Short Term Load Forecasting in Greek Interconnected Power System using ANN: a Study for Output Variables

G.J. TSEKOURAS[1], F.D. KANELLOS[2], CH.N.ELIAS[3], V.T. KONTARGYRI[3],C.D.TSIREKIS[2], I.S.KARANASIOU[3,4], A.D.SALIS[3], P.A.CONTAXIS[3], A.A. GIALKETSI[3], N.E. MASTORAKIS[1]

[1]Department of Electrical & Computer Science, Hellenic Naval Academy
Terma Hatzikyriakou, Piraeus

[2]Hellenic Transmission System Operator
Kastoros 72, Piraeus

[3]School of Electrical and Computer Engineering, National Technical University of Athens
9 Heroon Polytechniou Street, Zografou, Athens

[4]Department of Computer Science, Hellenic Army Academy
Hellenic Army Academy, Vari, Athens

GREECE

Email: tsekouras_george_j@yahoo.gr, kanellos@mail.ntua.gr, xelia@hlk.forthnet.gr, vkont@central.ntua.gr, ktsirekis@desmie.gr, ikaran@esd.ece.ntua.gr, anastasios.salis@gmail.com, pkont@mail.ntua.gr, agialketsi@gmail.com, mastor@wseas.org,

Abstract: - The purpose of this paper is to compare the performance of different structures of Artificial Neural Networks (ANNs) regarding the output variables used for short term forecasting of the next day load of the interconnected Greek power system. In all cases the output variables are the hourly actual loads of the next day. The classical ANN design adopts an ANN model with 24 output variables. Alternatively, 24 different ANN models can be implemented for each hour of the day. This solution can affect the selection of input variables indirectly. In this paper, various scenarios of the solution of 24 different ANN models are going to be studied with different sets of input variables using the scaled conjugate gradient training algorithm, for which a calibration process is conducted regarding the crucial parameters values, such as the number of neurons, the type of activation functions, etc. The performance of each structure is evaluated by the Mean Absolute Percentage Error (MAPE) between the experimental measurements and estimated values of the hourly load demand of the next day for the evaluation set in order to specify the optimal ANN. Next, the load demand for the next day of the test set (with the historical data of the current year) is estimated using the best ANN structure, to verify the behaviour of ANN load prediction techniques. Finally the classical design and different proposed structures are compared.

Key-Words: - artificial neural networks, output variables' analysis, short-term load forecasting, ANN training scaled conjugate gradient algorithm

1 Introduction

In a deregulated electric energy market, the accuracy of short-term load forecasting (for the next few hours to a week with the usual time step of an hour) affects operational issues of power systems, such as unit commitment, scheduling of spinning reserve, available transfer capability, system stability, application of load demand management programs, etc succeeding higher reliability and lower operational costs for power systems [1]. Several forecasting methods have been implemented with quite satisfactory results, such as regression models [2], auto-regressive models [3], ANNs [1], fuzzy logic [4], hybrid systems [5] etc. Specifically, in Greece, ANNs have been used successfully either for the interconnected power system [1, 6-7], or autonomous big islands [6, 8-9]. Some techniques belong to classical ANNs [1, 6-8] or specialized ones [9] or hybrid ones with fuzzy logic algorithms [10]. During last three years an analytical study of the application of ANNs models has been presented by the authors for the interconnected Greek power system with different aspects, which are: (1) the comparison of various training algorithms, such as stochastic training process, batch process, scaled conjugate gradient algorithm, resilient algorithm, quasi-Newton algorithm, Levenberg-Marquardt algorithm [1], (2) the effectiveness of the input variables [11], (3) the estimation of the confidence interval [12-13].

In this paper the effect of the output variables on the ANNs design is examined. The basic structure of the ANN proposed by Kiartzis et al [6-7] for the inputs and outputs variables for the Greek interconnected power system is used, while the training algorithm is the scaled conjugate gradient [14] according to the results obtained in [1]. Generally, the output variables are the hourly loads of the next day, so an ANN structure with 24 outputs is formed. Alternatively, Hippert et al [15] have proposed the investigation of different ANNs design: the development of different ANN per each output variable. This modification may also cause the respective changes to the set of the input variables. The main goals are:

- the modulation of the internal neural network structure (number of neurons of hidden layer, kind of activation functions, etc.) for each different combination of input and output variables with respect to the smallest *Mean Absolute Percentage Error* (*MAPE*) of the evaluation set and
- the comparison of the respective cases in terms of *MAPE* and computational time.

