Using Support Vector Machine to Improve Ultrasonic Pulse Velocity Test for Concrete

Loan Ngo Tu Quynh\textsuperscript{1, a} and Yu-Ren Wang\textsuperscript{1}

\textsuperscript{1}Department of Civil Engineering, National Kaohsiung University of Applied Sciences, Taiwan

Abstract. In the construction industry, to evaluate the compressive strength of concrete, destructive and non-destructive testing methods are used. Non-destructive testing methods are preferable due to the fact that those methods do not destroy concrete samples. However, they usually give larger percentage of error than using destructive tests. Among the non-destructive testing methods, the ultrasonic pulse velocity test is the popular one because it is economic and very simple in operation. Using the ultrasonic pulse velocity test gives 20\% MAPE more than using destructive tests. This paper aims to improve the ultrasonic pulse velocity test results in estimating the compressive strength of concrete using the help of artificial intelligent. To establish a better prediction model for the ultrasonic pulse velocity test, data collected from 312 cylinder of concrete samples are used to develop and validate the model. The research results provide valuable information when using the ultrasonic pulse velocity tests to the inputs data in addition with support vector model by learning algorithms, and the actual compressive strengths are set as the target output data to train the model. The results show that both MAPEs for the linear and nonlinear regression models are 11.17\% and 17.66\% respectively. The MAPE for the support vector machine models is 11.02\%. These research results can provide valuable information when using the ultrasonic pulse velocity test to estimate the compressive strength of concrete.

1 Introduction

In recent years, support vector machines have been widely used in various sectors. Non-destructive testing (NDT) methods used in the construction industry to examine the compressive strength of concrete are not only important alternatives but also feasible and economical as well. When using destructive methods in the lab, the concrete compressive strength test results in some cases do not represent the in-situ cast concrete because the strength and the quality of concrete could be affected through the transportation process, the placement, tamping and curing. Moreover, taking the core samples on site is not always feasible since the drilling process might damage the structure. Therefore, non-destructive tests are more suitable in measuring the concrete strength in cases where destructive tests are not. Typically, ultrasonic pulse velocity (UPV) is a non-destructive testing method to estimate concrete compressive strength without damaging the concrete structure. The testing method can effectively evaluate the uniformity and relative quality of concrete structures. Besides, UPV has some advantages such as low cost, easier to operate and convenient to carry.

The actual concrete compressive strength is a self-compact that refers to the unconfined compressive strength in geotechnical term [1]. It is a kind of self-flowing cementitious material, also known as the flowable fill, the un-shrinkable fill, the controlled density fill, and other names [2]. UPV testing is done by measuring the travelling period and distance of an ultrasonic pulse, the ultrasonic wave velocity can be obtained through the travelling time and distance. It is shown in [3] that the better the concrete quality, the higher the ultrasonic wave velocity.

To predict the concrete compressive strength, a prediction model is constructed. The author of [4] stated that Multiple Linear Regression cannot give a precise prediction for compressive strength. In [5], it is said that using Artificial Neural Network is more suitable to give the estimated displacement in concrete reinforcement building; in addition, the bearing capacity of shallow foundations also be determined by neuro-fuzzy models.

The investigation constructed in this paper is a process using simulation, which aims to improve the accuracy of the results of ultrasonic pulse wave velocity mode when measuring concrete strength. Then with the use of the neural network, the support vector machine SVMs, the prediction and the generation of strength will be enhanced and predict the ultrasonic pulse velocity, this research aims to get the intensity value closer to the actual compressive strength value. In addition, this research will provide proof that using SVMs can improve the regression results and artificial models created predict good results.
2 Proposed diagnostic method

2.1 Ultrasonic Pulse Velocity

UPV is used to evaluate the property of concrete. UPV test is done based on the principle of measuring the speed of a pulse of ultrasonic wave passing through concrete knowing the traveling time and the traveling distance [6]. Paper [7] has the full report on the correlation between the compressive strength, the pulse velocity and the elastic modulus. In [8], a relationship between pulse velocity and compressive strength is also stated. [9] proposes an empirical equation on the relationship between the unconfined compressive strength (UCS) and the pulse velocity. The equation as below:

\[
UCS = a e^{bV} 
\]  

where the properties of material will determine a and b, and V is the speed of ultrasonic pulse.

