2017 International Nuclear Atlantic Conference - INAC 2017
Belo Horizonte, MG, Brazil, October 22-27, 2017
Associação Brasileira de Energia Nuclear – ABEN

Agência Brasileira do ISBN
ISBN 978-85-99141-07-6

MONITORING SYSTEM FOR AN EXPERIMENTAL FACILITY USING GMDH METHODOLOGY

Iraci Martinez Pereira[1], Davi Almeida Moraes[1] and Elaine Inacio Bueno[2]

[1] Instituto de Pesquisas Energéticas e Nucleares (IPEN / CNEN - SP)
Av. Professor Lineu Prestes 2242
05508-000 São Paulo, SP
martinez@ipen.br
dmoraes@dk8.com.br

[2]Federal Institute of Science, Technology and Education of São Paulo
Rua Pedro Vicente, 625
01109-010 São Paulo, SP

ebueno@ifsp.edu.br

ABSTRACT

This work presents a Monitoring System developed based on the GMDH - Group Method of Data Handling methodology to be used in an Experimental Test Facility. GMDH is a combinatorial multi-layer algorithm in which a network of layers and nodes is generated using a number of inputs from the data stream being evaluated. The GMDH network topology has been traditionally determined using a layer by layer pruning process based on a pre-selected criterion of what constitutes the best nodes at each level. The traditional GMDH method is based on an underlying assumption that the data can be modeled by using an approximation of the Volterra Series or Kolmorgorov-Gabor polynomial. The Fault Test Experimental Facility was designed to simulate a PWR nuclear power plant and is composed by elements that correspond to the pressure vessel, steam generator, pumps of the primary and secondary reactor loops. The nuclear reactor core is represented by an electrical heater with different values of power. The experimental plant will be fully instrumented with sensors and actuators, and the data acquisition system will be constructed in order to enable the details of the temporal analysis of process variables. The Fault Test Experimental Facility can be operated to generate normal and fault data. These failures can be added initially with small magnitude, and their magnitude being increasing gradually in a controlled way. The database will interface with the plant supervisory system SCADA (Supervisory Control and Data Acquisition) that provides the data through standard interface.

1. INTRODUCTION

Group Method of Data Handling was applied in a great variety of areas for data mining and knowledge discovery, forecasting and systems modeling, optimization and pattern recognition. Inductive GMDH algorithms give possibility to find automatically interrelations in data, to select optimal structure of model or network and to increase the accuracy of existing algorithms [1].

This original self-organizing approach is substantially different from deductive methods used commonly for modeling. It has inductive nature - it finds the best solution by sorting-out of possible variants. By sorting of different solutions GMDH algorithms aims to minimize the influence of the author on the results of modeling. Computer itself finds the structure of the model and the laws which act in the system.

In mathematical statistics it is need to have *a priori* information about the structure of the mathematical model. In neural networks the user estimates this structure by choosing the number of layers and the number and transfer functions of nodes of a neural network. This requires not only knowledge about the theory of neural networks, but also knowledge of the object nature and time. Besides this the knowledge from systems theory about the systems modeled is not applicable without transformation in neural network world. But the rules of translation are unknown. These problems can be overcome by GMDH that can pick out knowledge about object directly from data sampling. The Group Method of Data Handling is the inductive sorting-out method, which has advantages in the cases of rather complex objects, having no definite theory, particularly for the objects with fuzzy characteristics.

The purpose of this work is to develop a GMDH - Group Method of Data Handling model applied to an experimental bench specifically developed to generate normal and faulty data.
This work is part of a study of the sensitivity and robustness of Fault Detection methods using Artificial Intelligence methodology.

2. GROUP METHOD OF DATA HANDLING - GMDH

The Group Method of Data Handling − GMDH method is composed by an algorithm proposed by Ivakhnenko [2]. The methodology can be considered as a self-organizing algorithm of inductive propagation applied at the solution of many complex practical problems. Moreover, it is possible to get a mathematical model of the process from observation of data samples, which will be used in identification and pattern recognition or even though to describe the process itself.

The network constructed using the GMDH algorithm is an adaptive, supervised learning model. The architecture of a polynomial network is formed during the training process. The node activation function is based on elementary polynomials of arbitrary order. This kind of networks is shown in Figure 1.

