Estimation of Cutter Wear of a Milling Machine Using a Support Vector Regression Method

Young Do KOO[1], Man Gyun NA*[1], and Jung-Taek KIM[2]

1. Department of Nuclear Engineering, Chosun University, 309 Pilmun-daero Dong-gu, Gwangju 61452, Republic of Korea (magyna@chosun.ac.kr)
2. Nuclear ICT Research Division, Korea Atomic Energy Research Institute, 989-111 Daedeok-daero Yuseong-gu, Daejeon 34039, Republic of Korea (jtkim@kaeri.re.kr)

Abstract: The integrity of various machineries and equipment, which are used in several industrial fields, can be considered as a factor determining safety and efficiency such as productivity and economy. In the sense, prognostics and health management (PHM) techniques are used to estimate the machine condition and life for effective maintenance and risk minimization. It is known that PHM has already been applied in a wide range of industrial fields such as automotive industry, aeronautics industry, several energy industries, and military.
PHM denotes diagnostics that monitors conditions of the mechanical system or device and detects its failure symptoms, prognostics of remaining useful life (RUL), and effective health management using sensors.
However, in a case of nuclear power plants (NPPs) that consist of large architecture and complex internal structures, it is known that the actual failure data of its system and device are difficult to be obtained, compared to other industrial fields. Thus, since the cutter in a milling machine can be considered as a rotor such as a pump and a turbine in a NPP, the cutter wear data from PHM 2010 society conference data challenge were used in an effort to study PHM technologies for NPPs.
In this study, support vector regression (SVR) as a data-driven approach was used to estimate a total of 6 cutter wear of a milling machine. The basic concept of SVR is to map the input data into a high-dimensional space by nonlinear mapping to solve a linear regression problem in this space. Among the cutters, 3 actual cutter wear data were compared with estimated values obtained from the SVR method using the signals from dynamometer, accelerometers, and acoustic emission sensors built in the experimental device. Consequently, the proposed method can accurately estimate the degree of cutter wear in the experimental device and it is expected that SVR has a capability to estimate wear of the rotating machines in NPPs in the future.
Keyword: Prognostics and Health Management (PHM), Remaining Useful Life (RUL), Sensor, Support Vector Regression (SVR)

1 Introduction

Nuclear power plants (NPPs) are comprised of a lot of various equipment and components. Among the total number of 24 NPPs operating in Republic of Korea, pressurized water reactors (PWRs) such as optimized power reactor (OPR) 1000 and advanced power reactor (APR) 1400 are main reactor types in Korea.

Generally, these PWRs are classified into a primary system exposed to radioactivity and a secondary system without radioactivity. The primary system in NPPs has a role to transfer heat energy of the reactor and this heat energy from the primary system is changed into potential energy as steam through steam generator (SG) in the secondary system.

There are facilities such as reactor vessel (RV), reactor coolant pump (RCP), pressurizer (PZR), SG primary side, and so on in reactor coolant system (RCS), which is another name of the primary system. They are under high pressure and temperature conditions to prevent H_2O coolant with radionuclides from vaporization. The secondary system consists of turbine, condenser, pump, the rest side of SG, and so on, which operate under less harsher conditions.

In an effort to keep the integrity of these various NPP components, a study on application of prognostics and health management (PHM) technology to NPPs was carried out[1] since the equipment have an important role on the safety and efficiency of NPPs.

PHM can be defined as accurately monitoring mechanical system, device, and facility, detecting the fault symptom, and predicting remaining useful life (RUL) using the features extracted from the data or sensor signals (refer to Fig. 1[1]). That is, PHM system contains functions of condition monitoring, state assessment, fault detection, prognostics, operational decision support, and so on[2].