The respective results are based on actual hour load data of the Greek interconnected power system for years 1997-2000.

2 Proposed ANN Methodology for Short-term Load Forecasting

The short-term load forecasting is achieved by the application of an ANN; its parameters are properly obtained with the scaled conjugate gradient algorithm. This methodology is similar to that presented in [11] with the modification of an additional step for the selection of the output variables, as it is shown in the flow chart of Fig. 1. The methodology includes the following basic steps.

(a) *Output variables selection*: In this step the selection of the output variables is done. Specifically, either one ANN model using 24 output variables for the hourly loads of the next day can be constructed, or 24 ANNs for each hourly load of the next day can be formed.

(b) *Data selection*: The input variables for load forecasting of the *d*-th day are formed. The 1st scenario is the basic one according to Kiartzis et al [6-7], Tsekouras et al [1, 11-13], where the output variables are the 24 hourly actual load demand of the current day, while the input vector comprises 71 elements:

(1) the hourly actual loads of the two previous days,

(2) the maximum mean temperature per three hours and the minimum mean temperature per three hours for Athens and for Thessalonica, for the current and the previous day,

(3) the temperature difference between the maximum mean temperature per three hours of the current day and the respective one of the last day for Athens and Thessalonica,

(4) the temperature dispersion from comfortable living conditions temperature for Athens and for Thessalonica, for the current and the previous day,

(5) seven-digit binary numbers, which express the kind of the week day,

(6) two sinusoidal functions $\cos(2\pi d/T)$ and $\sin(2\pi d/T)$), which express the seasonal behavior of the current day, where T is the number of the days of the current year.

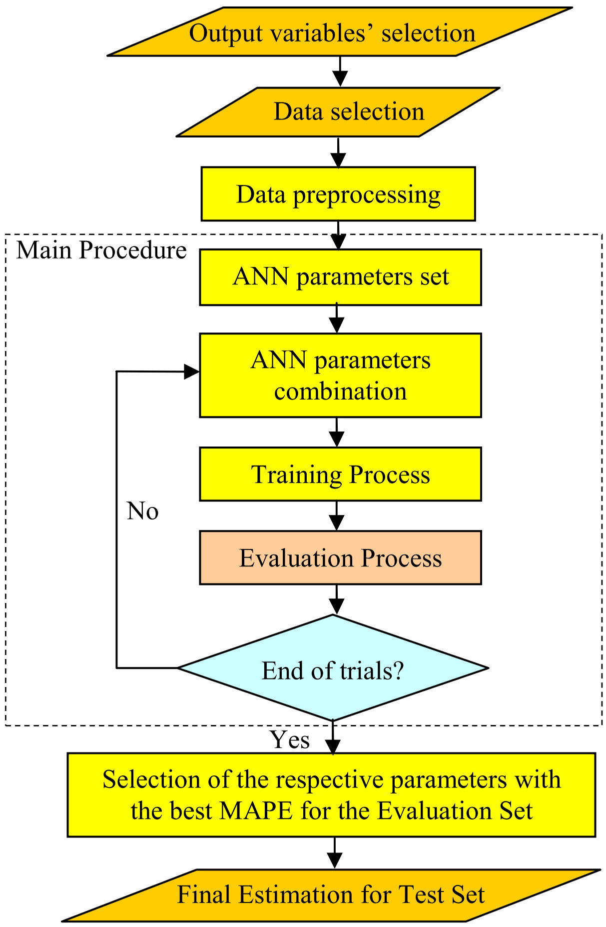

Fig. 1. Flowchart of the ANN methodology for the proper selection of ANN parameters for short-term load forecasting for different combinations of input variables – output variables

In all other scenarios 24 ANN models are constructed (one per each hourly load). The respective sets of the input variables are the following:

