

Improving the Prediction of Cement Compressive Strength by Coupling of Dynamical Models

D. Tsamatsoulis*

Halyps Building Materials S. A.,
17th Klm Nat. Rd. Athens – Corinth,
19300, Greece

doi: 10.15255/CABEQ.2015.2298

Original scientific paper
Received: September 15, 2015
Accepted: May 26, 2016

The dynamic approach of two well-known techniques has been used to predict a cement's 28-day compressive strength: Multiple linear regression (MLR), and artificial neural networks (ANN). The modeling is based on Portland cement data and utilizes daily physical, chemical analyses, and early strength results at days 1 and 7. Two kinds of models have been built, containing the 1-day strength as an independent variable, or both 1- and 7-day strength. The models are dynamic because they are applied to a movable past period of T_D days to calculate the parameters, and then used for a future period of T_F days. The comparison is based on the residual error of the testing period, and T_D , T_F have been optimized. Eight ANNs of different complexity have been developed, but some of them are suffering from over-fitting. A third model has also been created with the coupling of the initial two. The time parameters as well as the filtering and weighting coefficients of the coupled model have been optimized. The simple ANNs with one node in the hidden layer, sigmoid or hyperbolic functions and bias, show better performance. The combination of the coupled model with these two best ANN techniques provides an improved prediction of 28-day strength compared with the initial model containing the 1-day strength. The sensitivity also of T_F parameter is lower providing certain benefit in daily industrial application. The implementation of these methods in cement process control can contribute to quality improvement by maintaining a low variance of typical strength.

Key words:

cement, strength, neural networks, prediction, model, error

Introduction

Artificial neural networks (ANN) are an attractive tool for the modeling of non-linear processes and phenomena. In the cement industry, ANNs are mostly used to describe and control the main production operations: burning^{1–4} and grinding^{5–7}. Prediction of a cement's 28-day strength from earlier analysis results of the same sample, based on the quality control database for cement produced in the past, remains a challenging issue. Mainly linear and polynomial models have been developed, or algorithms that can be traced back to such models. Extensive reviews for these techniques can be found in literature^{8–9}. Neural networks have been used successfully for the prediction of concrete strength by building various ANN structures. Examples of such application of ANN in concrete are referred to in numerous studies^{10–17}. However, relatively few studies are based on ANN methodology or generally evolutionary and genetic algorithms (GA) that cor-

relate a cement's typical 28-day strength with other cement properties.

In 2011, Dolado *et al.*¹⁸ presented an excellent review of the recent efforts to describe cement-based materials by computational means. Their approach was to distinguish the models at several levels. They summarized sub-micro level simulations, as well as micro-, meso-, and macro-level models. They concluded that the trend to establish open platforms for modeling and for the use of simulation models illustrates the increasing demand for this type of tool for solving real engineering problems. Akkurt *et al.*¹⁹ developed a GA – ANN model of cement compressive strength by collecting and processing six months-worth of industrial data on the chemical, physical, and mechanical characteristics. Their results indicated that an increase in C_3S , SO_3 and specific surface leads to increased strength. Gene expression programming (GEP) and neural networks were used by Baykasoglu *et al.*²⁰ and Thamma *et al.*²¹ to predict the strength of Portland composite cement. Motamedi *et al.*²² predicted the compressive strength of cockle shell-cement-sand mixtures using and comparing the support vector

*Corresponding author: e-mail: d.tsamatsoulis@halyps.gr

regression (SVR) and adaptive –neuro-fuzzy inference (ANFIS) techniques. Firstly, the ANFIS network was used to find the parameters having a stronger impact on strength. Their findings showed that the capability of generalization could be improved by the ANFIS approach in comparison to the SVR estimation. Motamdi *et al.*²³ also applied the ANFIS technique to predict the strength of a pulverized fuel ash-cement-sand mixture. They concluded that ANFIS provided a suitable platform when the analysis was aimed at countering the uncertainties in a system. Zhang *et al.*²⁴ introduced an algorithm named Double-layer Multi-expression Programming (DMEP). They implemented the DMEP to the prediction of Portland cement's 28-day strength, and they compared the results with those of four other computational models, namely the Multi-Expression Programming model (MEP), Gene Expression Programming model (GEP), Neural Network model and Fuzzy Logic model (FL). Madsen *et al.*²⁵ applied FL and GA techniques to predict the strength of CEM I cement for all the strength classes. Ren *et al.*²⁶ applied generalized regression neural network (GRNN) techniques to predict the heat of hydration and compressive strength of cement. Yongzheng *et al.*²⁷ developed a model combining Principal Components Analysis (PCA) and ANN algorithms. The predictions were accurate in their field of application. If the value of a parameter that is not contained in the set of the independent variables changes noticeably during the cement production process, the predictive model could fail. Consequently, most of the models could be called “static”, since the parameters are estimated from a specified data set and the future strength is predicted.

Tsamatsoulis^{28–29} presented a comparison between static models and those with a movable time horizon, based on polynomial equations and long-term process results. The latter models incorporate the uncertainty due to the time variability of non-involved factors during the modeling, thus they are dynamic. The particularities of these two models have been explored thoroughly, and the superiority of the dynamic models have been proven. Apart from chemical and physical measurements, the results of early strength were also utilized, and two independent equations were applied for the prediction of 28-day strength: (i) The first where the 1-day strength constitutes an independent variable, named Str28_1. (ii) The second where the 7-day strength is also included, named Str28_7. This second model provides much better accuracy than the first one, but its bigger delay – more than 7 days – constitutes a severe drawback in daily practical application. Tsamatsoulis expanded his studies³⁰ by developing neural network structures incorporating the results

of 1-day strength, as well as physical and chemical characteristics. The aim of this study is dual. Firstly, to compare two types of dynamical techniques: The multiple linear regression (MLR) with a technique based on several types of ANNs. Both techniques use data of a predetermined period for training, namely for computing of optimal parameters. Afterwards the data belonging to a time interval that follows the training period are used for validating the model. The training and validation time interval constitute the past and future period, respectively. Throughout this text, the errors corresponding to past periods are characterized as training errors, while the ones computed from future period data are named test errors. The criterion of comparison is the capability of a model to predict the cement's future strength that can be characterized as the generalization ability³¹ of modeling. The time intervals of training and validation need optimization to achieve the best prediction. The minimal test errors of MLR and ANNs are compared in order to select the best technique for further examination. Special care has been given to the over-fitting problem that frequently appears in ANN techniques. Subirats *et al.*³² clearly stated that one of the strategies of avoiding over-fitting is the search of compact architectures. Other popular techniques include early stopping of training using a validation data set, weight decay, and exclusion of noisy instances from training data. The MLR and ANN techniques are implemented to data sets corresponding to the Str28_1 and Str28_7 models mentioned earlier. The second main objective of this study is to couple dynamically the predictions of the model based on the 7-day early strength with the ones of the model based on the 1-day strength, aiming to utilize the prior information derived from the Str28_7 model in order to improve the predictability after the 1-day strength has been measured. In this case, a more precise prediction of strength leads to an improved quality control of the production process. The study is restricted to Portland composite cement types.

Materials and testing methods

Two Portland cement types, produced according to EN 197-1:2011³³ were studied: CEM II A-L 42.5 N and CEM II B-M (P-L) 32.5 N. Each common cement type necessarily contains clinker and gypsum. The first cement type contains limestone as the main component, while the main components of the second type are natural pozzolane and limestone. The typical chemical characteristics of the raw materials are shown in Table 1. The oxides analysis was performed with XRF. The typical mineral composition of clinker is also presented according to Bogue equations (1)–(4). The subsequent no-

Table 1 – Materials chemical analyses and clinker mineral composition

	LOI	SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	CaO	MgO	SO ₃	K ₂ O	Na ₂ O	Ins. Res.
	%	%	%	%	%	%	%	%	%	%
Clinker	0.30	21.45	5.10	3.30	65.67	1.93	1.20	0.66	0.26	0.30
Gypsum	21.0	1.42	0.32	0.18	34.84	1.64	39.91	0.11	0.02	
Pozzolane	3.23	74.43	12.59	1.31	1.61	0.32	0.01	3.12	3.70	84.23
Limestone	42.70	0.61	0.18	0.10	54.22	1.24	0.02	0.01	0.01	
		C ₃ S		C ₂ S		C ₃ A		C ₄ AF		
		%		%		%		%		
Clinker		58.8		17.2		7.9		10.0		

Table 2 – EN 197-1:2011 requirements for cement compositions and 28 days strength

Type	Clinker	Limestone	Pozzolane	Minor constituents	28 days strength low limit	28 days strength high limit
	%	%	%	%	MPa	MPa
CEM II A-L 42.5 N	80–94	16–20		0–5	42.5	62.5
CEM II B-M (P-L) 32.5 N	65–79		21–35	0–5	32.5	52.5

tation is used for these formulae: $C = \text{CaO}$, $S = \text{SiO}_2$, $A = \text{Al}_2\text{O}_3$, $F = \text{Fe}_2\text{O}_3$. With CaO_f the clinker free lime is denoted.

$$C_3S = 4.07 \cdot (C - \text{CaO}_f) - 7.6 \cdot S - 6.72 \cdot A - 1.43 \cdot F - 2.85 \cdot \text{SO}_3 \quad (1)$$

$$C_2S = 2.87 \cdot S - 0.754 \cdot C_3S \quad (2)$$

$$C_3A = 2.65 \cdot A - 1.69 \cdot F \quad (3)$$

$$C_4AF = 3.04 \cdot F \quad (4)$$

The EN 197-1:2011 requirements as regards composition and 28-day strength limits are shown in Table 2. The modeling is based on the results of daily average samples obtained from the corresponding chemical analyses, physical, and mechanical measurements. The following data were utilized: (i) Residue at 40 μm sieve, measured with air jet sieving according to EN 196-6³⁴; (ii) Specific surface, measured according to EN 196-6; (iii) Loss on ignition and insoluble residue of the cement, measured according to EN 196-2³⁵; (iv) SO_3 , measured with XRF; (v) Compressive strength at 1, 7 and 28 days. The preparation, curing, and measuring of the specimens were made according to EN 196-1³⁶. The typical chemical, physical, and mechanical properties of the two cement types are depicted in Table 3. The modeling is based on more than 3400 sets of data corresponding to nine years' daily production results of Halys plant.

Table 3 – Average values of cements characteristics

	CEM II A-L 42.5 N	CEM II B-M (P-L) 32.5 N
LOI (%)	7.3	11.5
SO ₃ (%)	2.9	2.5
Insol. Res. (%)	0.5	6.9
R40 (%)	10.8	11.1
Spec. surf. ($\text{cm}^2 \text{q}^{-1}$)	3440	4040
Str ₁ (MPa)	10.5	7.7
Str ₇ (MPa)	36.1	27.4
Str ₂₈ (MPa)	49.8	38.0

Mathematical models predicting strength

The common independent variables in all models are: Loss on ignition, LOI , sulfates content, SO_3 , insoluble residue, Ins_Res , residue at 40 microns sieve, $R40$ and specific surface, S_b . The reason for using chemical analysis instead of the composition of cement, as was done in earlier modeling, is the fullest generalization of generated equations. The direct use of chemical analysis does not need prior knowledge of raw materials and clinker composition. Two basic and independent models initially applied to the prediction of 28-day strength: (i) The one named Str_{28_1} , where the 1-day strength – Str_I – constitutes an input variable, except the set of physical and chemical data. (ii) The second one

named Str_{28_7} , where the 7-day strength variable $-Str_{7-}$ is also included. In fact, weaker or stronger functions exist among the variables LOI , Ins_Res , SO_3 , $R40$, S_b , Str_{1-} , Str_{7-} : (a) A decrease in $R40$ or increase in LOI , Ins_Res and SO_3 generally causes an increase in S_b . Specific surface is also a function of clinker and raw materials grindability, condition of cement mill grinding media, and mill operation settings. (b) An increase in LOI , $Insol_Res$ and $R40$ leads to lower early strength. However, the early strength depends also on the clinker mineral composition, free lime, equivalent alkalis, and clinker activity in general. (c) An increase in Str_{1-} causes also an increase in Str_{7-} , but Str_{7-} is also related to clinker composition and activity. Thus, because the variables S_b , Str_{1-} , Str_{7-} depend on parameters not taken into account in the current state of modeling, they can be considered as independent, since they contain an “independent” part. In case of a model correlating Str_{28} with only physical and chemical properties, excluding the early strength, the model error is higher because of all the non-included variables. A coupling of these models was attempted by a third model containing all the independent variables of Str_{28_1} and the addition of a correction computed from the moving difference between the calculated 28-day strength from Str_{28_7} and the measured one. The model was named Str_{28_EW} .

