Kernel-based models for prediction of cement compressive strength

Article *in* Neural Computing and Applications · December 2017

DOI: 10.1007/s00521-016-2419-0

3 authors:

Mohit Verma
CSIR Structural Engineering Research Centre
45 PUBLICATIONS **480** CITATIONS
SEE PROFILE

Thirumalaiselvi A.
CSIR Structural Engineering Research Centre
18 PUBLICATIONS **140** CITATIONS
SEE PROFILE

J. Rajasankar
CSIR Structural Engineering Research Centre
76 PUBLICATIONS **598** CITATIONS
SEE PROFILE

Neural Comput & Applic
DOI 10.1007/s00521-016-2419-0

ORIGINAL ARTICLE

Kernel-based models for prediction of cement compressive strength

Mohit Verma[1,2] · **A. Thirumalaiselvi**[1,2] · **J. Rajasankar**[1,2]

Received: 14 December 2015 / Accepted: 8 June 2016

Abstract This paper employs three different kernel-based models—support vector regression (SVR), relevance vector machine (RVM) and Gaussian process regression (GPR)—for the prediction of cement compressive strength. The input variables for the model are taken as C_3S (%), SO_3 (%), Alkali (%) and Blaine (cm^2/g), while the output is 28-day cement compressive strength (N/mm^2) of the cement. The hyperparameters of the SVR are obtained using two different metaheuristic optimization algorithms—particle swarm optimization (PSO) and symbiotic organism search (SOS). Trial-and-error-based approach is used for arriving at the hyperparameters of RVM and GPR. The compressive strength predicted using different kernel-based models is also compared with that obtained from ANN and fuzzy logic models reported in the literature. The performance of the different kernel-based models is benchmarked using six different error indices and residual analysis. The performance of the kernel-based models is found to be at par with ANN. The better generalization capability and excellent empirical performance of the kernel-based models overcome the disadvantages associated with ANN and provide a good tool for the prediction of the cement compressive strength.

Keywords Support vector regression · Relevance vector machine · Gaussian process regression · Cement compressive strength · Particle swarm optimization · Symbiotic organism search

✉ Mohit Verma
mohitverma@serc.res.in

[1] CSIR – Structural Engineering Research Centre, Chennai, India

[2] Academy of Scientific and Innovative Research, Chennai, India

1 Introduction

Compressive strength is one of the most basic properties of the cement. The strength of the cement is affected by various physical and chemical parameters. The physical parameters which affect the compressive strength of the cement are fineness, particle size distribution and density, while the chemical parameters include phase composition (C_3S, C_2S, C_3A, C_4AF), SO_3 and MgO. The effect of the particle size distribution on the compressive strength of the cement was studied by Osbaeck and Johansen [24]. Celik [5] and Chen and Kwan [7] studied the strength of the cement with respect to the surface area and the particle size distribution. The Alkali effect was studied by Osbaeck [25]. Mechling et al. [21] studied the relationship between the chemical composition and the compressive strength of the cement. Each of these studies was carried out concentrating on either one or more parameters affecting the compressive strength of the cement. Various studies have been carried out to predict the compressive strength of the cement. The studies involved the use of mathematical modeling and regression analysis [9, 29, 31]. Some researchers explored artificial neural network (ANN) models to predict the compressive strength [22, 23, 28]. A model based on fuzzy logic was developed by Fa-Liang [11] for prediction of cement compressive strength. Extensive experimental investigations have been conducted by Akkurt et al. [1, 2] to develop a reliable fuzzy logic and ANN-based models.

Recently, many researchers have worked toward the development of kernel-based models in different areas of

Journal : Large 521 | Dispatch : 15-6-2016 | Pages : 18
Article No. : 2419 | LE | TYPESET
MS Code : NCAA-D-15-01328 | CP | DISK

engineering—health monitoring [36], concrete research [6, 18, 22, 23], hydrology [20], remote sensing [4, 17] and fracture mechanics [35]. In this paper, kernel-based models which are widely used—support vector regression (SVR) [32], relevance vector machine (RVM) [30] and Gaussian process regression (GPR) [27, 34]—are selected to predict cement compressive strength. The hyperparameters of SVR are obtained by using two different metaheuristic optimization algorithms—particle swarm optimization (PSO) [16] and symbiotic organism search (SOS) [8] algorithms. The hyperparameters of RVM and GPR are obtained using trial-and-error approach since they have reduced sensitivity to hyperparameters. The input variables for the model are taken as C_3S (%), SO_3 (%), Alkali (%) and Blaine (cm^2/g) based on the research carried out by Akkurt et al. [1]. The output of the model is 28-day compressive strength (N/mm^2) of the cement. The performance of the models is benchmarked based on six different error indices and residual analysis. The training and testing of the model are carried out based on the experimental results published in the literature [1]. The compressive strength predicted using different kernel-based models is also compared with that of the ANN and fuzzy logic models. The kernel-based models can be used as a predictive tool at the manufacturing stage of cement to design the compressive strength by adjusting the composition of C_3S, SO_3, Alkali and Blaine.

2 Kernel-based models

In this section, we give brief description of the different kernel-based models used in the present study for the prediction of the cement compressive strength.

2.1 Support vector regression with optimized hyperparameters

Consider the case of approximating the set of data represented by $\{(x_i, y_i), i = 1, \ldots, N\}$, $X \in \mathbb{R}^n, Y \in \mathbb{R}$. A formulation based on SVR was developed by Vapnik [32, 33] to estimate the output (y) for the known values of the input parameters (x). The problem of nonlinear SVR is expressed as:

$$y = f(x) = \langle w, \phi(x) \rangle + b \tag{1}$$

where $\langle . \rangle$ denotes the inner product, w and b are the parameters of the function, and $\phi(x)$ is the function which maps the inputs to the higher-dimensional feature space.

The distance measure between predicted and actual values is included in SVR using a loss function to penalize for the error. The loss function used in the present study is ε-insensitive which is mathematically defined as:

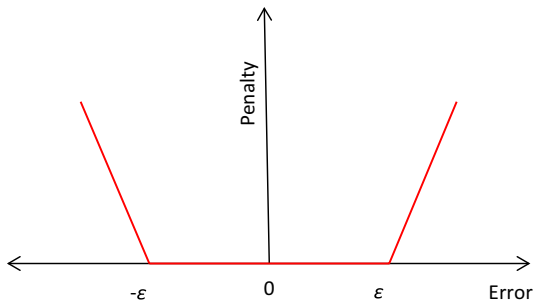

Fig. 1 ε-Insensitive loss function

$$L_1^{\varepsilon}(f(x), y) = \begin{cases} 0 & \text{for} \quad |f(x) - y| < \varepsilon \\ f(x) - y & \text{otherwise} \end{cases} \tag{2}$$

The errors between $-\varepsilon$ and ε are ignored (Fig. 1).

SVR involves determination of the parameters w and b such that the risk is minimized with respect to the ε-insensitive loss function. The objective of using SVR is to minimize the following expression by varying C:

$$\Omega = \frac{1}{2}||w||_2^2 + C\sum_{n=1}^{N} L_1^{\varepsilon}(y_n, \langle w, \phi(x_n) \rangle + b) \tag{3}$$

This is formulated as a quadratic minimization problem [32]:

$$\begin{aligned} &\min_{w \in \mathbb{R}^n, b \in \mathbb{R}} \quad ||w||_2^2 + C\sum_{n=1}^{N}(\xi_n^2 + \breve{\xi}_n^2) \\ &\text{subject to } (w, \phi(x_n) + b) - y_n \le \varepsilon + \breve{\xi}_n, \\ &\quad y_n - (w, \phi(x_n) + b) \le \varepsilon + \xi_n, \\ &\quad \xi_n, \breve{\xi}_n \ge 0 \quad n = 1, \ldots, N \end{aligned} \tag{4}$$

where the two slack variables ξ_n and $\breve{\xi}_n$ represent the upper and lower constraints on the output of the system. The above optimization problem can be solved by the method of Lagrange multipliers and its solution is given by:

$$\begin{aligned} y = &\sum_{n=1}^{l}(\alpha_n - \breve{\alpha}_n)\langle \phi(x_i), \phi(x) \rangle \\ &- \frac{1}{2}\sum_{n=1}^{l}(\alpha_n - \breve{\alpha}_n)[\langle \phi(x_r), \phi(x_n) \rangle + \langle \phi(x_s), \phi(x_n) \rangle] \end{aligned} \tag{5}$$

where α and $\breve{\alpha}$ are the Lagrange multipliers, l is the number of support vectors, and x_r and x_s are two typical support vectors.

The inner product $\langle \phi(x_m), \phi(x_n) \rangle$ involves complex computations. Based on the Mercer's condition [32], the inner product is replaced by simple arithmetic computation in input space using a kernel function. In the present study, a radial basis function is chosen to be the kernel function which is defined by the following equation:

Journal : Large 521 | Dispatch : 15-6-2016 | Pages : 18
Article No. : 2419 | LE | TYPESET
MS Code : NCAA-D-15-01328 | CP | DISK

$$K(x, x_n) = \exp\left\{-\frac{(x - x_n)^T (x - x_n)}{2\gamma^2}\right\} \quad (6)$$

The nonlinear SVR function is given by [32]:

$$f(x) = \sum_{n=1}^{l} (\alpha_n - \breve{\alpha}_n) K(x, x_n) - \frac{1}{2} \sum_{n=1}^{l} (\alpha_n - \breve{\alpha}_n)[K(x_r, x_n) + K(x_s, x_n)] \quad (7)$$

where l is the number of support vectors; x_r and x_s are two typical support vectors.

