Transactions of the Korean Nuclear Society Autumn Meeting
Jeju, Korea, October 21-22, 2010

A Development of Empirical Models for Equipment Condition Monitoring System

Song Kyu Lee [a*], Se Jin Baik [a], Sang Ha An [b]
[a] *KEPCO Engineering & Construction Company, Inc., Deokjin-dong, Yusung-gu, Daejeon, 305-353, Korea*
[b] *Korea Advanced Institute of Science and Technology, Yusung-gu, Daejeon, 305-701, Korea*
[*] *Corresponding author: sklee3@kepco-enc.com*

1. Introduction

A great deal of effort is recently put into on-line monitoring (OLM), specially using empirical model to detect earlier the fault of components or the calibration reduction/extension of instrument. The empirical model is constructed with historical data obtained during operation and it mainly relies on regression techniques. Various models are used in OLM and the role of models is to describe the relation among signals that have been collected. Ultimate goal of empirical models is to best estimate parameter as soon as possible close to actual value. Typically some of the historical data are used for model training, and some data are used for verification and assessment of model performance. Several different models for OLM of nuclear power systems are currently being used. Examples include the ANL Multivariate State Estimation Techniques (MSET) used in EPI center of SmartSignal, the expert state estimation engine (ESEE) used in SureSense software of Expert Microsystems, Process Evaluation and Analysis by Neural Operators (PEANO) OECD of Halden Reactor Project [1] and linear regression model used in RCP seal integrity monitoring system (SIMON) of KEPCO E&C.

2. Methods and Results

In this section the methodology to develop empirical models being used in OLM and the comparison of estimated results are described. The linear regression as a parametric model and Kernel regression as a non-parametric model are reviewed.

2.1 Empirical Modeling Techniques

The data based modeling is categorized into parametric and non-parametric techniques. The parametric models, for example linear regression, polynomial models are defined by functional relationships among parameters. The non-parametric model is also a form of regression analysis in which the predictor does not take a predetermined form. When the relationships among parameters are not linear, non-parametric models are commonly used. However, these non-parametric models are only accurate at similar operating conditions under which the data were collected. If the non-parametric model should be used to extrapolate outside the learned space, the results cannot be trusted. Non-parametric models include Neural Networks, Kernel Regression, MSET and ESEE [1].

On-line monitoring using empirical model is a data intensive process, therefore the data to build model shall be error-free and representative of the system and it shall cover the entire operating range. Operating range of data is significant for non-parametric model because it is unable to extrapolate outside the range in which they have been trained. The outlier, missing data, stuck data, sensor drifts, etc. are typical data error in nuclear power plant. The data to be used in model training shall be cleaned and selected by removing outlier, filtering, averaging, etc. Fig. 1 shows empirical modeling and fault-detection procedure in OLM.

2.2 Linear Regression Model

The multivariate linear regression model can be used in OLM system. The linear regression model shall satisfy some statistical assumptions [2]. If not, the model may not be robust. Once a linear regression model has been constructed, it shall be confirmed the goodness of fit and the statistical significance of the estimated parameters. This work is performed by statistical packages.

The following is general linear model.

$$y_i = b_0 + b_{1i} + \cdots + b_p x_{pi} + e_i, i = 1, \cdots, n, \quad e_i = error \tag{1}$$

It is relatively simple, but may be robust and accurate than non-parametric model because extrapolation is available. The RCP seal integrity monitoring (SIMON) operating at Ulchin NPPs 3&4 utilizes the multivariate linear regression model.

2.3 Kernel Regression Model

A Kernel regression is a local weighted regression method using a weighting function and is estimating a parameter's value by calculating a weight average from historical data. Kernel assigns a weight to each location based on its distance from observation point. The followings are equations for Kernel regression; Equation (2) is Euclidean distance, equation (3) is Gaussian Kernel function and equation (4) is estimation of y with given x [3].

$$d_i(X_i, x) = \sqrt{(X_i - x)^2} \tag{2}$$

$$K_h(d) = \frac{1}{\sqrt{2\pi h^2}} e^{-d^2/2h^2} \tag{3}$$

$$\hat{y}(x)=\frac{\sum_{i=1}^{n}\left\{K_h\left[d_i(X_i,x)\right]Y_i\right\}}{\sum_{i=1}^{n}K_h\left[d_i(X_i,x)\right]} \qquad 4)$$

h is commonly referred to as the kernel's bandwidth and is used to control what effective distances are deemed similar [3].

