Uncertainty Analysis of NPP LOCA Size Prediction Using SVR

Jae Hwan Kim [a], Sim Won Lee [a], Dae Seop Kim [a], Man Gyun Na [a*], Kwang-Il Ahn [b], Soo Yong Park [b]
[a]*Nuclear Engineering Dept., Chosun Univ., Seosuk-dong, Dong-gu, Gwangju 501-759, Republic of Korea*
[b]*Korea Atomic Energy Research Institute, Dukjin-Dong, Yuseong-gu, Daejoen 305-348, Republic of Korea*
[*]*Corresponding author: magyna@chosun.ac.kr*

1. Introduction

When transients or accidents occur in the nuclear power plants, it is important to figure out the initial event by observing major parameters during the beginning of event. However, if the severe accident happens, it is really hard for operators to predict the scenarios of event. Because, they have only few time to analyze the event and are offered limited information. Therefore, providing information such as a break size is essential to control this event successively. This study predicts the break size in the hot-leg, cold-leg and steam generator tubes using an SVR model when the Loss of Coolant Accident (LOCA) occurs, and performs the uncertainty analysis of this proposed model.

2. Prediction of the LOCA Size Using SVR

2.1 Support Vector Regression Method

The SVR model [1] which is used widely to solve the problem of non-linear regression is designed to provide valuable information about break size of LOCA for operators of the nuclear power plants. So, it makes successful performance of LOCA management possible when the LOCA occurred. The SVR model is a method which is mapping the source data for high-dimensional characteristic space non-linearly. This model can be expressed as follows:

$$y = f(x) = \sum_{i=1}^{N} w_i \phi_i(x) + b = w^T \phi(x) + b$$
$$w = [w_1 w_2 \cdots w_N]^T, \ \phi = [\phi_1 \phi_2 \cdots \phi_N]^T \quad (1)$$

Equation (1) is a non-linear regression model because the resulting hyper-surface is a non-linear surface hanging over m -dimensional input space. However after the input vectors x are mapped into vectors $\phi(x)$ of high dimensional kernel-induced feature space, the non-linear regression model is turned into linear regression model in feature space. The non-linear function is learned using a linear learning machine of which the learning algorithm minimizes a convex functional. The convex functional is expressed as the following regularized risk function, and the parameters w and b are the support vector weight and bias that are calculated by minimizing the risk function [2]:

$$R(w) = \frac{1}{2} w^T w + \lambda \sum_{i=1}^{N} \left| y_i - f(x) \right|_\varepsilon \quad (2)$$

where

$$\left| y_i - f(x) \right|_\varepsilon = \begin{cases} 0 & \text{if } \left| y_i - f(x) \right| < \varepsilon \\ \left| y_i - f(x) \right| - \varepsilon & \text{otherwise} \end{cases} \quad (3)$$

The constant λ is known as a regularization parameter. The regularization parameter determines the trade-off between the approximation error and the weight vector norm.

The regression model of the non-linear function has been learning by the linear learning model after mapping to the characteristic space. The learning algorithm optimizes an SVR model by minimizing the convex functional. In order to minimize upper bound on the expected risk, The SVR model uses the method of Structural Risk Minimization (SRM).

Existing methods use the principle of Empirical Risk Minimization (ERM), whereas the principle of Structural Risk Minimization (SRM) can optimize algorithm by searching the minimum of risk defined as the sum of empirical risk and confidence interval. The optimization of an SVR model is achieved by genetic operations, and the weighting vector and bias of an SVR function is calculated by using Lagrange multiplier technique and standard quadratic programming technique.

2.2 Application to the prediction of the LOCA size

To apply the proposed algorithm, it is essential to acquire the data which is required to train the SVR model from a number of numerical simulations, because there is few real LOCA data. The number of 540 accident simulations were carried out by using MAAP4 code to acquire data, and composed of 270 hot-leg LOCAs, 270 cold-leg LOCAs and 270 SGTRs.

3. Uncertainty Analysis of the Proposed Model

To develop a data-based model, uncertainty analysis is required to prove how accurate the prediction is. In this paper, we use statistical and analytical methods.

