

## Research Article

# Estimating Compressive Strength of High Performance Concrete with Gaussian Process Regression Model

Nhat-Duc Hoang,<sup>1</sup> Anh-Duc Pham,<sup>2</sup> Quoc-Lam Nguyen,<sup>3</sup> and Quang-Nhat Pham<sup>3</sup>

<sup>1</sup>*Institute of Research and Development, Faculty of Civil Engineering, Duy Tan University, P809-K7/25 Quang Trung, Danang 550000, Vietnam*

<sup>2</sup>*Faculty of Project Management, The University of Danang, University of Science and Technology, 54 Nguyen Luong Bang, Danang 550000, Vietnam*

<sup>3</sup>*Faculty of Civil Engineering, Duy Tan University, P809-K7/25 Quang Trung, Danang, Vietnam*

Correspondence should be addressed to Nhat-Duc Hoang; [hoangnhatduc@dtu.edu.vn](mailto:hoangnhatduc@dtu.edu.vn)

Received 3 June 2016; Revised 22 September 2016; Accepted 26 September 2016

Academic Editor: Ghassan Chehab

Copyright © 2016 Nhat-Duc Hoang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This research carries out a comparative study to investigate a machine learning solution that employs the Gaussian Process Regression (GPR) for modeling compressive strength of high-performance concrete (HPC). This machine learning approach is utilized to establish the nonlinear functional mapping between the compressive strength and HPC ingredients. To train and verify the aforementioned prediction model, a data set containing 239 HPC experimental tests, recorded from an overpass construction project in Danang City (Vietnam), has been collected for this study. Based on experimental outcomes, prediction results of the GPR model are superior to those of the Least Squares Support Vector Machine and the Artificial Neural Network. Furthermore, GPR model is strongly recommended for estimating HPC strength because this method demonstrates good learning performance and can inherently express prediction outputs coupled with prediction intervals.

## 1. Introduction

In construction industry, high-performance concrete (HPC) has been widely used in high-rise building/infrastructure projects for its superior strength, durability, and workability which exceed those of normal concrete [1, 2]. Generally, special ingredients are employed to make these specially designed concretes to satisfy a combination of performance requirements. Furthermore, the compressive strength is often considered as the most important property of HPC; other concrete properties such as elastic modulus, water tightness, and impermeability appear to have direct relationships with compressive strength [3]. Hence, the compressive strength is commonly utilized as the main criterion in defining the required quality of concrete [4].

The compressive strength is determined through a standard uniaxial compression test. If the test result does not meet the designed strength, remediation actions must be undertaken. Furthermore, corrective actions for underground concrete structures, such as concrete piles or foundations, can

be very costly. As a result, an accurate estimation of the compressive strength before the placement is a practical need of construction engineers.

As the relationships between concrete components and compressive strength are complex and highly nonlinear, mathematical modeling of HPC is very challenging and oftentimes inaccurate [5]. Consequently, traditional statistical methods are inadequate for modeling of HPC compressive strength. Herein, the main goal is to construct a system that can learn from a data set of different HPC mixes and can predict accurately the compressive strength based on the pattern of concrete components.

Accordingly, this research extends the body of knowledge by evaluating the capability of the Gaussian Process Regression (GPR) [6] for modeling compressive strength of HPC. GPR is an efficient and reliable learning approach for modeling nonlinear and complex functional mappings [7, 8]; therefore, an assessment of this model performance on HPC strength prediction is particularly useful for practicing

engineers. Additionally, this study can also be viewed as a comparative work since the performance of the GPR is benchmarked against those of other powerful nonlinear modeling methods including the Artificial Neural Network (ANN) and the Least Squares Support Vector Machine (LSSVM) [9]. Moreover, a data set of 239 HPC experimental tests recorded during the construction phase of the Nga Ba Hue highway overpass project in Danang City (Vietnam) has been collected for this research. The subsequent parts of the paper are organized as follows: Section 2 reviews pertinent research works in the literature. In Section 3, the research method is presented, followed by the experimental results. The conclusion of this study is stated in the final section.

## 2. Related Works

Due to the importance of the research topic, HPC compressive strength modeling has been a very active research area and various artificial intelligence (AI) techniques have been applied to tackle the problem of interest. Based on previous studies, AI techniques have proved its superior capability over traditional modeling methods. ANN is the most common modeling method [10, 11]. Yeh and Lien [12] applied the Genetic Operation Trees to establish empirical formulas which are more accurate than the nonlinear regression but less accurate than ANN models. Chou et al. [13] compare different data-mining techniques to identify the fittest method.

Sophisticated AI based systems have also been developed to fit particular HPC data sets. Słoiński [14] combined the ANN and the Bayesian evidence framework which helps to construct the ANN structure. Genetic Weighted Pyramid Operation Tree was developed by combining four individual Genetic Operation Trees with adaptive weights [15]; the method appears to be better than the individual tree. The  $K$ -nearest neighbor based regression was integrated with Differential Evolution by Ahmadi-Nedushan [3]; this work demonstrated that despite being simple, in certain learning circumstances, the instance-based regression method can outperform the ANN.

