



Default diagnosis and prognosis for a preventive and predictive maintenance. Application to a distillation column

Alaa Daher

► To cite this version:

Alaa Daher. Default diagnosis and prognosis for a preventive and predictive maintenance. Application to a distillation column. Automatic. Normandie Université, 2018. English. NNT : 2018NORMR090 . tel-01992688

HAL Id: tel-01992688

<https://theses.hal.science/tel-01992688>

Submitted on 24 Jan 2019

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.

THESE

Pour obtenir le diplôme de doctorat

Spécialité Automatique et traitement du signal

Préparée au sein de l'Université de Rouen Normandie

Diagnostic et pronostic des défauts pour la maintenance préventive et prédictive. Application à une colonne de distillation

**Présentée et soutenue par
Alaa DAHER**

Thèse soutenue publiquement le (date de soutenance) devant le jury composé de		
M / Abdessamad KOBİ	Professeur à l'Université d'Angers	Rapporteur
M / Maan EL-BADAOUI	Professeur à l'Université de Lille	Rapporteur
Mme / Maya KALLAS	MdC à l'Université de Lorraine	Examinatrice
M / Dimitri LEFEBVRE	Professeur à l'Université du Havre	Examineur
M / Yahya CHETOUANI	MdC à l'IUT de Rouen	Encadrant
M / Mohamad KHALIL	Professeur à l'Université Libanaise	Co-directeur
M / Ghaleb HOBLOS	EC-HDR à l'ESIGELEC	Directeur

Thèse dirigée par Ghaleb HOBLOS, laboratoire IRSEEM

Et codirigée par Mohamad KHALIL, laboratoire CRSI



Normandie Université



Normandie Université

Acknowledgments

I owe all my gratitude to those who contributed to the completion of this thesis work and who transformed my doctoral studies into an exceptional experience.

*First of all, I want to express my sincere gratitude to Mr. **Ghaleb HOBLOS**, professor at NORMANDY UNIVERSITY, UNIROUEN, ESIGELEC, IRSEEM for leading this work. I also thank him for the trust he has placed in me, his competence, and his involvement and all the hours he has devoted to the supervision of this thesis. I wish to express to him my sincere gratitude for his availability. I also appreciated his human qualities and his sincere friendship.*

*I express my deep gratitude Mr. **Mohamad Khalil**, Professor at the Lebanese University and Director of the Azm Center for Biotechnology Research and its Applications for co-direction this work. I also thank him for his advice, his support and the trust he placed in me.*

*I wish to express my most sincere gratitude to Mr. **Yahya Chetouani**, Lecturer at the University of Rouen for having welcomed me in the chemical engineering laboratory and thus having trusted me and entrusted this subject of thesis. I thank you, Yahya, for your help, and your involvement in my work. I would like to thank the IUT of Rouen (Chemical Engineering Department) for allowing me to carry out the experimental work of the thesis.*

I would like to thanks European Union and Region Normandy for the co-funding of this work. Europe is involved in Normandy through the European Funds for Regional Development. Also we would like to thank the Lebanese university (AZM center for biotechnology research), for funding this research.

I thank the members of the jury for their interest in my work. I extend my gratitude to:

- Professor **Maan EL BADAoui EL NAJJAR** and Professor **Abdessamad KOBi** for having accepted to report this work and to have made valuable comments.



- Professor **Dimitri LEFEBVRE** and Dr. **Maya KALLAS** for having agreed to examine this thesis work.

I would also like to thank the entire team of the Doctoral School for Sciences and Technology in Tripoli, Lebanon for their help. I warmly thank my colleagues and my friends at the AZM center for their kindness and the good atmosphere they have brought.

*I wish to express my gratitude and appreciation to my mother **Rachideh** and my father **Mahmoud** who have always supported and advised me throughout my studies and in my everyday life. This thesis is yours. I thank you for your love, for your trust and all that you have offered me and who has made me what I am now. Dear my sisters **Hanan, Jihan, Fatmeh, and Hiba** and my brothers **Kamel and Mohamad**, I thank you for your love, your encouragement, your reassuring words and your support throughout these years.*

Thanks to you! !

I love you all! ! ! !

*And finally, there is a person of great importance in my life, who accompanies me and who supports me all the time by the strength and energy available to him. You are my wonderful wife, **Racha** whose contribution I cannot measure in the accomplishment of this thesis. I thank you very much for you are my love and being always beside me. And I hope that we always stay together and have a better life full of happiness, success and, above all, good health. I dedicate to you from the bottom of my heart this thesis and all my life. ! ! ! !*

Contents

General introduction.....	17
Chapter 1 Notions on the industrial distillation processes-data collection for faults diagnosis and prognostics	23
1.1 Introduction	25
1.2 Accidents in the Chemical Industry	27
1.3 Terminology	28
1.4 Fault Type.....	31
1.4.1 Fault Classification According to the Extension	32
1.4.2 Fault Classification According to the Temporal Evolution	34
1.5 Principle of Distillation.....	35
1.5.2 Applications of Distillation.....	36
1.5.3 Industrial Distillation	36
1.5.4 Distillation Column.....	37
1.6 Description of the Used Distillation System	38
1.6.1 Automated Continuous Distillation Column	38
1.6.2 Distillation Column Accidents.....	42
1.6.2.1 Data Collection for Faults Diagnosis and Prognostic	43
1.7 Determination of the Representation Space	54
1.7.1 Data Pre-processing	54
1.7.2 Data Reduction	56
1.7.2.1 Data Reduction Based on Feature Extraction	58
1.7.2.2 Data Reduction Based on Features Selection	61
1.8 Maintenance of Complex Systems	64
1.8.1 Introduction.....	64
1.8.2 Definition.....	64
1.8.3 New Evolutions of the Maintenance Function.....	65
1.8.4 Navigating towards an Anticipation of the Phenomena of Failure.	66
1.8.4.1 Corrective Maintenance	66

1.8.4.2 Preventive Maintenance.....	67
1.9 Conclusion.....	70
Chapter 2 Industrial Diagnosis Methods Applied on Chemical Processes (nonlinear systems)-Application to Distillation Column	72
2.1 Introduction	73
2.2 Diagnosis of Dynamic Systems	74
2.3 Characteristics of a Diagnostic System	74
2.4 The Different Diagnosis Steps.....	75
2.5. Classification of Diagnostic Methods	76
2.5.1 Model-based Methods.....	77
2.5.2 Model Free Methods.....	78
2.5.2.1 Qualitative Methods.....	78
2.5.3 Comparative Analysis of Diagnostic Methods on Chemical Processes	81
2.6. Artificial Neural Network and Industrial Diagnostics.	88
2.6.1.2 Artificial Neural Network.....	88
2.6.2 Neural Architectures	89
2.6.2.1 Non-looped Neural Networks "Feedforward"	89
2.6.2.2 Looped Neural Networks (Recurrent)	90
2.6.4 Most Used Neural Networks.....	90
2.6.4.1 Simple Perceptron.....	90
2.6.4.2 Multi-layer Perceptron.....	91
2.6.4.3 Retro Propagation Network	92
2.6.4.4 Radial Basis Function Neural Network (RBF)	93
2.7. Fuzzy Logic and Industrial Diagnostics.....	94
2.7.1 History of Fuzzy Logic:.....	94
2.7.2 The Requirements of Fuzzy Logic.....	95
2.7.3 Fuzzy Sets	95
2.7.3.1 Introduction.....	95
2.7.4 Membership Function	97
2.7.6 Fuzzy Inference System [156]	97
2.7.6.1 Fuzzification (Fuzzy Quantization)	99



2.7.6.4 Defuzzification.....	99
2.8 Faults Diagnosis-Application to a Distillation Column.....	101
2.8.1 Introduction.....	101
2.8.2 Distillation Column Faults Diagnosis Based on Fuzzy Logic	103
2.8.2.1 Input Status Words.....	103
2.8.2.2 Output Action Words.....	103
2.8.2.3 Defining of the Inference Rules.....	104
2.8.2.4 Fuzzification of Inputs and Outputs:.....	104
2.8.2.5 Defuzzification.....	109
2.8.2.6 Results.....	110
2.8.3 Distillation Column Faults Diagnosis Based on Artificial Neural Network.....	112
2.8.3.1 Features Extraction Applied to the Distillation Column Data	112
2.8.3.2 Discussion	113
2.8.4 Modified Fuzzy c-Means Combined with Neural Network Based Fault Diagnosis as New Approach for a Distillation Column	114
2.8.4.1 Fuzzy c-Means Clustering (FCM)	114
2.8.4.2 Feed Forward ANN	115
2.8.4.3 Proposed Methodology	116
2.8.4.5 Proposed MFCM Algorithm.....	118
2.8.4.6 Classification with Multi-layer Feed Forward ANN	118
2.8.4.7 Results and Discussion	120
2.9 Conclusion.....	124
Chapter 3 Industrial Prognostics Methods Applied on Chemical Processes (nonlinear systems)-Application to Distillation Column	126
3.1 Introduction to Industrial Prognostics	127
3.2 Emergence of the Prognostics in the Maintenance Activity.....	127
3.2.1 Predictive Maintenance and Prognostics	128
3.2.1.1 Diagnostic vs Prognostic	128
3.2.1.2 Concept of Prognostics	129
3.3. Contribution to the Formalization of the Prognostics and to the Development of an Application Framework	132

3.3.1 Formalization of the Prognostic Process.....	132
3.3.2. Prognostic Indicators	134
3.4.2.1. The Prognostic Measures	134
3.4.2.2. Performance Measures for the Prognostic Process [180]	135
3.4.2.3 Prognostics Modeling Methods for RUL Estimation	138
3.5. Prognostic Approaches	139
3.5.1. Classification of Prognostic Approaches	139
3.5.1.1. Emerging Classifications in the Literature.....	139
3.5.1.2. Discussion.....	142
3.5.2. Prognostic Based on the Physical Model.....	143
3.5.2.1. The principle of Operation, Advantages, and Disadvantages	143
3.5.2.2. Techniques and Tools	144
3.5.2.3. Applications	144
3.5.3. Data-Driven Prognostics.....	145
3.5.3.1. The principle of Operation, Advantages, and Disadvantages	145
3.5.3.2. Techniques and Tools	146
3.5.4. Prognostics Based on Experience	147
3.5.4.1. The principle of Operation, Advantages, and Disadvantages	147
3.5.4.2. Techniques and Tools.....	147
3.6. Towards Predictive Systems	148
3.6.1. Typology of Prediction Systems	148
3.6.2. Criteria for Choosing a Predictive System.....	148
3.6.3. Restrictive Industrial Characteristics of Prognostic Tools.....	150
3.7. Selection Guide for a Family and a Prognostic Tool	150
3.7.1. Data-driven Tools	151
3.7.2. Information-Oriented Tools	152
3.7.3. Knowledge-Oriented Tools.....	152
3.8. Investigation Track to Choice of a Prognostic Tool	153
3.9. Neuro-Fuzzy Systems for Time Series Prediction	155
3.9.1. Neural Networks and Prediction.....	155

3.9.2. Towards a Hybrid Neuro-fuzzy Prognostic System	157
3.9.3 ANFIS (Adaptive-network-based Fuzzy Inference System)	158
3.9.3.1 ANFIS a Hybrid Tool for Prognostic.....	159
3.13 Frame of ANFIS in the Prognostics Domain	161
3.13.1 Data Acquisition	161
3.13.2 Data Pre-processing	162
3.14 New Prognostic Approach for Preventive and Predictive Maintenance- Application to a Distillation Column	162
3.14.1 Introduction.....	162
3.14.2 Data Acquisition and Measurements	163
3.14.3 Time Series Calculation.....	165
3.14.4 Fuzzy C-Means Clustering (FCM) for Health Assessment	166
3.14.5 Proposed Methodology	167
3.14.6 Results and Discussions.....	173
3.15 Parzen Window Distributions as New Membership Function for ANFIS Algorithm- Application to a Distillation Column Faults Prediction.	181
3.15.1 Introduction.....	181
3.15.2 Proposing Parzen Windows as a New Membership Function for ANFIS Algorithm	182
3.15.2.1 Membership Function (MF).....	182
3.15.3 Proposed Methodology	186
3.15.4 Results and Discussion	190
3.15.5 Conclusion	197
Chapter 4 Soft Computing Fault Prediction and Diagnosis Algorithm- Application to a Distillation Column	199
4.1 Introduction	200
4.2 problem formulation	201
4.3 Data acquisition.....	202
4.3.1 Signal acquisition.....	202
4.3.2 Data Pre-processing	204
4.3.3 Signal processing	205
4.4 Materials and Methods.....	210

4.4.1 Adaptive Neuro-Fuzzy Inference System (ANFIS).....	210
4.4.2 ANFIS Accuracy.....	210
4.4.3 ANN Classifier	211
4.4.4 Time Series Creation	211
4.5 Proposed Methodology	213
4.6 Results and Discussions	217
4.7 Conclusion.....	221
General conclusion.....	223
References:	228

List of figures

Figure 1.1: Distribution of Accidents by Sectors.....	25
Figure 1.2: Various Physical System Faults Types.....	32
Figure 1.3: The Effect of Different Fault Types on the Sensor Measurements. Dotted Lines Indicate the Measured Values of the Sensor; However, the Continuous Line Represents the Actual Values.	32
Figure 1.4: Graphical Representation of the Most Common Faults Actuator (Ducard, 2009) [36]. Dotted Lines Designate the Desired Values of the Actuator; However, the Continuous Line Lines Represent the Current Values. (a) Oscillation, (b) Blocking, (c) Saturation and (d) Loss of Efficiency.....	33
Figure 1.5: Temporal Evolution of the Different Types of Faults	34
Figure 1.6: Faults Classification: Multiplicative and Additive.	35
Figure 1.7: Design of a Binary Distillation Column	37
Figure 1.8: General View of the Distillation Unit	39
Figure 1.9: Supervision System ETP 200	41
Figure 1.10: Graphical Representation of the Normal Mode	43
Figure 1.11: Graphical Representation of the Timer in Accident Mode.....	45
Figure 1.12: Graphical Representation of the Heating Power in Accident Mode	46
Figure 1.13: Graphical Representation of the Preheating Power in Accident Mode	47
Figure 1.14: Graphical Representation of the Feed Flow Rate in Accident Mode	48
Figure 1.15: Graphical Representation of the Progressive Accident on Reflux Rate.....	49
Figure 1.16: Graphical Representation of the Heating Power in Progressive Accident	50
Figure 1.17: Graphical Representation of the Progressive Accident on Preheating Power	51
Figure 1.18: Graphical Representation of the Progressive Accident on Feed Flow Rate	52