3.1 Statistical Method

The statistical uncertainty analysis works by creating many bootstrap samples of the training data set and retraining the data-based model parameter. After sampling and training repeatedly, the resulting predictions provide the distribution for the output value. This distribution can be used to calculate the prediction intervals. In this study, the bootstrap pairs sampling algorithm is used to analyze the statistical uncertainty. The acquired data is divided into a development data (data pool) and test data (fixed value). Development data consists of large pool of the data which can extract training and verification samples. Calculation process of bootstrap pairs sampling algorithm is as follows:

$$Var(\hat{y}_0)\frac{1}{J-1}\sum_{j=1}^{J}\left[\hat{y}_o^j-\bar{\hat{y}}_0\right]^2 \tag{4}$$

where

$$\bar{\hat{y}}_0=\frac{1}{J}\sum_{j=1}^{J}\hat{y}_0^j \tag{5}$$

$$\text{bias}=\left\{\frac{1}{K}\sum_{k=1}^{K}\frac{1}{J}\sum_{j=1}^{J}\left[y_k^j-\hat{y}_k^j\right]^2\right\}^{1/2} \tag{6}$$

The estimate with a 95% confidence interval for an arbitrary test input x_0 can be expressed as follows:

$$\hat{y}_0\pm 2\sqrt{Var\left(\hat{y}_0\right)+\text{bias}^2}=\hat{y}_0\pm\delta \tag{7}$$

3.2 Analytic Method

Basically, the SVR model can be expressed as follows:

$$y_k=f(x_k,\theta)+\varepsilon_k \tag{8}$$

The matrix F is called the Jacobian matrix of first order partial derivatives with respect to the parameters determined from the least squares. The variance of the predicted output can be estimated as follows [3]:

$$Var(y_0-\hat{y}_0)\approx\sigma^2+f_0^TSf_0\approx s^2+s^2f_0^T(F^TF)^{-1}f_0 \tag{9}$$

The prediction with 95% confidence interval is calculated as follows:

$$\hat{y}_0\pm 2s\sqrt{1+f_0^T\left(F^TF\right)^{-1}f_0}=\hat{y}_0\pm\delta \tag{10}$$

4. Conclusions

In this paper, an SVR model was developed to estimate the NPP LOCA size. Also, prediction intervals were calculated by using analytic and statistical uncertainty analysis. The RMS errors of test data in hot-leg LOCA, cold-leg LOCA and SGTR are 0.7749%, 0.9509% and 3.2768%, respectively. As shown in the Fig. 1, prediction of hot-leg LOCA size using an SVR model fits very well. Also, if the SVR model is optimized by using a variety of data, it is possible to estimate the NPP LOCA size more accurately.

REFERENCES

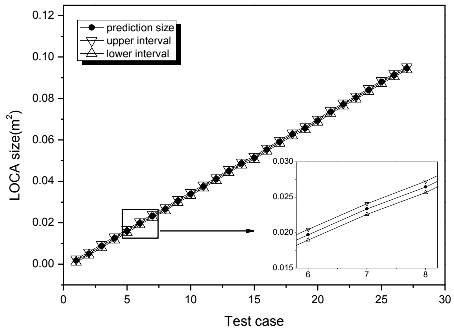
[1] M. G. Na, W. S. Park, and D. H. Lim, "Detection and diagnostics of loss of coolant accidents using support vector machines," IEEE Trans. Nucl. Sci., vol. 55, no. 1, pp. 628-636, Feb. 2008.

[2] V.N. Vapnik, Statistical Learning Theory. New York, NY: John Wiley & Sons, 1998.

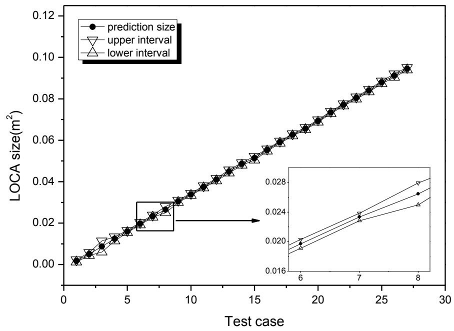
[3] R. Tibshirani, "A comparison of some error estimates for neural network models," Neural computation, vol. 8, pp. 152-163, 1996.

Table I. Performance of the proposed SVR

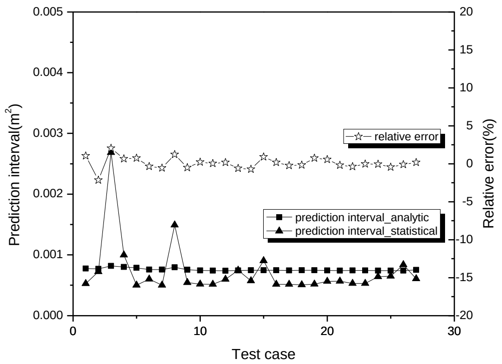
Event type	Data type	RMS error (%)	Relative Max. error (%)
Hot-leg LOCA	Training data	1.2014	10.4016
	Verification data	0.3683	1.0882
	Test data	0.7749	2.1472
Cold-leg LOCA	Training data	2.8866	26.5232
	Verification data	0.2643	0.6029
	Test data	0.9509	3.7441
SGTR	Training data	2.4582	22.4975
	Verification data	1.1754	4.0683
	Test data	3.2768	12.1289



(a) Analytic



(b) Statistical



(c) Prediction intervals and relative error

Fig. 1. Prediction intervals of hot-leg LOCA size using an SVR model