

# New Features for Fault Diagnosis by Supervised Classification

Sylvain Verron, Teodor Tiplica, Abdessamad Kobi

► **To cite this version:**

Sylvain Verron, Teodor Tiplica, Abdessamad Kobi. New Features for Fault Diagnosis by Supervised Classification. 18th Mediterranean Conference on Control and Automation (MED'10), 2010, Marrakech, Morocco. 2010. <inria-00517034>

**HAL Id: inria-00517034**

**<https://hal.inria.fr/inria-00517034>**

Submitted on 13 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# New Informative Features for Fault Diagnosis of Industrial Systems by Supervised Classification

Sylvain VERRON, Teodor TIPLICA, Abdessamad KOBİ

**Abstract**—The purpose of this article is to present a method for industrial process diagnosis. We are interested in fault diagnosis considered as a supervised classification task. The interest of the proposed method is to take into account new features (and so new informations) in the classifier. These new features are probabilities extracted from a Bayesian network comparing the faulty observations to the normal operating conditions. The performances of this method are evaluated on the data of a benchmark example: the Tennessee Eastman Process. Three kinds of fault are taken into account on this complex process. We show on this example that the addition of these new features allows to decrease the misclassification rate.

## I. INTRODUCTION

Nowadays, industrial processes are more and more complex. So, they include a lot of sensors giving measurements of some attributes of the system. A study of these measurements can allow to decide on the correct working conditions of the process. If the process is not in normal working conditions, it signifies that a fault has occurred in the process. If no fault has occurred, thus the process is in the fault-free case. An important research field is on the Fault Detection and Diagnosis (FDD) ([1]). The goal of a FDD scheme is to detect, the earliest possible, when a fault occurs in the process. Once the fault has been detected, the other important step is the diagnosis. The diagnosis can be seen as the decision of which fault has appeared in the process, what are the characteristics of this fault, what are the root causes of the fault.

One can distinguish three principal categories of methods for the FDD ([2]): the knowledge-based approach, the model-based approach and the data-driven approach. The knowledge-based category represents methods based on qualitative models (FMECA - Failures Modes Effects and Critically Analysis; Fault Trees; Decision Trees; Risk Analysis) ([3], [4]). For the model-based methods, an analytical model of the process is constructed based on the physical relations governing the process ([5]). The model gives the normal (fault free) value of each sensor or variable of the system for each sample instant, then residuals are generated (residuals are the differences between measurements and the corresponding reference values estimated with the model of the fault-free system). If the system is fault free, residuals are almost nil, and so their evaluations allow to detect and diagnose a fault. Theoretically, the best methods are the analytical ones, but the major drawback of this family of techniques is the fact that a detailed model of the process

is required in order to monitor it efficiently. Obtaining an effective detailed model can be very difficult, time consuming and expensive, particularly for large-scale systems with many variables. The last category of methods are the process history (or data-driven) methods ([6]). These techniques are based on rigorous statistical developments of process data. In literature, we can find many different data-driven techniques for FDD. For the fault detection of industrial processes many methods have been submitted: univariate statistical process control (Shewhart charts) ([7]), multivariate statistical process control ( $T^2$  and Q charts) ([8]), and some Principal Component Analysis (PCA) based techniques ([9]). [10] make comparisons between these different techniques. For the fault diagnosis techniques we can cite the book of [2] which presents a lot of them (PCA based techniques, Fisher Discriminant Analysis, PLS based techniques, etc).

The purpose of this article is to present a new contribution for the diagnosis of faults in industrial systems. Classically, when supervised classification is used for fault diagnosis, one has to select the most informative variables (features selection) in order to decrease the misclassification rate of new observations. Indeed, deleting the variables with no information concerning the separation of the different classes (faults) allows to increase the robustness of the classifier and increase the performance in classification. In this article, we propose an alternative solution which is to add some new features giving new informations for the classification. This new variables, and so new informations, are extracted from a comparison of the faults conditions to the normal operating conditions.

The article is structured in the following manner. In section II, we introduce the method to diagnose faults with supervised classifiers and present some classical supervised classifiers. The section III presents our new contribution allowing to compare the fault to the normal operating conditions and so to obtain new informations for the classification (diagnosis) of the fault. The section IV presents an evaluation of the proposed method for diagnosis of faults on the benchmark Tennessee Eastman Problem. Finally, we conclude on this method and present some perspectives.

