Advanced Method for
Small-Signal Stability Assessment
based on Neuronal Networks

S. P. Teeuwsen, I. Erlich, *Member, IEEE,* and M. A. El-Sharkawi, *Fellow, IEEE*

Abstract-- This paper deals with a new method for eigenvalue prediction of critical stability modes of power systems based on neural networks. The special interest of this paper is on inter-area oscillations of large-scale interconnected power systems. The existing methods for eigenvalue computations are time-consuming and require the entire system model that includes an extensive number of state variables. After reducing the input space and proper training of the neural network, the stability condition of the power system can be predicted with high accuracy. Furthermore, the new advanced method allows to predict a small variable wide band for the real part of eigenvalues.

Index Terms — Eigenvalues, Inter-Area Oscillations, Neural Networks, Small-Signal Stability

I. INTRODUCTION

Inter-area oscillations in large-scale power systems are becoming more common especially for the European Interconnected Power System UCTE/CENTREL. The system has grown very fast in a short period of time due to the recent eastward expansion. This extensive interconnection alters the stability region of the network, and the system experiences inter-area oscillations associated with the swinging of many machines in one part of the system against machines in other parts. Moreover, for certain load flow conditions, the system damping can widely change [1-2].

With the deregulation of the electricity markets in Europe, the utilities are allowed to sell their generated power outside their traditional borders and compete directly for customers. For economical reasons, the operators are often forced to steer the system closer to the stability limits. Thus, the operators need different computational tools for predicting system stability. These tools must be accurate and fast to allow on-line stability assessment.

The computation of the small signal stability is a time consuming process for large networks, which includes the load flow computation, the linearization at the operating point, and the eigenvalue computation [3].

An alternative method is to use a neural network (NN) trained with off-line data for different load flow and system conditions. By using NN, a fast computation of the eigenvalues is possible, provided that the network is properly designed and trained.

Recent NN approaches showed highly accurate results regarding the eigenvalue prediction of large-scale power systems [4]. However, the NNs are trained only for narrow operating conditions of the power system. In fact, when the operating point of a given power system changes, the corresponding eigenvalues may shift significantly. Small changes in the network topology, the load flow, or the voltage level may change the eigenvalues within the weakly damped area too.

The NN introduced in [4] predicts the eigenvalues themselves. But a NN that is able to manage all system variations must be designed for a variable number of eigenvalues. Hence, the NN must have a variable structure, which is a difficult task. To address this issue, the assessment method proposed in this paper is based not on the prediction of eigenvalues, but on the prediction of regions with one or more dominant eigenvalues of insufficient damping.

To generate the training data for the NN, an observation area is defined within the complex eigenvalue space. This area is located at a low damping level where inter-area eigenvalues typically occur. Then, the entire observation area is sampled. The distances between the sample points and the eigenvalues are computed and the sample points are activated depending on the existence of dominant eigenvalues. The closer the eigenvalue to a sample point the higher the activation. Once this is computed for the entire set of patterns, the NN is trained with these activations.

After proper training, the NN can be used in real time to compute the sample activations. Then, the activations are transformed into area ranges where the eigenvalues are located. These areas are characterized by activations higher than a given limit.

II. 16-MACHINE DYNAMIC TEST SYSTEM

The PST16 System used in this study is a dynamic 16-machine test network. The main focus is on modeling of power system dynamics and inter-area oscillations in the time range of a few seconds to minutes.

The PST16 System is modeled by using real data from generators and controllers for different units such as hydro,

nuclear and thermal power plants. The system consists of 3 strongly meshed areas, each having 5-6 generators. These areas are connected partly by weak transmission lines. Thus, inter-area oscillations can occur. The system is designed for power exchange between these areas. The first area (Area A) contains mostly hydro power plants and is considered to be a power exporting area. Area B and area C are load demanding areas that import power from area A. The total generation of the system is about 16 GW.

