PRE-PROCESSING ALGORITHM FOR THE IEA-R1 RESEARCH REACTOR DATA

Iraci Martinez Pereira Gonçalves
IPEN – Instituto de Pesquisas Energéticas e Nucleares
Travessa R-400, São Paulo – SP, CEP 05508-970 Brazil
Phone: 55-11-3816-9159 Fax: 55-11-3816-9432
E-mail: martinez@net.ipen.br

Abstract. *This paper presents a preprocessing algorithm divided in two steps: input variables are scaled independently of each other and then, using the stepwise regression algorithm, the initial number of variables are reduced. The preprocessing algorithm was applied to the IEA-R1 research reactor-monitoring algorithm. Data are collected from the SAD (Data Acquisition System), which has the objective of monitoring and registering the main reactor operational parameters. The SAD monitors a total of 58 operational variables, including temperature, flow, level, pressure, radiation, nuclear power and rod position variables. This database will be used to perform plant condition monitoring and sensor monitoring for fault detection. Using the preprocessing algorithm described in this work, the number of variables was reduced from 58 to only 8 significant variables. This reduced number of input variables allows a better performance in the model construction.*

Key Words: *pre-processing, plant monitoring, data-driven-model*

1. INTRODUCTION

Modern science and engineering are based on using *first-principle* models to describe physical, biological, and social systems. Such an approach starts with a basic scientific model and then builds upon them various applications in mechanical engineering or electrical engineering. Under this approach experimental data are used to verify the underlying first-principle models and to estimate some of the model parameters that are difficult to measure directly. However, in many applications the underlying first principles are unknown or the systems under study are too complex to be mathematically described. With the growing use of computers and low-cost sensors for data collection, there is a great amount of data being generated by such systems. In the absence of first-principle models, such readily available data can be used to derive models by estimating useful relationships among system's variables. Thus there is currently a paradigm shift from the classical modeling based on first principles to developing models from data.

The general experimental procedure involves the following steps, adapted from Dowdy and Wearden (1991):

1. State the problem.
2. Formulate the hypothesis.
3. Design the experiment.
4. Collect the data.
5. Performing preprocessing.
6. Estimate the model.
7. Draw the conclusions.

Generally a good preprocessing method provides an optimal representation for a problem, by incorporating a priori knowledge in the form of application-specific encoding and feature selection.

In this paper the input variables are scaled independently of each other. The preprocessing algorithm was applied to the IEA-R1 research reactor-monitoring algorithm. Data are collected from the SAD (Data Acquisition System), which has the objective of monitoring and registering the main reactor operational parameters. The SAD monitors a total of 58 operational variables, including temperature, flow, level, pressure, radiation, nuclear power and rod position variables. The SAD allows to record data bases containing the time history of all monitored process variables. This database will be used to perform plant condition monitoring and sensor monitoring for fault detection.