Multiple linear regression

The seven independent variables are named X_I with $I=1$ to 7, where: $X_1=LOI$, $X_2=SO_3$, $X_3=Ins_Res$, $X_4=S_b$, $X_5=R40$, $X_6=Str_{1-}$, $X_7=Str_{7-}$. The 28-day strength is a dependent variable, $Y=Str_{28}$ and then algorithm goes on as follows:

(i) For a given data set, the minimum and maximum values of X_I and Y , $X_{I,MIN}$, $X_{I,MAX}$, Y_{MIN} , Y_{MAX} respectively, are computed.

(ii) The variables X_p , Y are normalized. The set of the new variables XN_p , YN is calculated from equations (5) and (6). The normalized data belong to the interval $[0, 1]$.

$$XN_I = \frac{X_I - X_{I,MIN}}{X_{I,MAX} - X_{I,MIN}} \quad I=1..N \quad (5)$$

$$YN = \frac{Y - Y_{MIN}}{Y_{MAX} - Y_{MIN}} \quad (6)$$

(iii) The normalized 28-day strength, YN , is given by the formula (7).

$$YN = A_0 + \sum_{I=1}^N A_I \cdot XN_I \quad (7)$$

(iv) The calculated 28-day strength, Str_{28_Calc} is provided by equation (8).

$$Str_{28_Calc} = Y_{MIN} + YN \cdot (Y_{MAX} - Y_{MIN}) \quad (8)$$

(v) For a total number of data sets equal to M and actual strength Str_{28_Act} , the coefficients A_p , $I=0$ to N , are computed by minimizing the residual error s_{Res} which is calculated from the formula (9).

$$s_{Res}^2 = \sum_{J=1}^M (Str_{28_Calc,J} - Str_{28_Act,J})^2 / (M - N - 1) \quad (9)$$

Neural networks

Two main kinds of neural networks were developed: The usual feed-forward ANN with three layers, and the more complicated cascade ANN. The back propagation method is applied in batch mode. The hidden layer of the ANN with three layers contains one or two nodes. The non-linearity of the activation function is approached using the sigmoid, hyperbolic and radial basis functions. The modeling also involves ANNs with and without bias. The result of the combinations is an elevated number of structures with nomenclature presented in Table 4. The ANNs with three layers and one or two nodes in the hidden layer are depicted in Figure 1 as concerns the model St_{28_1} . The model Str_{28_7} includes an additional node in the first layer for the reception of Str_{7-} , thus it needs one or more additional synaptic weights, depending on the number of nodes in the hidden layer. The ANN with one node in hidden layer accepts and processes all the data. In the two nodes case, each one takes as inputs the data of each cement type (CEM B-M 32.5 and CEM A-L 42.5). The software identifies the cement type from the chemical analyses. At each node the linear combination of the inputs and of synaptic weights is performed and the result enters to the activation function. In the case of sigmoid or radial basis functions, equations (5) and (6) are applied for normalization. When hyperbolic tangent function applies, the normalization is made according to formulae (10) and (11). In this case, XN_p , YN belong to the interval $[-1, 1]$. Then Str_{28} is back calculated from its normalized value by applying equation (12).

$$XN_I = \frac{2 \cdot X_I - (X_{I,MIN} + X_{I,MAX})}{X_{I,MAX} - X_{I,MIN}} \quad I=1..N \quad (10)$$

$$YN = \frac{2 \cdot Y - (Y_{MIN} + Y_{MAX})}{Y_{MAX} - Y_{MIN}} \quad (11)$$

$$Str_{28_Calc} = \frac{Y_{MAX} + Y_{MIN}}{2} + YN \cdot \frac{Y_{MAX} - Y_{MIN}}{2} \quad (12)$$

The activation functions are described by equations (13) to (16).

Table 4 – Description of ANNs structure

Three Layers ANN				
	Number of nodes in hidden layer	Number of parameters	Activation function	Bias
S_1N	1	7	sigmoid	NO
S_1N_B	1	8	sigmoid	YES
HT_1N	1	7	hyperbolic tangent	NO
HT_1N_B	1	8	hyperbolic tangent	YES
RBF_1N	1	13	radial basis function	NO
S_2N	2	14	sigmoid	NO
RBF_2N	2	20	radial basis function	NO

Cascade ANN				
	Number of hidden layers	Number of parameters	Activation function	Bias
CASC	4	12	sigmoid	NO

Sigmoid function

$$o(J) = 1 / \left(1 + \exp \left(- \sum_{I=0}^N W_{I,J} \cdot XN_I \right) \right) \quad (13)$$

where $o(J)$ is the output of the node J . When the hidden layer contains one node, $J = 1$, while when two nodes exist, J is either 1 or 2.

Hyperbolic tangent function

$$o(J) = \left(1 - \exp \left(-2 \cdot \sum_{I=0}^N W_{I,J} \cdot XN_I \right) \right) / \left(1 + \exp \left(-2 \cdot \sum_{I=0}^N W_{I,J} \cdot XN_I \right) \right) \quad (14)$$

Radial Basis function

$$o(J) = \exp \left(- \sum_{I=1}^N (XN_I - X0_I)^2 / \sigma_I^2 \right) \quad (15)$$

where $(X0_1, X0_2, \dots, X0_N)$ is the vector of the center of the radial basis function, while $(\sigma_0, \sigma_1, \dots, \sigma_N)$ are variance parameters.

Output layer activation function

$$YN = \sum_{J=1}^{N_I} V_J \cdot o(J) \quad (16)$$

where $N_I = 1$ or 2 depending on the number of nodes in the hidden layer.

The more complicated structure of the cascade ANN is depicted in Figure 2. Cascade ANN has been applied for the prediction of concrete strength by Badde *et al.*³⁷ To develop this kind of ANN, the technique described by Shetinin³⁸ has been followed. All the hidden layers contain sigmoid functions, while the output layer combines linearly the outputs of each hidden layer with the weights V_p , $I = 1..4$. In the first hidden layer Str_1 and LOI , the two most significant inputs as regards their impact on Str_{28} , are fed. Their linear combination with W_{11} , W_{21} follows. The output $o(I)$ is multiplied with the weight V_1 and is fed to the output node. These three synaptic

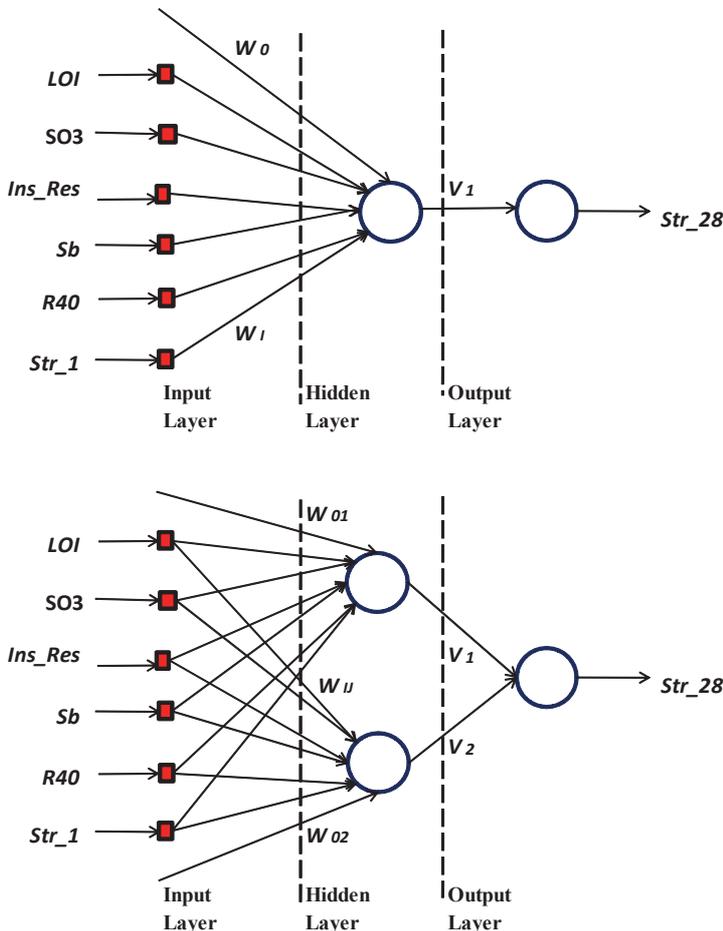


Fig. 1 – Three-layered ANN with one and two nodes in the hidden layer

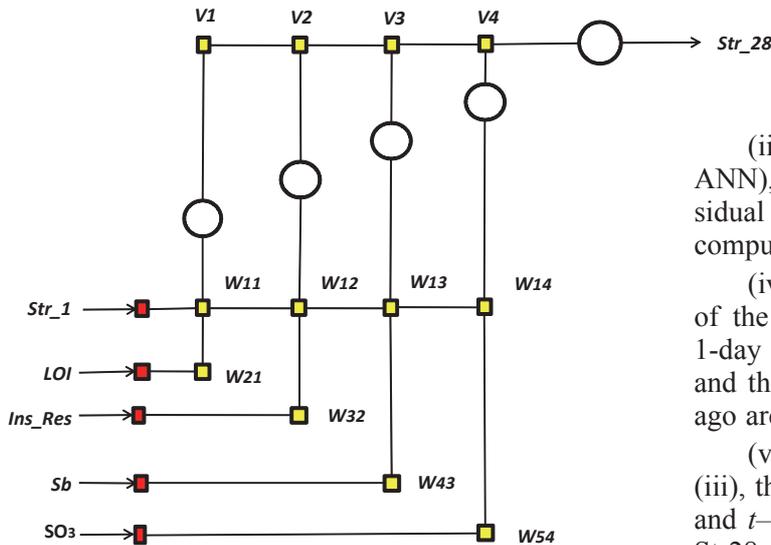


Fig. 2 – Cascade ANN with four hidden layers

weights are trained in batch mode and then tested. The second step involves a new layer where Str_1 and Ins_Res enter. The weights W_{12} , W_{32} , V_1 , V_2 are trained and tested as previously. The construction of new, hidden layers follows by adding each time Str_1 and one of the remaining variables. The algorithm stops when the addition of a new variable does not decrease the training error further, which means that the ANN is probably over-fitted. Applying this algorithmic logic, $R40$ was not added to the cascade ANN, as it causes worsening of the test error. An explanation is the expected descending function between $R40$ and S_b . Due to the fundamental impact that the specific surface has on the reaction rate of cement hydration, it is concluded that S_b is more significant compared to $R40$ at least for this ANN type.

Dynamic modeling

The common feature among the linear regression and neural network techniques is the dynamic modeling, which is described by the following algorithm:

(i) At date t , a new 28-day strength result occurs. The specimen was prepared 28 days ago. The production date is in distance $t-29$ days from the current date t .

(ii) A time interval of T_D days and the samples belonging to the period $[t-29-T_D, t-29]$ are considered. The dynamic data set contains the results of this population of daily average samples, i.e. all the daily results of LOI , Ins_Res , SO_3 , $R40$, S_b , Str_1 , Str_7 , Str_28 of both cement types for cement produced in each cement mill. An example of a data set is provided in Figure 3, for $T_D = 180$ days.

(iii) Using the selected technique (MLR or ANN), the sets of parameters that minimize the residual errors of models Str_28_1 and Str_28_7 are computed.

(iv) At day t , the chemical and physical results of the cement produced on the previous day, the 1-day strength of the cement produced 2 days ago, and the 7-day strength of cement produced 8 days ago are measured.

(v) With the set of parameters computed in step (iii), the 28-day strength of cement produced at $t-2$ and $t-8$ days is estimated, by applying the models $Str28_1$ and $Str28_7$, respectively.

(vi) Steps (iv), (v) are repeated for all the consecutive days up to $t+T_F-1$, where T_F is a predetermined time interval, without new parameters estimation.

(vii) According to step (vi), for the days belonging to the interval $[t, t+T_F-1]$, the future strength of the cement produced in the time intervals $[t-2, t+T_F-3]$, $[t-8, t+T_F-9]$ is computed according to the equation of step (iii). If the date becomes greater than $t+T_F-1$, a new parameter estimate is carried out starting from step (i), considering as initial time, the first date that is greater than $t+T_F-1$.

(viii) When the results of T_F days have been completed, the time interval of T_D length moves forward by T_F days. Thus, the future 28-day strength is calculated using models applied to movable data sets of time span T_D with step of length T_F . The above means that, when the data belonging to the interval $[t-29, t+T_F-29]$ are added, the data contained in the first T_F days of the interval $[t-29-T_D, t-29]$ are subtracted.

(ix) Parameters T_D and T_F shall be optimized considering the following two criteria: (a) minimum $MSRE_{Past}$ during modeling, and (b) minimum error $MSRE_{Futur}$ during future application of the models.

(x) For each model, each T_D and T_F , and for each past and future time interval, a set $(A_p, MSRE_{Past}, MSRE_{Futur})$ is computed from the samples belonging to this interval. A Newton-Raphson non-linear regression method has been used to determine the values of parameters A_i which minimize $s_{Res,TD}(I)$. Then the model with this set of parameters applies to the data of T_F length to obtain $s_{Res,TF}(I)$.