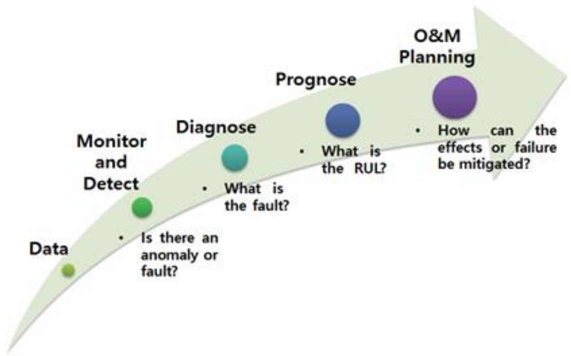

Fig.1 PHM systems.

PHM technologies have already been applied in various industrial fields such as defense, aircraft, automobile, and wind turbines. Accordingly, in the nuclear power generation field, it is considered that a PHM technology become a necessary need to enhance the safety and economy of NPPs moving forward with both long term operation (LTO) and new builds[1].

Unfortunately, however, it is known that there are rarely the real fault data, and thus it is difficult to be obtained. Therefore, three types of sensor signals such as force, vibration, and acoustic emission (AE) data were used to estimate the cutter wear of a milling machine in this study.

Although the cutter of a milling machine is not one of the components in NPPs, it was selected for estimation target in this study since it can be considered a rotating machine such as pump and turbine in NPPs.

Moreover, it can be regarded as a case study since the performance of the artificial intelligence technology used in this study was able to be checked due to the fact that cutter wear was estimated using the real sensor signals and compared with the actual cutter wear data.

Support vector regression (SVR), which is one of the machine learning method of artificial intelligence, was used in this study. It is known that this technique was applied to various regression analyses and showed good performance.

However, a lot of data were needed to guarantee the optimal performance using machine learning technique, and accordingly the problems on calculation time and overfitting were raised.

Therefore, subtractive clustering (SC) technique[3] and genetic algorithm[4][5] , which are for data selection and optimization, were applied to SVR to confirm the performance through proper generalization. For these reasons, the accuracy of the proposed SVR model increased and the precise estimation performance of the cutter wear was shown.

2 Machine learning algorithm

SVR used in this study is another name of support vector machine (SVM) in regression analysis[6]. SVM is a fundamentally machine learning method and data-driven approach. This method seeks to best fit the training data in establishing the mapping function, while maintaining the ability to generalize to unseen data.

However, since the efficacy of regression algorithm is determined by data type, quality, and quantity, a lot of data have to be used to gain good performance[7]. Additionally, the performance differs from the assumptions of error minimization or training method inherent in the algorithm[7].

2.1 Support vector machines in regression

SVM as a model generally used in event classification or regression problem is an algorithm with a neural network structure based on statistical learning theory. Current embodiment of these SVM was proposed by C. Cortes and V. Vapnik[6].

It is noted that SVM and artificial neural network (ANN) techniques have similar structure. However, they have differences on training method or risk minimization[8]. SVM uses a structural risk minimization (SRM) principle (refer to Fig. 2) to minimize the upper bound on the expected risk[8].

In other words, it can be possible to establish the optimized SVR algorithm by a SRM principle finding the minimum of the bound on risk $R(f^*)$ defined as the sum of empirical risk and the confidence interval. This SRM is depicted as follows[9]:

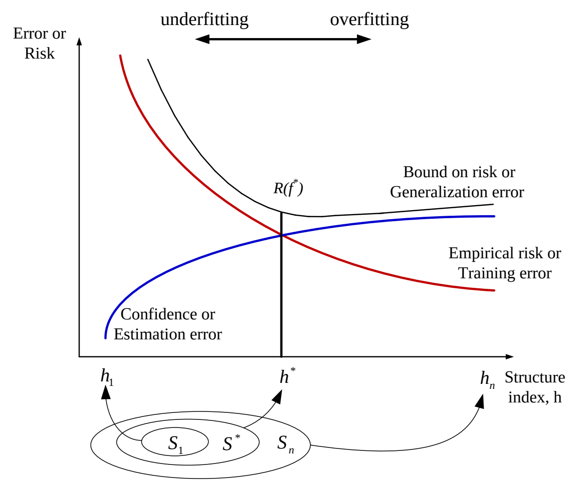

Fig.2 SRM principle used in SVM.