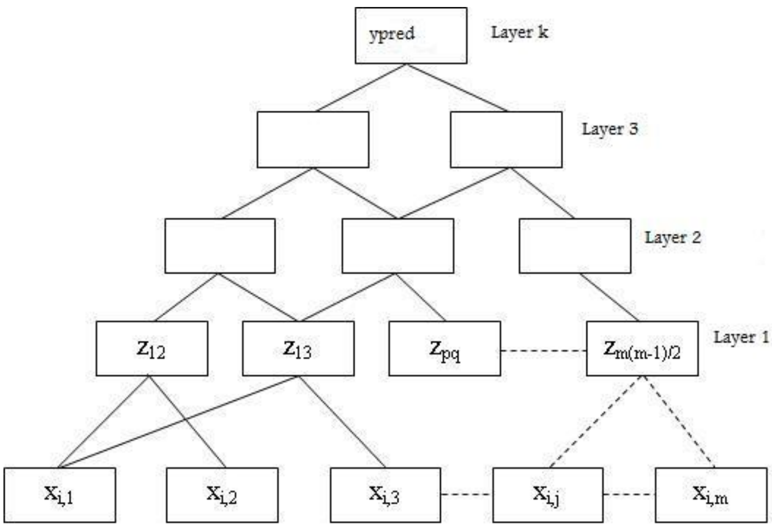

Figure 1: Self-organizing GMDH structure with m inputs and k layers.

This method solves the multidimensional problem of model improvement by the choice procedure and selection of models chosen from a set of candidate models in accordance with a supplied criterion. The majority GMDH algorithms use reference polynomial functions. A generic connection between inputs and outputs can be expressed by the series functions of Volterra which is the discrete analogous of the polynomial of Kolmogorov-Gabor [3], equation (1):

$$y = a + \sum_{i=1}^{m} b_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} c_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} d_{ijk} x_i x_j x_k + \cdots \tag{1}$$

Where:

$\{x_1, x_2, x_3 ...\}$: inputs
$\{a, b, c...\}$: polynomials coefficients
y: the node output

The components of input matrix can be changeable independent, functional forms or terms of finite differences, moreover, can be used other nonlinear reference functions. The methods still allow, simultaneously finding the model structure and the output system dependence as a function of the most important inputs system values.

The following procedure is used for a given set of ***n*** observations of the ***m*** independent variables $\{x_1, x_2, \ldots, x_m\}$ and their associated matrix of dependent values $\{y_1, y_2, \ldots y_n\}$ [2] [3] .

- Subdivide the data into two subsets: one for training and other for testing;
- Compute the regression polynomial using the equation (2), for each pair of input variables x_i and x_j and the associated output y of the training set which best fits the dependent observations y in the training set. From the observations, m(m-1)/2 regression polynomials will be computed from the observations;
- $$y = A + Bx_i + Cx_j + Dx_i^2 + Ex_j^2 + Fx_i x_j \tag{2}$$
- Evaluate the polynomial for all ***n*** observations for each regression. Store these ***n*** new observations into a new matrix Z. The other columns of Z are computed in a similar manner. The Z matrix can be interpreted as new improved variables that have better predictability than those of the original generation $x_1, x_2, \ldots, x_m$;
- Screening out the last effective variables. The algorithm computes the root mean-square value (regularity criterion – r_j) over the test data set for each column of Z matrix. The regularity criterion is given by the equation (3);

$$r_j^2 = \frac{\sum_{i=1}^{nt} (y_i - z_{ij})^2}{\sum_{i=1}^{nt} y_i^2} \tag{3}$$

- Order the columns of Z according to increasing r_j, and then pick those columns of Z satisfying $r_j < R$ (R is some prescribed value chosen by the user) to replace the original columns of X;
- The above process is repeated and new generations are obtained until the method starts overfitting the data set. One can plot the smallest of the r_j's computed in each generation and compare it with the smallest r_j's of the most recent generation start to have an increasing trend.

3. PLANT FOR DATA GENERATION AND ANALYSIS

Using a process plant to simulate a nuclear power reactor, it's possible to carry out the tests of control systems in different situations of normal and faulty operation.

Figure 1 shows the Engineering Flowchart of the PWR Power Reactor Plant. The plant will be composed of elements that will simulate the control behavior of the pressure vessel, steam generator, primary and secondary circuit pumps, an additional tank for heat exchange, sensor and actuator technologies, digital networks or connectivity using a Fieldbus system [4], signal processing devices (Programmable Logic Controllers), electro-electronic assembly of all system, as well as the mechanical structure in carbon steel, steel and aluminum.

