# Intelligent prognostic

## Towards an intelligent prognostic approach based on data mining and knowledge management

Safa Ben Salah, Imtiez Fliss and Moncef Tagina

{safa.bensalah, imtiez.fliss, moncef.tagina}@ensi-uma.tn
COSMOS Laboratory, National School of Computer Sciences, Manouba University, Tunisia

**ABSTRACT.** Due to the complexity of increasingly growing industrial processes, many of undetected faults can lead to catastrophic consequences for the entire system functioning. It is then crucial to detect and better more to anticipate the detection of faults. In this context, this paper presents an intelligent prognostic approach to anticipate the detection of faults that can affect a complex system. The proposed approach consists in proposing a multi-agent system using data mining and knowledge management techniques. It finally displays a list of faults that may appear to inform the human operator of the possible state of the system and help him to take the necessary preventive measures.

**KEYWORDS :** Fault prognostic, Complex Systems, Data Mining, Knowledge Management, Multi-Agent Systems.

## 1. Introduction

During their life cycle, industrial systems are prone to faults, which can cause a great damage or even disasters. The need to improve the availability, reliability and thus security lead to change the way of maintenance : Passing from corrective maintenance to predictive maintenance (called in literature Condition-Based Maintenance : CBM) [1, 2]. In this context, prognostic has become a crucial strategy to avoid "catastrophic" fault results. The term prognostic founds its origin in the Greek word "prognôstikos", which means "to know in advance" and it is defined as the estimation of the time to fault of a component (or a system) and the existence of risk or subsequent occurrence of one or more fault modes.

In order to predict different types of complex systems(continuous, discrete, hybrid, centralized and distributed) and guarantee reuse and better performance of our solution, we opted for the development of a multi-agent system[9]. Different approaches to perform fault prognostic have been developed, these methods may be associated with one of the three main categories according to [2, 3] ; model-based prognostic, data-driven based prognostic and experienced-based prognostic. In model based prognostic, the physical components or system and its degradation phenomenon are represented by a set of mathematical lows. Whereas, the data-driven approach aims at transforming the raw monitoring data into relevant behavior models of the system including its degradation. Finally, the approach of prognostic based on experience take into account the data and the knowledge accumulated by experience.

To ensure performance, computation cost, convenience and accuracy of prognostic, we propose to combine the use of data mining and knowledge management. In fact, in several cases, it is difficult to obtain a model that translates accurately the system. On the other hand the proper use of expert feedback and the historical data can lead to significant gains.

A list of faults that may appear are displayed to inform the user of the possible state of the system and help preventive actions. To validate the proposed approach, we rely on a simulation of a complex industrial system : Aircraft Elevator Control System. This paper is organized as follows : the second section is dealing with the prognosis approaches to justify our choice. The third section is devoted to present the new intelligent prognostic approach we propose. The forth section is dedicated to the validation of this approach. Finally, some concluding remarks will be made.

## 2. Prognostic approaches

The main goal of our work is to propose a prognostic approach in order to assist human operator to properly and timely manage faults. Fault prognostic consists of estimating the time before fault of a component (or a system) and the existence of risk or subsequent occurrence of one or more fault modes. Several methods are used to produce powerful solutions to anticipate detection of faults in complex systems [17, 16, 14, 12]. These methods may be associated with one of the three main categories according to [2, 3]namely : model-based prognostic, experience-based prognostic and data-driven based prognostic. The derived model of each approach is then used to predict the future evolution of the degradation of industrial system. To choose the best approach to use, we made a study of the three approaches.

### 2.1. Model based prognostic

This approach consists of representing the physical components or system and its degradation phenomenon by a set of mathematical low. There are several works using this approach for instance [11] and [12].

### 2.2. Data-driven prognostic

This approach aims at transforming historical data into relevant behavior models of the system including its degradation. The historical surveillance data is often the fastest and most reliable source of information to understand the degradation phenomena. Indeed, some previously experienced situations can breed and therefore the prognosis system will recognize it such as the deterioration of a parameter, system transition to monitor in a state of fault, the malfunction of a component, etc. Several prognostic works are based on data-driven [4, 5, 6], etc.