The model is general in its formulation, as no problem-specific parameter is included in it. Its performance depends purely on the choice of the values of the hyperparameters like C, ε and γ which in turn depend on the characteristics of problem-specific data. Owolabi et al. [26] suggest that the optimization of the parameter ε can be avoided as it does not have a profound effect on the performance of the model. On the other hand, there are a few research works which optimize all the three hyperparameters of SVR—C, γ and ε [3, 12, 14, 15, 19]. In the presence of two different opinions reported in the literature, the optimization of all the three hyperparameters of SVR has been taken up in the present study. The hyperparameters of SVR in the present study are determined using two different metaheuristic algorithms—symbiotic organism search (SOS) and particle swarm optimization (PSO) algorithm.

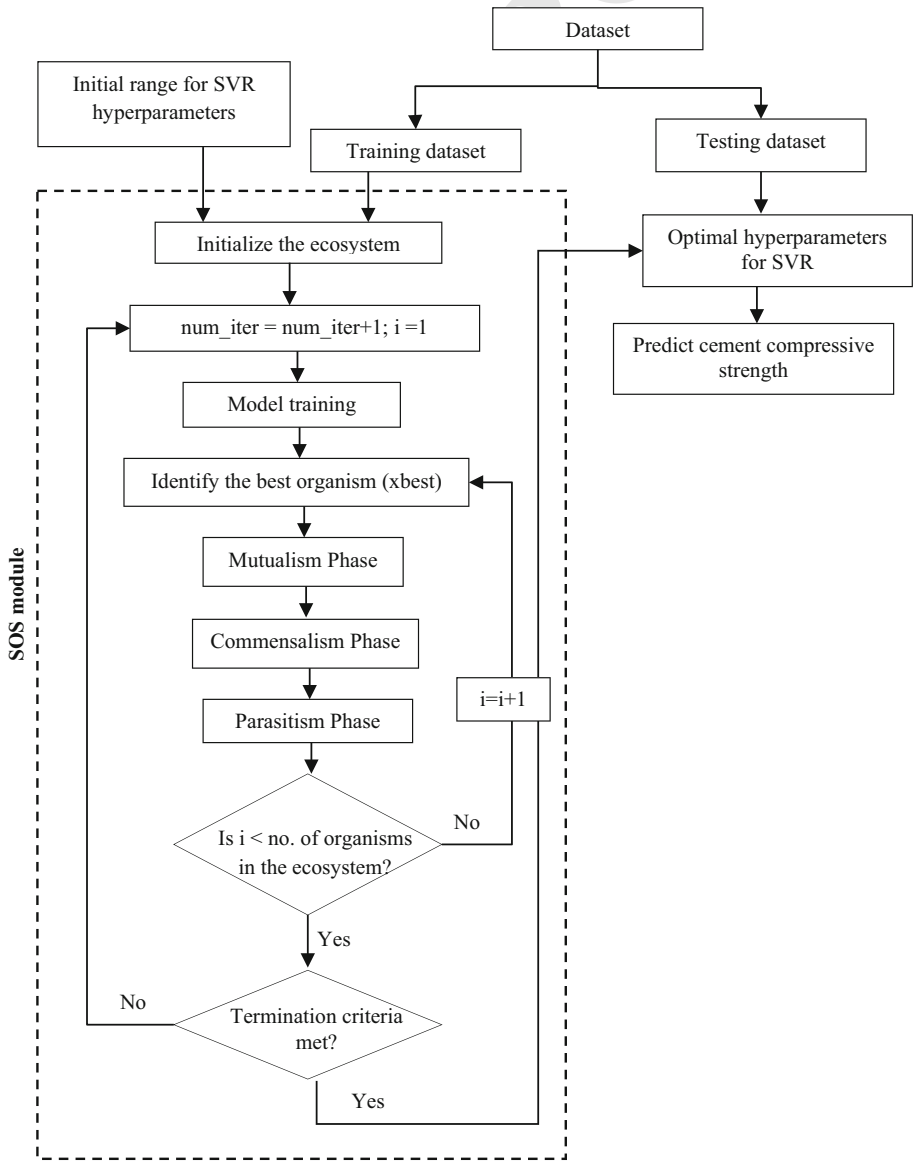

Fig. 2 Flowchart for optimization of the hyperparameters of SVR using SOS

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

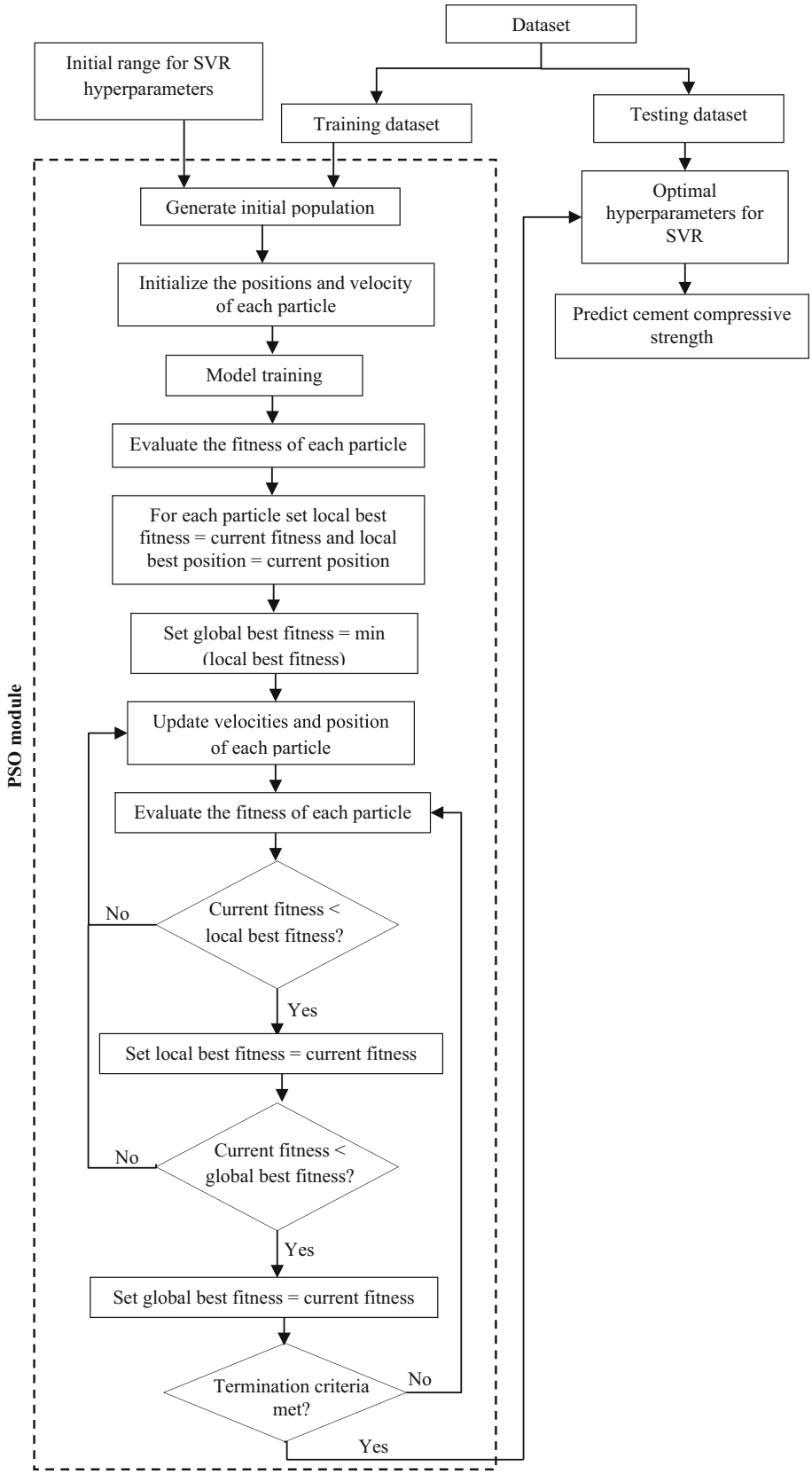

Fig. 3 Flowchart for optimization of the hyperparameters of SVR using PSO

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

Algorithm for optimization of hyperparameters of SVR

Step 1: Divide data into two parts—training and testing dataset.

Step 2: Set the initial values and the bounds for the parameter C, γ and ε.

Step 3: Optimization (PSO or SOS) is carried out using C, γ and ε as the design variables. Objective is to minimize error norm, which is the relative difference between the actual values (Y) and the predicted values (Yp) obtained from trained SVR for the training dataset

$$\text{Objective function} = \text{minimize } (\text{norm}((Yp - Y)/Y)).$$

Step 4: Calculate the optimized values for C, γ and ε. Using these values, obtain optimized SVR model.

Step 5: Predict the cement compressive strength for the testing dataset using optimized SVR model.

2.1.1 Symbiotic organism search (SOS)

SOS algorithm proposed by Cheng and Prayogo [8] is based on the concept of symbiosis. SOS involves the initialization of the population to form an ecosystem. Each organism in the ecosystem is a candidate solution for optimization. The organisms in the ecosystem are randomly generated within the upper and lowers bounds in the search space. The fitness value associated with each of the organism represents its degree of adaptation to the objective function. The best solution (Xbest) is selected as the organism with the best fitness value in the ecosystem. New organisms are generated in each iteration based on three different phases—mutualism, commensalism and parasitism. Mutualism represents a symbiotic relationship in which both the organisms are benefitted. In commensalism, one of the organisms is benefitted, while the other remains unaffected. Parasitism phase represents the relationship in which one of the organisms is benefitted, while the other is harmed. Each organism randomly interacts with the other organisms in the ecosystem through all the phases. This forms the inner loop of the algorithm. The process is continued till the termination criteria are achieved. This forms the outer loop for the algorithm. The main advantage of SOS over other metaheuristics algorithms is that it does not require any other specific algorithm parameters. The flowchart for the optimization of hyperparameters using SOS is shown in Fig. 2.

