Proceedings of the Korean Nuclear Society Spring Meeting
Gyeongju, Korea, May 2003

State Analysis of BOP Using Statistical And Heuristic Methods

Gyunyoung Heo and Soon Heung Chang
Korea Advanced Institute of Science and Technology
Department of Nuclear and Quantum Engineering
373-1 Kuseong-dong, Yuseong-gu, Daejeon 305-701, Republic of Korea

Abstract

Under the deregulation environment, the performance enhancement of BOP in nuclear power plants is being highlighted. To analyze performance level of BOP, we use the performance test procedures provided from an authorized institution such as ASME. However, through plant investigation, it was proved that the requirements of the performance test procedures about the reliability and quantity of sensors was difficult to be satisfied. As a solution of this, state analysis method that are the expanded concept of signal validation, was proposed on the basis of the statistical and heuristic approaches. Authors recommended the statistical linear regression model by analyzing correlation among BOP parameters as a reference state analysis method. Its advantage is that its derivation is not heuristic, it is possible to calculate model uncertainty, and it is easy to apply to an actual plant. The error of the statistical linear regression model is below 3% under normal as well as abnormal system states. Additionally a neural network model was recommended since the statistical model is impossible to apply to the validation of all of the sensors and is sensitive to the outlier that is the signal located out of a statistical distribution. Because there are a lot of sensors need to be validated in BOP, wavelet analysis (WA) were applied as a pre-processor for the reduction of input dimension and for the enhancement of training accuracy. The outlier localization capability of WA enhanced the robustness of the neural network. The trained neural network restored the degraded signals to the values within ±3% of the true signals.

I. INTRODUCTION

Now nuclear industry is under deregulation. On the basis of this background, the performance enhancement of secondary system of nuclear power plants (NPPs) called balance of plant (BOP) is being highlighted. The most interesting performance factor in BOP may be electrical output and heat rate. These performance factors are analyzed according to the performance test procedures provided from an authorized institution such as the American Society of Mechanical Engineers (ASME). However there are some problems in order to execute accurate performance tests even though qualified performance test procedures are used. It is difficult to measure the quality of saturated steam, so performance analysis is less accurate and more complicated. Moreover much more sensors are indispensable to calculate steam quality by mass and energy conservation, but even the quantity of sensors is not enough because minimum number of sensors for normal operation is installed in BOP [1]. This feature also makes signal validation difficult. In conclusion, the expansion of sensors is the first requirement for the enhancement of performance test accuracy. However the expansion of hardware may have concomitant economical and technical problems. We should trade off between the cost for new sensor installation and the benefit gained by performance tests. Considering cost-effectiveness, software-based solutions are noteworthy.

The signal validation using neural networks among software-based solutions may be one of the most general approaches in case of a black box model that is difficult to simulate analytically. The advantage of the black box model is that we don't have to know about the exact mechanism of a system. We do nothing but grasp the relation between the input and output signals. Many neural networks for signal validation have been proposed and they may be classified into two parts mainly. The one is for the selection of input parameters or the enhancement of training methods [2~5]. The other is the actual application [6~16]. The applications were focused on the estimation of the important parameters, for instance, feedwater flowrate or electric output. The analysis of system behavior, the diagnosis for trouble shooting, and the detection of fault signal were the important themes. These methodologies have the selection technique of input parameters, special training technique, and network structure to optimize training time and accuracy.

In this study, the statistical model based on linear regression will be proposed for the detection of sensor failure or degradation as a reference method. And then a neural network will be recommended to cover the shortcomings of the statistical model, especially robustness. The state analysis methodology using a neural network will be focused to compress input parameters because we must deal with a large quantity of sensors. We will be able to find strength of wavelet analysis as a compressing method.

II. DEVELOPMENT OF EMPIRICAL MODELS FOR STATE ANALYSIS

II.1. State analysis and BOP modeling

Signal validation is to inspect whether a signal represents the state of a system correctly or not. 'State' means any of various conditions characterized by definite quantities. A lot of performance indices in BOP such as electrical output or heat rate are calculated using state information collected by sensors. The typical process of signal validation is shown in Figure 1 [17]. The kernel of signal validation is to measure or estimate the same signal using various techniques, and compare them. If the deviation of more than a reference value is detected, the signal is regarded as 'failure' or 'degradation'. However it is difficult to carry out signal validation according to the scheme of Figure 1, because of the deficient sensors in BOP. Our unique alternative in Figure 1 is to use a BOP model and diverse signal sources. A BOP model is highly non-linear on the whole so that we cannot express its behavior by a simple state space model or an analytical model. Therefore authors considered BOP as a black-box model. In a black-box model, the only behavior of the input and output signals is investigated. This approach corresponds to 'empirical modeling' in Figure 1.

