Metaheuristic Computing and Applications, *Vol. 1, No. 2 (2020) 155*-163
DOI: https://doi.org/10.12989/mca.2020.1.2.155

Compressive strength prediction of high-strength concrete using support vector regression

Hai-Bang Ly and Binh Thai Pham*

University of Transport Technology, Hanoi 100000, Vietnam

(Received March 3, 2020, Revised August 19, 2020, Accepted December 22, 2020)

Abstract. Determination of compressive strength of concrete at early testing age is vital in many civil engineering applications. The strength at 7 or 14 days allows engineers to have confidence in the target strength and make a decision in case of unsuspected situations. In this study, the possibility to estimate the early compressive strength of concrete by a machine learning algorithm, namely the support vector regression (SVR), was investigated. To this aim, a database containing 324 data points was gathered from the available literature and use to develop the ML model. For the assessment of the accuracy, common statistical measurements, such as the Pearson correlation coefficient (R) and root mean square error (RMSE) were used. The results showed that the SVR model could successfully model the early compressive strength of concrete with R=0.92386 and RMSE=5.5089 MPa. The sensitivity analysis on the factors exhibiting a positive or negative effect on the early strength of concrete was conducted. The cement content was shown to have the most influential effect on the early development of concrete compressive strength.

Keywords: early compressive strength; concrete; support vector regression; machine learning

1. Introduction

High-performance concrete or High-strength concrete was developed in the late 1970s is now referred to as HPC (Köksal *et al*. 2008). It is preferable to be defined as a low water/binder concrete which receives an adequate water curing. HPC mixtures are composed of mainly the same materials as conventional concrete mixtures, but the proportions are designed or engineered to provide the strength and durability needed for the structural and environmental requirements of the project (Kim *et al*. 2007). It is designed to be more durable and, if necessary, more robust than conventional concrete (Chan *et al*. 1999). Therefore, the use of HPC is becoming more and more common in construction projects.

One of the most important mechanical properties of HPC is the early compressive strength (Öztaş *et al*. 2006). Accurate determination of this parameter affects significantly to safety and durability of the construction projects. Traditionally, this parameter is often determined in the laboratory by the destructive compression experiments on the samples at the required age (Gran *et al*., 1989). However, this kind of test takes much time for the preparation and maintenance of samples, and the accuracy of experimental results depends significantly on the quality of

*Corresponding author, Ph.D., E-mail: binhpt@utt.edu.vn

http://www.techno-press.org/?journal=mca&subpage=5 ISSN: 2713-5519 (Print), 2733-8053 (Online)