With the introduction of Vapnik's ε-insensitive loss function[8][9][10], SVM has been finally extended to be used in regression analysis. This can be applied to a time series forecast and a nonlinear regression. To be specific, the basic idea of SVR is to map the input data into a kernel-induced higher feature space, and then perform linear regression analysis[12]. That is, nonlinear regression in the input space can become linear regression in a feature space.

In this study, the SVR model is established using N training data indicated as $T = \{(x_k, y_k)\}_{k=1}^{N}$ in which x_k is a sample data vector and y_k denotes the actual output value, from which it learns a relationship between input and output values. The SVR function is expressed as follows:

$$\hat{y} = f(x) = \sum_{k=1}^{R} w_k \phi_k(x) + b = W^T \Phi(x) + b \tag{1}$$

where $\phi_k(x)$ is termed the feature which is the function nonlinearly from the input space x, $W = [w_1 \ \ w_2 \ \ \cdots \ \ w_N]^T$, $\Phi = [\phi_1 \ \ \phi_2 \ \ \cdots \ \ \phi_N]^T$. Parameters W and b are weighting coefficient and bias, respectively.

These weighting coefficient and bias are computed using the following regularized risk function with the ε-insensitive loss function[11]. To acquire them, the following regularized risk function has to be minimized.

$$R(W) = \frac{1}{2} W^T W + C \sum_{k=1}^{N} |y_k - f(x)|_\varepsilon \tag{2}$$

where

$$|y_k - f(x)|_\varepsilon = \begin{cases} 0 & \text{if } |y_k - f(x)| < \varepsilon \\ |y_k - f(x)| - \varepsilon & \text{otherwise} \end{cases} \tag{3}$$

The term of the ε-insensitive loss function in Eq. (2) is determined according to the condition such as Eq. (3). As shown in Fig. 3, the insensitiveness ε has a role to decide the ε-tube size and to stabilize the estimation by controlling the number of support vectors[11]. In addition, the parameter C as regularization parameter in Eq. (2) decides the trade-off between the weight vector norm and the approximation error[11]. In the SVR model, the parameters ε and C are design parameters and they are related to generalization and overfitting.

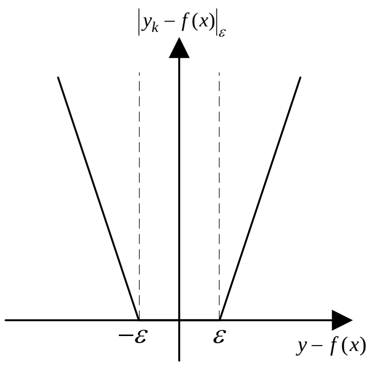

Fig.3 A linear ε-insensitive loss function applied to SVR.

The aforementioned generalized risk function is altered into a constrained risk function with the slack variables as follows:

$$R(W, \xi, \xi^*) = \frac{1}{2} W^T W + C \sum_{k=1}^{N} \left(\xi_k + \xi_k^*\right) \tag{4}$$

subject to the constraints

$$\begin{cases} y_k - W^T\Phi(x) - b \le \varepsilon + \xi_k, & k = 1, 2, \cdots, N \\ W^T\Phi(x) + b - y_k \le \varepsilon + \xi_k^*, & k = 1, 2, \cdots, N \\ \xi_k,\ \xi_k^* \ge 0, & k = 1, 2, \cdots, N \end{cases} \tag{5}$$

Likewise, Eq. (4) is used to compute W and b, and this can be solved using the Lagrange multiplier method and a standard quadratic programming technique.

The slack variables shown in Fig. 4 are located outside the ε-tube size and indicate the upper and lower constraints. These non-zero values have to be minimized.