## II. FAULT DIAGNOSIS BY SUPERVISED CLASSIFICATION

### A. Fault Diagnosis as classification task

Once a problem (fault) has been detected in the evolution of the process by the mean of a detection method, we need to identify (diagnosis) the belonging class of this fault. Thereby, the diagnosis problem can be viewed as the task to correctly

classify this fault in one of the predefined fault classes. The classification task needs the construction of a classifier (a function allocating a class to the observations described by the variables of the system). Two types of classification exist: unsupervised classification which objective is to identify the number and the composition of each class present in the data structure; supervised classification where the number of classes and the belonging class of each observation is known in a learning sample and whose objective is to class new observations to one of the existing classes. For example, given a learning sample of a bivariate system with three different known faults as illustrated in the figure 1, we can easily use supervised classification to classify a new faulty observation. A feature selection can be used in order to select only the most informative variables of the problem ([11]).

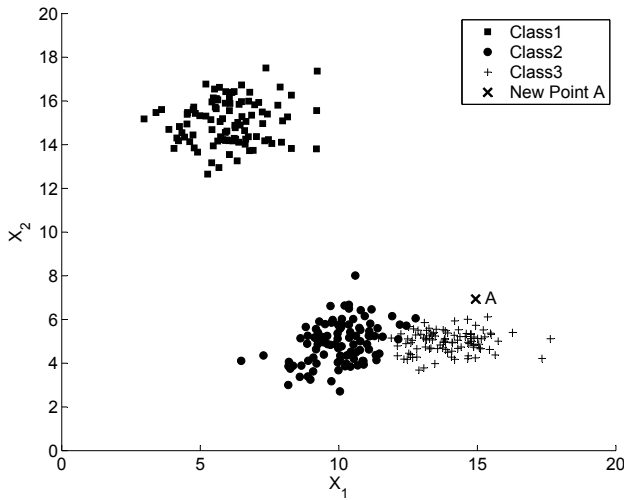


Fig. 1. Bivariate system with three different known faults

### B. Classical supervised classifiers

As we have seen how to diagnose a fault with supervised classification, we present here some classical classifiers. Indeed, there is not a classifier outperforming all other. It depends on the application. For a given application a classifier will give better results than an other one, but for an other application this will be the inverse, the first one will be less efficient.

1) *Linear Discriminant Analysis*: The Linear Discriminant Analysis (LDA) is a basic technique for supervised classification ([12]). This is a statistical classifier. Indeed, the classification rule is to assign a new observation to the group (class) with highest conditional probability. In order to deal with probability, a classical assumption is that the data follows Gaussian densities. In this case, LDA allocates to a new observation  $\mathbf{x}$  the class  $C_i$  with the minimal cost of the following equation:

$$\mathbf{x} \in C_i, \text{ if } i = \underset{i=1, \dots, k}{\operatorname{argmin}} \{K_i(\mathbf{x})\} \quad (1)$$

where  $K_i(\mathbf{x})$  is expressed by:

$$K_i(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - 2 \log(P(C_i)) \quad (2)$$

where  $\boldsymbol{\mu}_i$  is the mean vector of the fault  $F_i$ ,  $\boldsymbol{\Sigma}$  is the pooled variance-covariance matrix and  $P(C_i)$  is the a priori probability of the class  $C_i$ .

2) *The k-Nearest Neighbors algorithm*: The k-Nearest Neighbors (or kNN) algorithm ([13]) is a non-parametric (without parameters estimation) classification tool. The principle of this method is to observe the  $k$  nearest neighbors of a new observation in order to decide of his class. For a new observation to be classified, the algorithm computes the distance from this observation to each observation of a training set. Thus, given the distances, one can select the  $k$  nearest neighbors and their respective classes. Considering the classes of the  $k$  nearest neighbors, we conclude on the class of the new observation. In order to illustrate this rule, an example of classification (with 2 classes) in a 2 variables space is given on figure 2.