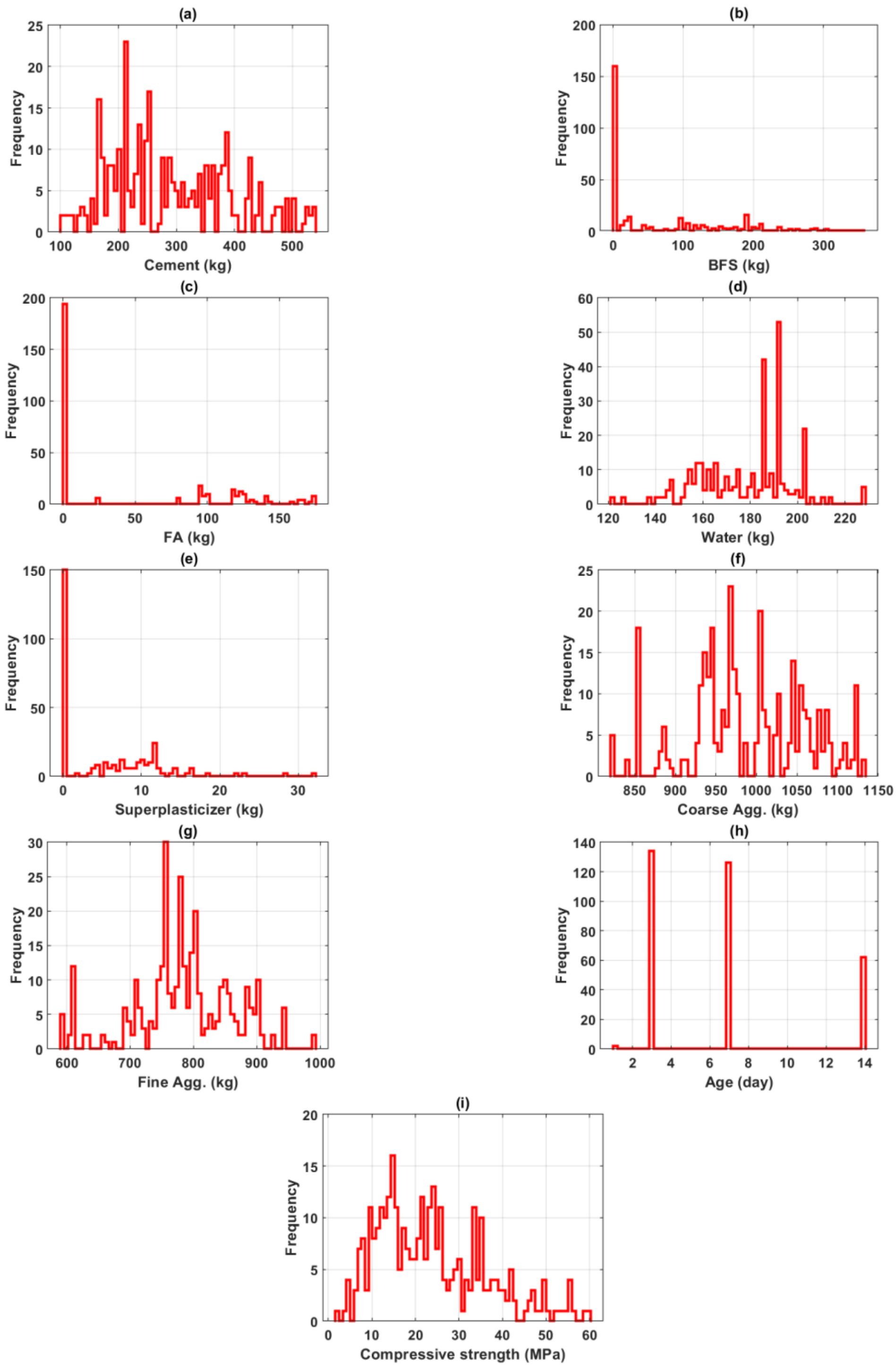

Fig. 1 Histogram including the ranges and frequencies of all the components in the database: (a) cement (b) blast furnace slag (c) fly ash (d) water (e) superplasticizer (f) coarse aggregate (g) fine aggregate (h) testing age (i) early compressive strength

experiments and testers, and the cost of these tests is also high (Rashid *et al*. 2002). Therefore, many researchers tried to find alternative ways to estimate the early compressive strength of HPC. Out of these approaches, machine learning is known as an objective and advanced technique for the prediction of mechanical properties of construction materials, including HPC. Mousavi *et al*. (Mousavi *et al*. 2012) used gene expression programming to predict the compressive strength of HPC. Prasad *et al*. (Prasad *et al*. 2009) used Artificial Neural Networks (ANN) to predict compressive strength HPC with high volume fly ash. In another study, one of the popular machine learning techniques, namely support vector machine (SVM) was used to predict concrete compressive strength (Chou *et al*. 2011). In general, the performance of these mentioned models is good and better than traditional approaches (Baykasoğlu *et al*. 2009).

In this study, the main aim is to apply the SVM model for the prediction of the early compressive strength of HPC. For this, data of a total of 324 experimental results were used to generate the training and testing datasets for the construction and validation of the model. Pearson correlation coefficient (R), Root mean square (RMSE) – common quantitative statistical indices were used to validate the predictive capability of the model. In addition, a sensitivity analysis was also carried out to evaluate the importance of input factors used in the modeling using Partial Dependence Plots.