2. IEA-R1 EXPERIMENTAL REACTOR

The monitoring system was applied to the IEA-R1 experimental reactor at IPEN, São Paulo Brazil. Figure (1) shows, schematically, the IEA-R1 reactor primary and secondary circuits and the sensors. For the development of the monitoring algorithm using the Group Method for Data Handling (GMDH), data is being collected from the IEA-R1 reactor's SAD (Data Acquisition System). The IEA-R1 Data Acquisition System (Tanomaru and Hiromoto, 1995) has the main objective of monitoring and registering the main reactor operational parameters. The monitoring function is independent of the Instrumentation and Control panels indications installed in the Reactor Control Room. The SAD is composed by a signal conditioning and processing module and PC-based man-machine interface software.
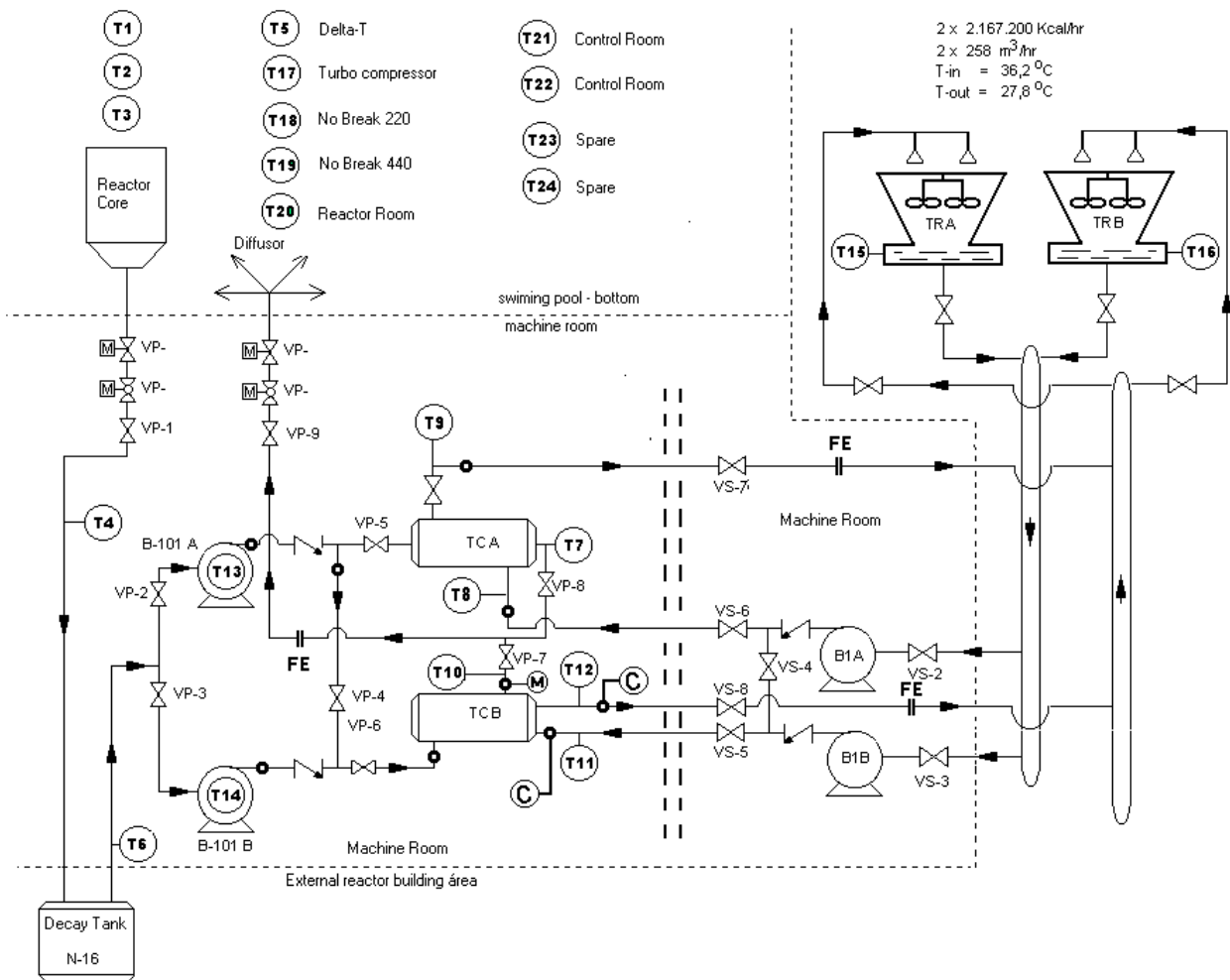

Figure 1. Schematic Diagram of the IEA-R1 Reactor

3. DATA ACQUISITION SYSTEM

The SAD signals are compared with level alarm setting points and when these levels are violated an indication of occurrence is produced. A total of 58 operational variables are monitored by the SAD, including temperature, flow, level, pressure, radiation, nuclear power and rod position variables. The variables monitored by the SAD are listed in Tables (1) to (4). The SAD allows to record data bases containing the time history of all monitored process variables. This database will be used to perform plant condition monitoring and sensor monitoring for fault detection.