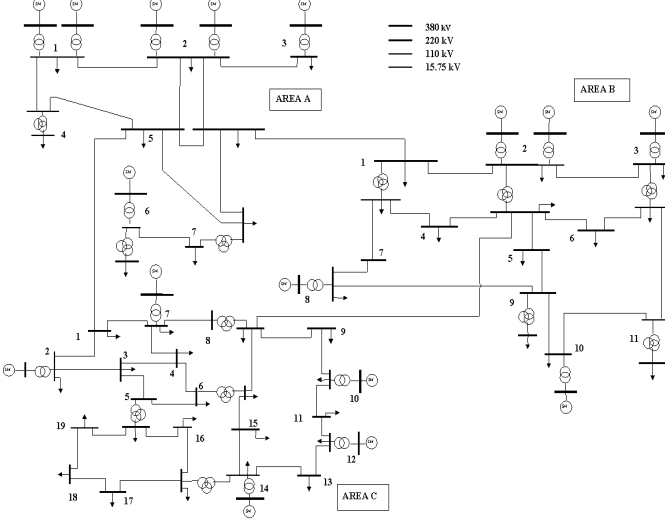The one-line diagram of the PST16 System is shown in Figure 1.

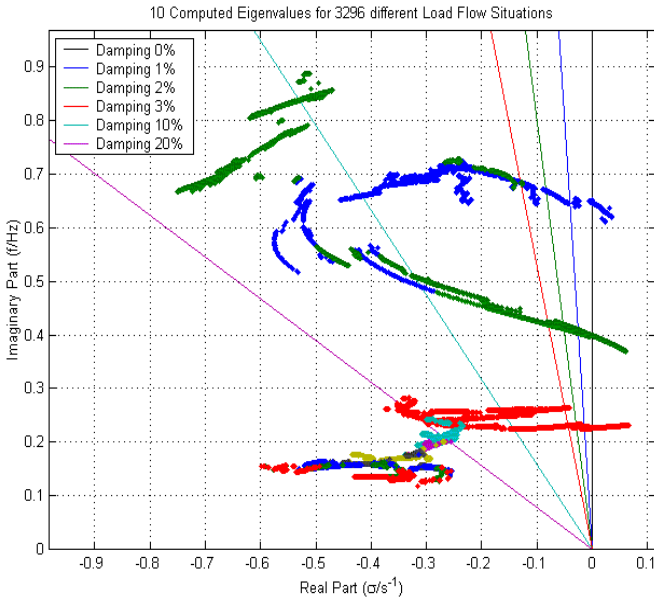Figure 1: One-Line Diagram of the PST 16-Machine Test System

Figure 2: Computed Eigenvalues of the PST 16-Machine Test System for 3,296 different Load Flow Conditions

To generate training data for the NN, different load flow conditions are considered. When the total generation in one area is changed, the change is distributed equally among all the generators in this area. Furthermore, if the generated power of a particular unit falls below a minimum set value, the unit is disconnected. The model also allows additional generating units to be connected when the demand increases.

The different load flow scenarios result in generating 3,296 patterns for NN training. The dominant eigenvalues for all cases are shown in Figure 2. The slant lines in the figure are for constant damping at 0% to 20%. As seen in the figure, most of the cases are for well-damped conditions, but in some cases the eigenvalues shift to the low damping region and can cause system instability.

III. FEATURE SELECTION

Large-scale power systems include many features such as transmission line power flows, generated powers and demands. For a large system, the feature set is too large for any effective NN training [5]. Therefore, feature extraction or selection techniques must be used.

In this study, a new multiple step selection method (MSS) is proposed. This technique is based on the principal component analysis (PCA) and the k-Means cluster algorithm.

The PCA technique is characterized by a high reduction rate and a minimal loss of information. However, the calculation of principal components (PC) requires all original physical features, which is not possible to provide in a large-scale power system. Therefore, for the proposed application only selection techniques are suitable, which are characterized by retaining the most important original feature vectors. However, in this study PCA is utilized for reduction of the dimensionality of feature vectors as shown below.

Let $\boldsymbol{F}$ be a feature matrix of dimension $p \times n$ where n is the number of the original feature vectors and p is the number of patterns. The empirical covariance matrix $\boldsymbol{C}$ of the normalized $\boldsymbol{F}$ is computed by

$$\boldsymbol{C} = \frac{1}{p-1} \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \tag{1}$$

Let $\boldsymbol{T}$ be a $n \times n$ matrix of the eigenvectors of $\boldsymbol{C}$. The diagonal variance matrix $\boldsymbol{\Sigma}^2$ is then given by

$$\boldsymbol{\Sigma}^2 = \boldsymbol{T}^T \cdot \boldsymbol{C} \cdot \boldsymbol{T} \tag{2}$$

$\boldsymbol{\Sigma}^2$ includes the variances σ_x^2. Notice that the eigenvalues λ_k of the covariance matrix $\mathbf{C}$ are equal to the elements of the variance matrix $\boldsymbol{\Sigma}^2$. The standard deviation σ_k is also the singular value of $\mathbf{F}$:

$$\sigma_k^2 = \lambda_k \qquad (k = 1,2,...,n) \tag{3}$$

The n eigenvalues of $\mathbf{C}$ can be determined and sorted in descending order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. While $\boldsymbol{T}$ is an n-dimensional matrix whose columns are the eigenvectors of $\boldsymbol{C}$,

T_q is a $n \times q$ matrix including q eigenvectors of **C** corresponding to the q largest eigenvalues of **C**.

