Improving Classification Accuracy for Single-loop Reliability-based Design Optimization

I-Tung Yang, and Willy Husada

***Abstract*—Reliability-based design optimization (RBDO) is to find the optimal structural design with minimum cost subject to the required level of reliability. Since the target failure probability is usually small, surrogate models have been created in the single-loop RBDO method to replace the time-consuming reliability analysis. The present study includes three parts: perform Monte Carlo sampling in the solution space, build a surrogate model based on the concept of data mining, and use the surrogate model to perform the reliability analysis. In this empirical study, we adopt three commonly used methods, Classification and Regression Tree (CART), Artificial Neural Network (ANN) and Support Vector Machine (SVM). The goal is to examine the performance of the methods in estimating the reliability states (failed or safe). A ten-bar truss project is adopted as the case study where the design options are selected from a commercial list. We also investigate the effect of the input types (discrete or continuous) on classification accuracy. The findings of the present study can assist in choosing a suitable surrogate model so the RBDO analysis can be performed with better accuracy and higher efficiency.**

***Index Terms*—Design optimization, reliability, data mining, simulation, classification**

I. Introduction

RELIABILITY-BASED design optimization (RBDO) aims to find the optimal design with minimum structure cost or weight subjected to specified reliability constraint such as maximum failure probability limit. RBDO includes two processes: design optimization and reliability analysis. In practical projects, RBDO involves highly non-linear limit state functions, non-normally distributed random variables and discrete design variables. All of these practical issues create significant challenges for RBDO [5]. Moreover, civil structures are usually designed to have a relatively small failure probability. Thus, it demands a long period of computation time to estimate the failure probability [17].

The RBDO analysis may be performed in various ways. The conventional approach for RBDO is the double-loop method. The double-loop method requires a full reliability analysis at every step of the design optimization process. It is too computationally expensive to be applied in practical applications [16]. In contrast, the single-loop method creates a surrogate model or an approximation function to replace the time-consuming reliability analysis process [12]. Despite the improved efficiency, the single-loop method may be inaccurate in estimating the structure failure probability because the surrogate model and approximation function may associate with certain errors [18].

This study enhances the single-loop RBDO method by providing a better surrogate model based on data mining approaches. The judgement criterion is being able to predict whether a structural design will fail with higher accuracy. This need leads to a binary classification problem where data mining is employed to find the hidden rules and to create a surrogate model. The surrogate model will replace the time-consuming reliability analysis in RBDO to achieve efficiency.

A benchmark ten-bar plane truss is used to examine and compare popular data mining methods in creating the surrogate model. The present study includes the following steps. First, Monte Carlo Simulation (MCS) is carried out to generate a data set where design variables (bar sizes) are input and binary reliability states (failed or safe) are output. Second, popular data mining methods are used to estimate the binary results given the design variables. Third, the comparisons among the data mining methods are made based on both accuracy and efficiency. More insights are given to whether the transformation of input data can improve the performance of the surrogate model.

A preliminary experiment is conducted to select the data mining methods for further examinations. We consider seven state-of-the-art methods: Classification and Regression Tree (CART), Artificial Neural Network (ANN), Support Vector Machine (SVM), CHi-Squared Automatic Interaction Detection (CHAID), Bayesian Network (BAYES), Quick, Unbiased and Efficient Statistical Tree (QUEST), and Logistic Regression (LOG). Among the seven methods, CART, ANN, and SVM, are selected because they yield the highest accuracy in predicting the failure probability. In the following section, we introduce MCS, along with the three data mining methods.

II. Methods

A. Monte Carlo Simulation

The Monte Carlo Simulation is a robust, but the most

Manuscript received December 1, 2016.

I-Tung Yang is with National Taiwan University of Science and Technology, TAIWAN and Taiwan Construction Research Institute, TAIWAN (phone: +886-2-27376569; fax: +886-2-27376606; e-mail: ityang@mail.ntust.edu.tw).