All variables are acquired at every 1-minute, during one cycle operation, from start up to shut down. The IEA-R1 reactor full cycle is one-week long, starting on Monday and ending on Thursday. During the full operating cycle there are non-stationary periods of time: the start-up, moving control rod positions and the shutdown period. Only the steady-state reactor operation period was considered to build the model.

To illustrate the IEA-R1 data, variables of one cycle operation are shown in Figures (2) to (5). Process variables are not shown because they are roughly constants. Each variable has 1500 samples and shows roughly one cycle operation, from start up to shut down.

Table 1. SAD Nuclear Variables

Z1	control rod position
Z2	safety rod 1 position
Z3	safety rod 2 position
Z4	safety rod 3 position
Period	period (log channel)
N2	% power (safety channel 1)
N3	% power (safety channel 2)
N4	% power (safety channel 3)
N5	logarithmic power (log channel)
N6	% power (linear channel)
N7	% load (automatic mode)
N8	core power (N16 channel)

Table 2. SAD Process Variables

F1M3	primary loop coolant flow rate
F2M3	secondary loop coolant flow rate (heat exchanger A)
F3M3	secondary loop coolant flow rate (heat exchanger B)
F23	emergency coolant flow rate
DP	core pressure drop
C1	pool water conductivity after treatment
C2	pool water conductivity before treatment
L1	decay tank water level

Table 3. SAD Radiation Variables

R1M3	core support bridge, left side
R2M3	core support bridge, right side
R3M3	reactor hall, north wall
R4M3	pool south wall, above beam tube 8
R5M3	pool west wall, above beam tubes 3 and 4
R6M3	reactor hall, east wall
R7M3	sump well with drained pool water
R8M3	Mg-27 and Na-24 activity between the resin columns in the primary water treatment system
R9M3	heat exchanger A
R10M3	exhaust duct between heat exchanger A and B
R11M3	air conditioning machine room, exhaust duct in the first floor
R12M3	reactor hall, air exhaust duct
R13M3	air exhaustion duct in the stack
R14M3	retention tank