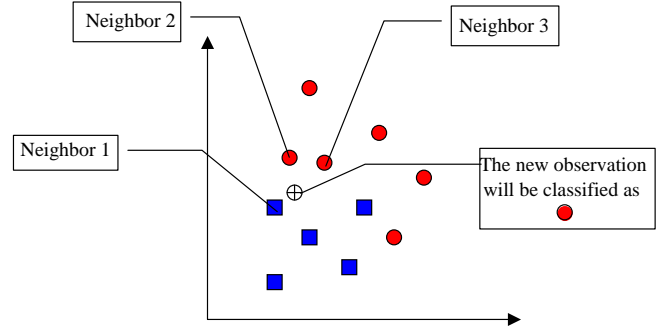


Fig. 2. Class attribution example with 3NN algorithm

3) *SVM*: Support Vector Machines (SVM) is a new machine-learning tool that is especially fit for classification, forecasting and estimation in small-sample cases ([14]). Originally, SVM is only applicable to linear problem, SVM search the linear separation maximizing the margin between two classes. But, with the kernel trick ([15]), it is possible to deal with non linear problems. For a better understanding, an example is given in figures 3 and 4. The first figure represents two classes non linearly separable. A transformation of the initial space ( $x$ ) to a bidimensionnal space ( $x; x^2$ ) allows to separate linearly and correctly (with the maximal margin) the two classes.

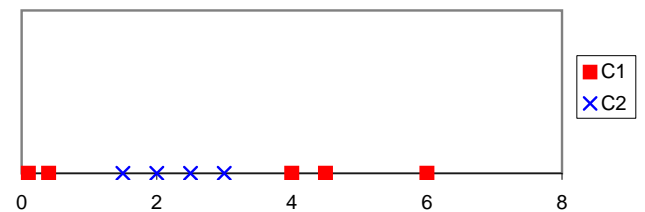


Fig. 3. SVM principle

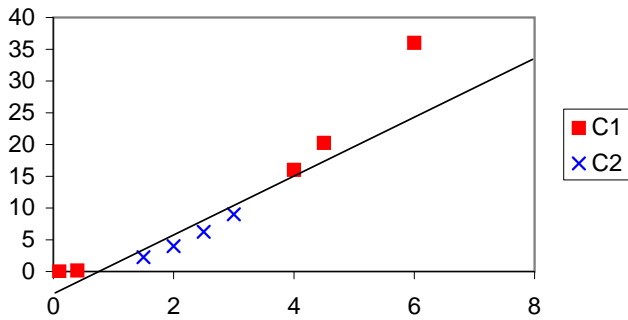


Fig. 4. SVM principle

4) *Bayesian Networks classifiers*: A Bayesian network (BN) is a graph ([16]). In this graph, each variable is a node that can be continuous or discrete. Edges of the graph represent dependence between linked nodes. Bayesian network classifiers are particular Bayesian networks ([17]). They always have a discrete node  $C$  (the class node) which codes the different classes. The other nodes represent the descriptors (variables) of the classification task. A particular type of BN classifier is the Naïve Bayesian Network (NBN). In a NBN, the class node is linked with all other variables of the system as indicated on the figure 5. This Bayesian classifier makes the strong assumption of non correlated variables is not realistic. In order to deal with correlated variables, several approaches have been developed like the Tree Augmented Naïve Bayesian networks (TAN) proposed by [17].

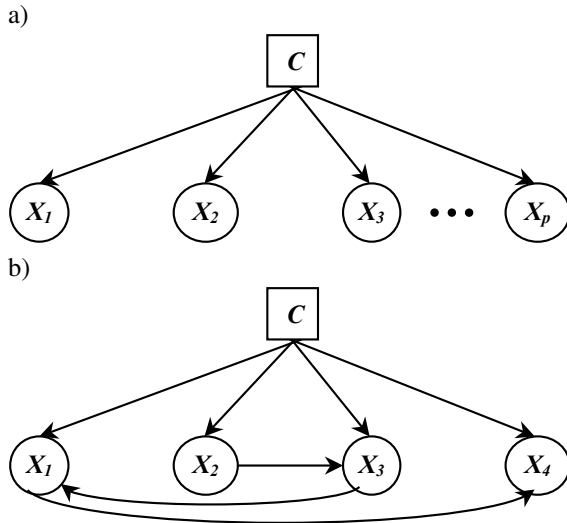


Fig. 5. Different Bayesian network classifiers: NBN (a) and TAN (b)

5) *Decision Tree*: Decision Tree (DT) is a hierarchical tree structure based on a series of questions (or rules) about the attributes of the class. The attributes of the classes can be of any type of variables (binary, nominal, ordinal, and quantitative values) while the classes must be qualitative type. Given a data of attributes together with its classes, a DT produces a sequence of rules (or series of questions) that

can be used to recognize the class ([18]). An example of a constructed DT is given on figure 6.

