Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM)
31 May – 1 June 2018, Singapore
Editor(s) Xudong Qian, Sze Dai Pang, Ghim Ping Raymond Ong, Kok-Kwang Phoon

APPLICATION OF MACHINE LEARNING TECHNIQUES TO EARTHQUAKE DAMAGE ESTIMATION AND EMERGENCY SHUTOFF OF LIFELINE SYSTEMS

NOBUOTO NOJIMA[1], TATSUYA MORIYAMA[2] AND KATSUSHIGE OHNISHI[3]

[1] *Department of Civil Engineering, Gifu University, Yanagido 1-1, Gifu, Japan.*
E-mail: nojima@gifu-u.ac.jp
[2] *Graduate School of Engineering, Gifu University, Yanagido 1-1, Gifu, Japan.*
[3] *Graduate School of Engineering, Gifu University, Yanagido 1-1, Gifu, Japan.*

Applicability of machine learning techniques has been examined for estimation of damage of low-pressure gas pipelines and decision making of emergency shutoff of city gas supply. A number of observation patterns of SI values, damage rate and shutoff patterns was generated by Monte Carlo simulation. The relationships between SI values and damage rate in training data was learned using support vector regression analysis. The relationships between SI values and shutoff status was learned using support vector machine. The results using test data suggests that the applied techniques can be promising tools for representing non-linear relationships among those factors related to damage estimation and shutoff decision.

Keywords: city gas supply system, SI sensor, supply shutoff judgement, machine learning, support vector machine, support vector regression

1 Introduction

When strong ground motion exceeding a certain level is detected and extensive damage to pipeline network is expected, emergency supply shutoff is conducted in city gas supply systems to prevent secondary disaster. Threshold of observed SI (spectral intensity) value to activate immediate supply shutoff is ordinarily set to SI*=60 cm/s. As shown in Figure 1, multiple SI sensors are equipped to large supply block to cover the spatial variability in wide area. In that case, when k out of n sensors in a supply block detect SI > SI*, the supply block is shutoff. Such system is referred to as "k-out-of-n shutoff system." The advanced real-time system for emergency shutoff system is composed of huge geospatial database of pipeline network configuration, pipe material, pipe joint, and pipe diameter, ground conditions, network of SI sensors, computer system and algorithm for evaluation of damage on the basis of observed SI values.

Considering the primary objective of emergency supply shutoff, truly effective activation of shutoff is that the supply block is isolated when the damage rate r exceeds a certain value r^* which is regared as a dangerous level. To accomplish this, real-time damage estimation must be performed after collecting observed SI values. In an emergency situation, however, such real-time operation should not require too much computer tasks. For this purpose, it is required to clarify the direct relationships between observed SI values and damage rate of pipeline, and also necessity of emergency shutoff.

In this study, applicability of machine learning techniques has been examined for estimation of damage to low-pressure gas pipelines and decision making of emergency shutoff of city gas supply. As indicated by dashed and solid frames in Figure 1, support vector regression (SVR) is applied for prediction of damage rate, and support vector machine (SVM) is applied for emergency shutoff judgement.

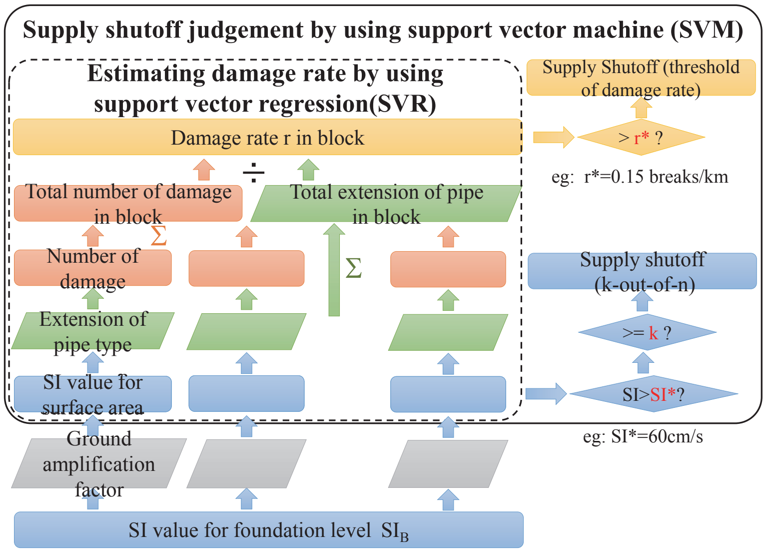


Figure 1. Damage rate and shutoff judgement and application area of SVR and SVM.

