An Intelligent Data Analysis-Base: Evaluation of Nuclear Power Plants Output Flow

Azizul Azhar Ramli, *Member, IACSIT*, Junzo Watada, and Witold Pedrycz

Abstract— **Abstract In order to realize stable electricity generation, nuclear power plant (NPP) generators are evaluated in their performance of generated output power in term of quality and quantity. Therefore, the evaluation is realized on the basis of several influential factors, which have to be analyzed via the exploitation of heterogeneous data sets obtained from scattered locations and different types of sources. In this paper, we stress the pivotal role of extended fuzzy switching regression analysis in handling this type of data, which come from real world of the NPPs industry. The key objective of this study is to implement the enhancement of a convex hull approach in the fuzzy switching regression analysis process which can be viewed as an intelligent data analysis (IDA) approach. This approach is concerned with the effective combination of fuzzy sets theory with the analysis of large amounts of online data. For deploying the multisource data problem, the fuzzy switching repression analysis is developed as an IDA by enhancing a fuzzy regression analysis based on convex hull, specifically Beneath-Beyond algorithm. The selected IDA becomes a potential analysis vehicle to successfully reduce the computing time as well as minimize the computational complexity. It is shown that the proposed approach becomes an efficient vehicle for the evaluation of produced output flow by NPPs. The study offers an interesting and practically appealing alterative platform to evaluate the quality and quantity of produced output flow of NPPs.**

Index Terms— **Convex hull, Fuzzy switching regression, nuclear power plant.**

I. INTRODUCTION

Highlight Nowadays, nuclear power plant industry (*NPP industry*) is one of power generation entities that provide an alternative of electricity for the society and industries. Over the past decade, managers of utilities such as NPPs have faced with an increasing number of new challenges. Considering the importance of nuclear energy, one has to emphasize highly optimized results of analysis related to the quality and quantity of produced output flow of nuclear

A. A. Ramli. is with Graduate School of Information, Production and Systems (IPS), Waseda University, 2-7, Hibikino, Wakamatsu-ku, Kitakyushu-shi, Fukuoka-ken, 808-0135 JAPAN (phone: +81-80-3981-9429; fax: +81-93-692-5179; e-mail: azizulazhar@ moegi.waseda.jp).

J. Watada, was with Graduate School of Information, roduction and Systems (IPS), Waseda University, 2-7, Hibikino, Wakamatsu-ku, Kitakyushu-shi, Fukuoka-ken, 808-0135 JAPAN (e-mail: junzow@osb.att.ne.jp).

W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, CANADA T6G 2V4 and Systems Research Institute, Polish Academy of Sciences, Warsaw, POLAND (e-mail: pedrycz@ece.ualberta.ca).

power realized in a timely manner becomes a key component for this industry [1].

In addition, an intelligent data analysis (IDA) is one of the pivotal elements which are implemented for different domains including engineering, management, and healthcare. Consequently, implementation of IDA purposely for NPPs related analysis was becomes necessity.

On the other hands, a fuzzy regression analysis exploits linear programming (LP) to describe dependencies among variables. In this context, LP is subject to constraints whose number is proportional to the number of samples (*data points*) and exponential to the number of attributes when constructing fuzzy regression model. The increasing number of attributes along with sample size might directly raise the computational complexity as well as the processing time. Fortunately, this limitation can be handled by using a convex hull approach [2]. Basically, the convex hull is defined as the smallest convex polygon located in multi-dimensional data space which contains all points (*vertices*). Therefore, the enhancement of convex hull approach can efficiently perform real-time fuzzy switching regression analysis implementation process.

Regarding the above problem related to the evaluation of produced output flow of NPPs and an enhancement of fuzzy regression analysis, this study elaborates on a potential vehicle to discover the most influential factors that significantly contribute to the performance of nuclear power output flow. In general, factors are revealed based on collected samples of NPPs data that come from several distributed sources, largely unknown and in many varying formats or in other words, this type of data is considered as heterogeneous sample of data. By this means, we employ an exploited concept and algorithm of convex hull in the realization of real-time fuzzy switching regression analysis [2]. In addition, the adaptation of a convex hull approach helps to alleviate the limitations of the standard switching regression when pursuing real-time data analysis. The main concern here is to decrease the data analyzing time as well as the computational complexity to efficiently support decision-making procedures, specifically for NPPs' industry.

With the intention to show the obtained results of analysis, the selected samples of data are originally obtained from a real world of NPPs' industry. Additionally, the considered intelligent hybrid approach will help to generate an optimal result of produced output flow of NPPs' analysis purposely for generating electricity; we also verify required time for each group of data samples (*initial and newly added samples of data*) during the implementation of the convex hull based fuzzy switching regression analysis.

The paper is organized as follows. Section 2 offers a

concise related literature review which focuses on the state of art in NPPs industry and its operation. Also we include an explanation related to the needs of the evaluation procedure for produced output flow of NPPs. Next, in Section 3, we present a basic idea of fuzzy switching regression analysis and the real-time processing of fuzzy switching regression model, realized with the use of the convex hull approach. Section 4 shows an illustrative example of the NPPs analysis, while Section 5 offers some interpretation of obtained results. Finally, Section 6 covers concluding remarks.

II. THE PRESENT STATE OF ART IN NUCLEAR POWER PLANTS INDUSTRY

The energy demands of the world are continuously increasing. Experts expressed concerns about the future of power generation because of the limitation of supplies of coal, water and gas to fulfill the needs of mankind in the long term future. Alternative sources of energy such as nuclear power energy are being developed. Nuclear power energy has several advantages over other sources of energy because it is not limited by space or location [3]. Nowadays, nuclear power energy became one of the most important electricity energy contributors. Many nations use NPPs to generate electricity for both civilian and military use and some nations also utilize nuclear power to run parts of their naval fleets, especially submarines.

