Estimation of a Reactor Core Power Peaking Factor Using Support Vector Regression and Uncertainty Analysis

In Ho Bae [a], Man Gyun Na[a*], Yoon Joon Lee [b], Goon Cherl Park [c]
[a] *Department of Nuclear Eng., Chosun Univ., 375 Seosuk-dong, Dong-gu, Gwangju, 501-759*
[b] *Department of Nuclear and Energy Eng., Cheju National Univ., 1 Ara-il-dong, Jeju-do, 690-756*
[c] *Department of Nuclear Engr., Seoul National Univ., 599 Gwanak-ro, Gwanak-gu, Seoul, 151-742*
[*] *Corresponding author: magyna@chosun.ac.kr*

1. Introduction

The monitoring of detailed 3-dimensional (3D) reactor core power distribution is a prerequisite in the operation of nuclear power reactors to ensure that various safety limits imposed on the LPD and DNBR, are not violated during nuclear power reactor operation. The LPD and DNBR should be calculated in order to perform the two major functions of the core protection calculator system (CPCS) and the core operation limit supervisory system (COLSS) [1]. The LPD at the hottest part of a hot fuel rod, which is related to the power peaking factor (PPF, F_q), is more important than the LPD at any other position in a reactor core. The LPD needs to be estimated accurately to prevent nuclear fuel rods from melting. In this study, support vector regression (SVR) and uncertainty analysis have been applied to estimation of reactor core power peaking factor.

2. Support Vector Regression and Uncertainty Analysis

2.1 Support Vector Regression (SVR) Model

SVR models are learning systems. They are optimized with a learning algorithm that originates from the theoretical foundations of statistical learning theory and structural risk minimization (SRM). SVR models use an SRM principle to minimize the upper bound on the expected risk, which is the sum of the empirical risk and of the confidence interval [2]. SVR models can be well applied to regression and classification problems. The regression problem is transformed to determine the coefficients of the basis function of linear expansion. The SVR models nonlinearly map the original input data $\mathbf{x}$ into higher dimensional feature space, $\boldsymbol{\varphi}(\mathbf{x})$. The SVR considers the following regression function:

$$y = f(\mathbf{x}) = \sum_{i=1}^{N} w_i \phi_i(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\varphi}(\mathbf{x}) + b \tag{1}$$

The function $\phi_i(\mathbf{x})$ is called the feature and the parameters $\mathbf{w}$ and $\boldsymbol{b}$ which are the support vector weight and bias. A nonlinear function is learned using a linear learning machine of which the learning algorithm minimizes a convex functional. The convex functional is expressed as a regularized risk function, and the parameters $\mathbf{w}$ and b are calculated by minimizing the risk function. The first line of Eq. (2) which is the regularized risk function is converted into the second line which is called a constrained risk function:

$$\begin{aligned} R(\mathbf{w},\boldsymbol{\xi},\boldsymbol{\xi}^*) &= \frac{1}{2}\mathbf{w}^T\mathbf{w} + \lambda \sum_{i=1}^{N} \left| y_i - f(\mathbf{x}) \right|_{\varepsilon} \\ &= \frac{1}{2}\mathbf{w}^T\mathbf{w} + \lambda \sum_{i=1}^{N} \left(\xi_i + \xi_i^* \right) \end{aligned} \tag{2}$$

subject to the constraints

$$\begin{cases} y_i - \mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}) - b \le \varepsilon + \xi_i, & i = 1, 2, \cdots, N \\ \mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}) + b - y_i \le \varepsilon + \xi_i^*, & i = 1, 2, \cdots, N \\ \xi_i,\ \xi_i^* \ge 0, & i = 1, 2, \cdots, N \end{cases} \tag{3}$$

where

$$\left| y_i - f(\mathbf{x}) \right|_{\varepsilon} = \begin{cases} 0 & \text{if } \left| y_i - f(\mathbf{x}) \right| < \varepsilon \\ \left| y_i - f(\mathbf{x}) \right| - \varepsilon & \text{otherwise} \end{cases} \tag{4}$$

The constrained optimization problem of Eq. (2) can be solved by applying the Lagrange multiplier technique to Eqs. (2) and (3), and using a standard quadratic programming technique. Finally, the regression function of Eq. (1) becomes

$$y = f(\mathbf{x}) = \sum_{i=1}^{N} \beta_i \boldsymbol{\varphi}^T(\mathbf{x}_i)\boldsymbol{\varphi}(\mathbf{x}) + b = \sum_{i=1}^{N} \beta_i K(\mathbf{x}, \mathbf{x}_i) + b \tag{5}$$

where $K(\mathbf{x}, \mathbf{x}_i) = \boldsymbol{\varphi}^T(\mathbf{x}_i)\boldsymbol{\varphi}(\mathbf{x})$ is known as the kernel function and the coefficient β_i is a function of the Lagrange multipliers α_i and α_i^*; $\beta_i = \alpha_i - \alpha_i^*$.

2.2 Uncertainty Analysis

There are several possible sources of uncertainty in predictions using data-based models; selection of training data, model structure including complexity, and noise in the input variables and the output variables [3]. Since an SVR model is developed using a given training data set, each possible training data set selected from the entire population of data will generate a different model and there will be a distribution of predictions for a given observation. Also, model misspecification takes place when a model structure is not correct, thereby introducing a bias.