2. Database preparation

The database in this study contained 324 experimental results from the works of Yeh (Yeh 1998a, Yeh 1998b, Yeh 2006, Yeh 2003, Yeh 1999). The cement type in the database was ordinary Portland cement (OPC), and all the concretes were under normal curing conditions. For the sake of simplicity, the HPC compressive strength values were unified to the unique 150-mm cylinders results using common guidelines (IS 516 1959, GB 50205 2001). The early HPC compressive strength was obtained using eight variables: cement content (kg), blast furnace slag content (kg), fly ash content (kg), water content (kg), superplasticizer content (kg), coarse aggregate content (kg), fine aggregate content (kg) and the testing age (day). Fig. 1 presents the ranges and frequencies of all the components.

The database was divided into two parts, subjected to the learning phase (the training dataset contained 70% of the data) and the validation phase (the testing dataset contained 30% of the remaining data). Each dataset was composed of 8 vectors representing eight input parameters, and the early compressive strength was the output. The chosen 70/30 ratio was selected following the suggestion in the work of Khorsheed *et al*. (Khorsheed and Al Thubaity 2013) or Leema *et al*. (Leema *et al*. 2016). It is easy to see the mixture components in this study covered a wide range of values, for instance, the cement content ranged from 100 to 560 kg, the slag content ranged from 0 to 380 kg with most of the values were 0, the fly ash content ranged from 0 to 180 kg with a high concentration at 0, the water content ranged from 120 to 230, the superplasticizer ranged from 0 to 33 kg, coarse aggregate ranged from 830 to 1150 kg, fine aggregate content varied from 580 to 1000 kg, the testing age composed of 3,7 and 14 days with one value of 1 day, and the compressive strength varied from2 to 60 MPa.

3. Support Vector Regression (SVR)

SVR is known as one of the most effective and popular tools in real-value function estimation.

It has been applied to predict travel-time (Wu *et al*. 2004), real-time flood stage (Yu *et al*. 2006), financial time series (Lu *et al*. 2009), and tourism demand (Chen and Wang 2007). In construction material science, it has been applied to predict the compressive strength of concretes containing metakaolin (Safarzadegan Gilan *et al*. 2012), atmospheric corrosion of metallic materials (Fang *et al*. 2008), and the unconfined compressive strength and permeability of concrete (Sun *et al*. 2019). Basically, a symmetrical loss function is used to train SVR, and a flexible tube of minimal radius is formed symmetrically around the estimated function utilizing Vapnik's -insensitive approach (Awad and Khanna 2015). An advantage of SVR is that its computational complexity does not depend on the input space dimensionality; thus, its predictive capability is excellent for prediction and regression (Smola and Schölkopf 2004). In this study, SVR was selected for investigation and prediction of the early compressive strength of concrete. To validate the performance of this model, Pearson correlation coefficient (R) is an important measurement in regression analysis, representing the correlation between the machine learning output and the targets. The values of R vary from -1 to 1, and higher absolute values of R close to 1 indicates better prediction accuracy. Root mean square (RMSE) is the second measurement used in this study, indicates the average squared difference between the machine learning output and targets. Differently, lower values of RMSE indicate better performance of the machine learning algorithm. The calculation of R and RMSE is as follow.

$$\mathrm{R} = \sqrt{1 - \frac{\sum_{j=1}^{n}(x_i - y_i)}{\sum_{j=1}^{n}(x_i - \bar{y})}} \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{2}$$

where n is the number of samples; x_j and y_j are the actual and predicted outputs, respectively; and $\bar{y}$ is the mean of the predicted output.

4. Results and discussion

4.1 Prediction performance

In this section, the performance of SVR in predicting the early compressive strength of HPC is investigated. Obviously, the predicted outputs (discontinuous red lines) given by SVR were in strong correlation with the experiment results (continuous black lines), as shown in Fig. 2. The corresponding errors in each case are presented in Fig. 3. The mean error values were 0.40012 and 1.9008 for the training, testing parts, respectively. The standard deviation error values were 4.4248 and 5.1974 for the training, testing parts, respectively. The computed values of RMSE were 4.4331 and 5.5089 for the training and testing parts, respectively. In both cases, the most frequented errors were in the range between -5 to 5 MPa, showing the high effectiveness of SVR in predicting the early compressive strength of HPC.