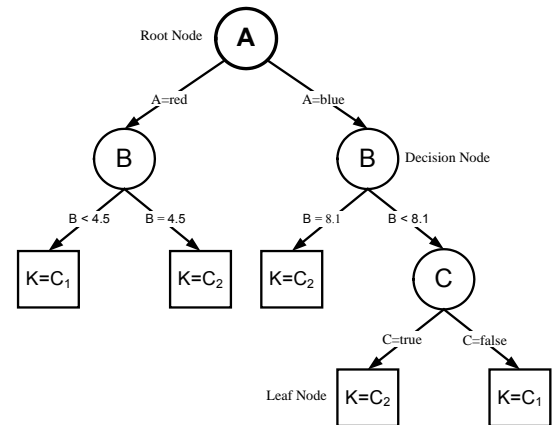


Fig. 6. A constructed Decision Tree

### III. NEW FEATURES FOR FAULT DIAGNOSIS

#### A. Obtaining new informations

Classically, misclassification in supervised classification is due to classes sharing a same place in the space, like in figure 7. On this figure, we can see that two faults (classes) have a similar space where the decision is not very clear. This is usually at this place that a wrong decision is taken.

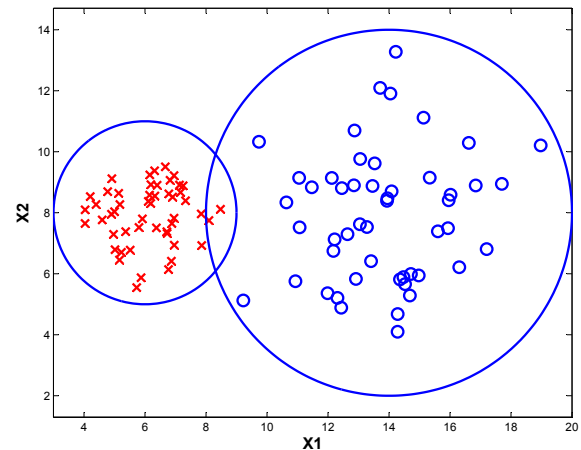


Fig. 7. Meeting place between 2 classes

In a classical supervised classification problem, there is no really solution to overcome this drawback. But, in the case of supervised classification applied to the fault diagnosis task, one can have a solution. Indeed, classification for fault diagnosis uses only data from previous observations of fault in order to define the classes. But, we generally have an other information: the data from the normal operating conditions (NOC). The idea of this paper is to think that if the information between the classes is not sufficient to obtain good classification, we can add to the model some informations of the difference between the fault and the

NOC. So, we have to search some new features (coming from the difference between the fault and the NOC) allowing to reduce (or to delete) the sharing place of figure 7, like represented on figure 8.

