

Neural Network based Multi-Dimensional Feature Forecasting for Bad Data Detection and Feature Restoration in Power Systems

S. P. Teeuwsen, *Student Member, IEEE*, I. Erlich, *Member, IEEE*,

Abstract--This paper deals with new methods for bad data detection and restoration, which can help to improve modern state estimation. Special interest is focused on the detection of bad data and outliers in the measurements from the power system. The new artificial neural network based method involves both the time series of the measured data and their physical correlation. The Neural Network (NN) inputs are power system measurements from the history, e.g. 5 – 30 minutes back. Then, the NN is utilized to predict the expected on-line measurements. When the measured data show significant differences to the NN prediction, the measurements are suspicious to contain bad data. Finally, identified bad data inputs will be restored by NN based optimization to their correct value. This paper shows that NNs implemented for bad data detection and restoration are a valuable addition for future state estimation.

Index Terms--Interconnected Power Systems, State Estimation, Measurements, Bad Data Detection and Restoration, Optimization, Neural Networks

I. INTRODUCTION

State estimation has been introduced and implemented in power systems already in the 60s as a single frequency, balanced and symmetric power system model under steady state conditions. The basic idea of state estimation is the estimation of information from the power system, which is not directly measured. Hereby, the voltages and the voltage angles are estimated based on the measured data. Then, voltage dependent features like transmitted real and reactive power are computed based on the estimated voltages and voltage angles. In order to obtain accurate estimates, the measured input data must contain redundancy. This redundancy leads to an over-determined system of equations, whose optimal solution is given by minimization of the estimation errors. The estimation errors are summed squared errors resulting from the measurement errors weighted by their variances [1]. In the case of non-linear relationships between the voltages and other electrical quantities in the power system, the solution must be obtained iteratively starting from an initial starting

condition [2]. However, presently implemented state estimation methods require a lot of redundancy and knowledge about fault tolerances. Moreover, they compute their estimates based on a set of instantaneous measurements without considering any trends over time. Also physical relationships between the variables for the given system cannot be considered a priori for bad data detection and error estimation. The methods for bad data detection and restoration proposed in this paper were originally developed for implementation in the field of Dynamic Stability Assessment (DSA) and in particular for Oscillatory Stability Assessment (OSA) of large interconnected power systems [3] – [5]. Such assessment tools require bad data detection independent of the state estimation, since some of the used system measurements are not involved in the state estimation process and therefore these data are not verified. For this reason, the authors developed new methods for data verification and restoration of bad data. These methods, introduced in [5] and [6], can also be implemented as additional tool for classical state estimation. Therefore, this paper suggests expanding the existing state estimation methods by a newly developed feature prediction tool.

A common method to verify measurements is the comparison to known limits of the measurements. If the measured value is between the limits, the value might be a correct measurement. If the value exceeds one of the limits, the value is a suspicious outlier. However, this way of verification does not consider that the measurement has a typical time curve over the day since the operating conditions change continuously. If the range of the limits is large, there is a high probability that the measured value is between the limits but though an outlier.

Therefore, this paper introduces a new approach to predict an expected value to be measured for the complete set of measurements. The prediction includes the time behavior of the measurements and therefore the changing operating conditions. The limit to recognize outliers is not the maximum and minimum values for the complete operation range, but a narrow band around the predicted measurement. The bandwidth is constant, but the absolute minimum and maximum values of this band are time dependent and change with the operating point. Figure 1 shows the time curve of one measurement and both methods of outlier detection. The common comparison to fixed feature limits is shown and the

S. P. Teeuwsen, University of Duisburg-Essen, Germany, (teeuwsen@uni-duisburg.de), since October 2005 with Siemens PTDErlangen (simonp.teeuwsen@siemens.com)
 I. Erlich, University of Duisburg-Essen, Germany, (erlich@uni-duisburg.de)
 M. A. El-Sharkawi, University of Washington, Seattle, USA, (elsharkawi@ee.washington.edu)

new approach by a variable tolerance band, which allows a much higher accuracy.