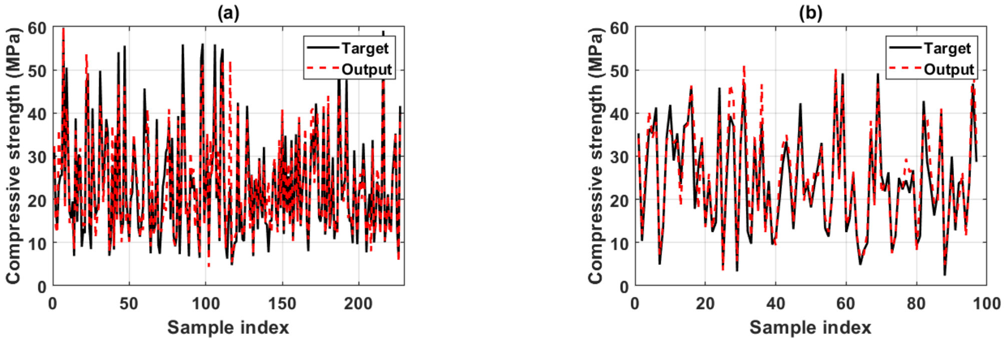

Fig. 2 Experimental and predicted values of early compressive strength of HPC with different datasets: (a) training part and (b) testing part

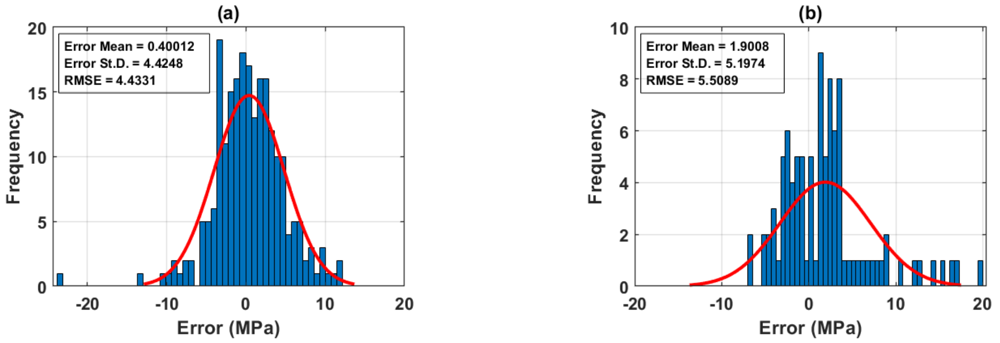

Fig. 3 Error histograms between experimental and predicted results of early compressive strength of HPC with different datasets: (a) training part and (b) testing part

The validation of SVR is presented in a regression analysis form in Fig. 4 for the training and testing models. The performance of SVR in predicting the early compressive strength values was satisfying with R=0.93364 and R=0.92386 for the training and testing datasets, respectively. It can be seen that, with respect to R values, the performance of the training part was slightly superior to the testing one. Besides, considering RMSE as the criterion, the training dataset was superior to the testing one. These results showed that a good fit of results was obtained by the developed SVR, as the errors from the testing part were somewhat higher than the training error.

Overall, the SVR model is a good predictor for the problem as such an algorithm could successfully predict the early compressive strength of HPC with high accuracy and low error.

4.2. Sensitivity analysis based on partial dependence plots

The sensitivity analysis using Partial Dependence Plots (PDP) is performed in this section. Two groups of input variables were identified, the first one consisted of all the variables exhibiting a positive effect on the early compressive strength, whereas the second one was composed of inputs that exhibiting a negative effect on the early compressive strength of HPC.

