Journal of Sciences, Islamic Republic of Iran 31(3): 259 - 275 (2020) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

An Artificial Neural Network Model for Prediction of the Operational Parameters of Centrifugal Compressors: An Alternative Comparison Method for Regression

S. H. Ebrahimi[1*], and A.J. Afshari[2]

[1] *Department of Industrial Engineering, Shomal University, Amol, Islamic Republic of Iran*
[2] *Department of Industrial Engineering, University of MehrAlborz, Tehran, Islamic Republic of Iran*

Received: 14 February 2020 / Revised: 1 June 2020 / Accepted: 21 July 2020

Abstract

Nowadays, centrifugal compressors are commonly used in the oil and gas industry, particularly in the energy transmission facilities just like a gas pipeline stations. Therefore, these machines with different operational circumstances and thermodynamic characteristics are to be exploited according to the operational necessities. Generally, the most important operational parameters of a gas pipeline booster station includes the compressor's input and output pressures, input and output temperatures and also the flow rate passing from the compressors. Different values of those parameters related to every point of operational conditions will exactly affect on the compressor poly-tropic efficiency and their driver fuel consumption. Although, calculating of the poly tropic efficiency and fuel consumption using the existing thermodynamic relations, would need to apply rather awkward equations for each operating point. In this research, a feed forward perceptron artificial neural network is presented to predict the output operational conditions. The network would be trained at least in two scenarios applying by practical data in the neuro solution software version.5 using the Levenberg-Marquadt algorithm and the optimum model is experimentally selected according to R^2, MSE and NMSE.

Keywords: Centrifugal compressor; Artificial neural network; Ridge regression; Performance prediction; Pipeline gas booster station.

Introduction

A gas transmission network generally consists of one or more operational gas compressor stations on the route of the main gas pipeline, in which a few turbo compressors are equipped according to the predefined operational demands, particularly the vital parameters of the network such as the required consumption flow rate and supplying the upstream pressure. The overall network monitoring of the consumption points is assigned to the gas transmission control and monitoring center just known as dispatching office that coordinates their demands with the gas compressor station operators to put some turbo compressors in service. Generally, providing the dispatcher demand based on the desirable passing flow rate and the output pressure of a gas compressor station are empirically met by two parameters of the number of the turbo compressor

* Corresponding author: Tel: +989111533909; Fax: +981144203755; Email: shebrahimi1978@gmail.com

running and their speeds notwithstanding the turbo compressor efficiencies in period of time. The absence of a unified database as a predetermined inferential information system of the important operational variables of turbo compressors can lead to imposition of additional operating costs specially the fuel consumption of gas turbine just used for the compressor driver. Moreover the ambiguity and confusion of the effectively recognition of how many turbo compressors to be run, would cause to depreciation of the hot parts and prematurity of the overhaul time.

Today, applying the predictive and modeling tools to create a robust database system to manage the data are common and practical in industry and it would be expectedly accessed to achieve the judgmental scope as a decision support system by a suitable design of the patterns and models. Data mining uses a variety of techniques to find hidden patterns and relationships in large pools of data and infer rules from them that can be used to predict future behavior and guide decision making [1]. The neural network is invented and also extended, just like a paralleled data mining analyzer tool, interpreting the phenomena's physical and arithmetic laws which modeled from the human nervous processing system.

The present study aims to develop a feed forward perceptron neural network to smoothly predict the effective operational parameter values of turbo compressors situated at gas transmission network's stations, such as the outlet pressure, the outlet temperature and the flow rate passing across each compressor. The model's input elements are also as the compressor's inlet pressure and temperature, ambient temperature, number of compressors (appointed to be run) just considered as a control variable and speed of each compressor running. Then it is to be continued about the relevant researches, model presentation and how to achieve the optimal network, model validation by real operational data and also the model input parameters sensitivity analysis.

Besides, the ridge regression model is deployed for a comparison to feed forward perceptron. Hence, there would be rather intensive collinearity between the variables used , then the ridge regression is selected. Ridge regression is a technique for analyzing multiple regression data that suffer from multicollinearity.

When multicollinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value. Actually, ridge regression is a remedial measure taken to alleviate multicollinearity amongst regression predictor variables in a model.

Paper structure

The research is followed by some literature of review associated with using the ANN, specially for compressor map prediction. Furthermore, the research is continued by methodology section. It includes of issues which discussing about the ANN structure, curve fitting of data and collinearity challenges among the input vector (consisting of P_{in}, T_{in} , T_{amb} ,P_{out} , T_{out}). After proofing the non orthogonality of input variables , the ridge regression model is deployed to have a model comparison with the ANN model selected earlier. Finally, a case study and sensitivity analysis is then applied for the ANN model output (P_{out} , T_{out} , q) and evaluating the compressor's vital characteristics including the H_P , η_P and FC.

Literature

The modern view of neural networks began in the 1940s with the work of warren McCulloch and Pitts who showed that networks of artificial neurons can in principle compute any arithmetic or logical function [2] . Moraal's results show that ANN modeling could be superior to other curve fitting techniques if the model is sufficiently trained [3]. Today the ANN is widely used to explain and model the physical phenomena of every kind of industrial machines. Thus interpreting the turbo compressor's thermodynamical behavior isn't excluded as a typical kind of industrial machines used at gas transmission pipeline.

Ghorbanian and Gholamrezaei used the ANN to predict the eight stages axial compressor performance map (pressure ratio- flow rate curve).They developed two hidden layers with ten neurons into their presented model. Although they also made use of two models to predict the operational performance curves. In the first model the pressure ratio is function of mass rate and vice versa, in the second, the mass rate is function of pressure ratio at constant speed. They ran 42 real operational points as the working data, into their model [4]. In another study, Bao et al. generated a back propagation neural network for modeling of helical and centrifugal compressor performance map [5]. The BPNN was also used to map the compressor working curve by Yu et al. [6]. Torabian and Karimian deployed a three layers ANN which trained by the BP algorithm to forecast the performance map of centrifugal compressor located at gas booster station. Their model calculates the outlet flow rate of compressor in terms of variety of different speeds and input pressures. They also used vendor's data to train the model [7]. Sanaye et al. simulated the mass rate and the output temperature of rotary blade compressor, by ANN according to its speed, input temperature and output pressure[8]. Sue young et al. also used ANN to design a three dimensional geometrical shape of centrifugal compressor impeller. The impeller figure characteristics are generally dependent on the input & output pressure profile on the compressor hub and impeller peak point [9]. Shaojun et al. made use of the back propagation neural network to forecast the separating centrifugal compressor map in ethylen utilities[10]. Thomas et al. applied a feed forward perceptron neural network for predicting of an axial two stages compressor. They simulated the compressor performance map, changing the spaces between the different stages of compressor and considering the pressure ratio as a function of mass flow rate and compressor speed and also the mass flow rate as a function of pressure ratio and speed. They assigned the 206 field data which divided into 60% for training , 20% for cross validation and 20% to test the model results. Their model consists of a hidden layer that emerged to the optimum structure by changing the neuron numbers from 1 neuron to 10 [11]. Pin Chen et al. used the back propagation neural network to model a predictive control system of supply centrifugal air compressor at the unit of nitric acid in China's petrochemical industry. The number of hidden layers 1, 3 and 17 were considered to determine the optimal model in their study [12]. Yang et al. offered a neural network model to predict the ison-tropic and volumetric efficiency of positive displacement compressor of refrigeration system. The compression ratio, the condenser temperature and the evaporator temperature are the inputs of the neural network in their model and the compressor speed is looked as the network output [13]. Vilalta et al. studied the accuracy and efficiency of various regression models and artificial neural networks in modeling the compressor pressure ratio, given the mass flow rate and rotational speed of the centrifugal compressor [14]. Fei et al. proposed an artificial neural network integrating feed-forward back-propagation neural network with Gaussian kernel function to predict the compressor pressure ratio[15]. Li et al. developed regression model to predict both the pressure ratio and the efficiency of a centrifugal compressor using partial least squares [16].