Cheng et al. [16] proposed a fuzzy Support Vector Machine (SVM) method which can be superior over the original SVM and be comparable to the ANN. Pham et al. [5] employ LSSVM and Firefly Algorithm to construct a hybrid model for HPC strength estimation. The capability of Genetic Programming based prediction models, which fuses the genetic algorithm and the symbolic programming, was investigated by Chen and Wang [17], Mousavi et al. [18], and Castelli et al. [19]. Extensive researches on ensemble learning of Erdal et al. [20] and Chou et al. [21] found that this form of learning may boost the predictive capability of individual AI models. As can be seen in the literature, the capability of GPR has rarely been investigated for the task of HPC strength modeling. Therefore, our study is an attempt to fill this gap in the literature.

## 3. Research Method

**3.1. Data Set of HPC Experiments.** This research employs a data set consisting of 239 testing results of HPC concrete

specimens. All the experimental tests were performed with 15 cm cylindrical specimens of HPC prepared according to the Vietnamese standard (TCVN 3105: 1993), which is relatively similar to the American standard ASTM C39. The amounts of cement ( $\text{Kg/m}^3$ ), sand ( $\text{Kg/m}^3$ ), small coarse aggregate ( $\text{Kg/m}^3$ ), medium coarse aggregate ( $\text{Kg/m}^3$ ), water ( $\text{liter/m}^3$ ), and superplasticizer ( $\text{liter/m}^3$ ) are batch components employed for expressing properties of a concrete sample. It is noted that the concrete age of each sample is measured in day.

Statistical descriptions of HPC test are reported in Table 1. It is worth noticing that the small coarse aggregate has diameter ranging from 5 to 10 mm; the medium coarse aggregate ranges in diameter from 10 and 20 mm. In addition, the ratio of water-to-cement of the concrete mixes in the data set ranges from 0.27 to 0.46. Such low water-to-cement ratio necessitates the utilization of superplasticizer to enhance the concrete workability. The minimum and maximum compressive strength of concrete specimens are 23.6 and 85.2 MPa, respectively.

**3.2. Gaussian Process Regression (GPR).** GPR presents a probabilistic, nonparametric supervised learning approach for generalizing nonlinear and complex function mapping hidden in data sets. This approach has recently received huge attention of researchers in various study disciplines [4, 22]. GPR is very efficient to handle nonlinear data due to the use of kernel functions. Furthermore, a significant advantage of GPR is that the model can provide a reliability response to an input data [23].

Given a training set  $D = \{(x_i, y_i) \mid i = 1, \dots, n\}$ , the input data  $X \in R^{D \times n}$  is called the design matrix and  $y \in R^n$  is the vector of desired output. The main assumption of GPR is that the output  $y$  is computed as [6, 24]

$$y = f(x) + \varepsilon, \quad (1)$$

where  $\varepsilon \sim N(0, \sigma_n^2) \in R$  represents a homoscedastic noise for all sample  $x_i$ .

In GPR methodology, the  $n$  observations in the data set of interest  $y = \{y_1, \dots, y_n\}$  are considered as a single point sampled from a multivariate Gaussian distribution. Moreover, it can be assumed that this Gaussian distribution has the mean of zeros. The covariance function  $k(x, x')$  dictates the relation of one observation to another observation. The squared exponential covariance function is often selected in GPR for the task of function approximation [22, 24]:

$$k(x, x') = \sigma_f^2 \times \exp\left(-\frac{(x - x')^2}{2l^2}\right) + \sigma_n^2 \delta(x, x'), \quad (2)$$

where the maximum allowable covariance is defined as  $\sigma_f^2$ . It is noted that  $k(x, x')$  reaches this maximum allowable covariance only when  $x$  is very closed to  $x'$  and therefore  $f(x)$  is almost perfectly correlated with  $f(x')$ . Meanwhile,  $l$  denotes the length parameter of the kernel function. In addition,  $\delta(x, x')$  represents a Kronecker delta function;  $\delta_{ij} = 1$  if  $i = j$  and  $\delta_{ij} = 0$  if  $i \neq j$ .

TABLE I: Concrete components and statistical descriptions.

