

# Prediction of the Silica Fume Concrete Corrosion in Sulfuric Acid by SVM-Based Method

Yue Zhan<sup>1,a\*</sup>, Zhigang Song<sup>2,b</sup>, Haiying Wang<sup>3,c</sup>

<sup>1</sup> Faculty of Architecture Engineering, Kunming University of Science and Technology, Kunming, P.R. China, 650500

<sup>a</sup>89448891@qq.com, <sup>b</sup>zhigang-s@163.com, <sup>c</sup>Hewenhui119@sina.com

**Keywords:** support vector machine; silica fume concrete; sulfuric acid; corrosion model

**Abstract.** A long time immersion test of silica fume concrete specimens in sulfuric acid is carried out. On the basis of experimental datum, a novel prediction model for sulfuric acid corrosion of concrete based on support vector machine (SVM) is proposed. The design steps and learning algorithm are also given. Five main influencing factors including the water and cement ratio, pH value of the soaking solution, cement and silica fume quantity, specimens fluidity are analyzed with the proposed model. By comparison with the measured results, the root mean square error (RMSE) and mean absolute percentage error (MAPE) of predicted result of the model is only 0.118 and 0.096. It implies that this new method of computation appears to be a useful tool for sulfuric acid corrosion of silica fume concrete research.

## Introduction

Sulfuric acid is one of the corrosion media to concrete. Due to its complex corrosion law and multiple influencing parameters, various factors will continue to exist in interactions[1,2], lead sulfate corrosion of concrete tend to present a complex nonlinear law. Therefore, for this problem, domestic and foreign scholars have carried out lots of research, common research methods can be broadly divided into different categories: (1)The traditional mathematics statistical forecasting methods[3];(2)Experience nonlinear prediction method[4]. But both are still some drawbacks. Only when the sample data tends to infinity, traditional statistical forecasting methods, tends to real model predicted results, but the actual number of samples is often limited, it is difficult to meet this requirement. Experience nonlinear prediction method lacks a unified mathematical theory, Still need to use the experience to build prediction model, so the predicted results are often partial optimal solution, rather than global optimal solution. Support vector machines (SVM), a new method proposed based on statistical theory by Vapnik[5,6] is widely adopted in recent years, the algorithm is a good way to the Structural problems of high-dimensional model of limited amount sample, and constructed model has good predictive performance.

A long time immersion test of silica fume concrete specimens in sulfuric acid is carried out in this study. On the basis of the experimental datum, the sulfuric acid corrosion database is established, and the prediction model of sulfuric acid corroding fly ash concrete based on SVM is proposed. Five main factors influencing the reaction rate, i.e. water-cement ratio, pH value of the soaking solution, cement and silica fume quantity and mixture fluidity, are analyzed with this model. The root mean square error (RMSE) and mean absolute percentage error (MAPE) of predicted result of the model were given in this article.

## The Experiment

In this paper, the research of silica fume concrete materials by (ordinary Portland) cement, silica fume, etc. The chemical compositions of cement and silica fume are shown in Table 1.

Table 1 Chemical composition of cement and silica fume

Material	SiO <sub>2</sub>	Fe <sub>2</sub> O <sub>3</sub>	Al <sub>2</sub> O <sub>3</sub>	CaO	MgO	SO <sub>3</sub>
Cement	24.3	4.8	3.8	55.3	4.2	2.4
Silica fume	50.1	1.7	0.7	45.0	1.2	1.3

Fluidity of mixture is measured according to Ref.[7] . The specimens are prepared according to Ref. [8]. After the cast of specimens, they are put in standard curing box (constant temperature 20 °C , relative humidity 95%) for 24 hours, and then put in saturated sodium hydroxide solution. 28 days later, they are naturally dried in the room for 7 days and then cleaned to remove the surface calcium hydroxide with a metal brush. 5 surfaces of each concrete cube are waxed and only 1 surface left for soaking. After waxed, the test groups, 3 specimens for each, are put into the plastic soaking boxes with a size of 467mm × 363mm × 215mm.