Most previous studies have been generally done in order to plot the performance map of a centrifugal compressor or to determine its geometrical designing parameters. The model presented in this study is flexible due to the number of paralleled units to be put into operation as a controlling input variable of the model and also is practical just using the real data, so it can be effectively applicable as a decision making support system to simulate the output data, by operators and dispatcher.

Research Methodology

In the late 1950s, Frank Rosenblatt and several other researchers developed a class of neural networks called perceptrons. The neurons in these networks were similar to those of McCulloch and Pitts. Rosenblatt's key contribution was the introduction of a learning rule for training perceptron networks to solve pattern recognition problems [2].

Up to now, generalizing of these kinds of perceptron networks in the proposed multi-layer models and also in increasing of the number of neurons in hidden layer, are progressed to explain the natural phenomena and scientific modeling.

Putting the turbo compressors of gas booster station into operation in a transmission network would be a challenging topic and lack of a predefined and authentic information system as an operational data bank in order to support the operators, is missing to best decide for the output parameters based on current conditions as well. Generally speaking, the most important input data of a gas booster station are the gas input temperature and pressure and the ambient temperature which often the operators would control the gas flow rate and the compressor's outlet pressure and temperature by means of the number of units to be candidates for service and the power turbine speed, as the two main degree of freedom. It would be explicable that the number of compressor to turn on, is selected by the operator according to existing conditions in order to supply the needed flow rate. All turbo compressors minus one can be commonly turned on to pass the demanded flow rate through a gas booster station .For example for a gas booster station with three compressors, two units usually run to service. Control room operators (no need to recycle the flow) can experimentally run a turbo compressor relatively at high speed or two compressors at lower speeds to maximize the gas flow rate through a typical station. These two distinct scenarios of operational conditions would lead to select the compressors with different poly-tropic efficiencies and also imposition of the cost of fuel consumed (in gas turbines) to drive the compressors. Consequently, the factor of number of compressors to be turned on, is as an input controlling variable of the proposed model.

In this research a feed forward neural network equipped with two hidden layers is used to predict the gas flow rate, outlet pressure and output temperature based on the input pressure and temperature, ambient temperature, compressor speed and the number of units in service (as a control variable). It must be said that all the pressures and temperatures are respectively scaled in Bar and Celsius and the gas flow rate is scaled in Million Cubic Meter standard per a day. It has been tested for each hidden layer by 4,6,8,12 neurons in the supposed model and also the transfer function of *tanh* and *sigm* has been used for these hidden layers. It is noteworthy that the transfer function of model's output layer is linear and the LM algorithm is also used for training of the model. The number of 200 operational field data which divided by 60% for training , 15% for network's validation and 25% for testing of the network have been imported into the model in batch. All data are really taken from the practical operating conditions of MAN turbo compressor located in neka gas booster station in Iran. The statistical information of data can be seen in the Table 1.

Moreover the information of Pearson correlation of data are gathered in Table 2.

Hence the Table 2 implies, the correlation values (green highlighted numbers) between the model's output parameters and input parameters are rather high. Thus all those mentioned input parameters have been initially applied into the model and also the parameter of units running (no) which is highly correlated with compressor flow rate (q), would be also considered as a controlling input variable of the neural network.

First the data are randomly sorted, then applied into the proposed neural network for each structure just selected by different layers, neurons and the transfer functions. Finally the software outputs are collected in the Table 3.

As it is obvious in the Table 3, the network structured in two hidden layers with which 12 neurons in each layer using the *sigm* transfer function, is the best constructed model due to the R^2 coefficient , the mean of standard deviation (MSE) and the normal mean of standard deviation (NMSE). It should be said that training of the model in neuro solution 5. lasts for almost 15 minutes (Fig. 1).

Moreover, the model's output for the parameters of P_{out} , T_{dis} and q are compared with the 50 primary real data which both graphs for each parameter are shown in Figure 2. The cutoff lines are designated for neural network that they are acceptably matches on the real data.

In this part the curve fitting module of SPSS spreadsheet is used to get a mathematical equation fitted on the scattered data points to have a comparison feature with neural network outputs. Notifying the correlation data of input parameters in the Table 2, it must be paid attention that there are no general prerequisites for inferring an authentic multivariable regression model ,consequently a new variable R

Table 1. Statistical information of data

	P_{in}	T_{in}	T_{amb}	Pt.rpm	P_{out}	T_{dis}	*q*
Mean	42.76	16.54	11.03	6420.25	52.64	40.44	16.24
Standard Error	0.50	0.29	0.39	34.42	0.66	0.44	0.18
Median	44.20	15.00	11.00	6300.00	55.15	41.00	17.15
Mode	34.00	15.00	12.00	6100.00	51.30	40.00	18.00
Standard Deviation	7.01	4.07	5.56	486.74	9.35	6.22	2.54
Sample Variance	49.19	16.57	30.92	236912.00	87.45	38.73	6.46
Kurtosis	-1.42	-1.25	0.44	-1.10	-0.99	-0.78	-0.42
Skewness	-0.25	0.56	0.74	0.36	-0.39	-0.50	-0.61
Range	22.80	12.00	26.00	1800.00	34.40	23.00	13.20
Minimum	31.60	11.00	0.00	5600.00	33.70	27.00	7.80
Maximum	54.40	23.00	26.00	7400.00	68.10	50.00	21.00
Count	200.00	200.00	200.00	200.00	200.00	200.00	200.00
Largest	54.40	23.00	26.00	7400.00	68.10	50.00	21.00
Smallest	31.60	11.00	0.00	5600.00	33.70	27.00	7.80

Table 2. Pearson correlation

	P_{in}	T_{in}	T_{amb}	Pt.rpm	P_{out}	T_{dis}	*q*
p_{in}	1.00						
t_{in}	0.76	1.00					
t_{amb}	0.65	0.62	1.00				
no	-0.46	-0.10	-0.06	1.00			
pt.rpm	-0.19	-0.09	-0.27	-0.38	1.00		
p_{out}	0.95	0.77	0.65	-0.28	-0.15	1.00	
t_{out}	0.59	0.79	0.48	0.03	0.20	0.76	1.00
q	-0.49	-0.12	-0.11	0.79	0.05	-0.37	0.01

Table 3. The software outputs

NMse	Mse	R Square	Transfer function for the second hidden layer	Neurons of the second hidden layer	Transfer function for the first hidden layer	Neurons of the first hidden layer
0.0227	0.0054	0.9754	Tanh	4	Tanh	4
0.0220	0.0052	0.9775	Sigm			
0.0138	0.0033	0.9847	Tanh	6		
0.0153	0.0037	0.9827	Sigm			
0.0088	0.0021	0.9910	Tanh	8		
0.0113	0.0027	0.9892	Sigm			
0.0050	0.0012	0.9944	Tanh	12		
0.0076	0.0018	0.9934	Sigm			
0.0263	0.0063	0.9712	Tanh	4	Sigm	
0.0316	0.0075	0.9655	Sigm			
0.0140	0.0033	0.9855	Tanh	6		
0.0209	0.0049	0.9771	Sigm			
0.0140	0.0033	0.9857	Tanh	8		
0.0112	0.0027	0.9866	Sigm			
0.0054	0.0013	0.9948	Tanh	12		
0.0055	0.0013	0.9948	Sigm			
0.0145	0.0034	0.9851	Tanh	4	Tanh	6
0.0189	0.0045	0.9799	Sigm			
0.0088	0.0021	0.9912	Tanh	6		
0.0065	0.0015	0.9930	Sigm			
0.0049	0.0011	0.9958	Tanh	8		
0.0066	0.0016	0.9936	Sigm			
0.0037	0.0009	0.9962	Tanh	12		
0.0036	0.00009	0.9970	Sigm			
0.0176	0.0042	0.9805	Tanh	4	Sigm	
0.0154	0.0037	0.9821	Sigm			
0.0103	0.0025	0.9888	Tanh	6		
0.0101	0.0024	0.9884	Sigm			
0.0064	0.0015	0.9763	Tanh	8		
0.0055	0.0013	0.9952	Sigm			
0.0038	0.0009	0.9956	Tanh	12		
0.0035	0.0008	0.9968	Sigm			
0.0085	0.0020	0.9916	Tanh	4	Tanh	8
0.0101	0.0024	0.9922	Sigm			
0.0047	0.0011	0.9952	Tanh	6		
0.0047	0.0011	0.9944	Sigm			
0.0036	0.0009	0.9978	Tanh	8		
0.0033	0.0008	0.9974	Sigm			
0.0013	0.0003	0.9986	Tanh	12		
0.0026	0.0006	0.9994	Sigm			
0.0074	0.0018	0.9930	Tanh	4	Sigm	
0.0083	0.0019	0.9914	Sigm			
0.0065	0.0016	0.9932	Tanh	6		
0.0144	0.0034	0.9859	Sigm			
0.0030	0.0007	0.9976	Tanh	8		
0.0043	0.0010	0.9970	Sigm			
0.0016	0.0038	0.9998	Tanh	12		
0.0129	0.0031	0.9932	Sigm			
0.0032	0.0008	0.9964	Tanh	4	Tanh	12
0.0050	0.0012	0.9952	Sigm			
0.0023	0.0005	0.9988	Tanh	6		
0.0023	0.0006	0.9992	Sigm			
0.0011	0.0003	0.9986	Tanh	8		
0.0012	0.0003	0.9988	Sigm			
0.0002	0.00005	0.9992	Tanh	12		
0.0003	0.00008	0.9988	Sigm			
0.0049	0.0018	0.9954	Tanh	4	Sigm	
0.0047	0.0012	0.9956	Sigm			
0.0036	0.0009	0.9978	Tanh	6		
0.0025	0.0006	0.9998	Sigm			
0.0026	0.0006	0.9964	Tanh	8		
0.0013	0.0003	0.9990	Sigm			
0.0008	0.0002	0.9992	Tanh	12		
0.0003	0.00006	0.9998	Sigm			