HPC input factor (IF)	Notation	Min	Mean	Std. dev.	Max
Cement (Kg/m <sup>3</sup> )	IF <sub>1</sub>	350.0	447.4	25.0	498.0
Fine aggregate (Kg/m <sup>3</sup> )	IF <sub>2</sub>	666.0	728.9	37.8	879.0
Small coarse aggregate (Kg/m <sup>3</sup> )	IF <sub>3</sub>	0.0	347.2	55.9	424.0
Medium coarse aggregate (Kg/m <sup>3</sup> )	IF <sub>4</sub>	626.0	721.3	61.4	1060.0
Water (liter/m <sup>3</sup> )	IF <sub>5</sub>	134.0	178.8	20.4	207.0
Superplasticizer (liter/m <sup>3</sup> )	IF <sub>6</sub>	3.5	5.1	0.6	7.0
Concrete age (day)	IF <sub>7</sub>	3.0	15.1	10.9	28.0
Compressive strength (MPa/m <sup>3</sup> )	CS	23.6	42.5	13.5	85.2

Given the training data set, the ultimate goal of the learning process is to predict the output value  $y_*$  of a new queried input pattern. To achieve such goal, it is necessary to establish three covariance matrices as follows:

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{bmatrix}, \quad (3)$$

$$K_* = [k(x_*, x_1) \quad k(x_*, x_2) \quad \cdots \quad k(x_*, x_n)],$$

$$K_{**} = k(x_*, x_*).$$

Due to the assumption that the data is sampled from a multivariate Gaussian distribution, we have the following expression:

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix}\right). \quad (4)$$

Since  $y_* | y$  is proved to be generated from a multivariate Gaussian distribution with the mean of  $K_* K^{-1} y$  and the variance of  $K_{**} - K_* K^{-1} K_*^T$ , the estimated mean and variance of the predicted output  $y_*$  are given as follows:

$$E(y_*) = K_* K^{-1} y, \quad (5)$$

$$\text{var}(y_*) = K_{**} - K_* K^{-1} K_*^T.$$

When the hyperparameters of the kernel function are specified, the model parameters, including  $k$  and  $\sigma_n$ , can be determined by Bayesian inference. This task boils down to maximizing a log-posterior function [24]. After the training phase terminates, the GPR model can be employed for predicting unknown input data sample.

## 4. Experimental Result and Comparison

*4.1. Estimating HPC Compressive Strength with Gaussian Process Regression.* In this experiment, the data set of HPC testing samples has been divided into two sets: the training set (90%) used for model construction and the testing set

(10%) employed for model testing. Prior to the training process, it is necessary to specify the hyperparameters of GPR model. These hyperparameters include the initial value for the standard deviation of the noise  $\sigma_n$ , the maximum allowable covariance  $\sigma_f$ , and the length parameter  $l$  of the kernel function. Based on recommendations in Mathworks [25], the initial value for the standard deviation of the noise is set as follows:

$$\sigma_n = \frac{\text{std}(Y_{\text{Training}})}{\sqrt{2}}, \quad (6)$$

where  $\text{std}(Y_{\text{Training}})$  denotes the standard deviation of the desired output in the training data set.

To select the maximum allowable covariance  $\sigma_f$  and the length parameter  $l$  of the kernel function, we have carried out a model selection process. These two parameters are allowed to be varied within the following parameter set (ParSet):

$$\text{ParSet} = \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100, 500, 1000\}. \quad (7)$$

The training set is further separated into two subsets: subset 1 (90%) and subset 2 (10%); and a grid search procedure described in Algorithm 1 is performed to identify the most suitable set of  $\sigma_f$  and  $l$ . Subset 1 herein plays the role of a training set. The set of parameters accompanied with the most desirable prediction performance of the model when it is used to predict data in subset 2 is selected. To quantify the model prediction performance, the Root Mean Squared Error (RMSE) has been used. Based on the experiment, the values of  $\sigma_f$  and  $l$  are selected to be 0.1 and 0.05.

With the three aforementioned hyperparameters, the training process of GPR can be executed. Accordingly the constructed model is utilized to predict the data instances in the testing set. The prediction outcome of the GPR testing phase is illustrated in Figure 1. To express the model accuracy, besides RMSE, the Mean Absolute Percentage Error (MAPE) and the Coefficient of Determination ( $R^2$ ) have also been employed. The experimental result is reported as follows: RMSE = 4.92; MAPE = 7.29%; and  $R^2 = 0.90$ . These outcomes demonstrate that GPR has successfully captured the non-linear function that determines the mapping between input factors of ingredients and the output of HPC compressive strength. As shown in Figure 1, the GPR model's outcomes

```

Establishing Subset1, Subset2
Establishing ParSet
PM = {∅} // Performance matrix
For i = 1 : Np // Np = number of hyper parameters in ParameterSet
    σf = ParSet(i)
    For j = 1 : Np
        l = ParSet(j)
        Train GPR model with Subset1
        GPR model prediction with Subset2
        PM(i,j) = RMSE of Subset2
        // RMSE denotes the Root Mean Squared Error
    End For
End For
End For
Finding the best set of {σf, l} based on PM

```

ALGORITHM 1: Grid search procedure for GPR hyperparameter determination.

