Computers and Concrete, Vol. 6, No. 2 (2009) 155-165
DOI: http://dx.doi.org/10.12989/cac.2009.6.2.155

Clustering-based identification for the prediction of splitting tensile strength of concrete

Bulent Tutmez*

School of Engineering, Inonu University, 44280 Malatya, Turkey

(Received September 11, 2008, Accepted December 12, 2008)

Abstract. Splitting tensile strength (STS) of high-performance concrete (HPC) is one of the important mechanical properties for structural design. This property is related to compressive strength (CS), water/binder (W/B) ratio and concrete age. This paper presents a clustering-based fuzzy model for the prediction of STS based on the CS and (W/B) at a fixed age (28 days). The data driven fuzzy model consists of three main steps: fuzzy clustering, inference system, and prediction. The system can be analyzed directly by the model from measured data. The performance evaluations showed that the fuzzy model is more accurate than the other prediction models concerned.

Keywords: compressive strength; splitting tensile strength; water/binder ratio; fuzzy clustering; modelling.

1. Introduction

High-performance concrete (HPC) is concrete having desired properties and uniformity which cannot be obtained routinely using only conventional constituents and normal mixing, placing, and curing practice. As examples, these properties may include (Carino *et al.* 1990):

- Ease of placement and compaction without segregation
- Enhanced long-term mechanical properties
- High early-age strength
- High toughness
- Volume stability
- Long life in severe environments.

The above definition was modified and adopted in 1998 as the ACI definition of HPC:

"Concrete meeting special combinations of performance and uniformity requirements that cannot always be achieved routinely using conventional constituents and normal mixing, placing, and curing practices (Russell 1999)". Current curing practices and standards are based on studies related primarily to strength development characteristics of conventional (ordinary) concretes. Most high-performance concretes, however, are fundamentally different from conventional concrete (Meeks and Carino 1999). HPC meets special mechanical performance that cannot always be achieved routinely via normal mixing and conventional materials (Zia *et al.* 1993).

Some important quality attributes of concrete materials include strength, volume stability, durability and permeability. Nevertheless, strength is often regarded as one of the most important properties. Tensile strength of HPC is mainly related to water/binder (W/B) ratio, compressive strength (CS)

* E-mail: btutmez@inonu.edu.tr

and concrete age. Many researchers (Zain *et al.* 2002; Choi and Yuan 2005) have investigated the relationships between tensile strength with these parameters and suggested different empirical expressions. In this paper, an application of clustering-based fuzzy modelling to the prediction of splitting tensile strength (STS) of concrete is presented.

In systems analysis, fuzzy logic has shown to be highly suitable for the modelling of complex and vague systems. Similarly, fuzzy logic is utilized to handle uncertainties and imprecision involved (Ayyub and Klir 2006). In practice, fuzzy models describe input-output relationships by fuzzy sets (membership functions) and if-then rules (fuzzy propositions). The most attractive characteristics of fuzzy models compared with other conventional methods, such as statistical models, are transparency and flexibility. There is an increasing interest in obtaining fuzzy models directly from experimental data (Setnes *et al.* 1998) and some recent works have been conducted for modelling the concrete parameters (Akkurt *et al.* 2004). This paper focuses on a data- driven soft computing method which is clustering-based fuzzy modelling for appraising splitting tensile strength of concrete. The aim of the present study is to obtain more accurate and transparent model structure in predicting STS of concrete by soft computing.

The rest of the paper is organized as follows. Section 2 presents a brief description for STS of HPC. Section 3 deals with clustering based fuzzy modelling structure. The application of the fuzzy model on experimental data sets is given in section 4. Section 5 presents the results and a brief discussion. Section 6 concludes the paper.