defined (equation 4) just considering the thermodynamical relationships between the input parameters (P_{in}, T_{in}, T_{amb}, no, Pt.rpm) and the values for each output parameters (P_{out}, T_{out} , q) are fitted on this

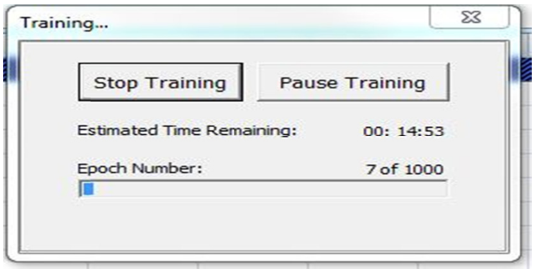

Figure 1. Training panel

new introduced variable.

$$p_{in} \propto t_{in} \quad (1)$$

$$p_{in} \propto t_{amb} \quad (2)$$

$$p_{in} \propto \frac{1}{pt.rpm \times no} \quad (3)$$

Then:

$$\left(R = \left[(p_{in} \times pt.rpm \times no) \Big/ (t_{in} \times t_{amb}) \right] \right) \propto 1 \quad (4)$$

Due to the scattered data points (Figs 3-5), three curves of logarithmic, inverse and quadratic equations have been nominated for data fitting which the typical model summaries and parameter estimation for output pressure fitness, is shown in the Table 4.

If the R square scale is just considered to choose the best model in the Table 4, the logarithmic curve will be finally selected for fitting on data. The logarithmic estimated equations for each of the output pressure, output temperature and flow rate are shown in equations 5, 6, 7.

$$p_{out}(R) = -8.623 \times Log(R) + 121.4 \quad (5)$$

$$t_{out}(R) = -4.145 \times Log(R) + 73.47 \quad (6)$$

$$q_{out}(R) = 1.555 \times Log(R) + 3.84 \quad (7)$$

To have a more closely and exactly examination of applying the multivariate regression for getting another comparison with the ANN proposed model, it should be first checked the collinearity of the input variables. As it mentioned before Table 2, the positive correlation values between the compressor's input pressure, temperature and also the input pressure and the ambient temperature are relatively high which they would be thermodynamically declared as well. It is because of that the temperature and pressure parameters are inherently interconnected and they have a strong direct relation. Moreover, there has been a negative correlation between the compressor's input pressure and its speed and the number of units running which this scenario is practically occurring, since the compressor's input pressure is reduced due to increasing speed and putting more compressors into operation. Therefore the input parameters for applying into the multivariate regression model are not algebraically orthogonal. Additionally, the affirmation of collinearity existence of model's input variables could be proved by special statistical analysis just like the correlation matrix's eigenvalues, variance inflation factor, condition index and checking the F&T statistics. Typically the multivariate regression in SPSS, for estimating of compressor's output pressure applying all the input parameters is shown in the Table 5.