### 2.3. Experience-based prognostic

The approach of prognostic based on experience take into account the knowledge accumulated by experience during the whole exploitation period of the complex system. In fact, the activity of supervision and control of complex systems is a very complex task that requires a great experience. This experience is gained by experts over the years.

### 2.4. The chosen approach

To choice the appropriate approach in our study case, we have done a comparison between the three approaches(summed up in table 1).

| | **Advantages** | **disadvantages** |
|---|---|---|
| **Data-Driven appoach** | -easy to implement<br>-Performance enhances over time.<br>-Low implementation cost | -Need a lot of data<br>-Abscence of physical implementation |
| **Model-based approach** | -Physical approach :quantification of the degradation<br>-Precise | -reduced applicability<br><br>-High implementation cost |
| **Experience-based approach** | -No physical model is required.<br>-Simple to develop and easy to understand | -Domain expert with strong experiential knowledge<br>-Domain expert required to develop rules. |

**Tableau 1.** *Comparative study of three prognostic approach*

It is obvious that none of approaches has only advantages as shown in table1. Therefore, it would be interesting to combine prognostic approaches to improve their prognosis result. The integration of various characteristics is a way to develop new hybrid approaches to overcome the limitations of individual strategies of each method. As it is difficult and expensive, in several cases, to obtain a model that translates accurately the system, we proposed a hybrid approach based on combining the data-driven prognostic and experience-based prognostic.

## 3. A hybrid intelligent approach for prognostic based on data mining and knowledge management

We recall that the objective of our work is to define a reliable prognostic approach for monitoring complex systems and predict the faults that may possibly appear. This a complex task. Moreover, we aim to predict faults in different types of complex systems : continuous, discrete, hybrid, centralized and even distributed ones. To guarantee reuse and better performance of our solution, it will be very interesting to exploit the Multi-agent paradigm [9]. Indeed, the contribution of Multi-agent systems in this perspective is to distribute intelligence across multiple entities which can cooperate in the resolution of the prognosis procedure combing the data mining and knowledge management. Each agent in our system is specialized and has a defined role and is able to communicate with others. The used agents are : the user interface agent, data mining agent, Knowledge management agent, the simulator agent and the predictor agent. The proposed approach is based on the definition of reactive and intelligent agents that can participate in the construction of a comprehensive prognosis solution.

### 3.1. Reactive agents

The user interface agent and the agent simulator are both reactive agents that perform their functions without intelligence.

#### 3.1.1. User Interface agent

The User Interface agent handles everything regarding the communication of the system with the external environment. It provides user with friendly graphical interface through which the prognosis procedure is initiated or stopped. Furthermore, this agent has interactions with different agents. Indeed, user interface agent takes care of sending the data and knowledge for the prognosis procedure and receipt of the final results.

#### 3.1.2. Simulator agent

The simulator will be used to simulate the system model to send the current state to data mining agent and knowledge management agent.

### 3.2. Cognitive agents

The thee other agents are cognitive : data mining agent, Knowledge management agent and the predictor agent.

#### 3.2.1. Data mining agent

The data provided by the User Interface agent will serve as input for the prognosis based on data mining. Data mining is a process of discovering unknown, hidden information from a large volumes of data, extracting valuable information, and then using the information to make critical business decisions. We have used very simple and easy but very powerful data mining technique for predicting the upcoming faults :decision tree [18, 20]. The decision tree is applicable to any type of data whether quantitative, qualitative or a combination of both. It allows the graphic representation of a classification procedure and it has an immediate translation in terms of decision rules. We have used C4. 5 algorithm developed by Quinlan[7] as part of our prognostic approach based on data mining. In our study, our data mining approach extracts information from the stored data by building a decision tree from which we can get decision rules as shown in figure 1.

