World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:6, No:1, 2012

Using Combination of Optimized Recurrent Neural Network with Design of Experiments and Regression for Control Chart Forecasting

R. Behmanesh, I. Rahimi

***Abstract*—**recurrent neural network (RNN) is an efficient tool for modeling production control process as well as modeling services. In this paper one RNN was combined with regression model and were employed in order to be checked whether the obtained data by the model in comparison with actual data, are valid for variable process control chart. Therefore, one maintenance process in workshop of Esfahan Oil Refining Co. (EORC) was taken for illustration of models. First, the regression was made for predicting the response time of process based upon determined factors, and then the error between actual and predicted response time as output and also the same factors as input were used in RNN. Finally, according to predicted data from combined model, it is scrutinized for test values in statistical process control whether forecasting efficiency is acceptable. Meanwhile, in training process of RNN, design of experiments was set so as to optimize the RNN.

***Keywords*—**RNN, DOE, regression, control chart.

I. INTRODUCTION

THE application of statistical methods to production quality control began in the early 1920. The Bell Telephone Company was the first to use statistical control charts and develop statistical acceptance sampling. Even though, importance of these techniques was really considered in the course of Second World War. Several companies adopted production control techniques because of their need to improve and control the quality of manufactured products. After creation American Society for Quality in 1946, it was observed clearly that ASQ persuaded companies to use quality improvement techniques not only for products but also services. Although, these techniques were not used in companies until 1960s in Japan and 1970s in Europe and America, the first companies to apply them were from the chemical manufacturing industry. Since the 1980s there have been major developments in statistical quality control techniques in numerous companies which have increased their competitive advantages considerably by applying these techniques [1].

One of the main used tools in statistical process control (SPC) is the control chart, also known as the Shewhart control chart that consists of a center line and two lines drawn parallel to it.

R. Behmanesh is with Khorasgan (Isfahan) Branch, Islamic Azad University, Isfahan, Iran (phone: +983115354001-9; fax: +983115354060; e-mail: rezaehs@yahoo.com).

I. Rahimi is with Department of Mathematical, University of Shahinshahr, Isfahan, Iran (e-mail: imanrahimi64@yahoo.com).

The center line represents the place where the characteristic measured should ideally be located and the parallel lines represent the control limits of the characteristic. The control limits are determined by statistical considerations. The use of control lines which group 99.7% of production data is very common when the production process is working correctly [1].We need an accurate knowledge of the production process to preserve product quality. This requires the automation of quality control systems and use of control charts as introduced by Shewart to observe the behavior of manufacturing process [2].

Control charting is the key point in SPC implementation. The correct application of these control charts requires satisfying statistical assumption such as the independence of random variable and symmetry in its probability distribution. If these assumptions are considered then the use of control charts is correctly applied since the upper and lower lines are established as 3sigma from the global mean of the X random variable [2]. In one study, the particleboard industry was used as a case study in prediction of variable process control by RNN, so that bending strength, modulus of elasticity, and internal bond strength were used as the most appropriate parameters for determining board quality [3], [4].

Recurrent neural networks are extensions of the multilayer feedforward neural networks, which employ feedback connections and have the potential to represent certain computational structures in a more parsimonious fashion [5]. Two fundamental ways can be used to add feedback into feedforward multilayer neural networks. Elman introduced feedback from the hidden layer to the context portion of the input layer. This approach pays more attention to the sequence of input values. Jordan recurrent neural networks used feedback from the output layer to the context nodes of the input layer and give more emphasis to the sequence of output values. On the other words, Close loop neural networks (recurrent neural networks) are related to the types of neural networks where a recurrent connection is implemented, taking into account the type of recurrent implemented [5].