As the Table 6 implies, the variation inflation factor

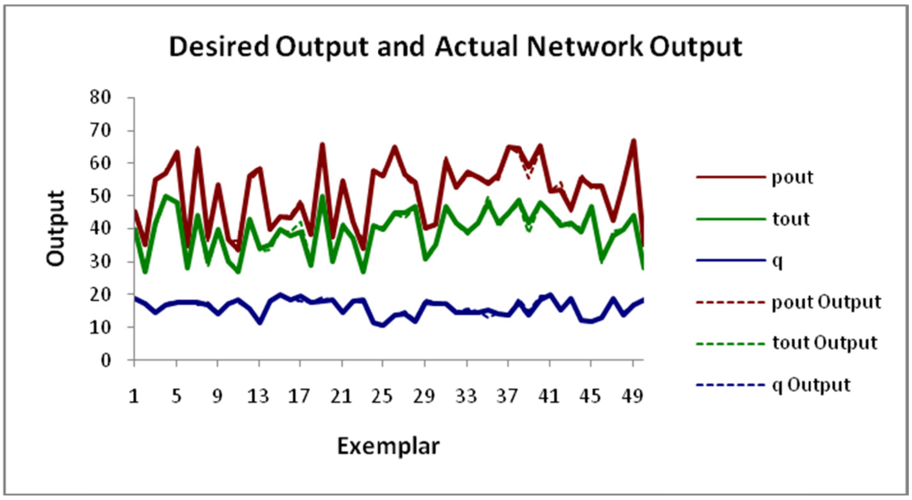

Figure 2. The model's output for the parameters

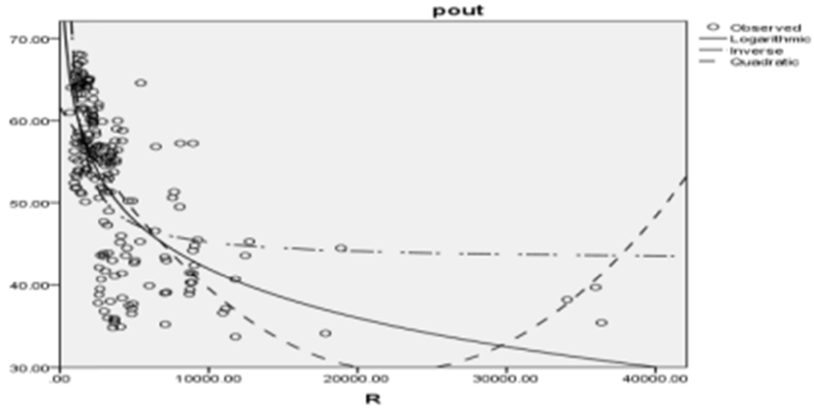

Figure 3. Scattered data points

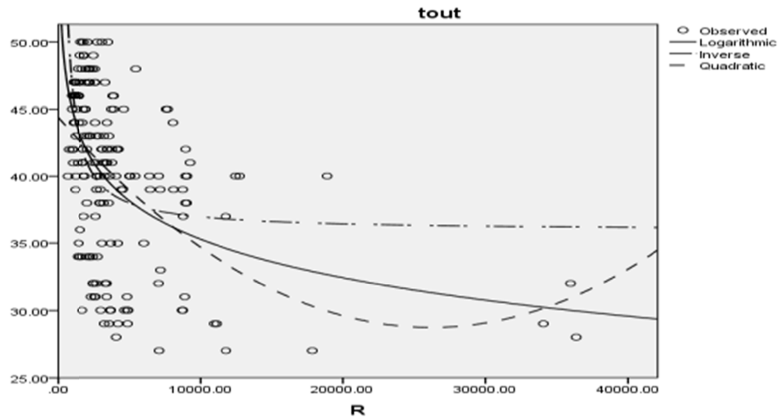

Figure 4. Scattered data points

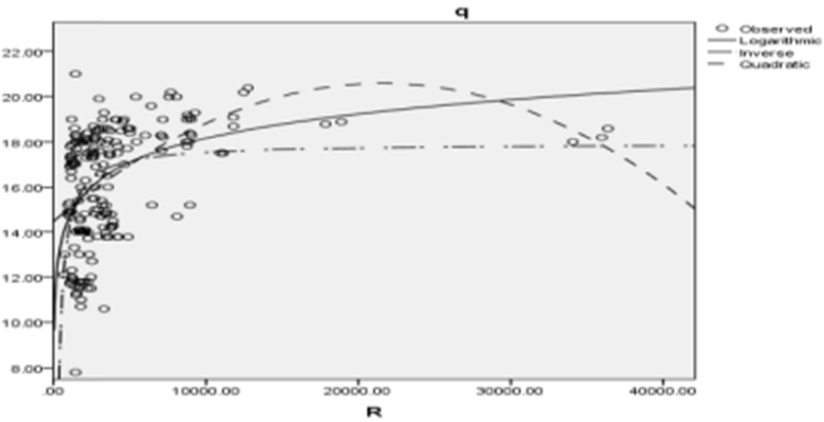

Figure 5. Scattered data points

(VIF) for all the input parameters except the ambient temperature are greater than 10. Collinearity is clear according to walker' opinion for VIF>2 [17] . In some references the collinearity would be clarified for VIF>10 [18] . In addition, the Table 7 could be referred in order to further investigation of collinearity existence between the input variables. According to green and walker's opinion if there is at least an eigenvalue close to zero, the collinearity among the model's input parameters would exist [19, 20] . Besides, due to Vinod and Ullah's proposed formula (Formula 8) the condition number values (by sequentially entrance of input

Table 4. Model summaries and parameter estimation for output pressure

Equation	R Square	F	Sig.	Constant	b1	b2
Logarithmic	0.469	174.974	0.000	121.385	-8.623	
Inverse	0.410	137.349	0.000	43.012	22215.498	
Quadratic	0.426	73.136	0.000	61.632	-0.003	6.251E-008

Table 5. The multi regression model for estimating of compressor's output pressure applying all the input parameters

Model	R	R Square	Adjusted R Square	Std.Error of the Estimate	F change	Sig.F Change	Durbin Watson
	0.999	0.998	0.998	2.6296	16494.944	.000	1.878

Table 6. Collinearity characteristics of input parameters

Model	Standardized Coefficeints	t	Sig.	Collinearity Statistics	
	Beta			Tolerance	VIF
Pin	0.954	28.001	0.000	0.010	95.932
Tin	0.053	2.249	0.026	0.021	46.581
Tamb	0.005	0.450	0.653	0.102	9.851
No	0.063	5.003	0.000	0.076	13.152
Ptrpm	-0.073	-2.876	0.004	0.019	53.217

Table 7. More Collinearity characteristics of input parameters

Dimension	Eigenvalue	Condition Index
1	4.747	1.000
2	0.162	5.407
3	0.063	8.700
4	0.021	14.876
5	0.007	26.666

parameters) are ranged in 5 to 30 which demonstrates the collinearity of variables [21] . It should be noticed that all the model's input parameters are orthogonal when the condition number would be 1.

Thus, it has been proved that there are rather strong collinearity between the model's input parameters so the ordinary linear regression (least square method) cannot be used to estimate the output variables including the output pressure, output temperature and flow rate.

Horrel and Kennard in 1970, offered the ridge regression method as a different procedure for modeling of non orthogonal variables [22]. Linearly independencies of the model's input parameter in the ordinary least square method (OLS) would cause the matrix X^TX (multiplication of correlation matrix by itself) be fully ranked, so it is invertible and estimating of regression's coefficients could be found by equation [10].

$$Y = \beta X + \varepsilon \tag{9}$$

$$\beta = (X^T X)^{-1} X^T Y \tag{10}$$

Variables being extremely collineared would result in non invertibility of the correlation matrix, therefore

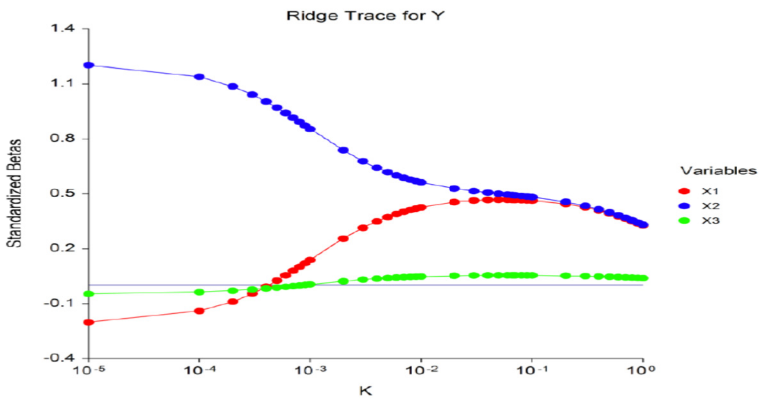

Figure 6. Typical betas for biased parameter k

the ridge regression adds a multiple (biased estimator) of identity matrix to the coefficients equation (equation 11) in order to make it mathematically invertible.

$$\beta = (X^T X + kI)^{-1} X^T Y \quad , \quad 0<k<1 \tag{11}$$

One of the existing challenges in this method is how to choose the biased parameter of k. Horrel and Kennard provided the curve of ridge trace to select it. From their point of view, estimating of the different (but nearly close to each other) regression's coefficients due to choosing of several values of biased estimator k (Fig. 6), would be ended up when the coefficients become stable.

In this way, the ridge regression coefficients for each of the output temperature, output pressure and the amount of flow rate passing through compressors, using MATALB software are gathered in the Tables 8-10.

The performance of each estimation approaches have been calculated in the Table 11.

As it could be viewed from the Table 11, the proposed ANN is noticeably able to predict the model's outputs and consequently it can be used as data explorer to produce the output variables.

Case study and sensitivity

As a case study, the supposed ANN model has been run for 20 operational points and its output values are used to calculate the compressor poly-tropic head, poly-tropic efficiency and the fuel volumetric consumption for compressor driver (gas turbine).These real operational raw data come from neka's gas booster station which has 3 turbo compressors. As mentioned earlier, there is usually one unit of all (compressors) remained off at each station. So the parameter of number of units running in our case study and numerical

Table 8. The ridge regression coefficients for each of the output temperature

k	b_0	b_1	b_2	b_3	b_4	b_5
1	-47.4159	1.6028	-0.1863	-0.0214	6.7990	0.0036
0.1	-50.9783	1.6554	-0.2381	-0.0335	7.2441	0.0038
0.01	-51.3530	1.6610	-0.2436	-0.0348	7.2909	0.0039
0.001	-51.3906	1.6615	-0.2441	-0.0349	7.2956	0.0039
0.0001	-51.3944	1.6616	-0.2442	-0.0349	7.2961	0.0039
0.00001	**-51.3948**	**1.6616**	**-0.2442**	**-0.0349**	**7.2962**	**0.0039**
0.000001	-51.3948	1.6616	-0.2442	-0.0349	7.2962	0.0039

$$Ridge\ (y,X,\ k,0)=y_{tout}=b_0+b_1P_{in}+b_2T_{in}+b_3T_{amb}+b_4NO+b_5PT.rpm \tag{12}$$

Table 9. The ridge regression coefficients for each of the output pressure

k	b_0	b_1	b_2	b_3	b_4	b_5
1	-6.8930	-0.0108	-0.0124	0.0452	5.2783	0.0022
0.1	-7.4495	-0.0055	-0.0178	0.0450	5.3531	0.0023
0.01	-7.5074	-0.0050	-0.0183	0.0450	5.3608	0.0023
0.001	-7.5132	-0.0049	-0.0184	0.0450	5.3616	0.0023
0.0001	-7.5138	-0.0049	-0.0184	0.0450	5.3617	0.0023
0.00001	**-7.5139**	**-0.0049**	**-0.0184**	**0.0450**	**5.3617**	**0.0023**
0.000001	-7.5139	-0.0049	-0.0184	0.0450	5.3617	0.0023

$$Ridge\ (y,X,\ k,0)=y_{pout}=b_0+b_1P_{in}+b_2T_{in}+b_3T_{amb}+b_4NO+b_5PT.rpm \tag{13}$$