2 Data generation by Monte Carlo Simulation

In this study, 10000 sets of distribution patterns of SI values (velocity spectra with 20% damping averaged over period of 0.1-2.5s), pipe damage patterns and supply shutoff patterns are generated by use of MCS. As an input ground motion, baserock SI value SI_B is assumed to be uniform in a supply block. SI_B values ranging from 5 to 100 cm/s are uniformly divided into the number of simulation trials of MCS, N_{sim}=10000. Surface SI values at each of n SI sensors in a supply block are simulated by multiplying lognormal random numbers representing amplification factor of surface ground to each SI_B value level.

Next, damage rate R[breaks/km] and number of damage DN are simulated by applying damage prediction equation for low-pressure gas pipelines given by the following equation.

$$R(SI) = C_p \cdot C_g \cdot R_0 \cdot \Phi\left(\frac{\ln SI - \lambda}{\zeta}\right) \qquad DN = R(SI) \cdot L \qquad (1)$$

where C_p is a correction coefficient for pipe material, C_g is a correction coefficient for ground type, Φ is probability distribution function for standard normal distribution, R_0=2.36, λ =4.298 and ζ =0.387 are model parameters, and L is extension of pipelines [km]. Figure 2 illustrates the standard damage function $R(SI)$ with C_p=C_g=1.

Two types of uncertainty are considered for the prediction. One is aleatory uncertainty associated with the randomness of the number of damage predicted for given damage rate. The probability that the number of damage becomes x is modeled by Poisson distribution with average damage rate λ. The other is epistemic uncertainty associated with damage prediction equation itself. In order to incorporate probabilistic variation of the parameter λ, a random variable Λ following gamma distribution as a conjugate distribution of Poisson distribution. The prediction distribution of Poisson distribution is obtained as negative binomial distribution whose PMF is given by the following equation.

$$\mathrm{NB}(x;k,\theta)=\int_0^\infty \mathrm{Po}(x;\lambda)\cdot f_\Lambda(\lambda;k,\theta)d\lambda=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{\lambda^{k-1}e^{-\lambda/\theta}}{\Gamma(k)\theta^k}d\lambda=\frac{\Gamma(x+k)}{x!\Gamma(k)}\left(\frac{1}{1+\theta}\right)^k\left(\frac{\theta}{1+\theta}\right)^x \quad (2)$$

The number of damage is obtained as a random number N_{NB} following the negative binomial distribution. The damage rate is obtained by $\lambda_{NB}=N_{NB}/L$. Figure 3 shows all the samples of damage rate generated by 10000 simulation trials. Three cases are compared: (a) Deterministic, (b) Poisson distribution, (c) Negative binomial distribution (CoV=20%).

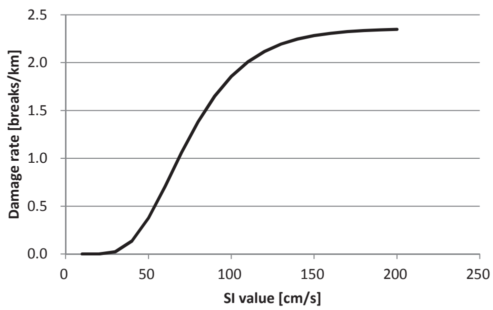


Figure 2. Standard damage rage function

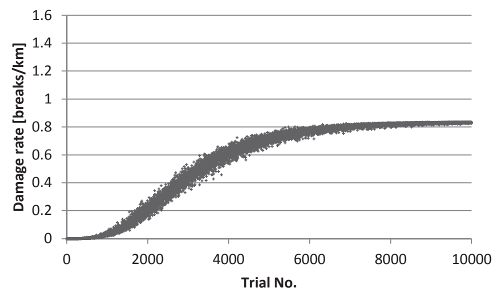
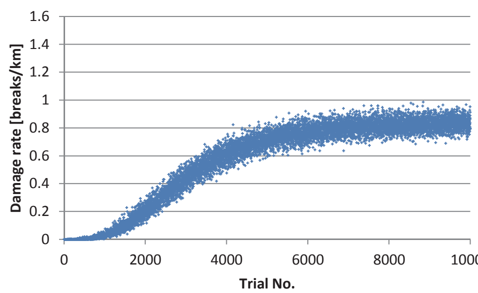
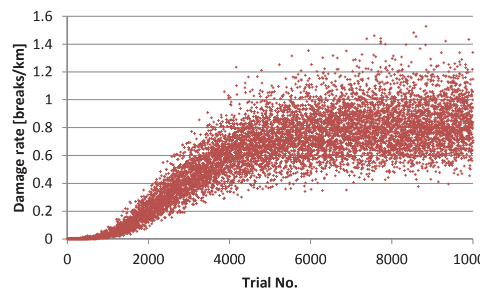


Figure 3. (a)Deterministic



(b) Poisson distribution



(c) Negative binomial distribution (CoV=20%)

Figure 3. Simulation samples of damage rate (Horizontal axis is the trial number for N_{sim}=10000 corresponding to S_{IB}=5-100 cm/s).