Basically, nuclear reactions fall into two major categories; fission and fusion. Fission refers to the nuclear reaction where a heavy nucleus is broken into nuclei of intermediate atomic number. Fusion refers to the nuclear reaction wherein light nuclei get combined to form a new nucleus. Energy can be either released or absorbed during the process depending on whether the final mass of the products is greater than or less than the initial mass of the reactants [4].

The above mentioned types of reactions are not of much use for generating electricity energy on their own. It requires something known as a controlled chain reaction if power is to be generated in an NPP. Berezna also highlighted that, when fission starts in a nuclear material, it can die out slowly, sustain itself constantly or develop into an uncontrolled reaction. The first and last options are not useful for generation of electricity. It is only when we have a sustained reaction, that we can utilize nuclear power energy in an effective manner.
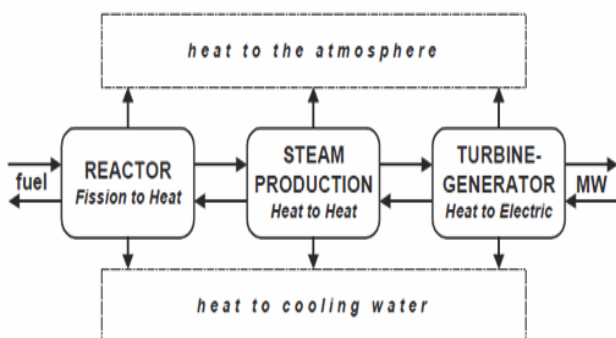

Figure 1 Basic flow of energy in a nuclear generating station [4]

Moreover, NPP is a thermal power station in which the heat source is one or more nuclear reactors. NPPs are base load stations, which work best when the power output is constant. Nuclear generating stations exist for the purpose of converting the energy obtained from the fission of certain nuclei to electricity power. This energy conversion takes place via a number of intermediate stages that require many pieces of equipment organized into several systems under the control and protection of both manual and automatic operations [5]. The basic nuclear generating station of energy cycle is shown in Figure 1.

Generally, fuel containing fissile material (*one that is capable of sustaining a chain reaction of nuclear fission*) (Uranium) is fed to a reactor where fission takes place. The liberated energy appears in the form of heat, which is used to boil water. The steam produced from the boiling water spins a turbine-generator set, where the heat is converted first to kinetic energy in the turbine and to electricity by the generator; the electricity produced (*denoted as megawatts-MW*) is supplied to the electric power system. The details of NPPs operation is discussed in next sub section.

Therefore, related with those situations above, IDA tools are required to support decision making activities especially in NPPs related industrial analysis. With combination of data analysis process and artificial intelligent (AI) approach, the IDA approach is able to efficiently deal with real-time data analysis as well as some level of data complexity such as related to the evaluation of produced output flow of NPPs. The study reported in [6] stresses the need to form a consistent methodology that supports the development of high quality models and offers their further maintenance.

One of the main requirements in a successful industrial data analysis is to perform fast computation with minimal computational complexity. This is because; obtained knowledge from this type of analysis became an important input for decision making procedure. This is an initial step for achieving companies' goal and in these case, companies are referred to the NPP generators [7].

Thus, the integrated methodology amplifies the advantages of the individual techniques, significantly reduces the computational time as well as the complexity of calculation procedure. Therefore, related to this explanation, we can wrap up here that by utilization of a selected convex hull based IDA approach, high risk industry such as NPPs which involve the processing of heterogeneous data sets gets their benefits as well as accomplishes customer needs and achieves company's profit. Overall, this approach will maximize the precision of the appropriate decision making towards successful of NPPs supervision process [8] [9].

A. Brief Overview of Nuclear Power Plants Operation Procedure

In the vast majority of NPPs in the world, the heat energy generated by uranium fuel is transferred to ordinary water and is carried away from the reactor's core either as steam in boiling water reactors (BWRs) or as superheated water in pressurized-water reactors (PWRs). In a PWR the superheated water in the primary cooling loop flows through a special heat ex-changer called a "steam generator" that is used to boil water and create steam in a secondary loop that

feeds the turbo-generator. The two loop design of a PWR keeps the radioactivity isolated and only clean steam is circulated through the turbine [3]. This helps minimize maintenance costs and radiation exposures to the plant personnel. Figure 2 illustrates the general diagram of nuclear power station.
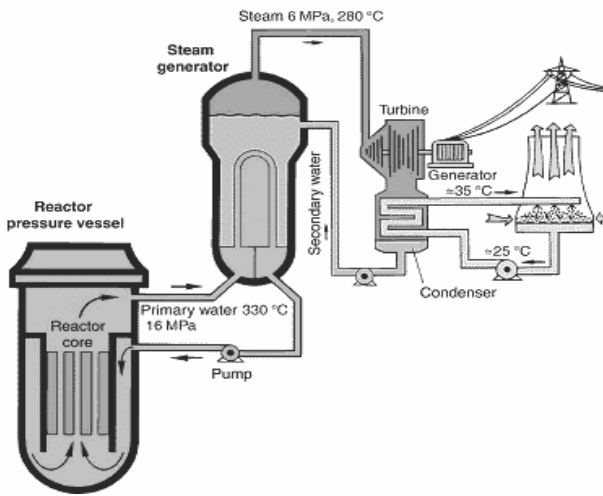

Figure 2 Basic flow of energy in a nuclear generating station

To turn nuclear fission into electrical energy, the most important NPP operators are to control the energy given off by the enriched uranium and allow it to heat water into steam. Enriched uranium is typically formed into inch-long (2.5cm long) pellets, each with approximately the same diameter as a dime. Next the pellets are arranged into long rods, and the rods are collected together into bundles. The bundles are submerged in water inside a pressure vessel. The water acts as a coolant. For the reactor to work, the submerged bundles must be slightly supercritical. Left to its own devices, the uranium will eventually overheat and melt [4] [10].