2.2 UPV Experimental Setup

In this study, the instrument used for ultrasonic pulse velocity test is a TICO concrete ultrasonic detector developed by Process Company in Zurich, Switzerland. As shown in Figure. 1, the instrument can perform the calculation and evaluation on functions of concrete uniformity, column hole, crack depth, elastic modulus, and concrete strength. The ultrasonic pulse velocity test is conducted in this study as follows: the measurement will be done four times for each of the cylindrical concrete sample with a circular top view of diameter 12 cm and height of 24 cm (refer to Figure. 2).

The ultrasonic pulse velocity tests are conducted for each of the 312 test samples. The travelling distance is the distance \(L\) between two ends of the transducer. The travelling time \(T\) is the time between the excitation to the moment that the first pressure wave is received by the probe. The ultrasonic pulse velocity is the quotient of distance and time.

The ultrasonic equipment consists of three parts: the ultrasonic pulser, the receive amplifier and the time measurement device that indicates the time of ultrasonic pulse travel through the medium. The detected data is displayed and recorded in the control Host.

\[
v = \frac{L}{T} \times 10^3 
\]  

where, \(l\) is the specimen length (peat–cement–sand brick with length measured in millimeters (mm)), \(v\) is the velocity of ultrasound (m/s) and \(t\) is the actual travelling time through medium of the ultrasonic pulse, measured in microseconds (\(\mu s\)).

![Figure 1. Ultrasonic pulse velocity test and TICO ultrasonic pulse velocity](image1)

![Figure 2. Model of schematic diagram of pulse velocity testing circuit](image2)

According to different settings of transducers location, there are three possible ways of measuring pulse of ultrasonic pulse velocity tests [10]. Measurements can be made by the following methods:

- Direct transmitter method and receiver are on two parallel surfaces was placed on the opposite sides of the concrete, pulse transmission path and the concrete surface perpendicular to the path length is the distance between two parallel transducers
- Indirect method or surface transmission where both transducers are on the same surface, pulse transmission path as the right triangle triangular oblique edge, the path length of the two electro-acoustical transducers center point distance.
- Semi-direct Methods, transmitters and receivers are on two mutually perpendicular surfaces or two different surfaces, need to put two electro-acoustical transducers at different distances repeated measurements, and then measured The return time is the average pulse wave velocity.

![Figure 3. Ultrasonic Pulse Velocity test arrangements](image3)
2.3. Support Vector Machines

Developed base on statistical learning theory, Support Vector Machines (SVMs) is an algorithm often used for small scale, nonlinear or high dimensional problems. The foundations of SVMs have been developed by Vapnik [11]. It is not only well founded theoretically, but also superior in application domains including object recognition problems, prediction formula, protein structure prediction, functional classification and also been successfully used in the construction-related areas in recent years [12]. SVMs in this research is used to create the prediction model for concrete compressive strength. The predicted result then will be used in comparison with the real test result.

In SVMs, the upper bound of generalization error has been minimized. This enables the ability to process data better than other artificial intelligence techniques even for unseen data [13]. Some basic concepts of SVMs model will be described as follow [14].

A dataset \( G = \{(x_i, d_i)\}_i \) is given with input vector \( x_i \), target value \( d_i \), and the size of the dataset \( n \). By mapping \( x \) into high-dimensional feature space, the nonlinear regression \( f(x) \) in low dimensional space is represented as \( f(x) = \omega \Phi(x) + b \) with \( \omega \) is the weight vector, \( \Phi(x) \) is the high dimensional feature space and \( b \) the bias of the hyper plane. In SVMs, by minimizing the risk penalty function, the parameters above can be found.

\[
f_{\omega,b}(x) = \arg \min_{\omega,b} R_{\exp}[f]
\]

(3)

The empirical risk minimizing makes sense only if:

\[
\lim_{n \to \infty} \min_{\omega,b} R_{\exp}[f] = R[f]
\]