- 2nd scenario: It is the same with the 1st scenario except of using two sinusoidal functions ($\cos(2\pi d/7)$, $\sin(2\pi d/7)$) instead of the seven digit binary numbers for the kind of the week day. In this case the input vector comprises 66 elements.
- 3rd scenario: It is the same with the 2nd scenario, but Principal Components Analysis is applied to the input variables except from the kind of the week day and the season of the year, decreasing their number from 62 to 15 for 99% description percentage of data dispersion according to Kaiser's criterion [16]. The total number of the input variables is 19.
- 4th scenario: For the formation of the ANN forecasting model of i^{th} hour, the hourly actual loads of i, i-1, i-2 hours of the two previous days are used, as well as the respective mean temperature per three hours for Athens and for Thessalonica, for the current and the previous day, two sinusoidal functions for the kind of the week day and two sinusoidal functions for the seasonal behavior of the current day. In this case the input vector comprises 14 elements.
- 5th scenario: It is the same with the 4th scenario, but Principal Components Analysis is applied to the input variables except the variables of the kind of the week day and the season of the year, obtaining 99% description percentage of data dispersion according to Kaiser's criterion. The total number of the input variables is different for each ANN per hour.

(c) *Data preprocessing*: Data are examined, in order to modify or delete the outliers. This process is known as noise suppression. Non linear activation functions, such as hyperbolic tangent, are preferably used because of the great non linearity of the problem. However, in this case, saturation problems may occur. In order to avoid saturation problems, the input and the output values are normalized, especially, the load and the temperature data.

(d) *Main procedure*: The ANN is trained using the scaled conjugate gradient algorithm, whose basic steps have been presented analytically in [11, 15]. The respective parameters of the neural network are selected after a set of trials. For each ANN parameter, such as the number of neurons of the hidden layer, the type of the activation functions (hyperbolic tangent, logistic, linear), the parameters of the activation functions, the maximum number of epochs, the SCGA algorithm parameters, the training algorithm is separately executed for the respective range of values (i.e. 20 to 70 neurons with step 1 for the 1st scenario) based on the error function (sum of the square of errors for all neurons per epoch) for the training set. Meanwhile the regions with satisfactory results (minimum Mean Absolute Percentage Error (*MAPE)* for evaluation set) are identified. It is noted that the *MAPE* index between the measured and the estimated values of hourly load demand for the evaluation set' days is given by:

$$MAPE_{ev} = 100\% \cdot \frac{1}{m_{ev}} \cdot \sum_{d=1}^{m_{ev}} \sum_{i=1}^{24} \frac{\left|\hat{L}(d,i) - L(d,i)\right|}{24 \cdot L(d,i)} \quad (1)$$

where $L(d,i)$ is the measured value of load demand for the i-th hour of d-th day of the evaluation set, $\hat{L}(d,i)$ the respective estimated value, m_{ev} the population of the evaluation set. For the case of *MAPE* of the i^{th} hourly load demand for separate ANN model per hour, the respective index is given by:

$$MAPE(i) = 100\% \cdot \frac{1}{m_{ev}} \cdot \sum_{d=1}^{m_{ev}} \frac{\left|\hat{L}(d,i) - L(d,i)\right|}{L(d,i)} \quad (2)$$

The *MAPE* index is a practical measure, which reflects the approximation of the actual load demand independently from its measurement units.

Following, the training algorithm is repeatedly executed, while all parameters are simultaneously adjusted to their respective regions, so that the combination with the smallest *MAPE* for the evaluation set is selected.

(e) *Final estimation for the test set*: The actual load demand for the days of the test set is finally estimated by using the respective ANN parameters.

Afterwards, the results of each scenario with the respective optimized parameters are compared, in order to choose the one leading to the smallest *MAPE* index within acceptable computational time.

3 Application of Short-term Load Forecasting in Interconnected Greek Power System based on ANN STLF Methodology

Following, the aforementioned method is applied for short-term load forecasting in Greek intercontinental power system. The training and the evaluation sets consist of the 90% and 10% of the normal days (no holidays) of the years 1997-1999, respectively, while the respective test set consists of the normal days of the year 2000.