The number of the consecutive sets $(A_p, MSRE_{Past}, MSRE_{Futur})$ is K_{TD} and it is a function of T_D and T_F values. The mean square residual errors $MSRE_{Past}$, $MSRE_{Futur}$ are calculated by equations (17) and (18), respectively.

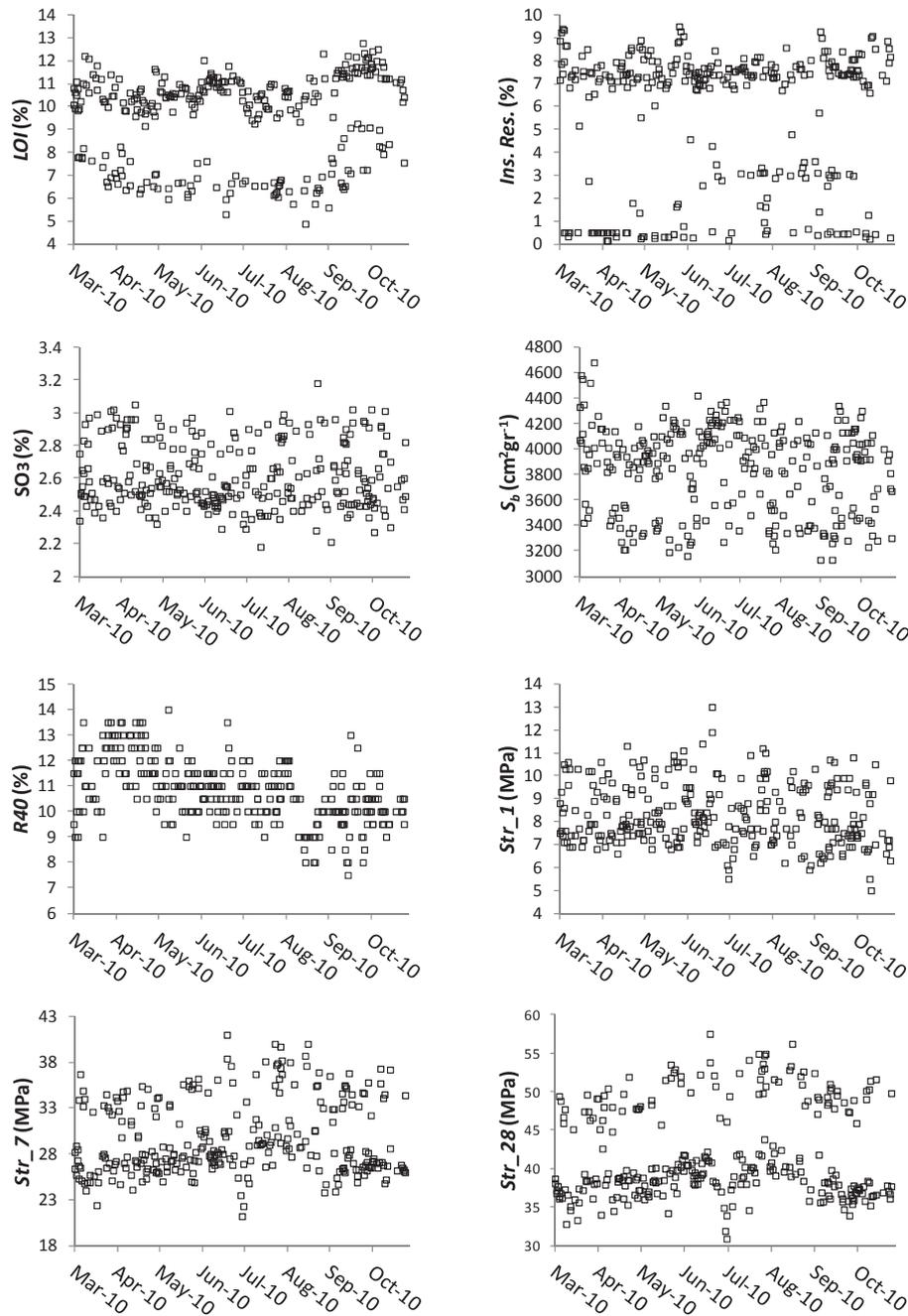


Fig. 3 – Example of data set for $T_D = 180$ days

$$MSRE_{Past} = \sqrt{\frac{\sum_{I=1}^{K_{TD}} S_{Res,TD}(I)^2}{K_{TD}}} \quad (17)$$

$$MSRE_{Future} = \sqrt{\frac{\sum_{TD} S_{Res TF}(I)}{TD}} \quad (18)$$

Coupling of models

The coupling of models Str_28_1 and Str_28_7 makes use of the exponentially weighted moving

average (EWMA) filter³⁹. For a variable X and discrete time I in days, the EWMA variable Y is defined by the procedure:

(i) For time $I = 0$, the initial moving average $Y(0)$ is expressed by the relation (19):

$$Y(0) = X(0) \quad (19)$$

(ii) For a parameter λ , where $0 < \lambda \leq 1$, the statistic $Y(I)$ is computed by the recursive formula (20).

$$Y(I) = \lambda \cdot X(I) + (1 - \lambda) \cdot Y(I - 1) \quad (20)$$

(iii) If $\lambda = 1$, the value of moving average becomes equal to the current one. The smaller the λ value is, the lower the rate of change becomes, and trends of longer duration can be revealed.

It is supposed that, at time I , the actual 28-day strength is $Str_28(I)$, and the computed one from Str_28_7 is $Str_28_7(I)$. The difference $Diff(I)$ is defined by formula (21).

$$Diff(I) = Str_28_7(I) - Str_28(I) \quad (21)$$

Afterwards, the following procedure applies:

(i) The moving average of $Diff(I)$, $EW_Diff(I)$, is calculated by applying equation (20) for a pre-defined value of λ .

(ii) The corrected value $Str_28_1(J)$, named $Str_28_EW(J)$, is computed from equation (22), where $J \geq I$.

$$Str_28_EW(J) = Str_28_1(J) + k \cdot EW_Diff(J) \quad (22)$$

(iii) The parameters k and λ of the coupled model need optimization using the mean square residual error ($MSRE$) as criterion. For each CEM type, different parameters are used, the set of (k_{BM}, λ_{BM}) for CEM II B-M 32.5, and the set (k_{AL}, λ_{AL}) for CEM II A-L 42.5.

Analysis of results and discussion

Preliminary processing of data

To examine the correlations among all of the input variables and Str_28 , the correlation coefficients were computed. The results are presented in Table 5. These correlations should be inspected with some criticism because some of them give a false signal. Therefore, the data have to be examined from a physical stand-point. From the typical characteristics of the cements shown in Table 3 and limits of composition and strength depicted in Table 2, the subsequent remarks can be made: (a) CEM II A-L 42.5 N has higher average SO_3 value compared with CEM II B-M (P-L) 32.5 N. (b) Due to the higher insoluble residue, the second cement contained pozzolan, which was not detected in the first. (c) The lower LOI value of the first cement means lower limestone content compared with the second. The negative correlation between Str_28 and S_b is false because limestone, pozzolane, and gypsum are materials of high grindability, consequently, in the

two CEM types the strength and specific surface were found in reverse order. The positive correlation between Str_28 and SO_3 is also doubtful and needs deeper investigation. A low correlation was found between 28-day strength and $R40$. The reason being, that $R40$ is similar for both cements because it is a process variable regulated by the separator of the cement mill, while strength differs considerably due to the different clinker content of each cement. All the above considerations need a deeper investigation. Strong correlations exist between Str_28 and LOI , Ins_Res , Str_1 , Str_7 . Early strength is not only a function of physical and chemical characteristics, but also of the clinker activity and grinding conditions. In the case of a model that relates 28-day strength only with physical and chemical properties of the cement, the modeling error is higher, because all other independent variables act as noise. Tsamatsoulis²⁹ performed such a comparison by developing second-degree polynomial models applied to a subset of data utilized in this study.

A detailed search of the shape of the functions between Str_28 and input variables follows based on the processing of results of two years. The reason for performing the analysis to a subset of the total population is to reduce the impact of independent variables not included. For early strength – Str_1 and Str_7 – the interval of their minimum and maximum values is partitioned into N equal intervals. In each sub-interval, the average values of early strength and Str_28 are computed and plotted in Figure 4. A noticeable non-linearity occurs in the function between Str_1 and Str_28 and the gain is higher for lower values of early strength. This is mainly attributed to the pozzolanic action as concerns the CEM II B-M cement. In CEM II A-L, an increase in early strength is mainly achieved by the higher fineness, but this characteristic contributes less to the strength development. A similar but more linear trend exists between Str_7 and Str_28 . Correlations between $\{SO_3, R_{40}, S_b\}$ and 28-day strength was studied by first separating the results of each CEM type using LOI and Ins_Res as criteria. Then the procedure implemented for early strength is followed for each independent variable and cement type. The trends are depicted in Figure 5. As concerns CEM B-M 32.5, the function between sulfates and Str_28 presents a clear maximum for SO_3 e [2.4, 2.6]. This result agrees with a laboratory study of Tsamatsoulis *et al.*⁴⁰ for the same clinker and raw materials, where parabolic equations were used to express this correlation. For the CEM 42.5, the function between the two variables is ascending, meaning that the results are located in the left branch of the parabola. The function between residue at 40 microns and Str_28 is descending. The rate of strength reduction increases as $R40$ aug-

Table 5 – Correlation coefficients between the output and input variables

	LOI	SO_3	Ins_Res	S_b	$R40$	Str_1	Str_7
Str_28	-0.86	0.68	-0.85	-0.74	-0.16	0.82	0.95

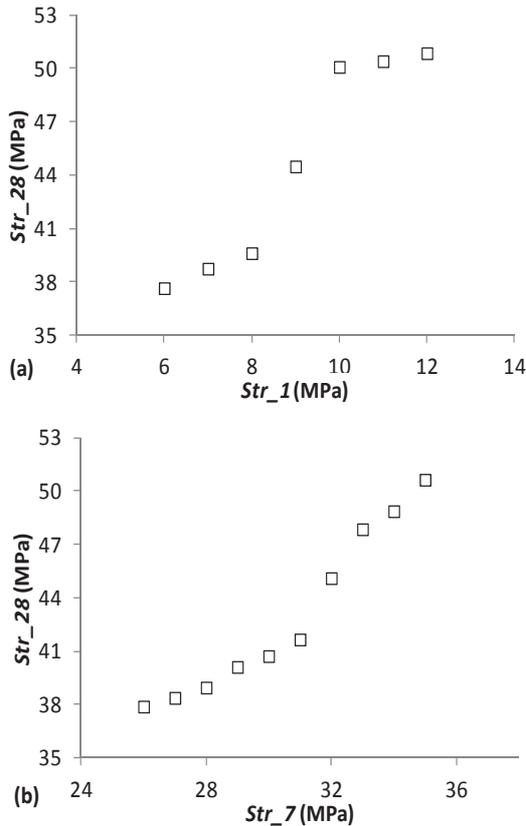


Fig. 4 – Effects of early strength on Str_{28}

ments. In Figure 5(c), constantly higher specific surface is observed in CEM II B – M compared with that of CEM II A-L. This is due to the higher content of softer materials in the first cement. The function between specific surface and Str_{28} passes from maximum value for both cements. Although this result seems to be in contradiction with the principle that the finer cement has higher strength, the explanation is as follows: In the left branch of each curve as S_b increases, the specific surface of the clinker also increases, leading to greater cement strength. A further increase in S_b is caused by the higher content of soft materials, leading into a strength reduction.