2.1.2 Particle swarm optimization (PSO)

PSO is a robust optimization technique inspired by the social behavior of the organisms that live and interact within large groups [16]. PSO constitutes a swarm in which each particle moves in the search space looking for the best solution. Each particle is treated as a point in a N-dimensional space whose movement is guided by its own best known position (local best) as well as the best known position of the entire swarm (global best). The basic concept of PSO is to accelerate each particle toward its local and global best locations using some random weighted acceleration for each time step. The objective function for the optimization was to minimize the difference between the experimental and the predicted compressive strength. The flowchart for the optimization of hyperparameters using PSO is shown in Fig. 3.

2.2 Relevance vector machine (RVM)

RVM proposed by Tipping [30] is based on Bayesian formulation of the generalized linear model of identical functional form to the SVM. The sparse representation in RVM is achieved by assuming a sparse distribution on the weights in the regression model whose most probable values are iteratively estimated from the data. The advantage of RVM is its ability to generalize and provide inferences at low computational cost. The outputs are predicted based on the following equation:

$$y(x) = \sum_{n=1}^{N} w_n K(x, x_n) + w_0 \tag{8}$$

where w_n are the model weights, $K(x, x_n)$ is a kernel function, and w_0 is the bias. Consider the case of approximating the set of data represented by $\{(x_i, y_i), i = 1, \ldots, N\}$. Assume that $p(t|x)$ is Gaussian $N(t|y(x), \sigma^2)$. The likelihood of the dataset is given by [30]:

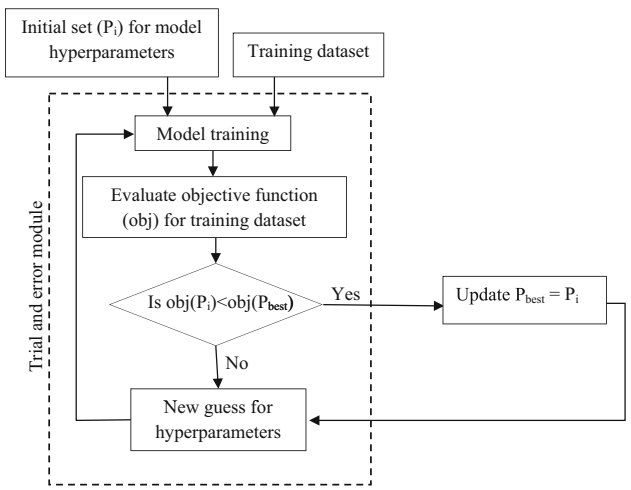

Fig. 4 Flowchart for evaluation of hyperparameters based on trial and error

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

Table 1 Training and testing datasets

Sl. no.	C_3S (%)	SO_3 (%)	Blaine (cm^2/g)	Alkali (%)	Compressive strength (MPa) [2]
Training dataset					
1	54.0	3.0	3530.0	1.1	53.9
2	54.8	2.9	3680.0	0.9	51.9
3	57.3	2.8	3560.0	1.0	53.9
4	64.6	2.6	3850.0	1.0	50.8
5	56.9	2.7	3580.0	0.8	54.5
6	61.3	2.3	3780.0	0.9	50.4
7	62.3	2.8	3640.0	0.9	55.4
8	62.4	2.8	3590.0	0.9	58.4
9	64.6	2.5	4090.0	0.8	54.8
10	59.3	2.8	3500.0	1.1	51.8
11	61.8	2.7	3630.0	1.1	51.3
12	61.3	3.0	3580.0	1.0	54.7
13	60.4	2.6	3680.0	1.0	54.1
14	55.6	3.1	3510.0	1.0	54.5
15	62.4	2.5	3590.0	1.1	51.5
16	63.1	2.6	3540.0	0.9	52.1
17	61.2	2.7	3610.0	0.9	51.7
18	55.6	2.7	3620.0	0.9	54.2
19	67.3	2.6	4020.0	0.8	53.8
20	58.7	3.0	3550.0	0.9	51.5
21	65.4	2.3	3730.0	0.9	48.9
22	58.0	2.7	3420.0	1.0	53.2
23	65.0	2.5	4070.0	0.8	54.7
24	62.0	2.9	3720.0	1.0	54.3
25	61.4	2.7	3840.0	0.9	52.5
26	63.5	2.5	3540.0	1.0	51.3
27	62.8	2.3	3580.0	0.9	51.1
28	56.4	3.0	3370.0	1.1	52.5
29	62.8	3.0	3750.0	1.1	54.1
30	58.9	3.0	3540.0	1.0	53.5
Testing dataset					
1	62.3	2.5	3910.0	0.9	53.6
2	57.7	2.7	3480.0	1.0	55.4
3	55.8	3.1	3420.0	0.9	53.7
4	55.9	2.8	3620.0	1.0	55.6
5	60.7	2.8	3740.0	1.1	55.2
6	59.3	2.5	3750.0	1.1	55.5
7	60.8	2.2	3520.0	1.1	49.8
8	60.7	3.0	3840.0	0.9	55.6
9	63.2	2.5	4010.0	0.9	52.1
10	59.3	2.6	3450.0	1.0	51.6
11	65.8	2.6	4050.0	0.9	53.0
12	57.4	2.5	3390.0	1.1	50.5
13	62.0	2.4	3490.0	1.0	54.0
14	59.7	2.2	3890.0	1.0	52.1
15	56.8	2.7	3620.0	1.0	53.8
16	61.7	2.4	3630.0	0.9	53.6
17	63.6	2.8	3680.0	0.9	53.0
18	61.6	2.8	3630.0	1.1	53.5

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

Table 1 continued

Sl. no.	C_3S (%)	SO_3 (%)	Blaine (cm^2/g)	Alkali (%)	Compressive strength (MPa) [2]
19	64.9	2.4	3900.0	1.0	49.9
20	61.0	2.8	3700.0	0.9	54.2

Table 2 Statistical properties of the dataset

	C_3S (%)	SO_3 (%)	Blaine (cm^2/g)	Alkali (%)	Compressive strength (MPa)
Total					
Maximum	67.30	3.10	4090.00	1.10	58.40
Minimum	54.00	2.20	3370.00	0.80	48.90
Mean	60.63	2.68	3668.20	0.97	53.14
Standard deviation	3.14	0.23	181.27	0.09	1.83
Median	61.25	2.70	3625.00	1.00	53.55
Variance	9.87	0.05	32,858.76	0.01	3.36
Training dataset					
Maximum	67.30	3.10	4090.00	1.10	58.40
Minimum	54.00	2.30	3370.00	0.80	48.90
Mean	60.70	2.72	3656.33	0.96	53.04
Standard deviation	3.38	0.22	172.13	0.10	1.87
Median	61.35	2.70	3600.00	0.95	53.35
Variance	11.42	0.05	29,629.89	0.01	3.48
Testing dataset					
Maximum	65.80	3.10	4050.00	1.10	55.60
Minimum	55.80	2.20	3390.00	0.90	49.80
Mean	60.51	2.62	3686.00	0.99	53.29
Standard deviation	2.75	0.24	192.81	0.08	1.77
Median	60.75	2.60	3655.00	1.00	53.60
Variance	7.54	0.06	37,174.00	0.01	3.15

$$p(t|w,\sigma^2) = (2\pi\sigma^2)^{-N/2}\exp\left\{-\frac{1}{2\sigma^2}t - \Phi w^2\right\} \quad (9)$$

where $t = (t_1,\ldots,t_n), w = (w_0,\ldots,w_n)$ and Φ is the $N \times (N+1)$ design matrix with $\Phi_{nm} = K(x_n, x_{m-1})$ and $\Phi_{n1} = 1$. Hyperparameters (α_i) are introduced for every weight in order to avoid severe overfitting and provide sparsity properties. Using Bayes' rule, the posterior over the weights is obtained as [30]:

$$p(w|t,\alpha,\sigma^2) = (2\pi)^{-\frac{N+1}{2}}|\Sigma|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(w-\mu)^T\Sigma^{-1}(w-\mu)\right\}$$
$$\Sigma = \left(\Phi^T B\Phi + A\right)^{-1} \quad \text{and} \quad \mu = \Sigma\Phi^T Bt, \quad (10)$$

where $A = \mathrm{diag}(\alpha_0,\ldots,\alpha_N)$ and $B = \sigma^{-2}I_N$. The marginal likelihood for the hyperparameters is given by [30]:

$$p(t|\alpha,\sigma^2) = (2\pi)^{-\frac{N}{2}}\left|B^{-1} + \Phi A^{-1}\Phi^T\right|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}t^T\left(B^{-1} + \Phi A^{-1}\Phi^T\right)^{-1}t\right\} \quad (11)$$

Table 3 Optimization parameters for SOS and PSO

Parameter	Value
SOS	
Size of the ecosystem	100
Maximum number of function evaluations	500
PSO	
Swarm size	25
Cognitive acceleration coefficient	2.8
Social acceleration coefficient	1.3
Maximum number of iterations	2500
Maximum duration of optimization	2500
Maximum difference between best and worst function evaluation in simplex	10^{-6}
Maximum difference between the coordinates of the vertices	10^{-3}

The value of σ^2 which maximizes the marginal likelihood for the hyperparameters is evaluated. The value of σ^2 is obtained from trial and error. The

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

flowchart for the procedure followed is shown in Fig. 4.

2.3 Gaussian process regression (GPR)

GPR is a powerful nonparametric technique with explicit uncertainty models for nonlinear regression. The major advantage of GPR is that the nonlinear problems can be solved in the Bayesian framework for learning, model selection and density estimation with the relatively simple linear algebra for the basic model. GPR tries to infer the correlation between the measured data rather than fitting the parameters of the selected basis function. Brief overview of GPR is given below. For details, readers are advised to refer Williams [34] and Rasmussen and Williams [27].

