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
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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{25} applied FL and GA techniques to predict the strength of cement at all the strength classes. Ren et al.\textsuperscript{26} applied generalized regression neural network (GRNN) techniques to predict the heat of hydration and compressive strength of cement. Yongzheng et al.\textsuperscript{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 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\textsuperscript{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 studies\textsuperscript{30} 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 ability\textsuperscript{31} 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.\textsuperscript{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 197-1:2011\textsuperscript{33} 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-
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 196-6; (ii) Specific surface, measured according to EN 196-6; (iii) Loss on ignition and insoluble residue of the cement, measured according to EN 196-2; (iv) SO₃, 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 196-1. 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 Halyps plant.

Mathematical models predicting strength

The common independent variables in all models are: Loss on ignition, LOI, sulfates content, SO₃, insoluble residue, Ins_Res, residue at 40 microns sieve, R40 and specific surface, S. 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_1 - constitutes an input variable, except the set of physical and chemical data. (ii) The second one

\begin{align*}
C₀S &= 0.47 \cdot (C - CaO_p) - 7.6 \cdot S - \\
& - 6.72 \cdot A - 1.43 \cdot F - 2.85 \cdot SO₃, \quad (1) \\
C₂S &= 2.87 \cdot S - 0.754 \cdot C₀S, \quad (2) \\
C₂A &= 2.65 \cdot A - 1.69 \cdot F, \quad (3) \\
C₂AF &= 3.04 \cdot F \quad (4)
\end{align*}

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 Halyps plant.
named Str_28_7, where the 7-day strength variable \(-\text{Str}_7\) is also included. In fact, weaker or stronger functions exist among the variables \(\text{LOI, Ins}_\text{Res, SO}_3, R40, S_p, \text{Str}_1, \text{Str}_7\). (a) A decrease in \(R40\) or increase in \(\text{LOI, Ins}_\text{Res, SO}_3\) generally causes an increase in \(S_p\). 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 \(\text{LOI, Ins}_\text{Res, R40}\) 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 \(\text{Str}_1\) causes also an increase in \(\text{Str}_7\), but \(\text{Str}_7\) is also related to clinker composition and activity. Thus, because the variables \(S_p, \text{Str}_1, \text{Str}_7\) 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 \(\text{Str}_28\) 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 \(\text{Str}_28\) and the addition of a correction computed from the moving difference between the calculated 28-day strength from \(\text{Str}_28\) and the measured one. The model was named \(\text{Str}_28\_\text{EW}\).

**Multiple linear regression**

The seven independent variables are named \(X_i\) with \(i=1\) to 7, where: \(X_i=\text{LOI, } X_i=\text{SO}_3, X_i=\text{Ins}_\text{Res, } X_i=S_p, X_i=R40, X_i=\text{Str}_1, X_i=\text{Str}_7\). The 28-day strength is a dependent variable, \(Y=\text{Str}_28\) and then algorithm goes on as follows:

(i) For a given data set, the minimum and maximum values of \(X_i\) and \(Y\), \(X_{1,\text{MIN}}\) \(<X_{1,\text{MAX}}\), \(Y_{\text{MIN}}\) \(<Y_{\text{MAX}}\), respectively, are computed.

(ii) The variables \(X_i, Y\) are normalized. The set of the new variables \(X_{N_i}\), \(Y_N\) is calculated from equations (5) and (6). The normalized data belong to the interval \([0, 1]\).

\[
X_{N_i} = \frac{X_i - X_{1,\text{MIN}}}{X_{1,\text{MAX}} - X_{1,\text{MIN}}} \quad I=1..N
\]  

\[
Y_N = \frac{Y - Y_{\text{MIN}}}{Y_{\text{MAX}} - Y_{\text{MIN}}} \quad (6)
\]

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

\[
Y_N = A_0 + \sum_{I=1}^{X} A_I \cdot X_{N_i}
\]  

(iv) The calculated 28-day strength, \(\text{Str}_{28}\_\text{Calc}\) is provided by equation (8).

\[
\text{Str}_{28}\_\text{Calc} = Y_{\text{MIN}} + Y_N \cdot (Y_{\text{MAX}} - Y_{\text{MIN}})
\]  