Table 4. SAD Temperature Variables

|	
T1	pool water temperature – at the surface
T2	pool water temperature – half length
T3	pool water temperature – above the core
T4	decay tank inlet
T5	delta T (T4-T3)
T6	decay tank outlet
T7	heat exchanger A primary side outlet
T8	heat exchanger A secondary side inlet
T9	heat exchanger A secondary side outlet
T10	heat exchanger B primary side outlet
T11	heat exchanger B secondary side inlet
T12	heat exchanger B secondary side outlet
T13	primary pump B1A motor casing
T14	primary pump B1B motor casing
T15	cooling tower A external temperature
T16	cooling tower B external temperature
T17	turbo compressor motor casing
T18	NO-BREAK –220V temperature
T19	NO-BREAK – 440V temperature
T20	control room ambient temperature
T21	control room ambient temperature
T22	control room ambient temperature
T23	core inlet coolant temperature (K type thermocouple)
T24	core outlet coolant temperature (K type thermocouple)
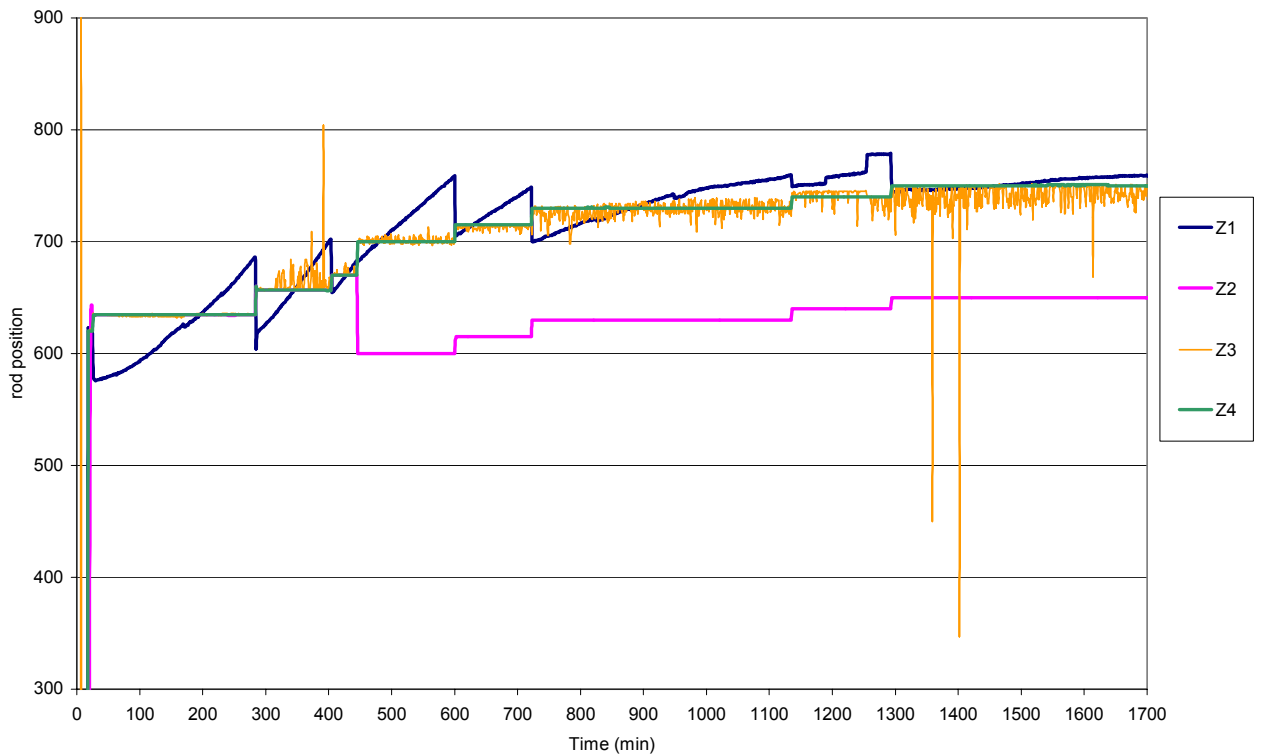
Figure 2. IEA-R1 Reactor Nuclear Variables
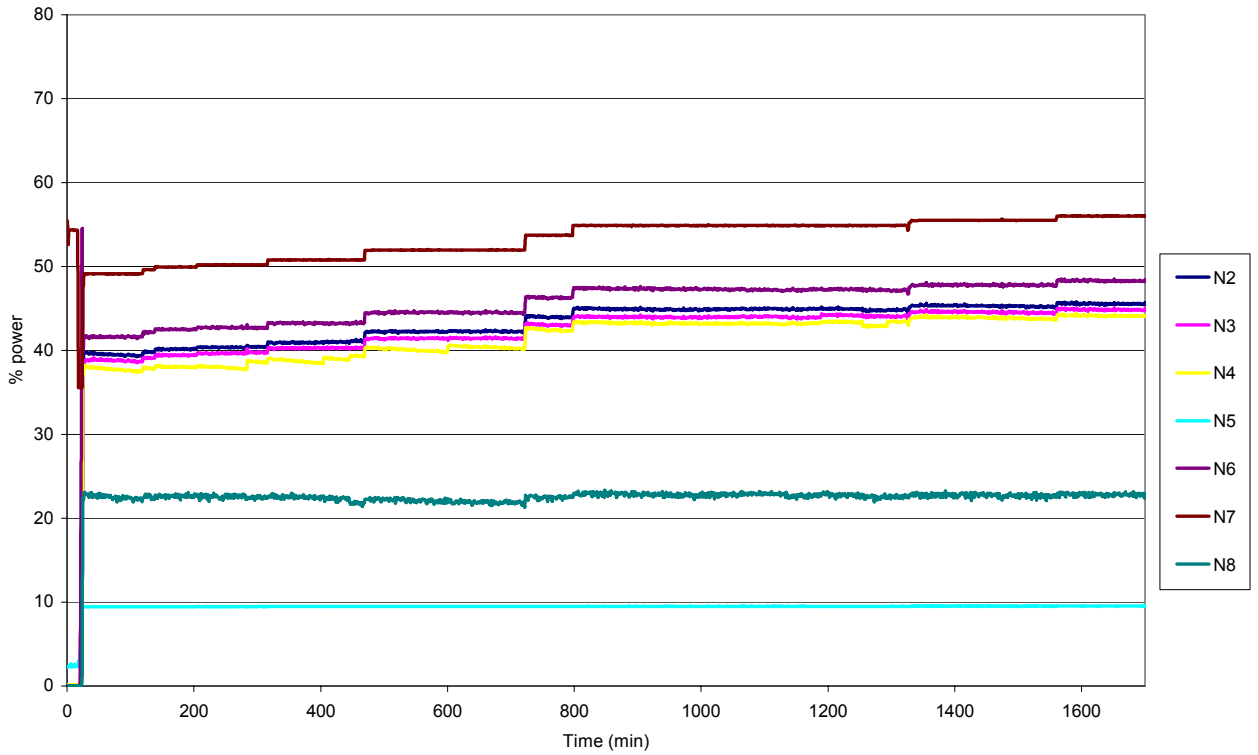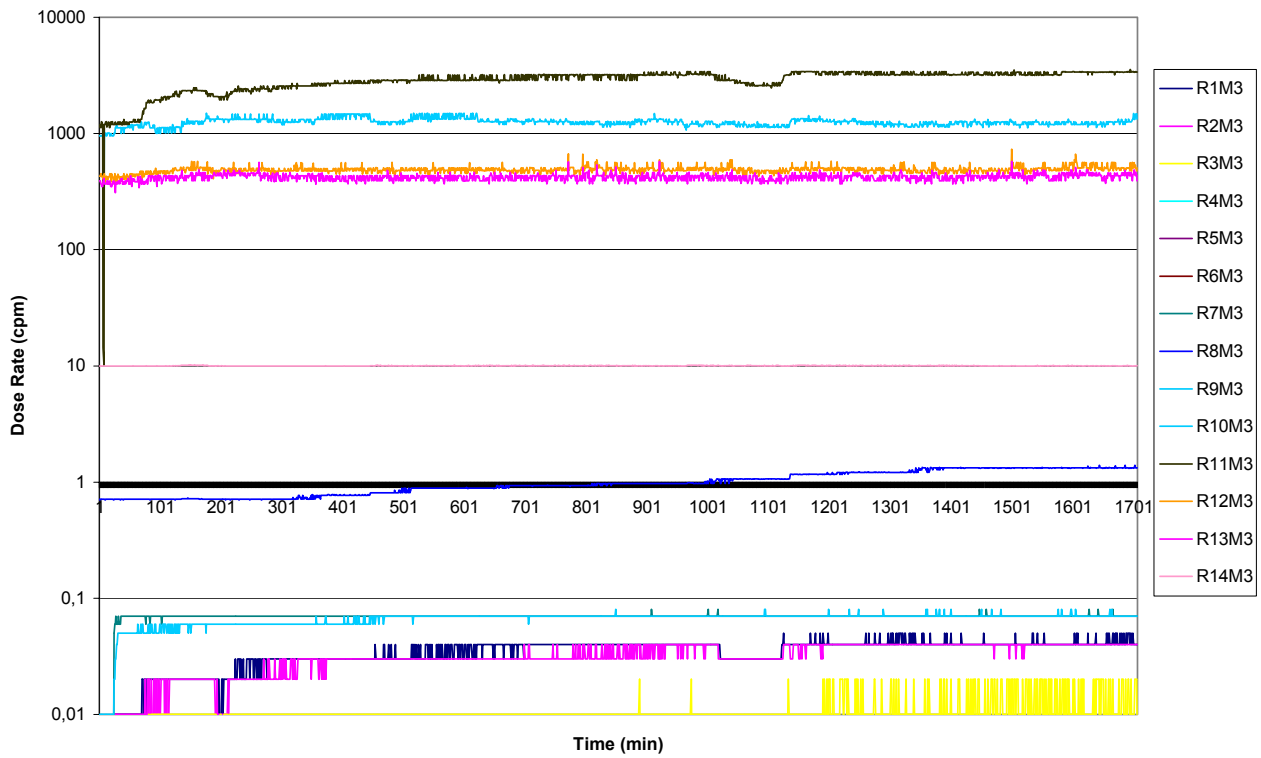
Figure 3. IEA-R1 Reactor Nuclear Variables
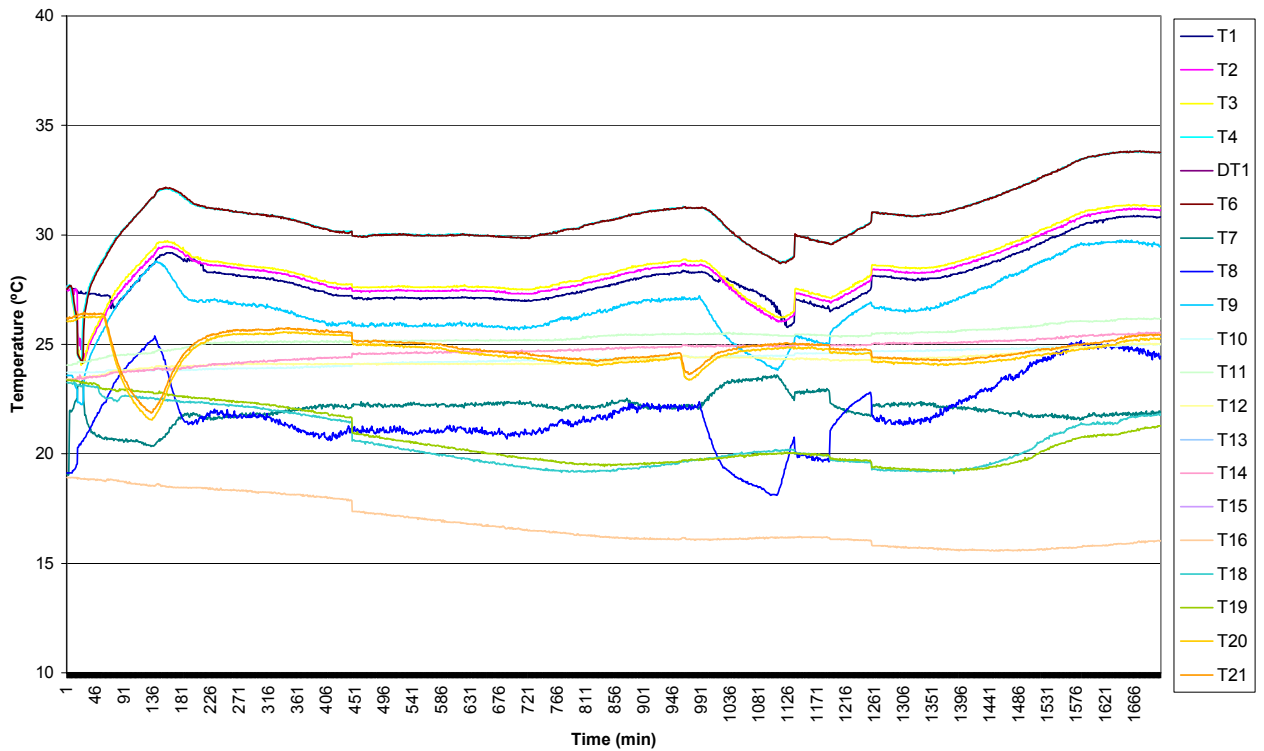