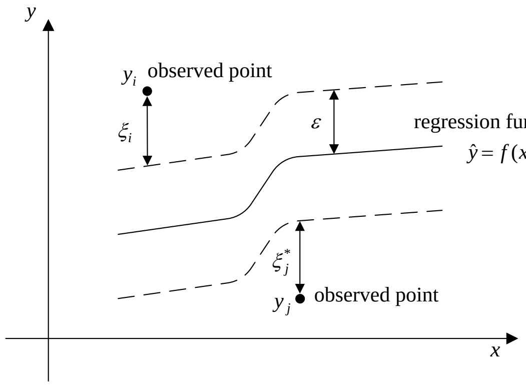

Fig.4 An illustration of the ε-insensitive loss function with slack variables.

Beyond linear regression, the SVR model can be applied to nonlinear regression analysis using the kernel. In this study, the radial basis function is used as the kernel function and defined as follows:

$$K\left(x_k, x\right) = \exp\left(-\frac{(x_k - x)^T(x - x_k)}{2\sigma^2}\right) \tag{6}$$

The parameter σ in Eq. (6) indicates the sharpness of the radial basis kernel function and it is one of the design parameters as the parameters ε and C, affecting on SVR's performance.

Finally, the SVR function using the kernel function becomes as follows:

$$\hat{y} = f(x) = \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)K\left(x_k, x\right) + b \tag{7}$$

The several Lagrange multipliers $\alpha_k - \alpha_k^*$ have non-negative and non-zero values and the corresponding training data are regarded as support vectors (SVs) lying on or outside ε-bound.

2.2 Optimization of SVR

To optimize the SVR model performing linear regression in a kernel-induced space, the genetic algorithm[4][5] was used in this study. The design parameters such as ε, C, and σ were optimized using the genetic algorithm as a technique imitating an evolutionary process of living organisms by the natural evolution mechanisms such as selection, crossover, and mutation (refer to Fig. 5).

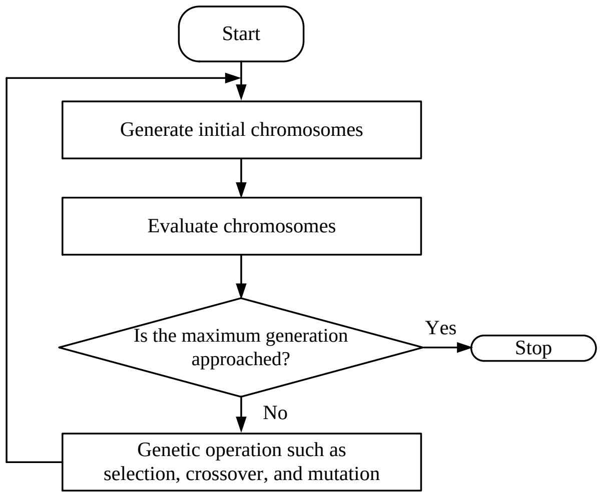

Fig.5 Optimization procedure using the genetic algorithm.

The genetic algorithm needs a fitness function used to minimize the root mean square error (RMSE) and maximum error by assigning a score to each chromosome in the corresponding population and evaluating how suitable a chromosome is. The fitness function is used as Eq. (8).

$$F = \exp(-\boldsymbol{\lambda}_1 E_1 - \boldsymbol{\lambda}_2 E_2) \tag{8}$$

where E_1 and E_2 are RMSE and maximum error for the development data set, respectively. $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are weights for each error.

$$E_1 = \sqrt{\frac{1}{N_{dev}}\sum_{k=1}^{N_{dev}}\left(y_k - \hat{y}_k\right)^2} \tag{9}$$

$$E_2 = \max_k (y_k - \hat{y}_k)^2, \; k = 1, 2, \cdots, N_{dev} \qquad (10)$$

where N_{dev} is the number of development data points, and y_k and $\hat{y}_k$ are a target value and an estimated value using the SVR model.

In this study, the development data were used to develop the SVR model for estimation of cutter wear. All of the data for the SVR model as well as the development data are described in Section 3.

3 Data application to SVR model

The machineries and facilities of NPPs are considered as important factors determining the safety and economy. Since the SVR model is known for a machine learning technique showing a good performance in various regression analysis, it can be used to estimate the state of NPP components.