(v) For a total number of data sets equal to \(M\) and actual strength \(\text{Str}_{28}\_\text{Act}\), the coefficients \(A_i, I=0\) to \(N\), are computed by minimizing the residual error \(s^2_{\text{res}}\), which is calculated from the formula (9).

\[
s^2_{\text{res}} = \sum_{J=1}^{M} (\text{Str}_{28}\_\text{Calc} - \text{Str}_{28}\_\text{Act})^2 / (M - N - 1)
\]

**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 \(\text{St}_28\_1\). The model \(\text{Str}_28\_7\) includes an additional node in the first layer for the reception of \(\text{Str}_7\), 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, \(X_{N_i}, Y_N\) belong to the interval \([-1, 1]\). Then \(\text{Str}_28\) is back calculated from its normalized value by applying equation (12).

\[
X_{N_i} = \frac{2 \cdot X_i - (X_{1,\text{MIN}} + X_{1,\text{MIN}})}{X_{1,\text{MAX}} - X_{1,\text{MIN}}} \quad I=1..N
\]  

\[
Y_N = \frac{2 \cdot Y - (Y_{\text{MIN}} + Y_{\text{MAX}})}{Y_{\text{MAX}} - Y_{\text{MIN}}}
\]

\[
\text{Str}_{28}\_\text{Calc} = \frac{Y_{\text{MIN}} + Y_N \cdot (Y_{\text{MAX}} - Y_{\text{MIN}})}{2}
\]  

The activation functions are described by equations (13) to (16).
Sigmoid function
\[ o(J) = \frac{1}{1 + \exp \left( -\sum_{i=0}^{N} W_{i,j} \cdot XN_{j} \right) } \] (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
\[ o(J) = \frac{1 - \exp \left( -2 \cdot \sum_{i=0}^{N} W_{i,j} \cdot XN_{j} \right) }{1 + \exp \left( -2 \cdot \sum_{i=0}^{N} W_{i,j} \cdot XN_{j} \right) } \] (14)

Radial Basis function
\[ o(J) = \exp \left( -\sum_{i=1}^{N} (XN_{j} - X0_{i})^2 / \sigma_{i}^2 \right) \] (15)

where \( (X0_{0}, X0_{1}, \ldots X0_{n}) \) is the vector of the center of the radial basis function, while \( (\sigma_{0}, \sigma_{1}, \ldots \sigma_{N}) \) are variance parameters.

Output layer activation function
\[ YN = \sum_{J=1}^{N} V_{J} \cdot o(J) \] (16)

where \( N \) is 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.\(^3^7\). To develop this kind of ANN, the technique described by Shetinin\(^3^8\) 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 \( V_{I}, I = 1..4 \). 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 \( W_{01}, W_{02} \) follows. The output \( o(I) \) is multiplied with the weight \( V_{I} \) and is fed to the output node. These three synaptic connections are crucial for the prediction of the cement compressive strength.
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 $W_{1p}, W_{1p}, V_1, V_2$ 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 $S_r$. Due to the fundamental importance of the specific surface in relation to the hydration rate of cement hydration, it is concluded that $S_r$ 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, $S_r$, 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 $A_p$ which minimize $s_{Res,T_D}(f)$. Then the model with this set of parameters applies to the data of $T_F$ length to obtain $s_{Res,T_F}(f)$.

The number of the consecutive sets ($A_p, MSRE_{Past}$, $MSRE_{Futur}$) is $K_{TP}$ 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.
The coupling of models Str_28_1 and Str_28_7 makes use of the exponentially weighted moving average (EWMA) filter. 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):

$$Y(0) = X(0)$$  \hspace{1cm} (19)

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

$$Y(I) = \lambda \cdot X(I) + (1 - \lambda) \cdot Y(I - 1)$$  \hspace{1cm} (20)
(iii) If $\lambda = 1$, the value of moving average becomes equal to the current one. The smaller the $\lambda$ 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 $\text{Str}_{28}(I)$, and the computed one from $\text{Str}_{28, 7}(I)$ is $\text{Str}_{28, 7}(I)$. The difference $\text{Diff}(I)$ is defined by formula (21).

$$\text{Diff}(I) = \text{Str}_{28, 7}(I) - \text{Str}_{28}(I) \quad (21)$$

Afterwards, the following procedure applies:

(i) The moving average of $\text{Diff}(I)$, $\text{EW}_j \text{Diff}(I)$, is calculated by applying equation (20) for a predefined value of $\lambda$.