The value of q is determined based on the retained variability of the features, which is the ratio between the first q eigenvalues and the sum of all n eigenvalues [6]. The number q has to be chosen so that υ covers nearly the total signal variability, i.e. υ has a value near to 1.0.

$$\upsilon = \frac{\sum_{i=1}^{q} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{4}$$

From the basic equation of the PC transformation

$$F^{(PC)} = F\,T \tag{5}$$

follows

$$F \approx F_q^{(PC)} T_q^{T} \tag{6}$$

where $F^{(PC)}$ contains the selected PC feature vectors. According to Equation (6) T_q can be interpreted as Loadings Matrix, and $F^{(PC)}$ as the non-normalized Factor Matrix. Now, the idea is to use the columns of T_q^{T} instead of the high dimensional original feature vectors **F** for clustering. In fact, the columns of T_q^{T} are equivalent to those of **F** but with a much lower dimension. Therefore, the used k-Means clustering algorithm [7] shows with T_q^{T} a better performance and accuracy than with the original features directly. This is why both techniques, principal component analysis and clustering, are used in combination. Nevertheless, it is even possible to skip the PCA computation altogether, which means that the original features are clustered directly from the beginning. This approach is preferable for features with smaller pattern number.

The k-means clustering algorithm forms groups with similar characteristics [7]. The criterion for the best grouping is the minimizations of the within cluster sum of distances to the cluster centroids. Features within a cluster show high correlations. On the other hand, independent features are located always in different clusters. The k-means algorithm requires the number of clusters from the beginning. Therefore, to form homogeneous clusters with proper within cluster correlations, some iteration varying cluster numbers are necessary. However, the number of clusters p has to be always greater than the number of PC q to obtain the same variability as the PC.

Because of the similarity between the features within a cluster, one of them can be selected and the rest can be treated as redundant information. The feature in one cluster, which is closest to the centroid of this cluster, will be chosen. Thus, a group of p features will be maintained.
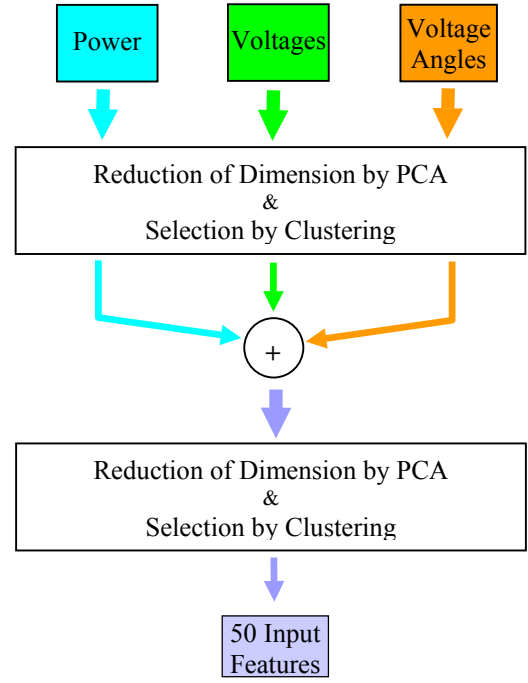

Figure 3: Applied MSS Feature Selection Method

In MSS, the process of selecting features is repeated for extremely inhomogeneous features. In this study, the original feature set was separated into 3 homogeneous subsets including power features (real power, reactive power), voltage and the corresponding voltage angle features. Then, the feature selection method is applied to each subset. The selected features are then combined to form a new group, from which the final features are selected applying the same method again. Finally, 50 of the original features remain as NN input variables. The MSS feature selection process is shown in Figure 3.

Although the feature clustering has been carried out based on the transformed features reduced dimensionality, for NN training as described in the next chapter always original features were used.

