsized ANNs that may suffer from overfitting.

Nonetheless, we seek to improve the selection of ANN parameters by leveraging power system behavioral knowledge. The power system exhibits patterns rooted in the physics of the various components and interconnections as well as specific topologies. For example, when a disturbance occurs, we know the oscillatory response of the system will be dictated by the modes of the system [10] and that the topology of the system will impact the stability of the system given such a disturbance. We also know that voltage disturbances, caused

The authors would like to thank the National Science Foundation (NSF): Award Numbers CNS 1446229 and CNS 1446471, ARPA-E: Award No. DE-AR0000233 with Smart Wire Grid, Inc., and PowerWorld Corporation for their support and sponsorship of this research.

Affiliation: Department of ECE at University of Illinois at Urbana-Champaign and Rutgers University*

Contact: {jajohns4, shossai2, uyenbui2, krogers6}@illinois.edu, {se260, saman.zonouz}@rutgers.edu*

978-1-5090-4000- 1/17/$31.00 ©2017 IEEE

by faults or control actions, are localized and studying only local bus voltage measurements is sufficient for classification methods, significantly reducing training set size and computation time [11]. We develop analytical methods to reconstruct such insight into power system behavior and (in this paper) to leverage it for ANN modeling applications, particularly to reduce trial-and-error.

To explore the connection between power system analyses and ANN parameter selection, this paper develops a systematic method for selecting the number of neurons for a model. Instead of relying on trial-and-error or inconsistent heuristics, we present an algorithm that is dependent on the power system being studied and the data set. For this investigation, we replace generator models with ANN in a post-fault system. The input data consists of the generator real power and exciter field voltages, and the generator rotor angles are obtained as output, assessing the system stability. Modal analysis is applied to determine dominant modes of the system and we hypothesize that the number of dominant modes can be equated to the number of neurons to be used. This idea is developed further in the remainder of this paper.

II. Literature Review

A. Artificial Neural Networks

An artificial neural network (ANN) processes a set of input data, usually referred to as training data, through its structure of weights, connections, and activation functions that are then adjusted using a specific training algorithm. Through these iterative adjustments, the ANN will increase its ability to correctly recognize patterns and classify or quantify an output. In this section, we will provide a brief description of ANNs, but comprehensive and detailed reviews can be found in [1], [6], [7], [12].

The basic unit of a neural network is an artificial neuron that receives input data information and processes it. The input values are linearly combined, using input weights and constant bias terms, and then an activation function is applied. An example is shown in Figure 1 (biases not shown). This nonlinear activation function is required to be non-decreasing and differentiable (e.g., sigmoid function).

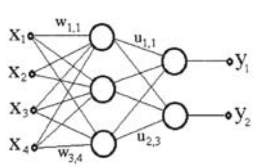


Fig. 1: Example two-layer neural network with four input nodes, various weights, and two output nodes [1].

For most power system applications, we utilize multilayer perceptron (MLP) networks that arrange the neurons, or more generally units, in layers [1]. In the feed-forward network, the outputs of one layer are inputs to the following layer; layers between are called hidden layers. The different weights on the connections and the bias terms are the parameters of the network, and estimating them is the focus of training the network using optimization functions such as gradient descent.

B. Power Systems ANN Applications

Since the early 1990s, the electric power system industry has seen a new movement toward artificial intelligence and machine learning to either model or predict certain phenomena within the systems. Due to prior successes, the most researched areas include load forecasting, fault diagnosis, economic dispatch, security assessment, and transient stability [12].

One notable application is in load forecasting and its effects on economic development and planning. Current models have had difficulty in many areas such as finding a relationship between variable and instantaneous load demand and ability to reevaluate the set of laws that govern the complex system and adjust themselves with rapid nonlinear system-load changes [13]. Another prominent application is the ability to diagnose faults, their locations, and how to most effectively clear them. An ANN in this setting has performed well in identifying problems and successfully diagnosing errors due to its flexibility in classification and its ability to handle noisy data which is generally produced during these events [13].

