

NEURAL NETWORK BASED DATA-DRIVEN PREDICTOR: CASE STUDY ON CLINKER QUALITY PREDICTION

Mihreteab Negash* and Getachew Alemu*
School of Electrical and Computer Engineering
Addis Ababa Institute of Technology, Addis Ababa University

ABSTRACT

Soft sensors are key solutions in process industries. Important parameters which are difficult or cost a lot to measure can be predicted using soft sensors. In this paper neural network based clinker quality predictor is developed. The predictor genuinely estimates LSF, SM, AM and C₃S values.

There is a time delay while physically measuring clinker quality parameters. This can be avoided and quick control action can be taken by predicting the parameters. Many neural network based predictors have been developed in different application areas. However, this paper has its own new contribution. First it has developed data synthesis strategy. Besides, multiple and advanced neural network architectures are used to get improved result. Moreover, it is of the first kind for the selected case (Mugher cement factory).

Key words: Soft sensor, neural network, clinker quality prediction.

INTRODUCTION

Problem Description

There are two major problems while measuring clinker quality parameters. The first one is measurement delay. The time delay of the laboratory analysis is around four hours and that of the X-Ray Fluorescence (XRF) technique is about fifteen minutes. These delays cause hindrance in communicating quality report, resulting in difficulties of making timely control adjustment whenever it is required. This in turn significantly affects producing good quality clinker. The second one is lack of backup for the XRF. In case of malfunction, quality measurement cannot be done using the XRF and there is no backup to supplement due to its high cost. Therefore, measurement delays and absence of backup indicate the need for clinker quality prediction.

Significance of the Paper Work

The paper has two sided importance. First, the neural network based quality prediction will avoid measurement delay and enable quick control actions to be taken. Therefore, the developed predictor can supplement clinker quality measurement and will aid on effective kiln operation. As a result, clinker quality can be further improved. Secondly, as the first of its kind, it will create motivation on applying neural network based soft sensor to inland process industries. Thus, it is a significant paper work.

Cement Manufacturing Process

Cement production is an energy intensive manufacturing process. Portland cement, the most common type of cement, is produced by grinding an intermediate product called clinker. This product is produced in a rotary kiln, which contains nonlinearity, lag and there is no precise mathematical model to represent it [1].

There are four fundamental stages, as shown in Fig. 1, in the production of Portland cement: quarrying, raw material preparation, clinkering and cement milling [2-4].

*E-mail: gmhrn3@yahoo.com, getachew.alemu@aait.edu.et

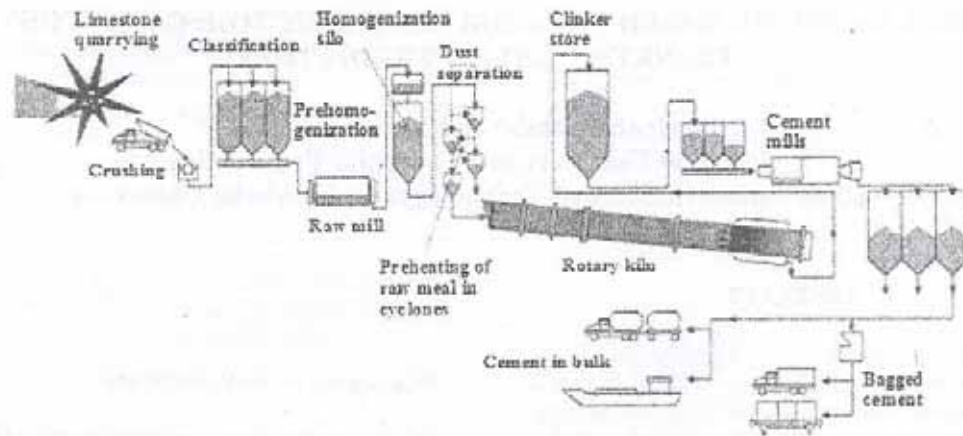


Figure 1 Cement manufacturing from limestone quarrying to cement bagging [11]

Step I: Quarrying

The raw materials (limestone, clay) for cement manufacture are quarried and stored separately.

Step II: Raw material preparation

The steps involved here depend on the process type used. For dry process, proportioned mix is milled without adding water. However, in wet process the mix is pulverized in the presence of water.