Performance tests are executed according to authorized procedures, for example, ASME performance test codes (PTCs). When ASME PTC 6 that is the highest accuracy procedure is used for a performance test, we must collect 2 ~ 400 signals [18]. As a result of actual investigation, there is only one half of or two-third of them installed permanently, and the others are installed and removed temporarily. If we have a suitable solution to solve the lack of sensors, we can reduce current workload. The signal validation explained above can be used as a solution because the empirical model in signal validation plays a role for the estimation of unmeasured signal.

In this study, authors expanded the concept of signal validation to state analysis by integrating above two topics, signal validation and signal estimation. Figure 2 describes the concept of state analysis [19]. It is an expanded figure of Figure 1. The function *f* in Figure 2 corresponds to an empirical model in Figure 1. 'Learned state' means the set of signals previously collected for the development of the empirical model. 'Observed state' means the set of signals collected each performance test. 'Estimated state' means the estimated output of the empirical model when an observed state is inputted. If an estimated state and an observed state are the same within a reference value, we regard that there is no failed or degraded sensor. Otherwise we should check the reliability of sensors in detail. The learned states should be collected to cover normal behavior as well as abnormal behavior of a system. In actual problems, it is difficult to collect these learned states because we cannot control normality or abnormality of a system. In this case we usually use a simulation code. For the data generation of steam turbine cycle, PEPSE (Performance Evaluations of Power System Efficiencies) code was used. Figure 3 is the lumped model of a typical BOP in NPPs, which is constructed by the visual tool of PEPSE code. The PEPSE code generated the learned states with changing boundary parameters according to the suggestion of performance engineer's guide [20]. From the results of the simulation, the data about total 28 parameters were collected.

II.2. Statistical state analysis model

We want to know the relation among signals that are collected in normal and abnormal system states. Of course, we assume all the learned states are reliable. In this model, the relation among signals is described using correlation matrix and linear regression. Let's assume the learned states gathered by the repeated PEPSE simulations are the universal set,

$$U = U(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p), \tag{1}$$

where $\vec{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iN})$,

p is the number of parameters,

N is the number of observations or the number of simulations.

We can calculate correlation coefficient r_{ij} between two parameters using Equation (2), (3), and (4).

$$\bar{x}_i = \frac{\sum_{k=1}^{N} f_k w_k x_{ik}}{\sum_{k=1}^{N} f_k w_k}, \tag{2}$$

$$s_{ij} = \frac{\sum_{k=1}^{N} f_k w_k (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{(\sum_{k=1}^{N} f_k) - 1}, \tag{3}$$

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii} s_{jj}}}, \tag{4}$$

where f_k is the frequency of the observations,

w_k is the weight.

Because the correlation coefficient is calculated for the universal set, r_{ij} represents the relation of parameters in all the system state. The magnitude of r_{ij} is proportional to the degree of correlation. In this study, the correlation among parameters is established by multivariate linear regression like Equation (5).

$$\hat{x}_j = b_0 + b_1 x_1 + \ldots + b_i x_i + \ldots + b_p x_p, \quad i \neq j \tag{5}$$

All of the x_is are the independent or explanatory variable and $\hat{x}_j$ is the dependent or response variable. We don't need to use all the parameters to construct a correlation. Considering computing capability and actual applications, we should reduce the number of explanatory variables. The selection of the best parameters is based on the correlation coefficients. In this study, maximum three parameters were selected as the explanatory variables, and we used the adjusted R^2, R_a^2 as a criterion for the selection of the best parameters. R_a^2 is defined as Equation (6),

$$R_a^2 = 100\left[1 - \left(\frac{N-1}{N-p}\right)\frac{SSE}{SST}\right], \tag{6}$$

where $SST = \sum_{j=1}^{N} w_j (x_j - \bar{x}_j)^2$, the corrected total sum of squares,

$SSE = \sum_{j=1}^{N} w_j (x_j - \hat{x}_j)^2$, the error sum of squares.