The negative correlation coefficients of LOI and Ins_Res with str_{28} denote that the examination of impact of each variable on the strength needs some classification of the results. The above is obtained with generation of two families of curves: (i) The data are classified into five groups with low variance of insoluble residue. The upper and lower limits of Ins_Res are (0, 2), (2, 4), (4, 6), (6, 8) and >8 %. For each group of data the average LOI and Str_{28} are calculated and plotted in Figure 6(a). (ii) Regarding LOI , the classification is made into five groups with limits (5, 8), (8, 9.5), (9.5, 11), (11, 12.5) and >12.5 %. The average Ins_Res and Str_{28} are found and plotted in Figure 6(b). An increase in LOI causes a decrease in Str_{28} , as may be ob-

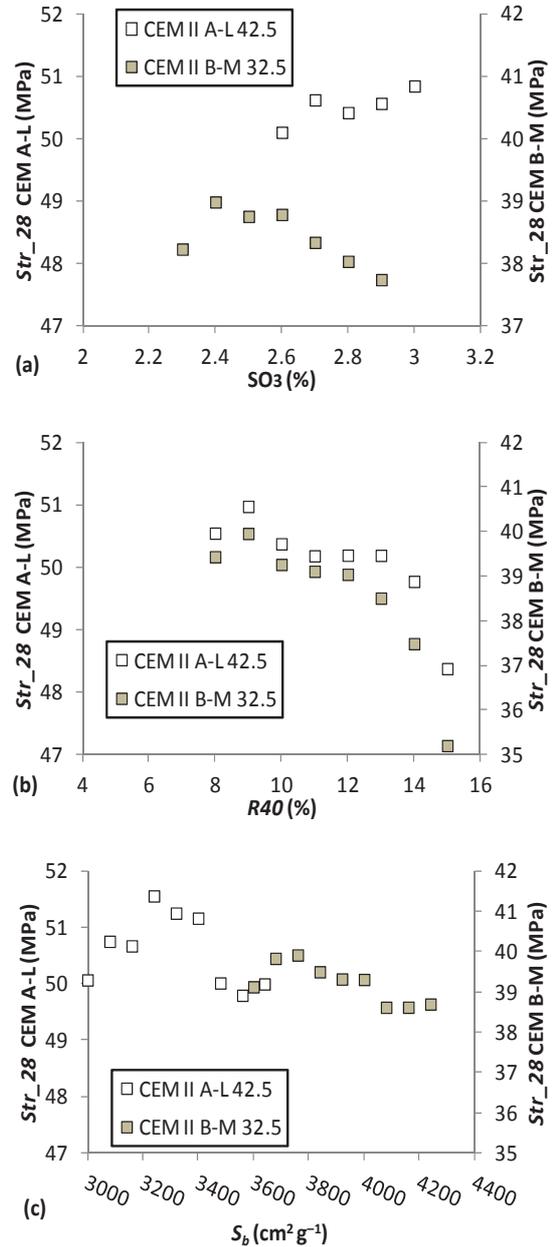


Fig. 5 – Effects of SO_3 , S_b and $R40$ on Str_{28}

served from the first Figure for all the Ins_Res groups. The slope of reduction is greater at higher LOI values. A similar trend occurs in the function between Ins_Res and Str_{28} , but the gradient of strength reduction is smaller. Especially for Ins_Res between 6 % and 9 %, the average decrease in strength is fairly low, demonstrating the positive effect of pozzolane addition on 28-day strength.

Residual errors of Str_{28_1} and Str_{28_7} models

The dynamical models initially apply to model Str_{28_1} for a movable training period of $T_D = 180$ days and the corresponding $MSRE_{past}$ are computed. Then the parameters of each model apply to the results of the next T_F days, which constitute the test-

ing set. This period starts at least 29 days after the last date of the training period. The calculated and actual results of 28-day strength during the testing period are compared. Thus, the errors $MSRE_{Future}$ are determined. The training and test errors, $MSRE_{Past}$ and $MSRE_{Future}$, of Str_28_1 for MLR and all ANN techniques for $T_D = 180$ and T_F ranging from 1 to 60 days are shown in Figure 7. Because the training

errors of each model are mainly the function of T_D , they remain almost constant for T_F ranging from 1 to 60 days. As concerns the test errors, a decrease in T_F generally causes a decrease in $MSRE_{Future}$. Especially for $T_F \geq 20$ days, a worsening of $MSRE_{Future}$ is observed, meaning that the ability of ANNs or MLR to predict future results is reduced. This result is attributed to two main causes: (a) Some of the param-

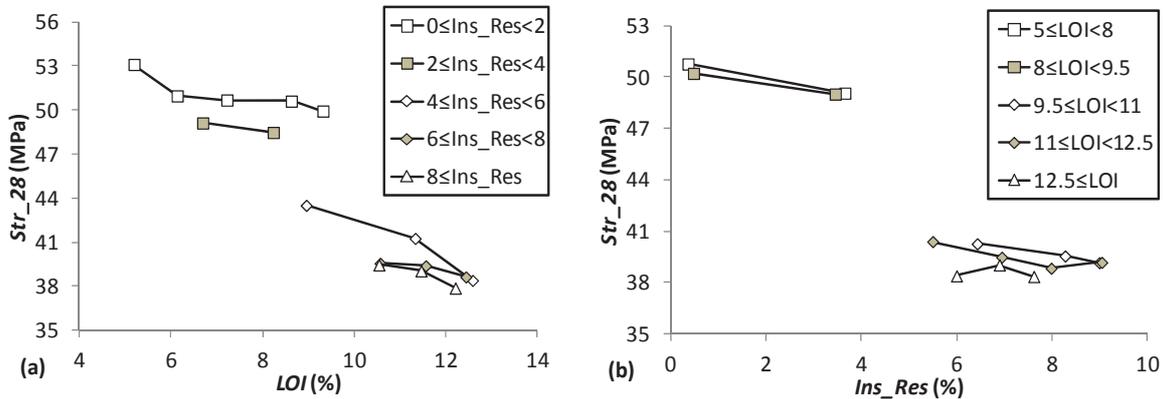


Fig. 6 – Effects of LOI and Ins_Res on Str_28

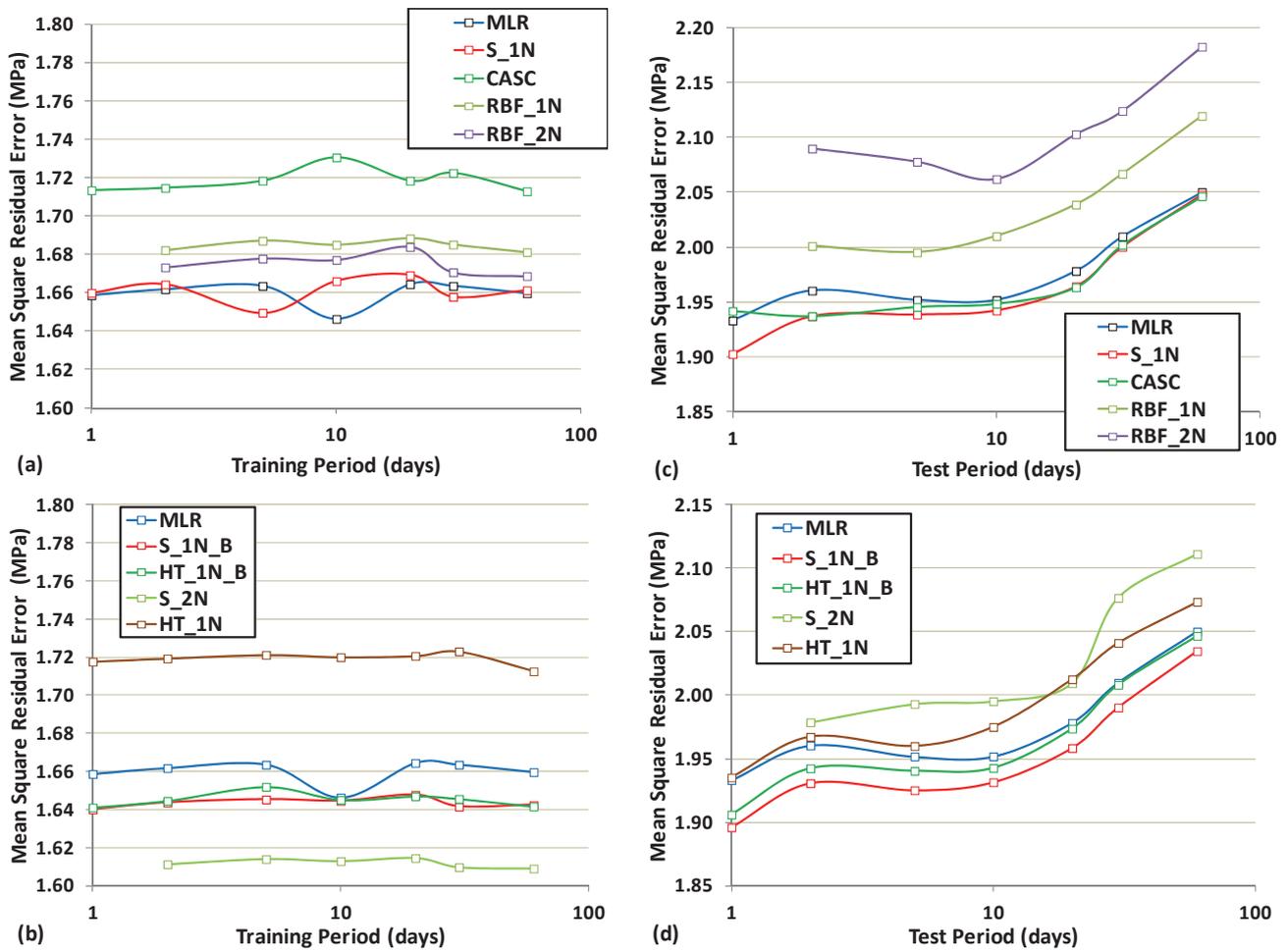


Fig. 7 – Training and test MSRE of Str_28_1 model

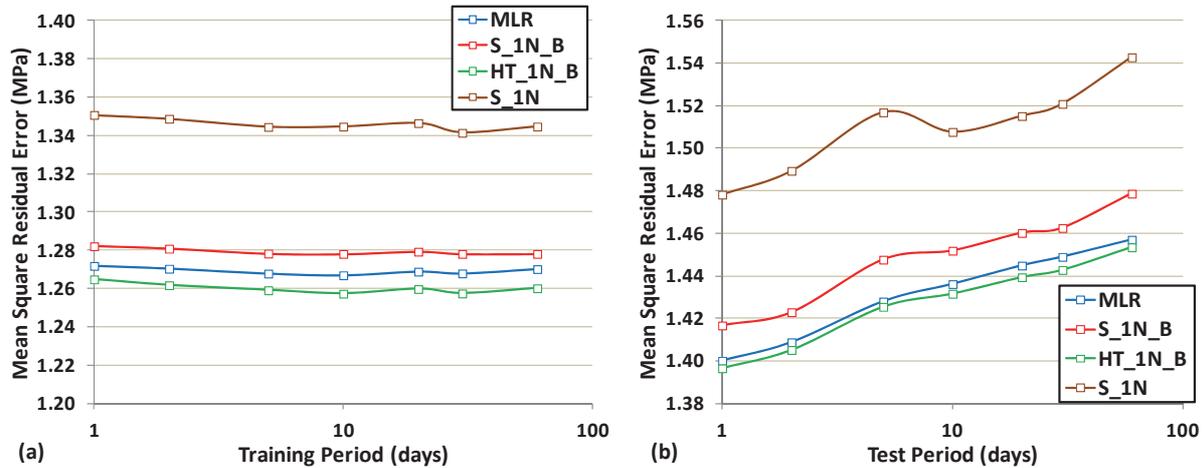


Fig. 8 – Training and test MSRE of Str_28_7 model

eters not included in the current models – clinker composition, mineralogy, and activity – have been modified from their previous state; (b) Some values of input variables do not belong to the range of these variables during T_D period. Therefore, the models are obliged to extrapolate the computation, which can lead to a worsening of prediction.

From Figure 7 it may be observed that the dynamic MLR model shows high efficiency: Only four out of the eight ANN models provide a lower test error: The S_1N, S_1N_B, CASC and HT_1N_B models behave better than MLR in future predictions. The more complicated ANNs with two nodes in the hidden layer or with the radial basis functions provide worse $MSRE_{Future}$. Generally, the addition of bias improves the generalization ability of the ANNs with one node in the hidden layer. The simple S_1N_B model provides the minimum test errors. Model S_2N, despite its lower training error, fails to predict the future strength better than MLR, a fact that is a clear indication of over-fitting. Over-fitting also appears in the case of ANNs based on radial basis functions (RBF). The RBF structure with two hidden nodes presents lower training error than that with one hidden node. However, the $MSRE_{Future}$ of the former model is noticeably higher than that of the latter for all the range of T_F . Cascade ANN behaves relatively better than MLR as concerns $MSRE_{Future}$, but it is the most complicated ANN and needs longer computational time. Therefore, it appears that the higher generalization ability is achieved with a more compact architecture, namely models with fewer parameters.

The MLR and ANN techniques have also been compared using the Str_28_7 model. This model has been applied only for three out of the eight ANN techniques, those with one node in the hidden layer, sigmoid and hyperbolic activation function σ , with or without bias: The S_1N, S_1N_B and

HT_1N_B models. As previously mentioned, the $MSRE_{Past}$ and $MSRE_{Future}$ curves were constructed for $T_D = 180$ and T_F ranging from 1 to 60 days. The results are shown in Figure 8. The MLR model shows an adequately small $MSRE_{Past}$, in the range of 1.27 MPa, almost equal to the $MSRE_{Past}$ of S_1N_B and HT_1N_B, while the training residual error of S_1N is relatively higher. Concerning the test errors, $MSRE_{Future}$ of MLR is lower than those of S_1N and S_1N_B, and almost equal to the error of HT_1N_B. Therefore, in the current level of ANNs development, the simple MLR technique is very reliable.