Consider the case of approximating the set of data represented by $\{(x_i, y_i), i = 1, \ldots, N\}$, $X \in \mathbb{R}^n, Y \in \mathbb{R}$, where N is the number of data point; x is input variable; y is output variable; and $\mathbb{R}^d$ is d-dimensional vector space. GPR assumes that the output of the regression model is dependent on the latent function ($f(x)$) and the Gaussian noise (ε) with zero mean and σ^2 variance ($N(0, \sigma^2)$). The output of the regression model is expressed as:

$$y = f(x) + \varepsilon \tag{12}$$

For a testing input x_t, the distribution of output y_t is given by

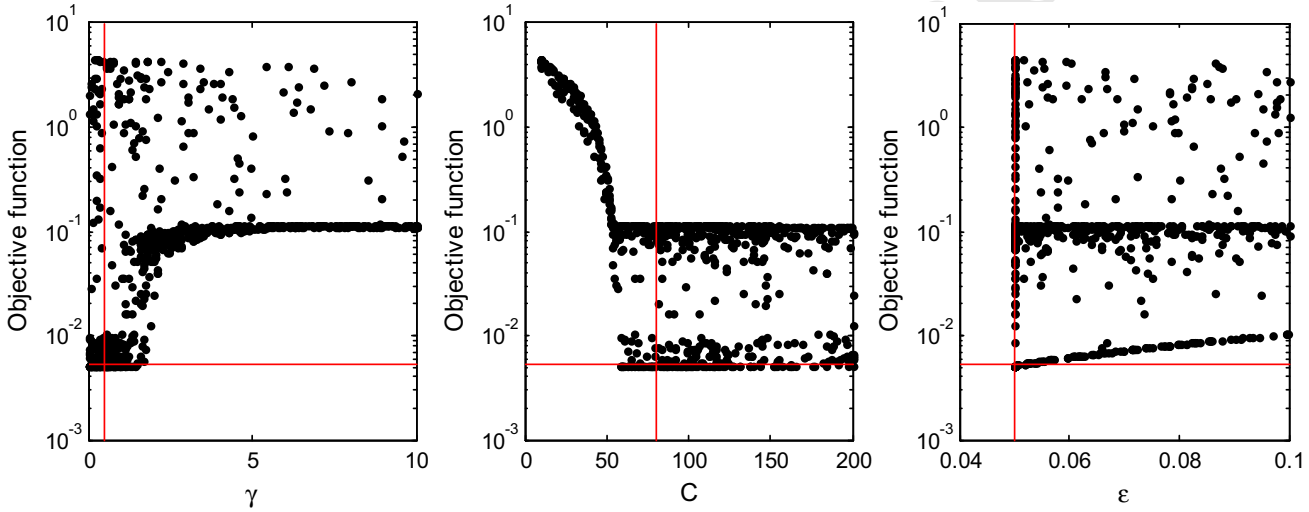

Fig. 5 Variation of objective function with the hyperparameters for SOS-SVR model

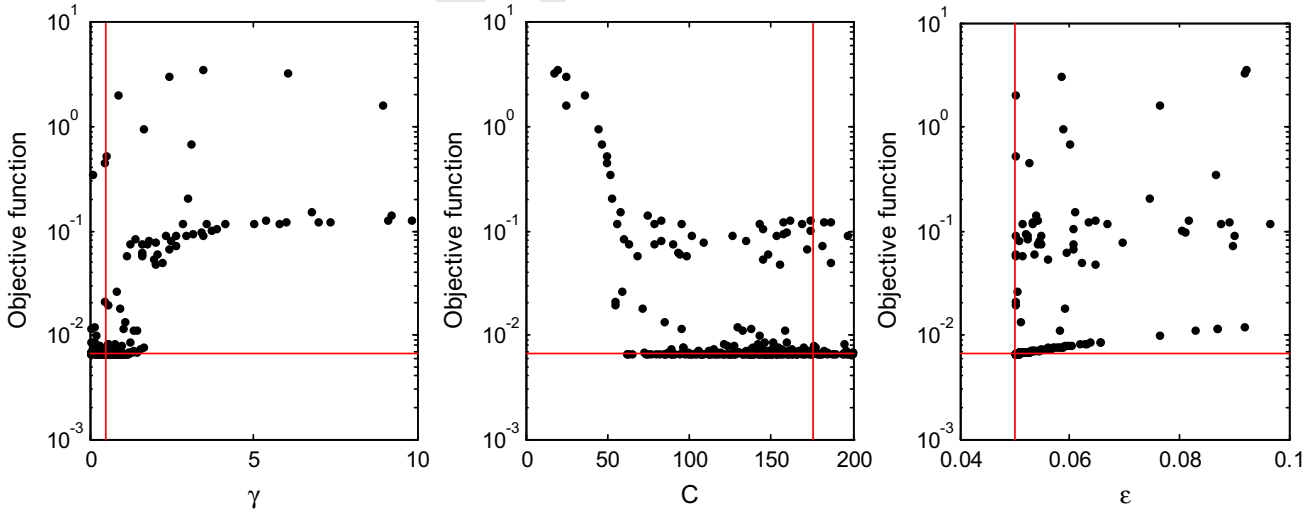

Fig. 6 Variation of objective function with the hyperparameters for PSO-SVR model

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

$$\begin{pmatrix} y \\ y_t \end{pmatrix} \sim N(0, K_t) \tag{13}$$

with covariance matrix, $K_t = \begin{bmatrix} k(x)+\sigma^2 I & K(x,x_t) \\ K(x,x_t)^T & k(x_t) \end{bmatrix}$ where $K(x, x_t)$ is the covariance between the training inputs and the test input; $k(x_t)$ is the autocovariance of the test input and $k(x)$ = autocovariance of the training input.

The output follows Gaussian distribution with a mean and variance:

$$\mu = K(x,x_t)^T \left(k(x)+\sigma^2 I\right)^{-1} y \tag{14}$$

$$\sigma = k(x_t) - K(x,x_t)^T \left(k(x)+\sigma^2 I\right)^{-1} K(x,x_t) \tag{15}$$

The model hyperparameter for GPR is obtained using the trial-and-error approach as summarized in Fig. 4.

3 Model application

SVR has been implemented using the SVM toolbox [13] in MATLAB. The toolbox simpleR [4, 17] is used for implementing RVM and GPR in MATLAB. Different kernel-based models are applied to predict 28-day compressive strength of cement. The application gains significance due to the profound influence of cement strength on the strength of structural members. Based on the investigations by Akkurt et al. [2], the input variables are chosen as C_3S, SO_3, Alkali and Blaine, while the output variable is 28-day compressive strength (N/mm^2) of the cement. The first three input variables are expressed in percentages, while the last variable is expressed in multiples of cm^2/g. The data used in the present study have

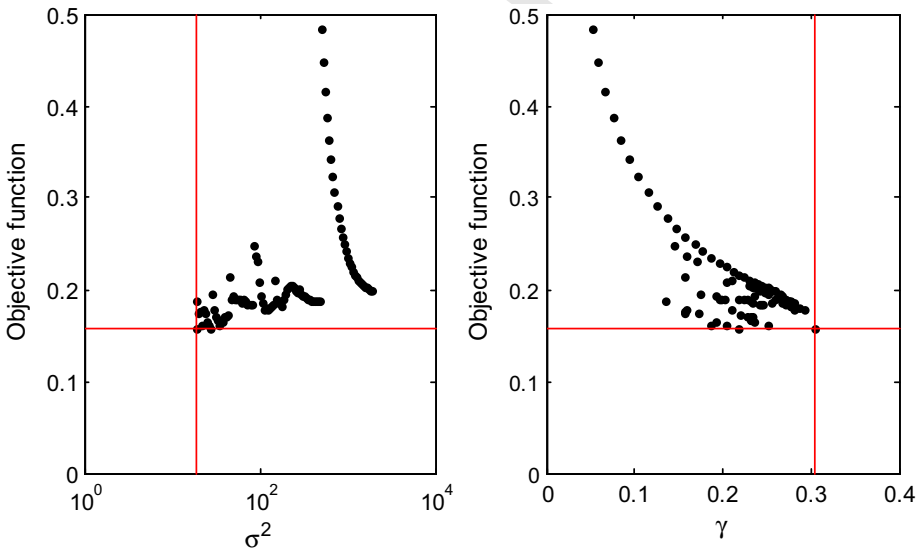

Fig. 7 Variation of objective function with the hyperparameters for RVM model

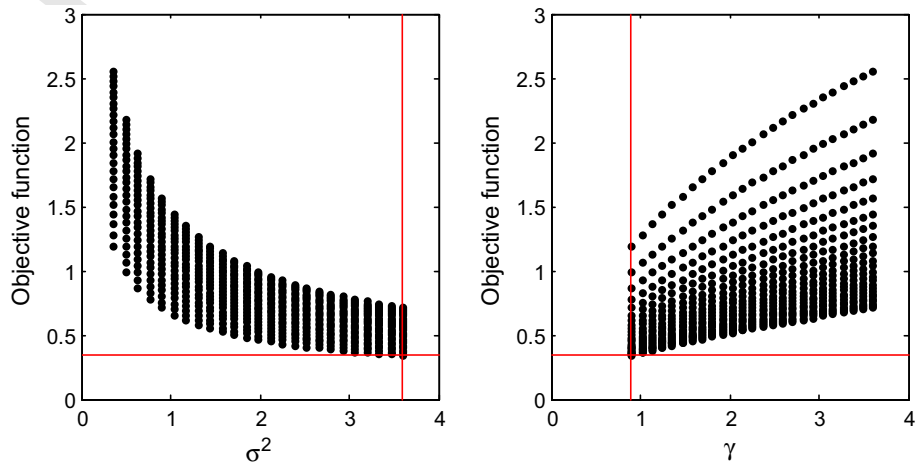

Fig. 8 Variation of objective function with the hyperparameters for GPR model

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

been adopted from Akkurt et al. [2]. The data used were collected from a local cement plant that uses strength testing for process control. The cement compressive strength testing was carried out according to the European standard [10] at 28th day. For further details, please refer Akkurt et al. [2].