Step III: Clinkering

The powder from dry process or the slurry from wet process is heated in a rotary kiln and then cooled down. While it is being heated chemical reactions take place to form mineral phases of the clinker.

Step IV: Cement milling

The final product, which is cement, is produced by grinding clinker with gypsum, in a cement mill.

Clinker Formation

There is a set of reactions in the kiln to form clinker [2,3]. The decomposition of calcite and clay, and reaction to give belite, aluminate and ferrite occur below about 1300°C. In the temperature range of 1300°C-1400°C, a melt is formed. In the presence of the melt, belite and lime react to give alite. During cooling, the liquid crystallizes to give aluminate and ferrite. Thus, in such chain of reactions clinker minerals are produced. Fig. 2 shows the variation in phase content, with respect to temperature, during clinker formation.

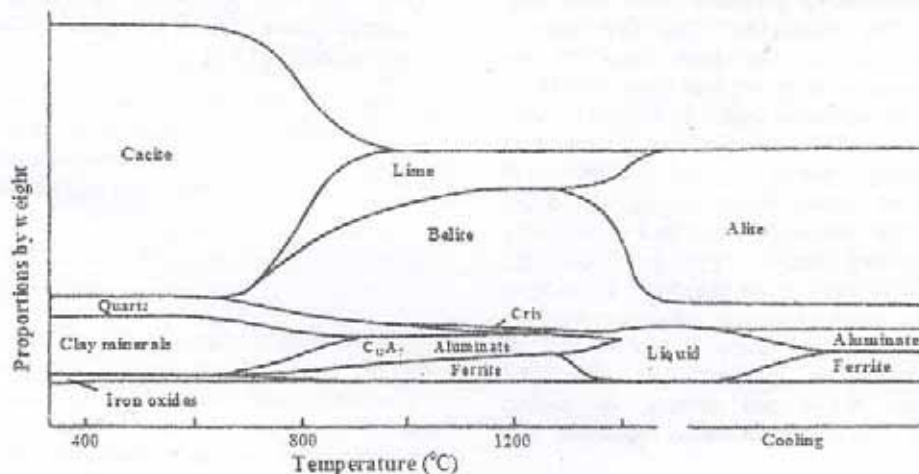


Figure 2 Phase variation during clinker formation [2]

LITERATURE REVIEW

Paper Works on Soft Sensors in Cement Factory

Soft sensors have application in predicting important process parameters. There are many papers on the application of soft sensors [1,5-10]. Among these, two papers are particular to cement factories and use neural network.

The first one is on temperature prediction. Neural network based temperature predictor for cement rotary kiln is reported in a paper [1]. In this predictor backpropagation and Elman neural network are used. This paper serves as a good illustration of the benefits of soft sensor in cement factory. However, it is not directly linked to clinker quality prediction, which is the focus of this work.

The other one is neural network based clinker quality parameters prediction [10]. It has used backpropagation neural network for simultaneous prediction of clinker quality parameters. Besides, the paper gives recommendation on using some other network architectures to further improve the result.

Data Collection and Variable Selection

An industrial database provides data of all the variables that are recorded. However, all the available variable data are not relevant to the process variable to be estimated. Thus, relevant variables are selected with the help of experts from Mughar cement factory. Table 1 shows the relevant variables.

METHODOLOGY

The basic steps in soft sensor design are data collection, data preprocessing, model selection & training and model validation [12]. These are shown in Fig. 3.