2 Splitting tensile strength

Different from the general concrete, high-performance concrete uses more cementious materials and has lower W/B ratio, and then it results in that the hydration heat and the early micro cracks are both increased, limiting the growth of tensile strength after HPC is hardened. On the other hand,

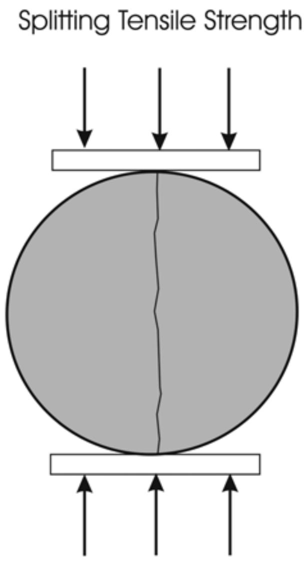

Fig. 1 Diagram of STS test

STS is an indirect indicator for understanding the tensile strength property of high-performance concrete (Song and Hwang 2004). STS tests involve compressing a cylinder or core on its side until a crack forms down the middle, causing failure of the specimen (Fig. 1).

Mindess and Young (1981) stated that concrete age, compressive strength, aggregate type, and degree of compaction influence STS. Therefore, STS can be estimated using these parameters. One of the empirical formulae developed for determining STS from compressive strength is given by De Larrard and Malier (1992) in connection with the French building code

$$f_{tj} = 0.6 + 0.06 f_{cj} \tag{1}$$

where f_{tj} and f_{cj} are average values of STS and CS, at j days (MPa). Similarly, another empirical formula has been given by CEB-FIP (1993) as follows

$$f_t = 0.301 (f'_c)^{0.67} \tag{2}$$

where f_t and denote f'_c STS and CS, respectively.

On the other hand, a recent study by Zain *et al.* (2002) suggested the equation (3) was suggested for predicting STS from CS and W/B ratio.

$$f_{ts} = 0.54\sqrt{f'_c}(W/B)^{-0.07} \tag{3}$$

3. Clustering-based fuzzy modelling

In this section, a data-driven fuzzy method, which is clustering based, by Takagi-Sugeno (TS) modelling (Takagi and Sugeno 1985) is presented. The method has three main steps; fuzzification (clustering), inference mechanism and defuzzification (prediction).

3.1. Fuzzy clustering

Fuzzy clustering is one of the methods used in fuzzification. Fuzzy c-means clustering (FCM) is a data clustering algorithm which employs fuzzy partitioning such that a given data point can belong to several groups with the degree of belongingness specified by membership grades between 0 and 1 (Jang *et al.* 1997). The objective function of this minimization algorithm is formulated by (Bezdek *et al.* 1984).

$$J_m(U, c) = \sum_{k=1}^{N}\sum_{i=1}^{c}(\mu_{ik})^m (d_{ik})^2 \tag{4}$$

where $m \in [1, \infty]$ is a weighting exponent, μ_{ik} is the membership of the kth data point in the ith class. The term, d_{ik} is a Euclidean distance measure (in 2-dimensional feature space, R^2) between the kth sample data x_k and ith cluster centre c_i, and is given by

$$d_{ik} = d(x_k - c_i) = \|x_k - c_i\| = \left[\sum_{j=1}^{2}(x_{kj} - c_{ij})^2\right]^{1/2} \tag{5}$$

Cluster means (prototypes) and elements of membership matrix are computed as follows.

$$c_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_k^m} \tag{6}$$

$$\mu_{ik} = \frac{1}{\sum_{p=1}^{c} \left(\frac{d_{ik}}{d_{pk}} \right)^{2/(m-1)}} \tag{7}$$

The FCM algorithm determines the cluster centres $\boldsymbol{c}_i$ and the membership matrix $\boldsymbol{U}$ using the following steps (Jantzen 2007):

1. Initialize the cluster centres $\mathbf{c}_i$ ($i = 1,2,\ldots,c$).
2. Determine the membership matrix $\boldsymbol{U}$ by (7).
3. Compute the objective function according to (4). Stop if either it is below a certain threshold level or its improvement over the previous iteration is below a certain tolerance.
4. Update the cluster centres according to (6).
5. Go to step 2.