IV. OBSERVATION AREA

The proposed stability assessment method requires that the observation area in the complex eigenvalue space is defined first. The observation area is located at the region of insufficient damping, where typical inter-area eigenvalues can be found. In this study, the damping range is chosen between 4% and −1.5%. Then samples are generated along the real axis (σ) for damping coefficients between 4% and 1.5%. This process is repeated for 4 different frequencies f.

The width of σ and f sampling steps are $\Delta\sigma$ and Δf, respectively. These step widths are constant for the entire sampling process at:

$$\Delta\sigma = 0.015\ s^{-1} \qquad \Delta f = 0.14\ \text{Hz}$$

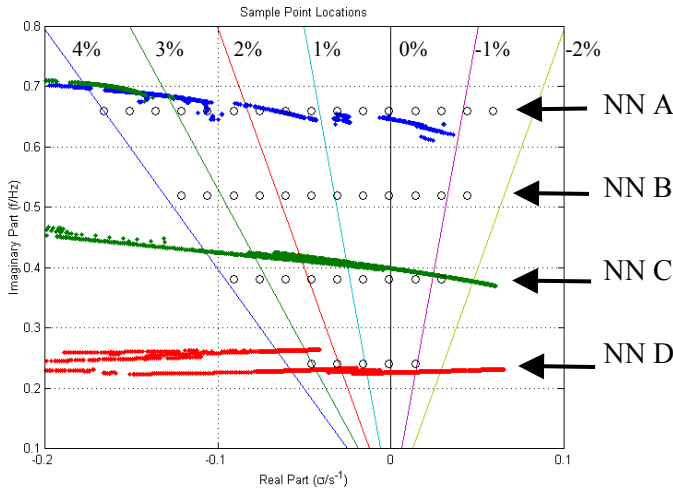The observation area of insufficient damping is shown in Figure 4. The sample points are marked by circles.

Figure 4: Computed Eigenvalues of the PST 16-Machine Test System for 3,296 different Load Flow Conditions. The Figure shows Sample Points in the Observation Area

The sampling method described in Figure 4 produced a set of 42 sample points. To increase the accuracy of the NN, the observation area is divided into smaller observation sections (rows of sample points) whereby each row of sample points uses its own NN. Hence, 4 independent NN are used. The 4 observed sections covered by the 4 NN are defined in Table I.

TABLE I
SAMPLED OBSERVATION AREA DIVIDED INTO 4 SECTIONS FOR NN TRAINING

NN	Outputs	σ_{min} / s^{-1}	σ_{max} / s^{-1}	f / Hz
A	16	-0.166	0.059	0.660
B	12	-0.121	0.044	0.520
C	9	-0.091	0.029	0.380
D	5	-0.046	0.014	0.240

After the observation area is sampled, the distances between the sample points and the eigenvalues are computed. Then, the distances are used to compute activation values of the sample points. The complex plane of the eigenvalues is defined by the real component σ_{ev} and the frequency f_{ev}. The sample points are defined by their location in this complex plane (σ_s, f_s). The Euclidean distance between a given eigenvalue and a given sample point is computed as follows:

$$d = \sqrt{\left(\frac{\sigma_s - \sigma_{ev}}{k_\sigma}\right)^2 + \left(\frac{f_s - f_{ev}}{k_f}\right)^2} \qquad (7)$$

Because σ and f use different units and cannot be compared directly, both are scaled by k_σ and k_f. These constants are chosen in relation to the step size in σ and f directions. Thus,

$$k_\sigma = \Delta\sigma \qquad (8)$$

$$k_f = \frac{\Delta f}{\sqrt{3}} \qquad (9)$$

Notice that this choice leads to a distance of 1 respective $\sqrt{3}$ between neighboring sample points in σ and f directions, respectively. The choice of these parameters is based on experience with NN training. This scaling puts emphasis on eigenvalue movement along sample point rows because the NNs are trained by sample point rows.