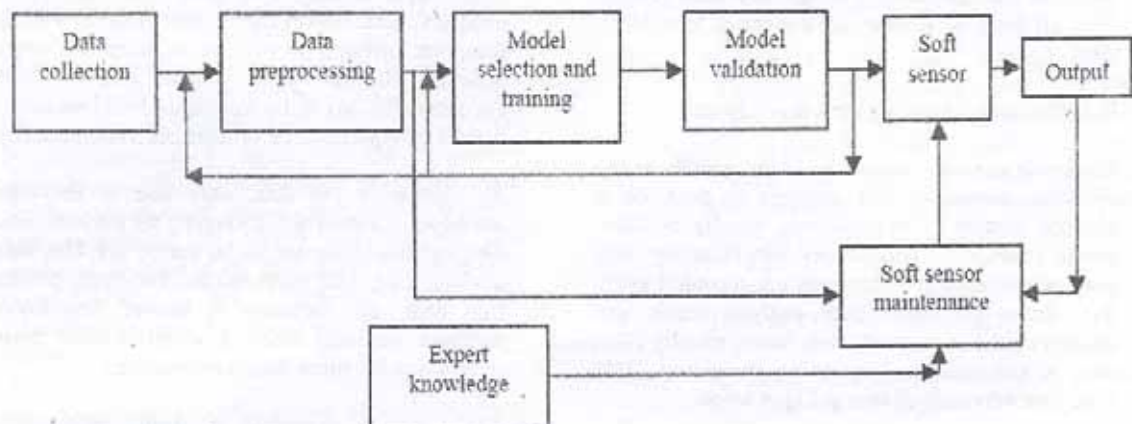


Figure 3 Soft sensor design steps

Table 1: The relevant variables for prediction

Input		Output
Operating parameters	Kiln meal variables	Clinker variables
Secondary air temperature, Calciner rising pipe temperature, Kiln gas flue chamber temperature, Quantity of oil before kiln burner, Quantity of oil before calciner burner, Kiln speed, Kiln meal flow rate and Hot gas ID fan power	SiO ₂ , Al ₂ O ₃ , Fe ₂ O ₃ and CaO	Lime saturation factor (L.SF), Silica modulus (SM), Alumina modulus (AM) and Alite (C ₃ S)

The Problem of Dimensionality Mismatch

The interval at which variables are measured creates a problem. Clinker data is measured everyday and kiln meal and operating parameters are measured every one hour. Thus, for hundreds of clinker data samples there will be thousands of kiln meal and operating parameters data samples.

Data Preprocessing

The collected historical data is not ready to use for training due to the presence of missing values, outliers and others. Thus, the data collected from the industrial database is subjected to appropriate treatment.

I. Correcting time format mismatch

The operating parameters, kiln meal and clinker variables are not recorded on uniform time format. Thus all the time formats are converted to military time.

II. Defining input-output time dependence

Clinker is analyzed once every 24hr, usually in the morning around 8. The analysis is done on a blended sample of many hours, usually of 24hr, where a sample is taken every 2hr. However, kiln meal and operating parameters are recorded every 1hr. Thus one set clinker analysis result will depend on the average of many hours, usually 24hr, data of kiln meal and operating parameters. This time dependence is shown in Fig. 4 below.

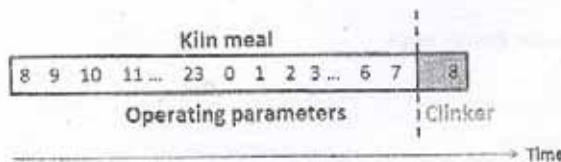


Figure 4 Input-output time dependence

III. Missing value imputation

Missing data corresponds to data values that should be present in a dataset but that, for various reasons, are absent [13]. Missing values are treated using known mathematical expressions and linear interpolation.

IV. Outlier detection

An outlier is an entry in a dataset that is anomalous with respect to the behavior seen in the majority of the other entries in the dataset [13]. Some outliers are detected using process knowledge. Whenever it is not possible to use process knowledge, the 3σ edit rule is used to detect outliers.

Data Synthesizing Strategy

After preprocessing the data, forty-eight cleaned input-output dataset is obtained from four month historical data. However, the smaller the size of the data, the difficult it will be to train the neural network. Besides, small dataset will not have enough variations to be representative. Therefore a dataset is required to be synthesized systematically.

To synthesize the data, only one of the input attributes is varied while keeping all the rest fixed. The variables that are to be varied are kiln meal attributes, i.e. kiln meal oxides. To obtain process like data, the variation is limited (employing different methods) within a sound possible range. In this way the input data is synthesized.