3.2. Inference system

Assume that an unknown system $y = f(x)$ is observed. Observed data can be expressed by a deterministic function as $y = F(x)$. This function can serve as a reasonable approximation of $f(x)$. In fuzzy modelling, the function F is represented by fuzzy if-then rules (Setnes *et al.* 1998). Fuzzy inference system of TS model contains rules that include all possible fuzzy relations between inputs and outputs. Mathematically, the 1*st* order Takagi-Sugeno fuzzy model takes the following form

$$R_i{:}\,If\; x_1 \text{ is } A_{i1} \text{ and} \ldots x_n \text{ is } A_{in} \text{ then } t_i = a_i^T x + b_i \quad i = 1,2,\ldots,K \tag{8}$$

where $x = [x_1,\ldots,x_n]^T \in X$ is input (antecedent) vector, $A_{i1},\ldots,A_{in}$ are fuzzy sets defined in the input space, t_i; the rule output, R_i is the ith rule, and a_i and b_i; unknown but estimated parameters for each rule i. In the rule base, K denotes the number of rules.

3.3. Prediction

The last step of the TS fuzzy model is parameter estimation and prediction. The consequent parameters for each rule are provided by the least square estimation (Sousa and Kaymak 2002). Let X_e denote a matrix with its elements consisting of input values and a column vector with ones: $X_e=[X;\ \mathbf{1}]$ and Γ_i be a diagonal matrix in $\mathcal{R}^{N*N}$ having the normalized weighted memberships $\gamma_i(x_k) = \beta_i(x_k)/\sum_{j=1}^{K}\beta_j(x_k)$ as its kth diagonal element. The final expression

$$[a_i^T, b_i]^T = [X_e^T \Gamma_i X_e]^{-1} X_e^T \Gamma_i y \tag{9}$$

where a_i and b_i are regression coefficients. The aggregated output of the model is calculated by taking the average of the rule contributions (Babuska 1998)

$$t^* = \frac{\sum_{i=1}^{K} \beta_i(x) t_i}{\sum_{i=1}^{K} \beta_i(x)} \tag{10}$$

where $\beta_i(x)$ is the degree of activation of the ith rule given by

$$\beta_i(x) = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j), \qquad i = 1, 2, \ldots, K \tag{11}$$

and $\mu_{A_{ij}}(x_j)$: $\mathcal{R} \to [0, 1]$ is the membership function of the fuzzy set A_{ij} in the antecedent of R_i.

4. Experimental studies

This section presents the application of the clustering based fuzzy modelling given in Section 3 for the prediction of STS. The input variables are compressive strength and water/binder (W/B) ratio. For this application, age of concrete is selected as a fixed parameter (28 days). The general structure of the model is depicted in Fig. 2.

The data set used in this case study was taken from (Zain *et al.* 2002). The data set (22 experiments) was divided into two subsets randomly: training set (70%) and the validation set (30%), respectively. In the first step, data clustering has been carried out. It is often suggested that the data should be appropriately normalized before clustering (Jain and Dubes 1988). Samples derived from experimental data were standardized by using a linear transformation between minimum and maximum STS values [3.6, 6.0]. For this study, an optimal number of clusters was determined experimentally using a new index method which has been proposed by Tutmez *et al.* (2007). The adopted method aims to reproduce STS variability of the sample data in STS of cluster centres with a minimum number of clusters as follows

$$\text{Minimize } n_c \quad under \quad Std[t(x)] \approx Std[t(c)] \tag{12}$$

where n_c is the optimal number of cluster, *Std.* is the standard deviation of STS values. The appropriate numbers of clusters are two.