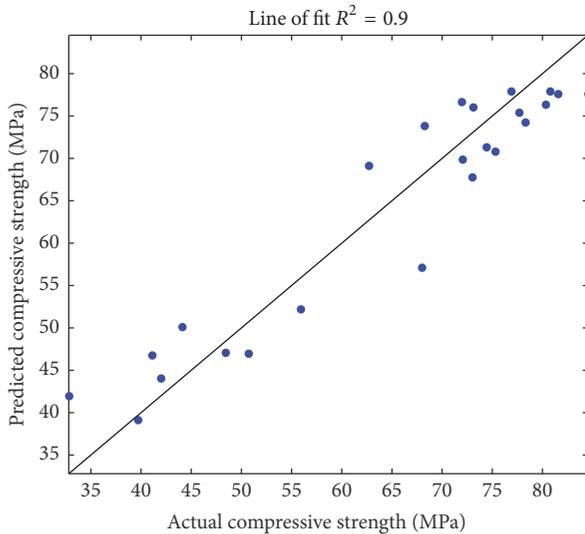


FIGURE 1: GPR prediction result in the testing phase.

have obtained a good fit to a straight line. Since GPR can express uncertainty associated with each predicted output, Figure 2 illustrates the GPR results of prediction interval with 95% level of confidence. It is noted that all actual compressive strengths are within the range of the lower and upper boundaries.

**4.2. Result Comparison.** In this section of the article, to better evaluate the performance of the GPR model, the ANN [26] and LSSVM [9, 27] are employed as benchmark methods. The reasons for selecting these two benchmark models are that ANN is widely accepted as an effective tool for nonlinear function approximation and this algorithm has been successfully employed for predicting concrete strength [10, 13]; LSSVM is also an advanced machine learning method featured by high modeling accuracy [28–31] and it has been recently used for modeling concrete compressive strength [5].

To establish an ANN model, number of neurons in the hidden layer should be determined in advance and

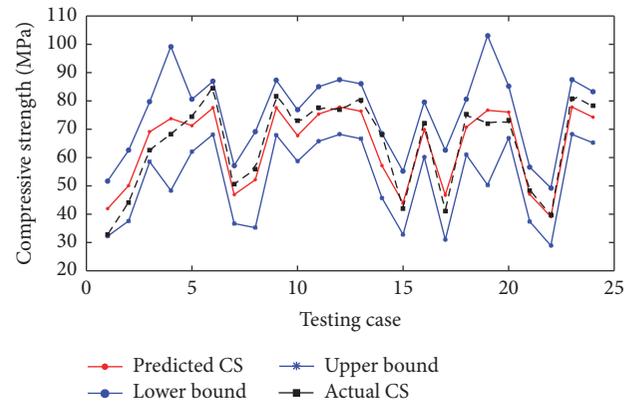


FIGURE 2: GPR prediction result with prediction interval.

this parameter significantly influences the ANN prediction capability. In order to specify an appropriate model structure for ANN, the hidden layer starts with seven neurons (which is equal to the number of input factors) and then gradually increased to the maximum value of 30 neurons. The log-sigmoid function is commonly employed as the activation function and the Levenberg-Marquardt algorithm is utilized to train the ANN [26, 32, 33]. Herein, the training set is also separated into two subsets: subset 1 (90% or equivalently 193 data samples) and subset 2 (10% or equivalently 22 data samples); the number of neurons in the hidden layer results in the best prediction outcome in the testing phase of ANN being chosen. On the other hand, the hyperparameters of LSSVM (the regularization parameter and the kernel function parameter) are automatically tuned by the Firefly Algorithm as described in [5].

As mentioned earlier, the data set is randomly divided into 2 sets: training set (90%) and testing set (10%). Accordingly, the training and testing sets consist of 215 and 24 cases, respectively. Nevertheless, to avoid the randomness in testing sample selection and to compare the performances of models reliably, a 10-fold cross validation process is performed [34, 35]. Accordingly, the whole data set is randomly divided