Willy Husada was with National Taiwan University of Science and Technology, TAIWAN (e-mail: willyhusada_srw@hotmail.com). He was also a graduate student at Institut Teknologi Sepuluh Nopember (ITS) Surabaya, INDONESIA.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

computationally expensive method to perform probabilistic reliability analysis of a component or a system [11]. According to the Law of Large Number, the average of the results obtained from a large number of trials in a simulation should be close to the expected value and tend to get closer as more trials are performed to the simulation. Using MCS to estimate the failure probability of a structural design is as follows

$$P_f = E(Z(\mathbf{D},\mathbf{X}) < 0) \approx \frac{1}{N}\sum_{i=1}^{N} I[Z^i < 0] \quad (1)$$

where P_f = failure probability;
E = expected value;
$Z(\mathbf{D},\mathbf{X})$ = limit state function with design variables **D** and random variables **X**;
N = number of samples;
$I[.]$ = indicator function;
Z^i = the i-th sample of Z.

The limit state function described in (1) is the condition where a structure exceeds a specific limit as the structure is unable to perform as required. The limit state function indicates the margin of safety between the resistance and the load of structures.

$$Z = R - L \quad (2)$$

where R is the resistance and L is the load. The structure is considered failed when Z is negative, i.e., when the load exceeds the resistance of the structure.

As the sample size increases, simulation accuracy also increases. When the sample size is large, MCS is quite accurate in estimating the failure probability yet very computationally expensive as the target failure probability is usually very small. The main advantage of MCS is that it is flexible enough in modeling non-linear limit state functions and non-Gaussian distributions.

B. Classification and Regression Tree

The classification and regression tree (CART) is a decision tree method by constructing a classification tree or regression tree according to its output variable data type, which can be either categorical or numerical [2]. The construction of the surrogate model using CART is extremely fast because it can provide a clear indication of which inputs are more important for making the prediction easier. This is done by constructing a decision tree.

In the beginning, the decision tree is started with a single parent node and then this parent node will split into two partition nodes. To split a parent node, we need to select a tree-split that gives the smallest impurity among possible splits. Impurity in the CART method is the measurement of how many records in a data set would be incorrectly classified associated with the chosen tree-split. Commonly used impurity measurements are Entropy and Gini Index.

C. Artificial Neural Network

Artificial Neural Network (ANN) is a computational method that is based on the neuron cell structure of the biological nervous system. Given a training set of data, the neural network can learn the pattern of the data set with a learning algorithm. Through propagation, the neural network forms a mapping between inputs and desired outputs from the training set by altering weighted connections within the network [1]. An ANN has one input layer, one output layer and one or more hidden layers with a certain number of neurons. Each neuron is associated with a weight and bias. The algorithmic parameters are the numbers of hidden layers and neurons in each layer.

This study adopts a feedforward backpropagation neural network that has input-to-unit and unit-to-unit connection modified by a weight. Each unit has an extra input that is assumed to have a constant value of one. The weight that modifies this extra input is called the bias. In the feedforward phase, all of the information from the input layer is fed to the network in forward direction from the first hidden layer to the output layer. This phase will activate the activation functions in the output layer. In the backpropagation phase, the activated output layer will propagate backward the error or difference between the predicted output and the actual output through the network. Then, the weight in every connection is adjusted by the error proportions that were propagated backward, thus improving the model [9].

D. Support Vector Machine

Support vector machine (SVM) is a supervised learning algorithm that builds a hyperplane of classification based on the training samples. The training data and the corresponding labels are expressed as follows:

$$D = \{(X_i, y_i) | X_i \in R^p, y_i \in \{-1,1\}\} \;;\; i = 1, \ldots, N \quad (3)$$

where D is the set of training data; X is the p-dimensional design vector; and y is the corresponding labels (safe or failure).

To perform classification, two boundary functions are defined to separate data points

$$w.X - b = 1$$
$$w.X - b = -1 \quad (4)$$

where w is the normal vector of the hyperplane of the classification and b is the distance between the origin and the hyperplane.

The goal of SVM is to maximize the margin between two support vectors [15]. This is equivalent to the minimization of w because the width of separation between two support vectors is $2/\|w\|$. The selection of parameters w and b can be formulated as a non-linear optimization problem, which has been discussed in [4].