Brain also stated that, in order to prevent overheating, control rods made of a material that absorbs neutrons are inserted into the uranium bundle using a mechanism that can raise or lower the control rods. Raising and lowering the control rods allow operators to control the rate of the nuclear reaction. When an operator wants the uranium core to produce more heat, the control rods are raised out of the uranium bundle (*thus absorbing fewer neutrons*). To create less heat, they are lowered into the uranium bundle. The rods can also be lowered completely into the uranium bundle to shut the reactor down in the case of an accident or to change the fuel.

The uranium bundle acts as an extremely high energy source of heat. It heats the water and turns it to steam. The steam drives a turbine, which spins a generator to produce power. Humans have been harnessing the expansion of water into steam for hundreds of years [10].

In some nuclear power plants, the steam from the reactor goes through a secondary, intermediate heat exchanger to convert another loop of water to steam, which drives the turbine. The advantage of this design is that the radioactive water/steam never contacts the turbine. Also, in some reactors, the coolant fluid in contact with the reactor core is gas (*carbon dioxide*) or liquid metal (*sodium, potassium*);

these types of reactors allow the core to be operated at higher temperatures [10].

Based on above description, we believe that, those related processes involved some uncertainty measure and it should be highly considered to generate an optimum NPPs output as well as produced electricity energy.

B. The Needs of Evaluation Procedure for Produced Output Flow of Nuclear Power Plants

Basically, the main purpose of evaluation related to the produced output flow of NPPs is to reduce power generation costs as well as fulfill customer needs. The main point of this reason is the achievement with condition-based maintenance and optimal operation control of NPPs. Nowadays, NPPs' operation consists of large available data and records related with their implementation and daily practices. These data are fully useful in order to produce an optimum output flow of NPPs. In addition, major contains of these data are acquired from general theory of process and procedure which related with nuclear power industry [9] [11].

Additionally, Miller and Hajek (1989) highlighted that, there are several behaviors of NPPs' which should be fully considered in order to achieve an excellent performance from the view point of NPPs management as well as increasing the quality and quantity of produced nuclear power energy. The following below is the required procedure [7]:
1) Monitor and comprehend the state of the plant,
2) Identify normal and abnormal plant conditions,
3) Diagnose the abnormal plant condition,
4) Predict the plant response to specific control actions,
5) Select the best available control action, and
6) Implement a feasible control action.

Based on these properties, Kang *et al.* (1992) proposed an expert system for performance evaluation, and diagnosis in NPPs [9]. This system has been developed to support the plant operator aiming at improving the produced output, plant efficiency, and operating flexibility. More specifically, this developed system accomplished several tasks including:
1) Monitor the component performance through the analysis of the performance effective factors,
2) Detect the deviation from the optimum performance,
3) Diagnose the cause of such deviation, and
4) Suggest the possible operational guides to eliminate malfunctions.

The necessary evaluation of produced output flow of NPPs is depending on several strategies, and objectives [5]. The objectives of this paper include:
1) Economy;
   a) Maximizing profit,
   b) Reducing the costs including operation, maintenance and waste management, and
   c) Increasing the load factor.
2) Safety;
   a) Maintaining/increasing safety and reliability level of plants,
   b) Protecting the environment and the public,
   c) Reducing occupational exposure, and
   d) Managing risk and uncertainty.
3) Other;
   a) Modernizing/upgrading a plant, and
   b) Enhancing public acceptance.

When we refer to the first task of Kang *et al.* proposed system and the objectives of output flow evaluation of NPPs, we recognized that, carefully monitor of related NPPs component through significant analysis tools becomes the most important job [9]. Therefore, in order to deal with NPPs data sets which contain fuzziness and ambiguity appearance, we need an IDA tool to successfully perform those related analysis process.

III. Review of Fuzzy Switching Regression Analysis

Regression models are statistical models, which describe the variation in one (or more) variable(s) when one or more other variable(s) vary. In other words, regression models are developed to statistically portray the relationship among the variables, that is, they explain one variable by making use of variation of some other (*independent*) variables. Variables, which are used to explain the other variable, are called explanatory ones [12] [13] [14]. Moreover inference based on such models is known as regression analysis.

In addition, a generic regression analysis is concerned with data, which originate from a single data source. A single functional relationship between the independent or input variables $\mathbf{x} \in \Re^K$ and the dependent or output variable $y \in \Re$ is assumed and this relationship holds for all the data being collected [15]. A general "standard" regression model is then described as follows

$$Y_i = h(\mathbf{x}_i) + \varepsilon_i \qquad 1 \le i \le c , \tag{1}$$

where $h(\cdot)$ is some function and ε_i are independent random variables with zero mean and some variance, $i = 1, 2, \ldots, n$.

In real world situation, we also encounter situations involving heterogeneous data. In addition, we might have prior information as to the split (*partition*) of the overall data set into some homogeneous subsets. Therefore, switching regression methods can be considered as a viable design alternative. Interestingly, switching regression was applied to various fields such as economics and bio-computing field.

An implementation of switching regression is realized for heterogeneous data set by forming c homogeneous subsets of data and determining a regression function for each subset $k(k = 1, 2, \ldots, c)$. In other words, a mixed distribution is given and aimed at splitting this distribution into c homogeneous sets. The performance criterion quantifies the squared differences between the estimated values y of the regression function observed in each subset and the corresponding experimental data. The criterion has to be minimized over all data subsets. Additionally, based on Hathaway and Bezdek in 1993, switching regression model has been discussed in a variety of detail. Assuming that as a single model, it can be drawn from c models in the following:

$$y = f_i(\mathbf{x}; \delta_i) + \varepsilon_i \qquad 1 \le i \le c , \tag{2}$$

where each $\delta_i \in \Omega_i \subset \Re_i^K$, and each ε_i is a random vector with mean vector $\mu_i = 0 \in \Re^T$ and covariance matrix, Σ_i.