Figure 1.19.b: Graphical Representation of the Accelerometer Signal (Normal & Fault Mode).....	54
Figure: 1.20: Class Concept in Pattern Recognition.....	55
Figure 1.21: Schematic Diagram of the Diagnostic Procedure	56
Figure 1.22: Principle of Feature Extraction	58
Figure: 1.23: Feature Extraction Techniques.....	59
Figure: 1.24: Principle of Feature Selection.....	63
Figure 1.25: General Algorithm for Feature Selection.....	63
Figure 1.26: Maintenance Forms	66
Figure 1.27: CBM Architecture	70
Figure 2.1: Diagnosis Steps	75
Figure 2.2: General Classification of Diagnostic Methods.	77
Figure 2.3: Example of an Artificial Neural Network	89
Figure 2.4: Some Usual ANNs	90
Figure 2.5: General Scheme of Simple Perceptron.....	91
Figure 2.6: Example of Multi-layer Perceptron Network.....	92
Figure 2.7: Non-looped Neural Network	92
Figure 2.9: Comparison between Classical and Fuzzy Set.	96
Figure 2.10: Different Form of Membership Function.	97
Figure 2.11: Fuzzy Inference System	98
Figure 2.12: Defuzzification Using Center of Gravity Method.....	100
Figure 2.13: Defuzzification Using Mean of Maximums Method	100
Figure 2.14: Graphical representation of the signals (normal & fault Mode)	102
Figure 2.15: Fuzzification of Pre-heated Temperature (S1).....	105
Figure 2.16: Fuzzification of the Timer: Reflux Rate (S2).....	105
Figure 2.17: Fuzzification of Pre-heating Power (S3).....	106
Figure 2.18: Fuzzification of the Loss of Charge (S4)	106
Figure 2.19: Fuzzification of the Heating Power (S5).....	106
Figure 2.20: Fuzzification of the Feed Flow Rate (S6).....	107
Figure 2.21: Fuzzification of the Boiler Temperature (S7).....	107
Figure 2.22: Fuzzification of the Column Head Temperature-TIC2 (S8)	107
Figure 2.23: Fuzzification of the System Output	108
Figure 2.24: Fuzzy Logic Applied to Distillation Column Data	109
Figure 2.25: Rule Viewer of the Whole System	109
Figure 2.26: Outputs of Fuzzy Logic when it applied to the Distillation Column Data	110
Figure 2.27: Difference between Real and Desired Outputs.....	111
Figure 2.28: Percentage of the Observation that Correctly Classified in Terms of the Observations that Incorrectly Classified	111
Figure 2.29: General Diagnosis System Based on ANN.....	112

Figure 2.30: Graphical Representation of the Signals (Normal Mode).....	113
Figure 2.31: Block Diagram of Our Proposed Methodology	117
Figure 2.32: Result of Clustering in the Same Mode (No-Fault)	117
Figure 2.33: Distance between Two Clusters in Different Modes (Fault Mode)	117
Figure 2.34: Neural Network Proposed Architecture	119
Figure 2.35: ANN Training Performance	119
Figure 2.36: ANN Training State	120
Figure 2.37: Degree of Variation in Processed Signals for Normal and Eight Error Modes. 1-Preheated Temp 2-Preheated Power 3-Timer 4-Charge Loss 5-Heating Power 6-Feed Rate 7- Boiler Temp 8- TIC2.....	122
Figure 2.38: Error Boxplot- (Error between Real and Desired Output)	123
Figure 2.39: Boxplot Showing the Difference between the Median of the Degree of Variation	123
Figure 2.40 Percentage of the Observation That Correctly Classified In Terms of the Observations That Incorrectly Classified	124
Figure.3.1: Complementarity of Detection, Diagnosis and Prognostics Activities	129
Figure: 3.2 (a) Prognostics as RUL estimation, (b) State probabilities.....	132
Figure 3.3: The Concept of Prognostic (Dragomir et al., 2007)	133
Figure 3.4: Prognostic- Process for Prediction and Assessment (Dragomir Et al., 2007)	133
Figure: 3.5 prognostic measures: RUL and confidence	135
Figure 3.6: Accuracy and Precision	137
Figure 3.7: Timeliness from (Goebel et al., 2005).....	137
Figure 3.9: Multivariate Degradation Based Modeling Method.....	139
Figure 3.10: Taxonomy of Prognostic Approaches	140
Figure 3.11: Prognostic Based on the Model.....	145
Figure 3.12: The Criteria for Choosing a Prognostic Tool	149
Figure 3.13: Other Classification of Prognostic Techniques	151
Figure 3.14: Towards Hybrid Systems.....	156
Figure 3.15: ANFIS Architecture.....	160
Figure 3.17: Graphical Representation of Accelerometer Signal (Normal & Fault Mode)	165
Figure 3.18: Graphical Representation of Time Series Creation	166
Figure 3.19: Graphical Representation of System Degradation Mode: Case of Aging of the Pump Impeller	168
Figure 3.20: Training of ANFIS	169
Figure 3.21.a: Graphical Representation of ANFIS Training, Testing and Checking	170
Figure 3.21.b: Graphical Representation of the Error between Target and ANFIS Output.....	170
Figure 3.22: Designing of Degradation Curve	171
Figure 3.23: Fitting and Clustering of Degradation Model	171
Figure 3.24: Graphical Representation of Fitting Curves to 5 Real Observations Degraded Curves.....	172

Figure 3.25: Testing, Prediction, and Clustering of New Set of Data.....	172
Figure 3.26.b: Graphical Representation of Absolute Value Of Error Calculated In Each Iteration	174
Figure 3.26.c: Graphical Representation of Error Absolute Value Distribution.....	174
Figure 3.27.b. Graphical Representation of Absolute Value of Error (Calculated Each Iteration)	175
Figure 3.27.c. Graphical Representation of Error Absolute Value Distribution.....	175
Figure 3.28.a. Vibration Signal extracted from Degraded Impeller Pump	176
Figure 3.28.b. Calculation of System Degradation Curve Using ANFIS.....	176
Figure 3.29.a. Graphical Representation of Results-Using ANFIS as a Predictor of the Degradation Curve D3	177
Figure 3.29.b. Graphical Representation of Absolute Value of Error Calculated in Each Iteration.....	177
Figure 3.29.c. Graphical Representation of Error Distribution	177
Figure 3.30: Fitting of Degradation Curve (D4).....	178
Figure 3.31: RUL Calculation (D8)	179
Figure 3.32: Flow Chart of Our Proposed Methodology.....	187
Figure 3.33: ANFIS Architecture that was used in this Study	189
Figure 3.34.a: Graphical Representation of ANFIS Results Applied to Distillation Column Data (Reflux Rate Signal) with Different MFs of ANFIS	191
Figure 3.34.b: Graphical Representation of Residual Errors for Different MFs.....	191
Figure 3.34.c: Graphical Representation of Mean Errors for Different MFs	191
Figure 3.35: Graphical Representation of Consumption Time in Seconds for Different MFs (ANFIS Applied to Normal Reflux Rate Signal)	192
Figure 3.36.a: Graphical Representation of ANFIS Results Applied To Distillation Column Data (Heating Power Degraded Signal) with Different MFs of ANFIS	194
Figure 3.36.b: Graphical Representation of Residual Errors for Different MFs.....	194
Figure 3.36.c: Graphical Representation of Mean Errors for Different MFs	194
Figure 3.37: Graphical Representation of Consumption Time in Seconds for Different MFs (ANFIS Applied to Degraded Heating Power Signal)	195
Figure 4.1: From Fault Diagnosis to Fault Prognostic	201
Figure 4.2: From fault prognostic to fault diagnostic	202
Figure 4.3: Graphical Representation of Signals (Normal & Faults Mode).....	203
Figure 4.4: Graphical Representation of Time Series Creation.....	212
Figure 4.5: classification system steps.....	213
Figure 4.6: Cancellation of feature extraction step.....	214
Figure 4.7: Block Diagram of Proposed Methodology.....	215
Figure 4.8: Neural Network Proposed Architecture	217
Figure 4.9.a: Graphical Representation of Results Using ANFIS as Predictor of Distillation Column Data (Reflux Rate Signal)	218
Figure 4.9.b: Graphical Representation of Absolute Value of the Error Calculated in Each Iteration	218

Figure 4.9.c: Graphical Representation of Error Absolute Value Distributions	218
Figure 4.10: Graphical Representation of Absolute Value of Errors Calculated on ANFIS Outputs (Accident E1).....	219
Figure 4.11: Graphical Representation of Increasing Degraded Reflux Rate Accident (E1)	220
Figure 4.12: Error Boxplot- Error between Real and Desired Output	221

List of tables

Table 1.1: Statistical Characteristics of the Acquired Signals in Normal Mode.....	44
Table 2.1 Comparative Analysis of Synthetic Fault Detection and Diagnosis Approaches.....	85
Table 2.2: The Range of Each Output Mode	108
Table 2.3: Statistical Analysis of Distillation Column Data in Normal Mode	113
Table 3.1: Emerging Approaches and Prognostic Techniques in the Literature.....	141
Table 3.2: The Information Required Deploying Prognostic Approaches	142
Table 3.3: Prognostic Support Tools (Vasile et al., 2008)	154
Table 3.4: Advantages and Disadvantages of ANNs and FL.....	157
Table 3.5: Industrial and Scientific Benefits of Neuro-Fuzzy Networks.....	158
Table 3.6: Performance of ANFIS when it applied on distillation column signals	174
Table 3.7: Performance of ANN When it applied on Whole Signals Extracted from Distillation Column ...	175
Table 3.8: Statistical Characteristics of Eight Degradation Modes.....	178
Table 3.9: Normal and Degraded Signals.....	190
Table 3.10: Statistical Characteristics of MFs of Acquired Signals in Normal Mode	193
Table 3.11: Statistical Characteristics of MFs of Acquired Signals in Abnormal Mode	196
Table 4.1: The Acquired Signals	203
Table 4.2: Statistical Characteristics of Acquired Signals in Normal Mode	203
Table 4.3 Artificial Neural Network Output Results on Reflux Rate Accident E1	220
Table 4.4: The Result of Artificial Neural Network Output for All Accidents	221

Acronyms

Accelerated Life Model	ALM
Adaptive neuro-fuzzy inference system	ANFIS
Artificial intelligence	AI
Artificial neural network	ANN
Azote Fertilizers	AZF
Boiler Temperature	BT
Calculated fitting curve ,	CFC
Center of Gravity	GoG



Normandie Université

Column Head Temperature	TIC2
Condition monitoring	CM
Decision trees	DT
End-of-life	EOL
Expert systems .	ES
Failure threshold	FT
Fast Fourier Transform	FFT
Fault Detection and Diagnosis	PFD&D
Fault Tolerance Control	FTC
Feed Flow Rate	FR
Feedforward neural networks	FFNN
French Nomenclature of Activities	FNA
Fuzzy c-mean	FCM
Fuzzy inference system	FIS
Fuzzy logic	FL
Fuzzy system ,	FS
Heating Power	HP
Institute of Electrical & Electronics Engineering	IEEE
International Federation of Automatic Control	IFAC
National Institute for Industrial Environment and Risks in France	INERIS
Kernel density estimation	KDE
Knowledge-based expert system ,	KBES
Liquid-liquid	L-L
Loss of Charge	LC
Mean Absolute Percentage Error	MAPE
Mean of Maximum	MoM
Mean of RES	MRES
Mean Power Frequency :	MPF
Mean time of good operation between two failures	MTBF
Membership Function	MF
Modified FCM	MFCM
Multi-layers Perceptron	MLP
Neuro-fuzzy	NF
Output	O
Partial Least Squares	PLS
Peak-to-Average Power Ratio	PAPR
Proportional Hazard Model	PHM
Preheated Power	PP
Pre-heated Temperature	PT
Principal component analysis	PCA
Prognostics and Health Management	PHM
Proportional–integral–derivative controller	PID
Radial basis function	RBF



Normandie Université

Real degradation curve	RDC
Recursive least square estimator	RLSE
Remaining Useful Life ,	RUL
Residual	RES
Root Mean Square	RMS
Root mean square errors	RMSE
Short-Time Fourier Transform	STFT
Standard deviation of RES .	SRES
Takagi–Sugeno	TS
Target ,	T
Trend Description Language	TDL
Widrow Adaptive Network	WAN



Normandie Université

General introduction

Nowadays, industrial processes have increased in a large and complex way, whether they are production units (petroleum refineries, nuclear power stations), means of transport (planes, trains), or others ... They often result from the interconnection of several subsystems. The study of these processes has received renewed interest from the scientific community in recent years, particularly with regard to the development of diagnostic and prognostic strategies, for several reasons: all of Firstly, the industrial competitiveness induced by the requirements of the economic market encourages the reduction of production costs. Secondly, the reliability and availability requirements of the systems have become indispensable for increasing the productivity of an industrial process. Finally, fault tolerance control (FTC) has been particularly developed in recent years, producing results applicable to the regulation of systems. In this diagnostic and prognostic context, manufacturers must meet binding specifications: satisfaction of quality and, safety constraints and compliance with environmental standards (noise, pollution, etc.). To meet these specifications, it is essential to ensure that industrial processes operate safely in relation to their human and material environment. It is also important to optimize production to reduce costs. However, faults can occur on these systems, whether on the instrumentation (sensors and actuators) or in the system itself. This can cause a degradation of the system performance (product quality, stability, availability etc.), which could defeat the mission of the system in question. This problem has a direct impact on the objectives stated above. Indeed, the determination of the operating conditions at each moment is necessary in order to know the moment when the system deviates from its normal behavior and to take the necessary measures so that there is no interruption of service. This is the purpose of the diagnosis of the system which provides fault detection and diagnosis. This concern must be ensured throughout the lifetime of a process. Despite the importance of fault diagnosis, but the most important issue, which is the focused work in the last years of industrial research, is the subject of prognostics or the prediction of the future status of the system.

In general, industrial processes are intended to produce objects, to synthesize components, or to perform services, which make it possible to obtain in large quantities products that would otherwise be relatively difficult or expensive to obtain. They help address important social challenges such as climate change, energy shortages, security issues, and health. Uncertainty, unpredictability, the dynamics of phenomena and a large number of data to be processed make the above tasks very complex and the decisions in real time are very difficult. In addition, failures of industrial processes have become difficult to detect and diagnose. To cope with these difficulties, it is essential to provide these processes with greater reliability and reliability. The chemical industry is a major player in the economy in many countries. This sector provides essential products (medicines, fertilizers, plastic,...). Despite its contribution to the rising standard of living, it faces a dangerous and polluting industry image because of the systems and raw materials it uses. Thus, the slightest anomaly in a chemical system can spread throughout the installation and thus causes serious consequences. Since the Second World War, the rapid growth and expansion of the chemical industry have led to serious accidents for the natural and industrial environment. On the night of December 3, 1984, in Bhopal (central India) (Stellman and Dufresne, 2000), the explosion of a plant of a subsidiary of the American firm Union Carbide producing pesticides, cleared 40 tons of methyl isocyanate in the atmosphere of the city. This accident killed 2,500 people and left 200,000 injured. The cause of the accident is a chemical reaction of the product with water remaining in a tank after cleaning.