The pH change of solution is monitored through a portable pH meter, model pHB-1 with an accuracy of ± 0.01pH unit. Titration sulfuric acid with concentration of 0.125mol / L or 0.5mol/L is added into the soaking solution to keep the original pH value of soaking solution. Experimental arrangement see Table 2.

Table2 Experimental arrangement

Group	Mix proportion(by mass ratio)				pH
	mw/mb	Cement	Silica fume	Sand	
1	0.53	2.472	0.697	5.072	4.6
2	0.66	2.444	0.102	5.697	4.6
3	0.47	3.146	0.429	4.565	4.4
4	0.59	2.079	0.769	5.310	4.4
5	0.43	3.204	0.703	4.439	4.2
6	0.64	2.231	0.394	5.779	4.2
7	0.51	3.228	0.066	4.958	4.0
8	0.69	1.826	0.609	5.895	4.0
9	0.56	2.700	0.300	5.099	3.8
10	0.46	2.593	1.059	4.418	3.8
11	0.58	2.665	0.232	5.279	3.6
12	0.61	2.286	0.468	5.513	3.6
13	0.41	3.852	0.246	4.294	3.4
14	0.67	2.006	0.501	5.738	3.4
15	0.50	2.554	0.806	4.797	3.2
16	0.63	2.640	0.027	5.733	3.2
17	0.44	3.284	0.535	4.547	3.0
18	0.55	2.138	0.916	4.952	3.0
19	0.70	2.088	0.312	6.013	2.8
20	0.49	3.120	0.309	4.756	2.8
21	0.62	2.086	0.623	5.583	2.6
22	0.54	2.956	0.156	5.215	2.6
23	0.42	2.960	1.040	4.315	2.4
24	0.57	2.387	0.560	5.160	2.4
25	0.68	2.298	0.173	5.859	2.2
26	0.52	2.714	0.517	5.002	2.2
27	0.45	3.621	0.112	4.413	2.0
28	0.65	1.861	0.724	5.545	2.0
29	0.48	2.765	0.735	4.626	1.8
30	0.60	2.492	0.308	5.444	1.8

The apparent diffusion coefficient value  $D_{OH^-}$  of concrete is analyzed through an analytical expression by Song et al[9]

$$V_{H^+}(t) = a / \sqrt{1 - \exp(-bt)} \quad (1)$$

In the formula,

$$\begin{cases} a = SC_{OH.S} \sqrt{D_{OH^-} KNPC_{H.S}} \\ b = 2KPC_{H.S} / M \\ M = \int_0^1 \phi(\eta) d\eta \\ N = -\phi'(0) \\ P = \int_0^1 \phi(\eta) \psi(\eta) d\eta \end{cases} \quad (2)$$

$\phi(\eta)$  is the shape function of hydroxyl ion concentration distribution in the boundary layer,  $\psi(\eta)$  is the shape function of hydrogen ion concentration distribution in boundary layer, K is a chemical reaction constant, S is the immersion area (0.0768m<sup>2</sup>),  $C_{OH.S}$  is the OH<sup>-</sup> concentration in saturated solution of Ca(OH)<sub>2</sub> is 0.045mol/L. So we can obtain the form of:  $D_{OH^-}$

$$D_{OH^-} = 3a^2 / bS^2 C_{OH.S}^2 \quad (3)$$

Suppose  $\phi(\eta)$  and  $\psi(\eta)$  meet the following form:

$$\phi(\eta) = (1 - \eta)^2, \psi(\eta) = \eta^2 \quad (4)$$

So we can get  $M=1/3, N=2$ , therefore, also obtain the value of  $D_{OH^-}$ . See table 3.