Generally, a basic idea of switching regression analysis has been mentioned on Hosmor research that described an example from fisheries application problem [16]. According to this research, the parameters defining the linear growth curves for sexes can be estimated by treating the data analysis as a switching regression problem. The general models read as follows

$$y = f_1(x; \delta_1) + \varepsilon_1 = \delta_{11}x + \delta_{12} + \varepsilon_1 , \tag{3}$$

$$y = f_2(x; \delta_2) + \varepsilon_2 = \delta_{21}x + \delta_{22} + \varepsilon_2 , \tag{4}$$

Moreover, a hybrid combination of discussed approach and fuzzy sets is referred to as fuzzy switching regression. Generally, fuzzy switching regression is a technique for estimating multiple fuzzy regression models for a data set and it has been used for capturing non-linear dependencies among selected input and output variables in many data mining application [16] [17] [18]. Initially, based on Wu *et al.* (2009), Hathaway and Bezdek first combined switching regression with FCM and referred to them as fuzzy c-regression (FCR) [19]. FCR is a fuzzy clustering-based switching regression model where regression errors are also used to form a clustering criterion in the FCM clustering, such as iterative optimization procedure [20].

The realization of fuzzy switching regression is completed in several phases. We start with the heterogeneous data that are divided into several fuzzy sets. For each fuzzy set, a weighted regression is completed, where the weights of the corresponding data are given as the membership degrees of the data to the corresponding subsets. Usually these models are applied when the results of a regression analysis are very poor for the overall data set, *e.g.* the multiple correlation coefficient, $\Re^2$ is far smaller than 1. On the other hand, it is assumed that a mixed distribution is given with a reasonable regression, which could be formed for each component of this distribution.

A. Convex Hull based Fuzzy Switching Regression Analysis

Related to this case, initially, Ramli *et al.* (2009) employed GA-optimized FCM [21]. Additionally, GA-optimized FCM was early proposed by Hall *et al.* [22]. Then, the second part uses of convex hull approach for the implementation of the fuzzy switching regression analysis. The complete procedure can be outlined as a series of the following steps as shown below, see also Figure 3.

Step 1- Data Selection: Select raw data samples which are retrieved from distributed local resources.

Step 2- Build Clusters: Given the data, construct the $C(1 < C < n)$ clusters. The number of clusters depends on the nature of the data. If the number of clusters is not sufficient to represent a structure in the data, it has to be increased. Given the relevance of the clustering phase, we use an augmented GA-based version of the FCM. In our case, we concentrate on Hall *et al.* (1999) in which genetically guided clustering approach was projected [22].

Step 3- Construction of Convex Hull: The adaptation of the convex hull approach is realized by considering the

outside points, which were obtained during the previous process. Such selected points will become vertices and bond each other to produce convex edges. The connected edges form a convex hull for the selected data.

Step 4- Carry Out Fuzzy Regression Analysis: Fuzzy regression is realized on a basis of the constructed clusters. In other words, each cluster comes with its own fuzzy regression models.

Step 5- Process Newly Added Samples: Related to real-time scenario, newly data are added to the current processed data. Therefore, if this group of data becomes huge or we have more processing time, the procedure will restart from **Step 2** else it moves to the next step.

Step 6- Solve LP: By utilizing the convex hull developed so far, the implementation of the LP is easier because of the slightly changes (*either increasing or decreasing*) number of vertices points that must be considered for further computing. In other words, the analysis takes into consideration of selected vertices points which were used for convex hull polygon construction. These vertices are treated as constraints in the LP formulation and used generate each of regression models.
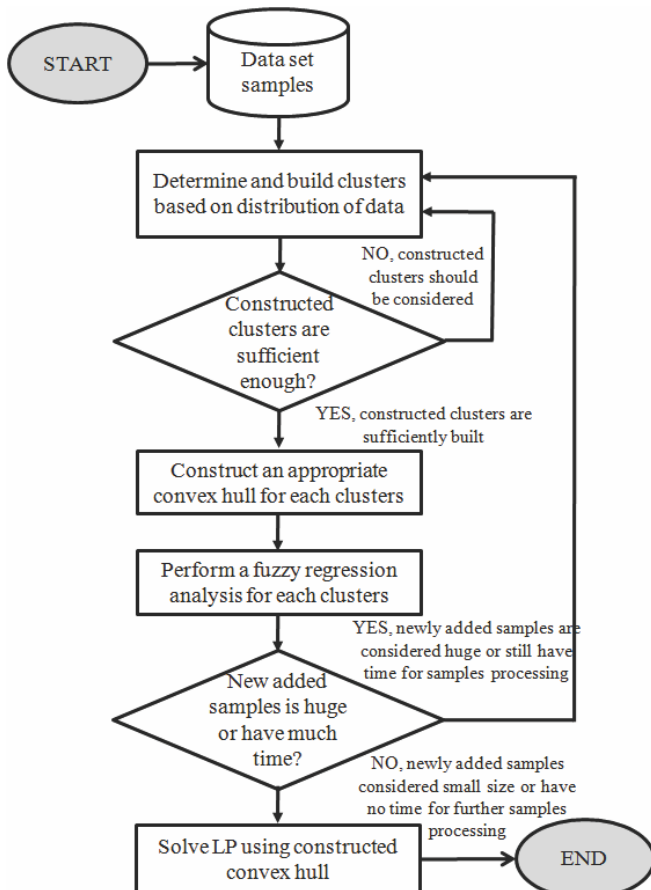

Figure 3 The main processing phases of the NPPs' sample of data

Here, we highlight main features of convex hull approach from the geometric point of view. Based on the original method of the determination of the convex hull, the process has to be realized using all given analyzed data. Ramli *et al.* (2009) propose a new method by building a convex hull for processed data [21]. Focusing on the construction of an appropriate convex hull for each cluster, see **Step 3** in Figure 3, the following sub-steps are completed as follows:

Sub Step 1: Select the outsider points among distributed analyzed data point. These points will become vertices of convex polygon.