Optimization of T_D and T_F parameters

The training and test periods, T_D and T_F respectively, need optimization as regards minimization of $MSRE_{Future}$. To achieve this objective, the models Str_28_1 and Str_28_7 were applied for T_F ranging from 1 to 60 days, using the MLR technique as well as the two best ANNs, S_1N_B and HT_1N_B, which show a minimal $MSRE_{Future}$. The dynamic models have been implemented for T_D values from the set $\{60, 90, 120, 180, 240, 300, 360, 540, 720\}$, while T_F takes values from the set $\{1, 2, 5, 10, 20, 30, 60\}$. For all the possible combinations of (T_D, T_F) , the errors $MSRE_{Past}$, $MSRE_{Future}$ are determined. The results for Str_28_1 model are depicted in Figure 9. The analysis of these results leads to the subsequent conclusions: (a) $MSRE_{Past}$ is an ascending function of T_D and not significantly dependent on T_F . (b) For the same T_D value, the $MSRE_{Past}$ results computed from both ANNs are ≈ 0.02 lower than that computed by MLR, for almost the entire range of T_F . (c) $MSRE_{Future}$ is a strong function of both T_D , T_F and for each T_F value, there is a T_D where $MSRE_{Future}$ becomes minimal. For the MLR and S_1N_B techniques, these T_D values of minimum future error

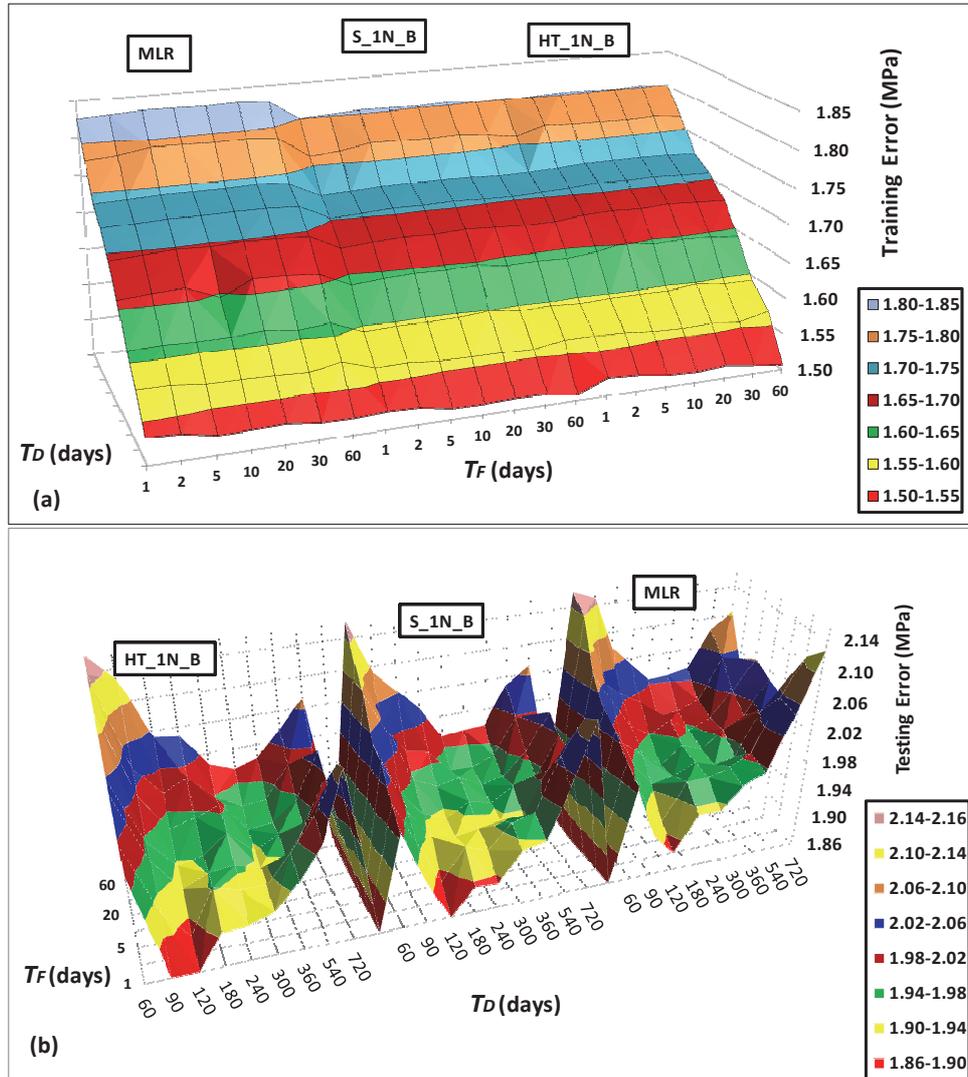


Fig. 9 – Training and test errors for $T_D \in [60, 720]$, $T_F \in [1, 60]$ and Str_{28_1} model

show an increasing trend as T_F increases: For small T_F , $T_D = 120$ days, while for higher T_F values, T_D increases to 180 days, reaching 240 days for $T_F = 60$ days. Therefore, a positive correlation between the time parameters exists for the two mentioned techniques. As concerns HT_1N_B technique, T_D is independent of T_F , remaining continuously equal to 120 days as T_F varies from 1 to 60 days. (d) Small values of $T_D - T_F = 60$ days – lead to a noticeable worsening of future prediction, meaning that short training periods are insufficient to train the models. (e) The minimum $MSRE_{Future}$ for the MLR technique appears for $(T_D, T_F) = (120, 1)$ and it is equal to 1.89 MPa. (f) For the same values of T_D and T_F the two ANNs provide a residual error continuously lower than that of MLR. Therefore, both ANNs behave better than MLR in predicting the future 28-day strength. (g) The minimum $MSRE_{Future}$ for S_1N_B and HT_1N_B appears for $(T_D, T_F) = (120, 1)$ and it is equal to 1.86 MPa. (h) Additionally the error of ANNs for $T_F = 2$ remains lower than that of MLR

for $T_F = 1$. The above advantage leads to a more robust implementation of the selected neural networks in the actual quality control of a cement plant. (i) Because all the points (T_D, T_F) in the grid $T_D \in \{60 \dots 720\}$, $T_F \in \{1 \dots 60\}$ have been scanned, the global minimum $MSRE_{Future}$ is assured. (j) For small $T_D = 60$ days or high $T_F \geq 20$ days, $MSRE_{Future}$ differs noticeably from optimum. The causes are explained in the previous paragraph as regards the variables not included in the current models, and values of input variables not belonging to the range of these variables during the training period. The sensitivity of the optimum $(T_D, T_F) = (120, 1)$ has also been studied by determining the pairs (T_D, T_F) providing $MSRE_{Future} \leq 1.02 MSRE_{Future, Min}$ using interpolation between the discrete time parameters. This area constitutes the optimum region of the time parameters. The results are demonstrated in Figure 10. The optimum areas are very similar for MLR and HT_1N_B techniques. In the case of S_1N_B, T_D is expanded to higher values, but max-

imum T_F is approximately one day shorter compared to HT_1_B. This is an indication of higher robustness of the HT_1_B technique.

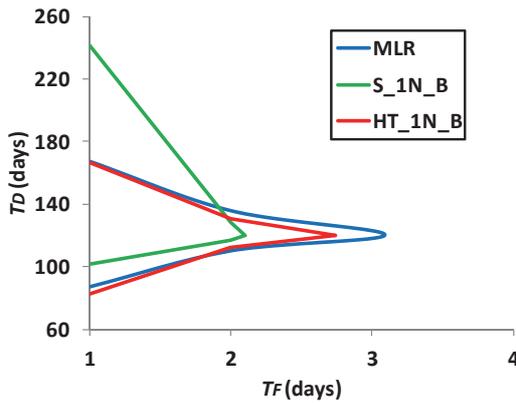


Fig. 10 – Optimum area of (T_D, T_F) for *Str_28_1* model

A similar analysis has been performed for the residual errors during the training and testing period for the model *Str_28_7*. The results are plotted in Figure 11, and the following trends were observed: (a) The $MSRE_{past}$ is an increasing function of T_D and independent of T_F . (b) The minimum $MSRE_{Future}$ appears for the highest T_D equal to 720 days, and for this T_D value it does not depend on T_F . (c). From (a), (b) it is deduced that smaller training periods are not adequate to train the models, leading to higher test error. (d) The lowest test errors for the full range of T_D, T_F are provided by applying the HT_1S_B neural network. The errors derived by applying MLR are not far from those of HT_1S_B. Third in this ranking is the S_1N_B model. (e) When T_D decreases from 720 to 60 days, the function between $MSRE_{Future}$ and T_F is ascending for the same T_D .

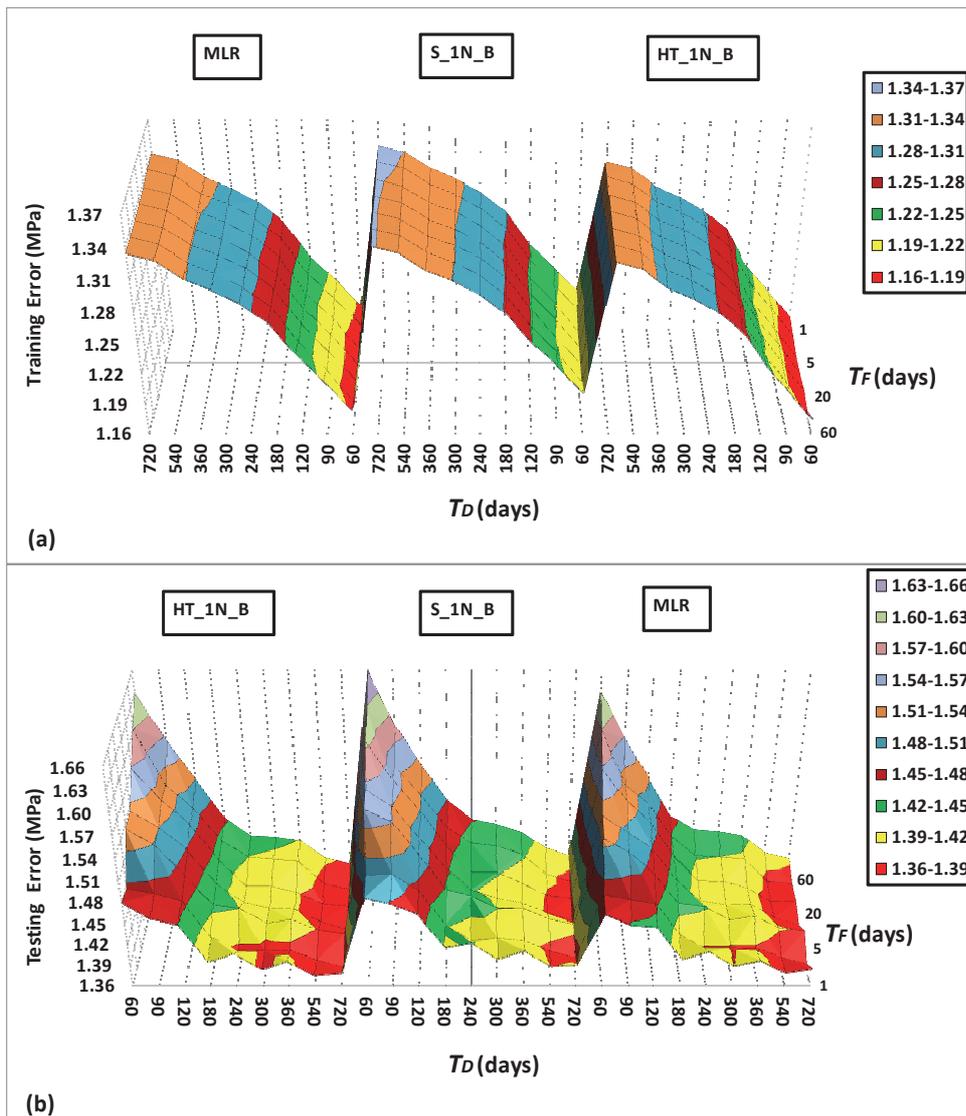


Fig. 11 – Training and test errors for $T_D \in [60, 720]$, $T_F \in [1, 60]$ and *Str_28_7* model