In a second level of communication, the communication system in a reactor, allows to make effective the safety in the status of the sensors and actuators, as well as makes effective the maintenance of the system.

The database will interface with the supervisory system of the plant, which will be a SCADA (Supervisory Control and Data Acquisition) and it will make the data available through a standard interface. SCADA system is a software that interfaces with industrial processes and / or machines of the most varied sectors. [5].

The system developed for plotting the graphics will bring the following benefits to the system:

• Recording and temporal analysis of the process variables of pressure, level, temperature and flow in the pressure vessel;

• Recording and temporal analysis of the variables of the process of pressure, level, temperature and flow in the steam generator;

• Analysis and storage of more than 40 electrical variables considering the pressure vessel and steam generator;

• Analysis of useful life and predictive maintenance of sensors and actuators;

• Application of SCADA supervision system for remote monitoring, via web of the control systems, by I / O variables.

As for the monitoring and interaction to the temperature variable, a system will be implemented that allows proportional operation of the signal through a PWM (Pulse Width Modulation) system.

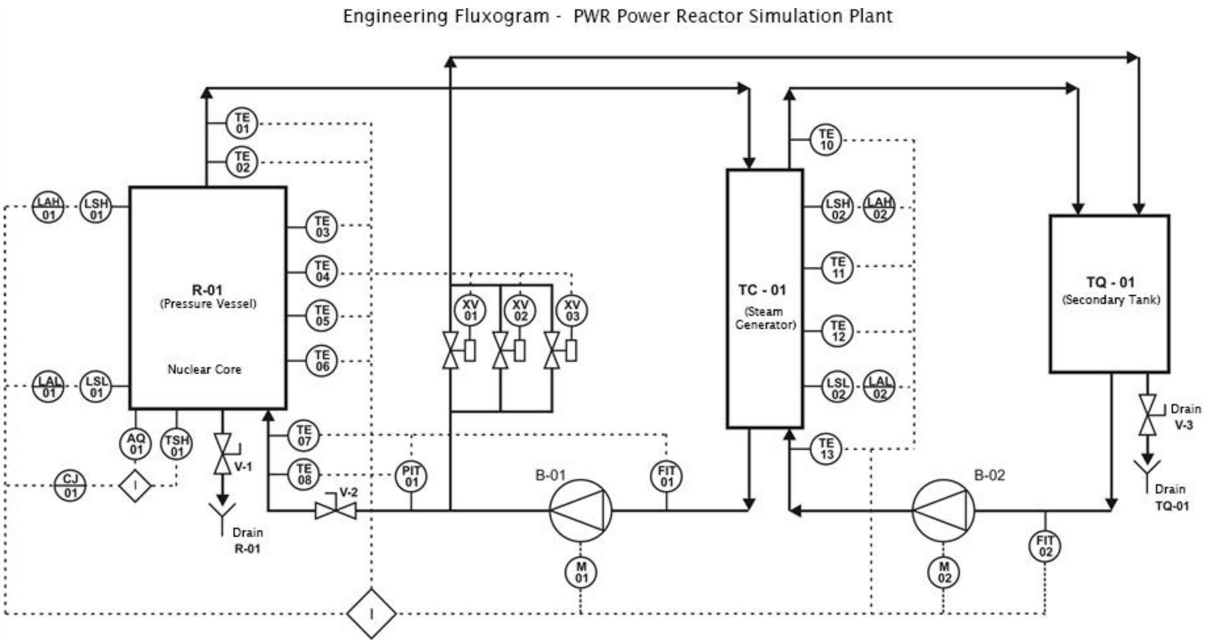

Figure 2. Engineering flowchart – PWR power reactor simulation plant

Control systems can occur through any of the variables, including pressure, level, temperature and flow (Figure 2).

- Temperature: it's possible to control the temperature through all the measuring elements installed in the Pressure Vessel and the Steam Generator.
  In the case of the Pressure Vessel, the performance for the development of the control will be through PWM power controller.

- Pressure: The measurement will occur through PIT-01 located between the pump and the pressure vessel inlet. The pressure in the system will act by varying the speed of the pump B-01.

- Level: Both LSH-01, LSL-01, LSH-02 and LSL-02 devices will be installed in both the Pressure Vessel and the Steam Generator. The actuation will be possible by pumping the pumps B-01 and B-02.