However, this method is not easy to be applied to a nuclear power generation field due to a lack of actual fault data and a difficulty on acquiring the data. Therefore, the data from "PHM 2010 society conference data challenge[13]" are applied to the SVR model to estimate the wear of three cutters among the six cutters with 3-flute of a computer numerical control (CNC) milling machine. A CNC milling machine is comprised of a cutter, workpiece, accelerometer sensor, dynamometer sensor, and AE sensor as shown in Fig. 6.

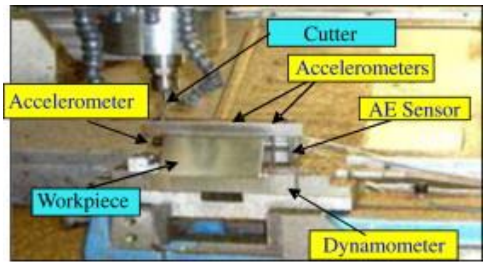

Fig.6 High speed CNC milling machine (Röders Tech RFM760).

Although the cutter of a milling machine as the estimation target is not directly related to the components of NPPs, the effort for diagnosis and estimation of the state for equipment using the sensor signal data can be considered as improving the safety of NPPs. In addition, the SVR model is verified since the estimated cutter wear data can be compared with the actual wear data.

For these reasons, the cutter wear data from PHM 2010 data challenge were applied to the SVR model, and the detailed description on the data and application result are stated as follows.

3.1 Composition of cutter wear data

The signals applied in the SVR model of this study were obtained from the built-in sensors of accelerometer, dynamometer, and AE in a milling machine testbed.

A total of 315 data files comprised of these types of real-time sensor signals were used to estimate the wear of three cutters. It is noted that the number of data files mean the amount of wear after each cut for six cutters. That is, each cutter made 315 cuts.

Table 1 Sensor signals of a CNC milling machine

No. of channels	Signals
1	Force in X
2	Force in Y
3	Force in Z
4	Vibration in X
5	Vibration in Y
6	Vibration in Z
7	Acoustic emission (AE)

The built-in sensor signals of a milling machine are concretely divided into seven channels such as force and vibration in X, Y, and Z directions, and AE signals as indicated in Table 1. There are more than 200,000 measured signals in every data file for each cutter.

Additionally, the actual wear data are given for three cutters (first, fourth, and sixth cutters) among the every cutter, and accordingly these cutters was able to be compared with the estimated cutter wear. However, only the sensor signals on the rest of cutters (second, third, and fifth cutters) were provided for competition participants to compare with their estimation results.

3.2 Data selection scheme for the SVR model

Among more than 200,000 points in the data, RMS, standard deviation (STD), and peak values of the points were used. This is to make the model track the trend of the cutter wear well.

That is, since one of the main factors to show the optimal performance on fitting is the data, the subtractive clustering (SC)[3] technique was used to effectively train the SVR model by selecting the data, which are considered informative, as a cluster center.

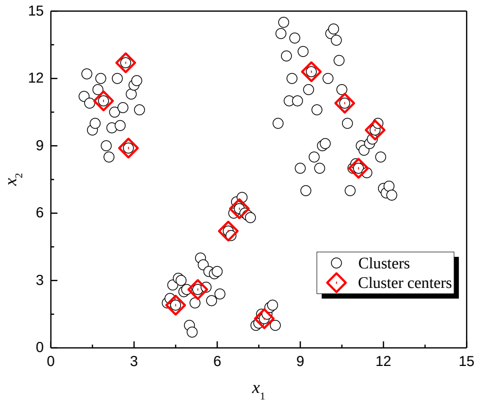

Fig.7 Graphical description of SC technique.

As expressed in Fig. 7, by calculating the potential of data points as a determining factor of instructive data, SC technique selects a data with the highest potential as a cluster center.