Table 3 Experiment result

Group	$\ln(D_{OH} \times 10^{-10})$	Group	$\ln(D_{OH} \times 10^{-10})$	Group	$\ln(D_{OH} \times 10^{-10})$
1	2.54	11	1.80	21	3.57
2	2.47	12	1.28	22	3.63
3	2.83	13	2.43	23	5.66
4	3.06	14	1.17	24	4.43
5	1.33	15	2.90	25	8.12
6	2.26	16	1.50	26	6.04
7	2.53	17	3.54	27	8.36
8	2.64	18	2.25	28	6.69
9	2.31	19	2.87	29	9.75
10	2.33	20	4.00	30	8.53

## Support Vector Machine

SVM is based on statistical learning theory developed by Vapnik. It is effective to improve generalization performance and can obtain globally optimal and unique solution. It is generally used for pattern recognition and regression estimation. In SVM regression, the basic objective is to minimize an upper bound of the generalization error based on the structural risk minimization (SRM). The detailed description of SVM can be found in the publications by Vapnik.

The main idea of SVM regression is to map the input data  $x$  into a high-dimensional feature space by a nonlinear mapping and to do linear regression in this space. The regression model is defined as  $y = f(x) + e$ , where  $x$  and  $y$  are input and output function, respectively, and defined in the high-dimensional feature space;  $e$  is the independently random error.

Given a dataset  $G = \{(x_i, y_i), i=1, 2 \dots N\}$ , where  $N$  is the size of the training data set,  $x_i$  is the value of the input and  $y_i$  is the value of output. The optimum regression function can be assumed as follows:

$$f(x) = w\phi(x) + b. \quad (5)$$

Where  $\phi(x)$  represents the high-dimensional feature spaces which are nonlinearly mapped from the input space  $x$ ,  $w$  is the weight vector and  $b$  is the scalar threshold. Introducing the  $\varepsilon$ -insensitive loss function

$$L_\varepsilon(x) = \begin{cases} 0 & \text{for } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & \text{otherwise} \end{cases}. \quad (6)$$

The Eq. 1 can be transformed to the dual optimization problem:

$$\max_{a, a^*} W(a, a^*) = \max_{a, a^*} \begin{cases} \sum_{i=1}^N a_i^* (x_i - \varepsilon) - a_i (x_i + \varepsilon) \\ -\frac{1}{2} \sum_{i=1, j=1}^N (a_i^* - a_i)(a_j^* - a_j) f(x_i, x_j) - \varepsilon \sum_{i=1}^N (a_i^* - a_i) + \sum_{i=1}^N y_i (a_i^* - a_i) \end{cases}. \quad (7)$$

Where  $0 \leq \alpha_i \leq C$ ,  $0 \leq \alpha_i^* \leq C$  and  $\sum_{i=1}^N (\alpha_i^* - \alpha_i) = 0$ . The Eq.3 can be solved by Lagrange multipliers, and its solution is given by:

$$f(x) = \sum_{i=1}^N (a_i - a_i^*) f(x_i, x_j) + b \quad (8)$$

The kernel function  $K(x_i, x_j)$  has been introduced to avoid computing explicitly the map  $\phi(x)$ . Thus, Eq. 4 can be written as:

$$f(x) = \sum_{i=1}^N (a_i - a_i^*) K(x_i, x_j) + b \quad (9)$$

## SVM for prediction of sulfuric acid corroding concrete

This study attempts to use SVM for the prediction of sulfuric acid corroding concrete. The experimental datum and results (see Table 2 and 3) are used in this study.

First, choosing 1 out of 24 groups for training data set, the rest are considered as testing data set. All the input data (water-cement ratio, pH, cement quantity, silica fume quantity, mixture fluidity) and output data ( $D_{OH}$ ) are normalized within the program (step 1 and 2). Then training patterns are selected and transformed into feature space using kernel function (step 3). By using the SVM, the predicted results are obtained, and then validated using the testing set (step 4). The predicted results are compared with the experimental results (step 5).