Sub Step 2: Connect each selected potential vertex points for constructing convex edges.

Sub Step 3: Connect constructed edges for producing boundaries of convex hull, H_β.

Sub Step 4: Omit points which are not included in the convex hull, H_β.

Sub Step 5: Perform the Beneath-Beyond method to formulate the convex hull, H_0 using one of the selected vertex points which were chosen for building the convex hull.

Overall the selected IDA hybrid approach purposely for produced output flow of NPPs analysis may reduce the computational complexity as well as computing time required to obtain the best solution for this problem especially for real-time implementation.

IV. Nuclear Power Plants Flow Analysis

Since the maximal operating thermal power of any nuclear plants is bounded by the specific licensing requirements and existing certain amount of uncertainty in related input components/items estimation has a direct effect on the maximum energy that can be produced [23]. In order to show an IDA implementation of the NPPs flow analysis, we consider the real-world NPPs data that come from the NPPs industry. Given the complicated generation process of electrical power by deployment of nuclear energy, this set of raw data was obtained from distributed sources and locations and as such it is highly heterogeneous.

The approximate flow of the produced output flow of NPPs, y is determined on the basis of three main input attributes x_1, x_2 and x_3. The details are shown in Table 1.

TABLE 1 Main Inputs and Output for NPPs' Flow Estimation.

i.	Pump differential pressure transmitter (x_1)	Pressure and differential pressure transmitters are currently employed in NPPs to provide a remote indication of critical pressure or pressure differentials, for example, pressure differences in the primary fluid within the nuclear reactor between the inlet and outlet headers of a nuclear steam generator. These measurements are used to monitor the performance of the NPPs' and the reliable and accurate operation of such detector is necessary to ensure safe operations management of the NPPs.
ii.	Pressure transmitter (x_2)	Pressure transmitter may reflect to the accuracy measure value of a process parameter. This value refers to the difference between the actual value of the process and the value that is indicated by the sensor. The static performance or accuracy of pressure transmitters depends on how well the transmitter is calibrated and on how long

		it can maintain its calibration.
iii.	Valve position (x_3)	A valve position indicating device with a movable permanent magnet attached to the valve spindle. Fixed "reed" type switches are permanently encased in epoxy in the switch housing. Two sets of switches provide redundancy, with each set consisting of three switches indicating valve closed mid and full open positions.
iv.	Estimated produced output flow of nuclear power (y)	The energy released from continuous fission of the atoms of the fuel is harnessed as heat in either a gas or water and is used to produce steam. The steam is used to drive the turbines which produce electricity (as in most fossil fuel plants).

As mentioned above, these particular attributes come from distributed resources and each attribute has its own weight as well as specific function. By considering these factors, greatly believe that their sound combination will successfully generate an optimum nuclear dispensation in producing electricity power.

There are two major processes, which are involved in the implementation of real-time fuzzy switching regression analysis. The first process concerns the determination of the clusters. Here we employ the GA-optimized FCM. The second one is utilization of convex hull approach. As presented in Figure 3, we first utilize a genetic based clustering approach to split data into several clusters. The NPPs samples are placed into two clusters. Due to the real-time data processing, the data dynamically changed; therefore the evaluation procedure must be performed carefully to achieve optimal solutions. Next, we execute the convex hull approach for each of the clusters constructed so far. By utilizing coordinates of the selected vertices points, we can easily form fuzzy regression models for each cluster. In real-time data processing, this might become an iterative process based on the frequency of data change.

To demonstrate the dynamical changes of data representation when carrying out real-time data analysis, we initially carry out the IDA for 300 data points. We observe that the constructed convex hull for the first cluster (*triangular shapes*) consists of 10 vertices while another one has 12 convex vertices, see Figure 4. Therefore, utilizing these vertex points as the constraint part of LP, we can directly form an appropriate regression models for each cluster of data.

Next we received 100 more data in real-time. The added samples were processed in the same manner as discussed in the previous example and the obtained results are shown in Figure 5.

There are changes which occurred to the initial constructed of convex hull due to the distribution of the newly received data. The reconstructed convex hull for the first cluster was modified and comes with one new vertex. The convex hull corresponding to the second cluster has been slightly changed and now the total number of vertices is equal to 15.

Finally, the selected vertices, which are used for the reconstruction of convex hull polygon in each cluster become

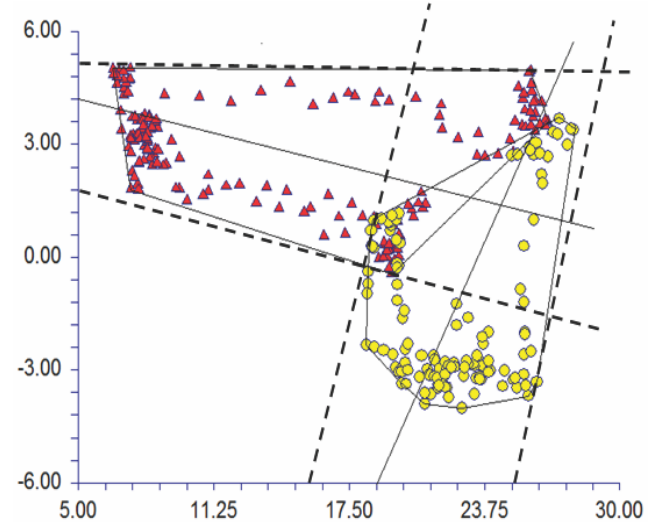the constrain part of LP formulation which will generate fuzzy regression models.