The SC technique regards every data point as a cluster center and the amount potential of input data is defined as a function of Euclidean distances to all other data points as follows:

$$P_1(k) = \sum_{i=1}^{N} \mathbf{e}^{-4\|x_k - x_i\|^2 / \mathbf{r}_\alpha^2}, \quad k = 1, 2, \cdots, N \qquad (11)$$

where r_α indicates a radius that defines the proximity between the points. This radius has a considerable influence on the input data potentials. Thus, after the potentials of all points were calculated, the data point with the highest potential was selected as a first cluster center. Generally, the potential of a data point is high in case that there are many adjacent data points.

Next cluster center was determined by defining the potential of a data point as well. However, considerable potentials were subtracted for each data point as a function of its distance from the pre-selected cluster center. Eventually, the potential for data points positioned near the pre-selected cluster center considerably decreased. This is to unlikely to make the data points near the pre-selected cluster center a next cluster center.

To find the data point with the highest revised potential as a next cluster center, the potentials of every data point are modified by the following function:

$$P_{i+1}(k) = P_i(k) - P_i^* \mathbf{e}^{-4\|x_k - x_i^*\|^2 / \mathbf{r}_\beta^2}, \quad k = 1, 2, \ldots, N \qquad (12)$$

where x_i^* indicates the point for the i-th cluster center and P_i^* is the corresponding potential value. Creating the cluster centers was repeated until the number of them is equal to the number of each data set, used to develop the estimation model of SVR.

3.3 Estimation result

To verify the proposed model to estimate the cutter wear of a milling machine using the chosen data, the data in this study were separated into the training data, the validation data, and the test data. This is an effort to prevent the SVR model from overfitting.

The training data set and the validation data set, which is to optimize the SVR model, were directly related to the development of the model. The test data set had no effect on training and was used to independently provide a measure of performance after training. Additionally, the development data set consisted of the training and verification data as stated above.

Table 2 Estimation performance of cutter wear using SVR

Data type	RMSE (%)	No. of points
Training	2.70E-05	833
Verification	5.60E-05	94
Test	3.59E-04	15
Development	3.16E-05	833+94

Table 2 shows the RMSE of the cutter wear estimation and the number of the used data assigned to the SVR model. The estimation performance of the cutter wear of a milling machine using the SVR model can be considered outstanding.

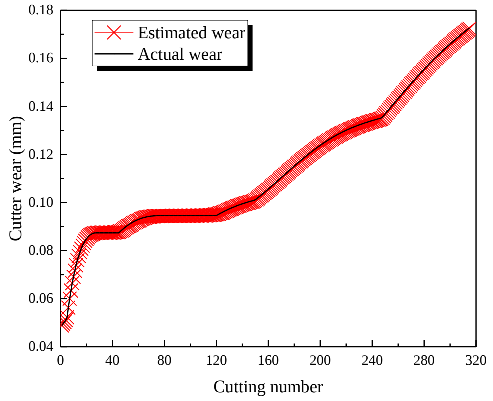

Fig.8 Estimation performance for cutter 1 using SVR.

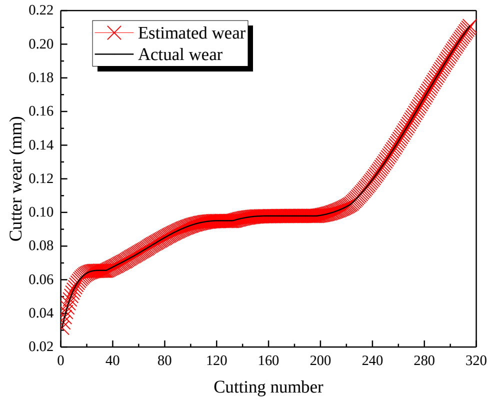

Fig.9 Estimation performance for cutter 4 using SVR.

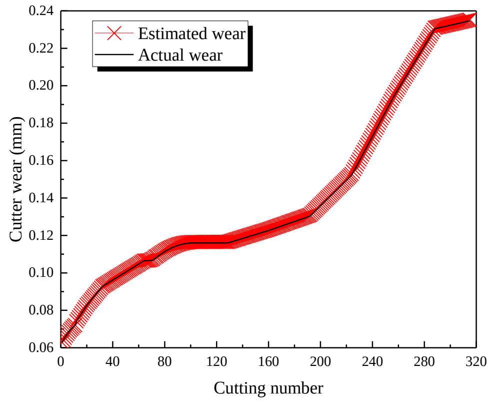

Fig.10 Estimation performance for cutter 6 using SVR.

