PREDICTING THE STRENGTH OF CEMENT MORTARS CONTAINING NATURAL POZZOLAN AND SILICA FUME USING MULTIVARIATE REGRESSION ANALYSIS

*Hany A. Dahish1,2, Mudthir Bakri1,3 and Mohammed Saleh Alfawzan1

1Civil Engineering Department, College of Engineering, Qassim University, Unaizah, Saudi Arabia
2Civil Engineering Department, Faculty of Engineering, Fayoum University, Egypt
3Civil Engineering Department, College of Engineering, Sudan University of Science and Technology, Sudan

*Corresponding Author, Received: 03 Nov. 2020, Revised: 01 Feb. 2021, Accepted: 11 Mar. 2021

ABSTRACT: In this study, mortars containing locally available natural pozzolan (NP) in Almadinah Almunawara, Kingdom of Saudi Arabia, were investigated as a partial substitute for sand or cement in mortars and silica fume (SF). The benefit of using local NP powder as a replacement for cement is that it reduces the carbon dioxide emission during the cement manufacturing process, whereas the benefit of using local NP as fine aggregates is that it reduces the density of the produced mortars and improves its properties because of its pozzolanic effect. Because of these reasons, there is a need to develop an effective predictive model to estimate the compressive strength of mortars with partial replacement of cement or sand with NP and with SF as a replacement for cement at 28 days. Data of 68 cubic specimens of 50 mm were established through experimental work with other researchers, and they were chosen to create a database for the proposed model. There were three input parameters: a) level of partial substitution of cement with NP powder, b) level of partial substitution of sand with NP, and c) level of partial substitution of cement with SF. The output parameter was compressive strength. Best correlations were obtained between the compressive strength and sand replacement with NP. To predict the compressive strengths of cement mortars containing NP and SF, multivariate regression models were proposed and compared to find the best one. It was concluded that the full quadratic model was the best model with highest correlation when compared with other proposed models.

Keywords: Compressive strength, Multivariate regression, Natural pozzolan, Prediction, Silica fume

1. INTRODUCTION

Concrete is considered a primary construction material because of its high compressive strength and durability. The partial substitution of cement in concrete with supplementary cementitious materials to improve its properties has complex behavior [1]. Silica fume (SF) is considered a highly effective pozzolanic material. The effect of partial substitution of cement with SF has been investigated by many researchers. Cement replacement with SF up to 7.5% improved the compressive strength of concrete [2].

The partial substitution of cement in concrete with SF and fuel ash leads to a higher strength than that with partial substitution of cement with iron filings [3]. Carmela et al. [4] used rice husk ash as a partial substitute for cement in mortars. Replacement level up to 10% of cement with rice husk ash is best for maximizing the strength of cement mortars. Arifi and Cahya [5] replaced 25% of cement in recycled aggregate pervious concrete with fly ash, leading to an improve in the mechanical properties. Lejano et al. [6] replaced 5% of cement in mortars with powdered eggshells, leading to an increase in the compressive strength by 36.4%. The 10% replacement of cement in concrete with coconut shell ash led to 92.1% of the strength of conventional concrete [7].

Many studies have been conducted on the possibility of utilizing different substances as partial substitutes for fine aggregate in concrete to improve its properties. The utilization of natural pozzolan (NP) as a partial replacement for sand in mortar cubes increased the compressive strengths for replacement levels up to 20%, 30%, and 40% for mortars with cement replacement with SF of 0, 5 and 10%, respectively [8]. Tampus et al. [9] utilized wood ash as a partial substitute for sand in mortars. The full substitution of sand with wood ash increased the compressive strengths of mortars.

The partial substitution of fine aggregate in concrete with plastic waste showed improvement in energy absorption under impact loading [10]. Granite quarry can be used as a substitute for sand in concrete at levels up to 60% to improve the compressive and flexural strengths [11]. The utilization of raw vermiculite as a substitute for sand in cement mortars decreased strength at high temperatures [12]. The replacement of 5% of cement and 15% of sand in concrete with limestone fines increased the compressive strength [13].
was recommended to replace sand in concrete with copper slag up to 50% to produce eco-friendly concrete [14]. Because of the complex behavior of concrete, many predictive models were proposed to study the effects of partial substitution of sand or cement in concrete with various supplementary cementitious materials on its compressive strength. A predictive model using multi-objective optimization method for the compressive strength of concrete containing SF have been proposed with achievement of 31 optimized Artificial Neural Networks (ANN) models [15]. A multivariate regression analysis model for studying the effect of cement replacement with various proportions of blast furnace slag and steel slag on the compressive strength of concrete at 28 days was developed [16].