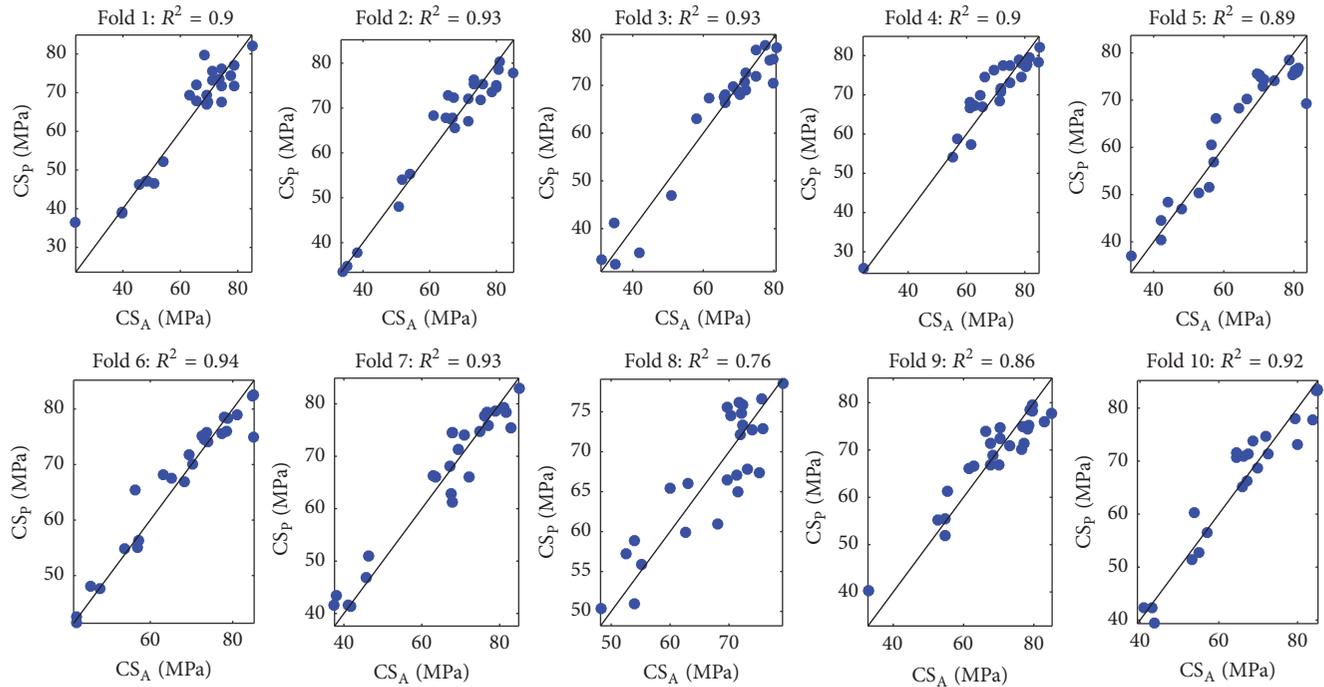


FIGURE 3: GPR prediction result obtained from the 10-fold cross validation.

TABLE 2: Prediction result comparison.

Model	Criteria	Training phase		Testing phase	
		Average result	Standard deviation	Average result	Standard deviation
GPR	RMSE	4.06	1.29	4.04	0.47
	MAPE	5.02	1.86	5.14	0.89
	$R^2$	0.90	0.07	0.90	0.05
LSSVM	RMSE	4.46	0.58	4.63	0.62
	MAPE	5.67	0.92	5.94	0.86
	$R^2$	0.89	0.02	0.87	0.05
ANN	RMSE	5.07	0.87	5.21	1.85
	MAPE	6.56	1.23	6.34	2.32
	$R^2$	0.85	0.14	0.81	0.15

into ten data folds in which each fold in turn serves as a testing set; and the performance of the three models (GPR, ANN, and LSSVM) can be quantified by averaging results of the ten folds. Because all of the subsamples are mutually exclusive, this cross validation process can reliably assess the GPR model and the other two benchmarking methods.

Prediction results of the GPR model and the two benchmark models obtained from the ten-fold cross validation process are reported in Table 2. It can be observed that the GPR has achieved the best prediction result in all of the performance evaluation criteria, followed by LSSVM and ANN. Particularly, in terms of RMSE, GPR achieves a 12.74% improvement compared with LSSVM and a 22.46% improvement compared with ANN.

The outcomes of the GPR, LSSVM, and ANN models attained from the cross validation process are graphically reported in Figures 3, 4, and 5, respectively. In these figures,

the horizontal axis measures the actual compressive strength noted as  $CS_A$ ; meanwhile, the vertical axis measures the predicted compressive strength ( $CS_p$ ) obtained from the prediction phases of each model. The performances of models in these figures can be appraised graphically with the line of best fit and quantitatively with  $R^2$  values; it is noted that a data point locating closely to the line of best fit indicates an accurate prediction outcome. Based on experimental results that are visually displayed in Figures 3, 4, and 5, it can be confirmed that the GPR model is best suited for modeling the data set at hand.