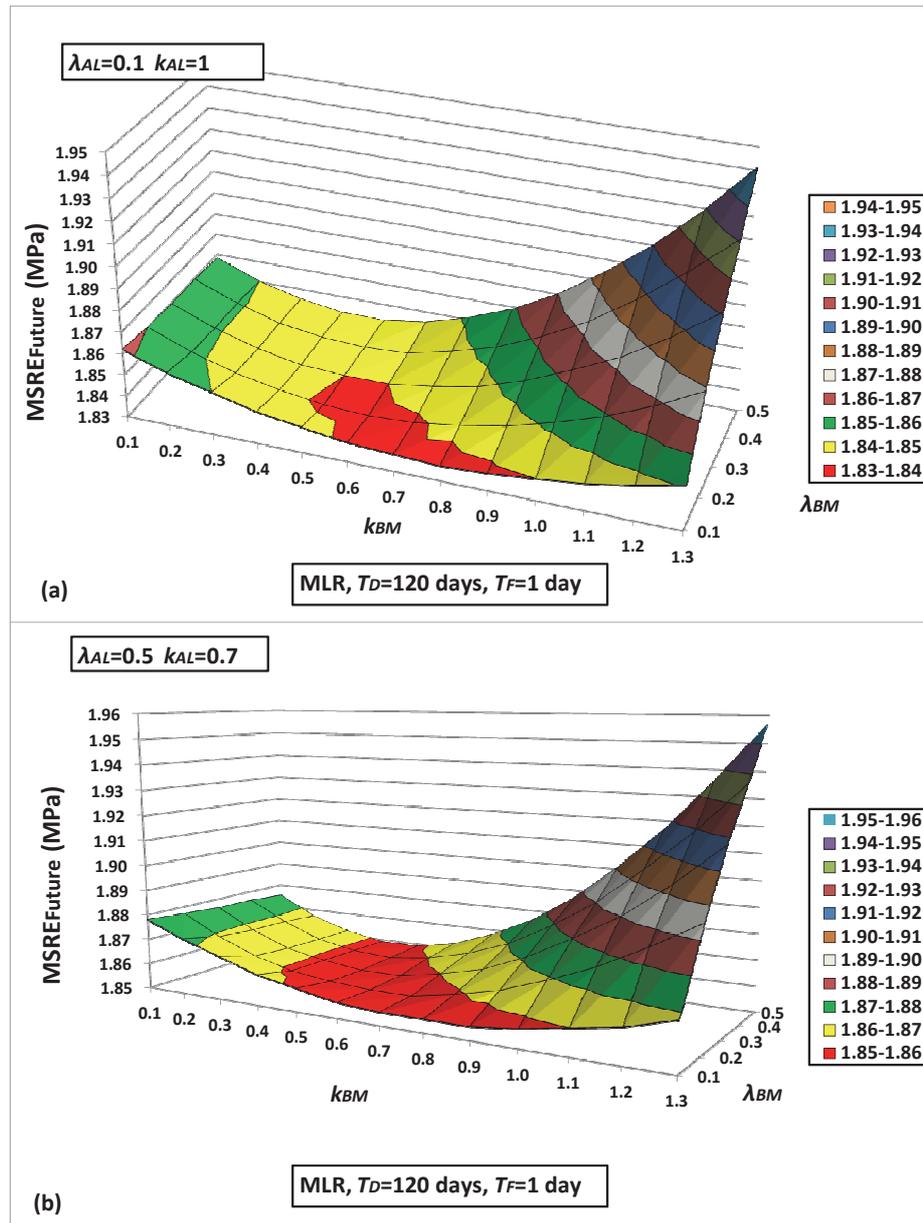


Fig. 12 – Test errors for Str_28_EW model, MLR technique and (a) $(\lambda_{AL}, k_{AL}) = (0.1, 1)$, (b) $(\lambda_{AL}, k_{AL}) = (0.5, 7)$

Residual errors of Str_28_EW model

The model Str_28_EW is composed by the superposition of models Str_28_1 and Str_28_7. An EWMA filter is applied to smooth the variable $Diff(I)$ defined by equation (17). Str_28_EW contains four additional parameters, named k_{BM} , λ_{BM} , k_{AL} , λ_{AL} whose values need determination. Two difficulties arose during the solving of this problem: (a) The residual errors do not diminish any value of k and l , but an optimization technique is required, using $MSRE_{Future}$ as criterion. (b) Some trials of application of non-linear regression technique showed convergence in local minima, thus the optimum values were not guaranteed. For these reasons, the following steps applied: (a) For each parameter k_{BM} , λ_{BM}

k_{AL} , λ_{AL} minimum and maximum values, k_{MIN} , k_{MAX} , λ_{MIN} , λ_{MAX} are selected. (b) Steps of change, dk , $d\lambda$ are also chosen. (c) The rectangular area defined by the vertices k_{MIN} , k_{MAX} , λ_{MIN} , λ_{MAX} is scanned with steps dk , $d\lambda$ for each CEM type. The errors during training and test provided by equations (13) and (14) are computed. (d) As optimization criterion, the minimum $MSRE_{Future}$ is selected. (e) The steps (a)-(d) are implemented for each (T_D, T_F) and the optimum $(k_{AL}, \lambda_{AL}, k_{BM}, \lambda_{BM})$, producing the minimum $MSRE_{Future}$ is found. The designed optimization technique provides the actual optimum as it performs calculations in all the lattice points.

The coupled model Str_28_EW has been applied using the MLR technique as well as the two best ANNs: S_1N_B and HT_1N_B and time pa-

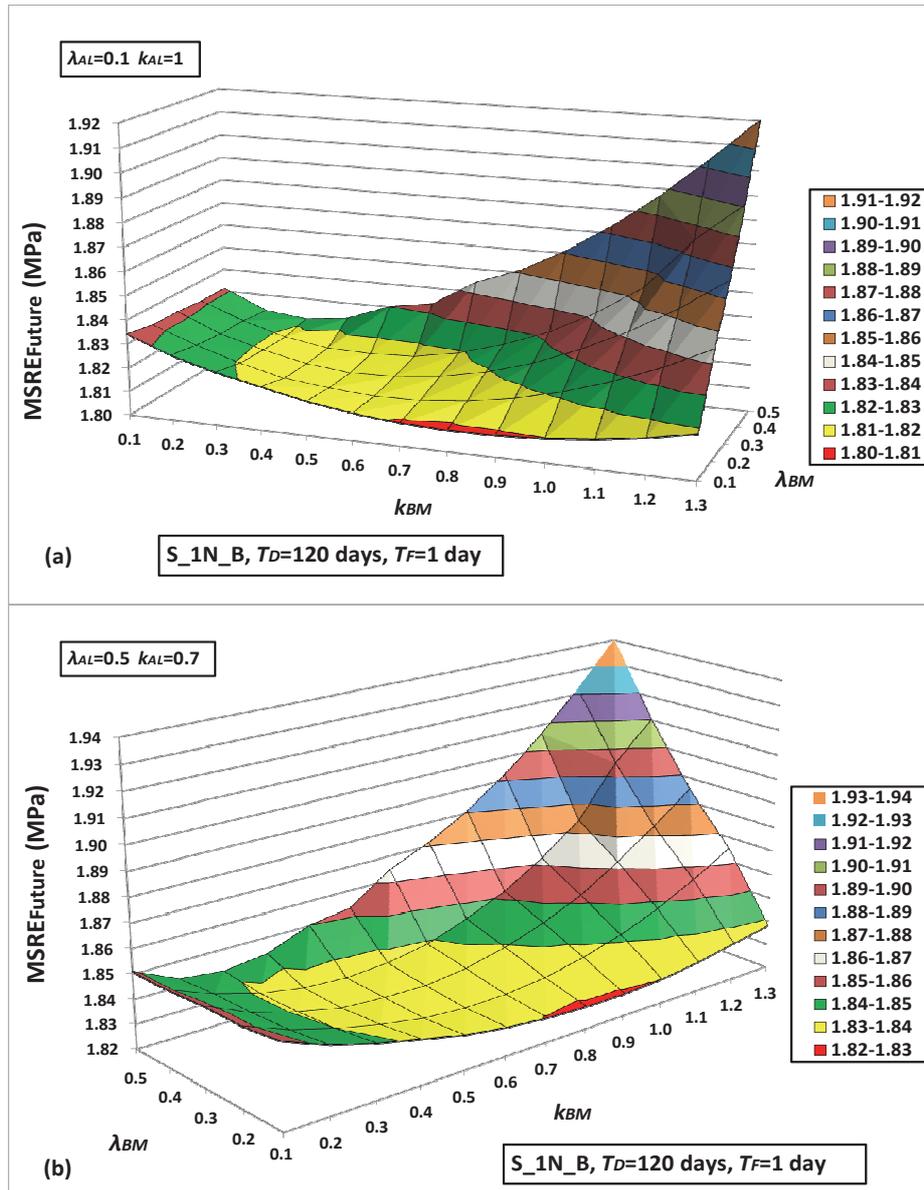


Fig. 13 – Test errors for Str_28_EW model, S_1N_B technique and (a) $(\lambda_{AL}, k_{AL}) = (0.1, 1)$, (b) $(\lambda_{AL}, k_{AL}) = (0.5, 7)$

rameters $T_D = 120$, $T_F = 1$. For these three models, $MSRE_{Future}$ as function of k_{BM} , λ_{BM} is plotted in Figures 12, 13, 14 for (k_{AL}, λ_{AL}) equal to (1, 0.1) and (0.7, 0.5). The $MSRE_{Future}$ of model Str_28_1 for the time parameters selected and for each technique are as follows: for MLR the error is 1.885, for S_1N_B it is 1.857, while for HT_1N_B it is 1.858. Figures 12, 13, 14 prove that Str_28_EW is able to provide an improved test error, in case the optimum parameters k, λ are determined and applied. For each λ value there is only one k value where $MSRE_{Future}$ becomes minimal. From the shape of the curves, it is concluded that increasing λ , minimal $MSRE_{Future}$ is obtained for lower k values, meaning that there is a negative correlation between these two parameters. Despite the differences in the shape of the surfaces in the three Figures, generally, the minimum error is

found at low values of λ and high values of k , in the intervals (0.1, 0.2) and (0.9, 1.0) correspondingly.

The sensitivity of the optimum (k, λ) is also studied by determining the pairs (k, λ) which provide $MSRE_{Future} \leq 1.005 MSRE_{Future Min}$ using interpolation between the discrete parameters and the results shown in Figures 12(a), 13(a), 14(a). This area of (k_{BM}, λ_{BM}) constitutes the optimum region. The respecting areas for MLR and the two ANN techniques are depicted in Figure 15, which shows that the sensibility of the optimal k_{BM}, λ_{BM} is low. For HT_1N_B neural network, the range of k, λ of optimum test error is larger compared with S_1N_B for almost the same $MSRE_{Future Min}$. From these results, it is concluded, that a slow EWMA filter provides a more effective correction and amelioration of the prediction. The test errors for the full range of T_D ,

T_F by implementing the MLR, S_1N_B, HT_1N_B techniques are shown in Figure 16. For each pair (T_D, T_F) the optimum values ($k_{AL}, \lambda_{AL}, k_{BM}, \lambda_{BM}$) have been determined. The results of this Figure compared to the results of Figure 9(b) concerning the test error of Str_28_1 model, lead to the following remarks.

(i) The coupled model is noticeably better than Str_28_1 for all three techniques as regards test error. This improvement is higher, increasing T_D and T_F . For both models, a small T_D period augments the test error meaning that the size of the training set is inadequate to train the model satisfactorily. The above could be compared with a small integral time in a PID controller, which creates oscillations around the set point.

(ii) Time parameters, T_D and T_F , have strong impact on test error for both models, independently of the technique applied: The optimum value of T_F is one day, while the test error as function of T_D passes from a minimum value found between 120 and 240 days depending on the model and technique applied. The optimum (T_D, T_F) and the minimum $MSRE_{Future}$ for both models and all the techniques are shown in Table 6.

(iii) As T_D increases, the surface $MSRE_{Future}$ is more flat in the case of Str_28_EW compared to that of Str_28_1, for all the three techniques, i.e. Str_28_EW is more robust as regards the selection of T_D .