Table 4 Optimal model hyperparameters

Method	Hyperparameters
SOS-SVR	
Loss function	ε-Insensitive
Kernel function	Radial basis function
C	79.8074
ε	0.05
γ	0.4827
PSO-SVR	
Loss function	ε-Insensitive
Kernel function	Radial basis function
C	175.6521
ε	0.05
γ	0.4733
RVM	
Kernel function	Radial basis function
γ	0.3048
σ^2	18.8592
GPR	
Covariance function	Radial basis function
γ	0.9003
σ^2	3.6012

3.1 Dataset preparation

The dataset adopted from Akkurt et al. [2] is randomly divided into two parts—training and testing. Out of the 50 datasets, 30 are used for training and 20 for testing. The training and testing datasets are given in Table 1. In order to avoid the clustering of the data, it was assured that the statistical properties of the original data are preserved in both training and testing datasets. The statistical properties of the dataset are given in Table 2.

3.2 Evaluation of hyperparameters

The values of the hyperparameters for different kernel-based models are obtained based on the flowchart described in the previous section. The details of the optimization parameters for SOS and PSO are given in Table 3. The variation of the objective function with respect to the whole set of hyperparameters explored is shown in Figs. 5, 6, 7 and 8 for different models. Each point in the plot corresponds to a single experiment. The horizontal red line in the plot corresponds to the minimum value of the objective function obtained. The vertical red line corresponds to the optimal value of the hyperparameter. The optimal values of the hyperparameters obtained for different models are summarized in Table 4.

3.3 Prediction of compressive strength

The model is trained with the optimal values of the hyperparameters using the training dataset. The

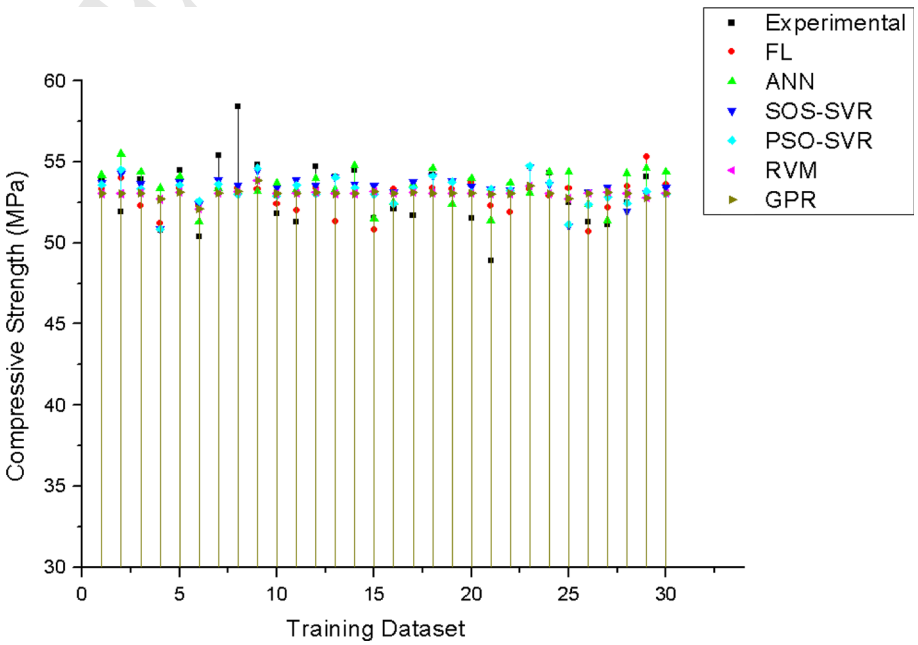

Fig. 9 Training phase of different models

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

Table 5 Compressive strength predicted from different models

Sl. no.	Compressive strength (MPa)						
	Experiment [2]	Fuzzy [2]	ANN [2]	SOS-SVR	PSO-SVR	RVM	GPR
1	53.9	53.3	54.2	53.72	53.59	53.04	53.04
2	51.9	54	55.5	54.29	54.53	53.04	53.04
3	53.9	52.3	54.4	53.64	53.34	53.08	53.08
4	50.8	51.2	53.4	50.85	50.85	52.70	52.70
5	54.5	53.4	54.1	53.79	53.57	53.14	53.14
6	50.4	52.3	51.3	52.42	52.58	52.10	52.10
7	55.4	53.3	53.4	53.91	53.63	53.05	53.05
8	58.4	53.4	53.2	53.52	52.99	53.16	53.16
9	54.8	53.3	53.2	54.48	54.61	53.87	53.87
10	51.8	52.4	53.7	53.38	52.91	53.04	53.04
11	51.3	52	53.3	53.88	53.57	53.06	53.06
12	54.7	53.4	54	53.52	53.02	53.15	53.15
13	54.1	51.3	53.3	54.05	54.05	53.04	53.04
14	54.5	53.4	54.8	53.61	53.38	53.04	53.04
15	51.5	50.8	51.5	53.52	52.99	53.16	53.16
16	52.1	53.3	52.5	53.16	52.43	53.05	53.05
17	51.7	53.4	53.6	53.76	53.40	53.11	53.11
18	54.2	53.4	54.6	54.15	54.15	53.08	53.08
19	53.8	53.3	52.4	53.85	53.75	53.04	53.04
20	51.5	53.7	54	53.50	53.09	53.06	53.06
21	48.9	52.3	51.4	53.33	53.32	53.02	53.02
22	53.2	51.9	53.7	53.25	53.25	53.04	53.04
23	54.7	53.3	53.1	54.69	54.75	53.51	53.51
24	54.3	52.9	54.4	53.63	53.70	53.04	53.04
25	52.5	53.4	54.4	51.07	51.13	52.74	52.74
26	51.3	50.7	52.4	53.13	52.37	53.05	53.05
27	51.1	52.2	51.4	53.41	52.81	53.15	53.15
28	52.5	53.5	54.3	51.95	52.45	53.04	53.04
29	54.1	55.3	54.6	53.07	53.18	52.78	52.78
30	53.5	53.6	54.4	53.45	53.00	53.05	53.05
31	53.6	53.2	53.3	50.80	50.68	53.04	53.07
32	55.4	51.7	53.8	53.50	53.19	53.04	53.11
33	53.7	53.6	54	53.36	53.49	53.04	53.50
34	55.6	52.4	54.4	54.14	54.12	53.08	53.46
35	55.2	52.2	53.9	53.34	53.50	52.94	52.44
36	55.5	50.7	50.7	53.18	53.39	52.78	52.39
37	49.8	50.4	49	53.28	52.68	53.04	52.94
38	55.6	54.3	54.8	51.08	51.15	52.74	52.74
39	52.1	53.3	53.3	53.58	53.46	53.04	53.09
40	51.6	51.4	53.2	53.39	53.11	53.04	53.57
41	53	53.3	53.6	54.55	54.54	53.13	52.25
42	50.5	50.8	50.8	52.65	52.93	53.04	52.75
43	54	50.7	51.6	53.21	52.58	53.04	52.29
44	52.1	50.4	49.6	50.68	50.63	53.01	52.77
45	53.8	51.8	53.8	54.09	54.02	53.08	53.32
46	53.6	52.6	52.2	53.88	53.58	53.06	52.44
47	53	53.4	53.2	53.89	53.75	53.04	56.00
48	53.5	52.4	53.9	53.89	53.59	53.06	52.54
49	49.9	50.6	51.6	50.67	50.55	53.03	52.22
50	54.2	53.6	54.1	53.90	53.95	53.04	55.95