Considering the first group, the effect of each variable is highlighted in Fig. 5. It can be observed that the early compressive strength varied from 9.794 to 42.15 MPa with the variation of cement, 21.73 to 28.09 MPa with the variation of blast furnace slag, 23.43 to 25.13 MPa with the

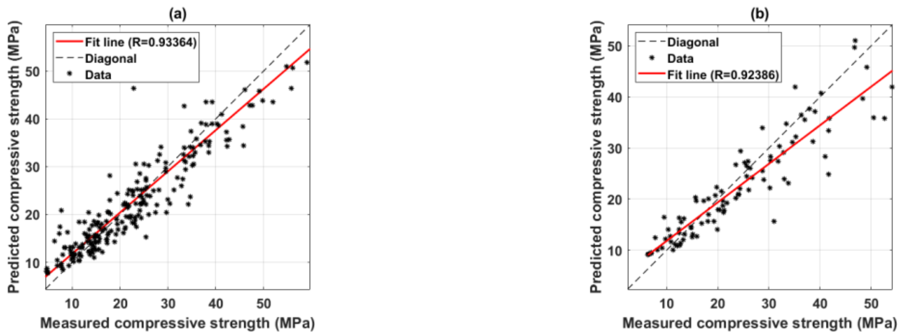

Fig. 4 Correlation graphs of SVR in predicting the early compressive strength of HPC: (a) training part and (b) testing part