However, in this scenario, it is critical to have best case computational efficiency to avoid serious damage to the power system and the consumers it supports. This is an extremely difficult problem in modern ANN applications. During the 1990s, expert systems were the main tool used; however, they had a major drawback of not being able to handle the complexity of growing power systems. Recently, ANNs have been used to handle this complex problem; however, they too have a setback. With many neural units, or neurons, to model such a complex system, the cost, both in time and computational power, to train the model, whether it be feedforward, backpropagation, etc., is so large that at times it makes this model obsolete. Advancements in processing power and development of new training algorithms in the past decade have greatly benefited the feasibility of these models [14].

Although the application of ANNs in these areas of power systems has proven useful, improved structure is needed. Of the many necessities of creating a reliable and accurate neural network, there are two main focuses for power system ANNs. They are efficiency and the handling of noisy data, both being mainly driven by the size of the ANN itself; efficiency calls for fewer neurons (units) and handling noisy data calls for more. For efficiency, it is important for a model to accurately and, in a timely manner, predict its output before any damage is incurred to a power system. Any additional units within an ANN will slow both training and prediction. Conversely, handling noisy data is a task that generally improves as the number of units increases. If many units are present in the network, a meaningless input can be pinpointed during training and have devalued weights so that the values have minimal effect on the generated output. However, with fewer units in the network, it is more difficult as each unit is influenced by many input variables. If one unit is influenced by both an important variable and an insignificant one, the average will be middle-tier influence on the generated output that devalues important information and overvalues meaningless information, which will clearly lead to increased error and inhibit successful modeling.

C. Selection of Number of Neurons

The goal of this paper is to use a deep knowledge of power system behavior in order to find a balance of the two main focuses provided earlier, to create a potential best case number of units in a hidden neuron layer within an ANN model that can generate timely and accurate predictions, even with a set of noisy data. Selection of a neural network structure, in many scenarios, is done through computation. The number of neurons can be set as an arbitrary value, and during training, a better number will be found through some algorithm [15]. In almost all cases, this selection increases the computational cost and time incurred by additional cross-validation. For power systems, this additional cost may not actually outweigh the small benefit from error minimization if there is a known value that can perform better under tighter time constraints.

In recent years, there have been many attempts to provide a standard for the number of neurons in the hidden layer of an ANN. Generally, these all relate the number of input neurons, output neurons, input variables, and a few other metrics to generate the appropriate number of units. However, they are unfortunately too general to be very accurate for all sets of data, as there are potentially many more unknown variables that go into this calculation [16].

As mentioned, the goal of this paper is utilize a deep understanding of power system behavior in order to set a value for the number of hidden neurons in a neural network model. Through many experiments and generated results, we seek to link modal analysis of power systems with the number of neurons in the hidden layer of a neural network as we feel there can be a qualitative reason to model an ANN using the number of most dominant power system modes [17].

III. Method

A. Modal Analysis

Small signal stability is the ability of a power system to maintain its synchronism after a small disturbance. Modal analysis is the analysis of small signal stability through the eigenvalues; it also looks at the eigenvectors, the participation factors, and the mode shapes [18]–[20]. To obtain those parameters, first the power system is described by a set of equations:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{y}) \qquad \mathbf{0} = \mathbf{g}(\mathbf{x},\mathbf{y}) \tag{1}$$

where $\mathbf{x}$ is the vector of state variables (such as the generator rotor angles δ_i and rotor speed ω_i) and $\mathbf{y}$ is the vector of the algebraic variables (primarily the bus complex voltages). Next, the system can be linearized about the equilibrium point as:

$$\mathbf{\Delta\dot{x}} = \mathbf{A\Delta x} + \mathbf{B\Delta y} \tag{2}$$

$$\mathbf{0} = \mathbf{C\Delta x} + \mathbf{D\Delta y} \tag{3}$$

The variable $\boldsymbol{\Delta y}$ in (2) and can be substituted using (3) to derive a differential equation of only variable $\boldsymbol{\Delta x}$ as follows:

$$\mathbf{\Delta\dot{x}} = (\mathbf{A} - \mathbf{BD^{-1}C})\mathbf{\Delta x} \tag{4}$$

$$\mathbf{A_{sys}} := \mathbf{A} - \mathbf{BD^{-1}C} \tag{5}$$

$$\mathbf{\Delta\dot{x}} = \mathbf{A_{sys}\Delta x} \tag{6}$$

Equation (6) represents the deviation of the system's state away from the equilibrium point. As the result, small signal analysis is done by looking at the eigenvalues and other properties of $\mathbf{A_{sys}}$. For simplicity, from here let us call matrix $\mathbf{A_{sys}}$ as $\mathbf{A}$. The eigenvalues λ_i, i = 1..n correspond to the modes of the system and are the solutions of the following equation:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0} \tag{7}$$