The purpose of the present study is to investigate and predict the complex effects of utilizing NP as a partial substitute for sand or cement in cement mortars containing SF. For reliable prediction of compressive strength of cement mortar containing SF with cement or sand replacement with NP, several models, such as the linear strength model, pure quadratic model, interaction strength model, and full quadratic strength model, were adopted as tools to model the complex behavior instead of conducting direct laboratory tests to save energy, cost, and time. Different conditions for cement replacement with NP powder, sand replacement with NP, and cement replacement with SF were used as major data inputs in the models.

2. MATERIALS AND METHODS

The data used in this study was based on experimental work with other researchers [8]. Locally available natural pozzolan was the primary material used in this research work. Besides that, for cement mortar mixing purposes, ordinary Portland cement, local natural sand, and silica fume were used. The properties of cement mortars ingredients are as follows.

2.1 Materials

2.1.1 Powders

The used powders in this study were cement of type I, SF conforming ASTM C1240, and ground NP powder of 2000 to 3000 cm²/g fineness. The chemical composition of cement, SF, and NP were determined using X-ray fluorescence analysis and the results are as given in Table 1.

Table 1 Chemical composition of powders

<table>
<thead>
<tr>
<th>Oxides</th>
<th>Cement</th>
<th>SF</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>19.97</td>
<td>92</td>
<td>41.12</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>5.85</td>
<td>1.0</td>
<td>15.44</td>
</tr>
<tr>
<td>Fe₂O₃</td>
<td>3.43</td>
<td>1.0</td>
<td>17.35</td>
</tr>
<tr>
<td>Ca O</td>
<td>64.13</td>
<td>0.3</td>
<td>11.21</td>
</tr>
<tr>
<td>Mg O</td>
<td>0.6</td>
<td>0.6</td>
<td>4.47</td>
</tr>
<tr>
<td>SO₃</td>
<td>2.8</td>
<td>0.3</td>
<td>0.18</td>
</tr>
<tr>
<td>K₂O</td>
<td>0.72</td>
<td>0.54</td>
<td>1.05</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.16</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>LOI</td>
<td>1.60</td>
<td>1.57</td>
<td>2.20</td>
</tr>
</tbody>
</table>

2.1.2 Aggregates

The bulk density, specific gravity, and water absorption of natural sand and NP were (1670 kg/m³, 2.60, and 0.21%) and (1100 kg/m³, 2.51, and 5.23%), respectively. The gradations of sand and NP with upper and lower limits are shown in Fig.1.

Fig.1 Gradations of sand and NP

2.1.3 Water

Normal tap water was used for both mixing and curing purposes for mortars. Table 2 lists the quantities of the used materials.
Table 2 Cement mortar constituents

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Sand (gm)</th>
<th>Cement (gm)</th>
<th>w/c ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1356</td>
<td>490</td>
<td>0.49</td>
</tr>
</tbody>
</table>

2.2 Experimental Work

Sixty-eight cubic specimens of 50 mm of different mortar mixes with different ratios of local natural pozzolan were prepared. The experimental program focused on studying the effect of cement replacement with local NP powder and SF, and the effect of sand replacement with local NP on the compressive strength of cement mortars.

The levels of replacement of cement with NP (CRNP) by weight ranged from 10% to 40%. The levels of cement replacement with SF (CRSF) were 5 and 10%. The levels of replacement of sand with local NP (SRNP) by volume ranged from 10% to 40%. The curing period was 28 days.

2.3 Compressive Strength

The compressive strength results of cement mortar cubes at different conditions of cement replacement with NP powder, sand replacement with NP, and cement replacement with SF at 28 days are presented in Table 3, Fig.2, and Fig.3.