(ii) The corrected value $\text{Str}_{28, 1}(J)$, named $\text{Str}_{28, 1}(J)$, is computed from equation (22), where $J \geq I$.

$$\text{Str}_{28, 1}(J) = \text{Str}_{28, 1}(J) + k \cdot \text{EW}_j \text{Diff}(J) \quad (22)$$

(iii) The parameters $k$ and $\lambda$ 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 $(k, \lambda_{CEM \, II \, A-M \, 32.5})$ for CEM II B-M 32.5, and the set $(k_{CEM \, II \, A-L \, 42.5}, \lambda_{CEM \, II \, A-L \, 42.5})$ for CEM II B-M 32.5.

### Analysis of results and discussion

#### Preliminary processing of data

To examine the correlations among all of the input variables and $\text{Str}_{28}$, 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 $\text{Str}_{28}$ and $S_i$ is false because limestone, pozzolan, 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 $\text{Str}_{28}$ 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 $\text{Str}_{28}$ and LOI, $\text{Ins}_{Res}$, $\text{Str}_{1}$, $\text{Str}_{7}$. 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. Tsamatsoulis$^{29}$ 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 $\text{Str}_{28}$ 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 – $\text{Str}_{1}$ and $\text{Str}_{7}$ – 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 $\text{Str}_{28}$ are computed and plotted in Figure 4. A noticeable non-linearity occurs in the function between $\text{Str}_{1}$ and $\text{Str}_{28}$ 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 $\text{Str}_{7}$ and $\text{Str}_{28}$. Correlations between {SO$_3$, $R_{40}$, $S_i$} 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 $\text{Str}_{28}$ presents a clear maximum for SO$_3$ ε [2.4, 2.6]. This result agrees with a laboratory study of Tsamatsoulis et al.$^{30}$ 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 $\text{Str}_{28}$ is descending. The rate of strength reduction increases as $R40$ aug-

### Table 5 – Correlation coefficients between the output and input variables

<table>
<thead>
<tr>
<th></th>
<th>LOI</th>
<th>SO$_3$</th>
<th>Ins Res</th>
<th>$S_i$</th>
<th>$R40$</th>
<th>$\text{Str}_{1}$</th>
<th>$\text{Str}_{7}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Str}_{28}$</td>
<td>-0.86</td>
<td>0.68</td>
<td>-0.85</td>
<td>-0.74</td>
<td>-0.16</td>
<td>0.82</td>
<td>0.95</td>
</tr>
</tbody>
</table>
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 observed 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 increase 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 \( T_f = 180 \) days and the corresponding MSRE_\\text{P} 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 $MSRE_{\text{Future}}$ are determined. The training and test errors, $MSRE_{\text{Past}}$ and $MSRE_{\text{Future}}$, 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 $MSRE_{\text{Future}}$. Especially for $T_F \geq 20$ days, a worsening of $MSRE_{\text{Future}}$ 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](image)

![Fig. 7 – Training and test MSRE of Str_28_1 model](image)
ters 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 $T_p$ 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 training 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 $T_p$. 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 $\sigma$, 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 $T_F$ values from the set \{60, 90, 120, 180, 240, 300, 360, 540, 720\}, while $T_D$ 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 $\approx 0.02$ lower that which computed by MLR, for almost the entire range of $T_F$. (c) $MSRE_{Future}$ is a strong function of both $T_D$ and $T_F$ values, and for each $T_D$ value, there is a $T_F$ where $MSRE_{Future}$ becomes minimal. For the MLR and $S_1N_B$ techniques, these $T_F$ values of minimum future error
show an increasing trend as \( T_p \) increases: For small \( T_p, T_D = 120 \) days, while for higher \( T_p \) values, \( T_D \) increases to 180 days, reaching 240 days for \( T_p = 60 \) 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 \( T_p \) remaining continuously equal to 120 days as \( T_p \) varies from 1 to 60 days. (d) Small values of \( T_D - T_p = 60 \) 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 \((T_p, T_D) = (120, 1)\) and it is equal to 1.89 MPa. (f) For the same values of \( T_p \) and \( T_D \) 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 \((T_p, T_D) = (120, 1)\) and it is equal to 1.86 MPa. (h) Additionally the error of ANNs for \( T_p = 2 \) remains lower than that of MLR for \( T_p = 1 \). 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 \((T_p, T_D)\) in the grid \( T_D \in \{60 \ldots 720\}, T_p \in \{1 \ldots 60\} \) have been scanned, the global minimum \( MSRE_{Future} \) is assured. (j) For small \( T_p = 60 \) days or high \( T_p \geq 20 \) 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 \((T_p, T_D) = (120, 1)\) has also been studied by determining the pairs \((T_p, T_D)\) providing \( MSRE_{Future} \leq 1.02 \) 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.

