International Journal of Engineering & Technology, 3 (2) (2014) 224-229
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v3i2.2011
Research Paper

Developing a forecasting model of concrete compressive strength using relevance vector machines

Mohammad Thalji Awwad

Civil Engineering Department, Al-Balqa Applied University, Amman University
**Corresponding author E-mail: thalji406@hotmail.com*

Abstract

We analyze results of two experiments that tested effect of adding Silica on the compressive strength of concrete at early stage and after long period. The two experiments evaluated different silica/cement ratios for different mixing periods. Adding Silica to concrete mix produce high early strength material which is highly desirable in airports and highways.
More than 90 samples of different silica/cement ratios are tested for compressive strength at 3 and 28 days. Test results showed high early up to 60 MPa. Strength increase is proportional with the increase of silica/cement ratio and mixing time with maximum at ratio of 15/100 and 30 minutes mixing time.
A relevance Vector Machine (RVM) model is developed to predict concrete compressive strength using concrete mixture inputs information. RVM model predictions matched experimental data closely. The developed model can be used to predict compressive strength in future periods based on initial information related to cement mixture.

Keywords: *Relevance Vector Machine, Silicate Percent, Prediction Model, Milling Time, Compressive Strength, Concrete.*

1. Introduction

Today, there is an increasing expansion in infrastructure development worldwide. This expansion creates demand for high performance concrete (HPC) with high durability. Virgin Silica is an additive proven to enhance concrete durability especially compressibility. Virigin Silica is a natural resource available in abundance in many regions in the world and it is considered attractive option for its environmental safety and economic efficiency [6], [7]. There are a limited number of experimental studies that evaluated the impact of virgin silica on durability. Since, experimental studies take long periods and require expensive costs alternative approaches are highly desirable. Therefore, developing statistical leaning models is desirable option because it evades long and expensive experimental studies.
While some researchers conducted experimental studies, there is a need for further research to understand the impact of silica on concrete mixture.
Data mining using statistical learning methods can be valuable tool in various fields [10]. Statistical learning methods can use observations to build a structure of the relations among variables and can be used to predict system response in a given set of conditions [9].

2. Experimental data

The data used in this study are summarized in Tables 1 and 2. The data set contains information about Cement Ratio (CR), Silicate Ratio (SR), Milling Time (MT), Sample Age (SA), and Compressive Strength (CS). Table 1 summarizes experimental data for samples with 3 days age.
Cement percent, silicate percent, Milling time, and sample age are used as input parameters to RVM model. The output is a prediction of compressive strength. This study uses the experimental work by Suliman and Awwad, 2000 and 2005 [6], [7].

Table 1: Data Set Used in This Study for Test after 3 Days. Data Source is Suliman and Awwad (2000, and 2005).

Cement %	SiO2 %	Milling Time (min)	Sample Age (days)	Compression (MPa)
100	0	0	3	210
100	0	0	3	209
100	0	0	3	211
85	15	0	3	180
85	15	0	3	179
85	15	0	3	181
70	30	0	3	170
70	30	0	3	169
70	30	0	3	171
100	0	30	3	240
100	0	30	3	239
100	0	30	3	241
85	15	30	3	325
85	15	30	3	324
85	15	30	3	326
70	30	30	3	260
70	30	30	3	259
70	30	30	3	261
100	0	60	3	255
100	0	60	3	254
100	0	60	3	256
85	15	60	3	340
85	15	60	3	339
85	15	60	3	341
70	30	60	3	325
70	30	60	3	324
70	30	60	3	326
100	0	90	3	265
100	0	90	3	264
100	0	90	3	266
85	15	90	3	340
85	15	90	3	339
85	15	90	3	341
70	30	90	3	325
70	30	90	3	324
70	30	90	3	326
100	0	120	3	220
100	0	120	3	219
100	0	120	3	221
85	15	120	3	345
85	15	120	3	344
85	15	120	3	346
70	30	120	3	325
70	30	120	3	324
70	30	120	3	326

Compressive strength data for 28 days are also used in the RVM model and summarized in Table 2. Data Source is Suliman and Awwad (2000, and 2005).