Table I shows the results when the response variable is the feedwater flowrate. Figure 4 shows the regression accuracy. Generally all of the correlations fall in $\pm 3\%$ of original feedwater flowrate.

The empirical model based on the statistical regression is easy to make and its derivation process is clear. Sensitivity study is also possible about each explanatory variable. Therefore it is effective for the estimation of some important parameters. However it is impossible to accomplish state analysis for all of the signals because we should check the abnormality of the explanatory variables repeatedly. Also the result of regression is sensitive to the outlier so we need another robust method for the support of the statistical model.

II.3. Hueristc state analysis model

A lot of researchers have used neural networks for state analysis. The most advantage of a neural network is that (1) we don't need to know the physical mechanism of a target system, (2) any system regardless of linear or nonlinear can be modeled, and (3) it has robustness. However we can use these advantages only when we select suitable input parameters, training method, neural network structure. First, we should reduce the dimension of input parameters for training efficiency and accuracy because there are a lot of sensors to be validated in BOP. Also we should prevent the loss of information during the reduction of input parameters. Additionally we need the method to localize the outlier that located outside a statistical distribution. The localization of the outliers means to prevent the deformation of a transformed domain by the outliers. This capability enhances the robustness of a neural network.

In this study, wavelet analysis (WA) was selected as a pre-processor of a neural network. We call the special functions satisfying some conditions as mother wavelets and father wavelets. From the mother and father wavelet, we can derive other wavelet functions and scale functions. They construct orthogonal bases. A wavelet function, $y_{b,a}$ and a scale function, $f_{b,a}$ are defined as Equation (7) and (8).

$$y_{b,a} = \frac{1}{\sqrt{a}} y\left(\frac{t-a}{b}\right), \qquad (7)$$

$$f_{b,a} = \frac{1}{\sqrt{a}} f\left(\frac{t-a}{b}\right), \qquad (8)$$

where y is the mother wavelet and f is the father wavelet,

b is a translation coefficient,

a is a scale coefficient.

If the translation coefficient and the scale coefficient are given as Equation (9), wavelets are discrete and construct orthogonal bases according to Equation (10) and (11).

$$a = 2^{-j},\ \ b = 2^{j} k\ , \qquad (9)$$

$$y_{j,k} = 2^{j/2} y\left(2^{j} - k\right), \qquad (10)$$

$$f_{j,k} = 2^{j/2} f\left(2^{j} - k\right), \qquad (11)$$

where $j, k \in Z$, Z is integer.

On the basis of Equation (10) and (11), we can do multi-resolution analysis. In multi-resolution analysis, we can analyze a function, f in an arbitrary dimension, specifically dyadic dimension, without any loss of information. In other words, we assume f is the approximation coefficients in a sufficiently higher resolution space. Then we determine the approximation and detail coefficients in projecting f to lower resolution spaces like Equation (12).

$$f = \sum_k c_{j,k} f_{0,k} + \sum_k \sum_j d_{j,k} y_{j,k} , \tag{12}$$

where $c_{j,k}$ is the approximation coefficient, $d_{j,k}$ is the detail coefficient.

The approximation coefficient is defined as Equation (13) using the father wavelet, and the detail coefficient is defined as Equation (14) using the mother wavelet.

$$c_{j,k} = \left\langle f, f_{j,k} \right\rangle = \int f \cdot f_{j,k} , \tag{13}$$

$$d_{j,k} = \left\langle f, y_{j,k} \right\rangle = \int f \cdot y_{j,k} . \tag{14}$$

The approximation coefficients represent the relation between a higher space and a lower space, and the detail coefficients represent the relation between a higher space and the compliment space of a lower space. A lower space has a dyadically reduced dimension comparing to a higher space according to Equation (10) and (11). By synthesizing these coefficients, the original data can be constructed completely without any information loss. We call the former method as wavelet decomposition and the later method as wavelet reconstruction. After all we do nothing but process the approximation and the detail coefficients instead of the original data.

In this study, 'haar' wavelet was selected as the mother and the father wavelet. In Figure 5, the scheme of wavelet decomposition is shown. According to Figure 5, 28-dimension of the original data is reduced to 4-dimension by selecting $j = 3$ and a total of 8 neural networks were prepared. Generally the only approximation coefficient is decomposed, but the detail coefficient is also decomposed to equalize the structure of the neural networks. This method is called wavelet packet analysis. We made 4-13-4 structure neural networks for training and testing. The training method was Levenberg-Marquardt back-propagation, and Bayesian regularization method was adopted to prevent overfitting. Figure 6 shows the results of training and testing for the feedwater flowrate. All the testing cases were nearly accurate.