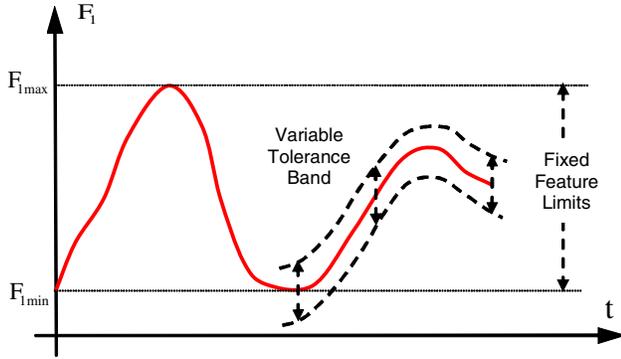


Fig. 1 Traditional Way of Outlier Detection (Fixed Feature Limits) and New Approach (Variable Tolerance Band)

As can be seen in Figure 1 the observed measurement changes continuously over time. The range between the maximum value and the minimum value is large and therefore, an outlier can only be detected, if the limits are exceeded. However, small outliers cannot be detected. The figure also shows the variable tolerance band. This band results from a predicted value of the measurement. If the measured value is near to the predicted, but inside the tolerance band, it is correct. If the measurement exceeds the limits of the tolerance band, it is an outlier. The tolerance band has any time the same band width, but is varies with the predicted value and therefore allows much higher detection accuracy.

The new method for feature forecast is based on a Neural Network (NN) and involves both the time dependency of the measurements and the physical correlation between the measurements. The short time history of measured data is used to predict or estimate the values for present measurements. If one measurement shows a significant difference to the estimate for this measurement, it is suspect to contain bad data. Once the measurement is identified, it can be restored to the original value by an optimization method, based on a NN, too. One advantage is the expansion of the existing state estimation methods. Another advantage is the fact that NNs can be trained for several different operating conditions, since the classical state estimator depends on one given operating condition.

II. 16-MACHINE DYNAMIC TEST SYSTEM

The PST 16-Machine Test System used in this study is a 400/230 kV dynamic test system shown in Figure 2. The main focus of this system is on the power system dynamics and inter-area oscillations in the time range of a few seconds to minutes [5] and [7].

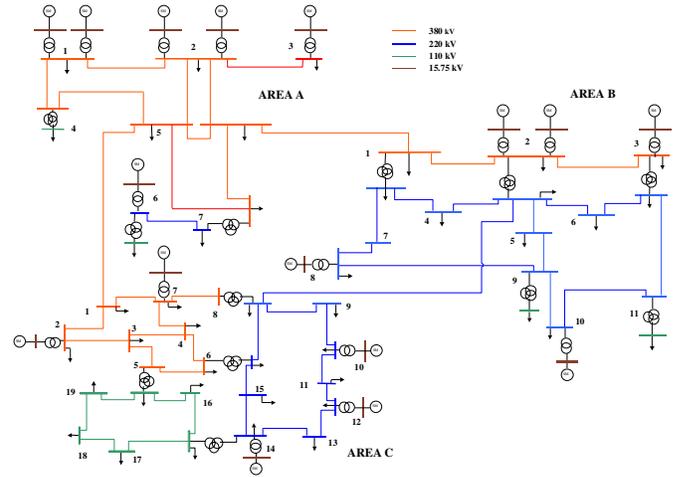


Fig. 2 One-Line Diagram of the PST 16-Machine Test System

The PST 16-Machine Test System is modeled by using real data from new generators and controllers for different units. The power system operates with 3 different types of generators, such as hydro power generators, thermal driven generators, and nuclear driven generators. For hydro generators, the rated generator power is 220 MW per block. For thermal driven generators, the block size is 247 MW, and for nuclear driven generators about 259 MW. However, a power plant in the network is modeled by connecting a particular number of generator units to the electric transmission system. The generator models are 5th order and the exciter models are obtained from IEEE standard types [8]. Hereby, the hydro generators are controlled by a slightly modified version of the ST1A excitation system. The thermal driven generators use the DC1A exciter model, and the nuclear driven generators use the AC1A exciter model. The system consists of 3 strongly meshed areas. Each has 5 – 6 generators.