![Fig.2 Compressive strengths (replacing C-28 days)](image)

Table 3 Compressive strength results

<table>
<thead>
<tr>
<th>Repl. Level</th>
<th>Silica Fume 0%</th>
<th>Silica Fume 5%</th>
<th>Silica Fume 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>45.8</td>
<td>60.4</td>
<td>62.6</td>
</tr>
<tr>
<td>Sand</td>
<td>10</td>
<td>48.7</td>
<td>56.1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>47.3</td>
<td>54.8</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>44.6</td>
<td>46.5</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>38.2</td>
<td>42.9</td>
</tr>
<tr>
<td>Cement</td>
<td>10</td>
<td>39.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>37.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>28.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>21.1</td>
<td></td>
</tr>
</tbody>
</table>

Fig.3 Compressive strengths (replacing S-28 days)

2.4 Multiple Linear Regression (MLR)

Regression analysis is an effective mathematical method that employs statistics to forecast the existence of the relationship among multiple variables. MLR attempts primarily to understand more about the interactions between multiple independent variables or predictors and dependent variable [25]. This method is commonly used to predict concrete compressive strengths at various ages.

The provision of a single value of a criterion derived from one variable predictor is a straightforward linear regression, whereas two or more variables forecast the criterion in the MLR. Thus, multiple regression explores the relationships between various independent and dependent variables. The general equation form of MLR is:

\[ Y = a + \sum_{i=1}^{n} b_i X_i + \epsilon \]  

(1)

Where:
- \( Y \) is the dependent variable
- \( X_i \) is the independent variable, and \( n \) is the number of variables
- \( b_i \) is the regression coefficient
- \( a \) is the constant
- \( \epsilon \) is error

2.5 Multiple Nonlinear Regression Analysis

The dynamic relationship among the independent variable and the function is estimated by utilizing nonlinear regression. The interaction of various independent and dependent variable parameters is estimated by nonlinear multivariable regression. This refers to various modelling and regression methods for certain factors, particularly factors where there is a limited data supply [26]. The most important variables are described in step-by-step regression, which may account for the strongest correlation between the response variable and independent variables. This approach is an algorithmic one intended to select the right model.
subsets to filter in a forward or reverse direction. The first direction involves choosing a permanent model and using model terms before maximizing their fitness. That is, step-by-step regression may be seen as a forward selection method that, at each step, tests the relevance of all variables used previously. If the partial square sums for previously used variables fail to conform to the minimal norm to remain in the model, improvements to the retroactive exclusion process will be introduced and variables will be eliminated one at a time before the other variables quantities fitting the minimum criteria. Step-by-step regression involves more calculations than forward or backward detailed calculations, but it comprises a better probability of discovering the right subset models [27].

A multiple nonlinear regression (MNR) analysis on the mechanical properties was shown with the second-order polynomial firstly suggested by Schef`e [28]. Various forms of MNLR, such as pure quadratic and complete quadratic versions, have been investigated.

The following sections discuss mathematical expressions as well as each model. Table descriptions are also given, such as root mean Square Error, R² (coefficient of determination), and F-value (p) of each model.

**Coefficient of Determination (R²)**

The degree to which the explanatory variables account for the calculated response variable is defined (R²), evaluates how solid the linear relationship is between two or more variables. The higher the R² value, the greater this model’s predictive power. The mathematical expression for R² is given as follows:

\[
R^2 = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})(\bar{X}_i - X)}{\sqrt{\left(\sum_{i=1}^{n} (X_i - \bar{X})^2\right)\left(\sum_{i=1}^{n} (Y_i - \bar{Y})^2\right)}}
\]

**Root mean square error RMSE**

A statistical indicator of predictive precision is the Root Mean Squared Error (RMSE). The RMSE calculates the sum of variation to be explained by a formula in the response variable. The equation of the RMSE is as follows:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2}
\]

In Eq. (3), \(y_i\) and \(\bar{y}\) are the actual measures from the laboratory experiments and the magnitude estimated by employed model, respectively, whereas \(N\) is the sum of data set A lower RMSE would mean lower values in a model’s overall errors and would result in a more beneficial predictive strength.

### 3. RESULTS

The summary of the data statistics of the proposed models is presented in Table 4.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>N</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP Repl. Cement</td>
<td>68</td>
<td>0</td>
<td>40</td>
<td>4.85</td>
<td>11.13</td>
</tr>
<tr>
<td>SF Repl. Cement</td>
<td>68</td>
<td>0</td>
<td>40</td>
<td>3.75</td>
<td>4.09</td>
</tr>
<tr>
<td>NP Repl. Sand</td>
<td>68</td>
<td>0</td>
<td>40</td>
<td>14.71</td>
<td>15.11</td>
</tr>
<tr>
<td>Stress, MPa</td>
<td>68</td>
<td>20.88</td>
<td>63.56</td>
<td>46.89</td>
<td>10.46</td>
</tr>
</tbody>
</table>