Figure 4. IEA-R1 Reactor Radiation Variables

Figure 5. IEA-R1 Reactor Temperature Variables

4. PRE-PROCESSING ALGORITM

A preprocessing technique is applied to the raw data to ensure consistency of the results. The preprocessing procedure can be divided in two parts:
- filtering and scaling
- variable reduction using Stepwise Regression

4.1. Filtering and scaling

Prior to scaling, the variables were centered to zero mean to provide an intercept-free model and to remove the correlation between the constant term and all other terms. In the current study we used only column methods for scaling the data matrices. The response variable Y was centered to zero mean. The type of variables scaling used was standard scaling, that divides each variable by its standard deviation. After scaling and filtering, we calculate the correlation coefficient matrix. The results for nuclear variables are shown in Tables (5) and (6).

Table 5. Nuclear Variables Correlation Coefficients

	Z1	Z2	Z3	Z4
Z1	1	0.9394	0.9693	0.9791
Z2	0.9394	1	0.9669	0.9564
Z3	0.9693	0.9669	1	0.9933
Z4	0.9791	0.9564	0.9933	1

Table 6. Nuclear Variables Correlation Coefficients

	PER	N2	N3	N4	N5	N6	N7	N8
PER	1	0,4367	0,4352	0,4408	0,4602	0,4369	0,4372	0,039
N2	0,4367	1	0,9172	0,9504	0,9695	0,9971	0,9987	0,1822
N3	0,4352	0,9172	1	0,9921	0,9468	0,9158	0,924	0,1563
N4	0,4408	0,9504	0,9921	1	0,9644	0,9501	0,9564	0,1615
N5	0,4602	0,9695	0,9468	0,9644	1	0,9726	0,9685	0,1806
N6	0,4369	0,9971	0,9158	0,9501	0,9726	1	0,9965	0,1808
N7	0,4372	0,9987	0,924	0,9564	0,9685	0,9965	1	0,178
N8	0,039	0,1822	0,1563	0,1615	0,1806	0,1808	0,178	1

4.2. Variable reduction using Stepwise Regression

In this work, we want to predict a variable and have several predictor variables to choose from. Stepwise regression is a function used to choose the important variables to use in the regression model. The function fits a regression model to the input variables in an interactive environment (Draper and Smith 1981). The model starts with all regressor variables and can exclude or add predictor variables at each step.

In the stepwise algorithm an interactive display allows the user to select and deselect variables to use as inputs to the regression model. After selection, the program shows whether the corresponding coefficients are statistically significant or not.

We can see from the analysis made in the previous section that there are many variables that are very highly correlated with each other. This means that these variables contain the same type of information, and the information is redundant. Then we do not need to use all the 58 variables to construct the model.