A problem regarding the choice of two parameters ( $C$  and  $\sigma$ ) for SVM was studied by Cherkassky[10]. The parameter  $C$  controls the smoothness or flatness of the approximate function. The Gaussian radial basis kernel[11] was chosen for the SVM because it is the most common and successfully applied kernel.  $\sigma$  is the parameter that determines the spread of the Gaussian function in the Gaussian radial basis kernel.  $K$ -fold cross validation is used in this study to determine the best values of the training parameters  $C$  and  $\sigma$ . The values of  $C$  and  $\sigma$  are 17 and 2.93.

For convenient comparison purpose, the experimental and predicted results are shown in table4

Table 4 Prediction result

Group	2	9	14	15	21	28
Experiment value	2.47	2.31	1.17	2.90	3.57	6.69
Predicted value	2.58	2.66	1.43	2.72	3.69	7.14
RMSE	0.118		MASE		0.096	

Seen from table 4, The root mean square error (RMSE) and mean absolute percentage error (MAPE) are used as the main criteria to evaluate the performance of the SVM model. There are only 0.118 and 0.096.

### Conclusions

This study has investigated the use of SVM for prediction of sulfuric acid corroding concrete. The performance of the proposed method is verified by comparing the predicted results with the experimental results. RMSE and MASE of the predicted results by SVM method are 0.188 and 0.096. It implies that SVM can be regarded as a very effective method for prediction of sulfuric acid corroding concrete. Selecting the parameters of the kernel function in SVM, has a great influence on the accuracy of the prediction results [12]. Consequently, how to get the optimal kernel parameters quickly, it remains to be further studied.

### Acknowledgements

The study of this paper is supported by the Natural Science Foundation of China and Yunnan Province (Grant No. 51078175, 2008E038M).

### References

- [1] B.E. Robin, W.D. Horst. Modelling Acid Attack on Concrete (Part I): The Essential Mechanisms. *Cement and Concrete Research*, 2005, 35(12): 2333-2339
- [2] E. Hewayde, M. Nehdi, E.N. Allouche. Experimental Investigations of the Effect of Selected Admixtures on the Resistance of Concrete to Sulfuric Acid Attack // *Proceeding of International on Pipeline Engineering and Construction*. Baltimore; ASCE, 2003:504-513.
- [3] CHEN Wei, WANG Zi-ting, The Relation Between the Traditional Regularization and Bayesian Method. *Journal of Hebei Polytechnic University (Natural Science Edition)*. 2011, 08(3), 1674-0262. (In Chinese)
- [4] WU Fangjun, Understanding Knowledge sharing activities in software fault-prone prediction: a transfer learning study. *Journal of Chinese Computer Systems*, 2014, 11(11), 11-2416-06 (In Chinese).
- [5] VAPNIK V N. The nature of statistical learning theory [M]. Berlin: Springer, 1995.
- [6] VAPNIK V N. An overview of statistical learning theory [J]. *IEEE Transaction Neural Networks*, 1999, 10(5):988-999
- [7] JGJ55-2011, Standard for Test Method of Performance on Ordinary Fresh Concrete [S]. (In Chinese)
- [8] GB/T 50080-2002, Specification for Mix Proportion Design of Ordinary Concrete [S]. (In Chinese)
- [9] Song Zhigang, Zhang Xuesong. Experimental Study of Mortar under Corrosion in Sulfuric Acid [J]. *Journal of Building Materials*, 2012(15):163-167. (In Chinese)
- [10] V. Cherkassky, Y. Ma. Practical Selection of SVM Parameters and Noise Estimation for SVM Regression. *Neural Network*, 2004(17):113-126.
- [11] Keerthi S S, Lin C J. Asymptotic behaviors of support vector machines with Gaussian kernel [J]. *Neural computation*, 2003, 15(7): 1667-1689.
- [12] Zhang W, Song Z. Prediction of Concrete Corrosion in Sulfuric Acid by SVM-Based Method [J]. *Electronic Mechanical Engineering and Information Technology*, Atlantis Press, 2012, (23): 400-404. (In Chinese)