Moreover, the maximum distance possible between an eigenvalue and the closest sample point occurs when the eigenvalue is located exactly in the geometrical center of 4 adjacent sample points. According to Equations (7) to (9), the maximum distance can be computed as

$$d_{max} = \sqrt{\left(\frac{\frac{\Delta\sigma}{2}}{\Delta\sigma}\right)^2 + \left(\frac{\frac{\Delta f}{2}}{\frac{\Delta f}{\sqrt{3}}}\right)^2} = 1 \qquad (10)$$

Based on this maximum distance, the activation value a for a sample is defined as a linear relationship depending on the distance by Equation (7)

$$a = \begin{cases} 1 - 0.5 \cdot \dfrac{d}{d_{max}} & 0 \le d \le 2d_{max} \\ 0 & d > 2d_{max} \end{cases} \qquad (11)$$

The activation a_i is computed for a given set of sample points with respect to all eigenvalues of one pattern. The final activation value act for the given sample point is the summation of all activations a_i

$$act = \sum_{i=1}^{n} a_i \qquad (12)$$

where n is the number of considered eigenvalues. The maximum distance in Equation (8) and the activation function in Equation (9) lead to the minimum activation for a sample point located near an eigenvalue

$$act \ge 0.5 \qquad (13)$$

V. EIGENVALUE PREDICTION BY NEURAL NETWORKS

The activations of the sample points are computed for all patterns. The data are then normalized and shuffled randomly before the 4 NN are trained independently. In this study, the number of training patterns is 2,966 and the number of testing patterns is 330. Once the NNs are trained, the NN output (activation value) is used to identify the locations of the eigenvalues. For all patterns, the activation values given from the sample points are used to set up an activation diagram, which is used to construct a range for the eigenvalue position as shown in Figure 5 for one pattern.

According to Table I, the observed section of NN (A) is within $\sigma_{min} = -0.166 \ s^{-1}$ and $\sigma_{max} = 0.059 \ s^{-1}$. The activation diagram is extended outside of these borders by step $\Delta\sigma$, where the activation values of the last points are kept constant.

The frequency of the imaginary part is $f = 0.66\,\text{Hz}$ and is bounded by $f \pm \dfrac{\Delta f}{2}$.

The activations of the sample points in Figure 5(a) are plotted in bold face and connected by lines. The maximum activation limit (margin) is set at 0.5, which is the minimum activation value according to Equation (11). The intersection of the activation border with the activation level curve determines the range r of the real part. The range r and the limits of f form a rectangle representing the prediction region. The function of the NN is to predict the eigenvalues of this region. In the example given in Figure 5(b), the position of the eigenvalue is $\sigma = -0.0068\,s^{-1}$ and $f = 0.6602\,\text{Hz}$.
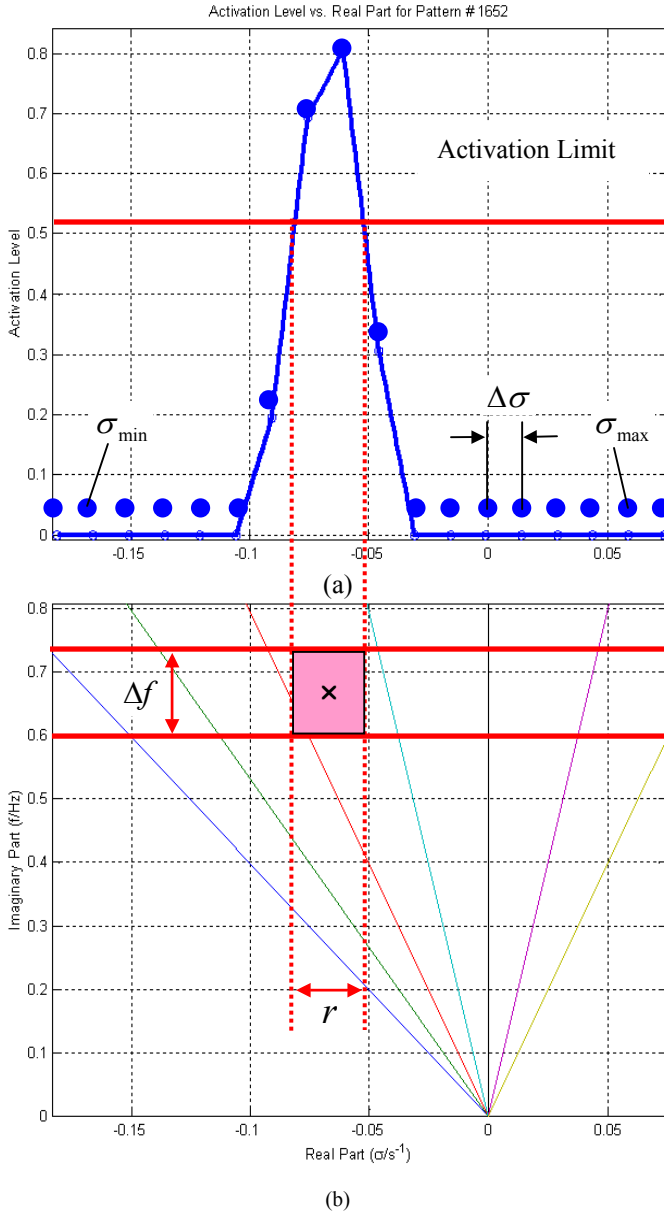

Figure 5: (a) Activation Diagram for one Pattern, (b) Complex Space with Corresponding Eigenvalue Identified

The identified eigenvalue is marked with a cross and matches the predicted region exactly.