3. RVM model overview

Statistical learning models have been successfully implemented in various applications. Relevance Vector Machines (RVM) is a sparse method for training generalized linear models [8]. It can be seen as probabilistic version of support vector machines (SVM). Bayesian regression models such as Support SVM and RVM view data as a chaotic system in which data are assumed to provide enough information about the behavior of the system to perform forecasting [1]. Successful applications of statistical learning algorithms have been reported for various engineering fields [3, 5].
RVM is a sparse method for training generalized linear models [9]. RVM have the capacity to consider uncertainty in data and model parameters [2]. RVM simplifies complex systems by producing "structured" models; therefore parameterization process that fits the information content. The key advantage of RVM is the generalization ability and

the sparse formulation of the resulting model that utilizes few kernel functions. RVM fits naturally into a regression framework and yield full probability distributions of the output.

Table 2: Data Set Used in This Study for Test After 28 Days.

Cement %	SiO2 %	Milling Time (min)	Sample Age (days)	Compression (MPa)
100	0	0	28	410
100	0	0	28	409
100	0	0	28	411
85	15	0	28	320
85	15	0	28	319
85	15	0	28	321
70	30	0	28	310
70	30	0	28	309
70	30	0	28	311
100	0	30	28	460
100	0	30	28	459
100	0	30	28	461
85	15	30	28	580
85	15	30	28	579
85	15	30	28	581
70	30	30	28	520
70	30	30	28	519
70	30	30	28	521
100	0	60	28	510
100	0	60	28	509
100	0	60	28	511
85	15	60	28	640
85	15	60	28	639
85	15	60	28	641
70	30	60	28	580
70	30	60	28	579
70	30	60	28	581
100	0	90	28	520
100	0	90	28	519
100	0	90	28	521
85	15	90	28	660
85	15	90	28	659
85	15	90	28	661
70	30	90	28	590
70	30	90	28	589
70	30	90	28	591
100	0	120	28	550
100	0	120	28	549
100	0	120	28	551
85	15	120	28	660
85	15	120	28	659
85	15	120	28	661
70	30	120	28	600
70	30	120	28	599
70	30	120	28	601

RVM uses the following equation for the prediction of output (y).

$$y = a'\varphi(x) = \sum_{i=1}^{n} a_i K\left(x_i\right) + a_0 \tag{1}$$

Where:
x, x_i = input variables
K(x, x_i) = kernel function
n = number of data
a_0 = weight.
In this study, inputs are cement percent (CR), milling time (MT), age of sample (SA), Silicate percent (SR), and compression strength (CS) is the output.

$$x = [CR, SR, MT, SA] \quad (2)$$

And

$$y = [CS] \quad (3)$$

The likelihood of the complete data set can be written as

$$p(y/a,\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\|y - a\phi\|^2\right\} \quad (4)$$

The posterior distribution over the weights is thus given by:

$$p\left(a/y,\alpha,\sigma^2\right) = \frac{p(y/a,\sigma^2)p(a/\alpha)}{p(y/\alpha,\sigma^2)} = (2\pi)^{-(n+1)/2}\left|\sum\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(a-\mu)'\sum^{-1}(a-\mu)\right] \quad (5)$$

Where the posterior covariance and mean are respectively:

$$\sum = \left(\sigma^{-2}\varphi'\varphi + A\right)^{-1}$$
$$\mu = \sigma^{-2}\sum\varphi' y \quad (6)$$

For uniform hyperpriors over α and σ2, one needs only to maximize the term:

$$p\left(y/\alpha,\sigma^2\right): \quad (7)$$

$$p(y/\alpha,\sigma^2) = \int p\left(y/a,\sigma 2\right)p(a/\alpha)$$
$$= (2\pi)^{-n/2}\left|\sigma^2 I + \varphi A^{-1}\varphi'\right|^{-1/2} \times \exp\left[-\frac{1}{2}y'\left(\sigma^2 I + \varphi A^{-1}\varphi'\right)^{-1} y\right] \quad (8)$$

Maximization of this quantity is known as the type II maximum likelihood method (Berger, 1985) or "evidence for hyper parameter" [4]. Hyper parameter estimation is carried out in iterative formulae, e.g., gradient descent on the objective function [9]. The outcome of this optimization is that many elements of this go to infinity such that we will have only a few nonzero weights that will be considered as relevant vectors.

4. Results

Development of the RVM model included training and testing of the model. Model training and parameters optimization was performed on 75% of the data.