Table 10. the ridge regression coefficients for each of the flow rate

k	b_0	b_1	b_2	b_3	b_4	b_5
1	-56.5727	0.5852	0.6194	-0.0014	8.0953	0.0075
0.1	-58.4367	0.6062	0.6036	-0.0070	8.3198	0.0076
0.01	-58.6309	0.6085	0.6019	-0.0076	8.3433	0.0076
0.001	-58.6504	0.6087	0.6017	-0.0077	8.3456	0.0076
0.0001	-58.6524	0.6087	0.6017	-0.0077	8.3459	0.0076
0.00001	**-58.6526**	**0.6087**	**0.6017**	**-0.0077**	**8.3459**	**0.0076**
0.000001	-58.6526	0.6087	0.6017	-0.0077	8.3459	0.0076

$$Ridge\ (y,X,\ k,0)=y_{q}=b_0+b_1P_{in}+b_2T_{in}+b_3T_{amb}+b_4NO+b_5PT.rpm \tag{14}$$

Table 11. Performance of each estimator

	Mse			Average Absolute error		
Estimator model	**Pout**	**Tout**	**Q**	**Pout**	**Tout**	**Q**
Logarithmic curve fitting	50.32	26.91	4.95	6.11	4.24	1.88
ANN (perceptron)	1.22	1.26	0.46	0.83	0.89	0.49
Ridge regression	4.43	5.54	1.68	1.57	1.95	1.01

solution will be 1 or 2. The proposed network outputs that determined by neuro solution 5. and also the 20 operational points can be seen in the Table 12.

For example, if an operator wants to run one compressor with which 6600RPM for the operational conditions including the input pressure 46 bar, input temperature 15°c and the ambient temperature 13°c, then the output pressure and temperature and the flow rate passing through the compressor by the ANN would be respectively modeled as 55.99 bar, 39.18°c and 11.36 MMCD and whether two compressors are run in 5800 RPM with the same those mentioned input conditions then the model output would be respectively 56.98 bar, 39.19°c and 15.46 MMCD.

The compressor head, poly-tropic efficiency and its driver's fuel consumption for each of the operational data given in the Table 12 can be thermodynamically calculated by the equations [15,16,17].

$$\frac{H_p}{S^2} = A + B\left(\frac{Q}{S}\right) + C\left(\frac{Q}{S}\right)^2 \tag{15}$$

$$\eta_p = D + E\left(\frac{Q}{S}\right) + F\left(\frac{Q}{S}\right)^2 = \frac{T_{suc} \times \left(\frac{P_{dis}}{P_{suc}}\right)^{k-1/k} - T_{suc}}{T_{dis} - T_{suc}} \tag{16}$$

$$FC = \frac{\alpha Q P_{suc} \times 10^5 \left[\left(\left(P_{dis}/P_{suc}\right)^{k-1/k} - 1\right) k/k-1\right]}{\eta_M \times \eta_T \times HV \times \eta_P} \tag{17}$$

Table 12. The proposed network outputs

Working point	p_n	t_{in}	t_{mb}	no	Pt.rpm	p_{out}	t_{out}	q
1	46	15	13	1	6600	55.99116	39.1847	11.35873
2	46	15	13	2	5800	56.97998	39.1929	15.45717
3	41.8	14	12	1	5700	45.9967	27.13791	10.93688
4	41.8	14	12	2	5500	46.88801	29.64939	17.80682
5	46.1	16	17	1	6800	57.40835	43.16096	11.60754
6	46.1	16	17	2	5700	57.87845	40.50694	14.27565
7	44.1	17	11	1	6800	56.26898	44.84222	10.75519
8	44.1	17	11	2	6300	59.61326	47.71606	14.48764
9	47.9	18	12	1	6200	56.93267	38.67555	11.56409
10	47.9	18	12	2	5800	61.96017	44.57932	13.99417
11	49.1	14	14	1	6100	54.65499	31.23981	12.67964
12	49.1	14	14	2	5600	56.64087	33.69328	17.49114
13	41.5	16	10	1	6000	50.24671	37.6776	8.0661
14	41.5	16	10	2	5700	53.53144	40.4583	13.00016
15	42.3	17	12	1	6400	53.56963	43.11186	8.595982
16	42.3	17	12	2	6000	57.4423	46.63514	12.11301
17	48.9	24	13	1	6700	56.94215	44.1238	13.12955
18	48.9	24	13	2	5800	59.59073	44.95882	16.29717
19	51.3	20	23	1	7100	57.48579	43.06732	16.73196
20	51.3	20	23	2	6000	65.4207	46.38069	18.10121

Table 13. Poly tropic efficiency regression coefficients of under studied turbo compressors

$\eta_p = 0.559 + 0.297\left(Q/S\right) - 0.122\left(Q/S\right)^2$	**Compressor 1**
$\eta_p = 0.669 + 0.135\left(Q/S\right) - 0.071\left(Q/S\right)^2$	**Compressor 2**
$\eta_p = 0.611 + 0.215\left(Q/S\right) - 0.092\left(Q/S\right)^2$	**Compressor 3**

Estimating of each under study compressor's poly-tropic efficiency regression coefficients (A_i,B_i,C_i) would be carried out via equation [16] and applying 100 operational working points (T_{suc},T_{dis},P_{suc},P_{dis},Q_{ac},S) associated to every turbo compressor by curve fitting tool in SPSS software. The poly-tropic efficiency of each compressor is listed in the Table 13.

It should be noticed that the flow rate (q) in the Table 12 is measured in MMSCD which must be converted to the actual volumetric cubic meter per an hour and then used it in Formula 17. The Formula 18 has been used to yield the conversion.

$$Q_{ac} = \frac{Q_{st} \times 1.013 \times (273.15 + T_{suc}) \times z}{0.0864 \times P_{suc} \times 288} \times 3600 \quad (18)$$

The obtained values of the compressor head, poly-tropic efficiency and their driver's fuel consumption for each of the 20 working points of the Table 12 are totally listed in the Table 14.

The following ideas are derived from the Table 14 at a glance:

1. Generally speaking, there has been relatively more head value that produced by one compressor than two compressors (working points marked by odd numbers in Table 14) which be put into operation under

Table 14. Values of the compressors head, poly tropic efficiencies and their driver's fuel consumptions

Working point	Flowrate	Comp.1			Comp.2			Comp.3		
		H_p	η_p	FC	H_p	η_p	FC	H_p	η_p	FC
1	9952.572	31774.52	0.73	1144.348	30976.57	0.71	1173.827	31624.92	0.73	1149.762
2	6771.82	24875.28	0.74	836.727	24551.61	0.73	847.758	24779.61	0.74	839.958
3	10509.22	22478.84	0.69	563.415	21981.15	0.68	576.171	22569.62	0.69	561.148
4	8555.27	21955.28	0.73	524.829	21392.86	0.71	538.628	21865.64	0.72	526.981
5	10183.74	33762.8	0.73	1309.053	32919.93	0.71	1342.57	33600.08	0.73	1315.393
6	6262.29	23978.95	0.62	826.045	23770.29	0.62	833.298	23917.93	0.64	828.154
7	9897.98	33885.78	0.73	1347.322	33065.03	0.72	1380.765	33710.25	0.73	1354.338
8	6666.47	29238.44	0.74	1116.936	29067.07	0.73	1123.521	29191.68	0.74	1118.725
9	9831.90	27799.14	0.72	1043.171	27082.36	0.70	1070.78	27699.47	0.72	1046.925
10	5948.99	24734.85	0.74	925.924	24650.5	0.73	929.092	24716.51	0.73	926.611
11	10372.39	26466.5	0.71	710.740	25796.51	0.69	729.199	26440.74	0.71	711.433
12	7154.18	23184.83	0.74	629.369	22754.46	0.73	641.273	23065.82	0.74	632.617
13	7861.10	26593.26	0.74	783.821	26063.91	0.72	799.741	26451.5	0.73	788.022
14	6334.88	23990.28	0.74	841.29	23761.22	0.73	849.400	23922.76	0.74	843.664
15	8247.50	30274.89	0.74	1034.71	29698.57	0.73	1054.789	30117.16	0.74	1040.129
16	5810.97	26359.53	0.73	953.906	26393.4	0.73	952.682	26385.55	0.73	952.966
17	11159.93	32106.28	0.72	1077.05	31282.94	0.70	1105.397	32045.62	0.71	1079.089
18	6926.18	24883.22	0.74	839.866	24522.35	0.73	852.225	24777.57	0.74	843.447
19	13,374.08	34559.45	0.69	1056.816	33843.86	0.67	1079.162	34760.37	0.69	1050.708
20	7,234.27	26630.64	0.7	1132.16	26228	0.73	1149.544	26513.42	0.74	1137.17