If an eigenvalue is located inside a region, the prediction is considered correct. If an eigenvalue is located outside the region the error is called false dismissal. If a range is identified, but no eigenvalue is located inside the observation area, the error is called false alarm. For each of the 4 NN, the false dismissal and the false alarm error is computed as given in Equation (14)

$$E[\%] = 100 \cdot \frac{\text{number of false dismissals or false alarms}}{\text{number of pattern}} \quad (14)$$

The results are given in Table II.

TABLE II
FALSE DISMISSAL AND FALSE ALARM ERRORS OF NN

NN	False Dismissal		False Alarm	
	Training	Testing	Training	Testing
A	0.00 %	0.00 %	0.00 %	0.00 %
B	0.00 %	0.00 %	0.00 %	0.00 %
C	0.03 %	0.00 %	0.00 %	0.00 %
D	0.00 %	0.00 %	0.00 %	0.00 %

The errors shown in Table II are almost always zero. False alarm did not occur and false dismissal occurs only once for one NN. In other words, the actual eigenvalue position is always within the predicted region.

However, it was observed that when the range is small, the accuracy of the NN is reduced. To address this issue, a relative range c is defined as the ratio of the absolute range r to the step width $\Delta \sigma$

$$c = \frac{r}{\Delta \sigma} \quad (15)$$

Table III shows the mean and the standard deviation of c for all constructed ranges r.

TABLE III
MEAN AND STANDARD DEVIATION OF NN ERRORS WHEN THE RELATIVE RANGE ACCORDING EQ. (15) IS USED

NN	Training		Testing	
	Mean	STD	Mean	STD
A	1.893	0.209	1.892	0.208
B	/	/	/	/
C	1.909	0.292	1.840	0.332
D	2.046	0.346	1.999	0.328

Another way for evaluating the accuracy of the NN is to check the range with respect to the damping coefficient. For this purpose, the real part of r is transformed into a damping range $\Delta \xi$ as shown in Figure 6. The mean and standard deviations of the damping range are shown in Table IV.
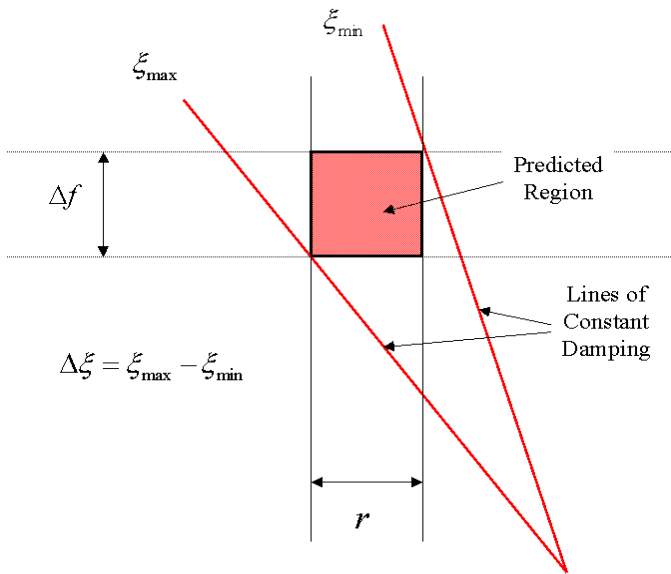
Figure 6: Construction of the Damping Range for the Predicted Eigenvalues

TABLE IV

NN ERRORS USING THE DAMPING RANGE $\Delta\xi$

NN	Training		Testing	
	Mean	STD	Mean	STD
A	1.3 %	0.3 %	1.3 %	0.3 %
B	/	/	/	/
C	1.9 %	0.4 %	1.9 %	0.4 %
D	3.5 %	0.7 %	3.4 %	0.7 %

VI. CONCLUSION

This paper introduces a new method of eigenvalue prediction for small-signal stability assessment. The method is flexible in terms of network conditions and variable number of eigenvalues. The errors, listed in Table II, are almost always zero. In other words, by the proposed method, the eigenvalues are located within the predicted region. By decreasing the sample step width, the accuracy can be improved even more, but the time for NN training increases.