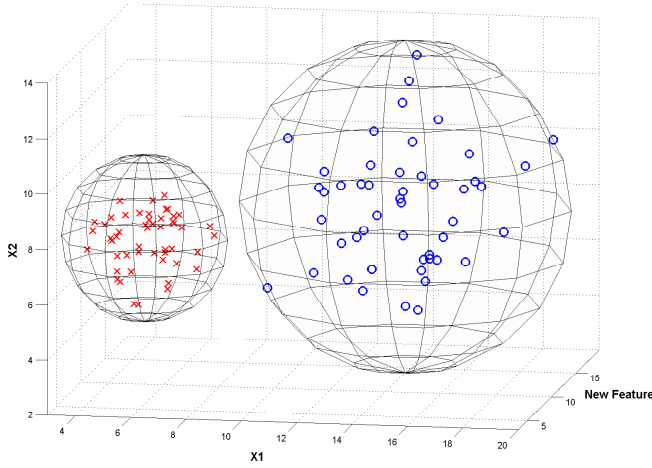


Fig. 8. Adding a new feature

### B. Computing the probabilities

In order to obtain new features, we exploit the data from the normal operating conditions. For that, we use a previous proposed method. Logically, this method has been designed in order to identify some variables responsible of a faulty observation. About the fault identification, one of the most interesting statistical technique is the MYT decomposition ([19]) which makes a decomposition of the  $T^2$  statistic in orthogonal components allowing to determine which variable or group of variables has contributed to the fault. Recently, Li et al. ([20]) have proposed an improvement to the MYT decomposition: the causal decomposition of the  $T^2$ . In order to make this decomposition, authors use a causal Bayesian network ([16]) representing the different variables of the process. We have proposed ([21]) an extension to this method, always based on a Bayesian network. In our method, a faulty (detected) observation (measurement of each variable at an instant) is given to the network, then it computes for each variable the probability to be responsible of the faulty situation. An example of such a Bayesian network is presented on figure 9.

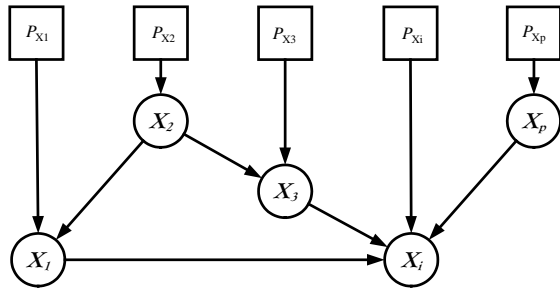


Fig. 9. Causal Bayesian network allowing to compute probabilities

On figure 9, we can see that each circle represents a variable of the system. The connection between two circles (two variables) denotes a regression of the first one to the second. Each square (one for each variable of the system) represents a bimodal variable coding the responsibility of his connected variable to a faulty situation.

### C. Proposed method

The idea of this article is to exploit the probabilities computed by the Bayesian network in order to integrate them as features for the classification task, and so to decrease the misclassification rate. For that, we make the assumption that we have some previous observed data from the process, for the known faults and for the normal operating conditions. From the free-fault data, the causal Bayesian network can be constructed. One can present to the network the data from the different faults. The network is then giving the different probabilities (one for each variables). So, we obtain  $2 \times p$  features: the original  $p$  features and the new  $p$  features (corresponding to the probabilities given by the Bayesian network). Finally, we can train a classifier (see section II-B) with the  $2 \times p$  features. Practically, given a new faulty observation, the  $p$  measurements are presented to the network in order to obtain the  $p$  probabilities. After that, the  $p$  measurements and the  $p$  probabilities are presented to the classifier which take the decision on the belonging class of the faulty observation. For a better understanding of the proposed method, we recall the principle on figure 10.

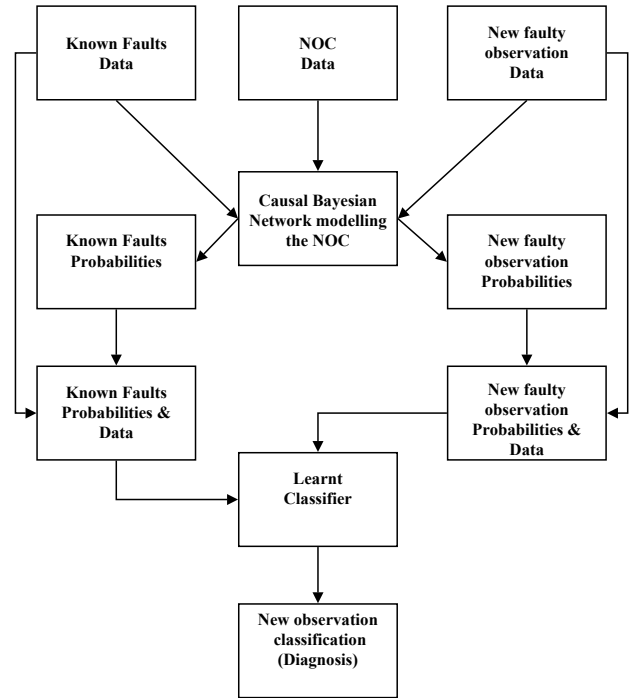


Fig. 10. The proposed method

Now, we will see an application of this approach on a benchmark problem: the Tennessee Eastman Process (figure 11).