Table 15. PEM values for each point

	PEM		
Working point	**Comp.1**	**Comp.2**	**Comp.3**
1	27.77	26.39	27.51
2	29.73	28.96	29.50
3	39.89	38.15	40.22
4	41.83	39.72	41.49
5	25.79	24.52	25.54
6	29.03	28.53	28.88
7	25.15	23.95	24.89
8	26.18	25.87	26.09
9	26.65	25.29	26.46
10	26.71	26.53	26.67
11	37.24	35.38	37.17
12	36.84	35.48	36.46
13	33.93	32.59	33.57
14	28.52	27.97	28.36
15	29.26	28.16	28.96
16	27.63	27.70	27.69
17	29.81	28.30	29.70
18	29.63	28.77	29.38
19	32.70	31.36	33.08
20	23.52	22.82	23.32

the same working conditions. It should be paid attention that more head generated by centrifugal compressors is very operationally important to be dynamically more stable.

2. It would be more flow rate passing through by running two compressors (working points marked by even numbers in Table 14) under the same operating conditions, although it will results in increasing the fuel consumed by gas turbines.

Consequently, concerning the two mentioned notifications above, it could be faced two scenarios just regarding the importance account of the head and supplying the flow rate. One scenario is about to select a compressor running at high speed for gaining higher head along with dynamically steady state performance ,and the other one is about to select two compressors to provide more flow rate. Expectedly, these two scenarios followed by their own practical benefit and loss. Generally, one compressor running at high speed, concludes the more flame temperature in turbine's combustor chamber which this can along with so much load on the compressor lead to parts deterioration and results in earlier system overhaul time. In the second scenario, each compressor running at lower speed could totally.Yielding more flow rate and additionally to get better poly-tropic efficiencies for compressors but the turbines fuel consumption would attentively increase.

In order to normally and easily comparing the compressor's behavior for every operating points, it is suggested to define a measuring scale (equation [19]) named as "positive energy meter" (just shown as PEM hence forth) which its values will be calculated for each working points. Consequently, based on the previous operational working conditions of each compressor, the machine with the more PEM would be selected to be turned on to service. Indeed, this index calculates the internal energy added to the unit mass of gas flow (head) per a cubic meter fuel consumption in an hour.

$$positive\ energy\ meter = \frac{Head}{Fuel\ Consumed}\ ,\ \ (joule\ hour\ per\ cubic\ meter\ kilogram) \tag{19}$$

PEM values for each point of the Table 14 are gathered in the Table 15.

Now, the compressor operating points could be mapped for each machine. Actually, this graph which is very important to control the steady state of the dynamic compressor would define the head produced by compressor for the amount of different passing flow rates from it. As an example, if the operating points of the Table 14 that labeled by odd numbers are considered for compressor no.1, then its head-flow graph would be plotted just shown as in Figure 7.

If this graph is supposed to be plotted for the operating points that labeled by even numbers in the Table 14 then the head-flow graph for compressors no.1and 2 would be plotted as indicated in Figure 8.

Thus, if those determined compressor's head-flow points in Figures 7, 8 are being compared with the compressor performance map (Fig. 9) then it could be perceived that the compressor working points are

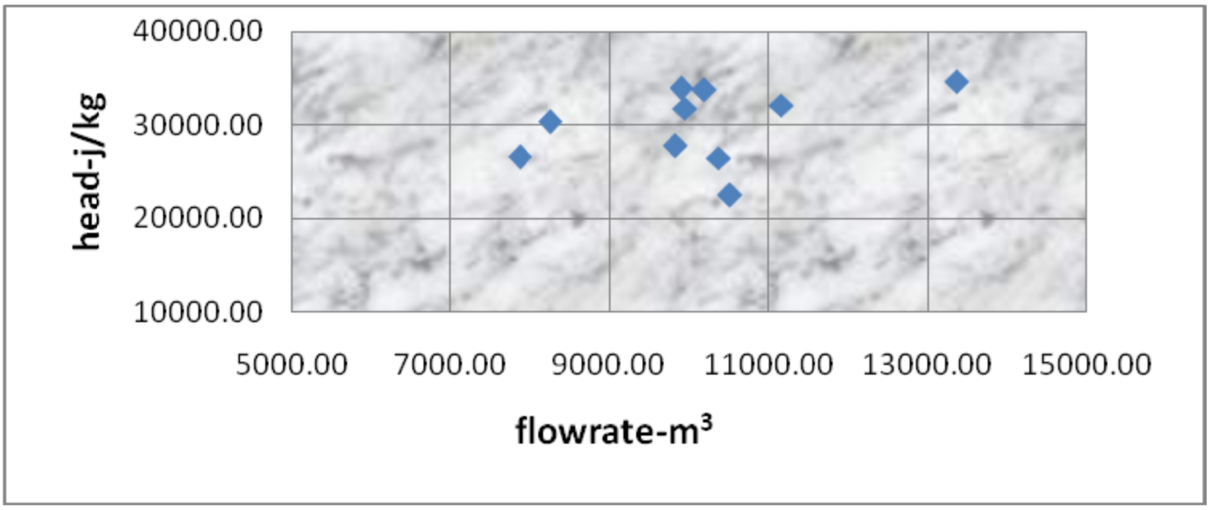

Figure 7. Head-flow graph

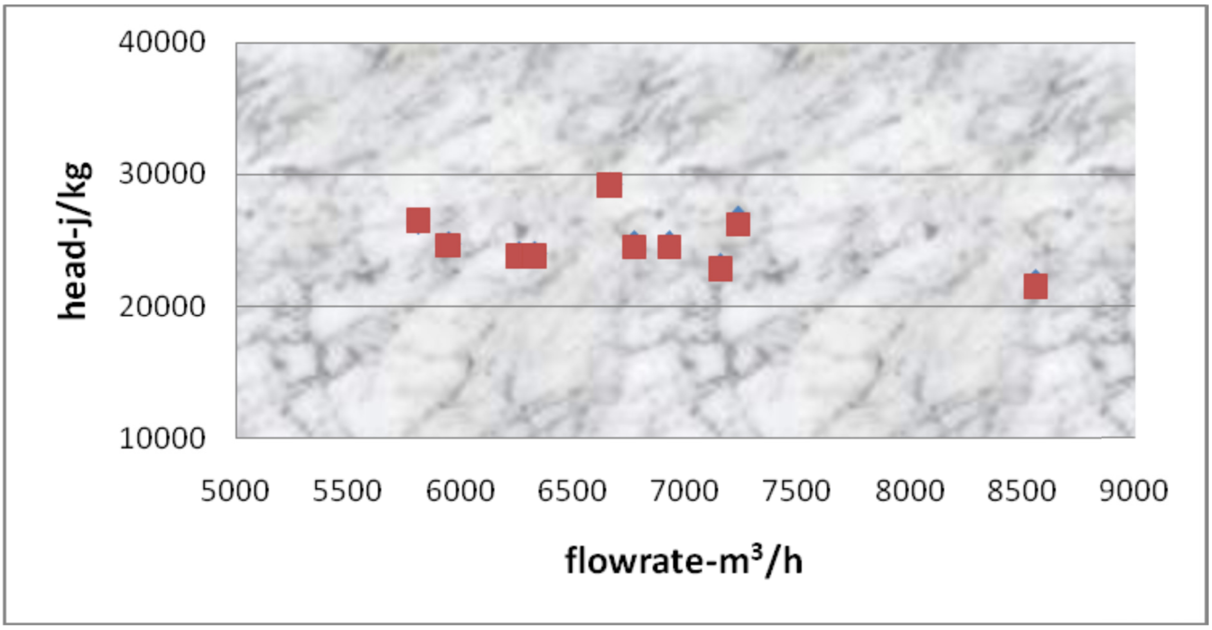

Figure 8. Head-flow graph

approximately in the limits of the stone wall for one machine running and in the limits of off-design for two machines running (Figure 9 -the rectangular boundary). The stone wall limit is generally assigned to conditions which passing rather maximum flow rate along with the lowest head applied by compressor. It is also defined that any sudden flow rate increment would results in rapid reduction in the compressor head which this phenomena could lead the gas velocity to mach number in the compressor inlet impeller [23].