III. INPUTS FOR MODERN STATE ESTIMATION

Typical state estimator inputs are voltages, real and reactive powers, and the network topology. However, in modern state estimation, it might be of interest to use as many data from the power system as possible. The more inputs, the more detailed is the information about the current state of the power system. Considering the terminology of Computational Intelligence, the variables treated in the state estimation can be defined as Features. Typical power system features such as the real and reactive power, voltages, and voltage angles are listed in Table I.

TABLE I
TYPICAL FEATURES IN LARGE INTERCONNECTED POWER SYSTEMS

| # | Feature Description |
|---|---|
| 1 | Generated Real and Reactive Power |
| 2 | Real and Reactive Power Demand |
| 3 | Real and Reactive Power on Transmission Lines |
| 4 | Power exchanged between interconnected Areas |
| 5 | Bus Voltages |
| 6 | Voltage Angles (requires phasor measurements) |

To implement methods for bad data detection, daytime data are generated first. Hereby, a Genetic Algorithm (GA) is used for generator unit commitment. The applied load curves are adapted from real load curves [9] to the PST 16-Machine Test System. The curves are applied once for an operating condition in winter and once for an operating condition in summer. About half of the generators in the power system are scheduled to generate constant power output at each time. The remaining generators are allowed to adapt the generation output to fit the current demand. In the following, the GA is used to schedule the generator commitment for 3 winter days and 3 summer days under the given constraints. Afterwards, the generation and load curves are sampled in time steps of 5 minutes in order to simulate data measurements in the power system at 5-minute intervals. The summed generation for each area in the power system for both real and reactive power is shown in Figure 3.

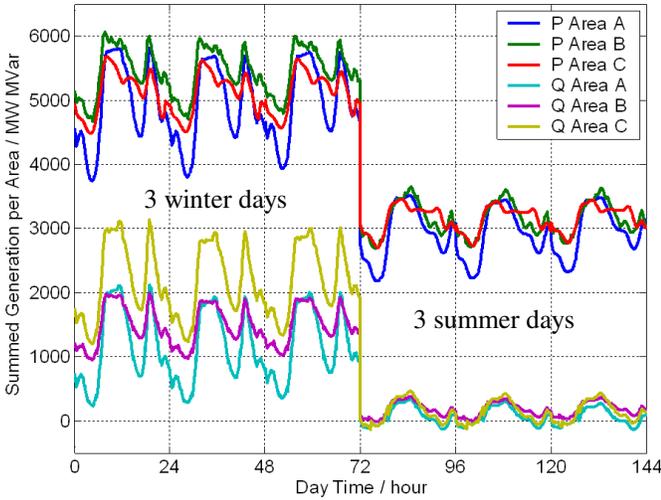


Fig. 3. Sum of Generated Real and Reactive Power per Area in Winter and Summer Operating Conditions over 6 Days

IV. MULTIDIMENSIONAL FEATURE FORECASTING

The method for non-linear, multidimensional feature forecasting is based on a NN and was introduced first in [6]. Hereby, a multilayer feed-forward NN is trained with the data history of selected input features. To keep the size of the network in a dimension where it can be trained easily, a small set of selected features (about 50 features) is used as NN input. Moreover, only the last 4 – 6 history data are used. When a time interval of 5 minutes for measurements is assumed, the information reaches about 20 – 30 minutes back into the past. Figure 4 shows the basic idea for n input features and k history data. Each of the history data inputs is averaged over N time steps to include more history information than of only the k presented inputs per feature. Representatively for all input features, the process of averaging of the time series is shown for feature 1 in Figure 5.

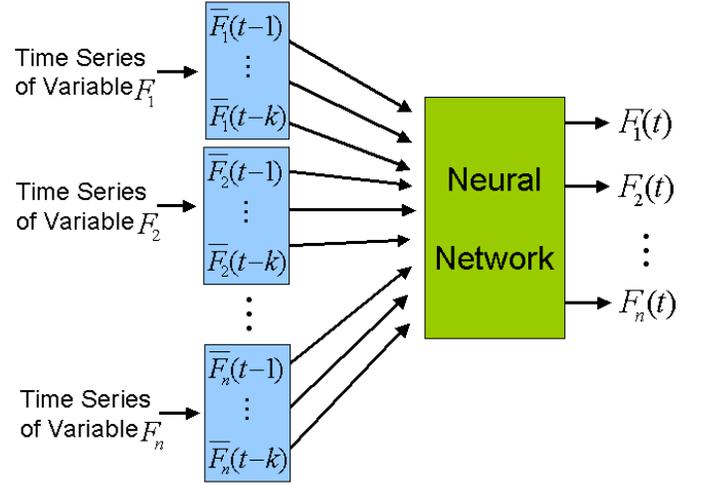


Fig. 4 Basic Scheme for non-linear, multidimensional Feature Forecasting using NN with n Input Features and k History Values