Bhopal is probably the worst industrial chemical disaster of all time. On September 21, 2001 (Souriau et al., 2002), an explosion largely destroyed the chemical complex AZF (Azote Fertilizers) belonging to the Grande Paroisse company (Total group), the leading French fertilizer manufacturer. Thirty people were killed and 2,500 injured and 27,000 homes damaged. The explosion occurred in a stockpile of about 300 tons of ammonium nitrate pellets, digging an oval-shaped crater 70 meters long and 40 meters wide, and 5 to 6 meters deep. The detonation was heard more than 80 km from Toulouse. An earthquake of magnitude 3.4 has been recorded. The most likely cause of this disaster was the mixing of incompatible products in the hangar, which contained manufacturing waste and where a dump had just been emptied thirty minutes before the explosion, where the explosion occurred. In addition, the yearly costs of these accidents are beyond billions of dollars [1]. In view of these and other disasters, it is essential to ensure the normal operation of industrial installations. Indeed, the detection of a dysfunction at the beginning of its appearance, Pre-knowledge and the predicting of the future state of the system can avoid serious consequences. In fact, the efficiency of the maintenance of industrial systems is a major economic stake for their commercial exploitation. The main difficulties and sources of inefficiency lie in the choice of maintenance actions. A maintenance action is to replace equipment in the system that is out of order and is no longer capable of performing its function. Maintenance operations are expensive for several reasons.

In this case, throughout the maintenance phase, the system is not operational. The longer the maintenance phase, the more expensive it is due to the unavailability of the system. Therefore the maintenance phase should ideally be reduced to the operations of replacing, without trial and error, equipment actually down. The decision of a maintenance action is very complex and must be based on intelligent monitoring and analysis of the state of the system. A fault diagnosis is then necessary to determine as precisely as possible the equipment that needs to be repaired. When the diagnosis is less ambiguous, more maintenance operations are effective. The second reason that maintenance can be costly for cases of emergency where the safety or performance of the function of the system is involved. In fact, when equipment suddenly breaks down and the system can no longer perform its function, maintenance actions must be automatically performed to bring the system back into working order. These unforeseen actions are naturally more expensive because the needs and services for maintenance have not been anticipated and must be quickly available. To minimize the occurrence of this type of situation, preventive maintenance may be considered. Equipment failures can be anticipated and corrected before generating too much damage that could cause an unexpected stopping of the system. Most often, preventive maintenance relies solely on reliability analyzes that do not take into account the demands that actually influence the system's equipment throughout its operation. Indeed, abnormal or unforeseen stresses can accelerate equipment degradation. Preventive maintenance can be improved by a prognostic of reasoning to estimate the impact of these demands on the lifetime of the equipment. By establishing any maintenance action is appropriate at a given time, the prognostic also helps to plan future maintenance phases. Although the plants are equipped with automatic systems, computer simulation and analysis is still limited to maintain process plant integrity and extremely relied on human operators. For instance, humans can't detect hidden faults or predict future problems. Previous Industrial statistics have shown that major catastrophes may be infrequent, and minor accidents are very common.

Based on all of the above and in order to put the work in its proper course, the main objective in this thesis is centered around an effective health monitoring technique must be adapted to determine the state of the system at all times. A diagnostic method determines the current state of the system and identifies the probable causes (interaction with the environment, faults, etc.) that can lead to this state by reasoning on the observations. A prognostic method uses the current mission plan and knowledge of system degradation to anticipate abnormal behaviors or faults and thus predict future states. Prognostic is usually associated with the term end-of-life (EOL) prediction of an in-service system when the system is no longer operational, or its Remaining Useful Life (RUL), that means the remaining time until the end of life. The diagnosis and the prognostic thus make it possible to have a report on the current health of the system as well as a prediction of the evolution of its state in the future. This information is used to reconfigure the system and update the mission and maintenance plans, hence the importance of diagnostic and prognostic in the strategy of the maintenance. This thesis focuses on optimizing the maintenance of complex industrial systems. It proposes to set up a supervision architecture that integrates diagnostic and prognostic capabilities with the aim of helping to make decisions on maintenance actions. The complex systems that are considered in this thesis are composed of totally heterogeneous equipment (hardware, software) and require several types of techniques to be monitored. This thesis presents an abstract and homogeneous description of a complex system from which it is possible to characterize an original coupling of diagnostic and prognostic problems. The main application of this thesis work concerns the monitoring (diagnosis and prognostics) of a distillation column. Such a system consists of separating the constituents of a liquid mixture. Distillation systems are nowadays used on a large scale, especially in the chemical industry. Due to their growing presence in increasingly diverse application areas, the security issues associated with this type of system are becoming more prevalent in operational constraints. The slightest failure on the distillation unit can limit the performance of the products and have serious consequences if it is not detected quickly. It is particularly relevant to be able to monitor these systems in real time so as to ensure the safety of goods and people in direct or indirect relation with the application while achieving the operational objectives set. Based on the literature, diagnostic methods differ according to different criteria: the dynamic process, complexity, online diagnostic implementing, the nature of information, depth, distribution... in this context, several classifications are also proposed in the literature. These classifications are influenced by terminologies and specific contexts of each community and are not always consistent.

The researchers that are applied to the monitoring of chemical engineering distillation system are not strongly addressed and the numbers of articles related to this topic are very limited. Researchers classified diagnostic methods into two large families: methods based on mathematical model and model-free methods [2]. Many researchers said that the use of particular techniques model free is more effective than the methods with a model, especially for the diagnosis and prognostic of faults in real time on a distillation process [3]. Concerning the prognostics part, in the previous works (Ciarapica and Giacchetta in 2006) and (Jardine et al., 2006), have proved that the prognostic topic represents a work main frame that ensures the safety for industrial environment, and is considered as a key process in maintenance strategies. According to ISO 13381-1 [ISO, 2004], the prognostic of faults corresponds to the estimation time of operation before failure and the risk of the existence or subsequent appearance of one or more modes of failure [4]. This duration of operation before failure is commonly called Remaining Useful Life (RUL) (Khelif et al., 2014)

[4]. In literature, prognostic methods are classified into two large families: data-driven methods and statistical techniques (Vasile-Dragomir, 2008) [5]. Usually, real systems are noisy, complex, nonlinear, and nonstationary, that's why, building a mathematical model for prognostic of a real industrial system is so difficult [3], Therefore, it is important to use methods that depend on algorithms without a model. Previously, hybrid systems have been shown as important and efficient algorithms in forecasting and prediction domain, therefore, in our research we rely on a hybrid neuro-fuzzy system called adaptive neuro-fuzzy inference system (ANFIS) that combine fuzzy logic (FL) and artificial neural network (ANN) in the same algorithm and them are dramatically used as prediction and forecasting approaches [6] [7] [8] [9] [10]. The most important factor that affects the prediction accuracy of ANFIS is the type of Membership Function (MF) used on the first layer of ANFIS architecture. The execution time is also important for real-time processing [11]. Referring to a comparative study in 2016, Ardhan et al. [11] suggested that the trapezoidal shape is the best MF that can be used for load forecasting. As well in 2011, Mayilvaganan et al. [12] proved that the Gaussian shape showed significant results for the prediction of the groundwater level of a watershed. Also in 2012, Singh et al. demonstrated that the Gaussian and the bell-shape are the best MFs for the estimation of the elastic constant of rocks [13]. In the absence of any study applied to a data extracted from the distillation column to find the best forecasting technique, it is, therefore, necessary, to do a comparative study between different types of membership functions (MFs) in order to conclude upon the best MF that can be used for forecasting the distillation process data. A thoughtful consideration of the way to find a new MF with a low number of parameters is the first incentive to propose the Parzen windows distribution as a new membership function to be used on the first layer of ANFIS algorithm. As it is known, only the standard deviation (h) can modify the Parzen shape form, and in this case, we have only one parameter that should be updated for each iteration. In a part of this thesis, we try to find the best type of MF, which has the smallest root mean square errors (RMSE) among actual data and forecasting data, when considering the execution time.

This thesis is divided into four chapters, whose content is organized as follows:

The first chapter presents an overview of the general industrial accidents then we will go towards the accidents that occur in the chemical industry especially in chemical reactors, then we will focus on the all type of distillation process with depth explanation of the automated continuous distillation process then we develop the different types of faults that may occur in distillation process then we will move to the data collected from the side of diagnosis and prognostic by clarifying the characteristic of the acquired signals. This chapter is ended by a pre-processing of the data used in this study followed by a state of art for data reduction include features extraction and features selection and a framework of the maintenance strategy of complex systems.

The second chapter presents the qualitative and quantitative methods most widely used in the literature for the diagnosis of chemical processes and the different methods of risk analysis. This state of the art is the result of a bibliographic search of about a hundred well-studied articles on the evolution of the monitoring of the chemical processes and the different methods developed. A synthesis on methods of monitoring chemical reactors without models is developed in this section. Then we move to a state of the art that

presents the realized work of detection and diagnosis of faults that may occur in the distillation column and clarify the weaknesses of the previous methods. The results of this biography aim to select the neuro-fuzzy as a better technique for the diagnosis of faults that occur in a distillation column. This chapter is finished by an application of fuzzy logic and ANN separately on the data extracted from the distillation column. Depending on the obtained results we decided to propose an approach more efficient in real-time analysis of distillation column system. It proposes a methodology that combines fuzzy c-mean (FCM) clustering and neural network for diagnosis, detection, and classification of many faults. Moreover, a modified FCM method (MFCM) is presented in place of a feature extraction and selection approach. MFCM is a clustering method that allows calculating the degree of variation between normal and abnormal modes. The output of the MFCM is considered as inputs for the neural network classifier. This proposed methodology is then tested via a real experimental data obtained from a distillation column, after a pre-processing step including filtering and smoothing of the signals. A database with normal and faulty observations is analyzed. The database is composed of eight different types of faults that may occur during the automated distillation process in the chemical industry. The results of the proposed method confirm the ability to classify between normal and eight abnormal classes of faults.

The third chapter is devoted to the positioning of prognostic activity in the context of industrial maintenance and to the study of potentially useful tools to support this process and gives an overview of prognostic methods. Then we will move to an overview of the classification of prognostics approaches applied to the chemical reactors and the distillation column will be presented including their pros and cons. Different RUL estimation strategies are also reviewed. The investigation track in this chapter is aim to choose the adaptive neuro-fuzzy inference system (ANFIS) approach as the best technique. ANFIS is able to calculate the RUL of the distillation column degradation (guide to choosing a prognostic tool). Then, this chapter will position our work in relation to the literature of prognostic. This chapter is ended by the developing of a new prognostic strategy, applied to a real experimental data acquired from distillation column and from a metric pump. This methodology is a new technique effective in determining the path of deterioration of the distillation column system and also predicts the future path (prognostic) of this system by determining their RUL. Also, this work presents a direct monitoring approach based on the technique of adaptive neuro-fuzzy inference system (ANFIS) combined with fuzzy C-means algorithm (FCM). The results of a comparative study between the results of our proposed methodology and other based on ANN are discussed at the end of this application. The Results demonstrate the validity of the proposed technique to achieve the needed objectives with a high-level accuracy, especially the ability to determine a more accurate Remaining Useful Life (RUL) when it applied on the automated distillation process in the chemical industry. To improve the performance of ANFIS algorithm, Parzen windows distribution is proposed as a new membership function for ANFIS algorithm. The aims of this proposing are to reducing the consumption time and make the processing closer to a real-time application or minimizing the root means square error (RMSE) between the real and predictive data. The methodology is tested on real experimental data obtained from a distillation column aiming to predict the failure that may occur during the automated continuous distillation process. A comparative study was needed to choose the better membership function can be used for ANFIS algorithm when ANFIS applied to distillation column data. The results obtained in this research demonstrated that Parzen window proved its worth as a new

membership function of ANFIS algorithm when it is applied to the distillation column data and it also proved to be successful in reducing the execution time of ANFIS. The results have also shown that Parzen MF is chosen as the best MF for three over eight types of normal signals and for five over eight degraded signals.

The last chapter, chapter 4 proposes a new methodology that works for fault prognostics and diagnosis at the same time as a full scanning system can be applied to distillation column faults. The Adaptive Neuro-Fuzzy Inference System (ANFIS), as a hybrid system, has been selected for the step related to prediction since it combines the advantages of fuzzy logic and ANNs in one simultaneous algorithm. In our research, we tested this methodology with real experimental data that was obtained from a real distillation column. This resulted in the analysis of a database with different types of faults that could potentially occur during the automated distillation process. The results that were observed proved the validity and strength of this proposed technique. It was also demonstrated that the technique achieved with a high level of accuracy, the objective of prediction and diagnosis especially when applied to the data obtained from automated distillation process in the chemical industry.

Thesis Contributions

In this section, we briefly summarize the main contributions of this thesis. We classify them according to the field to which they naturally belong. In the field of faults diagnosis and prognostics, the main contributions are:

- We propose a new methodology that combines a modified fuzzy c-mean (MFCM) clustering technics and artificial neural network in one algorithm for detection, diagnosis, and classification of faults for real-time analysis.
- We develop of a new fault prognostic strategy based on the technique of Adaptive Neuro-Fuzzy Inference System (ANFIS) combined with Fuzzy C-means algorithm (FCM). This methodology is a new technique used to determining the deterioration path of a system then predicts the future path (prognostic) of this system by determining their Remaining Useful Life (RUL) as a direct and real time monitoring approach.
- We propose Parzen windows distribution as a new membership function to improve the performance of ANFIS algorithm. This new methodology aims to reduce the consumption time and make the processing closer to a real-time application and minimizing the root means square error (RMSE) between the real and predictive data for more and more accurate prognosis.
- We propose a new methodology used for fault prognostics and diagnosis at the same time as a new maintenance strategy. The Adaptive Neuro-Fuzzy Inference System (ANFIS), as a hybrid system, has been selected for the step related to the data prediction where this step followed by the classification of faults based on ANN as a fault diagnosis step.

We examined the applicability of our proposals for fault detection, diagnosis, and prognostics in *real data extracted from the distillation column system*.



Normandie Université

Chapter 1

Notions on the industrial distillation processes-data collection for faults diagnosis and prognostics

1.1 Introduction

The chemical industry is one of the pillars of the global economy, but in recent years it has faced an unflattering image of a dangerous and polluting industry. The history of hazardous chemical accidents shows that this industry remains one of the major sources of serious incidents that are relatively more likely to happen than previously thought [14].