We choose the following 14 variables based on the correlation coefficient results: Z1, Z2, PERIOD, N2, N5, F1M3, F23, DP, L1, T1, T3, T6, T14 and T21. Using the stepwise algorithm, we can classify the input model variables according to its statistical significance. We start using the model with all regressor variables and excluding one predictor variables at each step according to the statistical results. From this analysis we find out that the less significant variables are Period, F1M3, F23, DP, L1 and Z1. The results are shown in Figure (6), where we can see that if these variables were not in the model, the changes in the regression coefficient would not be significant. From this analysis we can conclude that the more significant variables to use in the regression model are T6 T3 T14 N2 T14 N5 T21 and Z2.
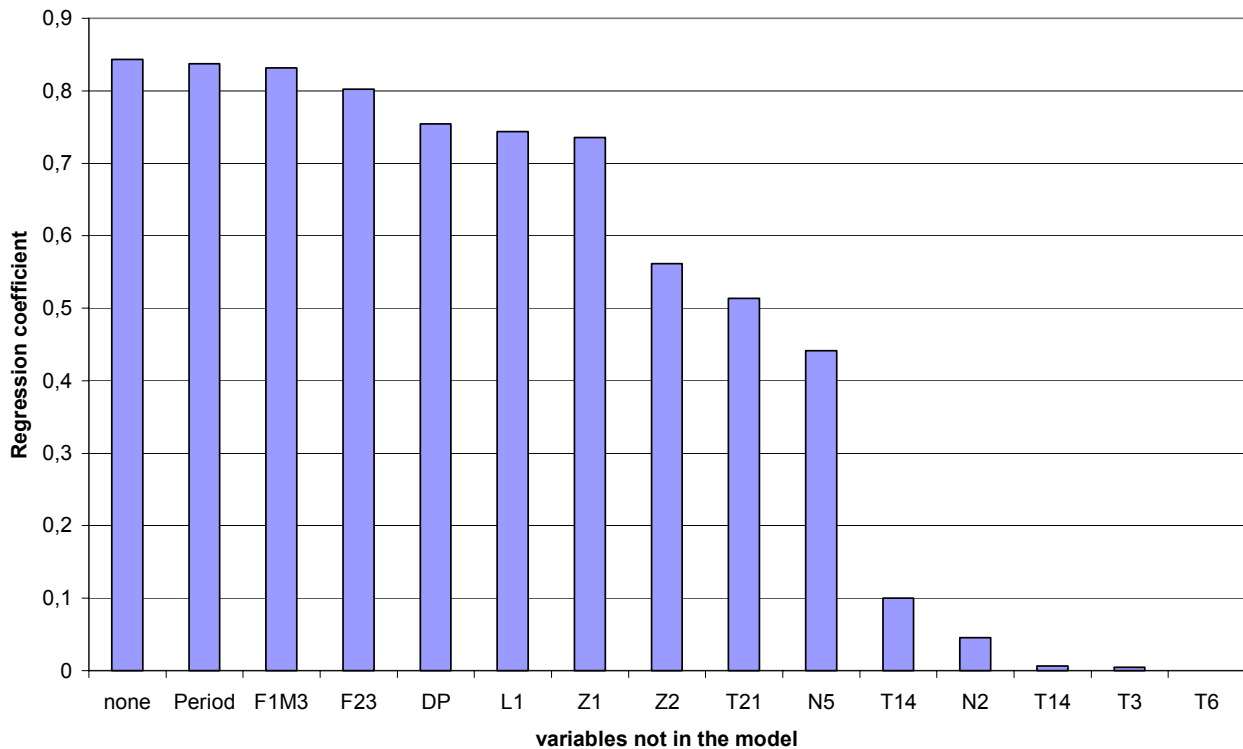
Figure 6. Preprocessing results.

5. CONCLUSIONS

The preprocessing algorithm is divided in two steps. In the first step, input variables are scaled independently of each other. After this first step, we can see that there are many variables that are very highly correlated with each other. This means that these variables contain the same type of information, and the information is redundant. From the initial 58 variables, we choose 14, making the first variable reduction. The second step was to use the stepwise regression algorithm. Using the 14 variables as input, we find out that the less significant variables are Period, F1M3, F23, DP, L1 and Z1. If these variables were not in the model, the changes in the regression coefficient are not significant. The more significant variables to use in the regression model are T6 T3 T14 N2 T14 N5 T21 and Z2. Using the preprocessing algorithm described in this work, the number of variables was reduced from 58 to only 8 significant variables. This reduction in the number of input variables allows a better performance in the model construction.

6. REFERENCES

Dowdy, S. M. and Wearden, S., 1991,"Statistics for Research", New York, Wiley.
Draper, N. and Smith H., "Applied Regression Analysis", Second Edition, John Wiley and Sons, Inc. 1981 pp. 307-312.
Tanomaru, N. and Hiromoto, Y., 1995, IEA-R1 Reactor's Acquisition Data System - Installation and Operation Manual, SP.