Another testing was carried out after changing parameters intentionally, for the validation of robustness. Figure 7 shows the testing results when 80% of the original feedwater flowrate is inputted. The neural network with WA estimated the feedwater flowrate within $\pm$3% of the original value even if the estimated values are under-estimated. To identify the effect of the outlier, Figure 8 shows the first approximation coefficient, AAA and the first detail coefficient, AAD of the original signals and the signals after changing feedwater flowrate. As shown in Figure 8, the detail coefficients are nearly the same in both cases. In case of the approximation coefficients, the only single approximation coefficient, the mark ▼ is changed. In the mean while, other approximation coefficients are also nearly the same. This means the outliers in the feedwater flowrate signals are localized. Therefore the neural network with WA could compensate the outlier on the basis of training effectively.

III. Discussions and Conclusions

This study is focused on state analysis for BOP performance tests of NPPs. State analysis includes expansive concept of signal validation. State analysis detects and isolates the signal collected from failed or the degraded sensors, and provides estimated signals instead of the original low-quality signals.

In case of NPPs, the adoption of an empirical model was a unique alternative to achieve state analysis. We recommended the statistical model as one of the empirical models. The statistical model is based on correlation and regression among signals. Its advantage is that it is possible to construct an empirical model systematically, to calculate model uncertainty, and to apply to an actual system easily. However we need another supporting model since the statistical model is difficult to apply to the validation of all of the sensors and is too sensitive to the outliers in signals.

The neural network was proposed as a supporting heuristic method of the statistical model. For the reduction of numerical calculation, WA was adopted as a pre-processor. WA was useful to reduce the number of the input and output dimension. WA was also superior from the viewpoint of the outlier localization. Because of this, the neural network with WA compensated the outlier in the signals collected from failed or degraded sensors effectively.

IV. References

[1] G. Heo, S.J. Lee, S.H. Chang, and S.S. Choi, "Thermal Performance Test through On-line Turbine Cycle Performance Monitoring in Nuclear Power Plants" in *Proc. of the Korean Nucl. Soci.*, 1999 Fall.

[2] Jaques Reifman and Javier E. Vitela, "Accelerating Learning of Neural Networks with Conjugate Gradients for Nuclear Power Plant Applications," *Nucl. Tech*., vol. 106, pp. 225-241, 1994.

[3] Mahesh Lunani, Agus Sudjianto, and Pearse L. Johnston, "Generating Efficient Training Samples for Neural Networks using Latin Hypercube Sampling," *Intell. Eng. Sys. through Artificial Neural Networks : vol. 5 Fuzzy Logic & Evolutionary Programming*, pp. 209-214, 1995.

[4] Shufeng Tan and Michael L. Mavrovouniotis, "Reducing Data Dimensionality through Optimizing Neural Network Inputs," *AIChE Journal*, vol. 41, no. 6, pp. 1471-1480, 1995.

[5] J. Sola and J. Sevilla, "Importance of Input Data Normalization for the Application of Neural Networks to Complex Industrial Problems," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 1464-1468, 1997.

[6] Myung-Sub Roh, Se-Woo Cheon, and Soon-Heung Chang, "Thermal Power Prediction of Nuclear Power Plant using Neural Network and Parity Space Model," *IEEE Trans. Nucl. Sci.*, vol. 38, no. 2, pp. 866-872, 1991.

[7] Zhichao Guo and Robert E. Uhrig, "Nuclear Power Plant Performance Study using Neural Networks," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 4, pp. 915-918, 1992.

[8] Belle R. Upadhyaya and Evren Eryurek, "Application of Neural Networks for Sensor Validation

and Plant Monitoring," *Nucl. Tech*., vol. 97, pp. 170-176, 1992.

[9] Mostafa Khadem, Ali Ipakchi, Frank J. Alexandro, and Robert W. Colley, "Application of Neural Networks to Validation of Feedwater Flow Rate in a Nuclear Power Plant," *Proc. of the 55th American Power Conference*, pp. 842-847, 1993.