- Flow: The flow shall be measured at two points in the primary circuit and the secondary circuit. In the primary circuit, the measurement will occur through the FIT-01 meter and the secondary circuit, through the FIT-02 meter. The actuation will be possible by controlling the motor speed of the B-01 and B-02 pumps.

With the insertion of one or more faults, heating problems may occur in the Pressure Vessel. The flexibility of varying the cooling water flow of the primary circuit will act as a simulation of refrigerant fluid loss, known as LOCA (Loss of Coolant Accident), which is an important consideration in design criteria for nuclear power reactors [6] [7].

4. METHODOLOGY

The GMDH model was built using data generated by the Experimental Test Facility model [8]. This model was built using the Matlab GUIDE (Graphical User Interface Development Environment) toolbox is a set of functions designed to develop interfaces in an easy and fast way. One can add plots, sliders, frames, editable texts and push buttons that are related to other Matlab functions. Using the GUIDE Layout Editor, you can graphically design your User Interface. GUIDE then automatically generates the MATLAB code for constructing the UI, which you can modify to program the behavior of your application. Figure 3 shows the developed interface.

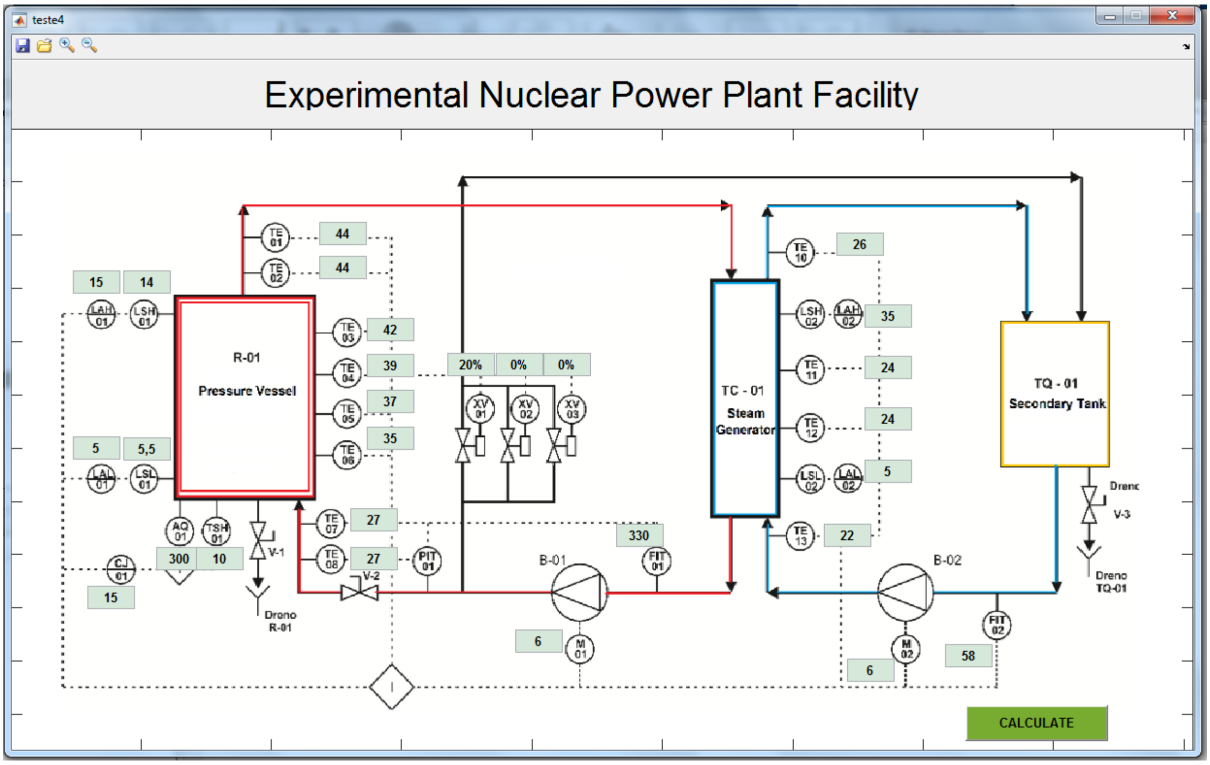

Figure 3: User Guide Interface developed to model the Test Experimental Facility.