3 Support Vector Regression (SVR) for damage rate estimation

The 10000 sets of surface SI values and pipeline damage rate mentioned in the previous chapter are divided into training data and test data. The number of training data is 9-fold: 50, 100, 200, 500, 1000, 2000, 5000, 7500 and 9000. The relationships between surface SI values and pipeline damage rate in the training data are learned, and regression model by use of SVR is constructed. Then the obtained SVR model is applied to the test data, and the results of predicted values of damage rate are compared with the hidden answers generated by MCS.

Three kinds of datasets of damage rate are considered: datasets A, B and C, corresponding to Figure 3 (a)-(c). Three patterns of learning and prediction are compared: pattern AA (learning A and predicting A), pattern CA (learning C and predicting A) and pattern CC (learning C and predicting C).

The RBF (Radial Basis Function) kernel is used as a kernel function for SVR. Wide range of hyper parameters, γ for the RBF kernel, the regularization parameter C and the insensitiveness parameter for the ε-intensive loss function have been explored. It has been found that the results are not so sensitive to ε and C, but are relatively sensitive to γ. In this paper, results are shown for the case where the number of training data is 2000 and the parameter C=5 and ε=0.05.

Figure 4 shows the results for multivariate regression analysis. In all patterns for learning and prediction, the relationships between the predicted and observed values are strongly skewed since the non-linear relationships in the damage function, complex relationships among pipeline configurations and ground condition are not adequately evaluated. Figure 6 shows the results for SVR with γ=0.02. In Pattern AA, prediction error is generally small. Non-linearity in the input-output relationship of input ground motion and resultant damage rate is adequately learned. In Pattern CA, the performance of SVR prediction is very well particularly at low range of damage rate. However, at high of damage rate, prediction error becomes large because of the aleatory uncertainty of the randomness of occurrence of damage considered in the training data. Furthermore, the prediction error becomes much larger in Pattern CC. In order to examine the effect of the parameter γ, Figure 6 shows the results for γ=0.5 which narrows the width of RBF kernel function than γ=0.02. The prediction error generally becomes larger with some exception in each pattern.

4 Support Vector Machine (SVM) for judgement of supply shutoff

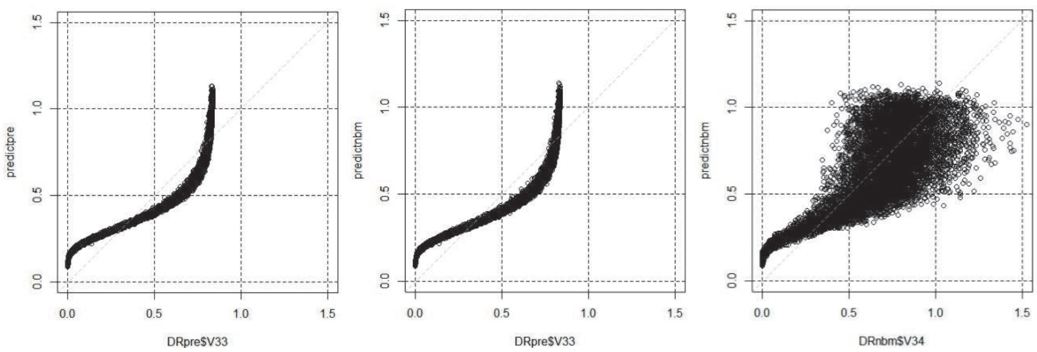
Using the same training data mentioned in the previous chapter, the relationship between surface SI values and whether damage rate is in dangerous level ($r>r^*$) or not ($r<r^*$) is learned, and binary classifier by use of SVM is constructed. Then the SV classifier is applied to the test data, and the results of binary classification associated with the damage level are compared with the hidden answers generated by MCS. Specifically, 2×2 contingency table compiling the case counts of combination of true/false and positive/negative judgement. The performance of classification is evaluated using four kinds of rate: True Positive (TP) rate, False Positive (FP) rate, Positive Predictive (PP) value and Negative Predictive (NP) value.