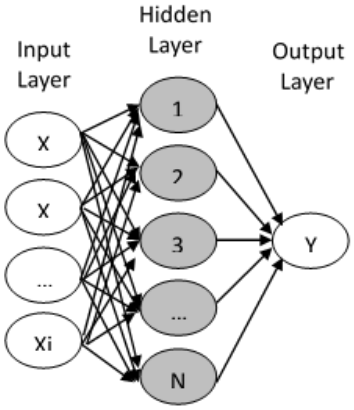

Fig. 1. Illustration of Artificial Neural Network.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

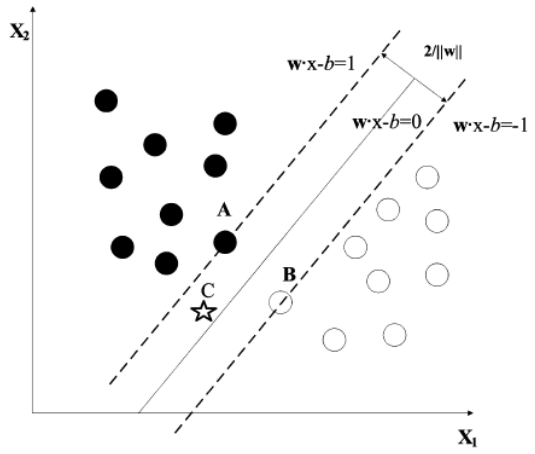

Fig. 2. Classification Margin of Support Vector Machine.

When the data points are not linearly separable, the original space should be transformed into a higher-dimensional space. The training process for SVM then requires the selection of proper kernel functions (linear, radial basis, polynomial, or sigmoid function) to do such transformation [7]. Default values of kernel parameters are highly dependent on their kernel type. The SVM model requires the user to specify two algorithmic parameters, the regularization parameter and the kernel parameter. The regularization parameter represents the tradeoff between model complexity and training error. The kernel parameter controls the kernel width used to fit the training data.

E. Cross Validation

Cross-fold validation is adopted to estimate the accuracy of the surrogate models. In the present study, a model is given a data set with known output which served as training data set and a data set with unknown output which served as testing data set. The training data set is used to construct the prediction model while the testing data set is used to estimate the output value based on the prediction model.

The k-fold cross validation has been widely used for validation because it can minimize bias associated with random feature of training data set sampling process [6]. Ten-fold cross validation is usually considered as the optimal k-fold cross validation in terms of computational time [10]. In the ten-fold cross validation method, a fixed number of data samples from a data set are divided into ten folds. Among these ten folds, nine folds will served as training data set to build surrogate model, while the rest one fold will served as testing data set to verify and validate the accuracy of surrogate model. To reduce variability, every round from ten-fold cross validation process is performed using different partitions. The accuracy of the surrogate model can be expressed as the average accuracy acquired by the ten rounds of validation process.

III. Procedure

As the focus is placed upon CART, ANN, and SVM, the procedure of the present study includes the following steps:
1. Use MCS to prepare a data set: design options as the input while the corresponding probability of failure as the output.

TABLE I
RANDOM VARIABLES

Variable	Distribution	Mean	Dispersion
P_1, P_2	Extreme Value Type I	150 kN	10% c.o.v.
$A_{1\sim10}$	Uniform	$\bar{A}$	±4%
F_y	Lognormal	517.28 N/mm2 (9th Bar), 172.43 N/mm2 (Others)	8% c.o.v.

TABLE II
ERROR

Error	Distribution	Mean	Dispersion
eP1, eP2	Uniform	0	±10%
eA1~10	Uniform	0	±3%
eF	Uniform	0	±20%
e_σ	Uniform	0	±5%

2. Divide the data set into ten folds and separate the training data set from testing data set.
3. Train the surrogate models based on the three methods.
4. Use the surrogate models to perform classification for the testing data set.
5. Compare the classification results with the actual output.
6. Repeat Steps 2 through 5 for the ten folds.
7. Report the classification statistics for each surrogate models.

IV. Benchmark Case

The benchmark case used in this study is a ten-bar truss structure. The shape, geometry and loading of the ten-bar truss structure are shown in [8]. The ten-bar truss is pin-jointed and subjected to two external loads, P_1 and P_2. Every bar is made of hollow carbon steel pipes and may have different sizes. The carbon steel pipes are selected from the local industry standard [3]. There are 36 options in the size list. The selection of bars represents a discrete set with three features: pipe outside diameter (D), wall thickness (t) and cross-sectional area (A).