Assume all the eigenvalues are distinct, for each λ_i there exists a right eigenvector $\mathbf{v_i}$ such that:

$$\mathbf{Av_i} = \lambda_\mathbf{i}\mathbf{v_i} \tag{8}$$

Similarly, for each eigenvalue there exists a left eigenvector $\mathbf{w_i}$ and the right and left eigenvectors are orthogonal.

$$\mathbf{w_i}^\mathbf{T}\mathbf{A} = \mathbf{w_i}^\mathbf{T}\lambda_\mathbf{i} \qquad \mathbf{A^T w_i} = \lambda_\mathbf{i}\mathbf{w_i} \tag{9}$$

Equation (6) needs to be decoupled to clarify the effect of the matrix $\mathbf{A}$'s parameters to the state vector $\mathbf{x}$. The decoupling can be conducted using the matrix of right and left eigenvectors. Let us define the modal matrices $\mathbf{V}$ and $\mathbf{W}$ as:

$$\mathbf{V} = \begin{bmatrix} \mathbf{v_1} & \dots & \mathbf{v_n} \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} \mathbf{w_1^T} \\ \dots \\ \mathbf{w_n^T} \end{bmatrix} \tag{10}$$

Equation (8) can be rewritten as:

$$\mathbf{A}V = V\mathbf{\Lambda} \tag{11}$$

where

$$\mathbf{\Lambda} = Diag(\lambda_i) \tag{12}$$

It follows that

$$\mathbf{V^{-1}AV} = \mathbf{\Lambda} \tag{13}$$

To decouple the variables, define vector $\mathbf{z}$ as

$$\mathbf{\Delta x} = \mathbf{Vz} \tag{14}$$

$$\mathbf{\Delta\dot{x}} = \mathbf{V\dot{z}} = \mathbf{A\Delta x} = \mathbf{AVz} \tag{15}$$

$$\mathbf{\dot{z}} = \mathbf{V^{-1}AVz} = \mathbf{\Lambda z} \tag{16}$$

Since $\mathbf{\Lambda}$ is a diagonal matrix, Equation (16) can be uncoupled as:

$$\dot{z}_i = \lambda_i z_i \tag{17}$$

Apply (17) on (14), the response $\mathbf{\Delta x}(t)$ can be rewritten as an equation of individual eigenvalues and right eigenvectors [18].

$$\mathbf{\Delta x}(t) = \sum_{i=1}^{n} \mathbf{v_i} z_i(0) e^{\lambda_i t} \tag{18}$$

B. Mode-Dependent Neuron Number Algorithm

As described in the previous section, modal analysis provides powerful insight into a specific power system and its response to various disturbances. We obtain information on the prominent modes of the system and how they dominate the response of certain system changes or events. These modes are unique to the component, system topology, and disturbance that are being studied. Thus, discovering the dominant modes for a particular situation provides information on the dominant behaviors and patterns present.

Within our neural network model, the neurons, or units, are "activated" depending on the input data presented. This activation is dependent on the type of function used as well as the input/output weights and biases, the latter two of which are the results of the neural network training and optimization. Ultimately, the activated units determine the output result. A specific set of units is activated for a particular set of input data. Different patterns or behaviors inherent within the input data are what differentiate the unit activation sets. The dominant modes of a dynamic system determine how the system will respond to a disturbance. The combination of these modes dictates the majority of the system response. In that case, what if the number of dominant modes could be equated to the number of units in the neural network?