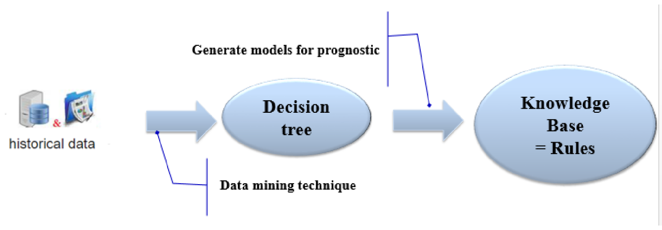


**Figure 1.** *Data mining technique*

#### 3.2.2. Knowledge management agent

Expert knowledge will be used for the prognosis based on knowledge management[8]. Thus, the rules extracted from the expertise will determine the future state of components of the complex system.
Knowledge handled and implemented in our project were acquired by the study of books, articles and reference documents concerning the complex system and also from expert interviews : we directly asked questions to the expert, which helped us to understand how the system works and define the functioning rules. In fact, in order to properly use expert knowledge, these knowledge are expressed as logical rules like "If Condition then Conclusion" as shown in figure 2.

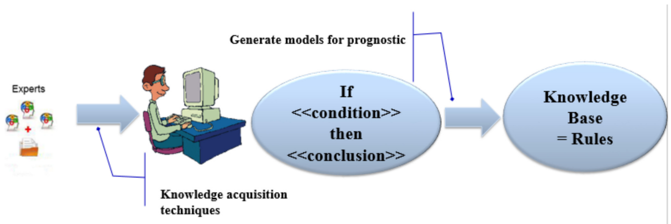


**Figure 2.** *Knowledge management technique*

#### 3.2.3. predictor agent

The Predictor Agent has as role to identify the overall future state of the system. It receives the result of prognosis based on data mining and the result of the prognosis based on knowledge management and subsequently sending the list of faults that can affect the system.

### 3.3. Communication between agents

A fundamental charactistic of multi-agent systems is that individual agents communicate and interact. This is accomplished through the exchange of messages. The figure 4 (in annex) presents a sequence diagram showing messages exchanged between the five agents.

## 4. Validation of the proposed approach

To validate our multi-agent fault prognostic system, we are based on the simulation of a complex industrial system : Aircraft Elevator Control System[13].

### 4.1. System description

The aircraft elevator control system consists of two elevators, the control surfaces. Each of these are controlled by one of two hydraulic actuators while the other one is operating as a passive load. The four actuators take their power from three hydraulic subsystems as depicted in figure 7 (in annex). Two primary flight control units are available to compute actuator control signals and modes.

### 4.2. Data description

The historical data used in our simulations are the data of 8 variables (7 independent variables+ 1 dependent variable). The independent variables are the measures that we have extracted from the simulated system which are data of system components show in figure7(C1 : The right inner actuator, C2 : the right outer actuator, C3 : the left inner actuator, C4 : the left outer actuator to the, H1 : the hydraulic circuit 1, H2 : the hydraulic2 circuit H3 : hydraulic3 the circuit). The dependent variable represents the state of the system (faulty or in not). All independent variables are digital and the dependent variable is nominal. In our study, more than 10 000 values for each variable were recorded every 0. 1 seconds ( as simulation lasts 100 seconds).

### 4.3. Knowledge Description

After studying the reference materials and the discussions with experts, the following rules are generated :

– If the aircraft is flying perfectly level, then the actuator position should maintain a constant value.

– If the position of an actuator increases or decreases by 10 cm from this zero point, then the fault detection system registers a fault in that actuator.

– The fault detection system also registers a fault if the change in actuator position is very rapid (i. e. , the position changes at least 20 cm in 0.01 seconds).

– the fault detection system registers a fault in one of the hydraulic circuits if the pressure is out-of-bounds or if the pressure changes very rapidly.

– the fault detection system checks that the pressure in the hydraulic circuit is between 500 kPA and 2 MPa, and that the pressure changes no more than 100 kPa in 0.01 seconds.