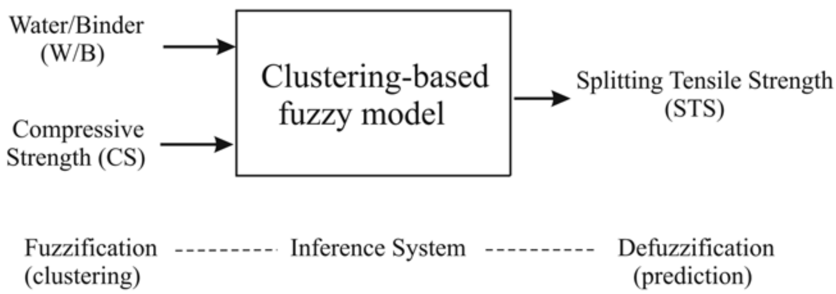

Fig. 2 Structure of the fuzzy model

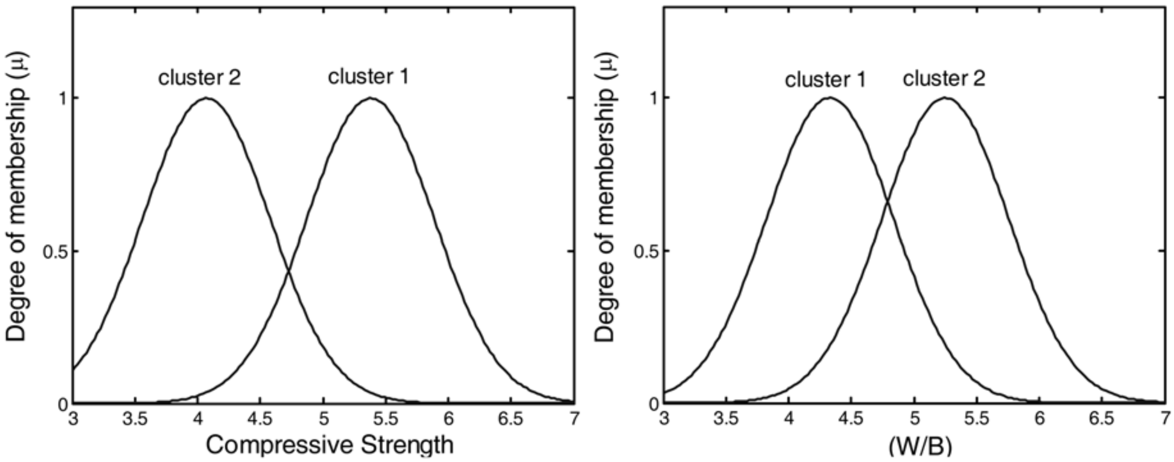

Fig. 3 Input membership functions

A cluster number is used in rule-based mechanism for defining the membership characteristics and parameter optimization. For this study, Gaussian type membership function was employed. The Gaussian functions facilitate smooth, continuously differentiable hypersurfaces of a fuzzy model (Piegat 2001). The memberships of the inference system were constructed as shown in Fig. 3. Rule mechanism and rule consequents are determined as follows:

Rule 1 : If x_1 coordinate is close to coordinate of c_{1X_1} and x_2 coordinate is close to coordinate of c_{1X_2}, then $STS = 4.820\ W/B - 9.934\ CS - 17.613$

Rule 2 : If x_1 coordinate is close to coordinate of c_{2X_1} and x_2 coordinate is close to coordinate of c_{2X_2}, then $STS = 11.087\ W/B - 5.104\ CS - 20.779$

where c_1-c_2 are the cluster centres. Each cluster is represented by a regression equation which is obtained from the weighted least squares estimation (see equation 9).

5. Results and discussion

The results of the fuzzy model have been presented in Tables 1 and 2 with the results of similar works (Zain et al. 2002; Song and Hwang 2004; Mindess and Young 1981). Figs. 4 and 5 show the performances of the fuzzy model and other approaches on the training and validation sets, respectively. As a performance index, r^2 denotes the coefficient of determination (COD). It is useful because it gives the proportion of fluctuation (variance) of one variable that is STS from the W/B and compressive strength. For the training set, $r^2 = 0.94$ means that 94% of the total variation in STS can be explained by TS fuzzy model.