Using the law of large numbers:

\[
\lim_{n \to \infty} \min_{\omega,b} R_{\exp}[f] = \min_{\omega,b} R[f]
\]

Next, this condition is less intuitive and requires that the minima also converge:

\[
R[f] \leq R_{\exp}[f] + \frac{\frac{1}{2} \ln(\frac{2n}{\delta}) - \ln(\frac{4}{\delta})}{l}
\]

By introducing the Lagrange multipliers: \( a_i \) and \( a_i^* \), the SVMs decision function becomes:

\[
f(x,a_i,a_i^*) = \sum_{i=1}^{n} (a_i - a_i^*) K(x_i, x) + b
\]

where \( K(x_i, x) \) is defined as the Kernel Function. It is the inner product of \( x \) and \( x_i \) in the feature spaces \( \psi(x_i) \) and \( \psi(x) \) respectively. There are four basic Kernel functions used in Support Vector Machines as follows [15].

**Linear Kernel:** \( K(x_i, x_j) = x_i^T x_j \)

**Polynomial Kernel:** \( K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \)

**Sigmoid Kernel:** \( K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \)

Radial Basis Function (RBF) Kernel:

\[
K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2), \gamma > 0
\]

With \( \gamma, r \) and \( d \) are Kernel parameters.

2.4 Regression SVM

Regression is performed by using support vector regression SVR, this supervised learning method generates an SVM by input-output mapping functions from a labeled training dataset. This function solves both classification and regression problems. Abubakar et al. (2013) stated that in a regression SVM, the functional dependence of the dependent variable \( y \) on a set of independent variables \( x \) has to be estimated. It assumes, like other regression problems, that the relationship between the independent and dependent variables is given by a deterministic function \( f \) plus the addition of some additive noise [16]:

\[
y = f(x) + \text{noise}
\]

(4)

The task is then to find a functional form for \( f \) that can correctly predict new cases that the SVM has not been presented with before. This can be achieved by training the SVM model on a sample set, i.e., training set, a process that involves, like classification (see above), and the sequential optimization of an error function [17,18]. Depending on the definition of this error function, two types of SVM models can be recognized:

For this type of SVM the error function is:

\[
\frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i + \gamma \sum_{i=1}^{n} \zeta_i
\]

The error function was minimized subject to:

\[
w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i, y_i - w^T \phi(x_i) - b_i \leq \epsilon + \zeta_i
\]

\[
\xi_i, \zeta_i \geq 0, i = 1, ..., N
\]

There are several numbers of kernels that can be used in Support Vector Machines models. These include linear, polynomial, radial basis function (RBF) and sigmoid:

That is, the kernel function, represents a dot product of input data points mapped into the higher dimensional feature space by transformation \( \phi \). Gamma is an adjustable parameter of certain kernel functions. The RBF is by far the most popular choice of kernel types used in Support Vector Machines. This is mainly because of their localized and finite responses across the entire range of the real \( x \)-axis.

The SVR uses \( \epsilon \) parameter (insensitive loss) to compute a linear regression function for a dimensional feature space while concurrently minimizing \( ||\omega||^2 \) to reduce model complexity.
Two new parameters $\xi_i, \xi_i^*$ are introduced to solve the minimization problem as shown in figure 4. Thus, SVR is formulated as minimizing following constraint-optimization problem:

Minimize $\frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{n} \xi_i + \xi_i^*$ subject to

\[
\begin{align*}
  y_i - \langle \omega, x_i \rangle - b & \leq \epsilon + \xi_i \\
  \langle \omega, x_i \rangle + b - y_i & \leq \epsilon + \xi_i^* \\
  \xi_i, \xi_i^* & \geq 0
\end{align*}
\]

determine the trade-off between the flatness of $f(x, w)$ and the tolerance for deviations larger than $\epsilon$.