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 $MSRE_{Train}$ 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. 10** – Optimum area of ($T_D$, $T_F$) for Str.28.1 model

**Fig. 11** – Training and test errors for $T_D$ $\in [60, 720]$, $T_F$ $\in [1, 60]$ and Str.28.7 model
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 $D_{B_2}(t)$ defined by equation (17). Str_28_EW contains four additional parameters, namely $k_{BM}, \lambda_{BM}, k_{AL}, \lambda_{AL}$ 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 $k_{BM}, \lambda_{BM}, k_{AL}, \lambda_{AL}$ minimum and maximum values, $k_{MIN}, k_{MAX}, \lambda_{MIN}, \lambda_{MAX}$ are selected. (b) Steps of change, $dk, d\lambda$ are also chosen. (c) The rectangular area defined by the vertices $k_{MIN}, k_{MAX}, \lambda_{MIN}, \lambda_{MAX}$ 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 $(T_{f}, T_{s})$ 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-
parameters \( T_D = 120 \), \( T_F = 1 \). For these three models, \( MSRE_{\text{Future}} \) as function of \( k_{BM}, \lambda_{BM} \) is plotted in Figures 12, 13, 14 for \((\lambda_{AL}, k_{AL}) = (1, 0.1)\) and \((0.7, 0.5)\). The \( MSRE_{\text{Future}} \) 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 \( k, \lambda \) are determined and applied. For each \( \lambda \) value there is only one \( k \) value where \( MSRE_{\text{Future}} \) becomes minimal. From the shape of the curves, it is concluded that increasing \( \lambda \), minimal \( MSRE_{\text{Future}} \) 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 \( \lambda \) and high values of \( k \), in the intervals \((0.1, 0.2)\) and \((0.9, 1.0)\) correspondingly.

The sensitivity of the optimum \((k, \lambda)\) is also studied by determining the pairs \((k, \lambda)\) which provide \( MSRE_{\text{Future}} \leq 1.005 \cdot MSRE_{\text{Future} \text{Min}} \) using interpolation between the discrete parameters and the results shown in Figures 12(a), 13(a), 14(a). This area of \((k_{BM}, \lambda_{BM})\) constitutes the optimum region. The respecting areas for MLR and the two ANN techniques are depicted in Figure 15, which shows that the sensitivity of the optimal \( k_{BM}, \lambda_{BM} \) is low. For HT_1N_B neural network, the range of \( k, \lambda \) of optimum test error is larger compared with S_1N_B for almost the same \( MSRE_{\text{Future Min}} \). 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 \( T_D \),
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 120 days, \(T_F = 1\) day.

---

Fig. 14 – Test errors for Str_28_EW model, HT_1N_B technique and (a) \((\lambda_{al}, k_{al}) = (0.1, 1)\), (b) \((\lambda_{al}, k_{al}) = (0.5, 7)\)
$T_F = 1$ are shown in Figure 17, for all the models and techniques. When $T_F$ increases, the rate that MSREFuture 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 MSREFuture 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.