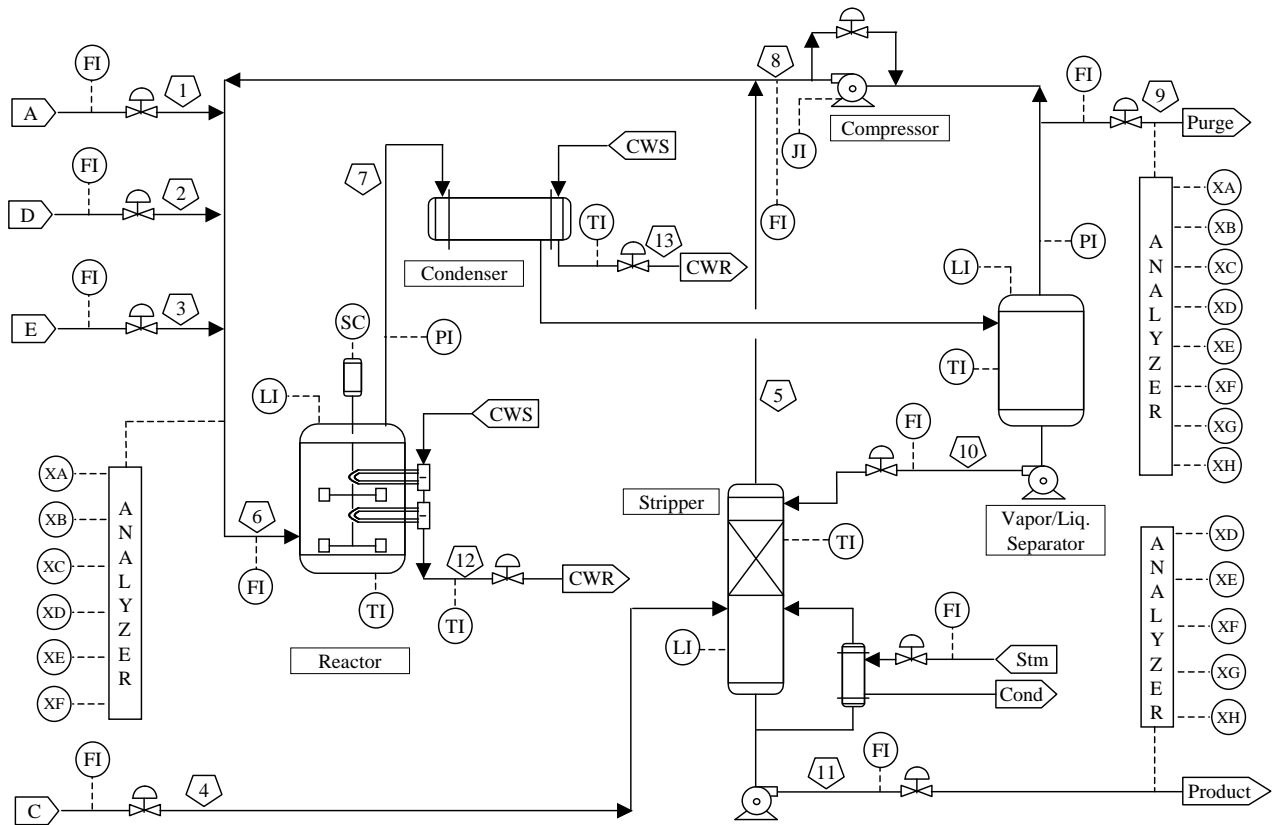


Fig. 11. Process flowsheet of the TEP

#### IV. APPLICATION TO THE TEP

##### A. Presentation of the TEP

The Tennessee Eastman Process (TEP) is a chemical process. It is not a real process but a simulation of a process that was created by the Eastman Chemical Company to provide a realistic industrial process in order to evaluate process control and monitoring methods. Article of [22] entirely describes this process. Authors also give the Fortran code of the simulation of the process. [23] has implemented the simulation on Matlab. The TEP is composed of five major operation units: a reactor, a condenser, a compressor, a stripper and a separator. Four gaseous reactant A, C, D, E and an inert B are fed to the reactor where the liquid products F, G and H are formed. This process has 12 input variables and 41 output variables. The TEP has 20 types of identified faults.