### 4.4. Results and discussion

To assess our fault prognosis multi-agent system (developed using the Jade environment [15]), we have made 127 tests. Whenever we launch system simulation using Matlab Simulink library [10] (the agent simulator handles the connection between our tool and the model simulation in Simulink), then we run the prediction of the data mining agent and the knowledge management agent simultaneously. Data mining agent uses the Weka environment[19] to treat the stored data based and generate the decision tree. The resulted decision tree is given in annex (figure5). The generated rules are given in annex (figure6). The knowledge management agent uses the rules previously presented (section 4.3). Finally, the predictor agent combines the results generated by data mining agent (which use the data mining technique) and knowledge management agent (which will operate knowledge management technique) by attributing to each of them a 50% probability. The result is transmitted to the user interface system to help the user to make the appropriate decision as shown in figure 3. In this example, a message appears informing the user that there will

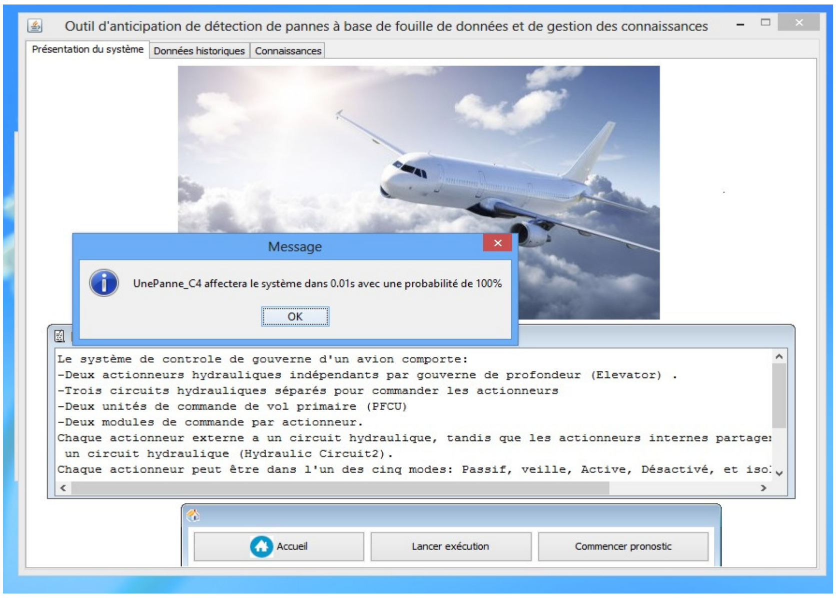


**Figure 3.** *Example of prognostic result*

be a breakdown in 0.01s with a probability of 100% (50% from prognostic result based on data mining and 50% prognostic result based on knowledge management). After the various tests of our multi-agent system, the results are very encouraging. Indeed, we have obtained the correct decision in 100% of tests with accuracy the expected timing and likelihood of occurrence of such faults is case of the example shown in the following figure. This can be explained by the combination of the results of data mining and knowledge management. Indeed, in our study the number of recorded data forms a very good basis for learning and the effectiveness of the used data mining technique. The step of acquiring knowledge also forms another crucial as a basis step for decision making.

## 5. Conclusion

The area of Intelligent Decision Support Systems is very interesting as it assist the decision maker to take the most appropriate decisions at the right time. In this context, we are particularly interested in intelligent prognosis which offers support to decision-maker in case of preventive maintenance. This paper proposed a multi-agent approach to predict faults that may appear in complex systems. This approach is based on the combination of data mining and knowledge management techniques. The simulation results of this approach for the case of the aircraft elevator control system are very encouraging. Future works aim to highlight the potential of such approach in real systems cases.

## 6. Bibliographie

[1] Jardine, A. K. , Lin, D. , & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical systems and signal proces-

sing, 20(7), p. 1483-1510.

[2] M. Lebold, M. Thurston (2001). Open standards for condition-based maintenance and prognostic systems, in :Maintenance and Reliability Conference (MARCON).

[3] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, G. Tripot (2010). A mixture of Gaussians hidden Markov model for failure diagnostic and prognostic, in :IEEE Conference on Automation Science and Engineering, CASE10.