Recurrent neural network applied to time series data. Network topology has n input units, m hidden neurons and one output neuron and q delays .The recurrent implemented back to the inputs r,is the output of the hidden layer for Elman network or back to the inputs r ,can be either the predicted output according to (1) or residual according to (2) for Jordan network.

$$\hat{x}_t \tag{1}$$

$$\hat{e}_t = x_t - \hat{x}_t \tag{2}$$

II. Methodology

A. Sampling and data collecting

In this study, the most key process of maintenance workshop in Esfahan Oil Refining Company (EORC) was considered as a case study of statistical process control and thereby response time of process from entering time of work order to delivering time of it in process and some effective factors on response time were collected as data. Aim of this paper is to predict variable process control by combination (association) regression and RNN models, in order that response time of process is variable in control chart, and also some factors such as distance entering time between work order and work piece, priority for maintenance, and maintenance workshop type were identified as the most effective independent variables on control of the response time quality as a service quality. For this purpose, regression and RNN models were taken. In the first place, regression equation of prediction the response time was obtained based upon three independent variables, then same factors and error obtained between actual response and predicted response were taken into account as input and output of RNN. Finally, the combination system in accordance with independent factors will predict control charts of process. Required data based on 1800 work orders were randomized as sample so that all of them determined regression model and then in the next stage, 900 out of them were used for training and 900 out of them for testing in RNN.

B. RNN training algorithm

For training step we have this algorithm:

1. We scaled all data in the range [-1, 1], that was the case for the test problem which we have used.
2. The next step was to divide the data up into training, validation and test subsets. We took one fourth of the data for the validation set, one fourth for the test set and one half for the training set.
3. After training network we converted the network output back into the original units.

The performance of a trained network was measured by the regression analysis between the network response and the corresponding targets.

For this option we measured three parameters m and b and r, pertain to the slope, the y-intercept of the best linear regression relating targets to network output, and correlation coefficient (R-value) between the outputs and targets, respectively. In case of the perfect fit (outputs exactly equal to the targets) the slope would be 1, and the y-intercept would be 0, and the R-value would be 1.

III. Results

A. Regression

As it is shown in table I aforementioned effective factors on response time constituted the linear regression model so that

TABLE I
LINEAR REGRESSION MODEL

Linear regression model	R^2	R^2(adj)	R
Resp = 0.343 + 0.130 F1 + 1.34 F2 - 0.006 F3	0.976	0.971	0.988

R^2 (ADJ): ADJUSTED DETERMINATION COEFFICIENT, R: CORRELATION COEFFICIENT

TABLE II
Significance of predictors in model

Predictor	Coef	SE Coef	T-value	P-value(sig)
Constant	0.3432	0.2674	1.28	0.2
F1	0.39963	0.02075	2.83	0.047
F2	1.33763	0.00494	270.72	0.000
F3	-0.0065	0.1603	-.0.04	0.968

adjusted determination coefficient in model indicates capability of forecasting for 97% of data. Also coefficient of every factor along with constant is shown in table II and according to these results it must be pointed out that F1 and F2 are significant in prediction of the response time whereas F3 is not helpful to predict it. Therefore coefficient of F3 is very low so as to predict the response time.

Analysis of variance is indicated in table III, according to

TABLE III
ANOVA (ANALYSIS OF VARIANCE)

Source	DF	SS	MS	F	P-value(sig)
Regression	3	930728	310243	24620.47	0.000
Residual Error	1796	22631	13		
Total	1799	953359			

the results of this, it is demonstrated that regression model is significant; on the other hand residual error is very low and hence it would be insignificant in predicting the response time.

TABLE IV
TRAINING SET AND STRUCTURE OF RNN FOR PREDICTION

Net	Parameters	Activation function
Number of hidden layers	2	tansig
Number of output layer	1	purelin
Train function	Trainbr	-
Performance function	MSE	-
Input delay	0	-
Recurrent connection	2	-
Input connection	All layers	-
Number neurons in first hidden layer	10	tansig
Number neurons in second hidden layer	10	tansig
momentum	0.8	-

B. RNN

Training set for recurrent neural network, structure of recurrent neural network and its parameters are shown in table IV. Also, existing activation functions in the structure have been determined.