## 5. Conclusion

This research has investigated the capability of the GPR model for the task of HPC compressive strength prediction. To construct and verify the machine learning model, a data set

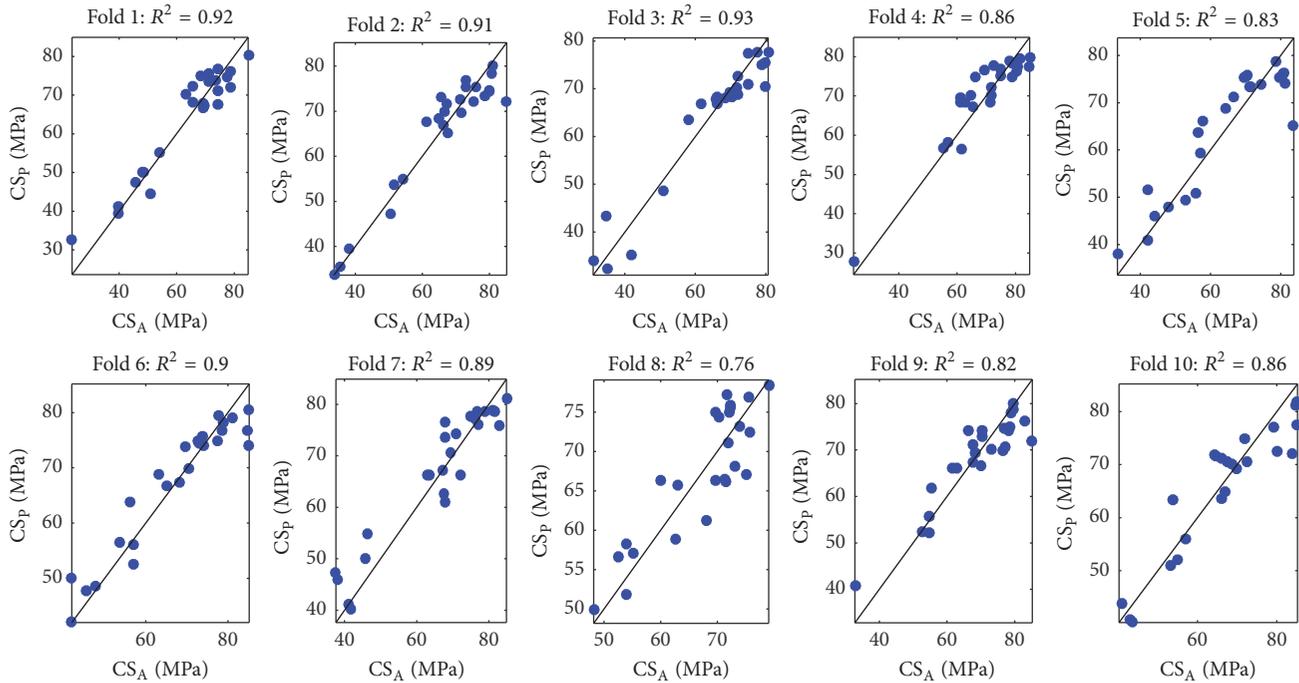


FIGURE 4: LSSVM prediction result obtained from the 10-fold cross validation.