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

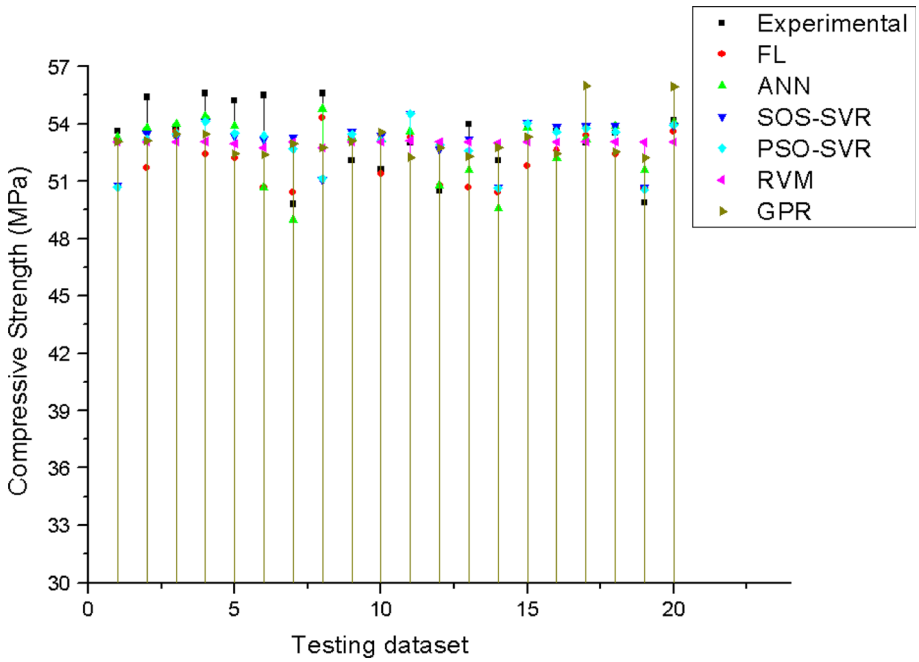

Fig. 10 Testing phase of different models

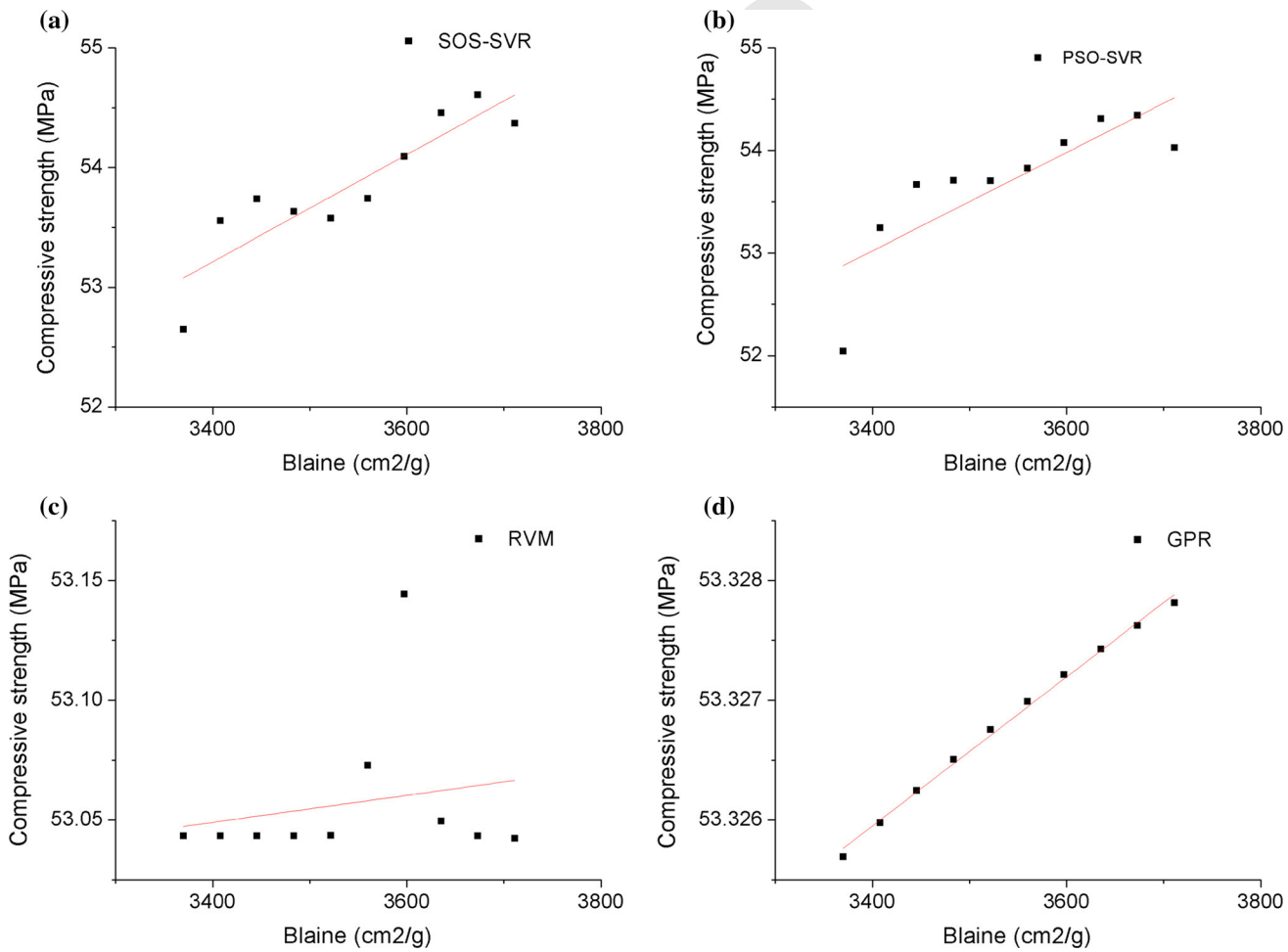

Fig. 11 Variation of compressive strength with Blaine for different models

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

Table 6 Statistical properties of the compressive strength predicted from different models

	Compressive strength (MPa)				
	Experimental	SOS-SVR	PSO-SVR	RVM	GPR
Maximum	58.40	55.30	55.50	54.69	54.75
Minimum	48.90	50.40	49.00	50.67	50.55
Mean	53.14	52.58	53.19	53.26	53.11
Standard deviation	1.83	1.15	1.37	1.01	1.01
Median	53.55	53.05	53.50	53.51	53.33
Variance	3.36	1.31	1.87	1.02	1.01

comparison of the experimental and the predicted cement compressive strength obtained from different models for training dataset is shown in Fig. 9. The trained model is then used to predict the cement compressive strength for testing dataset. The predicted cement compressive strength using artificial neural networks (ANN) and fuzzy logic (FL) is taken from Akkurt et al. [2]. The compressive strength predicted from different models is given in Table 5. The comparison of the experimental and the predicted cement compressive strength obtained from different models for testing dataset is shown in Fig. 10. The predicted values of cement compressive strength are found to have a variation within 10 % of the experimental cement compressive strength. The comparison confirms the excellent performance of the proposed model except for few datasets.

3.4 Model validation

The validation of the model is carried out by studying the effect of variations in the input variables on the compressive strength of the cement. The variation in the compressive strength of cement obtained from the kernel-based models is plotted for different values of input variables. The trends obtained are then compared with those reported in the literature. An acceptable model should not only be able to accurately predict the compressive strength, but it should also have the capability to accurately represent the variation in the compressive strength with respect to the input variables. The input variable selected in the present study is Blaine. However, the similar exercise can be carried out for other variables as well. The input variables C_3S, SO_3 and Alkali are kept at a constant value of 54, 3 and 1.1 %. Blaine is varied from 3370 to 3700 cm^2/g. The kernel-based models are then used to predict the compressive strength. The variation of the compressive strength obtained from kernel-based models to the varying Blaine content is shown in Fig. 11. A line of best fit is used to describe the general trend. The predicted compressive strength is found to increase with the increase in the Blaine content. With the increase in the Blaine content, more surface area is available for the hydration of cement. This results in the increase of cement compressive strength. The similar trend has been reported in the literature by Celik [5]. Hence, the kernel-based models are validated.

4 Error indices

The performance of the different kernel-based models is compared with that of the fuzzy [2] and the ANN [1] models. Six different error indices are adopted to compare the performance of the different models:

1. Mean absolute error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \tag{16}$$

Table 7 Error measures of different models

Method	Error indices					
	MAE	MAPE	MaxAE	MSE	RMSE	RRSE
Training						
FL	1.41	2.65	8.56	2.94	1.71	0.92
ANN	1.35	2.55	8.90	3.13	1.77	0.95
SOS-SVR	1.27	2.43	9.05	3.12	1.77	0.94
PSO-SVR	1.21	2.29	9.26	3.02	1.74	0.93
RVM	1.46	2.76	8.97	3.15	1.78	0.95
GPR	1.46	2.76	8.97	3.15	1.78	0.95
Testing						
FL	1.50	2.76	8.65	4.04	2.01	1.13
ANN	1.18	2.20	8.65	2.59	1.61	0.90
SOS-SVR	1.54	2.89	8.13	3.60	1.90	1.07
PSO-SVR	1.48	2.78	8.00	3.44	1.85	1.04
RVM	1.51	2.84	6.51	3.36	1.83	1.03
GPR	1.75	3.29	6.31	3.96	1.99	1.12
Total						
FL	1.44	2.69	8.65	3.38	1.84	1.00
ANN	1.28	2.41	8.90	2.91	1.71	0.93
SOS-SVR	1.38	2.61	9.05	3.31	1.82	0.99
PSO-SVR	1.32	2.49	9.26	3.19	1.78	0.97
RVM	1.48	2.79	8.97	3.24	1.80	0.98
GPR	1.58	2.97	8.97	3.48	1.86	1.02