The output is generated by using mathematical expressions [2,3,14]. MATLAB is used to accompany the synthesizing process. Clinker oxides (corresponds to the output) are calculated from kiln meal oxides using Eq. (1-4). From these result the outputs are calculated using Eqs. (5-8). Eq. 8 is obtained using multiple regression. This is because the original equation, in Bogue set of

equations, for calculating the alite value is involved with variables that cannot be obtained using an equation similar to Eqs.(1-4). Thus, finally 144 input-output dataset is prepared using the data synthesizing strategy.

$$\begin{aligned} \text{\%CaO in Clinker} \\ = \frac{(\text{\% CaO in kiln meal}) 100}{100 - \text{LOI}} \end{aligned} \quad (1)$$

$$\begin{aligned} \text{\% SiO}_2 \text{ in Clinker} \\ = \frac{(\text{\% SiO}_2 \text{ kiln meal}) 100}{100 - \text{LOI}} \end{aligned} \quad (2)$$

$$\begin{aligned} \text{\% Al}_2\text{O}_3 \text{ in Clinker} \\ = \frac{(\text{\% Al}_2\text{O}_3 \text{ kiln meal}) 100}{100 - \text{LOI}} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{\% Fe}_2\text{O}_3 \text{ in Clinker} \\ = \frac{(\text{\% Fe}_2\text{O}_3 \text{ kiln meal}) 100}{100 - \text{LOI}} \end{aligned} \quad (4)$$

Where: LOI means Loss On Ignition
Lime Saturation factor (LSF)

$$= \frac{100 \text{ CaO}}{2.8 \text{ SiO}_2 + 1.18 \text{ Al}_2\text{O}_3 + 0.65 \text{ Fe}_2\text{O}_3} \quad (5)$$

Silica Module (SM)

$$= \frac{\text{SiO}_2}{\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3} \quad (6)$$

Alumina Module

$$= \frac{\text{Al}_2\text{O}_3}{\text{Fe}_2\text{O}_3} \quad (7)$$

$$\begin{aligned} \text{Alite (C3S)} = & -38.2749 + 4.3795 \text{ CaO} - \\ & 7.0331 \text{ SiO}_2 - 8.0014 \text{ Al}_2\text{O}_3 + 1.0512 \text{ Fe}_2\text{O}_3 \end{aligned} \quad (8)$$

Neural Network Model Selection and Training

Neural network models are selected to be compared with a proposed benchmark. Hidden layer size is selected by a rule of thumb followed by trial and error. The benchmark network is multiple input multiple output feedforward network (12-20-4). Using the benchmark, two models are selected. The first one is multiple input single output feedforward network (12-20-1). For ease of predicting single output per network than multiples, it is expected that the model will give better result. The second one is modified Elman network. This model is chosen because it is derived from one of the most known neural network, its dynamic nature could fit the delay feature of the kiln system and its strength for complex system modeling is indicated in some papers [1,15,16]. The training, validation and test sets have 0.7:0.15:0.15 proportion respectively. The architectures of the models are given in Fig. 5-7 below.

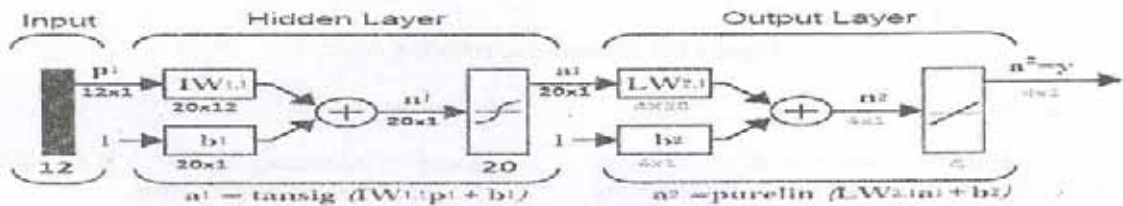


Figure 5 The benchmark network (The transfer function type is given in block)