In addition to COD, performances of the prediction models have been compared each other using the following performance indexes namely, the root mean square error (RMSE), and the variance account for (VAF).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^*)^2} \tag{13}$$

Table 1. Training set and predictions

Age of concrete (days)	W/B	Compressive strength (MPa)	Measured STS (MPa)	Zain *et al.*	French code	CEB/FIP code	Multiple Linear Regres.	TS fuzzy model
28	0.55	41.00	3.90	3.61	3.06	3.62	3.93	3.95
28	0.44	42.00	4.00	3.71	3.12	3.68	3.84	3.85
28	0.40	44.80	3.60	3.85	3.29	3.85	3.91	3.85
28	0.30	84.20	5.70	5.39	5.65	5.87	5.52	5.83
28	0.38	56.70	4.60	4.35	4.00	4.50	4.41	4.27
28	0.38	64.50	4.80	4.64	4.47	4.91	4.76	4.70
28	0.36	66.50	4.70	4.73	4.59	5.01	4.82	4.77
28	0.36	65.40	4.30	4.69	4.52	4.95	4.77	4.71
28	0.35	74.00	5.30	5.00	5.04	5.38	5.14	5.24
28	0.34	72.50	5.00	4.96	4.95	5.31	5.06	5.09
28	0.32	74.60	4.90	5.05	5.08	5.41	5.13	5.12
28	0.30	77.90	5.50	5.19	5.27	5.57	5.25	5.18
28	0.29	86.50	5.50	5.48	5.79	5.98	5.61	5.64
28	0.27	101.00	6.50	5.95	6.66	6.63	6.22	6.24
28	0.25	111.00	6.20	6.27	7.26	7.06	6.64	6.35

Table 2. Validation set and predictions

Age of concrete (days)	W/B	Compressive strength (MPa)	Measured STS (MPa)	Zain *et al.*	French code	CEB/FIP code	Multiple Linear Regres.	TS fuzzy model
28	0.22	118.00	6.20	6.52	7.68	7.36	6.91	6.17
28	0.35	73.80	4.90	4.99	5.03	5.37	5.13	5.01
28	0.25	94.50	5.80	5.79	6.27	6.34	5.91	5.87
28	0.30	79.10	5.20	5.22	5.36	5.63	5.30	5.44
28	0.40	56.20	4.40	4.32	3.97	4.48	4.41	4.38
28	0.28	102.00	5.50	5.96	6.72	6.67	6.28	5.58
28	0.30	83.30	5.90	5.36	5.60	5.83	5.48	5.44

$$VAF = \left(1 - \frac{\mathrm{var}(y - y^*)}{\mathrm{var}(y)}\right) x 100\% \tag{14}$$

where y_i is the measured, and y_i^* is the predicted STS values, respectively. Var denotes the variance and N is the number of experiments. Table 2 gives the RMSE and VAF measures for both the training and the validation data.

Figs. 4 and 5 and Tables 3 and 4 indicate that the fuzzy model is more accurate than the other prediction models used in this study**.** In addition to the numerical prediction power, the fuzzy model is more transparent and flexible in modelling the structure. The results also show that the traditional prediction models are influenced by high STS values. Note that for compressive strengths greater than 90 MPa, a large variation is observed with the CEB-FIP, French codes and regression model. The decreasing validation (testing) performances as seen in Fig. 5, may be related to randomly selected high CS values.