The units and their combination provide the output result in the neural network. Our intuition from the power system and modal analysis motivates the hypothesis that the number of dominant modes represents the most significant patterns in the system and thus could capture the different behaviors successfully, similar to the function of neural network units. At the very least, the number of dominant modes provides an estimate of the number of units. We achieve this estimate with the process illustrated in Figure 2. First, we obtain the model

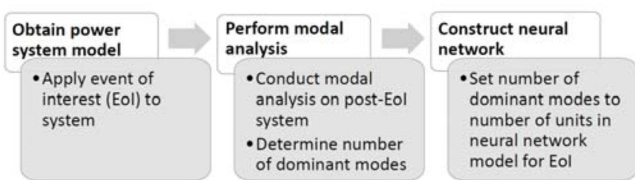


Fig. 2: Mode-dependent neuron number algorithm.

of the power system to analyze either mathematically or with power system simulation software. In our case, we model our system in PowerWorld [21], a simulation software, as we will study generator bus faults and the subsequent impact on rotor angle response.

The event of interest (EoI) that is being studied using the neural network, whether it be a change in load or a system fault, is applied to the system and modal analysis is conducted on the post-EoI system. Depending on the type of study being conducted, different components will vary the EoI parameters and modal analysis must be performed for each to obtain the most comprehensive result. In our example scenario, we only perform modal analysis once as we construct the neural network only for generators in various post-fault systems.

To determine the significant modes, methods using Prony analysis such as [22]–[24] can be used. We utilize the largest weighted percentage (LWP) values of each mode to determine the most dominant in the post-fault response, as calculated in PowerWorld. The LWP is the largest signal component in the mode weighted by time and is expressed as a percentage of total signal components. Subsequently, we analyze the LWP values by calculating the percent difference ($pDiff$) of the max LWP against the rest of the values, shown in (19); if $pDiff$ is less than the threshold difference of 50%, we count

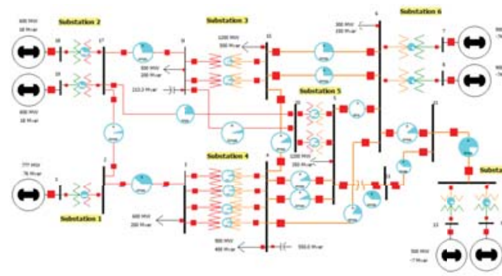


Fig. 3: EPRI 20-bus system [21].

that mode as dominant.

$$pDiff = \frac{LWP - max(LWP)}{max(LWP)} \tag{19}$$

Finally, we set the number of dominant modes as the number units in the hidden layer of our neural network.

IV. Evaluations

To demonstrate the mode-dependent neuron selection algorithm, we studied a neural network model developed for power system generators. The ANN takes input of the generator real power and electric field voltage and provides the rotor angle response after a balanced three-phase fault on a generator bus (all time-series data). Thus, we achieve the response without requiring the complex generator and system model. The input data selection is based on the experiment performed in [2]. The location of this fault varies, excluding the slack bus, and the clearing time also varies. For the training data set, a three-phase fault was applied at each generator bus and data was collected for 20 different clearing times (up to the critical clearing time). These faults were simulated in PowerWorld and the data was obtained with the transient stability toolbox [21].

A nonlinear autoregressive network with exogenous inputs (NARX) with one hidden layer is constructed with the training data set and the mode-dependent estimate of units is used and compared against other random or heuristic-based values. The NARX feedback neural network is often utilized for time-series prediction, as is our goal, and is described further in [25]. The post-fault generator rotor angle scenario is applied to the EPRI 20-bus system shown in Figure 4. The system has 7 generators of which Generator #1 (at bus 1) is the slack bus and is excluded. Therefore, faults are simulated and data is obtained from 6 generators at 20 different clearing times for each. The mode-dependent neuron number algorithm, summarized in Figure 2, is applied and Table I displays an example set of resultant modes for a fault at Generator #2 (at bus 7) (number of dominant modes similar across generator buses). The dominant modes are highlighted in yellow, as calculated with (19). With 4 dominant modes, we equate to the number of units in the NARX network, as illustrated in Table I.

We calculate estimates for number of units from known heuristics summarized and discussed in [16]; Table II lists these estimates for our generator bus fault scenario. Figure 5 illustrates the average mean squared error (MSE) that represents the difference between the actual rotor angle response

TABLE I: Resultant Modes and Largest Weighted Percentage

LWP (%)	pDiff (%)	Mode
69.945	0	0.0852
67.8717	0.029	-0.3936
61.021	0.128	-0.2199
52.0739	0.256	-2.0416
35.2362	0.496	-2.6339
29.078	0.584	-0.5904
8.9668	0.872	-1.5463

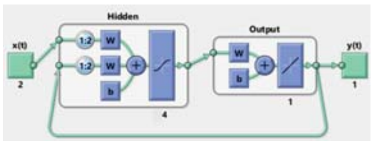


Fig. 4: NARX neural network with 4 units [26].