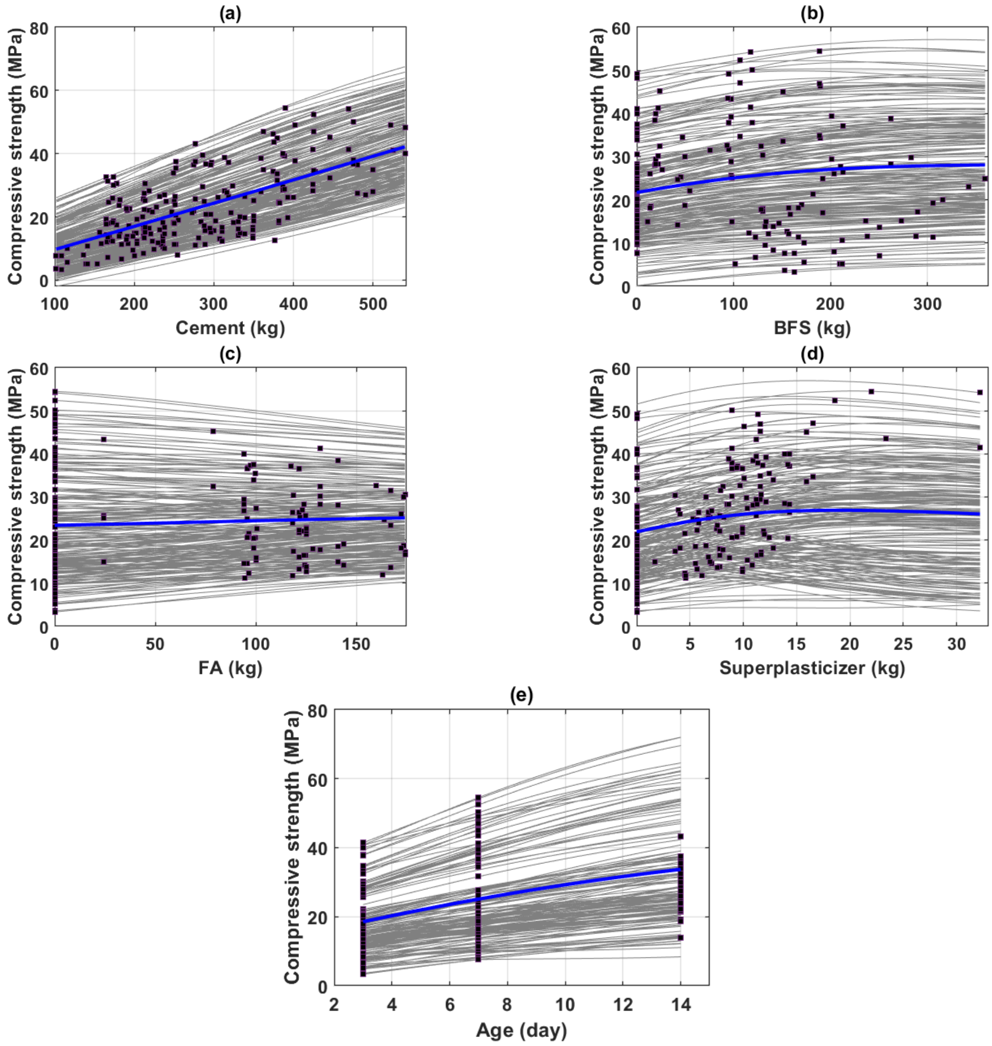

Fig. 5 Input variables exhibited a positive effect to the early compressive strength of HPC: (a) cement (b) blast furnace slag (c) fly ash (d) superplasticizer and (e) testing age

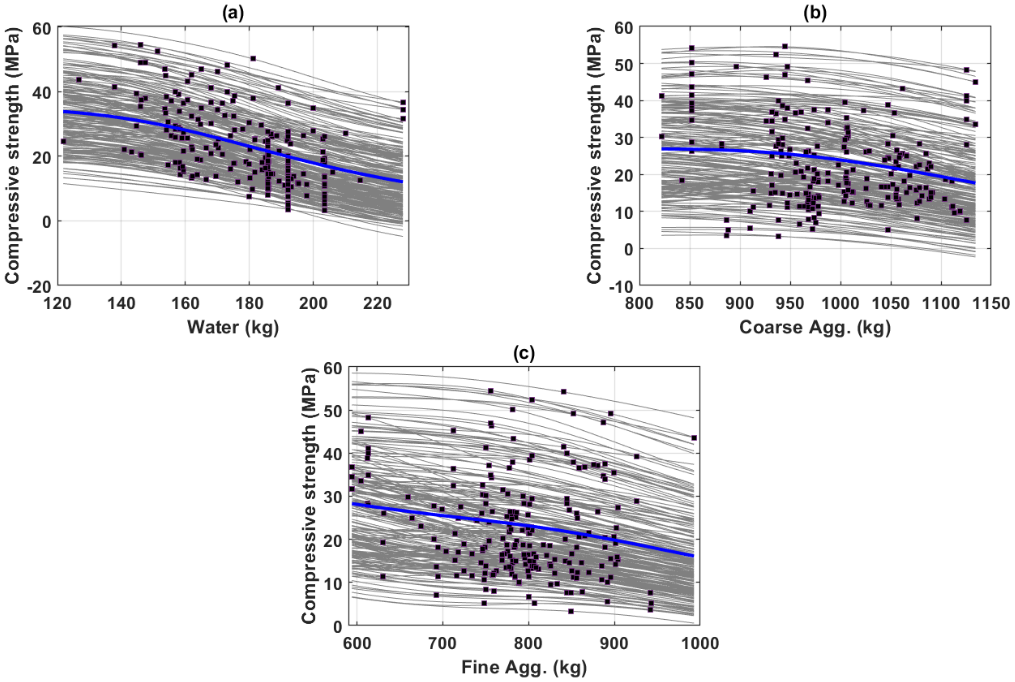

Fig. 6 Input variables exhibited a negative effect to the early compressive strength of HPC: (a) water; (b) coarse aggregate; and (c) fine aggregate

variation of fly ash, 21.92 to 26.06 MPa with the variation of superplasticizer content, 18.56 to 33.73 MPa with the variation of testing age. Thus, cement content was the most influential factor to the early compressive strength of HPC, following by testing age, blast furnace slag, superplasticizer content, and fly ash. The cement content is well-known the most affecting factor to the compressive strength. Besides, the effect of BFS or FA on the early strength development of HPC was experimentally confirmed (Nath and Sarker 2014).

Considering the second group, the effect of each variable is displayed in Fig. 6. It can be observed that the early compressive strength varied from 33.86 to 12.1 MPa with the variation of water content, from 26.92 to 17.17 MPa with the variation of coarse aggregate, from 28.28 to 16.14 MPa with the variation of fine aggregate. The water content was the most influential factor in this group, following by fine aggregate and coarse aggregate. As it would be expected and confirmed by several studies in the literature (Naik and Ramme 1989), stated that concrete containing cement replacement with fly ash allowed a decrease in the amount of water used.