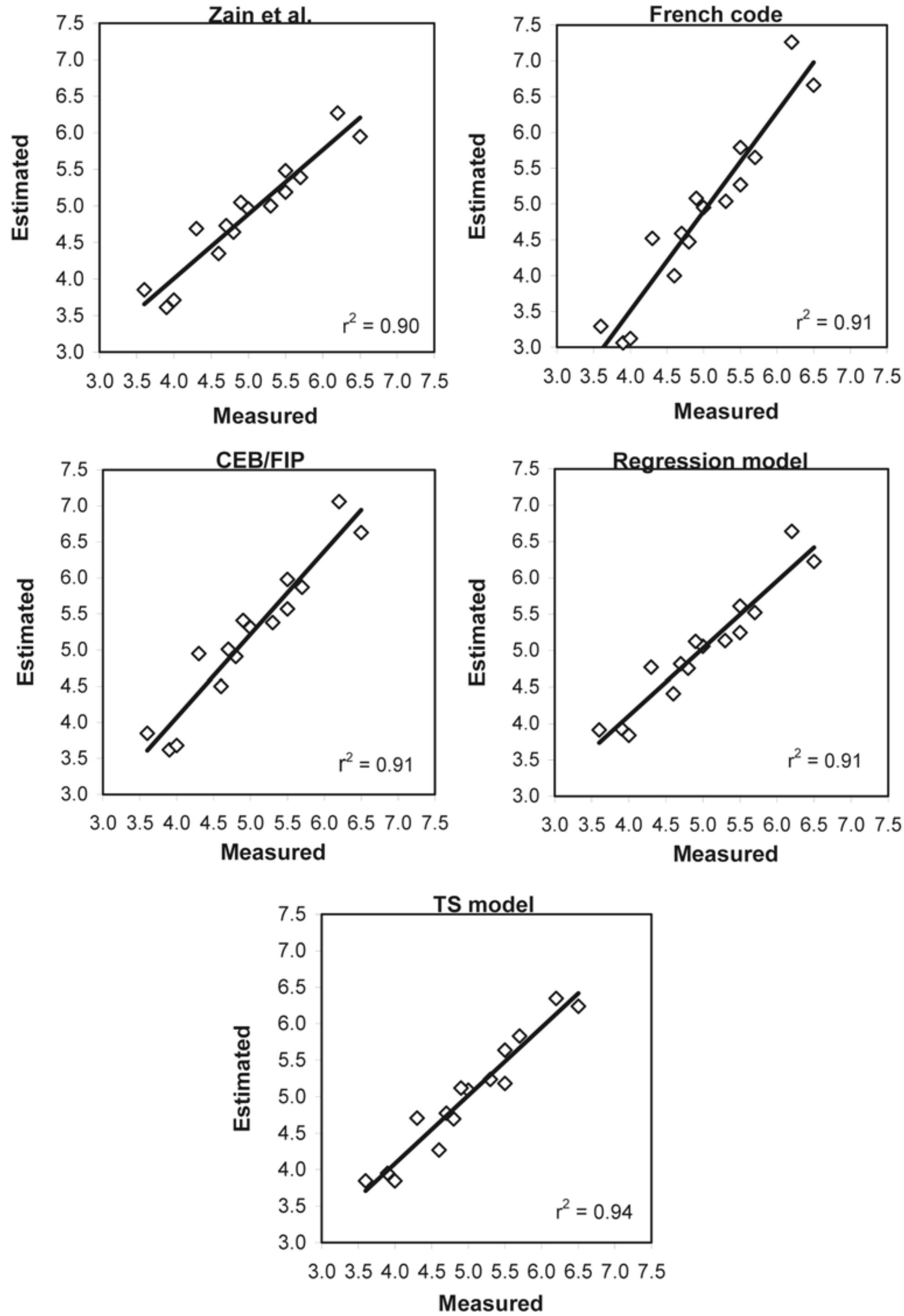

Fig. 4 Prediction models on standardized training set (70% of data)