So, this process is ideal to test monitoring methods. But, it is also a benchmark problem for control techniques because it is open-loop unstable. Many articles present the TEP and test new approaches on it. For example, in fault detection, we can cite [10]. Some fault diagnosis techniques have also been tested on the TEP [2], [24] with the plant-wide control structure recommended in [25]. In [24], authors focus on only 3 types of fault and give the datasets they used. We have taken into account these 3 types of faults: fault 4, 9 and 11 (see table I). These three types of fault are good representations of overlapping data and so, are not easy to

classify. As indicated on the table I, each type of fault have 2 datasets: a training sample and a testing sample, containing respectively 480 and 800 observations.

| Class | Fault type  | Train data | Test data |
|-------|---|------------|-----------|
| 1     | Fault 4: step change in the reactor cooling water inlet temperature       | 480        | 800       |
| 2     | Fault 9: random variation in D feed temperature                           | 480        | 800       |
| 3     | Fault 11: random variation in the reactor cooling water inlet temperature | 480        | 800       |

TABLE I  
DESCRIPTION OF FAULT DATASETS

##### B. Results and comments

We have tested different configurations for this experiments. Some authors ([24] and [11]) proved that for these 3 faults (F4, F9 and F11), only two variables are very informative: the variable 9 and 51. For that, we reduce the problem to the classification of this 3 classes (faults) in the space of 52 variables to a space of 2 variables (variables 9 and 51).

In order to demonstrate the interest of the adding of new informative features (the probabilities) to the model, we

have computed, for each case, the misclassification rate in the 2 variables space ( $X_9$  and  $X_{51}$ ), in the 2 probabilities space ( $P_9$  and  $P_{51}$ ), and in the 2 variables - 2 probabilities space ( $X_9$ ,  $X_{51}$ ,  $P_9$  and  $P_{51}$ ). Moreover, to showing the general application of the proposed method, we have used different classifiers: Linear Discriminant Analysis (LDA), Decision Tree (DT), k-Nearest Neighbors with 1 (1-NN) or 3 neighbors (3-NN), Naive Bayesian network (NBN) and Tree Augmented Bayesian network (TAN), Support Vector Machines with radial basis function (SVM). The misclassification rate (in percents) are presented in table II: the less is the misclassification rate, the better is the space.

| Classifier | $X_9, X_{51}$ | $P_9, P_{51}$ | $X_9, X_{51}, P_9, P_{51}$ |
|------------|---------------|---------------|----------------------------|
| LDA        | 31.58         | 19.2          | 15.33                      |
| DT         | 9.5           | 12.62         | 8.87                       |
| 1-NN       | 12.87         | 18.87         | 10.7                       |
| 3-NN       | 11.37         | 15.46         | 8.79                       |
| NBN        | 11.5          | 16.83         | 9.29                       |
| TAN        | 10.12         | 16.83         | 8.87                       |
| SVM        | 7.75          | 16.16         | 7.66                       |

TABLE II  
MISCLASSIFICATION RESULTS

In table II, we can see that for each classifier we can make the same conclusions. Firstly, we can observe that for all classifiers (excepted the LDA), the misclassification is lower in the original space than in the probabilities space. But, as proposed in the method, for each classifier, the misclassification rate is lower in the combined space ( $X_9, X_{51}, P_9, P_{51}$ ) than in the two other spaces (the original and the probabilities ones). So, we can tell that the two spaces ( $X_9, X_{51}$  and  $P_9, P_{51}$ ) contain different informations, and so that their combination is more efficient than each one separately.

## V. CONCLUSION

The main interest of this article is the presentation of a new method for the fault diagnosis of industrial processes. We have considered the case of fault diagnosis as a supervised classification. But, we have proposed to obtain more information than only the fault database. This extra information is contain in the difference of the faults to the normal operating conditions. This difference is expressed by probabilities computed by a causal Bayesian network representing the normal working of the process. So, these probabilities are seen as new variables by the supervised classifier. The performances of this approach have been tested on a concrete example: the Tennessee Eastman Process. The results of the method are good and well demonstrate the advantage of this method.

This first work opens some interesting perspectives. Indeed, as we have worked on the causal Bayesian network giving the probabilities for the combined space, we have only used it. But, viewed the results, it will be very interesting to discover which information give the network. Maybe this information can be expressed more precisely than a probability (a statistical distance, a regression distance) and gives again better results.