### 2.5 Least squares support vector machine (LSSVM)

The least square support vector machine (LSSVM) have been developed and introduced by Suykens [20]. By using the equality constraints instead of the inequality ones, the solution of LSSVM follows from solving a set of linear equations and the computational complexity is rather lower than that of SVM. The LSSVM helps the process of programming for constrained optimization in SVMs less heavy. The error square is used as a measure of the error condition, which represents the error between the output and the correct result. So the LSSVM guidelines is represented as

\[
\begin{align*}
  \text{minimize} \quad & \frac{1}{2} \| w \|^2 + \gamma \frac{1}{2} \sum_i \xi_i^2, \\
  \text{subject to} \quad & y_i (w \cdot \Phi(x_i) + b) = 1 - \xi_i, \quad i = 1, \ldots, n
\end{align*}
\]

where $\gamma$ is the same as the regular parameter in the traditional SVMs. The aim is also to margin maximization the spacing with error minimization between the transactions $\xi_i$ is the wrong variable.

The difference between the LMSVM and the traditional SVM is that the LSSVM is an equation and the SVM is an inequality, and the optimal solution for the LSSVM is not as sparse as the traditional SVM (Kruif & Vries, 2003). And solve the problem of (4) optimization is the same as using the Lagrangian method, so LSSVM optimization problem to solve all the above linear equations of the problem, reducing the computational complexity. The solving process of LSSVM is mainly composed of two steps. Firstly, the final solution to the right weight vector can be gotten by the gradient descent algorithm. Secondly, for the fixed right weight vector in each iteration, an analytical solution to the left weight vector can be gained. The detailed description about at LSSVM can be found in [21]. Because the search is not for the sparse solution, so LSSVM does not like SVM to find some influential information as a support vector (support vector), but all the data as a support vector.

In this research a total of 312 cylinder concrete samples are collected in collaborating with a local material testing laboratory, with 312 test samples collected in the study, 252 of them are randomly chosen to be the training dataset. 60 remaining samples are used as the test group data. The model calculates the mean absolute percentage error (MAPE) and uses the Microsoft Excel 2013 and MATLAB 2013 suite of software to construct the LSSVM model. The model predicts the concrete strength of the concrete column will be compared with the actual compressive strength of the concrete column specimen, and further verify the model is used to predict the accuracy of the actual compressive strength of the concrete cylindrical specimen, and whether the result is more accurate and the construction process of the model is as follows figure 5.

**Step 1:** Import training group and test group data.

The prediction model will use the ultrasonic wave velocity of the ultrasonic wave to obtain the average velocity $V$ value as the input, and the compressive strength of the concrete is the output. In the 312 cylinder concrete samples, the average wave velocity ($X$) and the concrete compressive strength ($Y$) data as the LSSVM model training. The optimal prediction model is found by the model training process. The remaining 60 remaining samples are set as the best prediction model for the test group ($X_t$) to calculate the average absolute error percentage. The model calculates the mean absolute percentage error (MAPE) and uses the Microsoft Excel 2013 and MATLAB 2013 suite of software to construct the LSSVM model as follows:

\[
X = \text{xlsread}('X.xls')
\]

\[
Y = \text{xlsread}('Y.xls')
\]

\[
X_t = \text{xlsread}('X_t.xls')
\]

**Step 2:** Adjust the values of $C$ (sigma) and $\gamma$ (gamma) to establish the LSSVM prediction model.

In the LSSVM model analysis to adjust the value of $C$ (sigma) and $\gamma$ (gamma), repeatedly test, adjust and record the relationship between sig and gam, the input item data set through the training group instruction training, and then use the test group instructions the prediction results are predicted and the predicted results are compared with the actual compressive strength and the average absolute error percentage is calculated until a best combination of the parameters is found. The import command syntax is as follows:
Training group instruction
\[ [\alpha, b] = \text{trainlssvm}(\{X,Y,\'f',\text{gam}, \text{sig2}\}) \] (8)

Test group instructions
\[ Y_t = \text{simlssvm}(\{X,Y,\'f',\text{gam}, \text{sig2}\}, \{\alpha, b\}, X_t) \] (9)

In this research, a revolution of support vector machine called least square support vector machine (LSSVM) is used to import data into MATLAB to predict the model for properties of concrete after testing those properties in the lab. LSSVM is modified from existed SVMs and then the mean absolute percentage error (MAPE) is calculated. In figure 6, MATLAB gives the prediction strength data. The value for C and Y need to be setup differently for a better prediction model. Support vector instruction inputs are shown in figure 7.