[10] Kadir Kavaklioglu and Belle R. Upadhyaya, "Monitoring Feedwater Flow Rate and Component Thermal Performance of Pressurized Water Reactors by means of Artificial Neural Networks," *Nucl. Tech.*, vol. 107, pp. 112-123, 1994.

[11] Yair Bartal, Jie Lin, and Robert E. Uhrig, "Nuclear Power Plant Transient Diagnostics Using Artificial Neural Networks That Allow 'Don't-Know' Classifications," *Nucl. Tech.*, vol. 110, pp. 436-449, 1995.

[12] Paolo F. Fantoni and Alessandro Mazzola, "Multiple-failure Signal Validation in Nuclear Power Plants using Artificial Neural Networks," *Nucl. Tech.*, vol. 113, pp. 368-374, 1996.

[13] Venkatramana N. Reddy, Paul M. Riley, and Michael L. Mavrovouniotis, "Analysis of Plant Measurements through Input-Training Neural Network," *Com. Chem. Eng*., vol. 20, suppl., pp. S889-S894, 1996.

[14] P.F. Fantoni and A. Mazzola, "A Pattern Recognition-Artificial Neural Networks Based Model for Signal Validation in Nuclear Power Plants". *Ann. Nucl. Energy*, vol. 23, pp. 1069-1076, 1996.

[15] Chang Hyun Roh, Hyun Sop Chang, Han Gon Kim, and Soon Heung Chang, "Identification of Reactor Vessel Failure Using Spatiotemporal Neural Network," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pp. 3223~3229, 1996.

[16] Robert E. Uhrig, Darryl J. Wrest, and J. Wesley Hines, "Intelligent Surveillance and Calibration Verification in Power Systems," *Int. Conf. on Intelligent System Application to Power Systems*, pp. 55-59, 1997.

[18] H.M. Hashemian, *On-line Testing of Calibration of Process Instrumentation Channels in Nuclear Power Plants*, NUREG/CR-6343, US NRC, 1995.

[19] W.A. Campbell, *Performance Test Codes 6 on Steam Turbine*, ASME, 1996.

[20] M.B. Caudill, R.D. Griebenow, E.J. Hansen, and G.C. Alder, "Application of Advanced Pattern Recognition to Power Plant Condition Assessment" in *Proc. of EI International*, 1996.

[21] Timothy Eckert, *Thermal Performance Engineer's Handbook: An Introduction to Thermal Performance*, TR-107422-V1, EPRI, 1998.