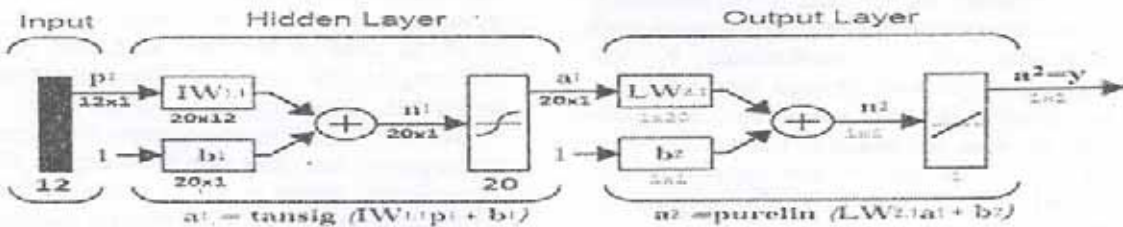


Figure 6 Multiple input single output feedforward network

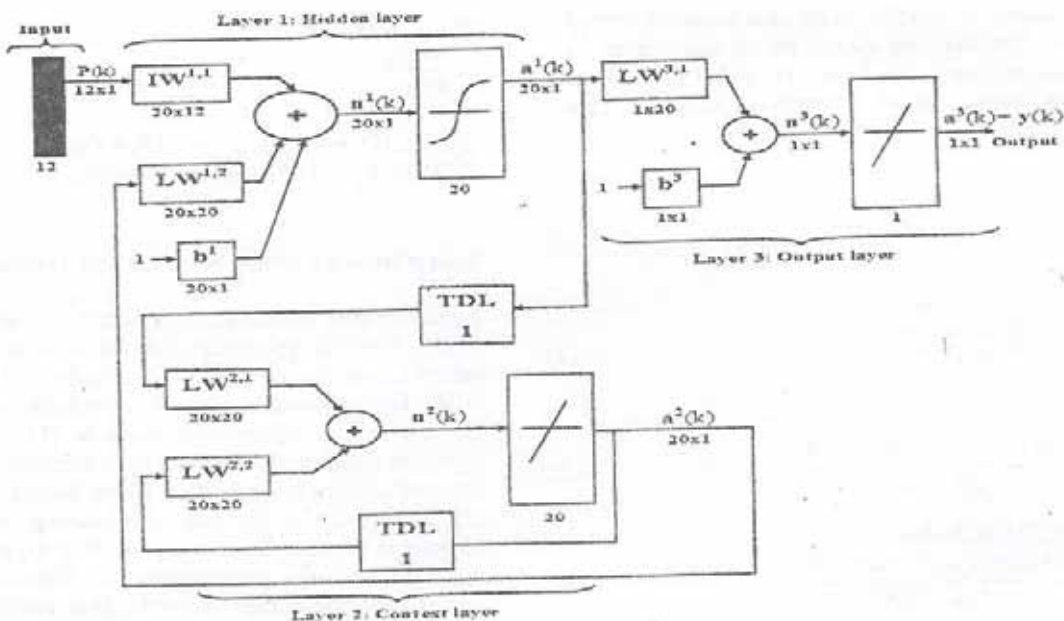


Figure 7 The modified Elman network

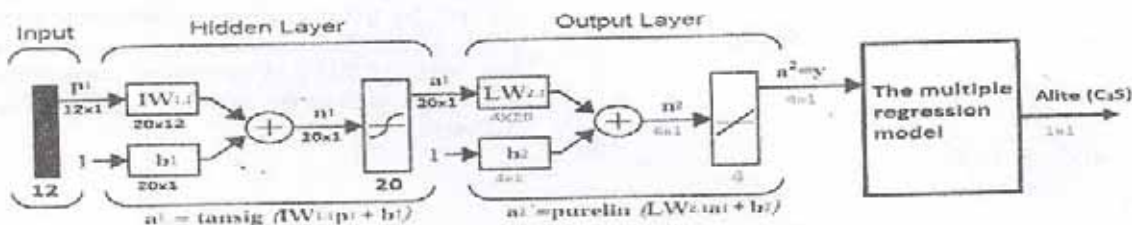


Figure 8 The improved alite predicting model

Network Features

The neural network models are developed using MATLAB. All the models share the following common features.

I. Normalization

It is important for the data to be normalized. Otherwise important process variables having small magnitudes will be overshadowed by less important variables having larger magnitudes. The normalization function 'mapminmax' is used to scale the inputs and targets so that they fall in the range [0, 1]. This function uses extreme values of the original data. The normalization is effectively part of the network.

II. Performance function

All of the network models use mean square error (MSE) as a performance function.