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

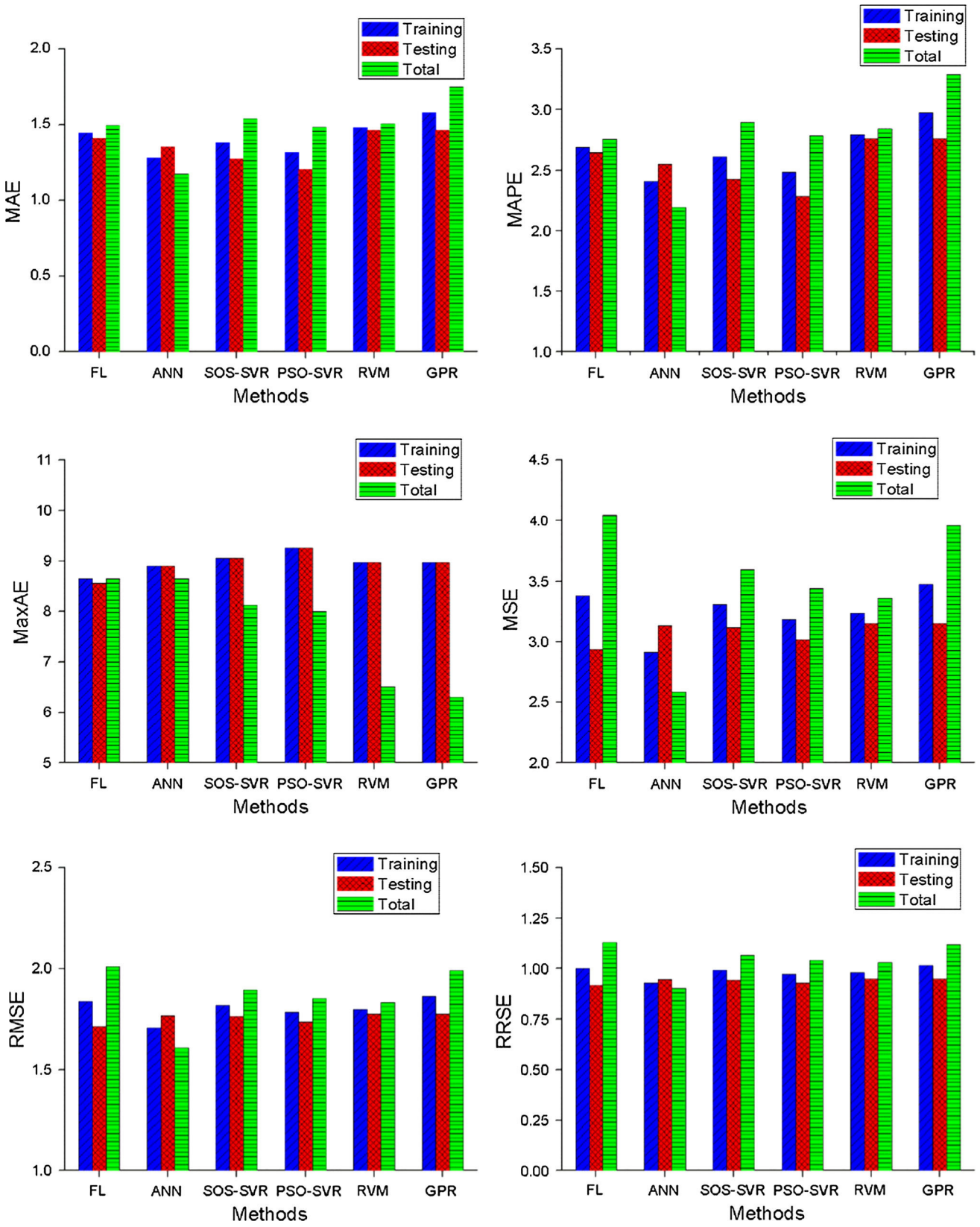

Fig. 12 Error indices

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

2. Mean of the absolute percentage error (MAPE)

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{y_i}\right| \times 100 \quad (17)$$

3. Maximum absolute percentage error (MaxAE)

$$\text{MaxAE} = \text{Max}\left(\left|\frac{x_i - y_i}{y_i}\right| \times 100\right) \quad (18)$$

4. Mean square error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \quad (19)$$

5. Root-mean-square of percentage error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad (20)$$

6. Root relative squared error (RRSE)

$$\text{RRSE} = \sqrt{\sum_{i=1}^{n}\frac{(y_i - x_i)^2}{(x_i - \bar{x})^2}} \quad (21)$$

where x_i and y_i are experimental and predicted values; $\bar{x}$ is the mean of the experimental values; and n is the total number of values used in the comparative study. The error indices are calculated using the output of the kernel-based models available in Table 5. The statistical properties of the output of different models are given in Table 6. The values of the error indices obtained are summarized in Table 7. It should be noted that the kernel-based models have been trained with only 30 datasets, while fuzzy and the ANN models are trained with more number of datasets.

5 Results and discussion

The statistical properties of the compressive strength obtained from the different kernel-based models are found to be similar to that obtained experimentally (Table 6). The comparison of the various error indices obtained for different models in training, testing and total phase is shown in Fig. 12. A marginal difference is found in the error indices for the models selected in the present study. In order to benchmark their performance w.r.t. each other, the models are first ranked for each of the error indices. The average rank for each method is obtained by averaging the rank of the method obtained for various error indices in training, testing and total phases. The order of performance of the various models is summarized below:

1. Training phase

   FL > PSO − SVR > SOS − SVR > ANN > RVM = GPR

2. Testing phase

   ANN > RVM > PSO − SVR > SOS − SVR > FL = GPR

3. Total phase

   ANN > PSO − SVR > RVM > SOS − SVR > FL > GPR

The box plot of the residuals (difference between predicted and experimental cement compressive strength) for testing phase is shown in Fig. 13. The variation of the residuals is found to be least for ANN. The median of the residuals is found to be almost same for all the methods. However, the difference between the maximum and minimum residual is found to be on the higher side for SOS-SVR and PSO-SVR compared to other models. The

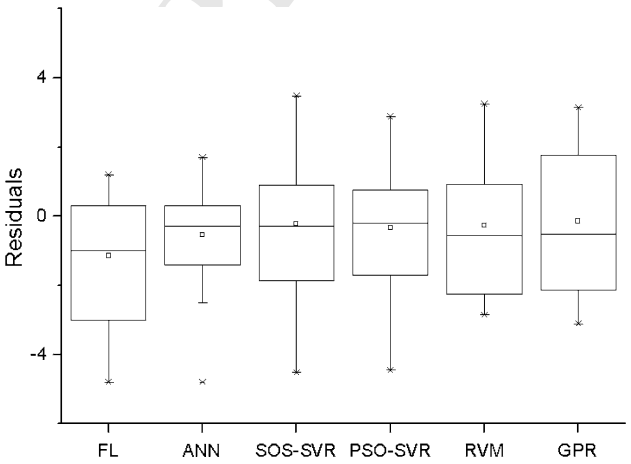

Fig. 13 Residual in testing phase

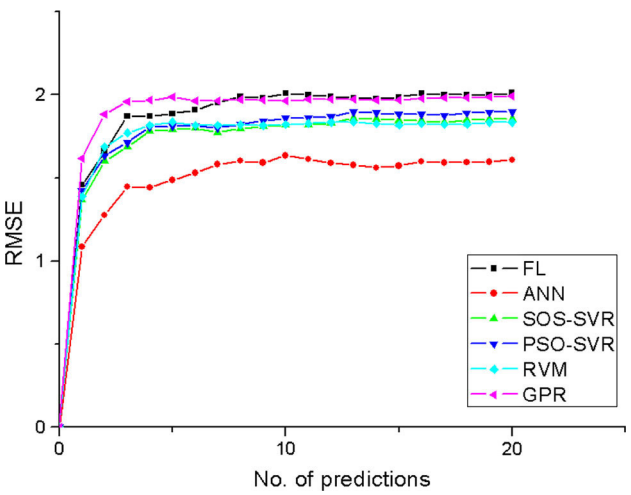

Fig. 14 Variation of RMSE with number of predictions

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

variation of RMSE with the number of predictions is shown in Fig. 14. The RMSE values for all the models are found to converge without any abrupt jump. This indicates the consistency of the models in prediction of the cement compressive strength. ANN is found to converge to a least value of RMSE. SOS-SVR, PSO-SVR and RVM are found to converge at the same rate to a similar RMSE value. GPR and FL converged to almost same RMSE value at the end of 20 iterations. Figure 15 shows the residual contours for different models. The contours give qualitative idea

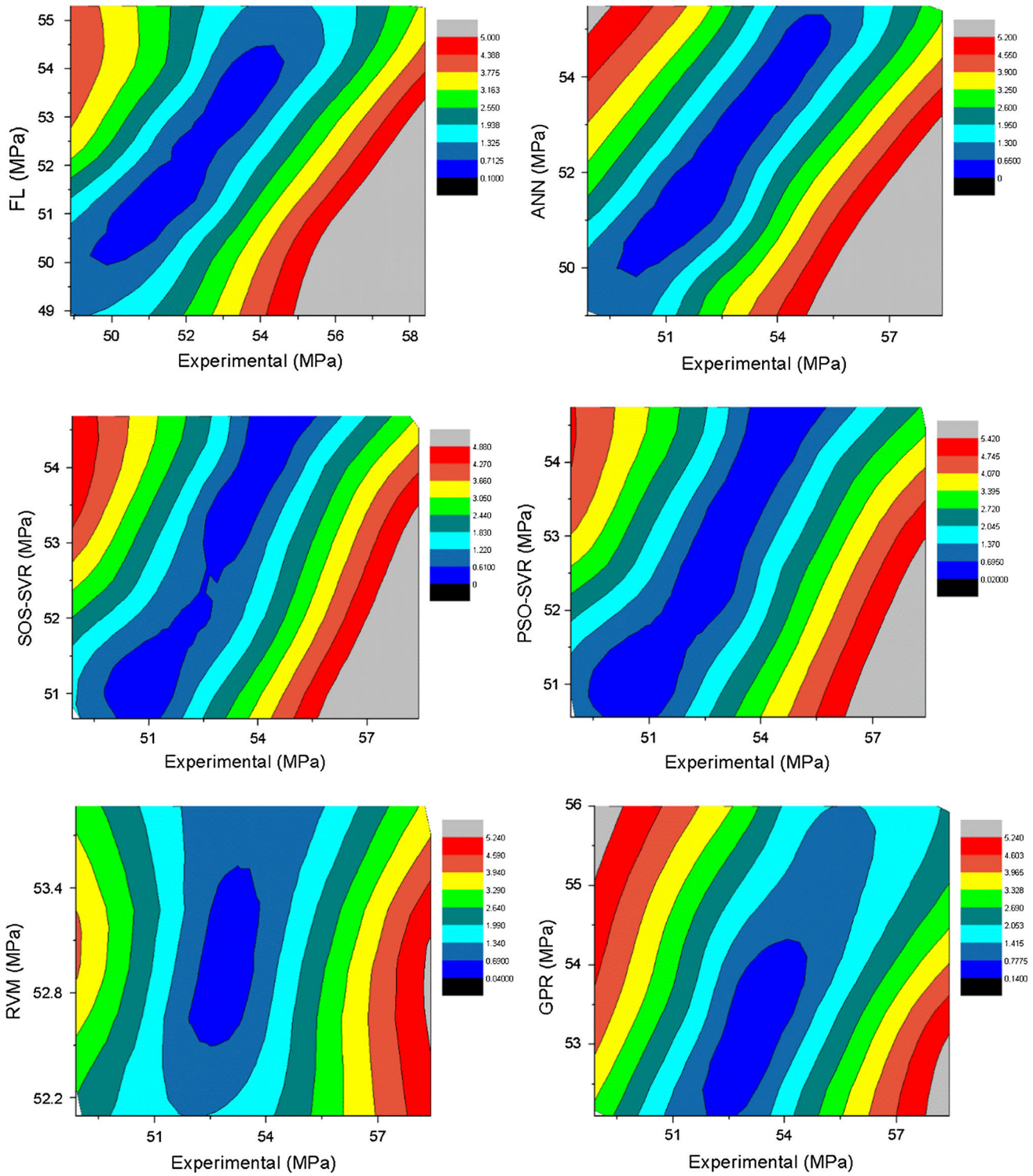

Fig. 15 Contour plots for residuals in testing phase

Journal : Large 521 Dispatch : 15-6-2016 Pages : 18
Article No. : 2419 LE TYPESET
MS Code : NCAA-D-15-01328 CP DISK

about the range of the output parameter where the model has the highest accuracy. Based on the contours, ANN, SOS-SVR and PSO-SVR are found to be accurate over wider range of output parameter compared to other models.