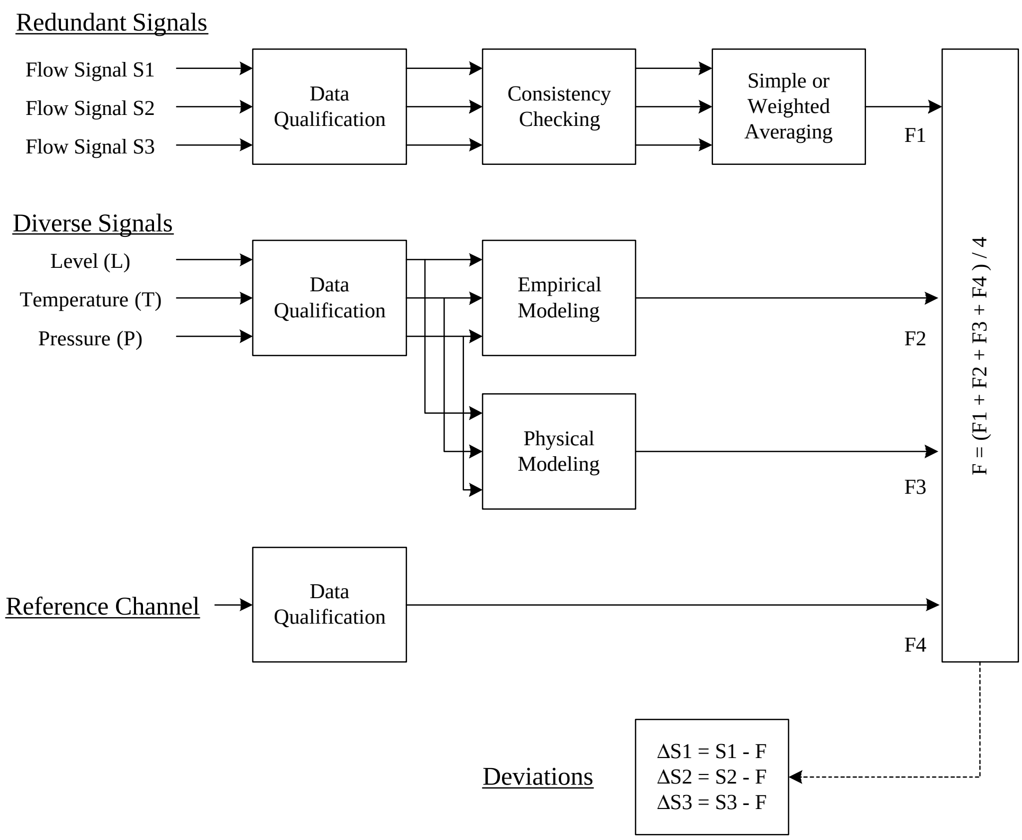

Figure 1. Concept of the signal validation proposed in NUREG/CR-6343

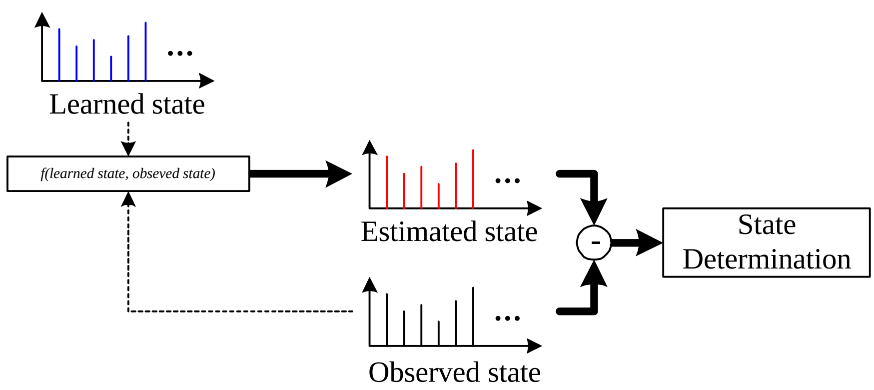

Figure 2. General concept of state analysis

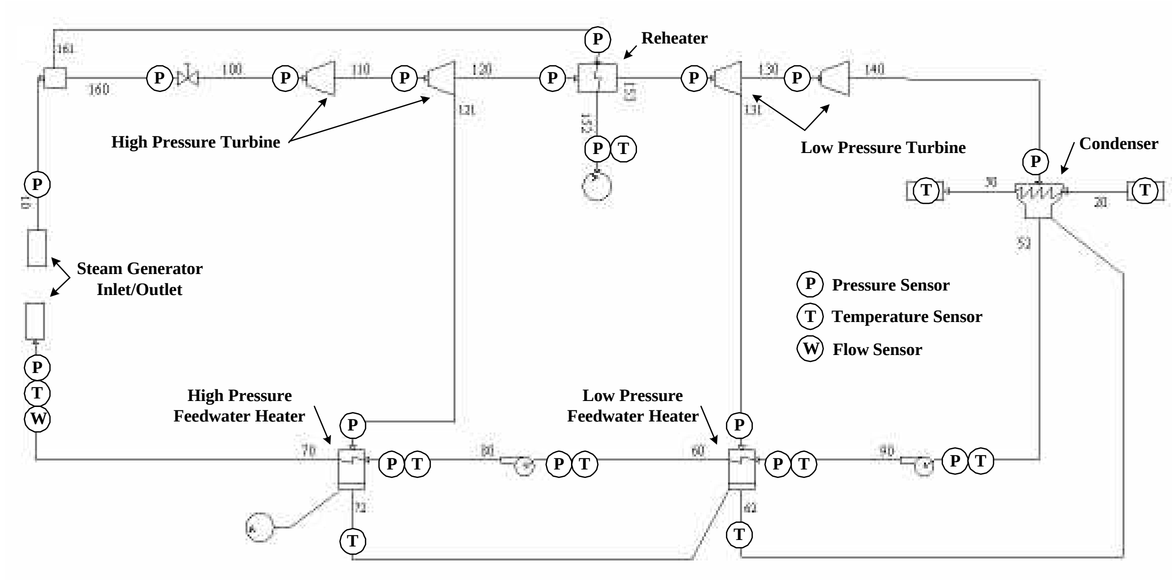

Figure 3. PEPSE model for turbine cycle simulation (the number means the stream number in PEPSE code)