Eight sectors of activity are distinguished in the accident sample studied with more than 6 recorded cases (Figure 1.1). Chemistry is leading with more than 54% of the accidents recorded (150 cases). This result can be explained by the high rate of automation of chemical processes, their diversity, and the high number of production sites in France, which is also the most widely represented in the ARIA database (12% of accidents of installations classified between 1992 and 2012). Many installations are versatile to produce a variety of products, which encourages the occurrence of unforeseen accidental situations in the design of the automated system or as a result of erroneous driving decisions. In this sector, a better integration of the treatment function would have prevented or reduced the probability of an accident in 21% of cases; slightly higher than the average rate of all sectors (16%) [15].

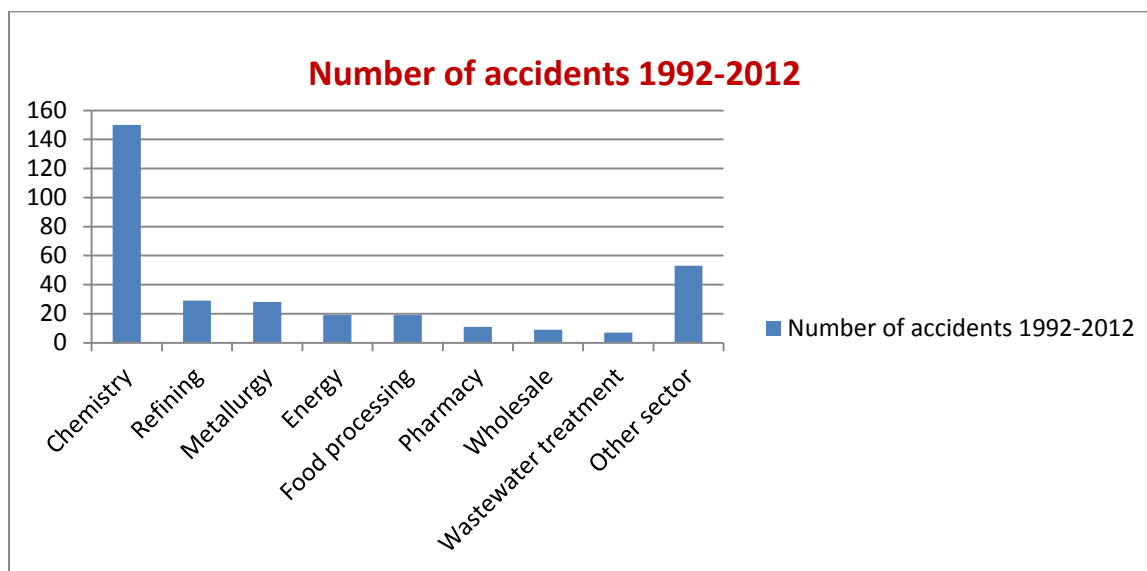


Figure 1.1: Distribution of Accidents by Sectors

Some recent examples of such accidents in the world that have caused significant loss of life and major damage to property and the environment include The explosion of the four reactors at the Fukushima-Daiichi nuclear power plant in North- East Japan in 2011, the destruction of the oxidation unit of the plant Flixborough FP Less. (1996) [16], the toxic cloud of dioxin dispersed in the commune of Seveso, and the 4000 victims and 200,000 injured in the Bhopal disaster or the explosion of the AZF plant. Gupta. (2005), [17]. Other major accidents can be cited, such as those in Mississauga, Ontario, St-Basile-le-Grand in

Quebec, Chernobyl in Ukraine, San-doz in Basel, Feyzin Rafferty in Lyon, and silos of Metz and Blaye, etc.

Therefore, many images will remain engraved in the memory of those who have lived it, and also in the memory of all the actors of security who are still trying to learn the lessons and avoid the mistakes of the past J.L. Gustin. (2002) [18] T.A. Kletz (2006) [19].

Many industrial accidents have been caused by reactions whose implementation has not been controlled: thermal runaway, uncontrolled secondary reaction ... These consequences are often important as much as an unwanted chemical reaction is liable, following the reactants employed, to give material both to an explosion and to the emission of toxic or flammable products into the environment [14].

The main causes and reasons for the development of uncontrolled reactions in the industry are due to lack of design or operation of the plants and manufacturing processes [20]. These accidents have demonstrated the extent to which the incidents caused by this type of industry can be significant and destructive. Formerly facing these events, we have become obliged to rely on chemical risk by questioning industrial practices and identifying new technological issues (diagnosis, dependability, supervision, prognostic ...) [21].

This development of the security aspects coincides with a change in the social context and the emergence of new themes such as ecology and environmental protection, which are now embedded in sustainable development [22][23]. From this period, this public awareness leads to a lower tolerance of the impact of industries on man and his environment (rejection of effluents, waste, noise pollution, etc.) [24].

The phenomena involved in the chemical process must be characterized in such a way as to ensure that it is the most suitable, both for the design of the installation and for the choice of the operating mode. In a chemical process, the non-linearity and sensitivity of certain parameters influence the course of the chemical reaction. A small variation in one of these parameters can dramatically change the course of the reaction and may even lead to an accident [14].

In the chemical industry, the consequences of accidents are listed in two types of effects:

- Thermal effects that correspond to the combustion of a flammable product or to an explosion.
- Mechanical effects that correspond to an overpressure caused by a shock wave (deflagration or detonation) from an explosion. The latter may result from an explosive, a violent chemical reaction, a sudden decompression of a pressurized gas, or a violent combustion or the ignition of a cloud of combustible dust. Specialists estimate the overpressure generated by the explosion in order to predict the associated effects on the bodies of living beings [25].

One of the important branches of process engineering is the chemical reaction field, which focuses on methods for the rational implementation of chemical transformations and particularly on the apparatus in which reactions are conducted: chemical reactors. Although the reactor represents only a modest share of the investment in an industrial process, its operation largely determines the upstream facilities (preparation

of reagent loads, choice of temperature and pressure conditions) and installations (products separating devices in particular). An improvement in the efficiency of the reactor by a few units can, therefore, result in a significant reduction in investment costs and consumption of materials and energy. In this sense, it can be said that the reactor is truly the heart of the process; thus, it requires the attention of the engineers.

In the case of a chemical reactor, the main factor which can lead to its monitoring is the assistance in the management of a process which provides the operator with the necessary tools to make decisions about actions to optimize its work (maximum production, safety, and non-degradation of equipment) [26]. This driving aid requires monitoring of the process in order to detect all the anomalies of the operation and to identify the causes as well as possible. Maintenance, which is aimed at the replacement or repair of worn or defective equipment, is carried out mostly off-line. It is important to note that while conduct and maintenance are operations that take time in different ways, the monitoring they involve must be in line for both purposes. In this light, the advantage of being able to determine in real time the occurrence of a malfunction during the implementation of a chemical process is justified in order to be able to effectively remedy the problem of faults detection and diagnosis of chemical faults.

On the other hand, the role of computer simulation and analysis is still restricted in the ability to preserve process plant safety and highly dependent on human operators. Humans could not be able to discover the hidden faults or predict future failures. Industrial statistics showed that the major disasters may be rare, but the minor accidents are very frequent with an annual cost that exceed billions of dollars [1][3]. Therefore, it's necessary for the industrialists to surround the severity status of the fault and to predict the ideal moment to intervene and stop the instrument. This is known as a fault diagnosis followed by a prognostic process [27].

1.2 Accidents in the Chemical Industry

The chemical industry is composed of several major sectors which include the activities of the French Nomenclature of Activities (FNA):

- Mineral chemistry: The main components of the inorganic chemistry are water, air, salt, sulfur, and phosphate to produce sulfuric acid and its derivatives. Derivatives of sulfuric acid are products obtained by electrolysis, such as chlorine or sodium hydroxide, compressed gases, and more sophisticated products such as fertilizers.
- Organic chemistry: it consists of petrochemical material such as rubber plastics and elastomers. The products that consist of this sector are ethylene, propylene, butadiene, benzene, ethanol, and acetone ...
- The specialty chemicals: This sector includes all products that have properties for a specific use such as:
 - Paints, varnishes, inks, glues, and adhesives
 - Plant protection products
 - Essential oils

- Soaps, cleaning products and perfumes: are all detergents, and personal beauty care. These products are used by the industrial sector and households.
- The Fine Chemicals: this sector conceptualizes complex molecules (called active ingredients) with many chemical reactions in cascades. Subsequently, these active ingredients are formulated into specific products which are the drugs [28].

The chemical activity in France is growing since 2013, which is a good sign because even with the economic crisis chemical industries manage to increase their activities.

Moreover, some sectors have not experienced in the last four years any increase in its activity, such as the area of mineral chemistry and specialty chemicals. We also see that the French chemical industries have better numbers than the European chemical industries.

A chemical accident is the accidental release of one or more substances which may be hazardous to health and/or the environment in the short or long term. These events may include fires, explosions, leaks or accidental releases of hazardous substances that can cause humans diseases, injuries, disabilities or deaths.

1.3 Terminology

As a step towards a unified terminology, the Technical Committee of International Federation of Automatic Control (IFAC) SAFEPROCESS suggested preliminary definitions in the field of fault diagnosis [29]. Some of the definitions are modified according to the terminology introduced by Blanke et al. in 2000 [30]; the following list shows common definitions presented by these two references:

As browsing through the literature, we realize immediately that the terminology in the field of diagnosis is not consistent. Many definitions of the same word are found. For example, the diagnostic term has several definitions, different by field of application:

- **In medicine**, the diagnosis means the approach taken to determine infection. It is based on research into the causes and symptoms of the infection.
- **In finance**, the diagnosis is a dynamic analysis tool for developing various papers anticipation of future funding needs of the association, establishment or service.
- **In automatic**, the diagnosis is a decision support system that can locate components or a system failing organs and eventually to determine its causes.

This inconsistency makes comparing different approaches tasks and the precision of the work and contribution of the objectives in this area elusive. In order to remove these ambiguities, the SAFEPROCESS technical committee of IFAC (International Federation of Automatic Control) has discussed this problem and it tried to standardize these definitions. In this context, it seems essential to recall the terminology used in this report. These definitions are based on the work of the technical committee SAFEPROCESS [31][29][32].

Structural analysis:

It is the analysis of structural properties of the models; properties that are independent of actual parametric values.

Fault:

This is an unacceptable drift of at least one characteristic property or variable or a deviation of the behavior of the system from its standard behavior; which is usually always acceptable. It does not cause the system to malfunction but lets us consider a probable failure.

Failure:

It may be a consequence of a fault. This is an alteration or interruption of proven performance features of a device.

Break-down:

It represents the consequences of failure in achieving nominal operation of a process. In other words, it is a state of non-operation or malfunction, hardware or software in the sense that a unit is unable to perform a required function, leading to failure. A breakdown can be regarded as permanent or intermittent:

- **Permanent failure:** it is a malfunction of a component that needs to be changed or repaired. It can be the result of gradual changes in the characteristics of a component, such as aging, for example, or a sudden change of sensitive material.
- **Intermittent failure:** it can allow a return of the process in its dynamic operation. These failures often precede permanent failure following a gradual degradation of system performance.

Malfunction:

A permanent interruption of the system's ability to perform a required function, under specified operating conditions.

Surveillance:

- **Definition 1:** Surveillance is a continuous, real-time task that aims to characterize the way the physical system works, by recording information, recognizing and indicating behavioral abnormalities [33].
- **Definition2:** Surveillance is a passive, informational device that analyzes the state of the system and provides indicators. Monitoring includes detecting and classifying failures by observing the evolution of the system and then diagnosing it by locating the failing elements and identifying the root causes [34].

Fault Detection:

It is the determination of the presence of faults and the time of their occurrences.

Fault Isolation:

It is the determination of the type, location and time of occurrence of the fault.

Fault Identification:

It is the determination of the size and the temporal behavior of a fault.

Diagnosis:

It is the determination of the type, size, location and time of occurrence of a fault; it follows the fault detection and includes the isolation and the identification.

Prognosis or Prognostics:

Prognostic is the ability to perform a reliable and sufficiently accurate prediction of the remaining useful life (RUL) of equipment in service.

Mode of operation:

It is a term used to describe the different operational situations of a process. There are three modes of operation: normal, degraded, and failure:

- The mode of operation is considered normal when the system performs its functions without decreasing performance.
- The mode of operation is considered degraded if the system partially fulfills its functions or if its performance suffers.
- The mode of operation is considered a failure when the system is no longer able to perform its functions or that its performance is greatly reduced.

Failure effect:

It is the consequence of a failure mode of the operation, function, or status of a variable.

Qualitative model:

It is a system model describing the behavior with relationships between variables and parameters of the system in terms of causalities such as heuristics or rules.

Quantitative Model:

It is a system model describing the behavior of relationships between system variables and parameters in analytical terms such as differential equations or differences.

Fault Modeling:

It is the determination of a mathematical model to describe a specific effect of the fault.

Reconfiguration:

It is changing of the structure and controller settings. The original control is achieved although the performance may be degraded.

Analytical Redundancy:

It is determining a variable by measurement or by use of a mathematical model of the process considered.

Residue:

The bearing information signals, based on the difference between measurements and calculations based on the model.

Fault tolerance:

Fault tolerance is the fact that the system continues able to perform the desired tasks or if necessary to achieve new goals in order to avoid catastrophic trajectories, even in the presence of one or more faults. It is based on two approaches to configuring and accommodation:

- **Reconfiguration:** it is the function of changing the control of the system or the substantive provisions of the system so that non-failed components can deliver acceptable service.
- **Accommodation:** the reconfiguration of the system without compromising its objectives or its structure. It is to correct or cancel the effects of a default or through a recovery procedure or offset errors.

Recovery:

It is to find a cure for failure (eg: replacing the faulty component).

Maintenance:

Is to replace or repair faulty equipment or users. In maintenance two types of service are possible:

- Preventive maintenance: it can be systematic or conditional (she speaks at the prediction of a future failure of the physical system).
- Corrective maintenance shall be carried out after the failure. It is either palliative (temporary emergency solution for the failing part to ensure at least some of its features) or curative (it corresponds to a replacement of the defective component and a reset state of the system).

Threshold:

The limit value of the deviation of a residue with zero, so if it is exceeded, a fault is declared as detected.

1.4 Fault Type

A fault is defined as not allowed deviation between the actual value of a characteristic of the system and its nominal value. As shown in Figure 1.2, three types of faults are distinguished: actuator failure, sensor failure and process failure (or component fault).

Each of these faults, as well as their influence on the process, is briefly described in the paragraph below.