C. Combination RNN and regression

It must be pointed out that predicted response time is obtained by adding the predicted variable from RNN and predicted variable from regression model according to (3). Equation (4) is based of literature [6]. To evaluate the result of the combination RNN and regression model, the prediction error was calculated according to (4).

$$V_p = V_{reg} + V_{ann} \tag{3}$$

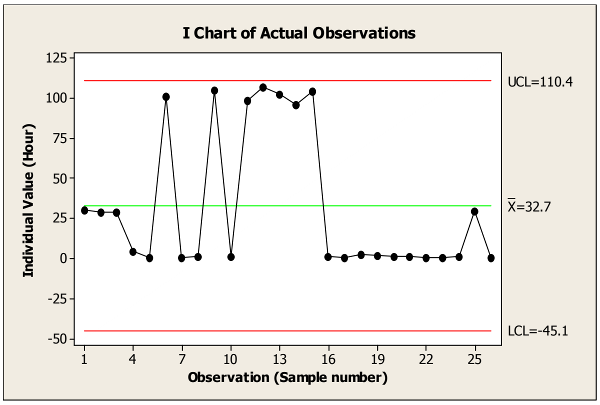
$$E\% = \frac{V_p - V_o}{V_o} \times 100 \tag{4}$$

E%: prediction error, V_{reg}: variable predicted by regression, V_{rnn}: variable predicted by RNN, V_p: variable predicted by RNN and regression combination, V_o: variable observed in testing.

It must be demonstrated that a prediction error of 15% was regarded as acceptable for a service process or production process and from 20 to 30% it was regarded as reject [7], [8]. As indicated in table V, due to this reason that prediction errors calculated by combination of RNN (on the testing group) and regression is less than 15%, it has to stated that this model can be regarded as valid. For this study, two samples of control chart with actual variables and prediction of RNN and regression were shown in Fig 1.

TABLE V
FORECASTING EFFICIENCY OF MODEL

Predictor models	Error%	Error range%
RNN and regression for X-bar chart	10.53	0.012 - 44.9



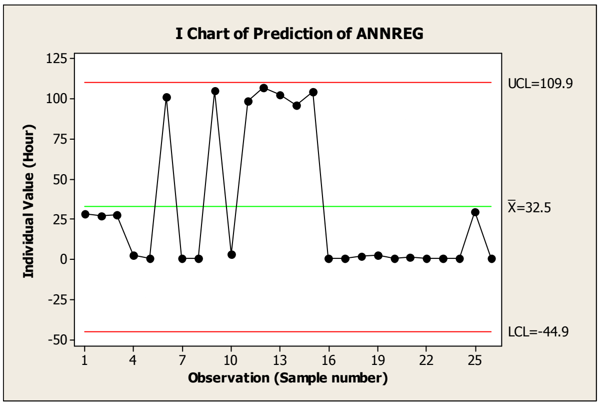


Fig. 1 comparisons of actual control chart and predicted.

D. Experimental setting

In this paper, the factors that affect the RNN's accuracy were studied so that they can be measured by root mean square error (RMSE), [9]. Three factors were taken into account in the experiments as they are factors that often found to be important as reported in literatures [9-10]. These factors and their setting are indicated in table VI.

TABLE VI
FACTORS AND THEIR LEVELS FOR OPTIMIZING RNN

Factor	Level 1(-)	Level 2(+)
Momentum (Factor A)	0.1	0.8
Number of neurons in hidden layer 1 (Factor B)	2	10
Number of neurons in hidden layer 2 (Factor C)	2	10

Two levels of number of neurons in hidden layer 1 are 2 and 10 neurons. Results shown those more than 10 neurons do not improve much of the network accuracy [10]. Consequently, this setting was selected. Likely first hidden layer, number of neuron in the second layer was set. Then momentum was set at 0.1 and 0.8 as last factor. Experimental setting is shown in Table VII. Full factorial design for three factors, namely 2^3 designs were carried out, results in 8 experimental runs. The experiments were run at four replicates per each setting. As a result, a total of 32 runs were conducted. Experiments were carried out according to run order. For example, the first experiment was carried out at 10 neurons in the first hidden layer, 2 neurons in the second layer, and momentum term equals to 0.8.