Figure 4 Constructed convex hulls for both clusters (*initial group of selected NPPs samples of data*)
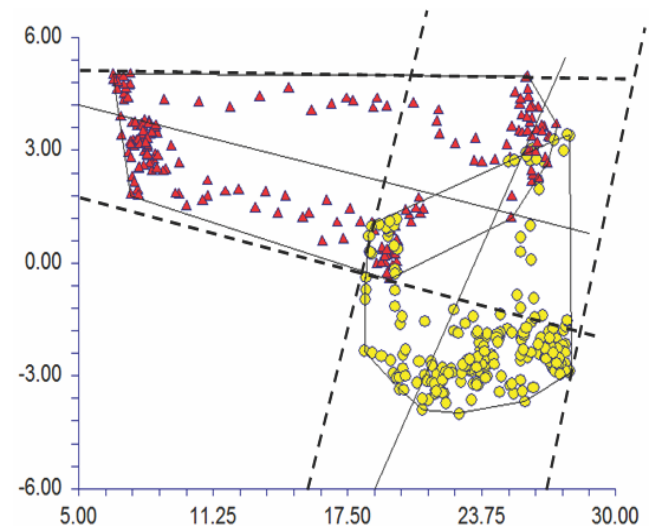

Figure 5 Reconstruction of convex hulls for both clusters (*initial group and newly added of selected NPPs samples of data*)

V. INTERPRETATION OF RESULTS

Ramli *et al.* (2009) noted that increasing the sample size might cause computational difficulties in the implementation of the LP [21]. Another problem may emerge as the entire set of constraints must be reformulated when changes occur with regard to the variables themselves. The increase of computing complexity is alleviated by implementing the IDA approach.

Several researchers have indicated that a good prediction model should not only be able to perform well in terms of sensitivity, specificity, and accuracy but should also be able to interpret and explain the decision arrived at [24]. For this reason, we summarize here, the results provided by the analysis process which was done by convex hull based fuzzy switching regression for selected NPPs' samples discussed in the previous section. In addition, in order to obtain the regression models for this study, an LP approach was used to estimate the fuzzy coefficients for each of the clusters determined clusters in fuzzy switching regression models. As

a result, the adaptive fuzzy regression produced a prediction formula, known also as the logistic score.

TABLE 2 FUZZY SWITCHING AND ORDINARY REGRESSION MODELS: DETAILS.

Regression Method	Group Sample	Regression Models (y_1 - *first cluster*, y_2 - *second cluster*)
Proposed Fuzzy Switching Regression	300 samples	$y_1 = -(18.01, 2.41) - (25.31, 2.22)x_1 - (25.86, 2.35)x_2 - (26.77, 2.56)x_3$ $y_2 = (19.65, 2.54) + (6.68E-02, 2.35)x_1 + (5.65E-02, 2.28)x_2 + (7.14E-02, 2.44)x_3$
	400 samples	$y_1 = -(19.82, 2.74) - (26.23, 2.69)x_1 - (26.43, 2.71)x_2 - (27.04, 2.84)x_3$ $y_2 = (20.52, 3.51) + (7.28E-02, 3.06)x_1 + (6.21E-02, 3.18)x_2 + (6.68E-02, 3.25)x_3$
Ordinary Switching Regression	300 samples	$y_1 = -18.66 - 2.82x_1 - 2.55x_2 - 6.19x_3$ $y_2 = 20.56 + 7.02E-02x_1 + 5.88E-02x + 6.97E-02x_3$
	400 samples	$y_1 = -18.71 - 3.38x_1 - 1.35x_2 - 7.49x_3$ $y_2 = 20.44 + 3.43E-02x_1 + 2.54E-02x + 6.91E-02x_3$

general model: $y = a + (m_1, c_1)x_1 + (m_2, c_2)x_2 + (m_3, c_3)x_3$

where a = intercept, y = prediction output NPPs' performance flow, x_1 = input variable *Pump differential pressure transmitter*, x_2 = input variable *Pressure transmitter*, x_3 = input variable *Valve position*, and m_j = regression coefficient, and c_j = regression spread.

Based on Table 2, we can generalize here, the obtained fuzzy switching regression models for entire sample (*including initial group of data and also plus with newly arrived group of data set*) are shown in a tiny spot differences. For this reason, indicated that, just a slightly changes of distributed data samples in term of spread points as well as their related constructed convex hull polygon. Therefore, we can observe here that the newly added data did not influence too much the regression models and the produced models are highly accurate because reconstructed convex hull covers automatically all points of the analyzed data.

Related to the obtained results interpretation, He *et al.* in 2005 noted that positive regression coefficients present in fuzzy regressions imply positive contribution to the prediction outcome [25]. As clearly stated in the previous section, this study purposely analyzed and evaluated the performance of produced nuclear power output flow with intention for electricity generation. When we refer to the obtained results, generally we can know here that first cluster (*triangular shape*), y_1 for both group samples are not so meaningful since the regression coefficients are both negative. Therefore, we concentrated on second cluster

(*round shape*), y_2 which provides more significant results which can provide good explanation of the related NPPs acquired knowledge.

Hence, prediction model implies that all input attributes have positive contribution to performance flow in NPPs since all of them have positive regression coefficients. Like that, we generally can assume here, all selected attributes (*Pump differential pressure transmitter, x_1; Pressure transmitter, x_1 and Valve position, x_1*) which related to NPPs performance flow evaluation. Moreover, each of them have its own weight to effectively produce electricity based on nuclear power flow, referring to the obtained regression coefficients, m_j seem that it should be a similar precedence level for each attribute and we highly believe that an optimum established of produced output nuclear power flow is supposed highly considered these factors.

Now, we move to the interpretation of obtained regression spread, c_j results. These interpretations mostly related with the fuzzy system theory which can consider for future enhancement of the implemented approach. Fully considered this factor, highly optimized regression models which represent the connection between independent and dependent attributes might be successfully resulted. In general, spread value of the fuzzy models depends on maximum and minimum value of a given sample. As shown in Table 2, the spread values of the regression coefficients have positive spreads. Therefore, it means that certain degree of fuzziness is indeed associated with these input attributes. To efficiently deal with an evaluation of NPPs' output flow, such uncertainty issue must be highly reflected on.

TABLE 3 TIME ANALYSIS.