[4] P. Wang, G. Vachtsevanos(2001). Fault prognostics using dynamic wavelet neural networks, AI EDAM-Artificial Intelligence for Engineering Design Analysis and Manufacturing 15, p. 349-365.

[5] HU, Jinqiu, ZHANG, Laibin, MA, Lin, et al. (2011). An integrated safety prognosis model for complex system based on dynamic Bayesian network and ant colony algorithm. Expert Systems with Applications, vol. 38, no 3, p. 1431-1446.

[6] WIDODO, Achmad et YANG, Bo-Suk(2011). Machine health prognostics using survival probability and support vector machine. Expert Systems with Applications, vol. 38, no 7, p. 8430-8437.

[7] Quinlan, J. R. (2014). C4. 5 : programs for machine learning. Elsevier.

[8] Qureshi, S. , V. Hlupic, et R. O. Briggs (2004). On the convergence of knowledge management and groupware. In Groupware : Design, Implementation, and Use, p. 25-33. Springer.

[9] Ferber, J. et J. -F. Perrot (1995). Les systèmes multi-agents : vers une intelligence collective. InterEditions.

[10] The Mathworks, I. , Simulink. Simulation and Model-Based Design. http ://www. mathworks. com/products/simulink/.

[11] Jianhui Luo, Krishna R Pattipati, Liu Qiao et Shunsuke Chigusa(2008). Modelbased prognostic techniques applied to a suspension system. Systems, Man and Cybernetics, Part A : Systems and Humans, IEEE Transactions on, vol. 38, no. 5, p. 11561168.

[12] Douglas E Adams et Madhura Nataraju(2002). A nonlinear dynamical systems framework for structural diagnosis and prognosis. International Journal of Engineering Science, vol. 40, no. 17, p. 19191941.

[13] Otter, Martin(2002), Simulation for Analysis of Aircraft Elevator Feedback and Redundancy Control, dynamics (typically not more than three continuous state variables), vol. 5, no. 13, 27.

[14] Kan, Man Shan and Tan, Andy CC and Mathew, Joseph (2015), A review on prognostic techniques for non-stationary and non-linear rotating systems, Mechanical Systems and Signal Processing, vol. 62, p. 1-20, Elsevier.

[15] Bellifemine, Fabio Luigi and Caire, Giovanni and Greenwood, Dominic(2007), Developing multi-agent systems with JADE, vol. 7, John Wiley & Sons.

[16] El-Koujok, Mohamed and Gouriveau, Rafael and Zerhouni, Noureddine(2014), Development of a prognostic tool to perform reliability analysis, Proc. of the ESREL-17th SRA-Europe Conf. , Valencia, Spain, sept. 22, vol. 25, p. 191-199.

[17] Widodo, Achmad and Caesarendra, Wahyu(2014), SUMMARY OF THE RECENT DEVELOPED TECHNIQUES FOR MACHINE HEALTH PROGNOSTICS, ROTASI, vol. 16, no. 1, p. 21-27.

[18] Quinlan, J. Ross(1986), Induction of decision trees, Machine learning, vol. 1, no. 1, p. 81-106, Springer.

[19] Hall, Mark and Frank, Eibe and Holmes, Geoffrey and Pfahringer, Bernhard and Reutemann, Peter and Witten, Ian H (2009), The WEKA data mining software : an update, ACM SIGKDD explorations newsletter, vol. 11, no. 1, p. 10-18, ACM.

[20] Kothari, RAVI and Dong, MING(2001), Decision trees for classification : A review and some new results, Pattern Recognit, vol. 171, p. 169-184, World Scientific.