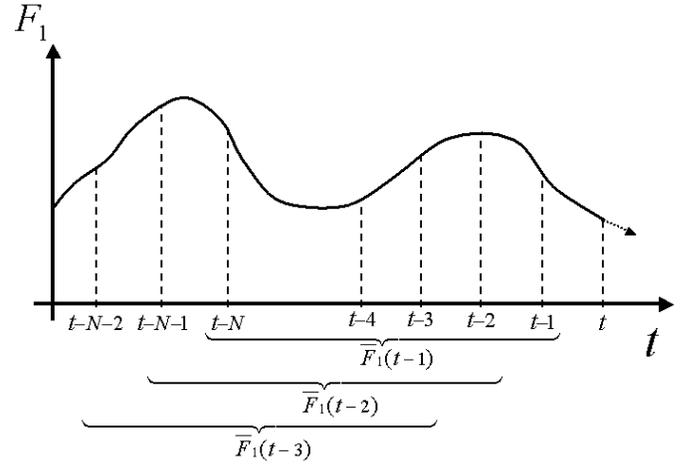


Fig. 5. Computation of Average Values at the Example of Feature 1

Average values are used as NN inputs. When t is the time in a discrete time curve, the weighted average value at time t is given by

$$\bar{x}(t) = \sum_{j=1}^N \frac{1}{2^j} \cdot x(t-j) + \frac{1}{2^N} \cdot x(t-j) \quad (1)$$

Hereby, N determines the number of time steps used from the past. The benefit of the weighted average is that much more history data influence the forecasting.

The NN is trained with time series data from the database. When the NN is used on-line, the NN outputs are compared to the measured features and those showing differences beyond a certain limit are flagged as outliers. The advantage of this method results from its multidimensional characteristic. Both the history of all features (time series) and the load flow conditions are considered and thus the NN is able to predict the correct reference value to check the measurements.

In the following, two examples are given. Example 1 in Figure 6 shows the NN results of feature forecasting for the real power on the transmission line between Area A and Area B.

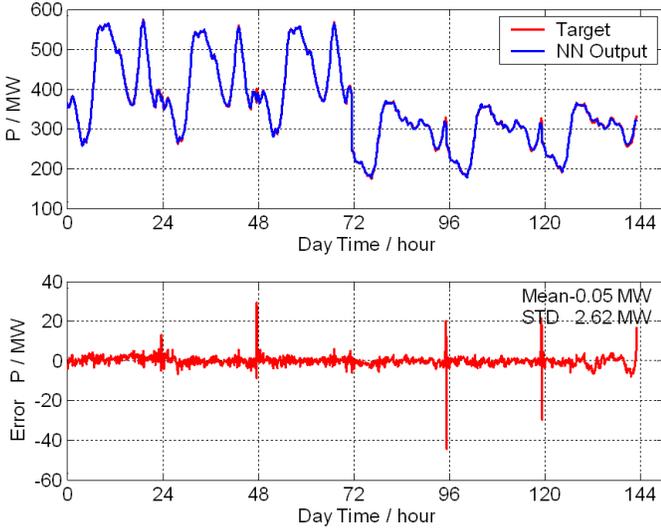


Fig. 6. Multidimensional Feature Forecasting by NN over 6 Days for the Real Power on a Transmission Line

The figure shows the comparison between the NN target and output (top figure) and the error as difference between the NN target and output (bottom figure). The NN training is carried out with a time series over 6 days. The NN inputs consist of $k = 4$ history data (each as a weighted average over $N = 50$ time steps back) for $n = 50$ measurements from the power system. Thus, the NN has a total number of 200 inputs.