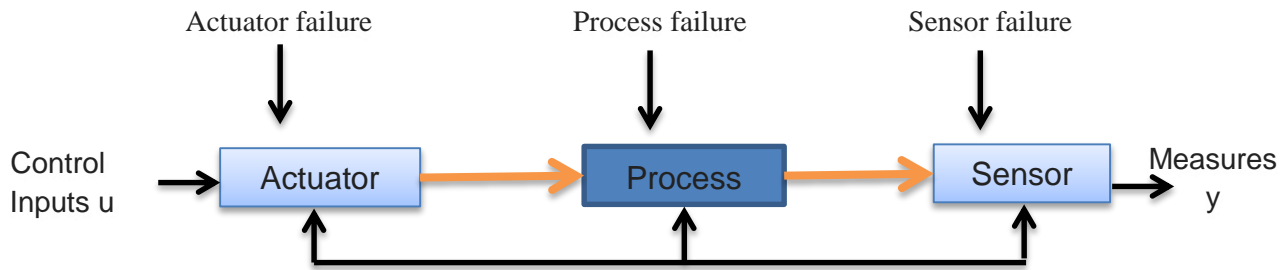


Figure 1.2: Various Physical System Faults Types

According to Isermann et al. (2002), the faults can be differentiated according to their form (systematic or random), according to the extension of the fault (local or global) or according to their temporal evolution [35].

1.4.1 Fault Classification According to the Extension

a- Sensor Fault

The sensors are instruments that transform a physical grandeur into voltage grandeur. The sensors are essentially the output interfaces of a system with the external environment. They help to communicate information concerning the status and the internal behavior of the process. Thus, a fault sensor characterizes a bad image of the physical quantity being measured. For closed-loop systems, the measurements from these sensors are used for generating the control signal. Therefore, the presence of a faulty sensor provides an inaccurate and inefficient control signal. The faults most common sensors are: a) bias b) drift, c) the loss of efficiency, d) the locking and e) the calibration fault. Figure 1.3 shows the effect of these faults on the measurements [36].

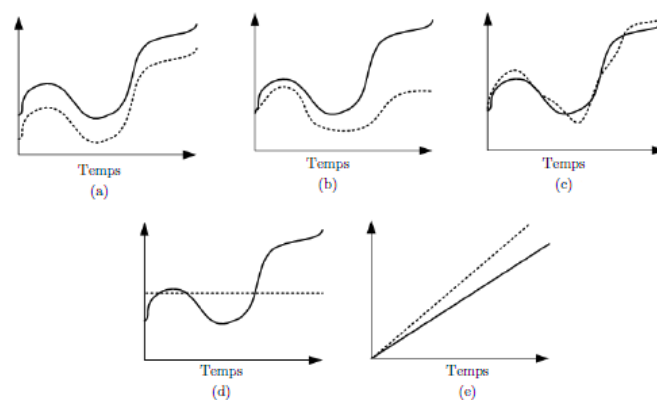


Figure 1.3: The Effect of Different Fault Types on the Sensor Measurements. Dotted Lines Indicate the Measured Values of the Sensor; However, the Continuous Line Represents the Actual Values.

In 2009 Sobhani et al. define the mathematical equations (eq.1.1) for these faults as:

$$y_i(t) = \begin{cases} x_i(t) & \forall t \geq t_0 \\ x_i(t) + b_i & b_i(t) = 0, b_i(t_{f_i}) \neq 0 \\ x_i(t) + b_i(t) & |b_i(t)| = c_i(t), 0 < c_i < 1 \\ x_i(t) + b_i(t) & |b_i(t)| \leq \bar{b}_i, \dot{b}_i(t) \in L^\infty \\ x_i(t_{f_i}) & \forall t \geq t_{f_i} \\ k_i(t)x_i & 0 < \bar{k} \leq k_i(t) \leq 1 \end{cases} \quad \forall t \geq t_{f_i} \quad (1.1)$$

b- Actuator Fault

The actuator is part of the operative part of a system which converts the control signals from the controller (processor) in a physical grander like the heat (electric heater) the movement (motor), or magnetic field (solenoid)...Thus, the actuators faults act at the operative part and destroy the system input signal. The consequences of actuators faults can vary from a high consumption of energy until the total loss of control. Actuators faults vary from one actuator to another, but not an exhaustive classification of the most common faults is given in Figure 1.4 [36].

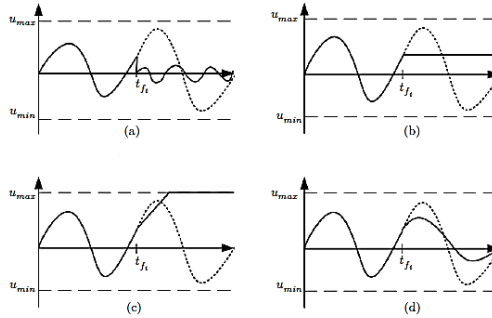


Figure 1.4: Graphical Representation of the Most Common Faults Actuator (Ducard, 2009) [36]. Dotted Lines Designate the Desired Values of the Actuator; However, the Continuous Line Lines Represent the Current Values. (a) Oscillation, (b) Blocking, (c) Saturation and (d) Loss of Efficiency.

Different types of actuator faults can be represented by the following expressions:

c- Process Fault

Components faults are faults that affect the system components itself. These are faults that cannot be classified either among the actuators faults or faults of the sensors. This type of fault causes a change in the dynamics of the system following a change in these parameters. The mathematical representation of the components faults is often difficult to determine and requires extensive experimental tests. In general, they result in a change in the state equation. This change can be parametric or structural/functional. These faults induce system instability.

1.4.2 Fault Classification According to the Temporal Evolution

The temporal evolution of the faults (Figure 1.5) is unpredictable: they can be of low or high amplitude, abrupt or gradual in the form of drifts. The following faults are distinguished in the literature:

- Sudden fault (or modeled by a bias) (1);
- Incipient fault (or drift) (2);
- Intermittent fault (with interruptions). (3)

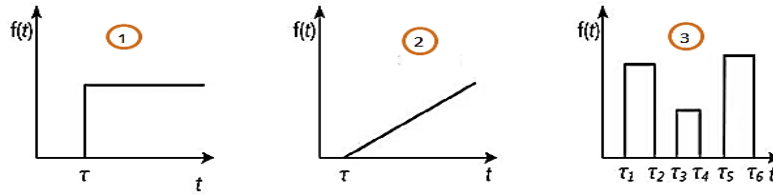


Figure 1.5: Temporal Evolution of the Different Types of Faults

a- Sudden Fault:

This type of default is characterized by a discontinuous temporal behavior. This development, if it does not match the expected normal dynamic changes to the variable (setpoint), is characteristic of a sudden failure of the element in question: total shutdown or partial connection [35]. A mathematical representation of this fault (eq.1.2) is given by:

$$f(t - t_{fi}) = \begin{cases} \delta & t > t_{fi} \\ 0 & t < t_{fi} \end{cases} \quad (1.2)$$

Where, $(t - t_{fi})$ is the temporal behavior of the fault and δ is a constant threshold.

b- Incipient Fault:

It is a characteristic absence of wear of a part or fouling. It is very difficult to be detected because of its temporal evolution that may be confused with slow parametric change representing the non-stationary process. Its evolution over time can be expressed by this relation (eq.1.3):

$$f(t - t_{fi}) = \begin{cases} \delta(1 - e^{-\alpha(t-t_{fi})}) & t \geq t_{fi} \\ 0 & t < t_{fi} \end{cases} \quad (1.3)$$

Where δ and α are two positive constants.

c- Intermittent Fault:

It is a default feature of false contact or intermittent failure of the sensors. This type of fault is a special case of abrupt or suddenly default with the special property that it is randomly its normal value.

The faults can be classified as additives and multiplicative faults, depending on their impact on system performance (Figure 1.6). Additives faults are interference signals which are superimposed at a point of functional. Les diagram faults sensors and actuators are typically modeled as additive faults, however, the

component conditions are modeled by multiplying faults. These bring about changes in the correlation of the output signal of the system, as well as a change in the spectral characteristics and the system dynamics.

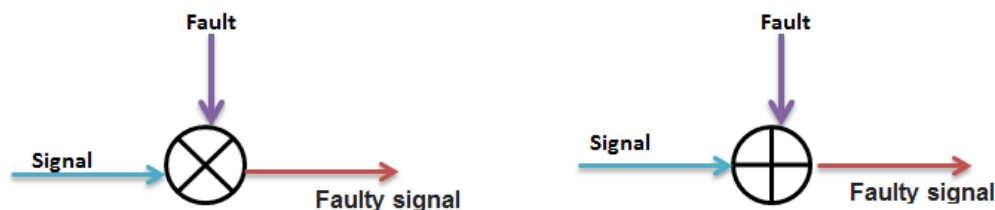


Figure 1.6: Faults Classification: Multiplicative and Additive.

In the chemical industry, the processes mainly used are reactors and distillation. In this work, we consider the problem of faults diagnosis and prognostic that occur in the distillation column.

1.5 Principle of Distillation

Distillation is one of the widest processes used in the industrial separation methods. This operation allows the separation and the purification of the mixtures of constituents whose boiling temperatures are different, taking advantage of the difference in volatility (capacity to evaporate according to the temperature) of the constituents and then carrying out a succession of condensations and vaporizations to finish by the recovery of one component. This separation procedure has been known for over 2000 years.

The distillation consists of heating to a boiling point a liquid mixture of constituting the most volatile, evaporates the former. By condensation of the vapor phase, a liquid called distillate or extract (also referred to as top product) is recovered with a high concentration of the most volatile compound. The un-vaporized liquid phase constitutes the residue or the raffinate (also called the bottom or bottom product).

The distillate is not a pure product; it is enriched with a lower boiling light component, while the bottom product is formed by the heavy component having a higher boiling point. In order to obtain a distillate of high purity, particularly when the components of the mixture have similar boiling temperatures, rectification is often employed.

Rectification is a technique which uses repeated distillation so as to cause exchanges between the upstream and liquid phases flowing back into a column intended for bringing the two phases into contact [37].

In the industry, distillation is the most widely used unit operation in terms of industrial processes, indeed it is very important. Distillation has many applications such as:

- Distillation makes it possible to separate the air into its constituents, in particular, oxygen, nitrogen, and argon for industrial use.

- In the field of industrial chemistry, raw liquid chemical synthesis products are distilled to separate them from other products.
- The distillation of fermented products produces beverages with high alcohol content.

So we see essentially that distillation plays a major role in the industry, and it affects practically all the sectors, because as previously stated it is a method of separation and not a reaction.

1.5.2 Applications of Distillation

The application of the distillation is divided into several groups:

- laboratory scale
- Industrial distillation
- Distillation of medicinal herbs and perfumery
- Food processing

The difference between distillation of the laboratory scale and industrial distillation is that the distillation method is not the same, on a laboratory scale it is often carried out batch-wise, while industrial distillation occurs continuously.

With batch distillation, the composition of the vapors of the compounds and of the distillate changes during the distillation. In discontinuous distillation, the feed is separated into its fractions which are recovered from the most volatile to the least volatile. In continuous distillation, the feed and distillate compositions are kept constant and at a constant temperature.

1.5.3 Industrial Distillation

Distillation plays a major role in unit operations, in fact, it is the process which is used most in terms of separation, rather than liquid-liquid (L-L) extraction or even with another absorption column, is not that its use is simple but it is less complex than other unitary operation. In addition, it allows a better separation than the L-L extraction process. In fact, it depends on the nature of the product to be separated, as for example, to separate the various constituents of petroleum we are not going to use the extraction process L-L but the fractional distillation which is more action to meet the requirements of separation [38].

Large-scale industrial distillation applications include fractional and continuous distillations. The most widely used industrial applications (continuous) are in petroleum refineries, petrochemical plants, as I mentioned earlier in my illustrations. The distillation is carried out in vertical cylindrical columns, called distillation columns, these columns can have diameters ranging from about 65 cm to 16 meters and the height ranging from about 6 meters to 90 meters, which seems a little large but on an industrial scale it is normal to view that one manipulates large quantity. When the feed of the process has a different composition of several components, as we have seen before in the distillation of the crude oil, we will have different liquid outlet fraction which allows the withdrawal of the various fractions of products which have

points d'ébullition différents. Les produits "légers" iront plutôt à la sortie du haut de la colonne et les produits "lourds" à la sortie du bas de la colonne [39].

1.5.4 Distillation Column

La distillation peut être réalisée de manière continue ou discontinue. Elle se fait dans un appareil appelé colonne de distillation. Cet appareil est composé d'une enveloppe cylindrique verticale contenant des dispositifs (trays ou packings) permettant d'augmenter la surface de contact entre la phase liquide et la phase gazeuse, améliorant ainsi les échanges entre les phases pour un volume de colonne donné. En plus de la colonne et de ces trays ou packings, deux échangeurs de chaleur permettent de fournir / retirer l'énergie nécessaire à la séparation : un chauffage situé au bas de la colonne et un refroidisseur au sommet de la colonne qui permet de liquéfier les vapeurs afin de récupérer le produit purifié sous forme liquide. Une partie des condensats est souvent réinjectée dans la colonne pour augmenter la pureté du produit souhaité, le reflux (Figure 1.7).

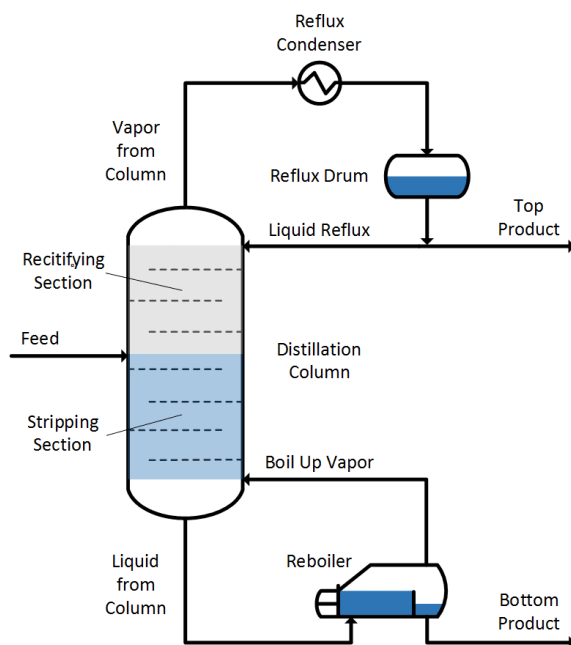


Figure 1.7: Design of a Binary Distillation Column

Le mélange liquide à séparer (aliment) est introduit dans la colonne. Ce liquide est ensuite porté à ébullition par le chauffe-vapeur. La vapeur s'échappant du chauffe-vapeur est condensée dans le condenseur. Le liquide non vaporisé est extrait au bas de la colonne et constitue le résidu. La vapeur produite est toujours plus riche en constituant le plus volatil, de sorte que la composition du condensat est à tout moment supérieure à celle du mélange initial. Une partie de la vapeur condensée est renvoyée au sommet de la colonne et constitue la phase liquide dans la partie supérieure de la colonne, qui est le reflux. En effet, le liquide se déplace dans la colonne par

gravity from top to bottom and the steam moves from bottom to top. The remainder of the steam is removed and forms the distillate [40].