Regression Method	Group Sample	Time Required (seconds)
Proposed Fuzzy Switching Regression	300 samples	00:01.85
	400 samples	00:02.51
Ordinary switching regression	300 samples	00:02.39
	400 samples	00:03.02

In term of the computing time point of view, it was reported in Table 3, we can discover here that, less time for obtaining regression models is required in comparison with the one used by the conventional method. In addition, the dynamical changes of data size do not influence on the computation because real-time processing execution did not affect the performance of used approach. Therefore, the adequate time of processing will becomes one of the major advantages in this highly important and critical situation.

Summarizing, we believe that the usage of IDA purposely for evaluation of NPPs performance flow is highly significant especially for real-time data analysis. The obtained models are of good quality in terms of their accuracy. Overall, the interpretation will become beneficial to NPPs' generators in order to evaluate the produced output flow as well as to improve the company performance.

VI. Concluding Remarks

Recently, the visibility of NPPs industry has increased due to the needs for alternative electricity sources. Therefore, the requirement of an adequate IDA becomes an important tool, which can be used as a supporting evidence to achieve an optimum produced output flow of NPPs. In this paper, we have reported on the implementation of IDA development of fuzzy switching regression model for analysis produced output flow of NPPs which can be regarded as a useful tool to an array of essential problems of real-time data analysis; especially those encountered in the industry and manufacturing field such as performance of NPPs' production.

We conclude here that the implementation of convex hull based (*specifically Beneath-Beyond algorithm*) fuzzy switching regression approach to the NPPs' samples of data which concern with the evaluation of produced output flow is highly significant. Based on the obtained results, We can note that, all input attributes (*Pump differential pressure transmitter-* x_1 , *Pressure transmitter-* x_2 *and Valve position-* x_3) have their own weight and role in order to support the quality as well as quantity of produced output flow of NPPs, y .

On the other hand, theoretically, real-time data analysis dealt with dynamically change of data appearance, therefore the use of the selected algorithm which deploy a fuzzy switching regression analysis by reconstruction of particular edges and considering new vertices for which the re-computing was realized. This situation is merely suitable for this evaluation of produced output flow of NPPs. The implementation of this IDA approach clearly highlights that the convex edges of the constructed convex hull become data boundaries under which the other analyzed data points are found inside the constructed convex hull. In addition, it is worth stressing the adaptation of the convex hull approach in order to improve a real-time fuzzy switching regression analysis procedure especially for this NPPs related evaluation case.

Moreover, we successfully showed that the number of obtained vertices of convex hull edifice will not drastically change. Therefore, retaining the computing effort is relatively constant in spite of an increasing number of data samples, this IDA approach can be applied to real-world large-scale systems, especially for real-time computing environment such as related on performance of produced output flow of NPPs'. In addition, this approach does not lead to repetition of computing which focuses only on newly arriving data, which potentially can become new vertices. This strategy becomes suitable for implementation of real-time fuzzy switching regression where the convex hull can effectively handle new data with low computational overhead thus decreasing the overall processing time.

In conclusion, our goal was to establish a practical approach to analyze a performance in term of quality and quantity of produced output flow of NPPs by solving a hybrid combination of fuzzy switching regression analysis with an enhancement of convex hull approach called Beneath-Beyond algorithm. We offered some evidence showing that an implantation of IDA purposely for evaluation of produced output flow of NPPs is successfully done where the ability this method performed as a randomized incremental algorithm that is truly output-sensitive to the number of vertices. In advance, this approach uses less space than most of the randomized incremental algorithm and executes faster for inputs with non-extreme points, especially when dealing with real-time data analysis. Therefore, this realization of IDA on the evaluation of produced output flow of NPPs were becomes an efficient vehicle for analyzing real world data where uncertainty issue cannot be avoided.

References

[1] E. D. Carvalho, R. A. Rixey, J. P. Shepley, J. O. Gomes and S. Guerlain, "Design of a nuclear power plant supervisory control system," The IEEE Systems and Information Engineering Design Symposium, April 28, 2006, pp. 251-256.

[2] A. A. Ramli and J. Watada, "A novel idea of real-time fuzzy switching regression analysis: A nuclear power plants case study," In V-N. Huynh, Y. Nakamori, J. Lawry and M. Inuiguchi, (Eds.), Integrated Uncertainty Management and Applications, Advances in Soft Computing, Springer Berlin/Heidelberg, Vol. 68, 2009, pp. 535-546.

[3] H. M. Hashemian, D. W. Mitchell, R. E. Fain and K. M. Petersen, "Long-term performance and aging characteristics of nuclear plant pressure transmitters," U. S. Nuclear Regulatory Commission, NUREG/CR-5851, 1993.

[4] G. Berezna, "Nuclear Power Plant Systems and Operation," CANDU 9 480/NU Technical Description, AECL document 69-01371-TED-001 Revision 1, 1995.

[5] Nuclear Energy Agency, Committee for Technical and Economic Studies of Nuclear Energy Development and Fuel Cycle, Status report on nuclear power plant life management, NEA/SEN/NDC(2000)6, 2000.

[6] L. Kordon, "Hybrid intelligent systems for industrial data analysis," International Journal of Intelligent Systems, Vol. 19, No. 4, 2004, pp. 367-383.

[7] D. W. Miller and B. K. Hajek, "The knowledge-based framework for a nuclear power plant operator advisor," American Nuclear Society Annual Meeting 1989, Atlanta, GA.

[8] M. R. Greenberg, "Risk analysis and port security: some contextual observations and considerations," Annals of Operations Research, Springer Netherlands, 2009, Available: DOI 10.1007/s10479-009-0631-3.

[9] K-S. Kang, S-W. Cheon and S-H. Chang, Development of an expert system for performance evaluation and diagnosis in nuclear power plants, IEEE Fifth Conference on Human Factors and Power Plants, June 7-11, 1992, 1992, pp. 308-313.