## REFERENCES

- [1] R. Isermann, *Fault Diagnosis Systems An Introduction from Fault Detection to Fault Tolerance*. Springer, 2006.
- [2] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault detection and diagnosis in industrial systems*. New York: Springer-Verlag, 2001.
- [3] D. H. Stamatis, *Failure Mode and Effect Analysis: FMEA from Theory to Execution*, Hardcover, Ed. ASQ Quality Press, 2003.
- [4] B. Dhillon, *Reliability, Quality, and Safety for Engineers*. CRC Press, 2005.
- [5] R. J. Patton, P. M. Frank, and R. N. Clark, *Issues of Fault Diagnosis for Dynamic Systems*. Springer, 2000.
- [6] V. Venkatasubramanian, R. Rengaswamy, S. Kavuri, and K. Yin, "A review of process fault detection and diagnosis part iii: Process history based methods," *Computers and Chemical Engineering*, vol. 27, no. 3, pp. 327–346, 2003.
- [7] D. C. Montgomery, *Introduction to Statistical Quality Control, Third Edition*. John Wiley and Sons, 1997.
- [8] J. Westerhuis, S. Gurden, and A. Smilde, "Standardized q-statistic for improved sensitivity in the monitoring of residuals in mspc," *Journal of Chemometrics*, vol. 14, no. 4, pp. 335–349, 2000.
- [9] E. J. Jackson, "Multivariate quality control," *Communication Statistics - Theory and Methods*, vol. 14, pp. 2657 – 2688, 1985.
- [10] M. Kano, K. Nagao, S. Hasebe, I. Hashimoto, H. Ohno, R. Strauss, and B. Bakshi, "Comparison of multivariate statistical process monitoring methods with applications to the eastman challenge problem," *Computers and Chemical Engineering*, vol. 26, no. 2, pp. 161–174, 2002.
- [11] S. Verron, T. Tiplica, and A. Kobi, "Fault detection and identification with a new feature selection based on mutual information," *Journal of Process Control*, vol. 18, no. 5, pp. 479–490, June 2008.
- [12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification 2nd edition*. Wiley, 2001.
- [13] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, pp. 21–27, 1967.
- [14] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995.
- [15] A. Aizerman, E. Braverman, and L. Rozoner, "Theoretical foundations of the potential function method in pattern recognition learning," *Automation and Remote Control*, vol. 25, pp. 821–837, 1964.
- [16] F. V. Jensen, *An introduction to Bayesian Networks*. Taylor and Francis, London, United Kingdom, 1996.
- [17] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, vol. 29, no. 2-3, pp. 131–163, 1997.
- [18] L. Breiman, J. Friedman, C. J. Stone, and R. Olshen, *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software, 1993.
- [19] R. L. Mason, N. D. Tracy, and J. C. Young, "Decomposition of  $T^2$  for multivariate control chart interpretation," *Journal of Quality Technology*, vol. 27, no. 2, pp. 99–108, 1995.
- [20] J. Li, J. Jin, and J. Shi, "Causation-based  $t^2$  decomposition for multivariate process monitoring and diagnosis," *Journal of Quality Technology*, vol. 40, no. 1, pp. 46–58, 2008.
- [21] S. Verron, T. Tiplica, and A. Kobi, "Réseaux bayésiens pour l'identification de variables hors-contrôle," in *5me Conférence Internationale Francophone d'Automatique*, Bucarest, Roumanie, 2008.
- [22] J. Downs and E. Vogel, "Plant-wide industrial process control problem," *Computers and Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.
- [23] N. Ricker, "Decentralized control of the tennessee eastman challenge process," *Journal of Process Control*, vol. 6, no. 4, pp. 205–221, 1996.
- [24] L. Chiang, M. Kotanchek, and A. Kordon, "Fault diagnosis based on fisher discriminant analysis and support vector machines," *Computers and Chemical Engineering*, vol. 28, no. 8, pp. 1389–1401, 2004.
- [25] P. Lyman and C. Georgakis, "Plant-wide control of the tennessee eastman problem," *Computers and Chemical Engineering*, vol. 19, no. 3, pp. 321–331, 1995.