III. Training function

'trainscg' is the training function. It is the scaled conjugate gradient training algorithm.

RESULT AND DISCUSSION

The modified Elman model is generally superior in performance than all the other models. However, all of the three models have large error while predicting the alite value. Thus, improved network architecture, shown in Fig. 8, is developed to predict this value. This improvement is a simple rearrangement that rather than predicting the alite value directly, clinker oxides are predicted. These predicted values are used by the multiple regression model, which is now part of the improved alite predicting model, to calculate the alite value.

As a means of observing model performance, network output (predicted) and target output is plotted for the training and testing sets. The plots are for the models with best result. These plots are given in Fig. 9-12.

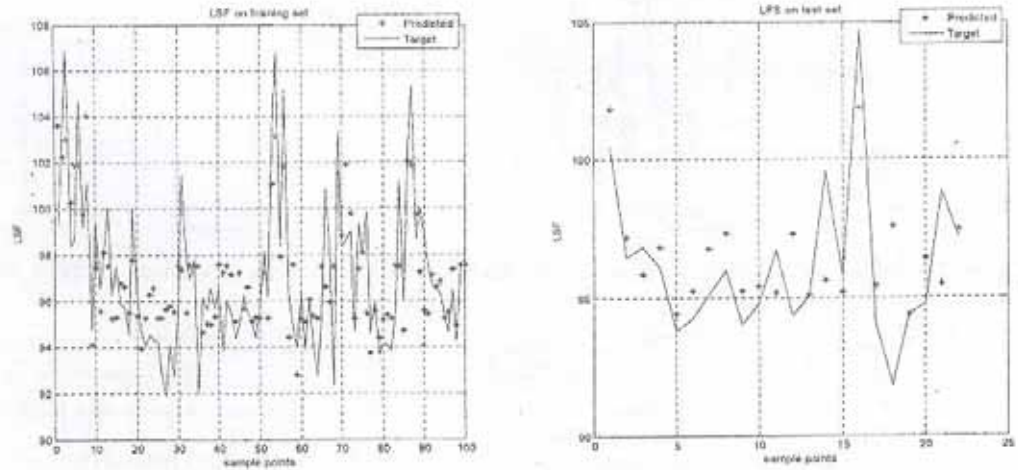


Figure 9 Target and predicted value of the modified Elman network. [Lime saturation factor sub model]

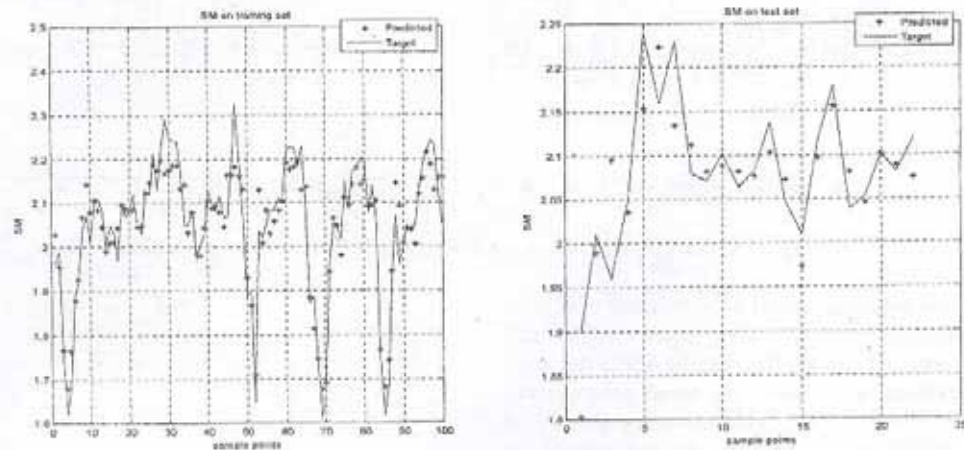


Figure 10 Target and predicted value of the modified Elman network.[Silica modulus sub model]