According to the 1st scenario the input vector comprises 71 input variables, while the output vector is formed by the normalized 24 output actual hourly load demands of the day under prediction. The ANN parameters, which are calibrated, are the

ISBN: 978-1-61804-023-7

following:

- the number of the neurons of the hidden layer, which ranges from 20 to 70 with incremental step of 1 neuron,
- the set of values of parameters σ and λ_0 which comprises 10^{-3}, 10^{-4}, 10^{-5} and 10^{-6}, 10^{-7}, $5 \cdot 10^{-8}$ respectively,
- the type and the parameters of the activation functions of the hidden and the output layers, where the type can be *hyperbolic tangent*, *linear or logistic*, while the a_1, a_2 parameters get values from the set {0.1,0.2,…, 0.5} and b_1, b_2 from the set $\{0.0, \pm 0.1, \pm 0.2\}$.

The parameters of the stopping criteria are defined after a few trials as $\max\_epochs$=5000, $\ell imit_1$=10^{-5}, $\ell imit_2$=10^{-5}.

Firstly, each parameter is calibrated i.e. the number of neurons varies from 20 to 70, while the remaining parameters are assigned with fixed values. The model with the smallest *MAPE* index is selected for the respective parameter, which can be not only a value of neuron but a region (here the respective region of neurons with satisfactory results is [43, 52]). This process is repeated for all ANN parameters and at the end the final calibration of the ANN model is done, where each ANN parameter is slightly varied in the respective region with satisfactory results. The best result for the *MAPE* index of the evaluation set is 1.487%, while the *MAPE* indexes for the training and the test sets are 1.294% and 1.781%, respectively. These are obtained for an ANN with 52 neurons in the hidden layer using hyperbolic tangent activation function in both layers (tanh(0.5·x) for the hidden layer, tanh(0.25·x) for the output layer, where x is the respective sum of the properly weighted inputs of the neuron). More details can be found in [11, paragraph 3].

This methodology is also applied to the four different scenarios for the short-term load forecasting in Greek interconnected power system, as it is described in section 2. A different ANN model is formed for each hour and each scenario. In Table 1, the respective results of *MAPE* of training, evaluation and test sets for the ANN models of 12:00 load for the four different scenarios and the respective calibrated parameters are given. It is noted that the parameters of the training algorithm namely σ and λ_0, the type and the parameters of the activation functions of the hidden and the output layers take the same values with the respective ones of the 1st scenario. On the contrary the range of the number of the neurons of the hidden layer has been extended properly, as the input vector's dimension has been modified.

In Fig. 2 and Table 2 the respective results of *MAPE* of training, evaluation and test sets from all scenarios are given. The mean *MAPE* is calculated for the scenarios with 24 different ANN models.

The analogy of the respective computational time (with the proper parameters calibration) for the five scenarios is: 1÷0.4÷0.05÷0.15÷0.03. It becomes apparent that the respective computational time is decreased significantly, as the dimension of each ANN model is decreased. The respective time could be decreased more for the last 4 scenarios, if 24 separate computers were used (parallel operation of computers), one per each ANN model.

The best results of *MAPE* for evaluation set are obtained with the 1st scenario. The second and the third scenario lead to slightly worse results than the 1st one. But the results of *MAPE* obtained for the test set are quite similar for the 1st and 2nd scenario, which means that the 2nd scenario is satisfying. The subtraction of input data (4th scenario) leads to worse results.

The use of compression technique (3rd and 5th scenarios) decreases the computational time significantly, but the performance of the *MAPE* for the evaluation set is worse, as the respective MAPE is decreased by 3% between 2nd scenario and 3rd one (2nd with PCA) and by 10% between 4th scenario and 5th one (4th with PCA). The respective performance of the *MAPE* for the test set is more worse. Because of this performance the data compression is not suggested.

TABLE 1
MAPE(%) OF TRAINING, EVALUATION & TEST SETS FOR 12:00 LOAD OF THE NEXT DAY FOR 4 DIFFERENT SCENARIOS WITH THE RESPECTIVE PROPERLY CALIBRATED PARAMETERS

No. of scenario	MAPE(%) of training set	MAPE(%) of evaluation set	MAPE(%) of test set	Neurons – Range of examined neurons	Activation functions
2	1.347	1.824	1.847	32 (10-60)	f_1=tanh(*0.50*x), f_2=tanh(*0.25*x)
3	1.492	1.876	1.924	17 (2-25)	f_1=tanh(*0.50*x), f_2=tanh(*0.25*x)
4	1.747	2.210	2.620	40 (10-60)	f_1=tanh(*0.25*x), f_2=tanh(*0.25*x)
5	2.120	2.404	2.953	14 (2-25)	f_1=tanh(*0.50*x), f_o=tanh(*0.25*x)