Table 6 – Comparison of minimum MSREFuture between Str_28_1, Str_28_EW models for the optimal techniques

<table>
<thead>
<tr>
<th></th>
<th>$T_D$</th>
<th>$T_F$</th>
<th>MSREFuture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Str_28_1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLR</td>
<td>120</td>
<td>1</td>
<td>1.885</td>
</tr>
<tr>
<td>S_1N_B</td>
<td>120</td>
<td>1</td>
<td>1.857</td>
</tr>
<tr>
<td>HT_1N_B</td>
<td>120</td>
<td>1</td>
<td>1.858</td>
</tr>
<tr>
<td>Str_28_EW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLR</td>
<td>240</td>
<td>1</td>
<td>1.826</td>
</tr>
<tr>
<td>S_1N_B</td>
<td>240</td>
<td>1</td>
<td>1.785</td>
</tr>
<tr>
<td>HT_1N_B</td>
<td>240</td>
<td>1</td>
<td>1.803</td>
</tr>
</tbody>
</table>

Fig. 15 – Optimum area of $(k_{str}, \lambda_{str})$ for Str_28_EW model

Fig. 16 – Test errors for Str_28_EW model as function of $T_D$, $T_F$ and optimum values of $k$, $\lambda$
ANN has been created. Three types of activation nodes in the hidden layer. In parallel, a cascade developed involving three layers with one or two moving forward periods. Various ANNs have been used for a future period of \( T_f \) because they are applied to a movable past period of 1-day and 7-day strength. The models are dynamic resulting of\( T_f \) and \( T_p \) as an independent variable or both 1-day and 7-day strength results are used to predict the typical 28-day cement strength: Multiple linear regression (MLR), and artificial neural networks (ANN). The techniques applied, the best estimations were obtained for \( T_f = 1 \).

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 \( \lambda \) 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 \( T_f, k, \lambda \) 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.

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 \( T_p \) 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 \( T_p \) and \( T_f \), has been performed. For Str_28_1 model and for all the techniques applied, the best estimations were obtained for \( T_f = 1 \).
### Nomenclature

- **$A_0, A_1$** – Coefficients of equation (3)
- **$C_{3S}$** – Tricalcium silicate, %
- **$\text{Diff}(I)$** – Difference between calculated and actual 28-day strength at time $I$, MPa
- **$dk$** – step of change of parameter $k$
- **$\text{EW}_{\text{Diff}}(I)$** – EWMA difference between calculated and actual 28-day strength at time $I$, MPa
- **$I$** – Discrete time in equations (16), (17), days
- **$\text{Ins}_{\text{Res}}$** – Insoluble residue, %
- **$J$** – Discrete time in equation (18), days
- **$k$** – Weight parameter of the coupled model
- **$K_{10}$** – Number of consecutive data sets
- **$\text{LOI}$** – Loss on ignition, %
- **$M$** – Total number of data sets in equation (5)
- **$\text{MSRE}$** – Mean square residual error, MPa
- **$N$** – Number of coefficients
- **$N_1$** – Number of nodes in the hidden layer
- **$o(J)$** – Activation functions of neural networks
- **$R_{40}$** – Residue at 40 microns sieve, %
- **$S_r$** – Specific surface, 10 m² kg⁻¹
- **$\text{SO}_3$** – Sulfates content, %
- **$s_{\text{err}}$** – Residual error, MPa
- **$\text{Str}_{1}$** – Compressive strength at 1 day, MPa
- **$\text{Str}_{7}$** – Compressive strength at 7 days, MPa
- **$\text{Str}_{28}$** – Compressive strength at 28 days, MPa
- **$\text{Str}_{28,1}(I)$** – Calculated strength at 28 days from model $\text{Str}_{28,1}$ at time $I$, MPa
- **$\text{Str}_{28,7}(I)$** – Calculated strength at 28 days from model $\text{Str}_{28,7}$ at time $I$, MPa
- **$\text{Str}_{28,\text{EW}}(I)$** – Calculated strength at 28 days from model $\text{Str}_{28,\text{EW}}$ at time $I$, MPa
- **$t$** – Time, days
- **$T_p$** – 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)
- **$X_N_i$** – Normalized independent variables in equations (1), (3), (6), (9), (10), (11)
- **$X_O_i$** – Parameters of the center of the radial basis function
- **$Y$** – Dependent variable in equation (2), (7)
- **$Y(I)$** – Dependent variable in equation (16)
- **$Y_N$** – 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

- **$\lambda$** – Parameter of the EWMA model
- **$d\lambda$** – Step of change of parameter $\lambda$
- **$\sigma_i$** – 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

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.