Example 2 is given in Figure 7 showing the NN results for the voltage at a 400 kV bus in Area C. The NN consist of 4 history data (each as a weighted average over 50 time steps back) for each measurement from the power system. In this example, there are only $n = 10$ measurements selected from the power system, which results in a total NN input number of 40.

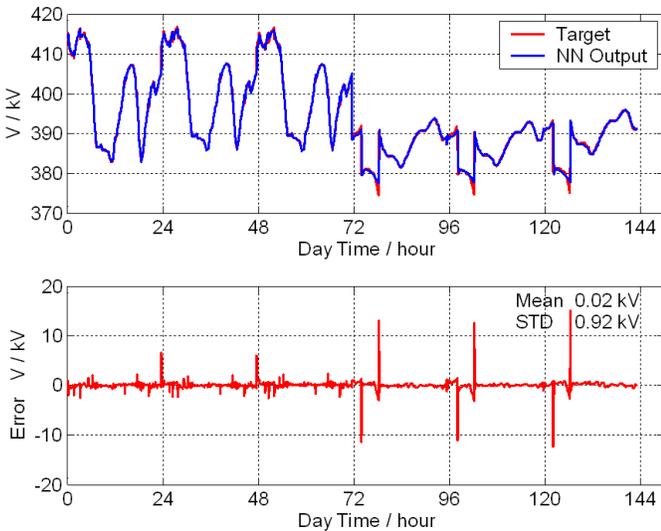


Fig. 7. Multidimensional Feature Forecasting by NN over 6 Days for a Bus Voltage

Both figures show highly accurate prediction results. Therefore, the NN outputs can be compared to real-time measurements in order to detect bad data.

V. OUTLIER RESTORATION

The outlier restoration method is implemented as non-linear sequential quadratic optimization algorithm using a cost function based on an NN auto encoder. This special type of NN is trained to recognize the load flow scenarios included in the database. When used for outlier restoration, the optimization algorithm minimizes the NN auto encoder error for a given load flow scenario with bad or missing data. The auto encoder network is a well-known type of NN and discussed in detail in [10]. This NN, shown in Figure 8, replicates the input vector at the output. In other words, the NN is trained to obtain identical input and output.

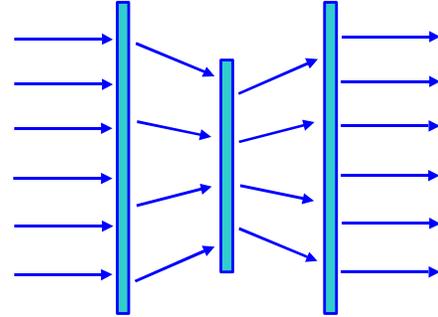


Fig. 8. NN Auto Encoder

The error between input and output vector is low for any trained load flow condition. When a new pattern contains bad data, the features of this pattern do not belong to a feasible load flow condition. Hence, the error is high when presented to the NN auto encoder. Therefore, abnormal differences between inputs and outputs of the auto encoder indicate patterns that are greatly suspicious to containing bad data. The auto encoder is highly sensitive to outlier patterns. However, it is not possible to determine the feature from the new pattern, which contains the bad data. Since all features inside the NN are cross-connected, a high auto encoder error of a given feature does not necessarily mean bad data. Therefore, the NN auto encoder can only be used to detect outlier patterns, but not outlier features.

Once a new input pattern is recognized as outlier pattern and the particular bad features are identified, the bad data have to be restored. This is possible using the NN auto encoder network. Hereby, the NN auto encoder is involved into an optimization loop where it describes the power-flow-conform relationship between the features. A sequential quadratic optimization, which is the most favorable algorithm for continuous problems, is utilized to adjust the elements of the feature vector. The algorithm restores the bad data elements by minimizing the error of the NN auto encoder.

Usually it can be assumed that most of the original measurements are correct and thus these variables can be kept constant. Hence, only some suspicious inputs need to be changed by the optimization. The restoration of bad features succeeds when the difference between the input and output vectors of the auto encoder is small. However, the predicted

feature value from the NN provides also information about the optimization constraints. In this case, the initial starting condition for optimization is not zero, but the predicted feature value and the maximum and minimum constraints are given by a small tolerance around the predicted value. The optimization diagram is shown in Figure 9.