## 7. Annex

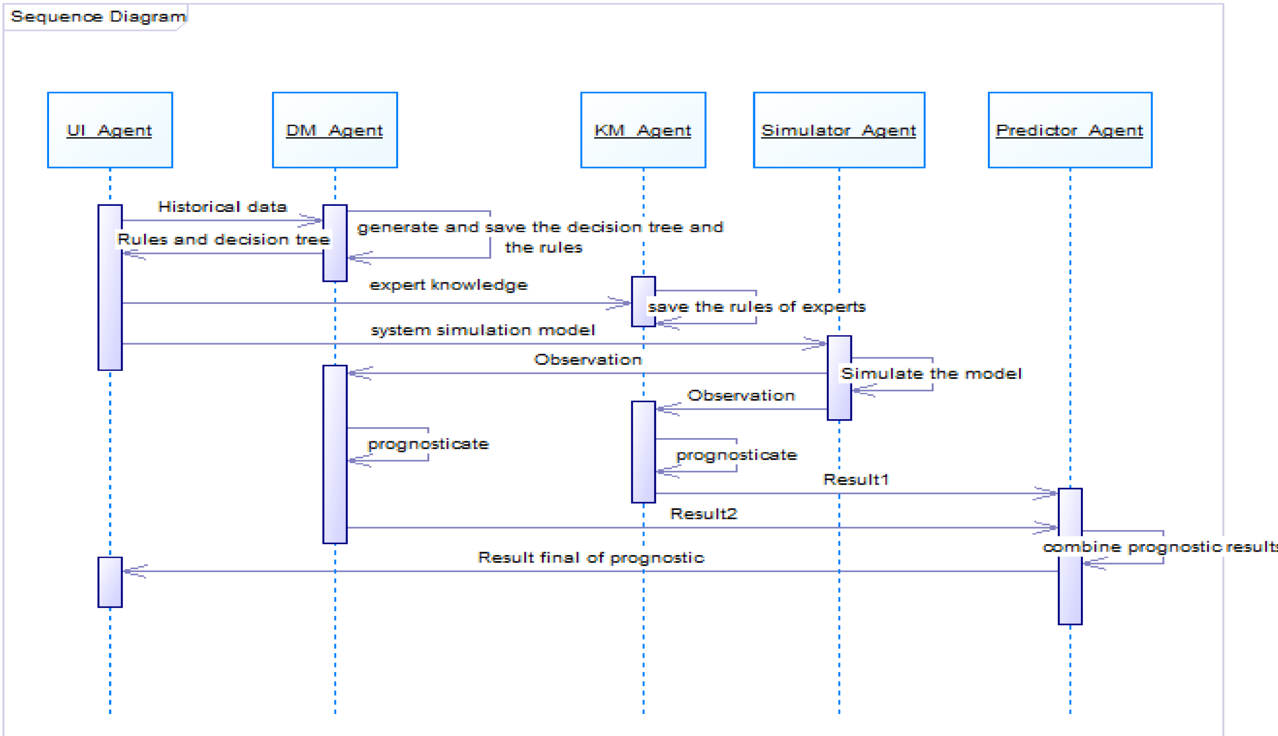


**Figure 4.** *sequence diagram*

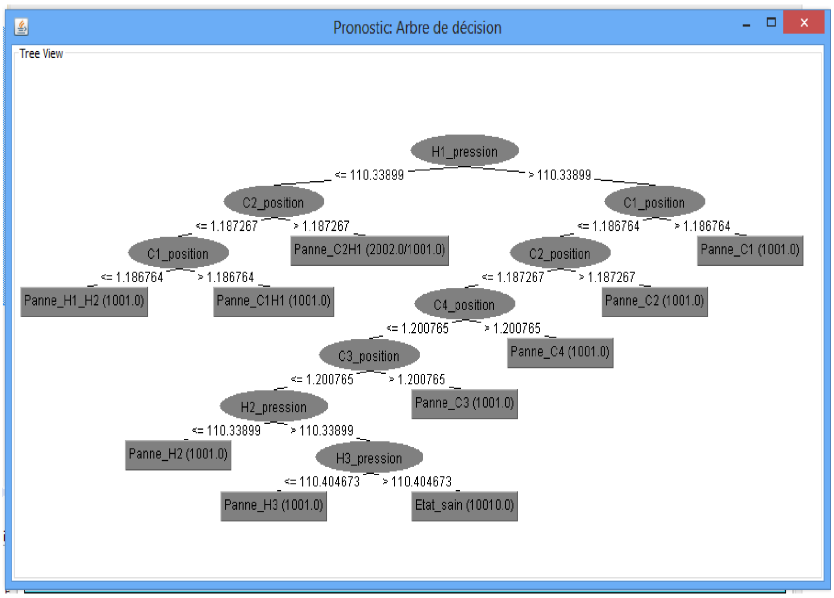


**Figure 5.