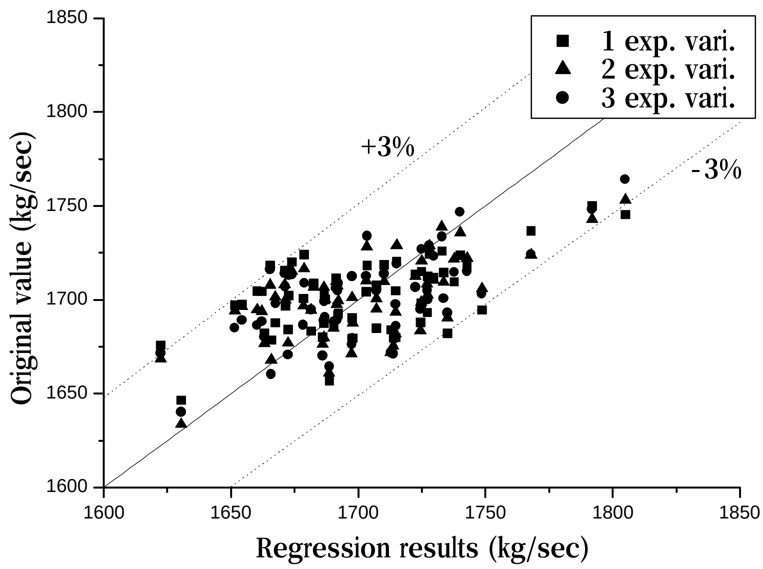

Figure 4. Regression accuracy of the feedwater flowrate correlation

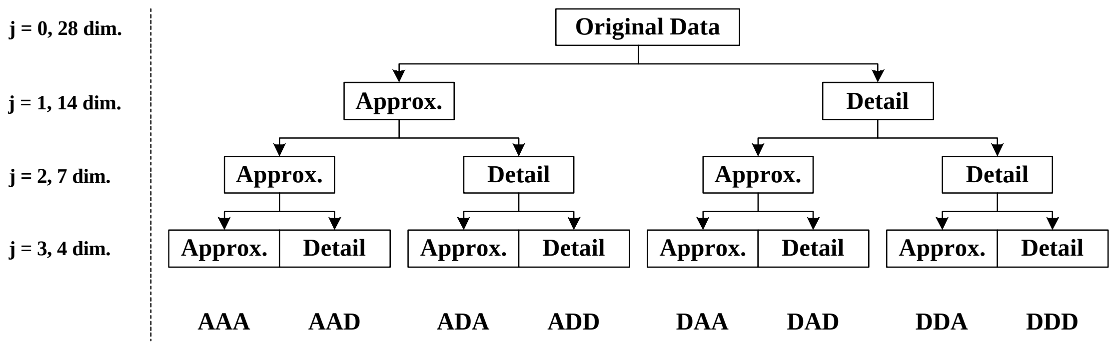

Figure 5. Wavelet packet analysis for the training and testing of a neural network