The trend of the compressive strength (\(F_{c28}\)) of cement mortars is presented in Fig.4 for cement replacement with NP powder at 28 days and in Fig.5 for sand replacement with NP at 28 days.

\[ f_{c28} = -0.6365 \text{CRNP} + 47.063 \]

\[ R^2 = 0.9294 \]

\[ f_{c28} = -0.0147 \text{SRNP}^2 + 0.4042 \text{SRNP} + 45.59 \]

\[ R^2 = 0.9923 \]

### 3.1 Multivariate Regression Models

In the context of concrete construction, strength at 28 days is the referred measurement of design strength and the critical quality control parameter.
Compressive strength at 28 days is a generally agreed indicator, which typically measures the strength of the concrete by means of a standard axial compression test. In section, hereinafter the Multilinear regression model to estimate compressive strength of mortar cubes is demonstrated. In addition to the Multilinear regression, the pure quadratic, complete quadratic and interaction models are fully explored and compared to find out the best model to predict the compressive strength of cement mortars with NP and SF as partial replacements.

3.1.1 Linear strength model

The linear power model only includes linear parameters and constant terms. Mathematical Equation (4) demonstrates the Multilinear compressive strength model:

\[
F_{cu_{28}} = 50.79 - 0.761 \times CRNP + 1.211 \times CRSF - 0.322 \times SRNP
\]  

Fig.6 displays a normal distribution plot of the standardized residues of the linear regression process. At 28 days, the regression line runs through 68 samples of compressive strength. The figure states independency of errors from each other. Residuals seem uniformly distributed around zero. The mathematical description of the model is shown in Table 4. It is obvious that the likelihood is less than 0.0005 (Table 4) based on the F statistic (Fisher statistic).

The proposed multilinear regression for compressive strength at 28 days is extremely statistically significant with more than 99.95% confidence. It shows the goodness of the model for the data. \( R^2 \) of the multilinear mortar strength model is 89.6%, which designates a good fit, and the RMSE is 3.45. Fig.7 displays the scattered plot of the experimental and expected compressive strengths on the data order of the experiments’ tests. It revealed a similarity between predicted magnitude of strength and laboratory results during the experimental program.

The coefficients of the equation for \( F_{cu_{28}} \) are assessed statistically to understand the contribution of each parameter used in the experiments; Table 5 shows that all parameters are significant, with their corresponding probability less than 0.0005 and their confidence level more than 99.95%.

3.1.2 Pure quadratic model

The pure quadratic model includes linear parameters and constant as well, and it can be used when the pattern does not seem linear and the relationship between parameters seems somewhat curved. The following equation illustrates the pure quadratic model for mortar compressive strength at 28 days.

\[
F_{cu_{28}} = 48.516 - 0.660 \times CRNP + 2.691 \times CRSF + 0.322 \times SRNP - 0.001 \times CRNP^2 - 0.162 \times CRSF^2 - 0.008 \times SRNP^2
\]  

Fig.6 Normal probability of standardized residuals of multilinear regression model

Fig.7 Experimental versus predicted strengths

Table 5 displays the quadratic strength model’s statistical summary. The model’s significance (p) value is extremely low, indicating a strong data model. The model determination coefficient (\( R^2 \)) is 93.30%, it is greater than that of the MLR strength model. The pure quadratic model’s RMSE is 2.83, less than the linear strength model (3.45). The pure quadratic model suits the data in a more improved form than the multilinear model because both \( R^2 \) and RMSE are better for the former model than for the multilinear one.

Fig.8 Normal probability of standardized residuals of pure quadratic model
3.1.3 Interaction strength model

The interaction mathematical model utilizes interaction (product), and parameters in a linear form and constant terms to establish interaction terms between quantitative predictors, allowing the response relationship to differ from the values of another predictor. Interestingly, this offers an alternative way to curvature and MLR model.

The following equation shows the interaction mortar compressive strength model.

\[
F_{cu28} = 48.443 - 0.682CRNP + 1.985CRSF - 0.177SRNP - 0.038CRNP \times SRNP \]

(6)

Table 5 shows the interaction model summary. This could be established that the model's significance (p) value is incredibly low, indicating a strong data model. The model's determination coefficient (R\(^2\)) is 92.90%, slightly less than in the pure quadratic model (93.30%). The interaction strength model's RMSE is 2.87, somewhat greater than value from the pure quadratic compressive strength model which is 2.83. Both models show minor variations in their predictions. This model equips data better than the multilinear model since the determination coefficient is larger, and the RMSE is smaller than the model of interaction power.