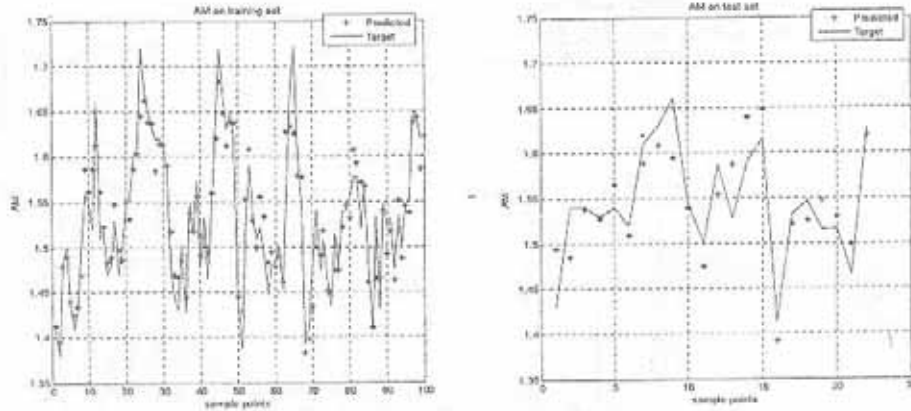


Figure 11 Target and predicted value of the modified Elman network. [Alumina modulus sub model]

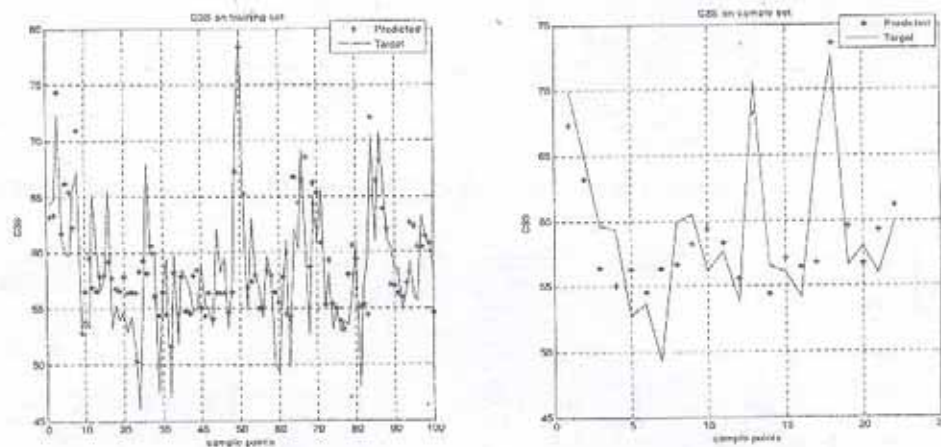


Figure 12 Target and predicted value of the improved alite predicting model

The neural network models are compared with each other. The comparison is done based on the mean square error on the test set for the set is not used while training. As a result, the mean square errors show the predicting performance of the models on new samples that the models are not trained on before. This performance comparison is presented in tabular form as shown in Tables 2-5.

Table 2: Mean square error [On lime saturation factor]

Models	Lime saturation factor (LSF)			
	MSE	LSF value		$mse(\%) = \frac{mse}{max} * 100$
		Minimum (min)	Maximum (max)	
Benchmark	9.0888	90.53	106.88	8.50
Multiple input single output (MISO)	4.5229	90.53	106.88	4.23
Modified Elman	4.3482	90.53	106.88	4.07

Table 3: Mean square error [On silica modulus]

Models	MSE	Silica modulus (SM)		$mse(\%) = \frac{mse}{max} * 100$
		SM value		
		Minimum (min)	Maximum (max)	
Benchmark	0.0459	1.61	2.33	1.97
Multiple input single output (MISO)	0.0039	1.61	2.33	0.17
Modified Elman	0.0027	1.61	2.33	0.12

Table 4: Mean square error [On alumina modulus]

Models	MSE	Alumina modulus (AM)		$mse(\%) = \frac{mse}{max} * 100$
		AM value		
		Minimum (min)	Maximum (max)	
Benchmark	0.0112	1.38	1.72	0.65
Multiple input single output (MISO)	0.0023	1.38	1.72	0.13
Modified Elman	0.0011	1.38	1.72	0.06

Table 5: Mean square error [On alit]