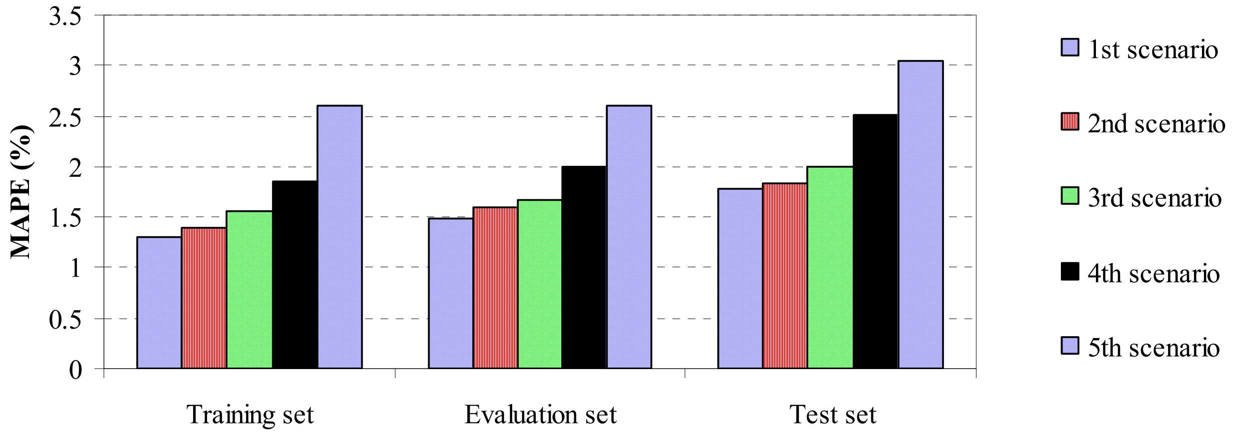

Fig. 2. MAPE (%) of training, evaluation and test set for load forecasting of the next day for interconnected Greek power system for 5 different scenarios of output variables

TABLE 2
MAPE(%) OF TRAINING, EVALUATION & TEST SETS OF THE NEXT DAY FOR 5 DIFFERENT SCENARIOS

No. of scenario	MAPE(%) of training set	MAPE(%) of evaluation set	MAPE(%) of test set
1	1.294	1.487	1.781
2	1.390	1.603	1.830
3	1.565	1.667	1.994
4	1.850	1.989	2.503
5	2.596	2.598	3.043

6 Conclusions

This paper compares the performance of five different ANN structures regarding the output variables for short-term load forecasting in Greek interconnected power system. The basic structure of ANN model is determined by Kiartzis et al [6-7] and Tsekouras [1, 11], where the output variables are the 24 hourly actual load demand of the current day, while the input vector has 71 variables. The other scenarios are based on the formation of 24 different ANN models: one per each hourly load. Afterwards, the input variables have been studied properly. Specifically, the input variables of the 2nd scenario are the same with the basic one except of using two sinusoidal functions for the kind of the week day. In the 3rd scenario the input variables of the 2nd scenario have been modified through principal component analysis decreasing the respective input dimension for ANN model. The input variables of the 4th scenario have been limited to the respective historical data of neighbourhood hourly load and temperature of the hour under prediction decreasing the dimension of the input vector to 19 variables. In the 5th scenario the input variables of the 4th scenario have been data suppressed through principal component analysis. The training algorithm used is the scaled conjugate gradient one, which has been calibrated properly through an extensive search for the number of neurons of the hidden layer, the type and the parameters of the activation functions, etc. The performance of each scenario is measured by the Mean Absolute Percentage Error of the evaluation set. Finally, the formations of a single ANN model with 24 output variables and of 24 ANN separate models (each one for hourly load) will give the same results practically, if the same input variables are used. In fact the basic scenario is slightly better, but it requires much more computational time.

Acknowledgements

The authors want to express their sincere gratitude to C. Anastasopoulos, D. Voumboulakis and P. Eustathiou from PPC for the supply of all the necessary data for this application.