1.6 Description of the Used Distillation System

1.6.1 Automated Continuous Distillation Column

The purpose of this distillation is to reproduce in large quantities the separation between two compounds by acting on the difference in boiling temperatures. One of the compounds being more volatile, it will vaporize to the top of the column where it will be condensed with a total condenser. There will, therefore, be an exchange between the rising steam and the descending liquid. This is called steam enrichment. Some of the condensates will be separated at the top of the column, it is the distillate. Another part will be separated at the bottom of the column; the latter is called the residue [41].

However, there is an optimization of this process; it is to make it automated. It can be automated by the control of:

- Control of flow and feed temperature
- A regulation of the pressure drop within the column
- Temperature control at the head of the column using a regulated reflux rate thanks to a timer
- A regulation of the heating capacity of the boiler

The system is equipped with a structure point of view (Figure 1.8):

- A Tray or packed column (1)
- A total condenser at the top of the column (2)
- A feed preheating (3)
- A Timer (4)
- A reservoir of distillate (5)
- A reservoir of residue (6)
- A boiler (7)



Figure 1.8: General View of the Distillation Unit

- 1- **The Timer** consists of an electromagnetic valve and thus regulates the rate of reflux. In addition, it is regulated by TIC2 (temperature at the top of the column). When the setpoint temperature of TIC2 is reached, the timer opens. Otherwise, it remains closed (automatic mode). The distillation unit has eleven sensors for continuously measuring the temperature throughout the column.
- 2- **The Metering Pump** is a volumetric pump. The configuration of the pump assembly is suction mounted. A non-return valve prevents the preheating tank from emptying. The volumetric pump is made up of a membrane allowing, firstly, the suction of the mixture to be distilled and then its discharge to the preheating tank. According to the construction data, the feed pump can deliver a flow rate of 4.32 L / h.
- 3- **The Boiler** is heat exchangers whose purpose is to allow the liquid to become depleted in volatile constituents and steam to enrich it. The amount vaporized is called the re-boiling rate. There are several types of boilers, tube boilers, and boilers with immersion resistors. The boilers of the first type have the operating principle of heating the mixture to be separated by means of a heat-transfer fluid circulating in the calendar of the exchanger (boiler). In general, the fluid most often used is water vapor. The principle of operation of the second type of boiler is to send a certain intensity

thus enabling to provide an adequate power to bring the mixture to a boil. The characteristics of the boiler of the system studied are the following:

- The boiler has a capacity of 2 liters and has a level sensor allowing the automatic stop of the heating if the level is insufficient.
 - The immersion heater has a heating power of 3.3 kW.
- 4- **The Condenser**, the main function of the condenser is to convert the rising steam into liquid on a cold surface, or via a heat exchanger kept cold by the circulation of a refrigerant fluid. However, the temperature of the liquid must not be lower than the boiling temperature of the mixture to be separated; otherwise, the reflux would be too cold and unbalanced the column.
- 5- **The Preheating** is made up of three cartridges of heating resistors with a power of 250 W each. It also has a level sensor that prevents starting if the level in the glass body is too low.

The monitoring system consists of the **ETP 200 software** (Figure 1.9). It monitors the evolution of physical quantities such as differential pressure or temperatures at a given point in the distillation column. This software has multiple functions for automated control of the distillation column.

The first feature is the modification of parameters. This allows the configuration of the control parameters of proportional–integral–derivative controller (PID). In addition, this menu allows setting the installation in the manual or automatic mode as well as the operating instruction of a process. The display of the measurement of each device is observed.

The second important feature is the group trend. The group trend makes it possible to follow the evolution of the measurements of each apparatus linked to the regulation. This evolution can be visualized in real time or in history.

The third important menu is the trend of a device. Like the previous functionality, it allows to follow the evolution of the measurement, the set-point as well as the output. There is a menu which allows having a history of the alarms that could be triggered during a manipulation. This software also has a synoptic visualization.

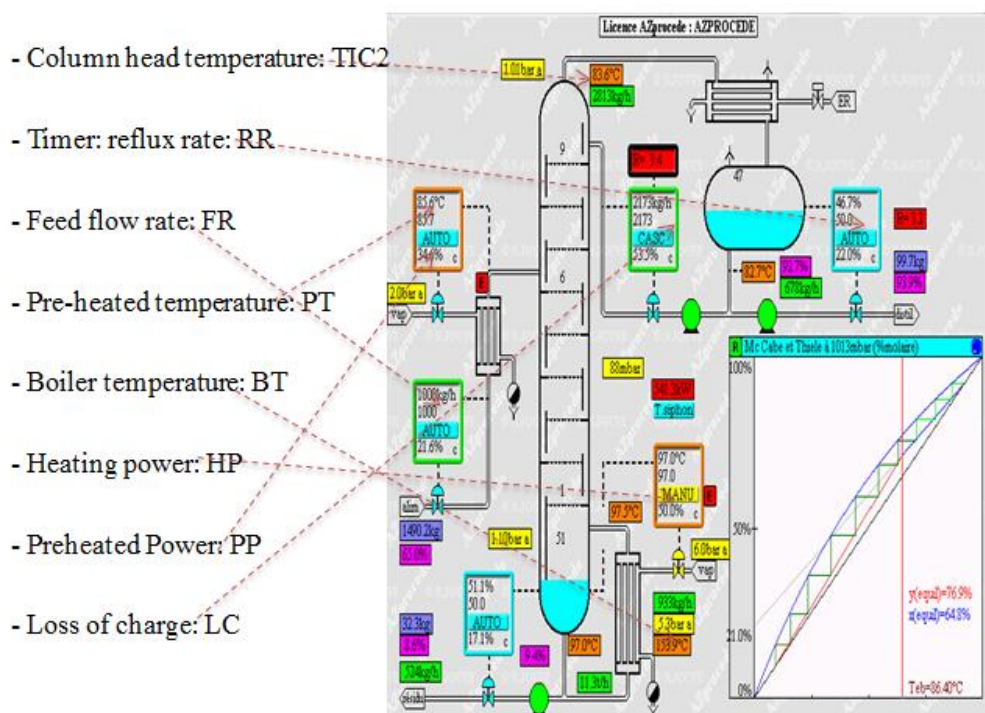


Figure 1.9: Supervision System ETP 200

The ETP 200 software allows monitoring of changes in parameters such as differential pressure or temperature at a given point of the distillation column. The signals obtained during each acquisition represent:

1. Heating Power (HP)
2. Timer: Reflux Rate (RR)
3. Feed Flow Rate (FR)
4. Preheated Power (PP)
5. Loss of Charge (LC)
6. Pre-heated Temperature (PT)
7. Boiler Temperature (BT)
8. Column Head Temperature (TIC2)

1.6.2 Distillation Column Accidents

The reactor where the biggest accident factor resides is the discontinuous reactor because all the addition is carried out at the beginning. If the reaction is exothermic, the reaction will produce energy called heat and the latter will increase the speed of the reaction and thus release a quantity of heat in shorter and shorter time and so on and cause a loss of control of the temperature within the reactor. This phenomenon is called thermal runaway. Thermal runaway can also occur when decomposing a substance. The loss of temperature control may be due to an increase in reaction rate, an adverse reaction that does not occur under normal process conditions (This reaction is often the decomposition reaction of the compound), an increase in the pressure inside the reactor caused by the evaporation of the reaction components or caused by the production of non-condensable gases. The increase in pressure can lead to serious accidents because it can cause the wall of the reactor to cause the projectile to emit a great distance. It can cause a fire because the mixture is found in the open air and fine particles, liquid or solid, can ignite. If the reactor is in a confined area then this may cause an explosion as the mixture of gases, droplets and air can explode and cause a large blast effect which causes the destruction of the confined area. This explosion continues with a secondary fire. But there is another "flagship" process in the chemical industry which can have serious consequences in the event of accidents, and it must be taken into consideration, it is the accidents of the distillation column.

On 24/04/2013, in LE-BAR-SUR-LOUP, an accident occurred on a distillation assembly at 100 ° C under vacuum in a glass column. The latter is building up pressure at 12:15 the laboratory of a plant of essential oils classified Seveso. Under the effect of pressure, the boiler and the column explode projecting the hot substance on 3 operators. The operators are lightly burned and are sent directly to the hospital where they will come out late in the afternoon.

The cause of the increase in pressure is due to the decomposition of the product or of a by-product during the reaction whose process has been used since 1995. The thermal runaway is due to insufficient cooling because the installation was bad resized because it had to change scale because of greater demand. Thermal runaway may also be due to a high column temperature, high vacuum, or a quality of a product that may vary.

Unfortunately, there are several malfunctions that come from the central unit like:

- Precise Column Head Pressure Control
- Hysteresis of the Control Valve
- Preheating the Power Supply
- Contamination of the Timer
- Loss of load too high in the column / too high flow of liquid which causes waterlogging
- Insufficient reflux rate impairing the quality of the distillate/residue

1.6.2.1 Data Collection for Faults Diagnosis and Prognostic

For faults diagnosis and prognostic a database with normal and faulty observations is acquired from the distillation column installed at IUT of Rouen-France. The database is composed of 120 observations, where 50 of them were used as data collection for faults diagnosis and the rest (70 observations) were used for faults prognostic; each observation has 5000 points with the sampling frequency, $SF=1\text{sample}/10\text{sec}$.

1.6.2.1.1 Normal Mode

The normal mode (Figure 1.10 & Table 1.1) is the common thread during all accidents. To do this, the column is left at atmospheric pressure and the chosen boiling temperature of the acetone set at 56.0°C . Thus in normal mode we obtained the following parameters:

- Feed flow rate set at 80% of its capacity.
- Pressure drop at 0.7 mbar
- Preheat temperature at 40°C
- Boiler temperature at 76°C
- Column head temperature at 56°C

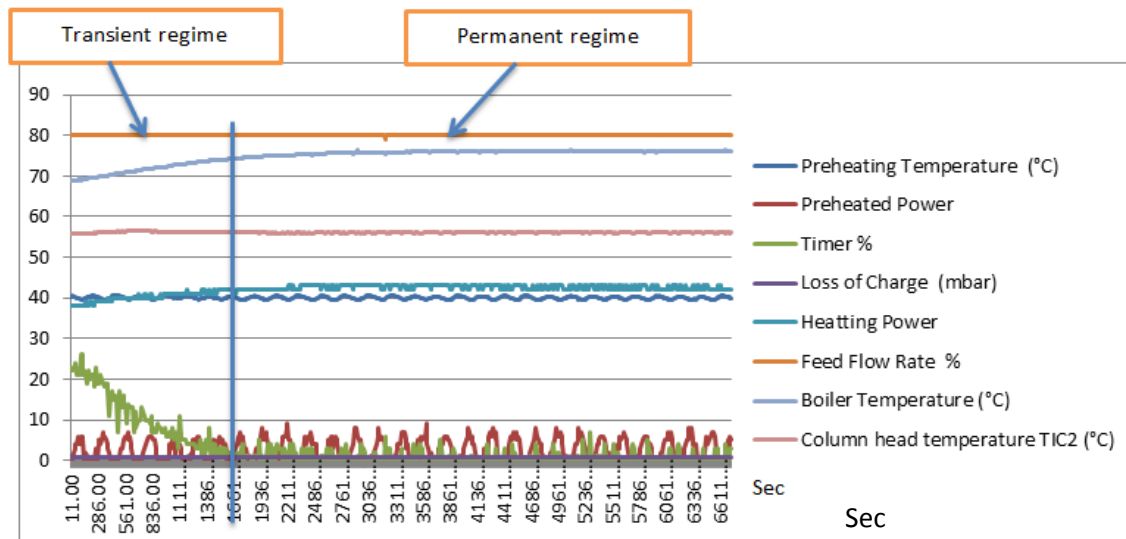


Figure 1.10: Graphical Representation of the Normal Mode

In the transient regime, the signals move a lot and the system is not stable for that we cannot count on this regime and include it in our study. So this scheme is deleted.

Signal	Mean	Max.	Min.	Variance	Frequency peak
Preheated temperature (°C)	40	40.6	39.6	0.006	0.0006
Timer %	0.5	7	0	1.6	0.001
loss of Charge (mbar)	0.7	0.7	0.7	0	0.0001
Heating power	42.7	42.9	42.3	0.009	0.0001
Feed flow rate %	80	80	80	0	0.0001
Boiler temperature (°C)	76	76.1	75.8	0.009	0.0001
TIC2 (°C)	56	56.1	55.9	0.001	0.0001
Preheated power	2.743044	9.1	0	6.354596	0.0005

Table 1.1: Statistical Characteristics of the Acquired Signals in Normal Mode

1.6.2.1.2 Data Collection for Faults Diagnosis

We shall now turn to the accidents of continuous automated distillation which may occur in the industry. To do this, we will provoke them by varying four parameters:

- The Reflux Rate (RR).
- The Heating Power (HP).
- The Preheating Power (PP).
- The Feed Flow Rate (FR).

1- Reflux Rate (Timer) Accident Mode

In the accident of reflux rate (Timer) (Figure 1.11), we first established the normal mode at time T1 then at time T2 we triggered the timer accident at **0%**. We can see that no parameter changes, the column is indeed in infinite reflux, which allows it to put itself at equilibrium. There is only one parameter that changes, it is the heating power. We can see that this parameter (heating power) decreases; it is to keep the loss of charge at 0.7 mbar because there is more condensed vapor which falls in the column which means that the loss of charge is likely to increase. It is therefore by decreasing the heating that there is less steam rising in the column, thus less loss of charge.

At time T3 we reset the system to the normal mode and then at time T4 we caused the accident **100%** of the timer. In this case, we can see that the temperature at the top of the column (TIC2) has increased very rapidly to 75 °, which means that it is near to the boiling temperature of ethanol which is 78.6 ° C. This is explained by the fact that there is no liquid falling back into the column. Indeed, to compensate for the decrease in the pressure drop, the boiler heater must send more steam into the column, so the heating increases, as we can see on the Figure 1.11. Now, there remains essentially only the vapor in the column, so the temperature TIC2 increases strongly. Because of this, the liquid level in the boiler decreases rapidly and the level of the distillate increases. We, therefore, interrupted the manipulation at the time T5 because there was a real danger.