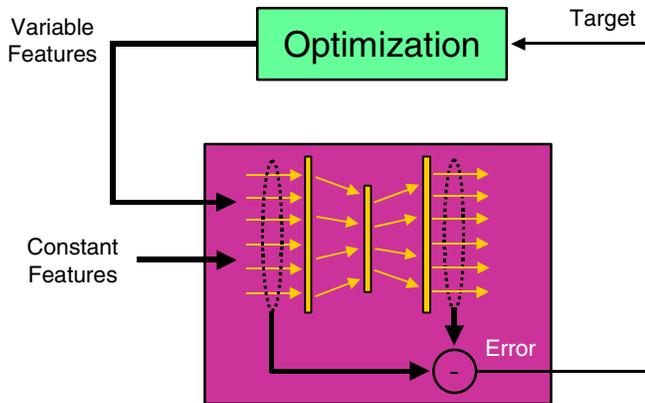


Fig. 9. Non-linear sequential quadratic Optimization Procedure for Feature Restoration using NN Auto Encoder

The two examples given above are continued in this subsection. The optimization algorithm is applied to restore features to their original values. In both examples, the optimization adjusts the feature determined as missing. The remaining inputs are kept constant to determine the operating point to be restored. Starting from zero as the initial value, the optimization restores the original value within a few seconds and with high precision. Table II shows the comparison of the original feature values and the restoration results.

TABLE II
RESTORATION RESULTS FOR TWO GIVEN EXAMPLES

| Example | Inputs | Original | Restored |
|---------|--------|-----------|-----------|
| 1 | 50 | 947.86 MW | 952.56 MW |
| 2 | 10 | 414.75 kV | 414.88 kV |

To test the ability of the optimization algorithm to restore more than one missing input, 8 voltage features from a set of 50 measurements are determined as “missing”. Figure 10 shows the progress of restoration of selected voltages as a function of the number of function calls during the optimization process.

The total number of iterations is 33, which is much smaller than the total number of calls since the optimization calls the cost function several times during each iteration step. The comparison between original and restored feature values is listed in Table III. As can be seen, the restoration is very accurate.

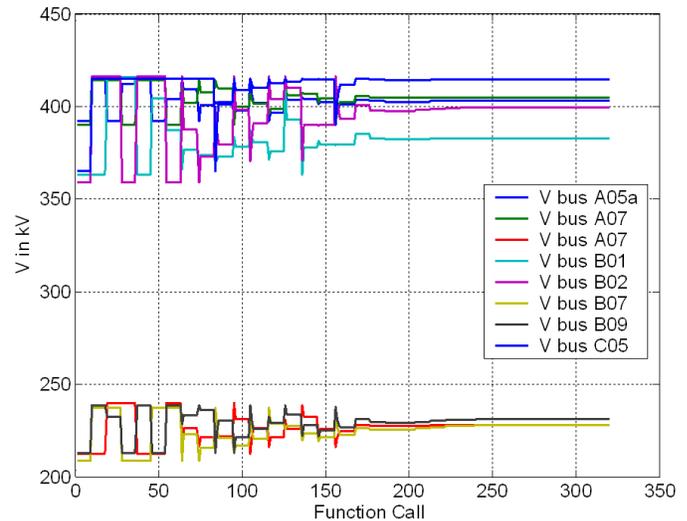


Fig. 10. Progress of Feature Restoration for 8 Voltage Features

TABLE III
COMPARISON OF 8 ORIGINAL AND RESTORED VOLTAGE FEATURES IN KV

| # | Original | Restored | Difference |
|---|----------|----------|------------|
| 1 | 402.96 | 402.88 | 0.02 % |
| 2 | 404.88 | 404.79 | -0.02 % |
| 3 | 227.92 | 227.95 | 0.01 % |
| 4 | 383.04 | 382.65 | -0.10 % |
| 5 | 399.74 | 399.33 | -0.10 % |
| 6 | 228.92 | 228.02 | -0.39 % |
| 7 | 232.01 | 231.05 | -0.42 % |
| 8 | 414.97 | 414.52 | -0.11 % |

The examples shown above are representative of optimization based restoration. Because the NN auto encoder is trained with various load flow scenarios, it always finds a solution that meets a feasible load flow condition, when the fixed variables are enough to represent this load flow situation. However, it requires that the input feature vector contains redundancy not only for a few variables, but for the whole vector. This issue must be considered in the feature selection process. When a small and large set of selected features both show low errors in OSA, the large set is preferred to be chosen since it contains definitely more redundancy than the small one. For a few missing features, the algorithm always finds the correct solution. In the investigated PST 16-Machine Test System, even for 50% missing data of 50 inputs, the algorithm is able to find a solution, which is close to the original values.