(from simulation) and predicted rotor angle response (from the NARX network) for up to 1000 training iterations, in 100 iteration intervals, for various numbers of units (comparing our modal estimate against heuristics), and testing 20 different clearing times at Generator #2 (at bus 7). The mean MSE for each given number of units indicates lower values for 1 and 2 units, similar values for 4 and 5 units, and high error for 3 and 12 units. The mean MSEs for 1 and 2 units are misleading as the model prediction for the rotor angle response is inaccurate, and the low MSE results from consistently producing a relatively flat line through the true rotor angle response as illustrated in Figure 6a;it is unsuccessful in capturing all variance in the data.

However, our estimate of 4 units provides a decent estimate, an example of which is shown in Figure 6c, with comparatively low MSE error while capturing the variations in the actual response. Figures 6a-6d show examples of the NARX model's prediction of the rotor angle response for faults at Generator 2 (at bus 7) for different clearing times, number of units, and training iterations. The average MSE of all clearing time and training iterations is represented in the comprehensive Figure 5. Figure 6d illustrates an overfitting situation with too many units, resulting in the spikes.

Yet, the NARX feedback neural network with 4 units does not provide good performance in predicting the rotor angle response and results in generally high MSE. The NARX network was a first-step selection to explore the mode-dependent neuron number algorithm and provided a good base in that respect. To improve our actual model prediction, to explore further in future work with the algorithm, we began testing with a layer recurrent neural network (LRNN) architecture in which each layer has a recurrent connection with a tap delay associated with it; essentially, the network is enabled to have infinite dynamic response to time-series input data [26]. The rotor angle response is greatly improved with this network, with our estimate of 4 units, and is represented in Figure 7a.

The overall performance of this network is shown Figure 7b where training is performed with 70% of the data, testing and validation with 15% each. This cross-validation allows for the elimination of overfitting or underfitting issues [27]. An example testing result (fault at Generator #2 (at bus 7)) is shown in Figure 7c where a close prediction of the rotor angle response is achieved as well as an acceptably low MSE. Finally, preliminary results comparing the average MSE between unit estimates are shown in Figure 7d where the LRNN achieves significantly lower MSE and the mode-dependent estimate of 4 units performs best.

TABLE II: Heuristic Unit Number Methods [16]

Heuristic Method	Unit Estimate
Li et al. method	2
Tamura and Tateishi method	1
Zhang method	3
Jinchuan and Xinzhe method	2
Shibata and Ikeda method	1
Hunter et al. method	3
Sheela and Deepa method	5

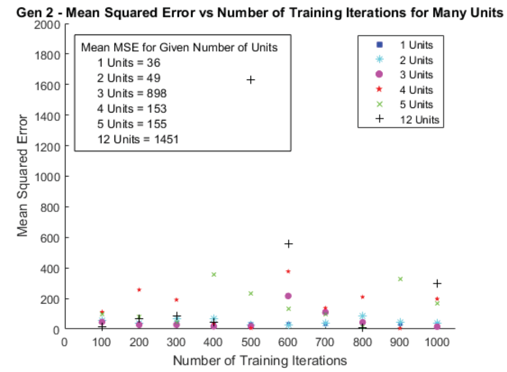


Fig. 5: NARX: Average MSE for different unit number estimates over many training iterations and clearing times for faults at Generator #2 (at bus 7); the mode-dependent algorithm estimate is 4 units.