In RVM, the trial and error approach has been adopted for determining the design value of σ. The developed RVM gives best performance at σ = 0.0643. The calibrated model performance is shown visually in Figure 1on the trained

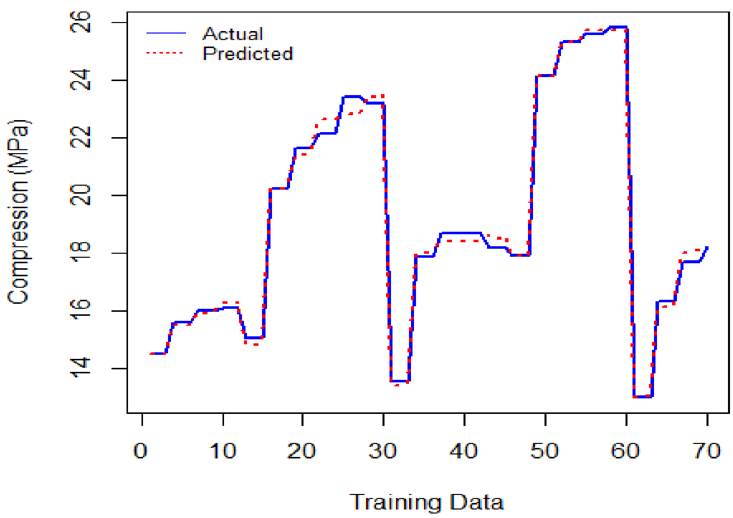

Fig. 1: Experimental and Predicted Compressive Strength Using RVM Model.

After training, testing was conducted on 25% of the data. The model predictions are shown to match data very closely with correlation (R^2) of 0.9995 and residual standard error of 0.0734 for testing data set which was not used in training phase. As shown in Figure 1, the trained model proved to be capable of predicting compressive strength over the entire experiment period with insignificant error.

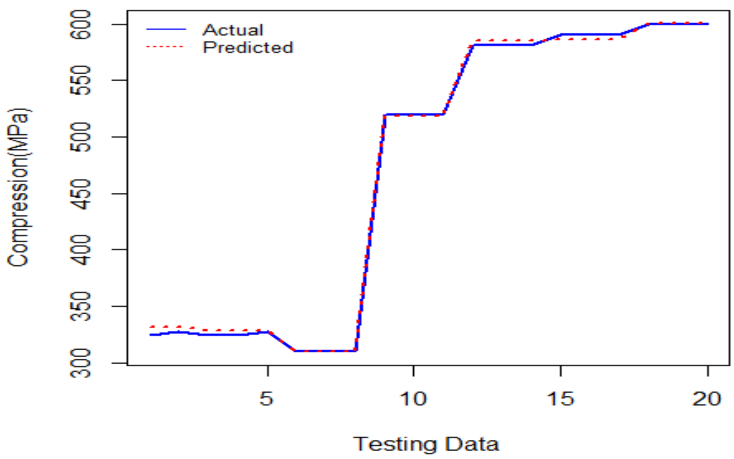

Fig. 2: Actual and Predicted Compressive Strength (MPA) for Testing Data Set.

The residuals distribution which is the distribution of error in model prediction (Predicted-Actual) are clustered within bands for normal distribution for testing and training data sets. Spread of error for testing data is larger than the spread of training data; however it is still within the normal distribution bounds. Therefore, the RVM model is capable of predicting compressive strength (CS) adequately with no unexplained trends in data.