Furthermore, the sensitivity analysis module of neuro solution is used to examine the effect of input parameters on the model's output clarification which shown in Figure 10.

According to the sensitivity analysis data shown in Figure 10 and Table 16, it is appeared that the input pressure and temperature parameters have the uppermost effect on the model output parameter's explanation and variation, vice versa the three other parameter namely the ambient temperature, number of compressor to be run and its/their speed have the lowest influence. Noticeably, the effect of higher input pressure

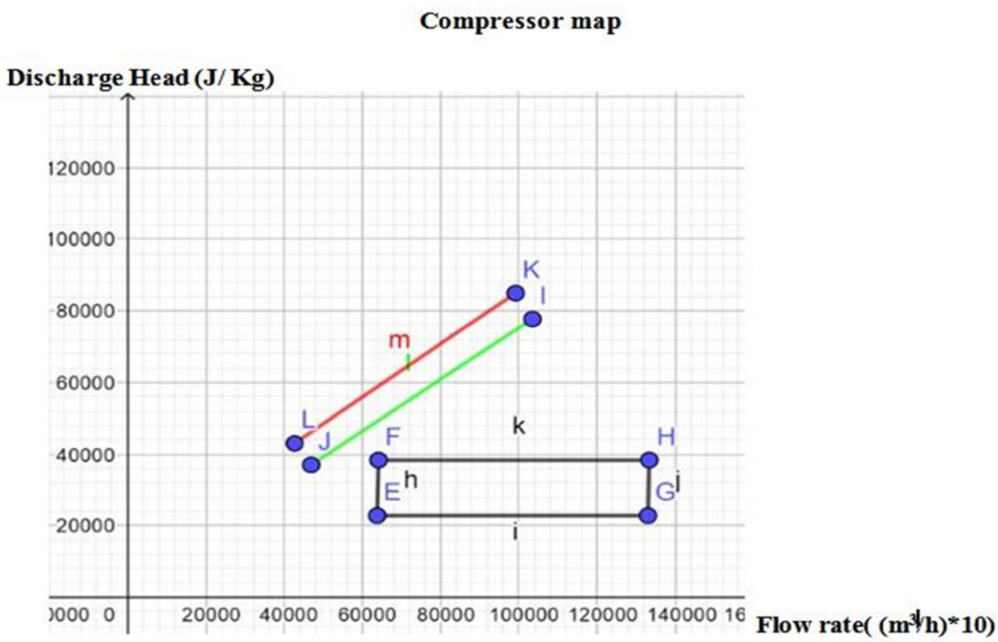

Figure 9. Compressor performance map

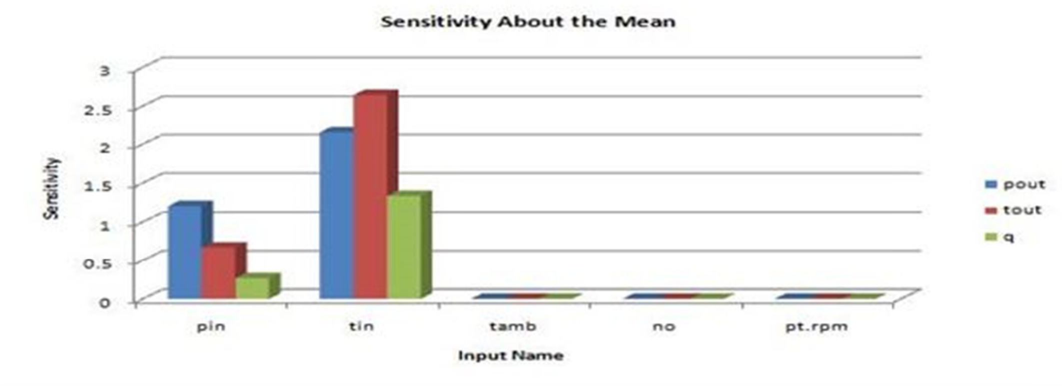

Figure 10. The sensitivity analysis graph

is a vital factor for which be emphasized by control room operators. The model's output trends due to changes of pressure and temperature parameters have been shown in Figures 11, 12.

Finally, the primary designed network is again reorganized due to the outcome of the sensitivity analysis. So, the two parameters of input pressure and input temperature just along with the number of units running as a control factor, are considered for the network's input vector. The neural network structure is as it was for the first offered network with two hidden layers with which 12 neurons in every layer and *Sigm* transfer function and also the linear transfer function for the output layer. The new network has been trained by the 200 operational data. The network's outputs are compared to the real data which could be seen in the following graph in Figure 13 along with the network performance characteristics in the Table 17.

Table 16. The sensitivity analysis of input parameters

Sensitivity	p_{out}	T_{out}	q
P_{in}	1.204255	0.667288	0.266593
T_{in}	2.152278	2.641997	1.333678
T_{amb}	1.58893E-14	0	5.29643E-15
No	1.8893E-13	0	6.29765E-14
Pt.rpm	1.78462E-16	0	5.94872E-17

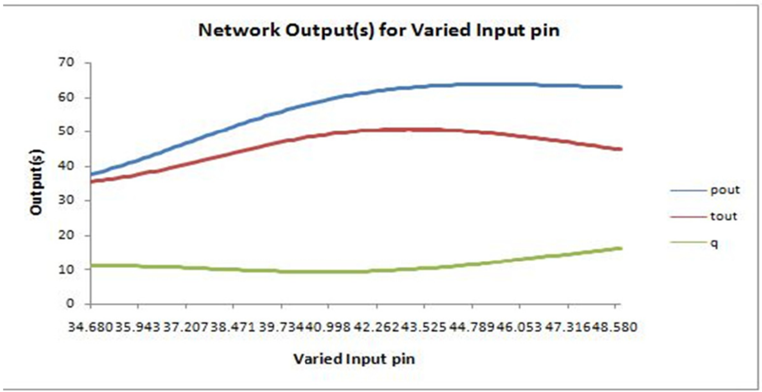

Figure 11. Trend due to changes of Pin

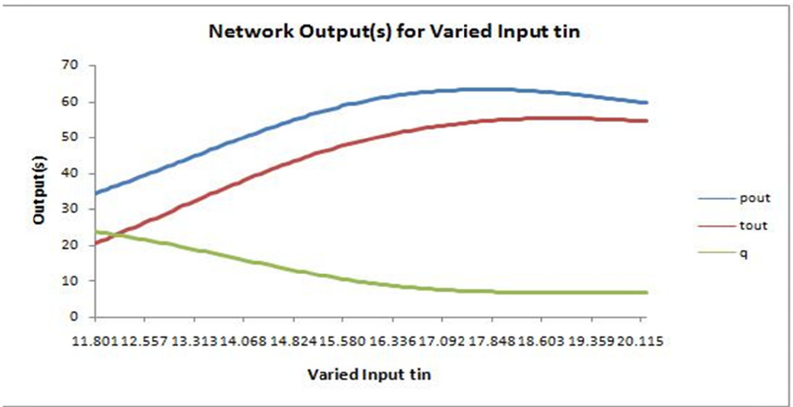

Figure 12. Trend due to changes of Tin

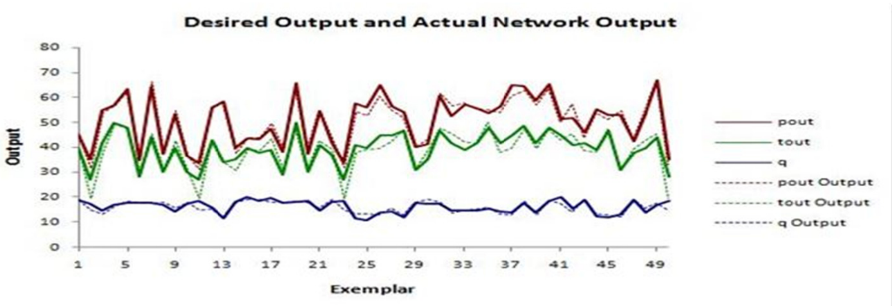

Figure 13. New network's outputs

Results

Simulation of the operational conditions of the centrifugal compressors which situated at a typical gas booster station is obviously an important issue due to the relevant high operational costs such as the fuel cost