The variables used as input for the GMDH model and the correspondence to the variable showed in figure 3 are shown at Table 1 where T are the temperature variables, pot is the heater power, mp is the primary loop flowrate and ms is the secondary loop flowrate.:

Table 1. GMDH model variables

GMDH	Matlab Model
T1	TE01
T2	TE02
T3	TE03
T4	TE04
T5	TE05
T6	TE06
T7	TE07
T8	TE08
T10	TE10
T11	TE11
T12	TE12
T13	TE13
pot	pot
mp	M01
ms	M02

Data were generated for three different power level: 100%, 75% and 50%. Figure 4 shows data generated by the Matlab Model.

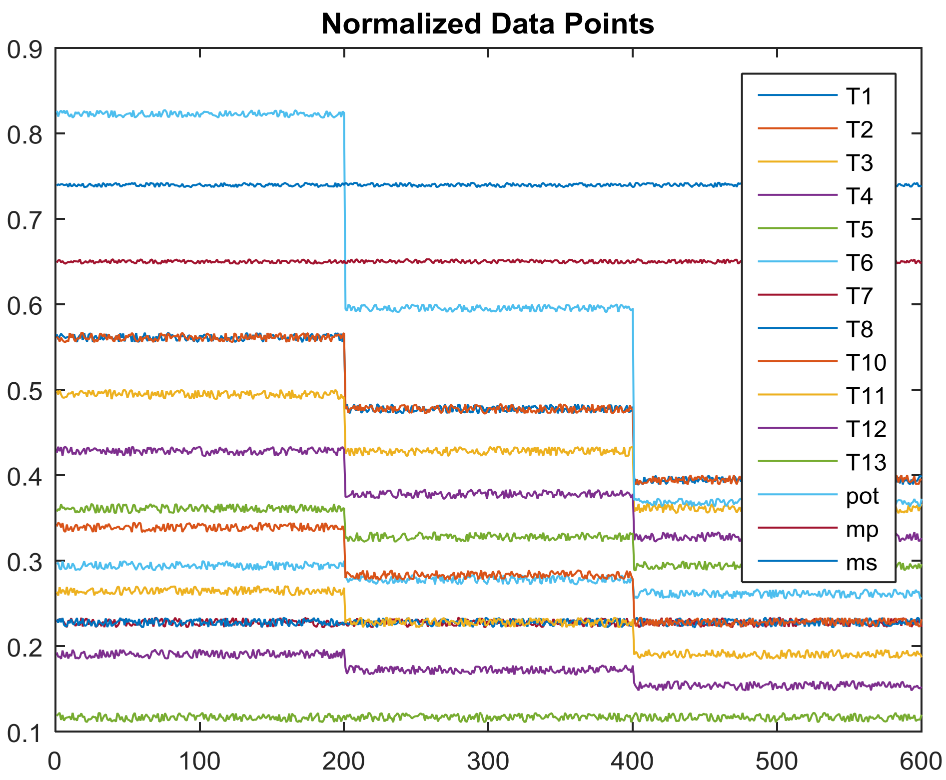

Figure 4. Normalized input data.

The GMDH Model was developed for monitoring the temperature T1. According to GMDH theory, the input variables are combined two by two and the coefficients of equation 2 are calculated. The procedure is repeated for 10 layers and the layer correspondent to the lowest error is chosen. Figure 5 shows the GMDH model development for the Experimental Test Facility variables.