TABLE VII
SETTING AND EXPERIMENTAL RESULTS (RMSE)

Run Order	Factor A	Factor B	Factor C	RMSE
1	1	-1	1	3.435113
2	-1	1	1	2.920616
3	1	1	1	2.901724
4	-1	-1	1	3.63318
5	-1	-1	1	3.646917
6	1	-1	1	3.405877
7	-1	1	-1	3.646917
8	-1	-1	-1	5.656854
9	1	-1	-1	4.41588
10	1	-1	1	3.577709
11	1	-1	-1	5.108816
12	1	1	-1	3.794733
13	-1	1	-1	3.674235
14	-1	-1	-1	5.700877
15	1	1	-1	3.820995
16	1	-1	-1	4.1833
17	-1	-1	-1	4.266146
18	1	-1	1	3.847077
19	-1	1	1	2.91719
20	-1	-1	-1	4.358899
21	-1	1	1	2.924038
22	1	1	1	2.75681
23	1	1	1	2.798214
24	1	-1	-1	5.09902
25	-1	-1	1	3.563706
26	1	1	-1	3.478505
27	1	1	-1	3.464102
28	-1	-1	1	4.38178
29	-1	1	-1	5.700877
30	1	1	1	2.796426
31	-1	1	1	3
32	-1	1	-1	3.464102

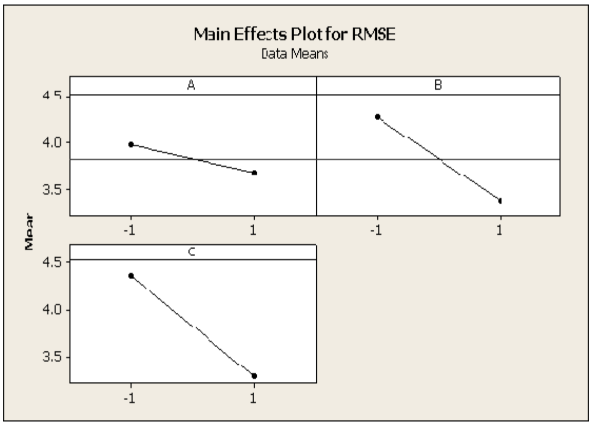
E. Experimental results and discussion

Analysis of variance (ANOVA) was taken based on the unseen testing data at 95% confidence level. The analysis was prepared by MINITAB software. Table VII shows the estimated effects of each parameters and coefficient for RMSE. The term A*C is the interaction between factors A and C In this analysis; only two-way and was used. Each column in table VIII includes information concerning the determination of the significant of each term to RMSE. The p-value in last column determines which of the effects are significant. In this study 95% confidence was used therefore terms that have p-value lower than 0.05 are significant. Table VIII showed that just factors B and C are significant.

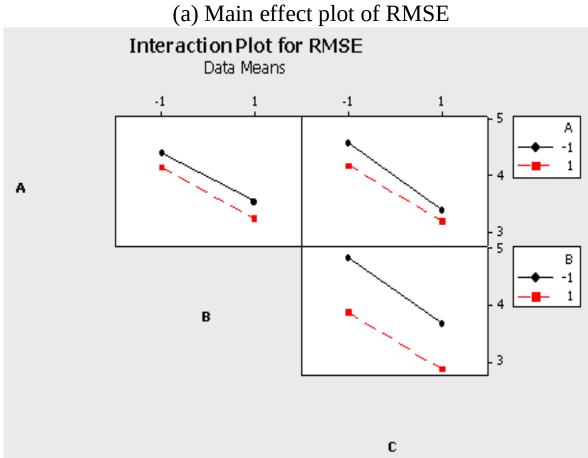
The effect of these factors is shown in Fig. 2. Fig. 2(a) shows the main effect plots of RMSE. The fig. indicates that number of neurons in hidden layer 1 at low level (2 neurons) prompt to lower RMSE. On the other hand, number of neurons in hidden layer 2 at high level (6 neurons) result in best accuracy. But, changing the level in momentum has no significant effect on response. Because, the best setting can not be determined from the main effect plot, hence the interaction plot (Fig. 2(b)) has to be considered. Fig. 2(b) shows the two-factor interaction plots among parameters. For example, the below subfigure indicates the interaction between number of neuron in hidden layer 1 (Factor B) and number of neuron in hidden layer 2 (Factor C). So in previous case, this plot shows that the effect of the number of neurons in layer 2 on the average RMSE is constant when the number of neurons in hidden layer 1 is at low level or high level. Therefore, simultaneous effect of these 2 factors on RMSE is insignificant. Consequently, in this study, the best setting of factors were found from the setting that provide lowest average RMSE (Table VII) and so the best setting in this case was 10 neurons in the first hidden layer, 10 neurons in the second layer and 0.8 momentum term. However, effect of momentum is very low on RMSE.