Models	MSE	Alite (C ₃ S)		$mse(\%) = \frac{mse}{max} * 100$
		C ₃ S value		
		Minimum (min)	Maximum (max)	
Benchmark	15.6418	44.72	79.60	19.65
Multiple input single output (MISO)	16.8973	44.72	79.60	21.23
Modified Elman	17.0127	44.72	79.60	21.37
Improved alite	10.8759	44.72	79.60	13.66

CONCLUSION AND RECOMMENDATION

In this paper a neural network based data driven clinker quality predictor is developed taking Mughar cement factory as a case study. This will avoid measurement delays in the laboratory analysis or X-Ray Fluorescence technique. As a result, quick control action can be taken which in turn contributes to quality improvement. The predictor developed by this work estimates LSF (mse=4.3482), SM (mse=0.0027), AM (mse=0.0011) and C₃S (mse= 10.8759) values of clinker.

The case factory is the source of data and it is recommended that the historical data to be digitalized. With the availability of digitalized data, data collection can be done easily. Thus, better result can be obtained by replacing synthesized data by the actual ones. Besides, it is recommended to use other neural network architectures with large input-output dataset.

REFERENCES

- [1] Baosheng, Y. and Xiushui, M., "Temperature Prediction and Analysis Based on BP and Elman Neural Network for Cement Rotary Kiln", Fourth International Seminar on Modern Cutting and Measurement Engineering, Proc. of SPIE, vol. 7997,79972I, 2011.
- [2] Taylor, H.F.W., "Cement Chemistry", Thomas Telford Publishing, 1997.
- [3] Hewlett, P.C., "Lea's Chemistry of Cement and Concrete", Elsevier, 1998.
- [4] Bhattu J.I., editor in chief, Miller, F. M. and Kosmatka S.H., editors, "Innovations in Portland Cement Manufacturing", Portland Cement Association, 2004.

- [5] Kadlec P., "On Robust and Adaptive Soft Sensors: A Thesis Submitted in Partial Fulfillment for the Degree of Doctor of Philosophy", Bournemouth University, December 17, 2009.
- [6] Najah A., Elshafie A., Karim O.A. and Jaffar O., "Prediction of Johor River Water Quality Parameters Using Artificial Neural Networks", European Journal of Scientific Research, ISSN 1450-216X, vol.28, no.3, 2009, pp.422-435.
- [7] Shi J. and Liu X., "Product Quality Prediction by a Neural Soft-sensor Based on MSA and PCA", International Journal of Automation and Computing 1, 2006, pp. 17-22.
- [8] Zhou Q., Xiong Z., Zhang J., and Xu, Y., "Hierarchical Neural Network Based Product Quality Prediction of Industrial Ethylene Pyrolysis Process", ISSN 2006, LNCS 3973, 2006, pp. 1132 – 1137.
- [9] Goyal, S. and Goyal, G.K., "Cascade and Feedforward Backpropagation Artificial Neural Network Models for Prediction of Sensory Quality of Instant Coffee Flavoured Sterilized Drink", Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition, vol. 2, no. 6, August 2011.
- [10] Pani, A.K., Vadlamudi, V., Bhargavi R.J. and Mohanta H.K., "Neural Network Soft Sensor Application in Cement Industry: Prediction of Clinker Quality Parameters", IEEE, 2011.
- [11] Kääntee U. et al., "Modelling a Cement Manufacturing Process to Study Possible Impacts of Alternative Fuels", Finland.
- [12] Pani, A.K. and Mohanta, H.K., "A Survey of Data Treatment Techniques for Soft Sensor Design", Chemical Product and Process Modeling, vol. 6, Iss. 1, article 2, 2011.
- [13] Pearson, R.K., "Mining Imperfect Data, Dealing with Contamination and Incomplete Records", Society for Industrial and Applied Mathematics, 2005.
- [14] Lackmariam G. and Leuel A., "Basic Concepts of Cement Production Technology: In house training material for front line operational personnel", Mugher Cement Factory, 2010.
- [15] Samek, D. and Bata, T., "Elman Neural Networks in Model Predictive Control", Proceedings 23rd European Conference on Modelling and Simulation, ISBN: 978-0-9553018-8-9.
- [16] Psent, D.T. and Llut, X., "Training of Elman Networks and Dynamic System Modelling", International Journal of Systems Science, vol. 27, no. 2, 1996, pp 221-226.