Furthermore, the estimation accuracy of the SVR model can be checked through Figs. 8-10 for the number 1, number 4, and number 6 cutters, respectively. A red line with 'X' symbol, which is an estimation line of a cutter wear, accurately catch up with a trend for an actual cutter wear of a black line. In addition, the actual cutter wear data as the target values were given for 3 flutes. Among them, the target wear was the wear data of the most damaged flute for each cut.

4 Conclusions

In an effort to apply a PHM technology to NPP industries, the study on estimation of the state of the equipment using the SVR technique, which is generally applied to various regression analyses, by sensor signals was carried out.

However, since it is known that the actual fault data for NPPs and are hard to be obtained, the data provided from "2010 PHM data challenge" were used to the SVR model with the genetic algorithm. Although a cutter wear of a milling machine as the estimation target is not directly related to the NPPs, it can be considered as a rotating machine such as a turbine and a pump in a NPP.

In addition, the SVR model can estimate a cutter wear and be verified since the data from sensor signals such as force, vibration, and AE, and the actual cutter wear data are given. RMSEs on every data set of the SVR model, have very low values. Consequently, the proposed SVR model of a machine learning technique can be a suitable model as an on-line monitoring (OLM) and a PHM technology to diagnose and prognose the state of the equipment using the sensor signals. Moreover, if the real data from NPPs are applied to the proposed model, it is expected that the present study will be helpful to increase the safety of NPPs.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant, funded by the Korean Government (MSIT) (Grant No. 2016M2A8A2953046).

References

[1] J. Coble, P. Ramulhalli, L. Bond, J. Hines, and B. Upadhyaya, Prognostics and Health Management in

Nuclear Power Plants: A Review of Technologies and Applications, Richland, Washington, Pacific Northwest National Laboratory (PNNL), 2012.
[2] J. Sheppard, M. Kaufman, and T. Wilmering, “IEEE Standards for Prognostics and Health Management”, Proceedings of AUTOTESTCON, Salt Lake City, Utah, United States, Sep. 8~11, 2008.
[3] S. Chiu, "Fuzzy Model Identification Based on Cluster Estimation", Journal of Intelligent & Fuzzy Systems, Vol. 2, pp. 267-278, 1994.
[4] D. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Reading, Massachusetts, 1989.
[5] M. Mitchell, An Introduction to Genetic Algorithms, Cambridge, Massachusetts, 1996.
[6] C. Cortes and V. Vapnik, “Support-Vector Networks”, Machine Learning, Vol. 20, pp. 273-297, 1995.
[7] J. Hines, J. Coble, and B. Upadhyaya, “Application of Monitoring and Prognostics to Small Modular Reactors”, Proceedings of Future of Instrumentation International Workshop (FIIW), Oak Ridge, Tennessee, United States, Nov. 07~08, 2011.
[8] S. R. Gunn, Support Vector Machines for Classification and Regression, Southhampton, University of Southampton, 1998.
[9] V. Kecman, Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models Cambridge, Massachusetts, 2001.
[10] V. Vapnik, The Nature of Statistical Learning Theory, New York, 1995.
[11] K. Y. Chen and C. H. Wang, "Support Vector Regression with Genetic Algorithms in Forecasting Tourism Demand", Tourism Management, Vol. 28, pp. 215-226, 2007.
[12] M. G. Na, W. S. Park, and D. H. Lim, "Detection and Diagnostics of Loss of Coolant Accidents Using Support Vector Machines", Nuclear Engineering and Technology, Vol. 55, pp. 628-636, 2008.
[13] PHM Society, PHM data challenge, Oct., 2010; https://www.phmsociety.org/competition/phm/10, 2010.