The steel pipe also has the following mechanical properties: modulus of elasticity (E) of 200,000 N/mm^2 and density (ρ) of 2.768 x 10^{-6} kg/mm^3. As there are 36 discrete options that can be selected for the 10 bars, there are 36^{10} discrete combinations, i.e., roughly 3.65 x 10^{15} options. This amount of possible options is considered huge even for a relatively small RBDO problem. This demonstrates the computational burden for the RBDO problem.

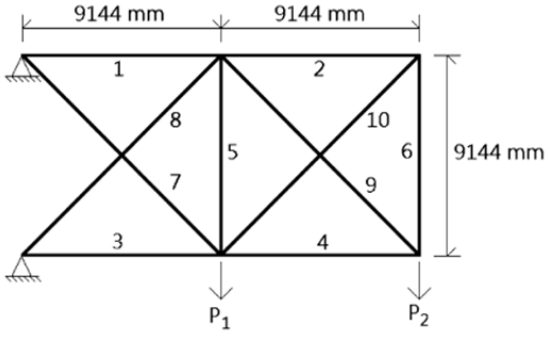

Fig. 3. Ten-Bar Truss Structure Geometry.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

There is variability and error of the random variables due to the randomness nature and manufacturing variation in real situation as follows:

- Randomness of external loads: P_1 and P_2
- Manufacturing variation of cross section area: $A_{1\sim10}$
- Randomness of steel yield stress: F_y
- Estimation error of external loads: e_{P1} and e_{P2}
- Estimation error of cross section areas: $e_{A1\sim10}$
- Estimation error of steel yield stress: e_F
- Error in the mathematical modeling of stress: e_σ

Tables I and II show the probabilistic distributions with mean value and dispersion for modeling the random variables and error. The loads are assumed to follow the extreme value distribution. The cross-sectional areas are uniformly distributed within a certain range caused by manufacturing imperfections. The yield stress is assumed to follow the lognormal distributions.

Two structural failure modes are considered in this case: yield stress and buckling stress. The truss is considered failed when any of the members fails either in terms of yield stress or buckling stress. For yield stress, a failure occurs when the axial stress exceeds the yield strength of the bar. Buckling stress takes place when the bars are subjected to high compressive stress. The bar would lose stiffness and strength capacity.

We consider two criteria for the performance of surrogate models: classification accuracy and computational efficiency. The classification accuracy is calculated by using a confusion matrix. A confusion matrix is a specific table layout that allows visualization of the classification model performance from a typically supervised learning algorithm [13]. Each column of the matrix represents the predicted output class while each row of the matrix represents the actual output class. The idea of the confusion matrix is to separate and count the total number of actual output class correspond to the predicted output class. Computational efficiency is measured by the total time spent in the prediction process, including the time for tuning the algorithmic parameters.

V. COMPARISONS

Grid search is performed to tune the algorithmic parameters. The selected ANN model has 3 hidden layers, each of which has 5 neurons with the training functions being the Scale Conjugate Gradient method. The transfer function is Log-Sigmoid. For SVM, we choose Least Square Support Vector Machine [14] with the radial basis function being the kernel function.

Table III compares the classification accuracy of the three methods: CART, ANN, and SVM after ten-fold cross validation. All the methods can achieve more than 90% accuracy. The performance of ANN is the best while the difference between ANN and CART is moderate. All the three methods are quite consistent because the standard deviation is relatively small. Both ANN and CART are significantly better than SVM. However, ANN and SVM take a lot more computation time (7.96 and 14.31 minutes, respectively) than CART (0.3 minute) in performing the prediction of reliability states.

TABLE III
CLASSIFICATION ACCURACY

Method	Mean	Standard deviation
CART	94.57%	1.44%
ANN	95.83%	1.64%
SVM	92.67%	1.05%

VI. TRANSFORMATION OF INPUT

It has been realized that the data types of input data may influence classification accuracy. In the present study, we investigate the effect of transforming the types of input. This is to suggest the suitable type of input to achieve higher classification accuracy.

Two input types are considered: discrete input represented by bar option (choice within the option list) and continuous input represented by bar cross-sectional area. Bar option is a discrete input because we cannot choose an option outside the list. In contrast, bar cross-sectional area is a continuous input.