References:

[1] G. J. Tsekouras, F.D. Kanellos, V.T. Kontargyri, C.D. Tsirekis, I.S. Karanasiou, Ch.N. Elias, A.D. Salis, N.E. Mastorakis. A comparison of Artificial Neural Networks algorithms for short term load forecasting in Greek intercontinental power system. *WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing,* (CSECS '08), Puerto De La Cruz, Canary Islands, Spain, December 15-17, 2008.

[2] T. Haida, S. Muto. Regression based peak load forecasting using a transformation technique, *IEEE Trans. on Power Systems,* Vol. 9, No. 4, 1994, pp. 1788 – 1794.

[3] H.T. Yang, C.M. Huang, C.L. Huang. Identification of ARMAX model for short term load forecasting: an evolutionary programming approach. *IEEE Trans. on Power Systems,* Vol. 11, No. 1, 1996, pp. 403 – 408.

[4] P.A.Mastorocostas, J.B.Theocharis, A.G. Bakirtzis. Fuzzy modeling for short-term load forecasting using the orthogonal least squares method. *IEEE Trans. on Power Systems,* Vol. 14, No. 1, 1999, pp. 29–36.

[5] N. Amjady, F. Keynia. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. *Energy*, 34 (1) pp. 46-57, 2009.

[6] S.J.Kiartzis. *Artificial Intelligence Applications on Short-term Load Forecasting*. Dissertation, Thessaloniki, A.U.Th., 1998 (in greek).

[7] A.G. Bakirtzis, V. Petridis, S.J. Kiartzis, M.C. Alexiadis, A.H. Maissis. A neural network short term load forecasting model for the Greek power system. *IEEE Trans. on Power Systems*, Vol. 11, No.2, 1996, pp. 858-863.

[8] S.J. Kiartzis, C.E. Zournas, J.M. Theocharis, A.G. Bakirtzis, V. Petridis. Short term load forecasting in an autonomous power system using artificial neural networks. *IEEE Trans. on Power Systems*, Vol. 12, No.4, 1997, pp. 1591-1596.

[9] K. Kalaitzakis, G.S. Stavrakakis, E.M. Anagnostakis. Short-term load forecasting based on artificial neural networks parallel implementation. *Electric Power Systems Research*, Vol. 63, 2002, pp. 185-196.

[10] A.G. Bakirtzis, J.B. Theocharis, S.J.Kiartzis, K.J.Satsios. Short term load forecasting using fuzzy neural networks. *IEEE Trans. on Power Systems*, Vol. 10, No. 3, 1995, pp. 1518-1524.

[11] G. J. Tsekouras, F.D. Kanellos, V.T. Kontargyri, C.D. Tsirekis, I.S. Karanasiou, Ch.N. Elias, A.D. Salis, P.A. Kontaxis, N.E. Mastorakis. Short Term Load Forecasting in Greek Intercontinental Power System using ANNs: a Study for Input Variables. *10th WSEAS International Conference on Neural Networks*, (NN '09), Prague, Czech Republic, March 23-25, 2009.

[12] N.E. Mastorakis, G. J. Tsekouras. Short Term Load Forecasting in Greek Intercontinental Power System using ANN: The Confidence Interval. *Advanced Aspects of Theoretical Electrical Engineering, Sozopol 2010*, 19.09.10 – 22.09.10, Sozopol, Bulgaria.

[13] G. J. Tsekouras, N.E. Mastorakis, F.D. Kanellos, V.T. Kontargyri, C.D. Tsirekis, I.S. Karanasiou, Ch.N. Elias, A.D. Salis, P.A. Kontaxis, A.A. Gialketsi. Short term load forecasting in Greek interconnected power system using ANN: Confidence Interval using a novel re-sampling technique with corrective Factor. *WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing*, (CSECS '10), Vouliagmeni, Athens, Greece, December 29-31, 2010.

[14] M.F. Moller. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, Vol. 6, 1993, pp. 525-533.

[15] H.S. Hippert, C.E. Pedreira, R.C. Souza. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. on Power Systems,* Vol. 16, No. 1, 2001, pp. 44-55.

[16] L.M. Saini, M.K. Soni. Artificial neural network-based peak load forecasting using conjugate gradient methods. *IEEE Trans. on Power Systems*, Vol. 17, No 3, 2002, pp.907-912.

ISBN: 978-1-61804-023-7