3 Results Analysis

SVM is developed and highly valued for combined NDTs (SonReb) for using UPV tests to the inputs data. The actual compressive strengths are set as the target output data to train the model. Total 312 test samples are collected in the lab experiment, which 252 of them are randomly selected as the training dataset and 60 test patterns and reloaded to the testing dataset. The 60 samples in the testing dataset are “unseen” that remains in the data will not be seen or new data for the support vector machine model is trained.

The prediction accuracy is measured by the mean absolute percentage error (MAPE) illustrated in the equation below:

\[ \text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{A_i - P_i}{A_i} \right| \] (10)

where, \( A_i \) is the actual concrete sample compressive strength and \( P_i \) is the predicted strength.

MAPE is calculated by comparing the data of the test group into the equation with the actual compressive strength. The outputs predict the estimated concrete compressive strength are compared directly to the actual compressive strength obtained. The result for model prediction are shown in Table 1.

In UPV test, to detect the ultrasonic wave velocity of the ultrasonic wave, the setup are as follows: the transmitting probe and the receiving probe are placed on the upper and lower sides of the concrete cylindrical test specimen (distance 24cm); the speed of transmission of the pulse in the concrete specimen then calculated as mentioned in the introduction section. MAPE was calculated by comparing the data of the test group into the equation with the actual compressive strength. The linear and non-linear regression are illustrated as in figure 8 and 9, respectively.

<table>
<thead>
<tr>
<th>Forecasting model</th>
<th>Prediction formula</th>
<th>Standard deviation</th>
<th>MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>( fc = 0.3713 V - 1162.8 )</td>
<td>16.80</td>
<td>11.17</td>
</tr>
<tr>
<td>Nonlinear regression</td>
<td>( fc = 0.8648e^{0.0015x} )</td>
<td>23.13</td>
<td>17.66</td>
</tr>
</tbody>
</table>

Shown in figure 8, the linear regression model is
\[ y = 0.3713x - 1162.8 \]
where \( x \) is the average ultrasonic wave velocity and \( y \) is the actual compressive strength, and using 60 test group data bands. The MAPE calculated by comparing the predicted formula with the actual compressive strength is 11.17%. Similarly for the non-linear regression in figure 9, the model is
\[ y = 0.8648e^{0.0015x} \]
MAPE calculated is 17.66%.

As shown in Table 1, the compressive strength of concrete predicted by linear and nonlinear regression model is 11.17% and 17.66% respectively. It can be
found that the linear regression model is better than the non-linear regression one. In this study, we try different kinds of artificial intelligence methods to correct the inaccuracy of ultrasonic wave velocity prediction, and to establish different types of concrete strength estimation system for improving the reliability of concrete strength when using the ultrasonic wave velocity.

![Linear Regression Graphs](image1)

**Figure 8.** Linear regression graphs.

In this research a total of 312 cylinder concrete samples are collected in collaborating with a local material testing laboratory, 252 cylinder of them are randomly chosen to be the training dataset. 60 cylinder remaining samples are used as the test group data. The model calculates the mean absolute percentage error (MAPE) and uses the Microsoft Excel 2013 and MATLAB 2013 suite of software to construct the LSSVM model. And the data according to the program required to be organized into the variable X (252 cylinder training group data) the average velocity and strain Y (252 cylinder training group data of the actual strength) xlsx. File into the Matlab software by adjusting the C (Sigma) value to 2 to 512 in order to searching and comparing with the actual compressive strength to get the best parameter (Table 2).

A total of 312 samples, the settings of input variable were randomly selected as the test samples for the support vector machine (SVM) methods. For this models, the actual compressive strength is set up as output variable. The model prediction results are summarized in Table 3.

The results show that the best MAPE values of 11.02% respectively was obtained by SVMs test with the remaining 60 data as the model. The results show that the prediction results are improved by means of artificial intelligence. The model does improve the prediction accuracy of concrete compressive strength. When attempting to increase the input of the model, the results show that the maximum error rate of the concrete is less than 9.996% and 10.495% respectively for SVMs methods. The results are better than the predictions of the one input and two input, which proves that the increase in the input of the model can reduce the mean absolute percentage error.