TABLE VIII
EFFECTS AND COEFFICIENTS FOR RMSE (CODED UNITS)

Term	Effect	Coef	SE Coef	T	P
Constant		3.8231	0.09267	41.26	0.000
A	-0.2858	-0.1429	0.09267	-1.54	0.136
B	-0.8889	-0.4444	0.09267	-4.80	0.000
C	-1.0830	-0.5415	0.09267	-5.84	0.000
A*B	-0.0188	-0.0094	0.09267	-0.10	0.920
A*C	0.1022	0.0511	0.09267	0.55	0.586
B*C	0.0793	0.0397	0.09267	0.43	0.673
A*B*C	0.0752	0.0376	0.09267	0.41	0.689



(a) Main effect plot of RMSE



(b) Interaction plot of RMSE
Fig. 2 Main effect and interaction plot

IV. CONCLUSION

In accordance with results and discussions, it can be concluded that integrated model consist of regression and RNN with high ability and accuracy on predicting the control charts servicing is efficient forecasting model with 89.47% accuracy. This combination was conducted in order to use more efficiency and accuracy in forecasting process. Also, for improving RNN so as to be more accurate in predicting, design of experiments was set based on the most important factors which had been introduced before in several literatures. As a result, best factors and their level setting such as number of neurons in hidden layer 1, number of neurons in hidden layer 2, and momentum with level settings of 10, 10, 0.8 respectively, were determined that lead to lowest RMSE in RNN training process and hence, optimizing the RNN in this case.

REFERENCES

[1] D. MONTGOMERY, "Control estadístico de la calidad", 2004, 3rd Ed. *Limusa Wiley*, Mexico, pp.797.

[2] J. A. Vazquez-Lopez, I. Lopez-Juarez, M. Peña-Cabrera, " On the Use of the Fuzzy ARTMAP Neural Network for Pattern Recognition in Statistical Process Control using a Factorial Design ", Int. *J. of Computers, Communications & Control*, ISSN 1841-9836, Vol. V (2010), No. 2, pp. 205-215.

[3] A.J. MORRIS, G.A. MONTAGUE, M.J. WILLIS, "Artificial neural networks : studies in process modeling and control", 1994, *Chem. Eng.* Res. Des. 72(A1), pp. 3-19.

[4] D.F. COOK, C.T. RAGSDALE, R.L. MAJOR, "Combining a neural network with a genetic algorithm for process parameter optimization", 2000, *Eng.Appl.Artif. Intell.* 13, pp. 391-396.

[5] L.R. MedsKer, L.C. Jain, "Recurrent Neural Networks : Design and Application", *CRC Press, Boca Raton London New York Washington, D.C.*, 2001.

[6] L.G. Esteban, F. García Fernández, P. de Palacios, M. Conde, "Artificial neural networks in variable process control: application in particleboard manufacture", *Sistemas y Recursos Forestales* 2009 18(1), pp. 92-100.

[7] D.F. COOK, C.C. CHIU, "Predicting the internal bond strength of particleboard: utilizing a radial basis function neural network", 1997, *Eng. Appl. Artif. Intell*. 10 (2), pp. 171-177.

[8] S. MALINOV, W. SHA, J.J. MCKEOWN, "Modelling the correlation between processing parameters and properties in titanium alloys using artificial neural network", 2001, *Comput. Mater.* Sci. 21, pp. 375-394.

[9] W. Laosiritaworn, N. Chotchaithanakorn, "Artificial Neural Networks Parameters Optimization with Design of Experiments: An Application in Ferromagnetic Materials Modeling", *Chiang Mai J.* Sci. 2009; 36(1), pp.83-91.

[10] M.S. Packianather, P.R. Drake, H. Rowland, "Optimizing the Parameters of Multilayered Feed Forward Neural Networks through Taguchi Design of Experiments", *Qual. Reliab. Eng. Int*., 2000; 16: pp. 461-473.