The RBF kernel function with γ=0.5, which shows lesser performance in Figure 6 than γ=0.02 in Figure 5, is used for demonstration. The other two parameters are C=5 and ε=0.05 as in the previous chapter. The boundary value of damage rate for binary classification of damage level is set to r^*=0.15.

As shown in Figure 7 (a), all of the three indices, TP rate, PP value and NP value, exceed 95%. The value of FP rate is larger than 10% with the number of training data up to 100. However, FP rate becomes smaller with increasing number of training data; FP rate becomes less than 5% with more than 1000 test data. These results suggest that fairly well performance of SV classification can be extracted when enough training data are provided.

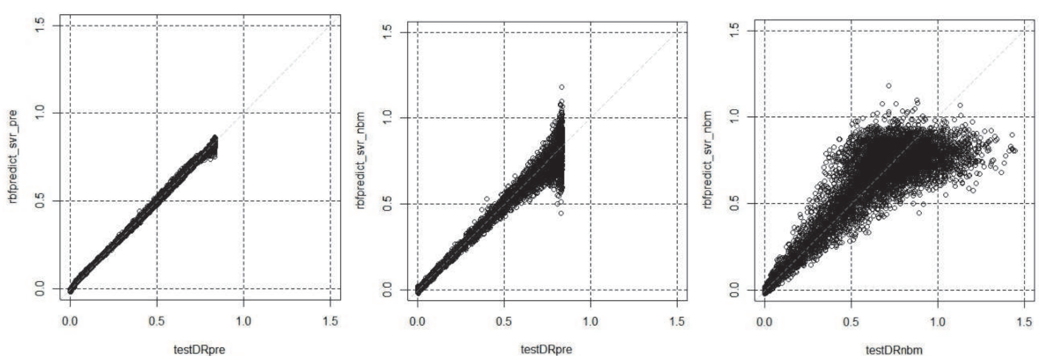
For comparison purpose, the results obtained by applying surface SI values directly to the damage function is shown in Figure 7 (b). In this particular case, there is no effect of machine learning. Essentially the indices should take the same value regardless of the number of training data. Small fluctuations are due to randomness in dividing the training data and test data from the entire Monte Carlo simulation samples. The difference between Figure 8 (a) and Figure 7 (b) is small except for FP rate with test data less than 200. In other words, the performance of SVM is almost equivalent to the damage function without uncertainty.

The accuracy of shutoff judgement by SVM is considered to be better than that of estimated damage rate by SVR. This fact implies that the reliability of k-out-of-n shutoff system is enhanced by collective observations by multiple SI sensors in a block. The SV classifier may pave the way for a new type of shutoff system without using SI* or k but using damage rate r^*.



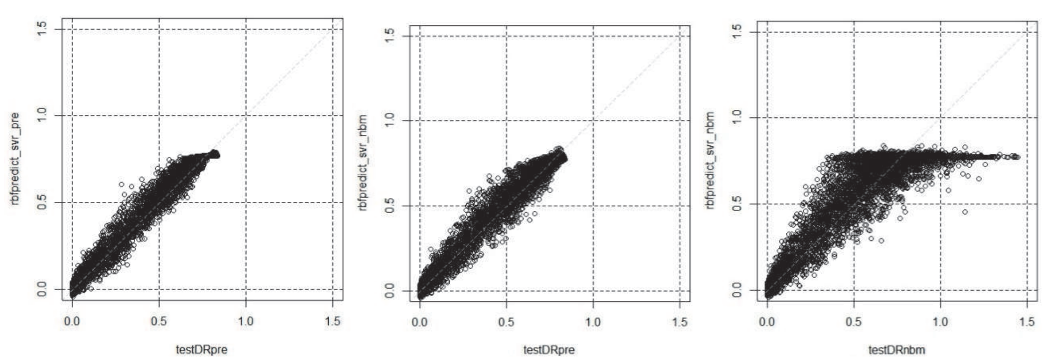
(a) Pattern AA (b) Pattern CA (c) Pattern CC

Figure 4. Prediction of damage rate by multivariate regression (X: observed, Y: predicted).