2.4 Comparison of Estimated Results

Linear regression model and Kernel regression model were constructed with RCP seal operation data and their accuracy are compared. Fig. 1 shows the residuals by linear regression model (blue) and by Kernel regression model (red) for RCP seal No. 2 pressure. The residual is a difference of actual value and estimated value and zero value of residual means model accurately estimates. The residual by kernel model (standard deviation: 0.043) is slightly smaller than by linear model (standard deviation: 0.0575). However, it will not make a difference for OLM performance. Fig. 2 shows the residuals for RCP controlled bleed-off flow (RCR156); blue by linear regression model and red by Kernel regression. The residual by kernel model (standard deviation: 0.1673) is slightly bigger than by linear model (standard deviation: 0.1531) in contrast to the pressure. However, the difference will not also affect for OLM performance.

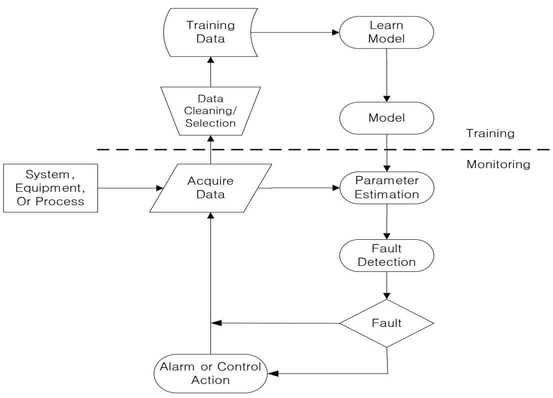

Fig. 1 Model estimation and fault-detection procedure

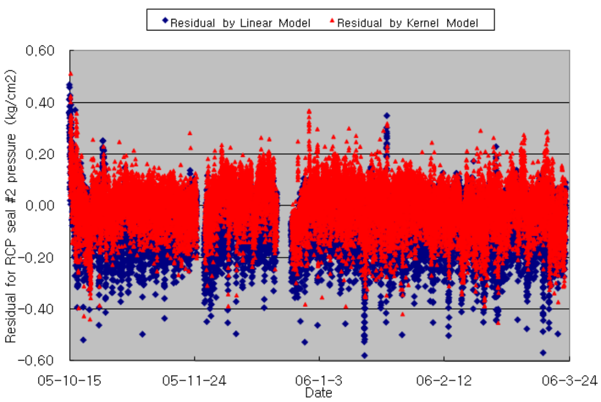

Fig. 2 Residual comparison for RCP seal No.2 pressure ((blue: linear model, red: Kernel model)

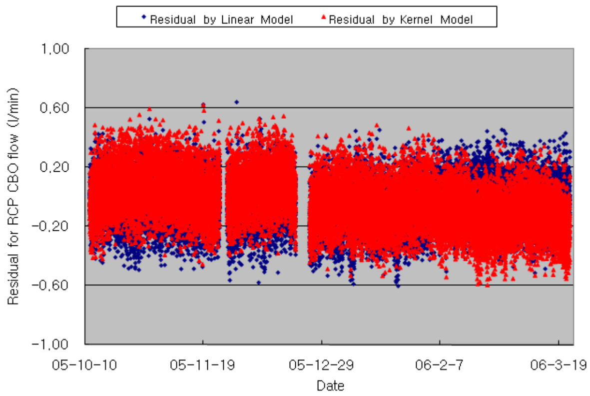

Fig. 3 Residual comparison for RCP CBO flow (blue: linear model, red: Kernel model)

3. Conclusions

This study was carried out to review the empirical model for OLM system. The linear regression (as a parametric) model and Kernel regression (as a non-parametric) model are constructed and their accuracy were compared. These models do not show much differences in estimation accuracy for RCP seal operation data. This is considered by that the relationship of RCP seal operation parameters (pressure, flow and temperature) represent a linear characteristic and the operating ranges for training data and observed data are same. Therefore, both linear regression model and Kernel regression model may be effectively applied to OLM for simple systems like RCP seal system. In some cases, the linear regression model may be advantageous than non-parametric model even though non-parametric models are being used widely in foreign OLM systems.

REFERENCES

[1] Technical Review of ON-Line Monitoring Techniques for Performance Assessment, U.S.NRC, NUREG/CR-6895, Vo;. 1, 2, 3, 2006.
[2] G.Y. Heo, Condition Monitoring Using Empirical Models: Technical Review and Prospects for Nuclear Applications, Nuclear Engineering and Technology, Vol. 40 No.1, February 2008.
[3] S.H. An, G.Y. Heo, S.H. Chang, Detection of Peocess Anomalies using an Improved Statistical learning Framework, Expert System Applications, Article in Press, 2010.