V. Conclusions and Future Work

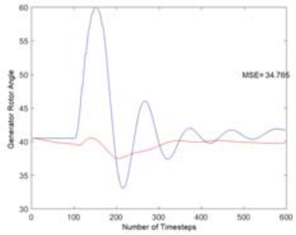
Through experimental testing, this paper tested a hypothesis that the number of dominant power system modes for a particular event can be equated to the number of units, or neurons, for construction of a neural network modeling the event and its characteristics. Our results indicate that the mode-dependent estimate of number of units provides promising performance, especially compared to known, generalized heuristics. We seek to develop systematic methods for power system neural network construction that leverages power system analyses and known behaviors. This initial work will be extended to mathematically formulate the relationship between the neural network model and power system modes; future work will also explore varying disturbance types and subsequent dominant mode impact, different neural network architectures, increased data set size and input sources, and larger systems. In this manner, trial-and-error methods can be eliminated and enable systematic, domain-dependent construction of effective artificial neural network models in power systems.

References

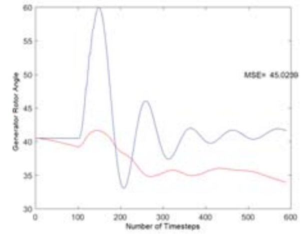
[1] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, Feb 2001.

[2] A. Bahbah and A. Girgis, "New method for generators' angles and angular velocities prediction for transient stability assessment of multimachine power systems using recurrent artificial neural network," in *IEEE Power Engineering Society General Meeting, 2004.*, June 2004, pp. 117 Vol.1–.

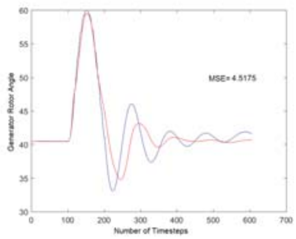
[3] E. M. Voumvoulakis, A. E. Gavoyiannis, and N. D. Hatziargyriou, "Application of machine learning on power system dynamic security assessment," in *2007 International Conference on Intelligent Systems Applications to Power Systems*, Nov 2007, pp. 1–6.



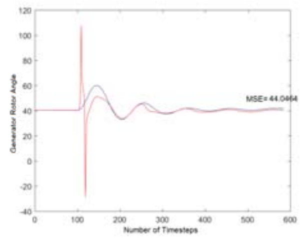
(a) Number of Units = 1



(b) Number of Units = 2

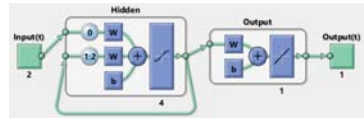


(c) Number of Units = 4

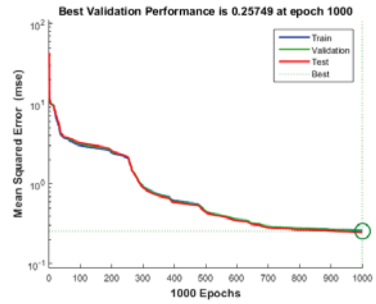


(d) Number of Units = 12

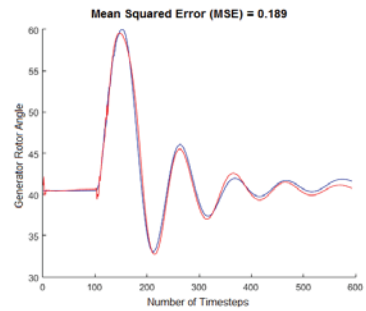
Fig. 6: Comparison of true rotor angle response (blue line) with NARX network prediction (red line) for various numbers of units.



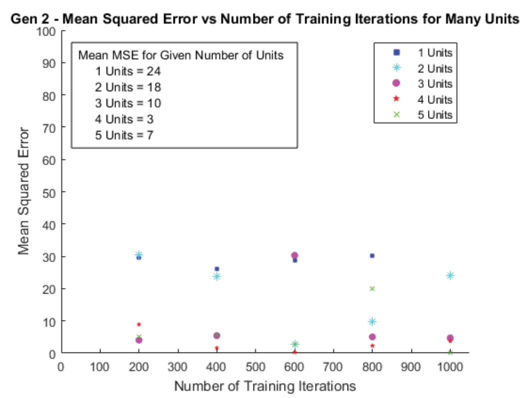
(a) Layer recurrent neural network with 4 units [26].



(b) Overall performance of LRNN with 4 units.



(c) True rotor angle response and LRNN prediction for 4 units.



(d) LRNN: Average MSE for different unit number estimates for faults at Generator #2 (at bus 7); the modal estimate is 4 units.