The statistical bootstrap method works by generating many bootstrap samples of the training data set and retraining the SVR model parameters on each bootstrap sample. After repetitive sampling and training, the resulting predictions provide a distribution for the LPD value. This distribution can be used to calculate prediction intervals. There are two general algorithms for the bootstrap method: bootstrap pairs sampling and bootstrap residual sampling.

In this study, the bootstrap pairs sampling algorithm was used. The available data is divided into development data and test data. The development data consists of a large pool of data from which training and verification samples can be drawn. The test data is fixed. Uncertainty is separated into two types: variability and bias.

The pool of development data represents all available data, excluding the defined set of fixed test data. Since bias estimates based on the training data can be much lower than bias estimates based on an independent set of data, especially in case of an overfit model, one should compute bias estimates based on the data pool rather than the training data. The estimate with a 95% confidence interval for an arbitrary test input $\mathbf{x}_o$ using a bootstrap method is

$$\hat{y}_0 \pm 2\sqrt{Var(\hat{y}_0) + bias^2} = \hat{y}_0 \pm \delta . \quad (6)$$

The estimate with a 95% confidence interval using an analytical method is

$$\hat{y}_0 \pm 2s\sqrt{1 + \mathbf{f}_0^T \left(\mathbf{F}^T \mathbf{F}\right)^{-1} \mathbf{f}_0} = \hat{y}_0 \pm \delta . \quad (7)$$

3. Application to reactor core power peaking factor

The proposed SVR was applied to the first fuel cycle of the YGN-3 PWR plant. Two kinds of data sets consist of the positive ASI cases (12,765 points) and the negative ASI cases (12,766 points), Two SVR models are optimized for two kinds of data sets. The data obtained from simulations of the MASTER code comprises a total of 25,541 input-output data points $(x_1, x_2, \cdots, x_{11}, y_r)$ or $(x_1, x_2, \cdots, x_{14}, y_r)$, depending on whether the SPND signals are used or not. In OPR1000 nuclear power plants, the CPCS and the COLSS calculate the LPD to protect and monitor nuclear plants. When the SPND signals are not used, the proposed algorithm can be utilized as a protection algorithm. The training data was selected using subtractive clustering (SC) scheme after the test data was removed from the pool of acquired data. An SVR model can be optimized well by using informative data.

To conduct an uncertainty analysis by the bootstrap method, 100 sample sets for training and verification are selected by randomly adjusting the radius r_α of the SC scheme in a specified range. The prediction intervals are very small, which means that the predicted values are very accurate. Also, the prediction intervals of the analytic method are about 2 or 3 times larger than those of the bootstrap method. The prediction intervals of the bootstrap method can have very high peak values at several test data points because the SVR models developed by 100 random sample training data sets have a large variance at the test data points. Fig. 1 shows the estimation errors and their prediction intervals. Table 1 shows other test results to compare the PPF values calculated from the proposed SVR method and the COLSS method.

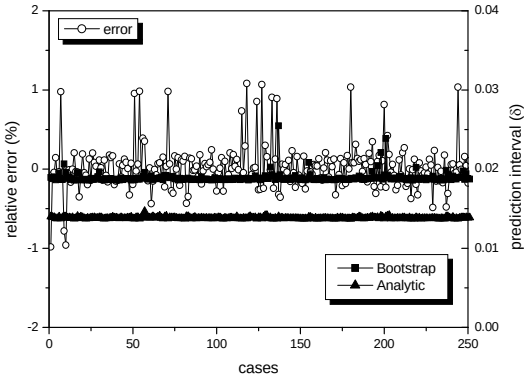

Fig. 1. Prediction Intervals of the SVR Model for Positive ASI (With SPND Signals)

Table I: Comparison of Calculated PPF Values.

ASI value	Power	MASTER (target)	SVR Model (with SPND)	COLSS
0.081	80	1.968	1.964	2.133
0.094	90	1.959	1.955	2.135
0.069	100	1.952	1.947	2.137
0.073	103	1.949	1.946	2.138
-0.525	80	2.778	2.779	3.000
-0.504	90	2.718	2.716	2.961
-0.483	100	2.663	2.658	2.918
-0.520	103	2.646	2.642	2.905

Also, PPFs were estimated by SVR models when the SPND signals are used and not used respectively. It is known that the RMS error calculated by the SVR model for the test data is similar to the RMS error for the verification data. Therefore, if the SVR models are first optimized using data for a variety of operating conditions, they can accurately estimate PPFs for other operating conditions.

4. Conclusions

The RMS error of the estimated PPF values is about 0.15%. In addition, their uncertainty was analyzed by a bootstrap method using 100 sampled training data sets and verification data sets and analyzed by an analytical method. The prediction intervals are very small, which means that the predicted values are very accurate. As a result, the SVR models are accurate enough for use in core protection using power peaking factors and the monitoring of power peaking factors.

REFERENCES

[1] ABB Combustion Engineering Inc., "Overview Description of the Core Operation Limit Supervisory System (COLSS)," CEN-312-P, Revision 01-P, 1986.
[2] V. N. Vapnik, Statistical Learning Theory. New York, NY: John Wiley & Sons, 1998.
[3] J.W. Hines, B. Rasmussen, "Online sensor calibration monitoring uncertainty estimation," Nuclear Technology, vol. 151, pp. 281-288, Sept. 2005.