5. Conclusions

In this study, the feasibility of using support vector regression (SVR) to predict the early compressive strength of HPC was investigated. A database containing 324 data points was used for the development and validation of the proposed SVR model with the training dataset (70% of data) and the testing dataset (30% of the remaining data), respectively. The assessment of model accuracy was performed thanks to common statistical criteria such as the Pearson correlation coefficient (R) and the root mean square error (RMSE). It was shown that the SVR model could successfully predict the early compressive strength of HPC with R=0.92386 and RMSE=5.5089 MPa for the testing dataset. In addition, two groups of input factors were identified via PDP analysis to deduce the positive and negative effects on the early compressive strength of HPC. Cement content was found as the most important exhibiting a positive effect, whereas the water content was the most critical factor that exhibiting a negative effect on the early compressive strength of HPC. Overall, the results of this study could pave the way for engineers to quickly and accurately estimate the strength behavior of HPC.

References

Awad, M. and Khanna, R. (2015), "Support vector regression", *In Efficient learning machines*, 67-80, Apress, Berkeley, C.A.

Baykasoğlu, A., Öztaş, A. and Özbay, E. (2009), "Prediction and multi-objective optimization of high-strength concrete parameters via soft computing approaches", *Expert Syst. Appl*., **36**, 6145-6155. https://doi.org/10.1016/j.eswa.2008.07.017.

Chan, Y., Peng, G. and Anson, M. (1999), "Residual strength and pore structure of high-strength concrete and normal strength concrete after exposure to high temperatures", *Cement Concrete Compos*., **21**, 23-27. https://doi.org/10.1016/S0958-9465(98)00034-1.

Chen, K.Y. and Wang, C.H. (2007), "Support vector regression with genetic algorithms in forecasting tourism demand", *Tourism Managem*., **28**, 215-226. https://doi.org/10.1016/j.tourman.2005.12.018.

Chou, J.S., Chiu, C.K., Farfoura, M. and Al Taharwa, I. (2011), "Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques", *J. Comput. Civil Eng*., **25**, 242-253. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000088.

Fang, S., Wang, M., Qi, W. and Zheng, F. (2008), "Hybrid genetic algorithms and support vector regression in forecasting atmospheric corrosion of metallic materials", *Comput. Mater. Sci*., **44**, 647-655. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000088.

Gran, J., Florence, A. and Colton, J. (1989), "Dynamic triaxial tests of high-strength concrete", *J. Eng. Mech*., **115**, 891-904. https://doi.org/10.1061/(ASCE)0733-9399(1989)115:5(891).

Khorsheed, M.S. and Al Thubaity, A.O. (2013), "Comparative evaluation of text classification techniques using a large diverse Arabic dataset", *Lang Resources Evaluation*, **47**, 513-538. https://doi.org/10.1007/s10579-013-9221-8.

Kim, H.S., Lee, S.H. and Moon, H.Y. (2007), "Strength properties and durability aspects of high strength concrete using Korean metakaolin", *Construct. Build. Mater*., **21**, 1229-1237. https://doi.org/10.1016/j.eswa.2008.07.017.

Köksal, F., Altun, F., Yiğit, İ. and Şahin, Y. (2008), "Combined effect of silica fume and steel fiber on the mechanical properties of high strength concretes", *Construct. Build. Mater*., **22**, 1874-1880. https://doi.org/10.1016/j.conbuildmat.2007.04.017.

Leema, N., Nehemiah, H.K. and Kannan, A. (2016), "Neural network classifier optimization using differential evolution with global information and back propagation algorithm for clinical datasets", *Appl. Soft Comput*., **49**, 834-844. https://doi.org/10.1016/j.asoc.2016.08.001.

Lu, C.J., Lee, T.S. and Chiu, C.C. (2009), "Financial time series forecasting using independent component analysis and support vector regression", *Decision Support Syst*., **47**, 115-125. https://doi.org/10.1016/j.dss.2009.02.001.