** *Decision tree*

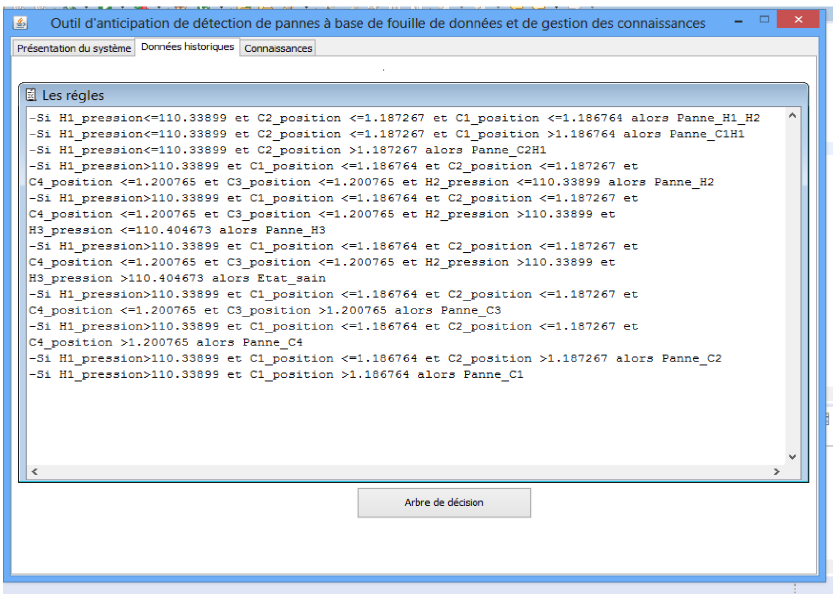


**Figure 6.** *Decision rules*

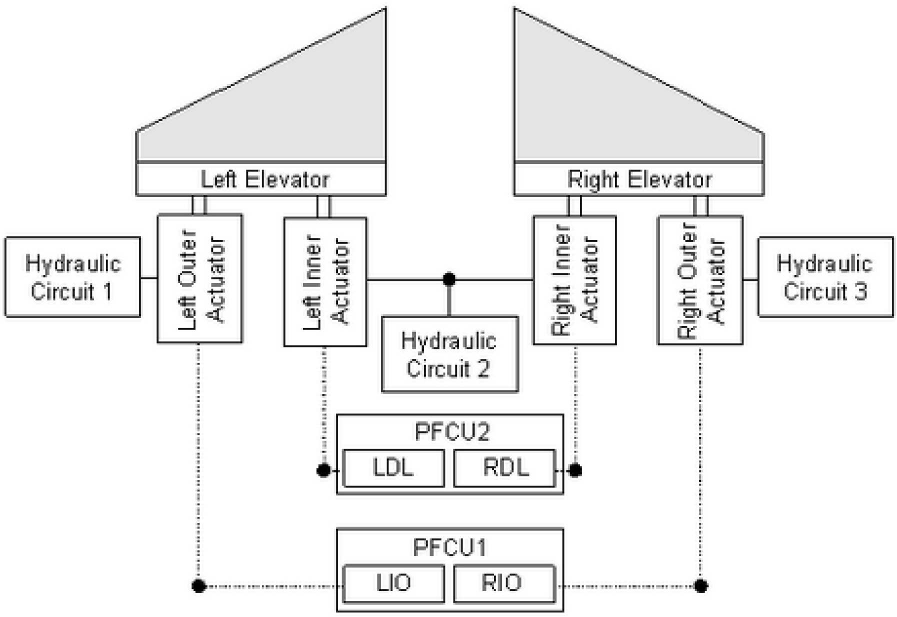


**Figure 7.** *The aircraft elevator control system[13]*