R\(^2\) = 0.933

3.1.4 Full quadratic strength model

The suitability of the complete quadratic model has been examined to predict the strength of cement mortar with NP and SF replacements, in this model each parameter will be in a squared, interaction (product), linear and constant terms, Equation (7) shows the revealed model. H Wang and CM Cortés [29] apply the same model to find 28 days' compressive strength of mortar and pervious concrete with co-utilization of coal fly ash and waste glass powder as partial cement replacements, in which it was successfully predict the compressive strength at 28 days. Fig.12 shows that all-quadratic strength model residuals are uniformly distributed about zero. There appear to be no signs that error terms are associated with model predictions.

\[
F_{cu28} = 47.452 - 0.577CRNP + 2.907CRSF - 0.014SRNP - 0.002CRNP^2 - 0.125CRSF^2 - 0.005SRNP^2 - 0.026CRSF \times SRNP \]

(7)

RMSE of the full quadratic model is 2.58, which is the least of all strength models. The results of prediction from the full quadratic strength model are in well agreement with the experimental results as shown in Table 5 and model's determination coefficient (R\(^2\)) is 94.5% in which it the highest among all models. Fig.13 shows the scatter plot of the experimental and predicted compressive strengths versus the experimental data order. It
shows that the predicted values are very close to values revealed experiments. This model better fits the data of all the strength models discussed in this study.

Fig.13 Experimental versus predicted strengths

Table 5 Statistical summary of strength models

<table>
<thead>
<tr>
<th>Model</th>
<th>R² (%)</th>
<th>RMSE</th>
<th>F-value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>89.6</td>
<td>3.45</td>
<td>183.61</td>
<td>2.20E-31</td>
</tr>
<tr>
<td>Pure quadratic</td>
<td>93.3</td>
<td>2.83</td>
<td>141.53</td>
<td>7.01E-34</td>
</tr>
<tr>
<td>Interaction</td>
<td>92.9</td>
<td>2.87</td>
<td>206.31</td>
<td>1.91E-35</td>
</tr>
<tr>
<td>Full quadratic</td>
<td>94.5</td>
<td>2.58</td>
<td>148.02</td>
<td>2.10E-35</td>
</tr>
</tbody>
</table>

3.2 Model Validation

In model validation using cement mortars with various replacements, built models were utilized to predict 28-day strength, integrating replacement with marginally different replacement ratios. A total of 10% of the data were picked to validate each model. Figures 7, 9, 11, and 13 demonstrate that the variability between the predicted and experimental values for strength ranged from 3.5% to 8%. These differences can be attributed to differences in the properties of the constituents’ material and experimental conditions. At the same time, the differences were not important. Thus, the models can be used to predict various replacement abilities, but within the range of replacements considered for the production of the models.

4. CONCLUSION

Four predictive models for the compressive strength of mortar cubes with replacement of cement or sand with NP and SF were proposed. The models were the linear strength model, pure quadratic strength model, interaction strength model, and full quadratic strength model. The trainings were made based on the results of 68 tests. The following conclusions were obtained:

1. The multilinear regression for the compressive strength of cement mortars with sand or cement replacement with NP at 28 days is highly statistically significant, with confidence level more than 99.95%.
2. The significance (p) values of the pure quadratic model, interaction model, and full quadratic model are close to zero, indicating that these models are good models for the data.
3. R² of linear strength model, pure quadratic model, interaction strength model, and full quadratic strength model are 89.6%, 93.30%, 92.9% and 94.5%, respectively. The highest value of determination coefficients is 94.5% for the full quadratic strength model.
4. RMSE of linear strength model, pure quadratic model, interaction model, and full quadratic strength model are 3.45, 2.83, 2.87 and 2.58, respectively. The lowest value of RMSE is 2.58 for the full quadratic strength model.
5. RMSE of the interaction model is slightly greater than that of the pure quadratic model, which indicates that both interaction and pure quadratic models show no significance difference in their predictions.
6. The predicted values are close to the experimental test results for all models.
7. The high level of correlation for the prediction of cement mortar compressive strength of sand or cement replacement with NP and SF is obtained from the full quadratic strength model with less prediction error and with best data fit compared with other models.

5. REFERENCES