Instead of the outlier restoration by optimization, such as shown in this section, it is also possible to use the feature value predicted by the NN for outlier detection. The predicted feature value is already a good approximation for the real feature value. Therefore, the measured value (outlier) is replaced by the predicted value (approximation), which is very close to the real feature value.

VI. CONCLUSION

The traditional state estimation based on the LS method depends on one given operating condition and therefore, it is not flexible in terms of operating conditions and data history. It does not take any advantage of the time series of measurements, which is already given in the database. Moreover, the physical relationship between the variables is not regarded in classical state estimation. When a NN is implemented for multidimensional feature forecasting, any measurement can be verified based on the prediction and bad data are identified easily. The new method can be used for any type of measurement such as voltage angles, which are not included in classical state estimation procedures.

The NNs in this study introduced for feature forecasting and restoration are highly accurate in their results. Both, the new feature forecasting method and the recently proposed methods for Computational Intelligence based Oscillatory Stability Assessment [5] are excellent additional tools for Transmission System Operators. They can be implemented easily in existing state estimators and provide the system operator with valuable extra information.

VII. REFERENCES

- [1] G.W. Stagg, et. al., "Techniques for the real-time monitoring of power system operation", IEEE Trans. Power Appar. Syst. 89, 1970
- [2] F.C. Schweppe, J. Wildes, "Power system static state estimation, Part I: Exact model", IEEE Tans. Power Appar. Syst. 89, 1970
- [3] S.P. Teeuwsen, I. Erlich, M.A. El-Sharkawi, "Fast Eigenvalue Assessment for Large Interconnected Powers Systems", IEEE PES General Meeting, San Francisco, USA, June 12-16, 2005
- [4] S.P. Teeuwsen, I. Erlich, M.A. El-Sharkawi, "Small-Signal Stability Assessment for large Power Systems using Computational Intelligence", IEEE PES General Meeting, San Francisco, USA, June 12-16, 2005
- [5] S.P. Teeuwsen, Oscillatory Stability Assessment of Power Systems using Computational Intelligence, Aachen, Shaker, 2005
- [6] S.P. Teeuwsen, I. Erlich, M.A. El-Sharkawi, "Robust Oscillatory Stability Assessment for Large Interconnected Power Systems", IEEE PES General Meeting, Denver, USA, June 6-10, 2004
- [7] S.P. Teeuwsen, I. Erlich, M.A. El-Sharkawi, "Neural Network based Classification Method for Small-Signal Stability Assessment", IEEE PowerTech, Bologna, Italy, June, 2003
- [8] P. Kundur, Power System Stability and Control, McGraw-Hill, New York, 1994
- [9] RTE France, Transmission System Operator, Load Curves available at <http://www.rte-france.com/jsp/an/courbes/courbes.jsp>
- [10] R.D. Reed and R.J. Marks, II, Neural Smithing, MIT Press, Cambridge, Mass., 1999

VIII. BIOGRAPHIES



Simon P. Teeuwsen (1976) is presently research assistant at the Department of Electrical Power Systems at the University of Duisburg-Essen/Germany. He started his studies at the University of Duisburg in 1995. In 2000, he went for one year as exchange student to the University of Washington, Seattle, where he performed his Diploma Thesis. He received the Dipl.-Ing. and the Dr.-Ing. degrees in electrical power engineering from the University of Duisburg-Essen in 2001 and 2005, respectively. In June 2005, he received the Basil Papadias Award on the IEEE PowerTech Conference in St. Petersburg, Russia, for the best student paper. He is a member of VDE, VDI, and IEEE.



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-Essen/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.