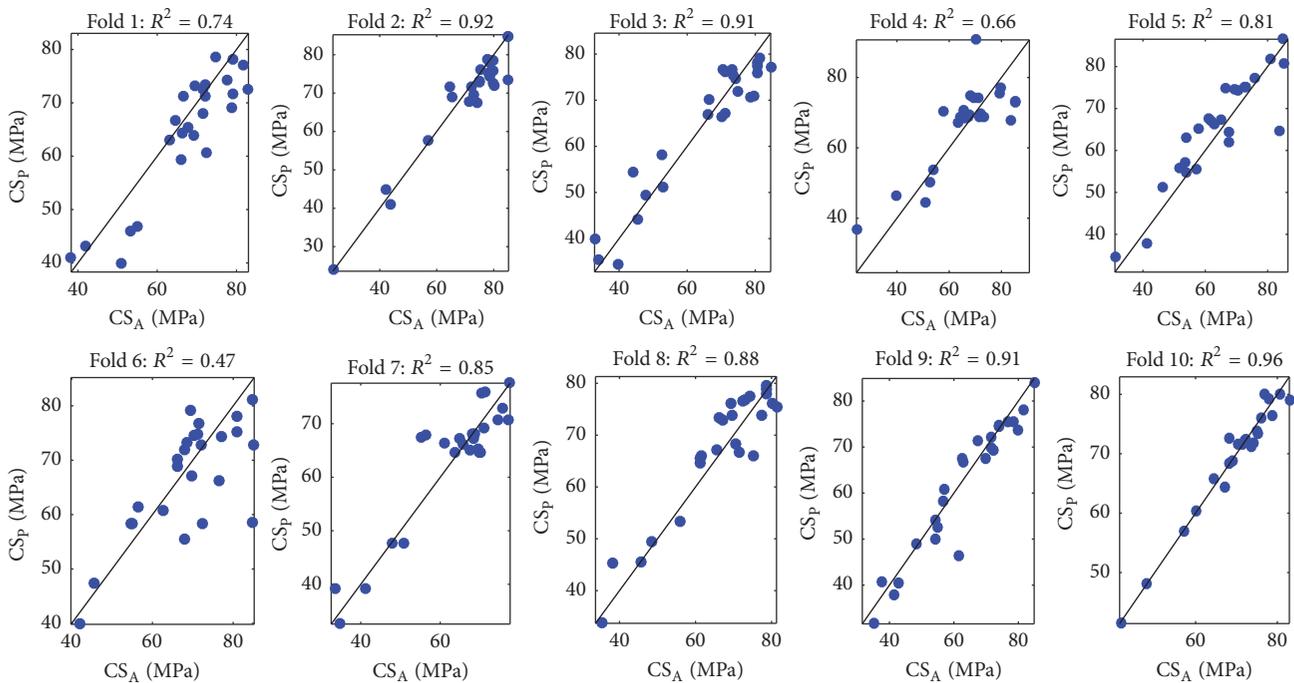


FIGURE 5: ANN prediction result obtained from the 10-fold cross validation.

of actual HPC compressive tests has been collected for this study. Based on experimental results, the GPR model has achieved the most desirable performance with comparatively low prediction errors (RMSE = 4.04, MAPE = 5.15%) and a high coefficient of determination  $R^2 = 0.90$ . These are

very desirable because the modeling HPC strength is widely known to be a highly complex task.

One significant advantage of GPR over other benchmark methods is that the GPR can deliver estimated compressive strength coupled with prediction interval. This property is

also of great usefulness for construction engineers to reliably assess the strength of HPC concrete mixtures. Therefore, the GPR model is recommended as a promising alternative to assist construction engineers in concrete mixture design.

Despite the aforementioned advantages of GPR, one limitation of the study is that the employed approach is a black-box prediction model; hence, this may impose certain hindrance for civil engineers to understand the model structure. In addition, the size of the current data set should be expanded by collecting more testing results of HPC samples to further enhance the generalization of the prediction model.

Therefore, future extensions of this research may include applications of GPR for solving other prediction/modeling tasks in civil engineering, investigation on the effects of novel covariance functions on the GPR model performance, and discovering new techniques to improve the model learning capability. On the other hand, studying the potentiality of other machine learning techniques with transparent model structures such as instance-based learning or regression trees to meliorate the model interpretation is also a worth-investigating research direction.

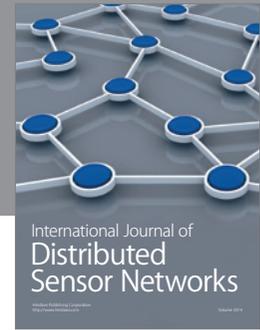
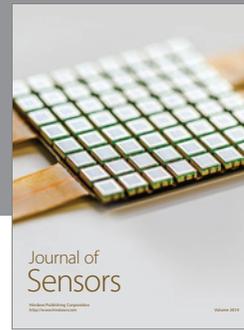
## Competing Interests

The authors (Nhat-Duc Hoang, Anh-Duc Pham, Quoc-Lam Nguyen, and Quang-Nhat Pham) declare that there is no conflict of interests regarding the publication of this article.