**4 Conclusion**

To improve the result of non-destructive tests, SVMs is used in order to predict the concrete compressive strength. Within 312 test samples, 252 test samples are randomly chosen to be the train data set for SVMs. The remained 60 test samples are set as the data set to evaluate the prediction model accuracy. The model the support vector machine is developed.

The results shows that MAPE is about 20%, and the compressive strength of the concrete predicted by the linear and nonlinear regression model is 11.17% and 17.66% respectively. Therefore, the study show the establishing a predictive strength model with the goal for the intensity of ultrasonic wave velocity detection method becomes more reliable, and also increasing the prediction accuracy. Comparing the prediction of three models mentioned, the best MAPE values of 11.02% is obtained by the SVM test. In addition, the prediction results from one input (average ultrasonic pulse wave velocity) and two input (average ultrasonic pulse wave velocity and standard deviation) show that the maximum error rate of the concrete is less than 9.996% and 10.495% the mean absolute percentage error of the model. Which imply that the results are better than the predictions of the three input (3 velocity ultrasound V values). For the collected sample, the research results have shown that the increase in the input of the model can reduce the absolute error percentage.
Table 2. The LSSVM models prediction accuracy (one input, two input, three input).

<table>
<thead>
<tr>
<th>One input</th>
<th>Two input</th>
<th>Three input</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ</td>
<td>C</td>
<td>MAPE</td>
</tr>
<tr>
<td>200</td>
<td>2</td>
<td>9.595%</td>
</tr>
<tr>
<td>200</td>
<td>4</td>
<td>9.685%</td>
</tr>
<tr>
<td>200</td>
<td>8</td>
<td>9.700%</td>
</tr>
<tr>
<td>200</td>
<td>16</td>
<td>9.772%</td>
</tr>
<tr>
<td>200</td>
<td>32</td>
<td>9.873%</td>
</tr>
<tr>
<td>200</td>
<td>64</td>
<td>9.955%</td>
</tr>
<tr>
<td>200</td>
<td>128</td>
<td>9.996%</td>
</tr>
<tr>
<td>200</td>
<td>256</td>
<td>10.029%</td>
</tr>
<tr>
<td>200</td>
<td>512</td>
<td>10.051%</td>
</tr>
<tr>
<td>200</td>
<td>1000</td>
<td>10.195%</td>
</tr>
<tr>
<td>1500</td>
<td>2</td>
<td>9.452%</td>
</tr>
<tr>
<td>3000</td>
<td>2</td>
<td>9.476%</td>
</tr>
<tr>
<td>6000</td>
<td>2</td>
<td>9.471%</td>
</tr>
<tr>
<td>20000</td>
<td>2</td>
<td>9.483%</td>
</tr>
<tr>
<td>50000</td>
<td>2</td>
<td>9.489%</td>
</tr>
<tr>
<td>60000</td>
<td>2</td>
<td>9.634%</td>
</tr>
<tr>
<td>75000</td>
<td>2</td>
<td>9.719%</td>
</tr>
<tr>
<td>100000</td>
<td>2</td>
<td>9.720%</td>
</tr>
</tbody>
</table>

Table 3. The results of the test group data into the SVM model.

<table>
<thead>
<tr>
<th>SVM Model</th>
<th>C</th>
<th>γ</th>
<th>Standard deviation</th>
<th>Training MAPE (%)</th>
<th>Testing MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One input</td>
<td>2</td>
<td>1500</td>
<td>14.41</td>
<td>9.452</td>
<td>10.495</td>
</tr>
<tr>
<td>Two input</td>
<td>4</td>
<td>3000</td>
<td>14.11</td>
<td><strong>9.565</strong></td>
<td>9.996</td>
</tr>
<tr>
<td>Three input</td>
<td>4</td>
<td>1000</td>
<td>16.38</td>
<td><strong>10.855</strong></td>
<td>11.02</td>
</tr>
</tbody>
</table>

References