For CART, the average classification accuracy is 91.6% for the discrete input (bar option) and 93.9% for the continuous input (bar area). A hypothesis test is conducted to determine the significance of the difference between two cases. After conducting the t-test, we get the p-value of 0.011. Since the p-value is smaller than 5%, we should reject the null hypothesis. Thus, the continuous input (bar area) is significantly better than the discrete input (bar option).

For ANN, the average classification accuracy is 90.7% for the discrete input (bar option) and 93.7% for the continuous input (bar area). The confidence interval of the discrete input type is completely above that of the continuous input type while both confidence intervals are not overlapped. Therefore, the continuous input (bar area) is significantly better than the discrete input (bar option).

For SVM, the average classification accuracy is 86.7% for the discrete input (bar option) and 83.8% for the continuous input (bar area). The confidence intervals between two input types are overlapped and the mean value of one input type is within the confidence interval of the other. As the result, the discrete input (bar option) is not significantly better than the continuous input (bar area).

VII. CONCLUSION

In the present study, three data mining methods (CART, ANN, and SVM) are adopted to be the surrogate models in the single-loop RBDO analysis. The concept is using the surrogate models to replace time-consuming Monte Carlo Simulation in the classification of reliability levels (fail or safe) for structural designs. A ten-bar truss structure is used to compare the performance of the three methods in terms of classification accuracy and computational efficiency. The empirical results show that ANN is superior to SVM and the difference between ANN and CART is moderate. With respect to computational efficiency, CART takes much less time than both ANN and SVM. We also examine the effect of transforming the input data from discrete options to continuous values. The transformation is shown to be advantageous for CART and ANN.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

REFERENCES

[1] Adeli, H., & Hung, S. (1995). Machine Learning: Neural Networks. *Genetic Algorithms And Fuzzy Systems, John Wiley & Sons, Inc*.
[2] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.
[3] CNS 4435, G3102 (1995). Carbon steel pipes for general structures. *Chinese National Bureau of Standards*, Ministry of Economic, Taiwan.
[4] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.
[5] Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms* (Vol. 16): John Wiley & Sons.
[6] Geisser, S. (1993). *Predictive inference* (Vol. 55): CRC Press.
[7] Genton, M. G. (2001). Classes of kernels for machine learning: a statistics perspective. *The Journal of Machine Learning Research*, *2*, 299-312.
[8] Haftka, R. T., & Gürdal, Z. (1992). *Elements of structural optimization* (Vol. 11): Springer Science & Business Media.
[9] Hagan, M. T., Demuth, H. B., & Beale, M. H. (2002). *Neural network design*: University of Colorado Bookstore.
[10] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th international joint conference on Artificial intelligence* - Volume 2, 1137-1143.
[11] Rubinstein, R. Y., & Kroese, D. P. (2011). *Simulation and the Monte Carlo method* (Vol. 707): John Wiley & Sons.
[12] Steenackers, G., Versluys, R., Runacres, M., & Guillaume, P. (2011). Reliability-based design optimization of computation-intensive models making use of response surface models. *Quality and Reliability Engineering International*, *27*(4), 555-568.
[13] Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote sensing of Environment*, *62*(1), 77-89.
[14] Suykens, J. A., Van Gestel, T., De Moor, B., & Vandewalle, J. (2002). Basic Methods of Least Squares Support Vector Machines, *Least Squares Support Vector Machines*: World Scientific.
[15] Vapnik, V. (1995). *The nature of statistical learning theory*: Springer Science & Business Media.
[16] Yang, R., & Gu, L. (2004). Experience with approximate reliability-based optimization methods. *Structural and Multidisciplinary Optimization*, *26*(1-2), 152-159.
[17] Yang, I.T. and Hsieh, Y.H. (2013). "Reliability-based design optimization with cooperation between support vector machine and particle swarm optimization." *Engineering with Computers*, 29(2), 151-163.
[18] Yang, I.T., Husada, W., and Adi, T.J.W. (2016). Data Mining Methods in Structural Reliability Estimation, *Proceedings of the 16th International Conference on Computing in Civil and Building Engineering*.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)