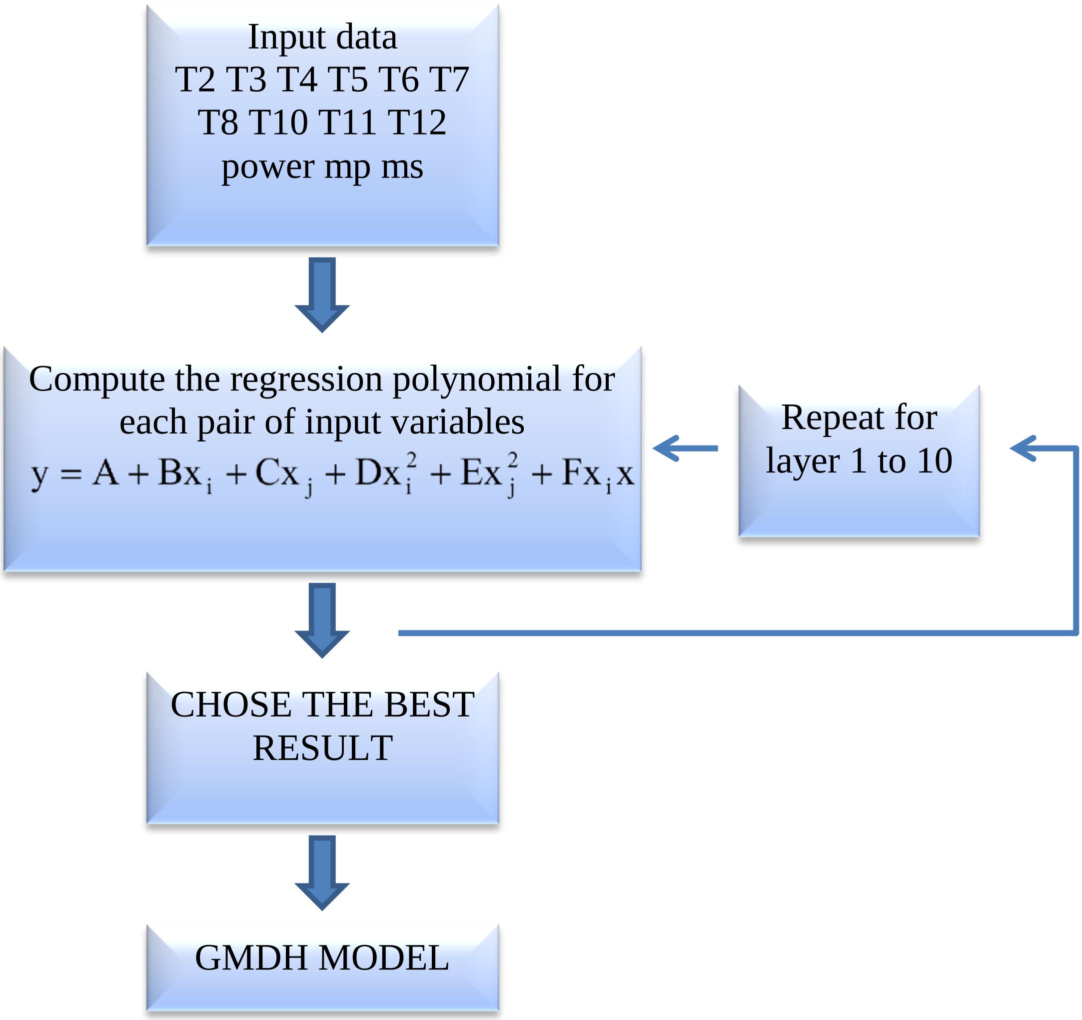

Figure 5. Determination of the GMDH Model

For this set of input data, the results of the 3rd layer showed the best result, that is, the lowest residual. This is shown in Figure 6.

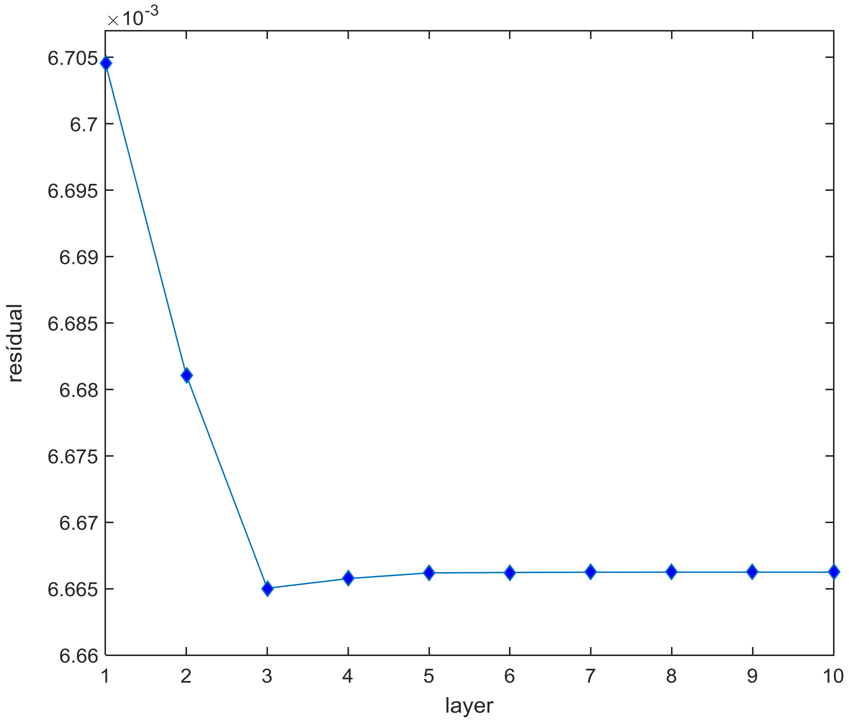

Figure 6. Residual calculated for each layer.