Based on the error indices and analysis of residuals, ANN is found to perform best among all the models. However, there are some disadvantages associated with ANN such as slow conversion speed, low generalization capabilities, absence of probabilistic output, over fitting and arrival at local minima. SVR and RVM provide alternate options which can overcome the problems associated with ANN. In SVR, the solution obtained is unique as it involves convex optimization. The appropriate choice of the model parameters makes SVR robust even if the training data have some bias. SVR linearizes the data on the implicit basis by means of kernel function transformation, and therefore, the accuracy of the model does not rely on the quality of human expertise judgement for the optimal choice of the linearization function. The only disadvantage of SVR is that it requires tuning of the model parameters which is taken care by optimization in the present study and the results obtained are not probabilistic in nature. On the other hand, RVM has reduced sensitivity to the hyperparameters and has ability to use non-Mercer kernel functions. The output obtained from RVM is probabilistic in nature and required definition of only one parameter. RVM has good generalization performance and yields sparse models with fewer relevance vectors which provides the inference at moderate computational cost. However, training phase in RVM involves solving of highly nonlinear optimization problem which makes it unsuitable for larger datasets.

6 Conclusions

Kernel-based models namely support vector regression (SVR), relevance vector machine (RVM) and Gaussian process regression (GPR) were adopted for predicting the 28-day compressive strength of the cement. The input variables for the model were taken as C_3S, SO_3, Alkali (in percentages) and Blaine (in cm^2/g). The hyperparameters of the SVR were obtained by carrying out optimization using two different types of metaheuristic algorithms—particle swarm optimization (PSO) and symbiotic organism search (SOS). For RVM and GPR, trial and error were used to arrive at the model parameters. Six different error indices were used to benchmark the performance of the different models. Residual analysis was also carried out to get better idea about the consistency and the range of output parameter where the model is accurate. Artificial neural network (ANN) is found to perform best in predicting cement compressive strength among all the methods. The performance of PSO-SVR, SOS-SVR and RVM was found to be comparable to ANN. PSO was found to be a suitable algorithm for obtaining the optimized hyperparameters of SVR. Though fuzzy logic (FL) method performed well in the training phase, the overall performance was not good compared to other models. GPR was ranked last in the order of performance in all the phases. The error indices and residual analysis should not be used as the sole criterion for the selection of the method. The computational cost associated with the solving of optimization problem and large storage space requirements might be a concern for larger datasets. The unique features of each of the models should be kept in mind before selection of appropriate method. For instance, ANN does not have generalizing capabilities. On the other hand, SVR has greater ability to generalize and provides promising empirical performance. Therefore, even though ANN performed well, SVR should be preferred. The desired compressive strength of the cement can be achieved by altering the levels of C_3S, SO_3, Alkali and Blaine at the manufacturing stage. The proposed model can be further improved by introduction of additional input variables such as C_3A, C_4AF and lime, which can influence the compressive strength of the cement to a lesser extent. The future program will aim at the improvement of the model by incorporation of such variables.

Acknowledgments Authors would also like to thank and acknowledge the help received from their colleagues of Shock and Vibration Group, CSIR-SERC. This paper is being published with the kind permission of the Director, CSIR-SERC, Chennai.

References

1. Akkurt S, Ozdemir S, Tayfur G, Akyol B (2003) The use of GA-ANNs in the modelling of compressive strength of cement mortar. Cem Concr Res 33:973–979
2. Akkurt S, Tayfur G, Can S (2004) Fuzzy logic model for the prediction of cement compressive strength. Cem Concr Res 34:1429–1433
3. Bermolen P, Rossi D (2009) Support vector regression for link load prediction. Comput Netw 53(2):191–201
4. Camps-Valls G, Gómez-Chova L, Muñoz-Marí J, Vila-Francés J, Amorós-López J, Calpe-Maravilla J (2006) Retrieval of oceanic chlorophyll concentration with relevance vector machines. Remote Sens Environ 105(1):23–33
5. Celik IB (2009) The effects of particle size distribution and surface area upon cement strength development. Powder Technol 188(3):272–276
6. Chen BT, Chang T-P, Shih J-Y, Wa J-J (2008) Estimation of exposed temperature for fire-damaged concrete using support vector machine. Comput Mater Sci 44:913–920
7. Chen JJ, Kwan AKH (2012) Superfine cement for improving packing density, rheology and strength of cement paste. Cem Concr Compos 34(1):1–10
8. Cheng MY, Prayogo D (2014) Symbiotic organisms search: a new metaheuristic optimization algorithm. Comput Struct 139:98–112

Journal : Large 521 | Dispatch : 15-6-2016 | Pages : 18
Article No. : 2419 | LE | TYPESET
MS Code : NCAA-D-15-01328 | CP | DISK

9. De Siquera Tango CE (1998) An extrapolation method for compressive strength prediction of hydraulic cement products. Cem Concr Res 28(7):969–983
10. European Committee for Standardization (CEN), Methods of testing cement part 1: determination of strength. European Standard EN 196-1
11. Fa-Liang G (1997) A new way of predicting cement strength-fuzzy logic. Cem Concr Res 27:883–888
12. Fattahi H, Gholami A, Amiribakhtiar MS, Moradi S (2015) Estimation of asphaltene precipitation from titration data: a hybrid support vector regression with harmony search. Neural Comput Appl 26(4):789–798
13. Gunn SR (1998) Support vector machines for classification and regression. Technical report, Image Speech and Intelligent Systems Research Group, University of Southampton
14. Hasanipanah M, Shahnazar A, Amnieh HB, Armaghani DJ (2016) Prediction of air-overpressure caused by mine blasting using a new hybrid PSO-SVR model. Eng Comput 1–9
15. Huang CL, Tsai CY (2009) A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. Expert Syst Appl 36(2):1529–1539
16. Kennedy J, Eberhart RC, Shi Y (2001) Swarm intelligence. Morgan Kaufmann, New York, pp 303–308
17. Lázaro-Gredilla M, Titsias MK, Verrelst J, Camps-Valls G (2014) Retrieval of biophysical parameters with heteroscedastic Gaussian processes. IEEE Geosci Remote Sens Lett 11(4): 838–842
18. Lee JJ, Kim DK, Chang SK, Lee JH (2007) Application of support vector regression for the prediction of concrete strength. Comput Concr 4(4):299–316
19. Lu CJ (2013) Hybridizing nonlinear independent component analysis and support vector regression with particle swarm optimization for stock index forecasting. Neural Comput Appl 23(7–8):2417–2427
20. Maity R, Bhagwat PP, Bhatnagar A (2010) Potential of support vector regression for prediction of monthly stream flow using endogenous property. Hydrol Process 24(7):917–923
21. Mechling J-M, Lecomte A, Diliberto C (2009) Relation between cement composition and compressive strength of pure pastes. Cem Concr Compos 31(4):255–262
22. Motamedi S, Shamshirband S, Hashim R, Petković D, Roy C (2015) Estimating unconfined compressive strength of cockle shell–cement–sand mixtures using soft computing methodologies. Eng Struct 98:49–58
23. Motamedi S, Shamshirband S, Petković D, Hashim R (2015) Application of adaptive neuro-fuzzy technique to predict the unconfined compressive strength of PFA–sand–cement mixture. Powder Technol 278:278–285
24. Osbaeck B, Johansen V (1989) Particle size distribution and rate of strength development of Portland cement. J Am Ceram Soc 72(2):197–201
25. Osbaeck B (1979) The influence of Alkalis on the strength properties of Portland cement. ZKG 32:72–77
26. Owolabi TO, Faiz M, Olatunji SO, Popoola IK (2016) Computational intelligence method of determining the energy band gap of doped ZnO semiconductor. Mater Des 101:277–284
27. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press, Cambridge
28. Sebastia M, Olmo IF, Irabien A (2003) Neural network prediction of unconfined compressive strength of coal fly ash–cement mixtures. Cem Concr Res 33:1137–1146
29. Svinning K, Høskuldsson A, Justnes H (2010) Prediction of potential compressive strength of Portland clinker from its mineralogy. Cem Concr Compos 32(4):300–311
30. Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 1:211–244
31. Tsivilis S, Chaniotakis E, Badogiannisa E, Pahoulasa G, Iliasa A (1999) A study on the parameters affecting the properties of Portland limestone cements. Cem Concr Compos 21(2):107–116
32. Vapnik V (2000) The nature of statistical learning theory. Springer, New York
33. Vapnik V (1999) An overview of statistical learning theory. IEEE Trans Neural Network 10(5):988–999
34. Williams CKI (1998) Prediction with Gaussian processes: from linear regression and beyond. In: Jordan MI (ed) Learning and inference in graphical models. Kluwer Academic Press, Dordrecht, pp 599–621
35. Yuvaraj P, Murthy AR, Iyer NR, Sekar SK, Samui P (2013) Support vector regression based models to predict fracture characteristics of high strength and ultra high strength concrete beams. Eng Fract Mech 98:29–43
36. Zhang J, Sato T, Iai S (2006) Support vector regression for on-line health monitoring of large-scale structures. Struct Saf 28(4):392–406

Journal : Large 521 | Dispatch : 15-6-2016 | Pages : 18
Article No. : 2419 | ☐ LE | ☐ TYPESET
MS Code : NCAA-D-15-01328 | ☑ CP | ☑ DISK