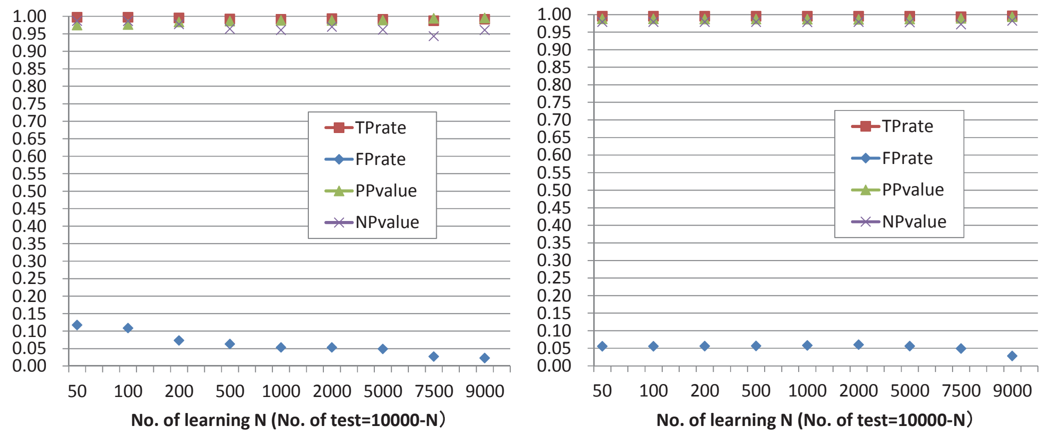
(a) Pattern AA (b) Pattern CA (c) Pattern CC

Figure 5. Prediction of damage rate by SVR (γ=0.02, X: observed, Y: predicted).



(a) Pattern AA (b) Pattern CA (c) Pattern CC

Figure 6. Prediction of damage rate by SVR (γ=0.5, X: observed, Y: predicted).



(a) SVM (b) Damage function without uncertainty

Figure 7. Evaluation indices for shutoff judgement by SVM and damage function (r^*=0.15).

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP17H0206804. Basic data for Super dense Real-time Monitoring of Earthquakes (SUPREME) were provided by Tokyo Gas Co,. Ltd. The authors gratefully acknowledge their assistance.

References

Ministry of Economy, Trade and Industry, Subcommittee for Safety of City Gas, Working Group for Disaster Countermeasures, *A Report on Disaster Countermeasures in City Gas Supply Systems in Consideration of the Great East Japan Earthquake Disaster*, 2012. (in Japanese)

Nojima, N. and Kato, H., Analysis of City Gas Supply Emergency Control Criteria Modeled As a k-out-of-n Shutoff System, *Journal of Japan Society of Civil Engineers, Ser. AI, Structural Mechanics & Earthquake Engineering*, Vol.71, No.4, pp.I_1-I_12, 2015. (in Japanese)

Moriyama, T. and Nojima, N., Performance Evaluation of Emergency Shutoff System of City Gas Supply System in Earthquake Disaster, *Journal of Japan Society of Civil Engineers, Ser. AI, Structural Mechanics & Earthquake Engineering*, Vol.73, No.4, pp.I_187-I_196, 2017. (in Japanese)

Nojima, N. and Moriyama, T., Application of Machine Learning Techniques to Earthquake Damage Estimation and Emergency Shutoff in City Gas Supply System, *Journal of Japan Society of Civil Engineers, Ser. AI, Structural Mechanics & Earthquake Engineering*, Vol.73, No.4, pp.I_197-I_207, 2017. (in Japanese)

Motoda, H., Tsumoto, S., Yamaguchi, T. and Numao M., *Fundamentals of Data Mining*, Ohmsha, Ltd., pp.201-208, 2008. (in Japanese)

Inomata, W., Norito, Y., Ishida, E., Tsukamoto, H. and Yamazaki, F., Summary of Supply Facilities Damage of Tokyo Gas Caused by the Great East Japan Earthquake and Verification of the Accuracy of Pipeline Damage Estimation, *Journal of Japan Association of Earthquake Engineering*, Vol.13, No.2, pp.2_37-2_44, 2013. (in Japanese)

Takeuchi, I. and Toriyama, M., *Support Vector Machine*, Kodansha, 2015. (in Japanese)

Ang, A. H-S. and Tang, W. H., *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*, 2nd Ed., John Wiley & Sons, Inc., 2007.

Rubinstein R. Y., *Simulation and the Monte Carlo Method*, 2nd Ed., John Wiley and Sons, Inc., 1981.

Shimizu, Y., Ishida, E., Isoyama, R., Yamazaki, F., Koganemaru, K. and Nakayama, W., Development of Real-time Earthquake Disaster Mitigation System for City Gas Network and Utilization of Regional Geological Information, *Journal of Structural Mechanics & Earthquake Engineering*, Japan Society of Civil Engineers, No.738, I-64, pp.283-296, 2003. (in Japanese)