(iv) For the optimum values of T_D , the differences between $MSRE$ of $1 < T_F \leq 30$ and $MSRE$ of

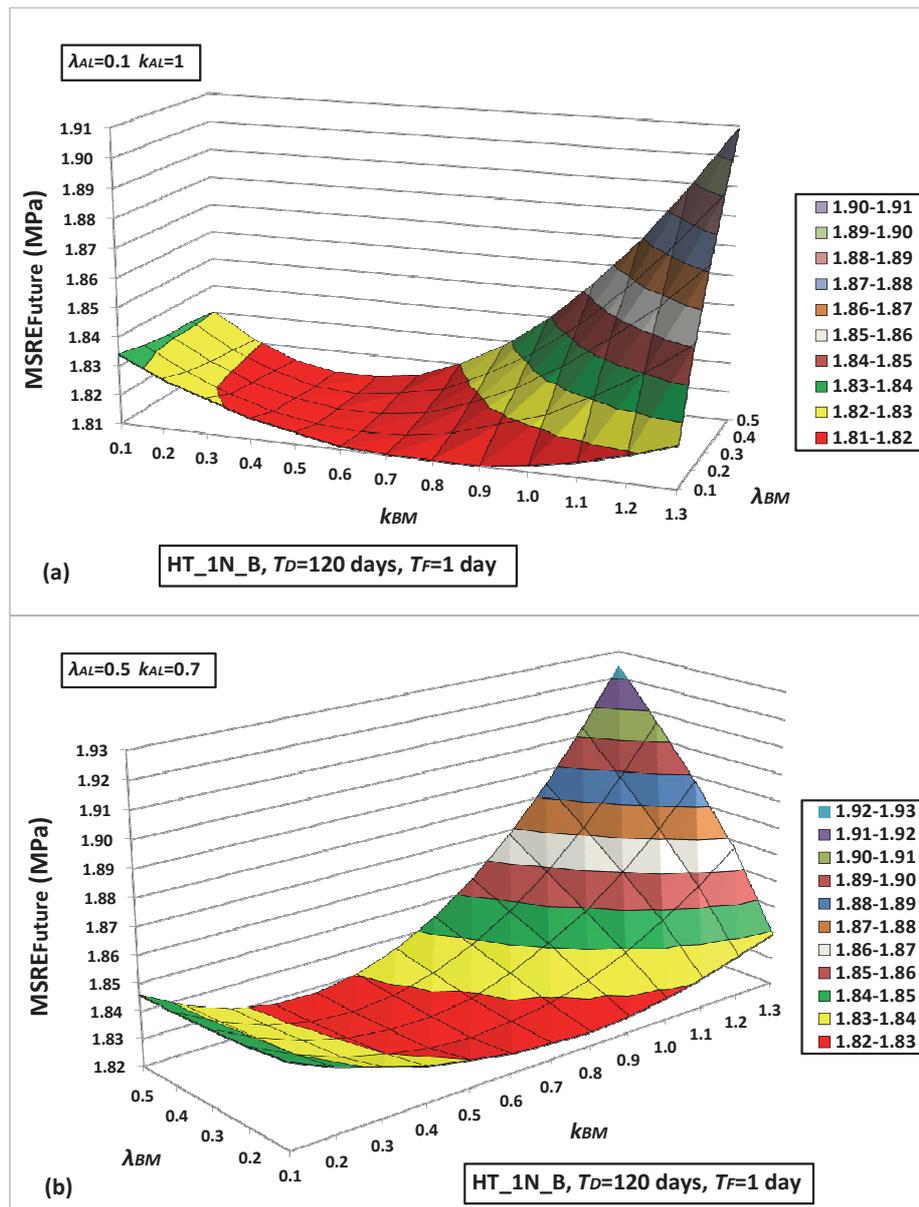


Fig. 14 – Test errors for Str_28_EW model, HT_1N_B technique and (a) $(\lambda_{AL}, k_{AL}) = (0.1, 1)$, (b) $(\lambda_{AL}, k_{AL}) = (0.5, 7)$

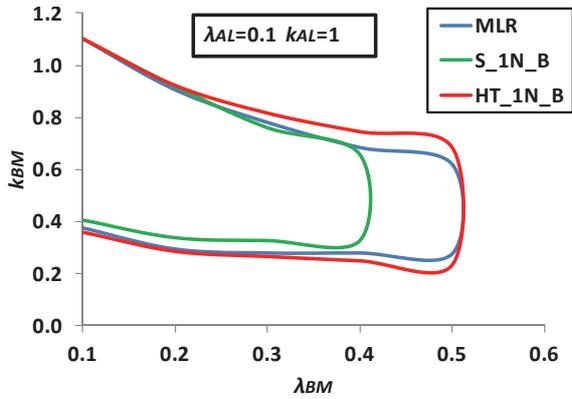


Fig. 15 – Optimum area of (k_{BM}, λ_{BM}) for *Str_28_EW* model

Table 6 – Comparison of minimum $MSRE_{Future}$ between *Str_28_1*, *Str_28_EW* models for the optimal techniques

	T_D	T_F	$MSRE_{Future}$
Str_28_1			
MLR	120	1	1.885
S_1N_B	120	1	1.857
HT_1N_B	120	1	1.858
Str_28_EW			
MLR	240	1	1.826
S_1N_B	240	1	1.785
HT_1N_B	240	1	1.803

$T_F = 1$ are shown in Figure 17, for all the models and techniques. When T_F increases, the rate that $MSRE_{Future}$ augments is much lower in the case of coupled model compared to the *Str_28_1*, meaning that the *Str_28_EW* model shows high robustness to this time parameter. The benefit of this significant characteristic during the implementation of the model in the daily quality control is that, if for some reason the model is not updated daily – $T_F = 1$ – but in some cases a few days later, the worsening of prediction is much lower for the *Str_28_EW* model compared to the *Str_28_1*.

(v) The implementation of neural networks assures a lower test error, compared with the linear technique for both models. The minimum error for the *Str_28_EW* model is obtained by using the S_1N_B technique for $T_D = 240$ and $T_F = 1$.

The above remarks prove that the combination of ANN techniques and of the models coupling, leads to a serious improvement of the 28-day strength prediction. Such an improvement becomes more important because already the *Str_28_1* model combined with the MLR technique shows a high ability to predict the future strength. Next, a basic *Str_28_1* model with MLR was considered for $T_D = 120$ and $1 \leq T_F \leq 60$. The percentage of reduction of $MSRE_{Future}$ by using the optimum *Str_28_EW* model with S_1N_B neural network, for $T_D = 240$ and the same range of T_F appears in Figure 18. The reduction of error starts from ~5 % for $T_F = 1$ and reaches ~10 % for $T_F = 30$. Therefore, there is a noticeable amelioration, which can lead to an enhanced quality control and reduction of the cement strength variance.

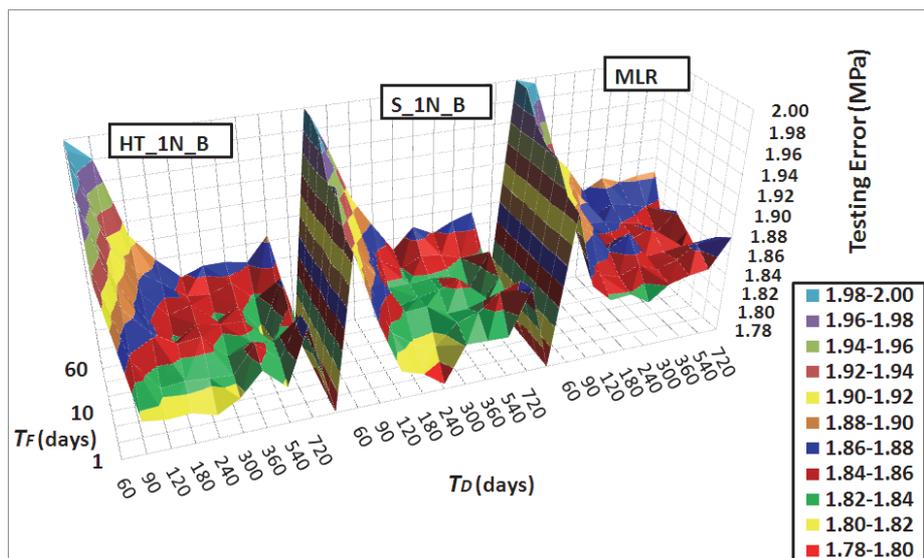


Fig. 16 – Test errors for *Str_28_EW* model as function of T_D , T_F and optimum values of k , λ

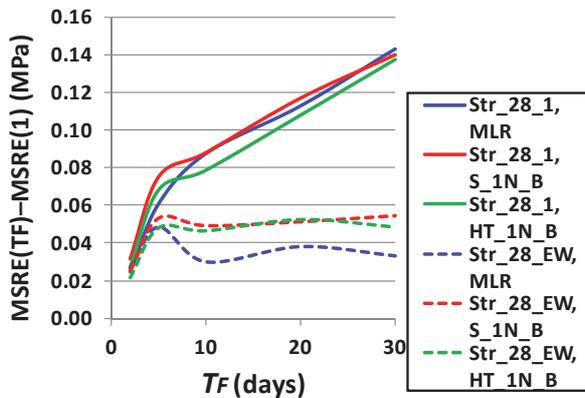


Fig. 17 – Difference $MSRE_{Future}(T_F) - MSRE_{Future}(1)$ as function of T_F

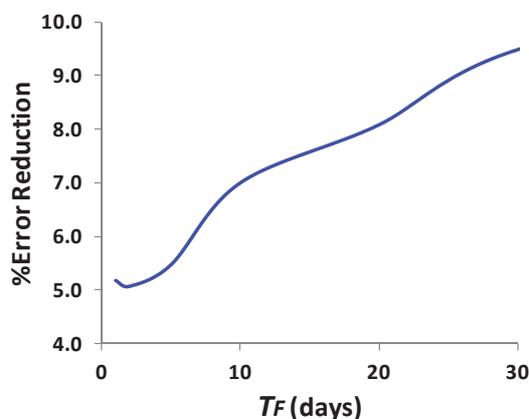


Fig. 18 – % Reduction of $MSRE_{Future}$ between Str_{28_1} , MLR and Str_{28_EW} , S_{1N_B}

Conclusions

A dynamic approach of two well-known techniques has been used to model and predict the 28-day cement strength: Multiple linear regression (MLR), and artificial neural networks (ANN). The modeling includes data for Portland cement produced according to EN 197-1:2011 and utilizes analyses of daily average samples of cement produced industrially. Physical, chemical, and early strength results are used to predict the typical 28-day strength by developing the models Str_{28_1} and Str_{28_7} . Two kinds of models have been built – Str_{28_1} and Str_{28_7} – which include only the 1-day strength as an independent variable or both 1-day and 7-day strength. The models are dynamic because they are applied to a movable past period of T_D days to calculate the parameters, and then used for a future period of T_F days. Upon completion of the future period, the process is repeated by moving forward periods. Various ANNs have been developed involving three layers with one or two nodes in the hidden layer. In parallel, a cascade ANN has been created. Three types of activation

functions have been utilized; sigmoid, hyperbolic tangent, and radial basis functions. The comparison is based on the $MSRE$ of testing sets. The linear technique shows high performance as only four out of the eight ANNs provide a lower test error for the Str_{28_1} model: Two ANNs with one hidden node, sigmoid transfer function, with and without bias, one ANN with one hidden node, hyperbolic tangent function and bias, as well as the cascade ANN. The other more complicated ANNs suffer from over-fitting, and test error fails to be lower than that of MLR. Using this error as a criterion, an optimization of the training and testing periods, represented by the time parameters T_D and T_F , has been performed. For Str_{28_1} model and for all the techniques applied, the best estimations were obtained for $T_F = 1$.

The model Str_{28_7} , taking into account the 7-day strength results, generates significantly smaller training and test $MSRE$ compared to Str_{28_1} , but it has longer delay time. Therefore, it is difficult to use for direct control purposes, although it provides more accurate information. For this reason, a third model has been developed, called Str_{28_EW} , with coupling the Str_{28_1} model and a filtered version of Str_{28_7} . This model involves two additional parameters per CEM type: The weight λ of the EWMA filter and the weight k of coupling. These parameters have been optimized to reach the minimum test error. By applying both MLR and two of the best ANN techniques, Str_{28_EW} leads to errors continuously smaller than those of Str_{28_1} . These best ANN techniques include only one node in the hidden layer, sigmoid and hyperbolic functions and bias. Their implementation in the coupled model provides a smaller error than the combination of the Str_{28_EW} with MLR. Consequently, by combining the coupled model with the two best ANNs and selecting the optimal parameters of time, of filtering and weighting, a minimal test error is assured. For the optimal values of T_D , k , λ the coupled model is less sensitive in the selection of T_F . During the implementation of the model in process control, the benefit is that if the model is not updated daily – for the optimum $T_F = 1$ day – but in some cases a few days later, the worsening of prediction is much lower for the Str_{28_EW} model, compared with the Str_{28_1} . The further improvement of these techniques should follow the directions below:

(a) The ANN inputs should include the chemical characteristics of the clinker, such as C_3S , C_3A , equivalent alkalis, free lime.

(b) Use the ANN coupled model for the development of robust controllers to regulate the cement composition for a predetermined strength target.