[10] M. Brain, and R. Lamb, "How nuclear power works," HowStuffWorks.com. 09 October 2000, Available: http://science.howstuffworks.com/nuclear-power.htm, 26 June 2010.

[11] S. Genta and J. C. Muoz, "On assigning drivers for a home-delivery system on a performance basis," Annals of Operations Research, Springer Netherlands, Vol. 155, Number 1, 2007, pp. 107-117.

[12] G. Peters, "A linear forecasting model and its application to economic data," Journal of Forecasting, Vol. 20, Issue 5, 2001, pp. 315-328.

[13] C. Tofallis, "Multiple neutral data fitting," Annals of Operations Research, Springer Netherlands, Vol. 124, Nos. 1-4, 2003, 69-79.

[14] A. Bargiela, W. Pedrycz, and T. Nakashima, "Multiple regression with fuzzy data," Fuzzy Sets and Systems, Vol. 158, No. 19, 2007, pp. 2169-2188.

[15] D. J. Pedregal, F. P. Garca and C. Roberts, "An algorithmic approach for maintenance management based on advanced state space systems

and harmonic regressions," Annals of Operations Research, Springer Netherlands, Vol. 166, No. 1, 2009, pp.109-124.

[16] R. J. Hathaway and J. C. Bezdek, "Switching regression models and fuzzy clustering," IEEE Transactions on Fuzzy Systems, Vol. 1, No.3, 1993, pp. 195-204.

[17] K. Jajuga, "Linear fuzzy regression. Fuzzy Sets and Systems," Vol. 20, Issue 3, 1986, pp.343-353.

[18] R. E. Quandt, and J. B. Ramsey, "Estimating mixtures of normal distributions and switching regressions," Journal of the American Statistical Association, Vol. 73, No. 364, 1978, pp. 730-738.

[19] K-L. Wu, M-S. Yang and J-N. Hsieh, "Alternative fuzzy switching regression," International Multi-Conference of Engineers and Computer Scientists (IMECS 2009) Vol. I, March 18-20, 2009, Hong Kong.

[20] K. Honda, T. Ohyama, H. Ichihashi and A. Notsu, "FCM-type switching regression with alternating least squares method," IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008), Hong Kong, June 1-6, 2008, 2008. pp. 122-127.

[21] A. A. Ramli, J. Watada and W. Pedrycz, "New perspectives and applications of real-time fuzzy regression," IEEE International Conference on Fuzzy Systems, 2009, (FUZZ-IEEE 2009), Jeju Island, Korea, August 20-24, 2009, 2009, pp. 1451-1456.

[22] L. O. Hall, I. B. Ozyurt and J. C. Bezdek, "Clustering with a genetically optimized approach," IEEE Transactions on Evolutionary Computation, Vol. 3, Issue 2, 1999, pp. 103-112.

[23] C. Chaton and J. A. Doucet, "Uncertainty and investment in electricity generation with an application to the case of Hydro-Qubec," Annals of Operations Research, Springer Netherlands, Vol. 120, Nos. 1-4, 2003, pp. 59-80.

[24] M. E. Matheny and L. Ohno-Machado, "Generation of knowledge for clinical decision support: Statistical and machine learning techniques, Clinical Decision Support," The Road Ahead. Academic Press, ISBN 0123693772, 2007.

[25] Y-Q. He, L-K. Chan and M-L. Wu, "Balancing productivity and consumer satisfaction for profitability: Statistical and possibilistic regression analysis," European Journal of Operational Research, Vol. 176, Issue 1, 2005, pp. 252-263.

Azizul Azhar Ramli received the B. Sc. Degree in Computer Science from *Univerti Teknologi Malaysia* (UTM), Johor Bahru, Johor Darul Takzim, Malaysia, and M. Sc. in Intelligence System form *Universiti Utara Malaysia* (UUM), Sintok, Kedah Darul Aman, Malaysia in 2002, and 2004, respectively. He is currently pursuing his study at Graduate School of Information, Production and Systems (IPS), Waseda University, Kitakyushu, Fukuoka, Japan. His research interests include management information system, data mining, expert system, fuzzy modeling, and real-time analysis.

Junzo Watada received his B.Sc. and M.Sc. degrees in electrical engineering from Osaka City University, Japan, and his PhD degree from Osaka Prefecture University, Japan. Currently he is a professor of Management Engineering, Knowledge Engineering and Soft Computing at Graduate School of Information, Production and Systems, Waseda University, Japan. He was a recipient of Henri Coanda Gold Medal Award from Inventico in Romania in 2002 and a fellow of both SOFT and Biomedical Fuzzy Systems Association (BMFSA). He is a contributing principal editor, a co-editor and an associate editor of various international journals, including International Journal of Biomedical Soft Computing and Human Science (IJBSCHS) and JSCE of I Mech E. His professional interests include soft computing, tracking system, knowledge engineering, and management engineering.

Witold Pedrycz received the M. Sc., Ph.D., and D. Sci. from the Silesian University of Technology, Gliwice, Poland. He is currently a Professor and Canada Research Chair (CRC) in Computational Intelligence in the Department of Electrical, and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland. He is the author or co-author of numerous papers and 13 research monographs. His current research interests include computational intelligence, fuzzy modelling, knowledge discovery, and data mining, fuzzy control including fuzzy controllers, pattern recognition, knowledge-based neural networks, granular and relational computing and software engineering. Dr. Pedrycz has been a member of numerous program committees of IEEE conferences in the area of fuzzy sets and neuro-computing. He is an Associate Editor of the IEEE Transactions on Fuzzy Systems. He is an Editor-in-Chief of IEEE Transactions on Systems, Man, and Cybernetics, Part A, Information Sciences and the past president of the International Fuzzy Systems Association (IFSA) and North American Fuzzy Information Society (NAFIPS). He is a recipient of the Norbert Wiener Prize of the IEEE Systems, Man, and Cybernetics Society.