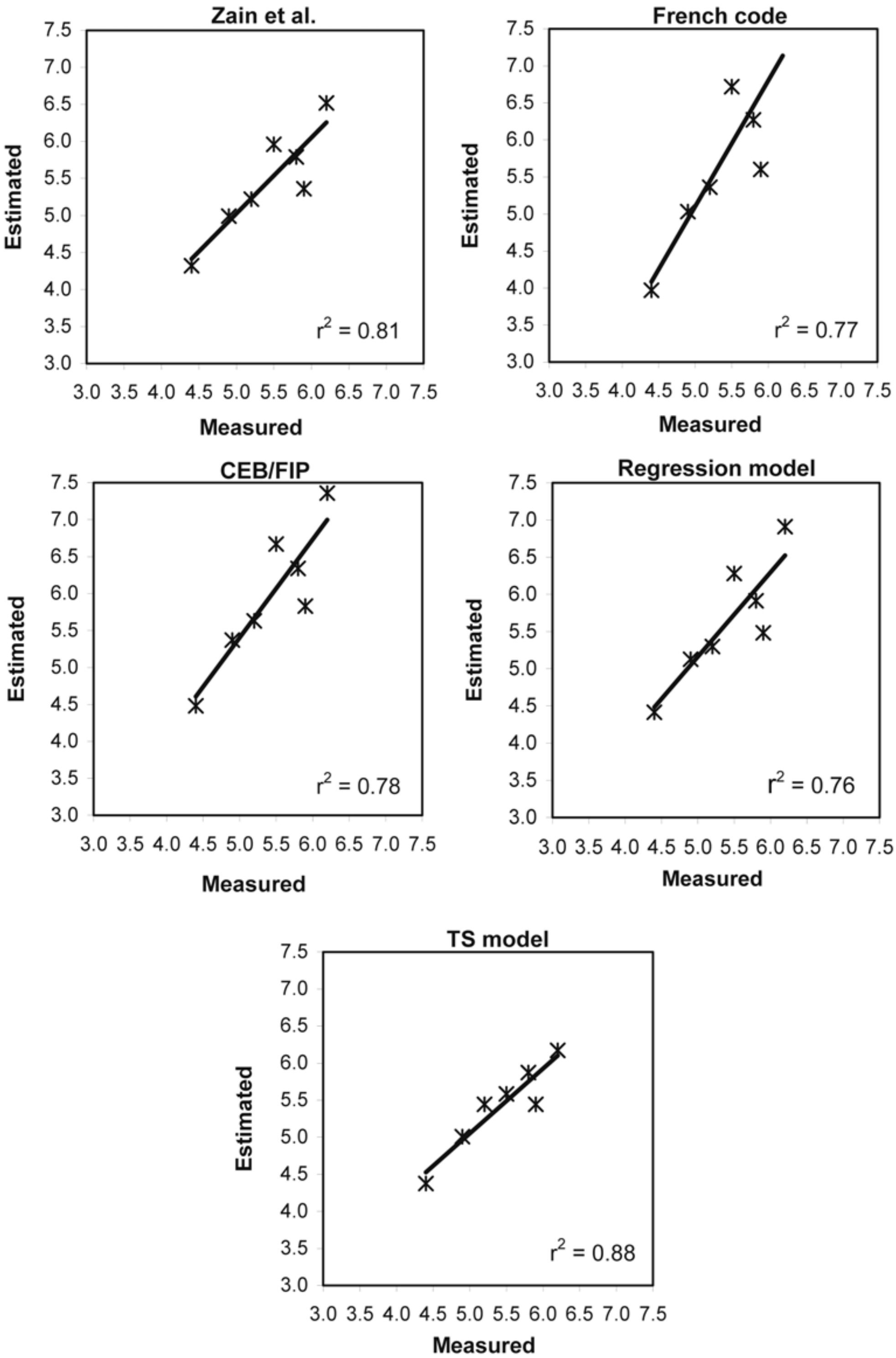

Fig. 5 Prediction models on standardized validation set (30% of data)

Table 3. Performance indices of the prediction models for training set

Method	VAF (%)	RMSE
Zain *et al.*	90.64	0.271
French code	58.29	0.483
CEB/FIP	84.85	0.381
Regression	90.3	0.241
TS fuzzy model	93.32	0.208

Table 4. Performance indices of the prediction models for validation set

Method	VAF (%)	RMSE
Zain *et al.*	74.10	0.297
French code	3.24	0.776
CEB/PIP	41.16	0.698
Regression	56.0	0.442
TS fuzzy model	87.52	0.206

6. Conclusions

A data driven fuzzy modelling approach has been employed for predicting STS (28-day HPC data) based on W/B and CS. It has been observed that performance of the clustering-based fuzzy model found to be better than other methods concerned within this study. The most prominent advantage of the fuzzy model is that it does not need to have well defined formulation of the system. The system can be analyzed and evaluated directly by the model from measured data.