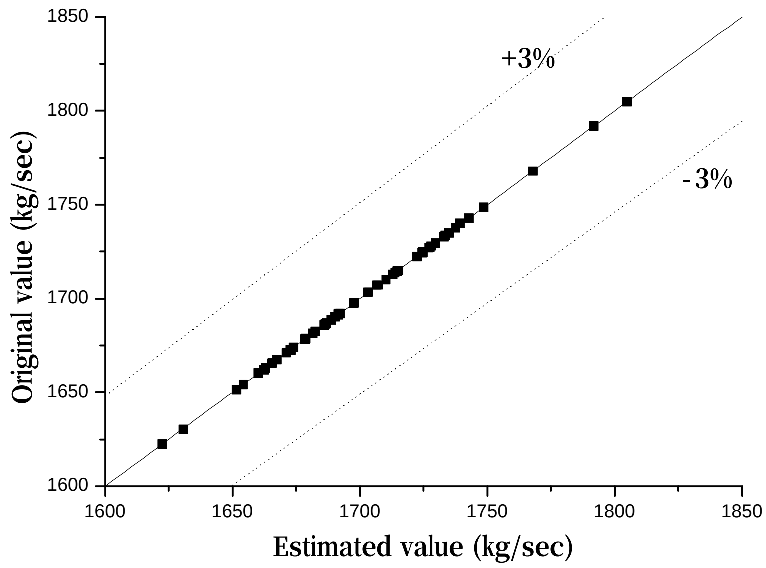

Figure 6. Testing accuracy for the feedwater flowrate

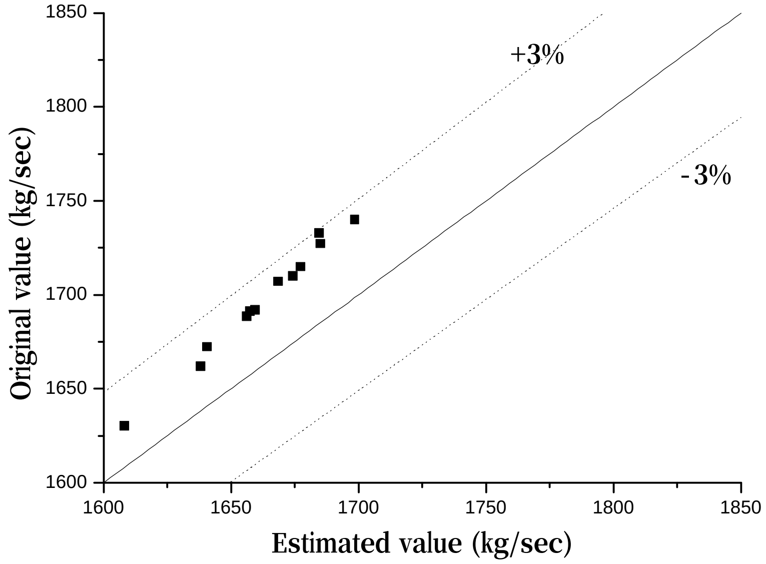

Figure 7. Testing accuracy for the 80% feedwater flowrate

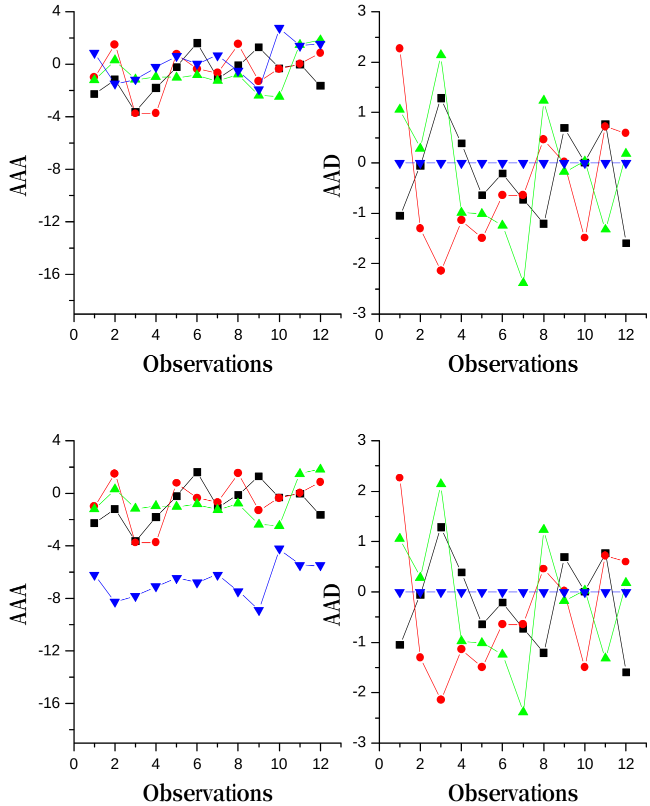

Figure 8. Comparison of the approximation and the detail coefficients under the outlier (Top: original data, Bottom: 80% of the feedwater flowrate)

Table I. Results of correlation analysis for the feedwater flowrate (flowrate of stream 70, an explanatory variable consists of a stream number and sensor type)

*	b_0	b_1	Exp. Vari.	b_2	Exp. Vari.	b_3	Exp. Vari.	R_a^2	SSE
1	864.247	0.713	153 Pres.					31.731	48,583
2	1369.243	0.693	153 Pres.	0.145	60 Pres.			38.120	44,037
3	1087.324	0.664	153 Pres.	0.160	60 Pres.	2.603	72 Temp.	41.788	41,426

* Number of explanatory variables