Table 13. New network performance characteristics

Performance	pout	tout	q
MSE	4.31187	10.61469	2.07809
NSME	0.04464	0.23522	0.31877
MAE	1.67246	2.44884	1.06180
Min Abs Error	0.04152	0.00731	0.04009
Max Abs Error	5.58789	8.58521	4.01005
r	0.97854	0.91959	0.83632

consumed for driver (gas turbine) and the depreciation cost of the parts. The ANN modulation tool with a suitable structure as a reliable predictor data bank can be helpful for operators to optimally select the machines and it would be also effective to determine the compressor's dynamical performance conditions. In this research, it has been clearly justified that the real existing operational conditions is not generally good enough for the steady state dynamical operation. So, it needs to be more paid attention to specially supply the input pressure for the steady state operation of centrifugal compressors. The following research could be pursued for future study:

1. It could be appointed to calculate the input pressure of compressor by fuzzy neural network to meet the dispatcher demand based on supplying the output pressure limit and passing the needed flow rate from each compressor ,thus considering the power equation (in order to pass the load) it then would be estimated the number of machines to be run along their dedicated speeds.

2. Hence there is usually the operational input pressure lower than the needed enough to steady state working condition for centrifugal compressor which it will result to increase the gas velocity to the sonic velocity in the compressor's impeller ,therefore it is offered to determine the erosion velocity variation due to input pressure fluctuation by the ANN.

References

1. Lauden K.C. and Lauden J.P., Management information system, 13th edition, published by Pearson Education, **258** (2014).
2. Hagan M., Neural Network DeSign,2nd edition, ebook: 1-3.
3. Moraal P. and Kolmanovsky I., Turbocharger modeling for automotive control application. *SAE Trans,***108**:1324–38 (1999).
4. Ghorbanian K. and Gholamrezaei M., An artificial neural network approach to compressor performance prediction. Applied Energy. *JAE*, **86**:1210–1221(2009).
5. Bao C., Ouyang M. and Yi B., Modeling and optimization of the air system in polymer exchange membrane fuel cell system. Journal of Power Sources. *JPS*, **156** (2):232–43(2006).
6. Yu Y., Chen L., Sun F. and Wu C., Neural network based analysis and prediction of a compressor's characteristic performance map. Applied Energy. *JAE*, **84**(1):48–55(2007).
7. Torabian. and Karimian., Prediction of compressor's map using the artificial neural network(Into Persian). First conference of developing of civil engineering ,architecture and mechanic, iran. *DCEAEM01*, **1**: 111-117(2014).
8. Sanaye S., Dehghandokht M., Mohammadbeigi H. and Bahrami S., Modeling of rotary vane compressor applying artificial neural network. International Journal of Refrigration. *Int. J. Refrig*, **34**: 764-772(2011).
9. Soo-Yong C., Kook-Young A., Young-Duk L. and Young-Cheol K., Optimal Design of a Centrifugal Compressor Impeller Using Evolutionary Algorithms.Mathematical Problems in Engineering. *Math Probl Eng*, **2012:**1-22 (2012).
10. Shaojun L. and Feng L., Prediction of Cracking Gas Compressor Performance and Its Application in Process Optimization. Process systems engineering. Chinese Journal of Chemical Engineering. *CJCE*, **20**:1089-1093(2012).
11. Palmé T., Waniczek P., Hönen H., Assadi M. and Jeschke P., Compressor Map Prediction by Neural Networks. Journal of Energy and Power Engineering. *JEPE*, **6**:1651-1662 (2012).
12. Chen P., Chang H. and Armada H., A Study of Using Artificial Neural Network in a Non-linear Centrifugal Compressor System. International Journal on Computer Science and Engineering. *IJCSE,* **4**:1890-1896 (2012).
13. Yang L., Zhao L., Zhang, C. and Gua B. ,Loss-efficiency model of single and variable-speed compressors using neural networks.International Journal of Refrigretaion. *Int.J.Refrig*, **32**:1423-1432 (2009).
14. Vilalta P.C., Wan H. and Patnaik S.S., Centrifugal compressor performance prediction using gaussian process regression and artificial neural networks. Proceedings of the ASME 2019. International Mechanical Engineering Congress & Exposition. *IMECE, 2019.* **11936**:1-10 (2019).
15. Fei J., Zhao N., Shi Y., Feng Y. and Wang Z., Compressor performance prediction using a novel feed-forward neural network based on Gaussian kernel function. Advances in Mechanical Engineering. *AIME*, **8**: 1-14 (2016).
16. Li X., Yang C., Wang Y., Wang H., Zu X., Sun Y. and Hu, S., Compressor map regression modeling based on partial least squares. Royal Society Open Science. *RSOS*, **5**:1-11 (2018).
17. Walker E., Ridge Regression as an Alternative to Ordinary Least Squares: Improving Prediction Accuracy and the Interpretation of Beta Weights., Association for Institutional Research Enhancing knowledge. Expanding networks. Professional Development, Informational Resources & Networking. *AIR*, **92:**1-12 (2004).

18. Sana M. and Eyup C., Efficient Choice of Biasing Constant for Ridge Regression. International Journal of Contemporary Mathematical Sciences. *IJCMS*, **3**:527–536 (2008).

19. Green W., Econometric Analysis, 7th edition,Mac-Millan, New York University:1039 (2012).

20.Walker E., Detection of Collinearity–Influential observations. Journal of Communications in Statistics, Theory and Methods. *JCS-TM* , **18**:1675–1690 (1989).

21.Vinod H. and Ullah A., Recent Advances in Regression Models. International Journal of Forecasting. *IJF*, **2**:246-361 (1981).

22.Hoerl A. and Kennard R., Ridge Regression: Biased Estimation for non orthogonal Problems. Journal of Statistics for the Physical, Chemical, and Engineering Sciences. *Technometrics*, **42**:80-86 (2000).

23. Hanlon, P.C., Compressors Handbooks manuals, McGraw-Hill publication. **4**:22 (2001).

Nomenclature

A,B,C,D,E,F	Regression coefficients	P_{out}	Gas station outlet pressur	T_{suc}	Compressor input temperature
FC	Fuel consumption	P_{suc}	Compressor suction pressure	Z	Compressibility coefficient
H_P	Compressor poly tropic head	Pt.rpm	Lp turbine speed	α	Power corrective coefficient
HV	Heating value	q, Q	Flow rate	η_m	Lp tubine mechanical efficiency
k	Poly tropic index	Q_{ac}	Real flow rate	η_p	Compressor poly tropic efficiency
MSE	Mean of standard error	Q_{st}	Standard flow rate	η_t	Hp turbine thermal efficiency
MAE	Maximum absolute error	R^2	Regression R square	λ_{max}	Maximum eigenvalue
NMSE	Normal mean of standard error	S	Compressor speed	λ_{min}	Minimum eigenvalue
no	Number of compressor	T_{amb}	Ambient temperature		
P_{dis}	Compressor output pressure	T_{dis} T_{out}	Compressor output temperature		
P_{in}	Gas station inlet pressure	T_{in}	Gas station inlet temperature		