## References

- [1] B.-I. Bae, H.-K. Choi, and C.-S. Choi, "Flexural strength evaluation of reinforced concrete members with ultra high performance concrete," *Advances in Materials Science and Engineering*, vol. 2016, Article ID 2815247, 10 pages, 2016.
- [2] O. Gunes, S. Yesilmen, B. Gunes, and F.-J. Ulm, "Use of UHPC in bridge structures: material modeling and design," *Advances in Materials Science and Engineering*, vol. 2012, Article ID 319285, 12 pages, 2012.
- [3] B. Ahmadi-Nedushan, "An optimized instance based learning algorithm for estimation of compressive strength of concrete," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 5, pp. 1073–1081, 2012.
- [4] B. A. Omran, Q. Chen, and R. Jin, "Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete," *Journal of Computing in Civil Engineering*, 2016.
- [5] A. Pham, N. Hoang, and Q. Nguyen, "Predicting compressive strength of high-performance concrete using metaheuristic-optimized least squares support vector regression," *Journal of Computing in Civil Engineering*, vol. 30, no. 3, Article ID 06015002, 2016.
- [6] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning, The MIT Press, Cambridge, Mass, USA, 2006.
- [7] J. Hu and J. Wang, "Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression," *Energy*, vol. 93, part 2, pp. 1456–1466, 2015.
- [8] L. Zhou, J. Chen, and Z. Song, "Recursive gaussian process regression model for adaptive quality monitoring in batch processes," *Mathematical Problems in Engineering*, vol. 2015, Article ID 761280, 9 pages, 2015.
- [9] J. Suykens, J. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle, *Least Square Support Vector Machines*, World Scientific, Singapore, 2002.
- [10] S. Chithra, S. R. R. S. Kumar, K. Chinnaraju, and F. Alfin Ashmita, "A comparative study on the compressive strength prediction models for High Performance Concrete containing nano silica and copper slag using regression analysis and Artificial Neural Networks," *Construction and Building Materials*, vol. 114, pp. 528–535, 2016.
- [11] R. Gupta, M. A. Kewalramani, and A. Goel, "Prediction of concrete strength using neural-expert system," *Journal of Materials in Civil Engineering*, vol. 18, no. 3, pp. 462–466, 2006.
- [12] I.-C. Yeh and L.-C. Lien, "Knowledge discovery of concrete material using Genetic Operation Trees," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5807–5812, 2009.
- [13] J.-S. Chou, C.-K. Chiu, M. Farfoura, and I. Al-Taharwa, "Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques," *Journal of Computing in Civil Engineering*, vol. 25, no. 3, pp. 242–253, 2011.
- [14] M. Słowski, "A comparison of model selection methods for compressive strength prediction of high-performance concrete using neural networks," *Computers and Structures*, vol. 88, no. 21-22, pp. 1248–1253, 2010.
- [15] M.-Y. Cheng, P. M. Firdausi, and D. Prayogo, "High-performance concrete compressive strength prediction using Genetic Weighted Pyramid Operation Tree (GW POT)," *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 104–113, 2014.
- [16] M.-Y. Cheng, J.-S. Chou, A. F. V. Roy, and Y.-W. Wu, "High-performance concrete compressive strength prediction using time-weighted evolutionary fuzzy support vector machines inference model," *Automation in Construction*, vol. 28, pp. 106–115, 2012.
- [17] L. Chen and T.-S. Wang, "Modeling strength of high-performance concrete using an improved grammatical evolution combined with macrogenetic algorithm," *ASCE Journal of Computing in Civil Engineering*, vol. 24, no. 3, pp. 281–288, 2010.
- [18] S. M. Mousavi, P. Aminian, A. H. Gandomi, A. H. Alavi, and H. Bolandi, "A new predictive model for compressive strength of HPC using gene expression programming," *Advances in Engineering Software*, vol. 45, no. 1, pp. 105–114, 2012.
- [19] M. Castelli, L. Vanneschi, and S. Silva, "Prediction of high performance concrete strength using Genetic Programming with geometric semantic genetic operators," *Expert Systems with Applications*, vol. 40, no. 17, pp. 6856–6862, 2013.
- [20] H. I. Erdal, O. Karakurt, and E. Namli, "High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 4, pp. 1246–1254, 2013.
- [21] J.-S. Chou, C.-F. Tsai, A.-D. Pham, and Y.-H. Lu, "Machine learning in concrete strength simulations: multi-nation data analytics," *Construction and Building Materials*, vol. 73, pp. 771–780, 2014.
- [22] M.-Y. Cheng, C.-C. Huang, and A. F. V. Roy, "Predicting project success in construction using an evolutionary gaussian process inference model," *Journal of Civil Engineering and Management*, vol. 19, supplement 1, pp. S202–S211, 2013.
- [23] M. Pal and S. Deswal, "Modelling pile capacity using Gaussian process regression," *Computers and Geotechnics*, vol. 37, no. 7-8, pp. 942–947, 2010.

- [24] M. Ebden, "Gaussian processes: a quick introduction," <https://arxiv.org/abs/1505.02965>.
- [25] Mathworks, *Statistics and Machine Learning Toolbox*, The MathWorks, 2016.
- [26] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox User's Guide*, The MathWorks, 2012.
- [27] K. De Brabanter, P. Karsmakers, F. Ojeda et al., "LS-SVMlab toolbox user's guide version 1.8," Internal Report 10-146, ESAT-SISTA, KU Leuven, Leuven, Belgium, 2010.
- [28] D.-T. Vu and N.-D. Hoang, "Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach," *Structure and Infrastructure Engineering*, vol. 12, no. 9, 2016.
- [29] W. Sun and M. Liu, "Prediction and analysis of the three major industries and residential consumption CO<sub>2</sub> emissions based on least squares support vector machine in China," *Journal of Cleaner Production*, vol. 122, pp. 144–153, 2016.
- [30] M.-Y. Cheng, N.-D. Hoang, and Y.-W. Wu, "Cash flow prediction for construction project using a novel adaptive time-dependent least squares support vector machine inference model," *Journal of Civil Engineering and Management*, vol. 21, no. 6, pp. 679–688, 2015.
- [31] M. Cheng and N. Hoang, "A self-adaptive fuzzy inference model based on least squares SVM for estimating compressive strength of rubberized concrete," *International Journal of Information Technology & Decision Making*, vol. 15, no. 3, pp. 603–619, 2016.
- [32] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.
- [33] T.-H. Tran and N.-D. Hoang, "Predicting colonization growth of algae on mortar surface with artificial neural network," *Journal of Computing in Civil Engineering*, 2016.
- [34] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010.
- [35] P. Zhang, "Model selection via multifold cross validation," *The Annals of Statistics*, vol. 21, no. 1, pp. 299–313, 1993.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