Mousavi, S.M., Aminian, P., Gandomi, A.H., Alavi, A.H. and Bolandi, H. (2012), "A new predictive model for compressive strength of HPC using gene expression programming", *Advan. Eng. Softw*., **45**, 105-114. https://doi.org/10.1016/j.advengsoft.2011.09.014.

Naik, T.R. and Ramme, B.W. (1989), "High early strength concrete containing large quantities of fly ash", *ACI Mater. J*., **86**, 111-116.

Nath, P. and Sarker, P.K. (2014), "Effect of GGBFS on setting, workability and early strength properties of fly ash geopolymer concrete cured in ambient condition", *Construct. Build. Mater*., **66**, 163-171. https://doi.org/10.1016/j.conbuildmat.2014.05.080.

Öztaş, A., Pala, M., Özbay, E., Kanca, E., Caglar, N. and Bhatti, M.A. (2006), "Predicting the compressive strength and slump of high strength concrete using neural network", *Construct. Build. Mater*., **20**, 769-775. https://doi.org/10.1016/j.conbuildmat.2005.01.054.

Prasad, B.R., Eskandari, H. and Reddy, B.V. (2009), "Prediction of compressive strength of SCC and HPC with high volume fly ash using ANN", *Construct. Build. Mater*., **23**, 117-128. https://doi.org/10.1016/j.conbuildmat.2008.01.014.

Rashid, M., Mansur, M. and Paramasivam, P. (2002), "Correlations between mechanical properties of high-strength concrete", *J. Mater. Civil Eng*., **14**, 230-238. https://doi.org/10.1061/(ASCE)0899-1561(2002)14:3(230).

Safarzadegan Gilan, S., Bahrami Jovein, H. and Ramezanianpour, A.A. (2012), "Hybrid support vector regression - Particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin", *Construct. Build. Mater*., **34**, 321-329. https://doi.org/10.1016/j.conbuildmat.2012.02.038

Smola, A.J. and Schölkopf, B. (2004), "A tutorial on support vector regression", *Statistics Comput*., **14**, 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

Sun, J., Zhang, J., Gu, Y., Huang, Y., Sun, Y. and Ma, G. (2019), "Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression", *Construct. Build. Mater*., **207**, 440-449. https://doi.org/10.1016/j.conbuildmat.2019.02.117.

Wu, C.H., Ho, J.M., Lee, D.T. (2004), "Travel-time prediction with support vector regression", *IEEE Transact. Intel. Transport. Syst*., **5**, 276-281. https://doi.org/10.1109/TITS.2004.837813.

Yeh, I.C. (1998), "Modeling concrete strength with augment-neuron networks", *J. Mater. Civil Eng*., **10**, 263-268. https://doi.org/10.1061/(ASCE)0899-1561(1998)10:4(263).

Yeh, I.C. (1998), "Modeling of strength of high-performance concrete using artificial neural networks", *Cement Concrete Res*., **28**, 1797-1808. https://doi.org/10.1016/S0008-8846(98)00165-3.

Yeh, I.C. (1999), "Design of high-performance concrete mixture using neural networks and nonlinear programming", *J. Comput. Civil Eng*., **13**, 36-42. https://doi.org/10.1061/(ASCE)0887-3801(1999)13:1(36).

Yeh, I.C. (2003), "A mix proportioning methodology for fly ash and slag concrete using artificial neural networks", *Chung Hua J. Sci. Eng*., **1**, 77-84.

Yeh, I.C. (2006), "Analysis of strength of concrete using design of experiments and neural networks", *J. Mater. Civil Eng*., **18**, 597-604. https://doi.org/10.1061/(ASCE)0899-1561(2006)18:4(597).

Yu, P.S., Chen, S.T. and Chang, I.F. (2006), "Support vector regression for real-time flood stage forecasting", *J. Hydrology*, **328**, 704-716. https://doi.org/10.1016/j.jhydrol.2006.01.021.

PA