In order to determine the GMDH Model performance, the results of the GMDH model calculated for the variable T1 is compared with the variable actual value. The result is show at Figure 7 where we have also the percent mean error.

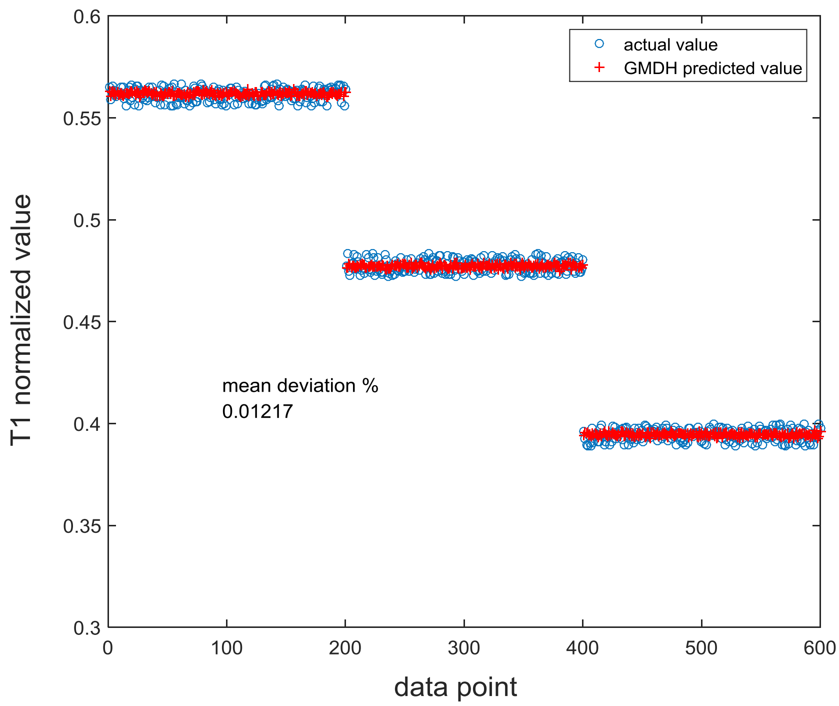

Figure 7. Comparison between actual value and GMDH predicted value

5. CONCLUSIONS AND FUTURE WORK

The GMDH model developed for monitor temperature variable T1 showed a good result, with a mean deviation error of 0.01217%. This is expected, as the actual T1 value was compared with the valued calculated for the model using the same variables used to build the model.

As future work, different GMDH models will be developed for all the Experimental Facility variables. New data can also be generated to be monitored using GMDH model.

REFERENCES

1. Farlow, S. J. *Self-organizing methods in modeling: GMDH-type algorithms*. New Your: M. Dekker, 1984.
2. 11. IVAKHNENKO, A. G. **Self-teaching Systems of Recognition and Automatic Control**. Moscou: *Tekhnika*, 1969.
3. IVAKHNENKO, A. G.; YARACHKOVSKIY, Yu. P. Self-organization at a System of Complex Models. *Radio and Signal*. Moscou, 1981.
4. FRENCH ASSOCIATION FOR STANDARDIZATION (FIP). “Bus for Exchange of Information Between Transmitter, Actuators and Programmable Controllers”, NF C46 601-607, Mar. 1990.
5. http://www.scadabr.org.br/?q=node/1 O que é um sistema SCADA. Accessed March, 2014.
6. NA, M. G., SHIN, S. H., JUNG, D. W. H., KIM, S. P., JEONG, J. H., LEE, B. C., “Estimation of break location and size for loss of coolant accidents using neural networks”, *Nuclear Engineering and design*, **v. 232(3),** pp. 289-300, 2004
7. On-line Monitoring for Improving Performance of Nuclear Power Plants, Part 1: Instrument Channel Monitoring, IEA Nuclear Energy Series No. NP-T-1.1 , Vienna, 2008.
8. PEREIRA, I. M., MORAES, D. A. Development of a fault test experimental facility model using Matlab. International Nuclear Atlantic Conference, São Paulo, 2015.