Nomenclature

A_0, A_1	– Coefficients of equation (3)
C_3S	– Tricalcium silicate, %
$Diff(I)$	– Difference between calculated and actual 28 – day strength at time I , MPa
dk	– step of change of parameter k
$EW_Diff(I)$	– EWMA difference between calculated and actual 28 – day strength at time I , MPa
I	– Discrete time in equations (16), (17), days
Ins_Res	– Insoluble residue, %
J	– Discrete time in equation (18), days
k	– Weight parameter of the coupled model
K_{TD}	– Number of consecutive data sets
LOI	– Loss on ignition, %
M	– Total number of data sets in equation (5)
$MSRE$	– Mean square residual error, MPa
N	– Number of coefficients
N_I	– Number of nodes in the hidden layer
$o(J)$	– Activation functions of neural networks
$R40$	– Residue at 40 microns sieve, %
S_b	– Specific surface, $10\text{ m}^2\text{ kg}^{-1}$
SO_3	– Sulfates content, %
s_{Res}	– Residual error, MPa
Str_1	– Compressive strength at 1 day, MPa
Str_7	– Compressive strength at 7 days, MPa
Str_{28}	– Compressive strength at 28 days, MPa
$Str_{28_1(I)}$	– Calculated strength at 28 days from model Str_{28_1} at time I , MPa
$Str_{28_7(I)}$	– Calculated strength at 28 days from model Str_{28_7} at time I , MPa
$Str_{28_EW(I)}$	– Calculated strength at 28 days from model Str_{28_EW} at time I , MPa
t	– Time, days
T_D	– Set training period, days
T_F	– Set testing period, days
$X(I)$	– Independent variable in equation (16)
X_I	– Independent variables in equations (1), (6)
XN_I	– Normalized independent variables in equations (1), (3), (6), (9), (10), (11)
XO_I	– Parameters of the center of the radial basis function
Y	– Dependent variable in equation (2), (7)
$Y(I)$	– Dependent variable in equation (16)
YN	– Normalized dependent parameter in equations (3), (4), (7), (8), (12)
$V(J)$	– Synaptic weights between hidden and output layers in equation (12)
$W_{I,J}$	– Synaptic weights between input and hidden layers in equations (9), (10)

Greek symbols

λ	– parameter of the EWMA model
$d\lambda$	– step of change of parameter λ
σ_I	– variance parameters of the radial basis function

Subscripts

Act	– actual value
AL	– CEM II A-L 42.5
BM	– CEM II B-M 32.5
Calc	– calculated value
Future	– future period
Min	– minimum value
Max	– maximum value
Past	– past period

Abbreviations

ANN	– artificial neural networks
CEM	– cement type
DMEP	– double-layer multi-expression programming
EWMA	– exponentially weighted moving average
FL	– fuzzy logic
GEP	– gene expression programming
GA	– genetic algorithms
GRNN	– generalized regression neural network
MEP	– multi – expression programming
MLR	– multiple linear regression
PCA	– principal components analysis
RBF	– radial basis function

References

1. Steege, F. F., Grob, H. M., Comparison of Long-Term Adaptivity for Neural Networks, in Villa, A., Duch, W., Érdi, P., Masulli, F., Palm, G. (Eds.), *Artificial Neural Networks and Machine Learning, Proceedings of 22nd. Int. Conf. on Artificial Neural Networks (ICANN 2012)*, Lausanne, Switzerland, Part II, Springer, 2012, pp. 50–57.
2. Olewi, M. A., Aljaseem, Z. A., Neural Network Control for PPDCV Clinker Cooler System, *Int. J. Comput. Appl.* **85** (2014) 29.
3. Karikov, E. B., Rubanov, V. G., Klassen, V. K., Construction of a Dynamic Neural Network Model as a Stage of Grate Cooler Automation, *World Appl. Sci. J.* **25** (2014) 227.
4. Kadri, O., Mouss, L. H., Mouss, M. D., Fault diagnosis of rotary kiln using SVM and binary ACO, *J. Mech. Sci. Technol.* **26** (2012) 601.
doi: <http://dx.doi.org/10.1007/s12206-011-1216-z>
5. Singh, V., Banerjee, P. K., Tripathy, S. K., Saxena, V. K., Venugopal, R., Artificial Neural Network Modeling of Ball Mill Grinding Process, *J. Powder Metall. Min.* **2** (2013) 1.
doi: <http://dx.doi.org/10.4172/2168-9806.1000106>
6. Gómez Sarduy, J. R., Monteagudo Yanes, J. P., Granado Rodríguez, M. E., Quiñones Ferreira, J. L. Miranda Torres,

- Y., Determining cement ball mill dosage by artificial intelligence tools aimed at reducing energy consumption and environmental impact, *Ing. Invest.* **33** (2013) 49.
7. U. Yildiran, U., Kaynak, M. O., Neural network based control of a cement mill by means of a VSS based training algorithm, in Proceedings of the 2002 IEEE International Symposium on Industrial Electronics, L'Aquila, Italy, 2002, Vol. 1, IEEE, Piscataway, New Jersey, 2002, pp. 326–331.
 8. Odler, I., Cement Strength, *Mater. Struct.* **24** (1991) 143. doi: <http://dx.doi.org/10.1007/BF02472476>
 9. Tsamatsoulis, D., Prediction of cement strength: analysis and implementation in process quality control, *JMBM*, **21** (2012) 81.
 10. Yeh, I. C., Modeling of Strength of High-Performance Concrete Using Artificial Neural Networks, *Cem. Concr. Res.* **28** (1998) 1797. doi: [http://dx.doi.org/10.1016/S0008-8846\(98\)00165-3](http://dx.doi.org/10.1016/S0008-8846(98)00165-3)
 11. Wang, J. Z., Wang, X. J., Ni, H. J., An Algorithm of Neural Network and Application to Data Processing in Concrete Engineering, URL: <http://www.mii.lt/informatica/pdf/INFO430.pdf> (14.3.2016).
 12. Özturan, M., Kutlu, B., Özturan, T., Comparison of Concrete Strength Prediction Techniques with Artificial Neural Network Approach, URL: https://www.academia.edu/472056/COMPARISON_OF_CONCRETE_STRENGTH_PREDICTION_TECHNIQUES_WITH_ARTIFICIAL_NEURAL_NETWORK_APPROACH, (14.3.2016).
 13. Özcan, F., Atis, C. D., Karahan, O., Uncuoglu, E., Tanyildizi, H., Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume concrete, *Adv. Eng. Software* **40** (2009) 856. doi: <http://dx.doi.org/10.1016/j.advengsoft.2009.01.005>
 14. Aggarwal, Y., Aggarwal, P., Prediction of Compressive Strength of SCC Containing Bottom Ash using Artificial Neural Networks, *World Acad. Sci. Eng. Technol.* **5** (2011) 735.
 15. Gupta, S., Using Artificial Neural Network to Predict the Compressive Strength of Concrete containing Nano-silica, <http://www.hrpub.org/download/201310/cea.2013.010306.pdf>, (14.3.2016)
 16. Omran, B. A., Chen, Q., Jin, R., Prediction of Compressive Strength of “Green” Concrete Using Artificial Neural Networks, in Sulbaran, T. (Ed.), Proceedings of 50th ASC Annual International Conference, Blacksburg, Virginia, Associated Schools of Construction, Washington, 2014.
 17. Trtnik, G., Kavčič, F., Turk, G., Prediction of concrete strength using ultrasonic pulse velocity and artificial neural networks, *Ultrasonics* **49** (2009) 53. doi: <http://dx.doi.org/10.1016/j.ultras.2008.05.001>
 18. Dolado, J. S., van Breugel, K., Recent advances in modeling for cementitious materials, *Cem. Concr. Res.* **41** (2011) 711. doi: <http://dx.doi.org/10.1016/j.cemconres.2011.03.014>
 19. Akkurt, S., Ozdemir, S., Tayfur, G., Akyol, B., The use of GA–Anns in the modeling of compressive strength of cement mortar, *Cem. Concr. Res.* **33** (2003) 973. doi: [http://dx.doi.org/10.1016/S0008-8846\(03\)00006-1](http://dx.doi.org/10.1016/S0008-8846(03)00006-1)
 20. Baykasoğlu, A., Dereli, T., Taniş, S., Prediction of Cement Strength Using Soft Computing Techniques, *Cem. Concr. Res.* **34** (2004) 2083. doi: <http://dx.doi.org/10.1016/j.cemconres.2004.03.028>
 21. Thamma, P., Barai, S. V., Prediction of Compressive Strength of Cement Using Gene Expression Programming, in Mehnen, J., Koepfen, M., Saad, A., Tiwari, A. (Eds.), *Advances in Intelligent and Soft Computing* **58** (2009) 203. doi: http://dx.doi.org/10.1007/978-3-540-89619-7_20
 22. Motamedi, S., Shamshirband, S., Hashim, R., Petković, D., Roy, C., Estimating unconfined compressive strength of cockle shell–cement–sand mixtures using soft computing methodologies, *Eng. Struct.* **98** (2015) 49. doi: <http://dx.doi.org/10.1016/j.engstruct.2015.03.070>
 23. Motamedi, S., Shamshirband, S., Petković, D., Hashim, R., Application of adaptive neuro-fuzzy technique to predict the unconfined compressive strength of PFA-sand-cement mixture, *Powder Technol.* **278** (2015) 278. doi: <http://dx.doi.org/10.1016/j.powtec.2015.02.045>
 24. Zhang, Q., Yang, B., Wang, L., Zhu, F., Predicting cement compressive strength using double-layer multi-expression Programming, in Proceedings of Fourth International Conference on Computational and Information Sciences (IC-CIS), Kuala Lumpur, Malaysia, IEEE Piscataway, New Jersey, 2012, pp. 94–97.
 25. Madsen, M., Thyregod, P., Popentiu, F., Albeanu, G., Coarna, M., Serbanescu, L., Mirza, D., Tautan, A., Considerations concerning a fuzzy – genetic algorithm with application to CEM I type cements quality optimization, in Zinn, D., Savoie, M. J., Kuo-Chi Lin, El-Badawy E., Zenga, G. (Eds), Proceedings of the 9th World Multi-Conference on Systemic, Cybernetics and Informatics, Orlando, Florida, Vol. X, 2005, International Institute of Informatics and Systemic, pp. 350–353.
 26. Ren, S., Yang, D., Ji, F., Tian, X., Application of generalized regression neural network in prediction of cement properties, in Proceedings of International Conference on Computer Design and Applications (ICCCA), Qinhuangdao, China, Vol. 2, IEEE, Piscataway, New Jersey, 2010, pp 385–388.
 27. Yongzheng, Z., Biyun, X., Yong, Z., The prediction research about the strength of cement based on principal component neural network, in Proceedings of IEEE 3rd International Conference on Communication Software and Networks (ICCSN), Xi’an, China, IEEE, Piscataway, New Jersey, 2011, pp. 238–241.
 28. Tsamatsoulis, D., Application of the static and dynamic models in predicting the future strength of pozzolanic cements, in Proceedings of 18th International Conference on Circuits, Systems, Communications and Computers, Latest trends in Systems, Santorini Island, Greece, Vol. I, INASE, 2014, pp. 138–146.
 29. Tsamatsoulis, D., Application of the static and dynamic models in predicting the future strength of Portland cements, Proceedings of International Conference on Mathematics and Computers in Sciences and Industry, Varna, Bulgaria, 2014, IEEE, pp. 75–85. doi: <http://dx.doi.org/10.1109/mcsi.2014.38>
 30. Tsamatsoulis, D., Comparison of dynamic linear and neural network models in predicting cement compressive strength, URL: https://www.researchgate.net/publication/280125745_Comparison_of_Dynamic_Linear_and_Neural_Network_Models_in_Predicting_Cement_Compressive_Strength (14.3.2016)
 31. Haykin, S., *Neural networks. a comprehensive foundation*, Prentice Hall, Delhi, 2005, pp. 24.
 32. Subirats, J. L., Franco, L., Molina, I., Jerez, J. M., Active learning using a constructive neural network algorithm, Franco, L., Elizondo, D., Jerez, J. M., (Eds.), *Constructive Neural Networks*, Springer, Chennai, India, 2009, pp. 193–206.
 33. EN 197-1:2011, Cement. Part 1: Composition, specifications and conformity criteria for common cements, CEN, Brussels, 2011, pp. 10–15.

34. EN 196-6:2010, Methods of testing cement – Part 6: Determination of fineness, CEN, Brussels, 2010, pp. 7–18.
35. EN 196-2:2013, Methods of testing cement – Part 2: Chemical analysis of cement, CEN, Brussels, 2013, pp. 25–26, 53.
36. EN 196-1:2005, Methods of testing cement – Part 1: Determination of strength, CEN, Brussels, 2005, pp. 18–21.
37. *Badde, D. S., Gupta, A. K., Patki, V. K.*, Cascade and feed forward back propagation artificial neural network models for prediction of compressive strength of ready mix concrete, in Second International Conference on Emerging Trends in Engineering, Civil Group, Vol. 3, IOSR-JMCE, Special Issue, 2013, pp 1–6.
38. *Shetin, V.*, A learning algorithm for evolving cascade neural networks, *Neural Process. Lett.* **17** (2003) 21. doi: <http://dx.doi.org/10.1023/A:1022935810223>
39. *Tsamatsoulis, D.*, Optimizing the cement compressive strength prediction by applying coupled linear models, *Mastorakis, N., Rudas, I.* (Eds.), Proceedings of WSEAS 17th International Conference on Automatic Control, Modeling & Simulation (ASMOS '15), Advances in Electrical and Computer Engineering, WSEAS Press, Tenerife, Spain, 2015, pp. 98–105.
40. *Tsamatsoulis, D., Nikolakakos, N.*, Optimizing the sulphates content of cement using multivariable modeling and uncertainty analysis, *Chem. Biochem. Eng. Q.* **27** (2013) 133.