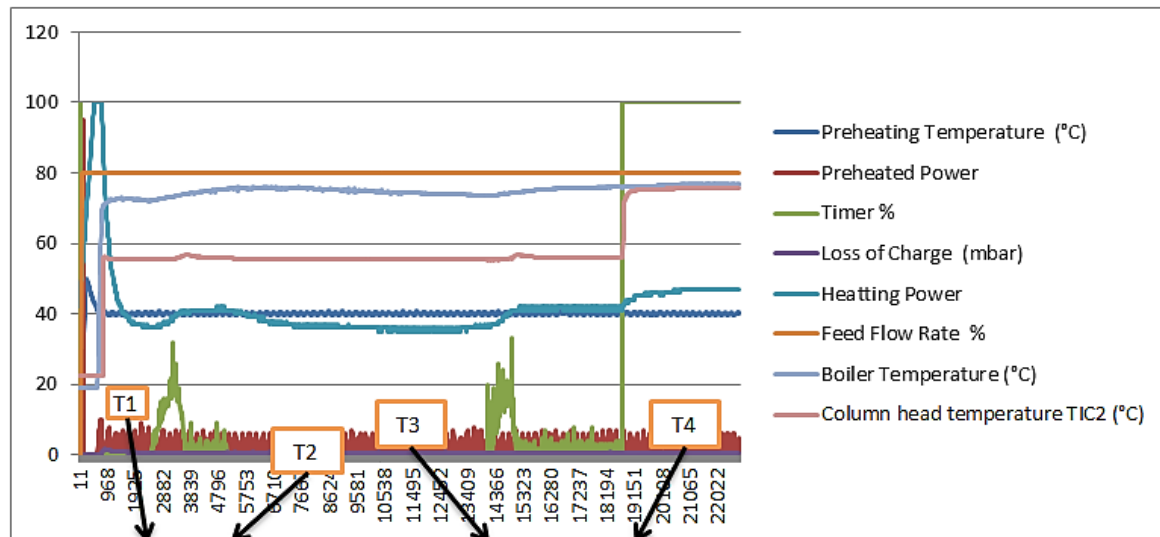


Figure 1.11: Graphical Representation of the Timer in Accident Mode

2- Heating Power Accident Mode

In the heating power accident (Figure 1.12), we first established the normal mode at T1 then at the time T2 we caused the accident with 100% heating. We can notice that the loss of charge increases to 3.6 mbar; this is due to the amount of vapor going into the column that cannot be compensated by the condensed vapor. Indeed we note that the timer closes as soon as the heating reaches 100%. This allows all condensed vapor to return to the column. On the other hand, we can see that the temperature at time T2 rises to 61 °C then returns to 56 °C. This shows that there is a return to equilibrium of the column but with a high-pressure drop. At time T3 we returned to the normal mode then we simulated the accident of 0% heating at the time T4. The temperature TIC2 returns to the outside temperature, the loss of charge decreases to 0.2 mbar.

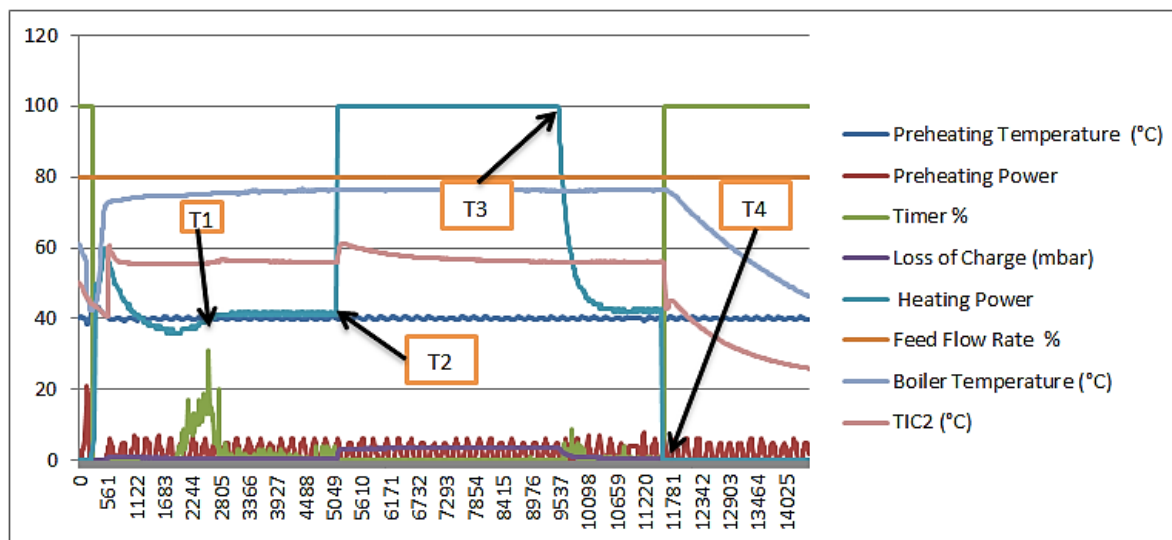


Figure 1.12: Graphical Representation of the Heating Power in Accident Mode

3- Preheating Power Accident Mode

In the accident of the preheating power (Figure 1.13), we first established at T1 the normal mode and then we caused the accident of 0% preheating at the time T2. At time T3 we returned again to the normal mode and then simulated the 95% preheating accident at the time T4. We had two accidents. The first accident, we simulated a failure of the heating resistance of the preheating. The second accident, we simulated a loss of control of the preheating and we put a preheating power to 95%. We were not able to put a 100% power because the preheating is oversized for the process because it is made to preheat heavy components while our components are light. If we put 100%, we could blow up the preheating reserve because the pressure would be too high because of the steam. It can be seen that on the first accident, stopping the preheating caused a small difference in the temperature in the boiler and the heating power because it had to increase to bring energy to the liquid coming from food. The temperature in the boiler has decreased because there is a temperature gradient between the vapors rising in the column and the liquid coming from the feed and thus this has caused a heat transfer between these two currents.

In the second accident, a small increase in the loss of charge can be seen therefore a logical decrease in the heating power. The loss of charge increases because when preheating has a power of 95%, the temperature of the feed mixture is increased to 70 ° C., a temperature above the boiling temperature of acetone at atmospheric pressure. Thus, the feed fluid is composed of steam and liquid. At the arrival of this fluid at the level of the distillation column, the steam rises while the liquid descends into the column. Moreover, the liquid is at a temperature close to the temperature of the vapor, so that there is little vapor which condenses so that there is a larger flow of steam rising in the column. Since there is a higher steam flow then the loss of charge is higher. Now, the boiler is regulated by the pressure drop, so we have a lower heating power than when we are in a normal mode. A slight increase in TIC2 can be seen, i.e. 1 or 2 ° C. We do not have a big increase even if we could believe it because the boiler and the timer are regulated so

as to keep a constant temperature. The risk that can occur is to have too high a vapor flow rate with respect to the flow of liquid and this could lead to congestion, that is to say, that the liquid stream could be blocked by a large current of steam. The risk is to have an explosion of the column head by the pressure of the vapor.

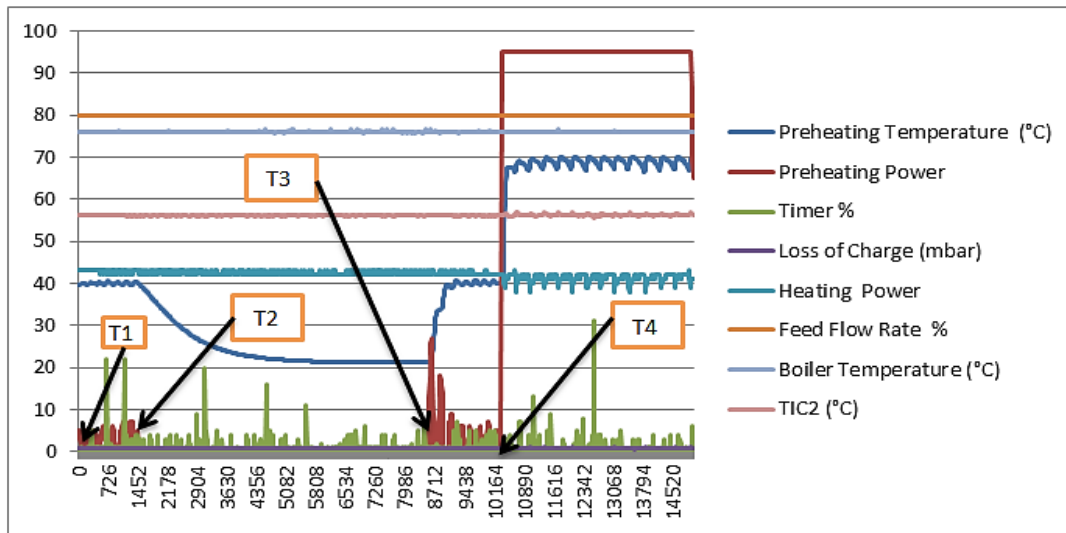


Figure 1.13: Graphical Representation of the Preheating Power in Accident Mode

4- Feed Flow Rate Accident Mode

In the accident of feed flow rate (Figure 1.14), we first established the normal mode at T1 and then we caused at the time T2 the 100% feed flow rate accident. At time T3 we went back to the normal mode and then simulated at the time T4 the 0% feed flow rate. At time T5 we went back to the normal mode and then simulated another 0% feed flow rate accident at time T6.

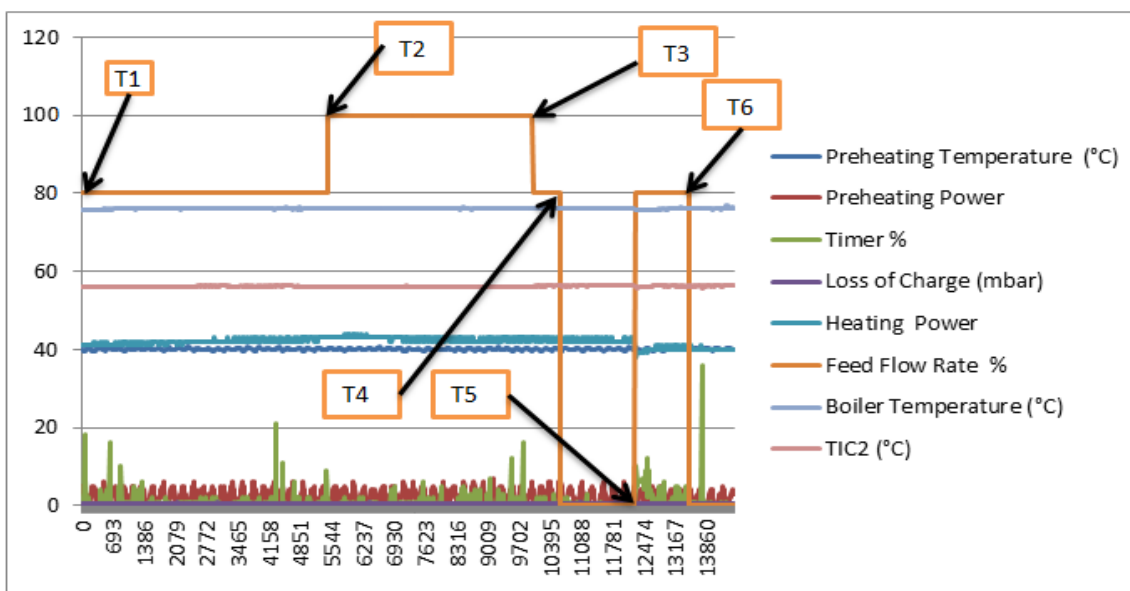


Figure 1.14: Graphical Representation of the Feed Flow Rate in Accident Mode

We can see that when the pump flow is triggered at 100% (**Accident with 100% feed flow rate**) at time T2, that there is no great consequence that occurs, we can see that the various parameters remained intact; this accident is comparable to the normal mode. The only consequence of a high feed rate is to fill the boiler level, but this would be possible with a pump that is more efficient than the one used.

When we apply the **Accident with 0% feed flow rate on the system** we see that when the accident occurred at the instant T4 and T6, shortly after we had a greater capacity to use the timer than the normal mode, we went from 19% to 39%, small modification we see no other modification that could have been due to a feed rate at 0%. The most serious consequence would be that the timer remains open and thus all the liquid present in the boiler turns into vapor and condenses at the top of the column and that everything is recovered in the distillate. This would cause the boiler to lower its level until the heater is no longer in contact with the liquid and this could cause a break in the boiler, but for that, it would take a very long time, does not open regularly.

1.6.2.1.3 Data Collection for Faults Prognostic

A. Data Collection via ETP200

In the part of fault prognostic, a database composed of seven 10 observations from each type of the increasing or decreasing fault over time (degradation of the system). While an accident occurs in the automated distillation process, it causes a cumulative increasing or decreasing over time on the following parameters:

- The Reflux Rate (RR).

- The Heating Power (HP).
- The Preheating Power (PP).
- The Feed Flow Rate (FR).

1- Progressive Accident on Reflux Rate (Timer)

In the progressive accident of reflux rate (Timer) (Figure 1.15), we first established the normal mode at time T1 then at time T2 we triggered the timer increasing progressive accident from 0% to 100%, the experience had stopped on T3 (100% reflux rate). In the case of increasing reflux rate accident (between T2 and T3), we can see that the temperature at the head of the column (TIC2) has increased very rapidly to 75°C, and the heating power also increase from 45°C to 50°C. As we mentioned before in the diagnosis accident of reflux rate, the liquid level in the boiler decreases rapidly and the level of the distillate also increases, and this is very dangerous.

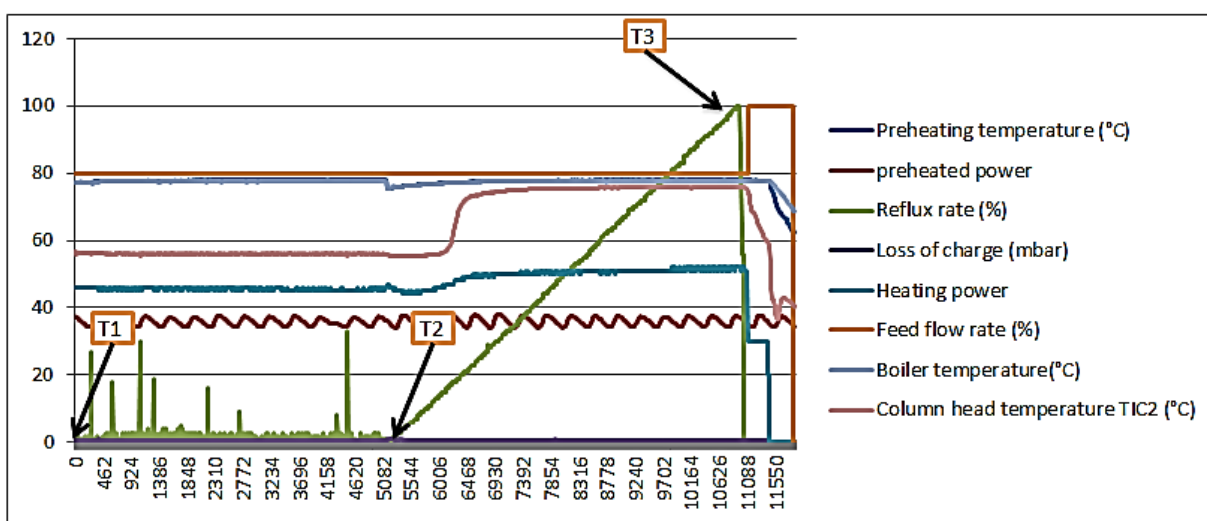


Figure 1.15: Graphical Representation of the Progressive Accident on Reflux Rate

2- Progressive Accident on Heating Power

In the progressive (increase and decrease) heating power accident (Figure 1.16), we first established the normal mode at T1 then at the time T2 we caused the increasing progressive accident of from normal mode (heating power $\approx 45\%$) to 100% heating power. We can notice that the loss of charge increases to 3.6 mbar. At the time T3, we caused the decreasing power heating accident (from 100% heating power to 0% heating power). We can see that the TIC2, preheating temperature and boiler temperature quickly falls when the heating power less than 20% contrariwise the reflux rate stay decrease when the heating power less than the normal value ($\approx 45\%$).