The clustering-based fuzzy modelling approach can easily be extended in many construction and building domains. The fuzzy approaches can have a large impact in analyzing of mechanical properties of construction and building materials.

Acknowledgements

The author would like to extend his appreciation to anonymous referees and the editor Professor Christian Meyer due to their valuable comments and contributions.

References

Akkurt, S., Tayfur, G. and Can, S. (2004), "Fuzzy logic model for the prediction of cement compressive strength", *Cement Concrete Res.* **34**, 1429-1433.

Ayyub, B.M. and Klir, G.J. (2006), *Uncertainty Modelling and Analysis in Engineering and the Sciences. CRC Press*, Boca Raton.

Babuska, R. (1998), *Fuzzy Modelling for Control*, Kluwer Academic, USA.

Bezdek, J.C., Ehrlich, R. and Full, W. (1984), "FCM: the fuzzy c-means clustering algorithm", *Comput. Geosci.*, **10**(2-3), 191-203.

Carino, N.J. and Clifton, J.R. (1990), "Outline of a National plan on High-Performance Concrete: Report on the

NIST/ACI Workshop", *NISTIR 4465, Nat. Inst.of Stds. and Tech*., Gaithersburg, MD.

Choi, Y. and Yuan, R.L. (2005), "Experimental relationships between splitting tensile strength and compressive strength of GFRC and PFRC", *Cement Concrete Res.* **35**, 1587-1591.

Committee Euro-International du Beton (CEB-FIP) (1993), *CEB-FIP Model Code 1990*, London: Thomas Telford.

Jain, A. and Dubes, R. (1988), *Algorithms for Clustering Data,* Prentice Hall.

Jang, J-S.R., Sun, C.T. and Mizutani, E. (1997), *Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall International (UK), London.

Jantzen, J. (2007), *Foundations of Fuzzy Control*. John Wiley and Sons, Chichester.

Larrard, FDe. and Malier, Y. (1992), "Engineering properties of very high performance concrete", in: Y. Malier (ed.), *High Performance Concrete: From Material to Structure*, E&FN Spon, London, 85-114.

Meeks, K.W. and Carino, N.J. (1999), "Curing of High-Performance Concrete: Report of the State-of-the-Art. NISTIR 6295", *Nat. Inst. of Stds. and Tech*., Gaithersburg, MD.

Mindess, S. and Young, J.F. (1981), *Concrete,* Englewood Cliffs, NJ: Prentice-Hall.

Piegat, A. (2001), *Fuzzy Modelling and Control*, Physica-Verlag, Heidelberg.

Setnes, M., Babuska, R. and Verbruggen, H.B. (1998), "Transparent fuzzy modelling", *Int. J. Hum.-Comput. St.* **49**, 159-179.

Song, P.S. and Hwang, S. (2004), "Mechanical properties of high-strength steel fiber-reinforced concrete", *Constr. Build. Mater*., **18**, 669-673.

Sousa, J.M.C. and Kaymak, U. (2002), *Fuzzy Decision Making in Modelling and control*, World Scientific, Singapore.

Russell, H.G. (1999), "ACI Defines High-Performance Concrete", *Concrete Int.* **21**(2), 56-57.

Takagi, H. and Sugeno, M. (1985), "Fuzzy identification of systems and its applications to modelling and control", *IEEE T. Syst. Man. Cy.,* **15**, 116-132.

Tutmez, B., Tercan, A.E. and Kaymak, U. (2007), "Fuzzy modelling for reserve estimation based on spatial variability", *Math. Geol.,* **39**(1), 87-111.

Zain, M.F.M., Mahmud, H.B., Ilham, A. and Faizal, M. (2002), "Prediction of splitting tensile strength of high-performance concrete", *Cement Concrete Res.* **32**, 251-1258.

Zia, P., Leming, M.L., Ahmad, S.H., Schemmel, J.J., Elliott, R.P. and Naaman A.E. (1993), "Mechanical behavior of high-performance concretes", *Volume 1: Summary Report SHRP-C-361*, National Research Council, Washington.