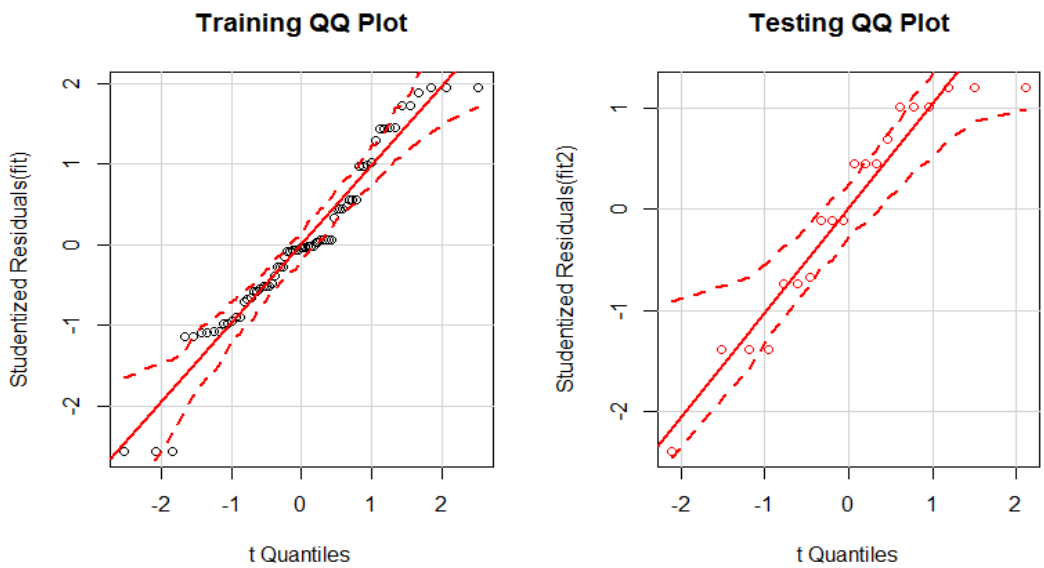

Fig. 3: Error Plots and Residuals Distribution Compressive Strength for Training and Testing Data Sets.

5. Conclusion

This study developed RVM model for the prediction of compressive strength of concrete. 90 data points have been utilized to train, optimize and test the prediction model. The developed model has practical application to forecast changes in CS with time up to 28 days based on input data which is unique to this study. The model can be an asset in concrete mix studies especially that it can reduce the extent of expensive laboratory studies to measure CS for a different combinations of virgin silica and additives in general. The model can be used to optimize Virgin Silica percent without additional laboratory experiment with adequate accuracy.

Generally, RVM produce sparse solutions and it can be used to simulate different applications in concrete studies. RVM are suitable models for concrete applications since they can save tremendous experimental efforts and costs once they are developed using appropriate experimental data sets.

References

[1] Khalil, A., McKee, M., Kemblowski, M., and Asefa, T., 2005. “Sparse Bayesian learning machine for real-time management of reservoir releases.” Water Resources Research. Vol. 41 (11), W11401.

[2] Khalil, A., McKee, M., Kemblowski, M., Asefa, T., and Bastidas, L., 2006. “Multiobjective analysis of chaotic dynamic systems with sparse learning machines.” Advances of Water Resources. Vol. 29(1), pp.72–88.

[3] Matić D., F. Kulić, M. Pineda-Sánchez, and I. Kamenko, “Support vector machine classifier for diagnosis in electrical machines: Application to broken bar,” Expert Syst. Appl., vol. 39, no. 10, pp. 8681-8689, 2012.
[4] MacKay D. J., “Bayesian methods for adaptive models,” Ph.D. Thesis, California Institute of Technology, 1992. Online at: http://resolver.caltech.edu/CaltechETD:etd-01042007-131447.
[5] Suetani H., Ideta A. M., and J. Morimoto, “Nonlinear structure of escape-times to falls for a passive dynamic walker on an irregular slope: Anomaly detection using multi-class support vector machine and latent state extraction by canonical correlation analysis,” Proceeding IEEE Int. Conference on Intelligent Robots and Sys., pp. 2715-2722, 2011.
[6] Suliman M.R. and M. Awwad, 2000. Utilizing of Silica in Early –High Strength Concrete. Cement and Concrete Technology in the 2000s, Second International Symposium, 6-10 September, 2000, Istanbul, Turkey.
[7] Suliman K. M.R. And M. T. Awwad, 2005. Virgin Silica improves the durability of Portland cement concrete. Journal of Engineering Scuence, Assiut University. Vol. 33, No. 1, pp 1-9. January.
[8] Tipping M. E., “The relevance vector machine,” *Advances in Neural Information Processing Systems*, Vol. 12, pp. 652-658, 2000.
[9] Tipping M. E., “Sparse Bayesian learning and the relevance vector machine,” J. of Machine Learning Research, vol. 1, pp. 211-244, 2001.
[10] Vapnik V. N., the Nature of Statistical Learning Theory, New York: Springer-Verlag, 1995.