For high and medium frequencies (NN A to NN C), the damping of eigenvalues can be predicted in a range, which is smaller than 1.9%. Only for low frequencies, the average width of the damping range increases to 3.5%. The reason is the nonlinear relationship that exists between the real part and the damping coefficient. The lower the frequency, the larger the damping difference for a constant interval.

The proposed method allows the fast assessment of small-signal stability for large power systems based on a restricted number of process variables. It is so fast that it can be implemented as an on-line tool. A further advantage of this method is that only the selected variables must be exchanged and transmitted between the different parts and owners of the system.

VII. REFERENCES

[1] U. Bachmann, I. Erlich and E. Grebe, "Analysis of interarea oscillations in the European electric power system in synchronous parallel operation with the Central-European networks", IEEE PowerTech, Budapest 1999

[2] H. Breulmann, E. Grebe, M. Lösing, W. Winter, R. Witzmann, P. Dupuis, M.P. Houry, T. Margotin, J. Zerényi, J. Dudzik, J. Machowski, L. Martín, J.M. Rodríguez, E. Urretavizcaya, "Analysis and Damping of Inter-Area Oscillations in the UCTE/CENTREL Power System", CIGRE 38-113, Session 2000

[3] P. Kundur, Power System Stability and Control, McGraw-Hill, New York, 1994

[4] S.P. Teeuwsen, A. Fischer, I. Erlich, M.A. El-Sharkawi, "Assessment of the Small Signal Stability of the European Interconnected Electric Power System Using Neural Networks", LESCOPE 2001, Halifax, Canada, June 2001

[5] S.P. Teeuwsen, I. Erlich, M.A. El-Sharkawi, "Feature Reduction for Neural Network based Small-Signal Stability Assessment", PSCC 2002, Sevilla, Spain, June 2002

[6] I.T. Jolliffe, Principal Component Analysis, Springer-Verlag, New York, 1986

[7] Visual Numerics: IMSL Fortran 90 MP Library Help, Stat/Library, Volume 2, Chapter 11: Cluster Analysis

VIII. BIOGRAPHIES

Simon P. Teeuwsen (1976) is presently PhD student in the Department of Electrical Power Systems at the University of Duisburg/Germany. He started his studies at the University of Duisburg in 1995. In 2000, he went as exchange student to the Unversity of Washington, Seattle, where he performed his Diploma Thesis. After his return to Germany in 2001, he received his Dipl.-Ing. degree at the University of Duisburg. He is a member of VDE and VDI.

Istvan Erlich (1953) received his Dipl.-Ing. degree in electrical engineering from the University of Dresden/Germany in 1976. After his studies, he worked in Hungary in the field of electrical distribution networks. From 1979 to 1991, he joined the Department of Electrical Power Systems of the University of Dresden again, where he received his PhD degree in 1983. In the period of 1991 to 1998, he worked with the consulting company EAB in Berlin and the Fraunhofer Institute IITB Dresden respectively. During this time, he also had a teaching assignment at the University of Dresden. Since 1998, he is Professor and head of the Institute of Electrical Power Systems at the University of Duisburg/Germany. His major scientific interest is focused on power system stability and control, modelling and simulation of power system dynamics including intelligent system applications. He is a member of VDE and IEEE.

Mohammed A. El-Sharkawi received the B.Sc. degree in electrical engineering in 1971 from Cairo High Institute of Technology, Egypt, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of British Columbia, Vancouver, B.C., Canada, in 1977 and 1980, respectively. In 1980, he joined the University of Washington, Seattle, as a Faculty Member. He served as the Chairman of Graduate Studies and Research and is presently a Professor of Electrical Engineering. He is the Vice President for Technical Activities of the Neural Networks Society. He co-edited an IEEE tutorial book on the applications of neural network to power systems. He organized and taught several international tutorials on intelligent systems applications, power quality and power systems, and he organized and chaired numerous panel and special sessions in IEEE and other international conferences. He published over 120 papers and book chapters in these areas and holds seven licensed patents. He is a member of the editorial board and Associate Editor of several journals, including the IEEE TRANSACTIONS ON NEURAL NETWORKS and *Engineering Intelligent Systems*.