Fig. 7: Layer recurrent neural network (LRNN) results.

[4] M. Kazerooni, H. Zhu, and T. J. Overbye, "Literature review on the applications of data mining in power systems," in *2014 Power and Energy Conference at Illinois (PECI)*, Feb 2014, pp. 1–8.

[5] R. Aggarwal and Y. Song, "Artificial neural networks in power systems. iii. examples of applications in power systems," *Power Engineering Journal*, vol. 12, no. 6, pp. 279–287, Dec 1998.

[6] M. A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2015.

[7] M. Caudill, "Neural networks primer, part i," *AI Expert*, vol. 2, no. 12, pp. 46–52, Dec. 1987. [Online]. Available: http://dl.acm.org/citation.cfm?id=38292.38295

[8] T. Xu, B. Venkatesh, C. Opathella, and B. N. Singh, "Artificial neural network model of photovoltaic generator for power flow analysis in PSS SINCAL," *IET Generation, Transmission Distribution*, vol. 8, no. 7, pp. 1346–1353, 2014.

[9] A. Qian, S. Shande, and Z. Shouzhen, "Transient stability study using artificial neural networks models of generator, excitation system, governor," in *Power System Technology, 1998. Proceedings. POWERCON '98. 1998 International Conference on*, vol. 2, Aug 1998, pp. 1331–1335 vol.2.

[10] G. Rogers, *Modal Analysis of Power Systems*. Boston, MA: Springer US, 2000, pp. 31–73. [Online]. Available: http://dx.doi.org/10.1007/978-1-4615-4561-3_3

[11] S. Hossain-McKenzie, K. K. Han, and K. Davis, "Generator control action classification based on localized voltage measurements," in *2016 North American Power Symposium (NAPS)*, Sept 2016, pp. 1–6.

[12] R. Aggarwal and Y. Song, "Artificial neural networks in power systems. iii. examples of applications in power systems," *Power Engineering Journal*, vol. 12, no. 6, pp. 279–287, Dec 1998.

[13] D. P. Kothari, "Application of neural networks to power systems," in *Proceedings of IEEE International Conference on Industrial Technology 2000 (IEEE Cat. No.00TH8482)*, vol. 1, Jan 2000, pp. 621–626 vol.2.

[14] J.-W. Park, R. Harley, and G. Venayagamoorthy, "Comparison of mlp and rbf neural networks using deviation signals for on-line identification of a synchronous generator," in *Power Engineering Society Winter Meeting, 2002. IEEE*, vol. 1. IEEE, 2002, pp. 274–279.

[15] L. Thomas, M. K. M. V, and A. B, "Discovery of optimal neurons and hidden layers in feed-forward neural network," in *2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech)*, Aug 2016, pp. 286–291.

[16] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[17] M. D. Blechner, "Behavior of various machine learning models in the face of noisy data," *Harvard-MIT Division of Health Sciences and Technology, Final Project*, 2005.

[18] P. W. Sauer and M. Pai, "Power system dynamics and stability," *Urbana*, 1998.

[19] J. Hauer, "Application of prony analysis to the determination of modal content and equivalent models for measured power system response," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 1062–1068, 1991.

[20] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[21] PowerWorld Corporation. (2016) Exercises for Students. http://www.powerworld.com/solutions/excercises.

[22] J. F. Hauer, C. Demeure, and L. Scharf, "Initial results in prony analysis of power system response signals," *IEEE Transactions on power systems*, vol. 5, no. 1, pp. 80–89, 1990.

[23] A. Cagnano, E. D. Tuglie, and F. Torelli, "Estimation of power system dominant modes," in *2009 IEEE Bucharest PowerTech*, June 2009, pp. 1–7.

[24] L. Ding, A. Xue, F. Han, J. Li, M. Wang, T. Bi, and J. Wang, "Dominant mode identification for low frequency oscillations of power systems based on prony algorithm," in *2010 5th International Conference on Critical Infrastructure (CRIS)*, Sept 2010, pp. 1–6.

[25] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent narx neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 27, no. 2, pp. 208–215, 1997.

[26] MathWorks Toolbox Release 2017a. (2017) Neural network toolbox. The MathWorks, Inc. Natick, Massachusetts, United States.

[27] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.