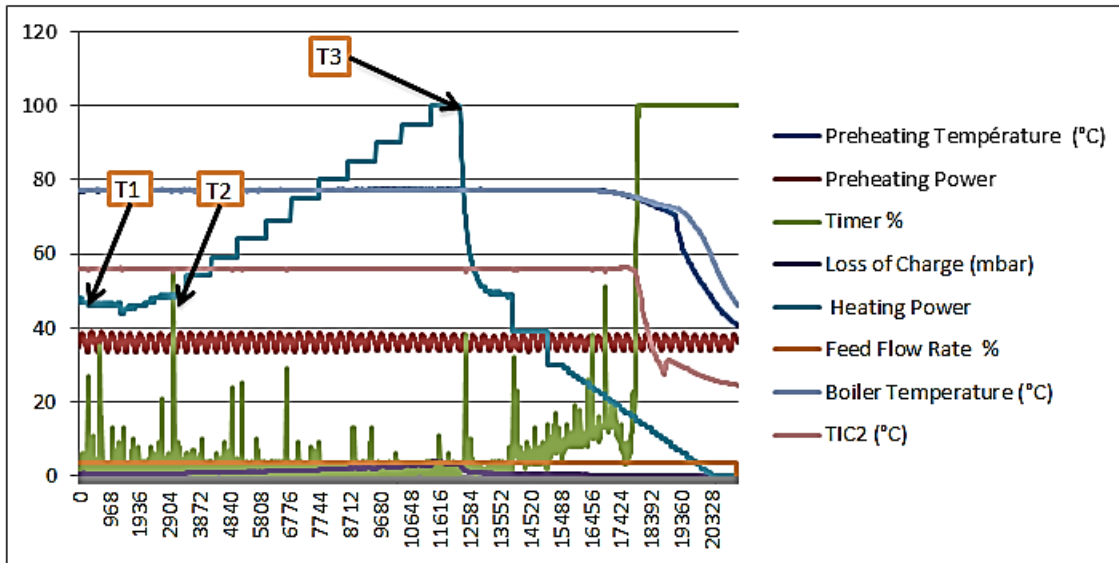


Figure 1.16: Graphical Representation of the Heating Power in Progressive Accident

3- Progressive Accident on Preheating Power

In the progressive accident of the preheating power (Figure 1.17), we first established at T1 the normal mode and then at time T2 we caused the decreasing accident (from normal range $\approx 37\%$ to 0%) of preheating power, at this stage we can see that the preheating temperature decrease also from their normal range $\approx 40^\circ\text{C}$ to $\approx 20^\circ\text{C}$. At the time T3 we moved to the increasing stage of preheating power accident (from 0% to normal range $\approx 37\%$), we can see that the preheating temperature returned to the normal range ($\approx 40^\circ\text{C}$). at the time T4 we back again to the normal mode (preheating power $\approx 37\%$), then at T5 we caused the decreasing preheating power accident (from normal range $\approx 37\%$ to 72%), at this stage we can see also the preheating temperature increase (from normal range $\approx 40^\circ\text{C}$ to 70°C).

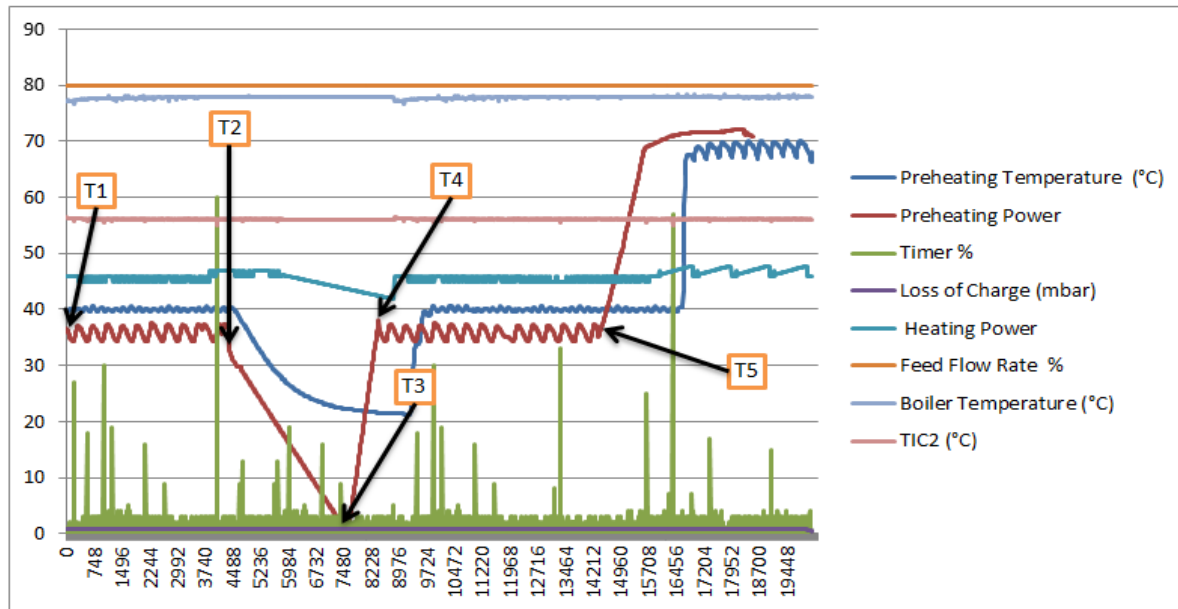


Figure 1.17: Graphical Representation of the Progressive Accident on Preheating Power

4- Progressive accident on feed flow rate

In the progressive accident of feed flow rate (Figure 1.18), we first established the normal mode at T1 and then we caused the increasing feed flow rate (from normal range $\approx 80\%$ to 100%) at the time T2, at this stage we have noticed that the boiler temperature and the heating power decrease. At time T3 we starting the decreasing feed flow rate (from 100% to the normal range $\approx 80\%$), in this case, we can see that the boiler temperature and the heating power come back to their normal range. At the time T4 we returned again to the normal mode and at T5 we simulate the decreasing feed flow rate (from normal range $\approx 80\%$ to 0%), at this stage, we can see that there is no noticeable change in the whole acquired signals.

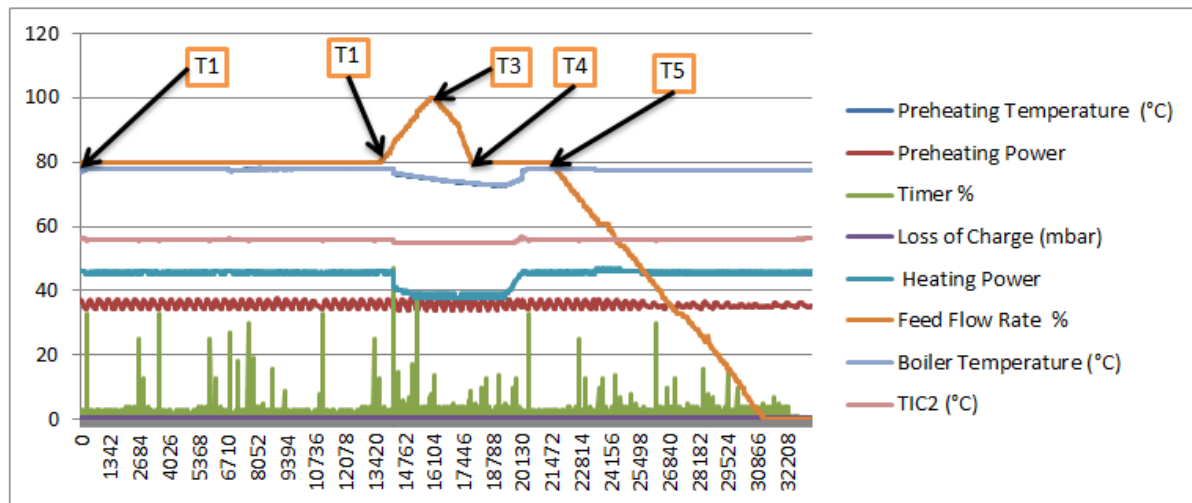


Figure 1.18: Graphical Representation of the Progressive Accident on Feed Flow Rate

B- Data Collection Via External System

The aging of the distillation column components should be considered, especially the aging of the metering pump because it controls the input flow rate. When we have a problem with the pump Impeller, the flow rate will decrease in this case. The most serious consequence would be when the Timer remains open and thus all the liquid in the boiler turns into vapor. This vapor will condensate when it is in contact with the condenser which is installed at the top of the column. The result is a liquid stocked at the distillate reservoir. The result of all this it will be a rapid drop in the level of the liquid that is found in the boiler, which means the heater is no longer in contact with the liquid and this could cause a break in the boiler.

In practice, good design will produce low vibratory levels in a rotating machine. However, the aging machine, the foundation's work, the parts are deformed and wear out, and slight changes in its dynamic properties appear. The shafts are out of alignment, the bearings are worn out, the rotors are unbalanced, the belts are relaxed and the clearance increases. All these factors result in an increase in vibratory energy that excites the resonances and adds considerable dynamic load to the bearings.

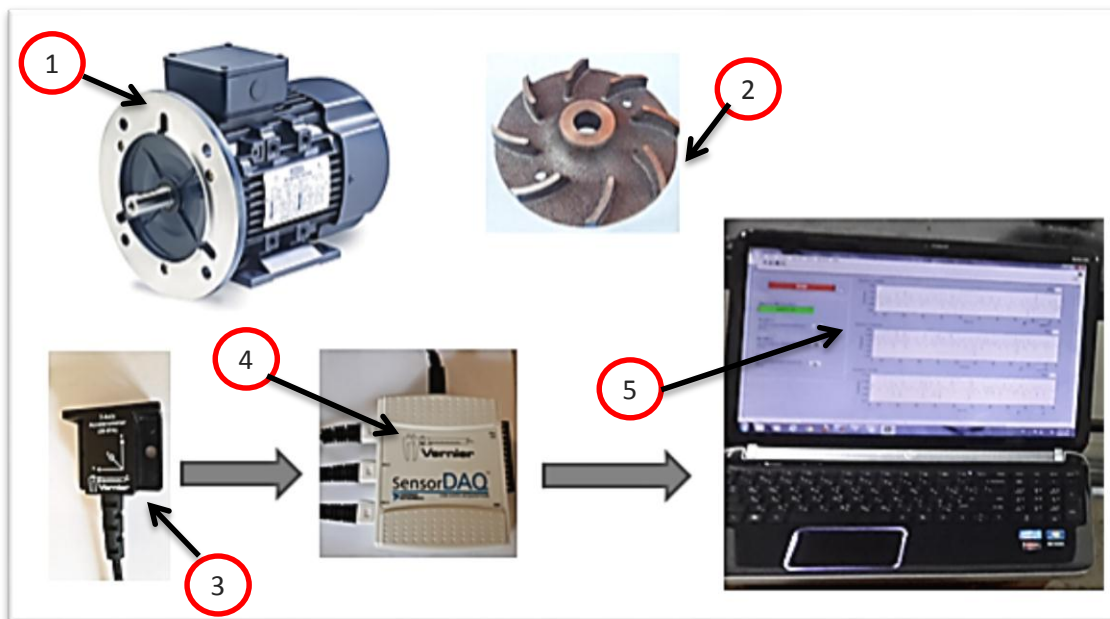
The vibrations collected during the measurement campaigns carry information which characterizes the operating state of certain mechanical components constituting the machine analyzed. It is thanks to the analysis of these vibrations that it is possible to detect defective components and possibly to locate them. When a certain threshold (corresponding to a fixed limit vibration level) is reached, it is possible to estimate the residual life of the component under the given operating conditions from knowledge of the damage laws [42].

All the machines in operation process produce vibrations, images of the dynamic forces generated by the moving parts, thus, a new machine in excellent working condition produces very little vibration.

The deterioration of the operation generally leads to an increase in the level of the vibrations, by observing the evolution of this level; it is, therefore, possible to obtain very useful information on the condition of the machine. These vibrations occupy a privileged place among the parameters to be taken into consideration for robust and logical prognosis; the modification of the vibration of a machine is often the first physical manifestation of an anomaly, a potential cause of degradation or even breakdowns. These characteristics make monitoring by vibration analysis an indispensable tool for modern maintenance, since it makes it possible, through appropriate faults diagnosis and prognosis, to avoid unexpected accidents that would disrupt the production process [43].

The corrosion of the pump impeller is generally caused by fatigue, it progresses at each loading, starting from an initial point almost always at the foot of the tooth, it appears mainly on fine steels, hardened by heat treatment, which is very sensitive to concentrations the appearance of these cracks is the consequence of a stress at the tooth's foot which exceeds the fatigue limit of the material and is generally located on the side of the tooth stressed in tension, as well as due to the quality of liquids entering it, such as lime water [44].

To simulate the pump aging, we scratch multi times the Impeller of the pump and measure each time the generated vibration via an accelerometer fixed on the pump chassis and a data acquisition system connected to Labview software (Figure 1.19.a)



Figure

1.19.a: External Data Acquisition System

- 1- Pump
- 2- Impeller
- 3- Tri-Accelerometer

- 4- NI Data Acquisition
- 5- Labview Instrumentation Programming Code

Figure 1.19.b shows a signal representative of the acceleration dimension z , where the x -axis is the sample index and y -axis is the amplitude of z acceleration signal. This procedure (introducing more scratches) is repeated many times. A database with normal and faulty observations is analyzed. This database is composed of 10 observations, each observation is 5×10^4 points with the sampling frequency, $SF=500$ samples/sec. A pre-processing step including filtering, normalization, and smoothing is applied to the data before processing.

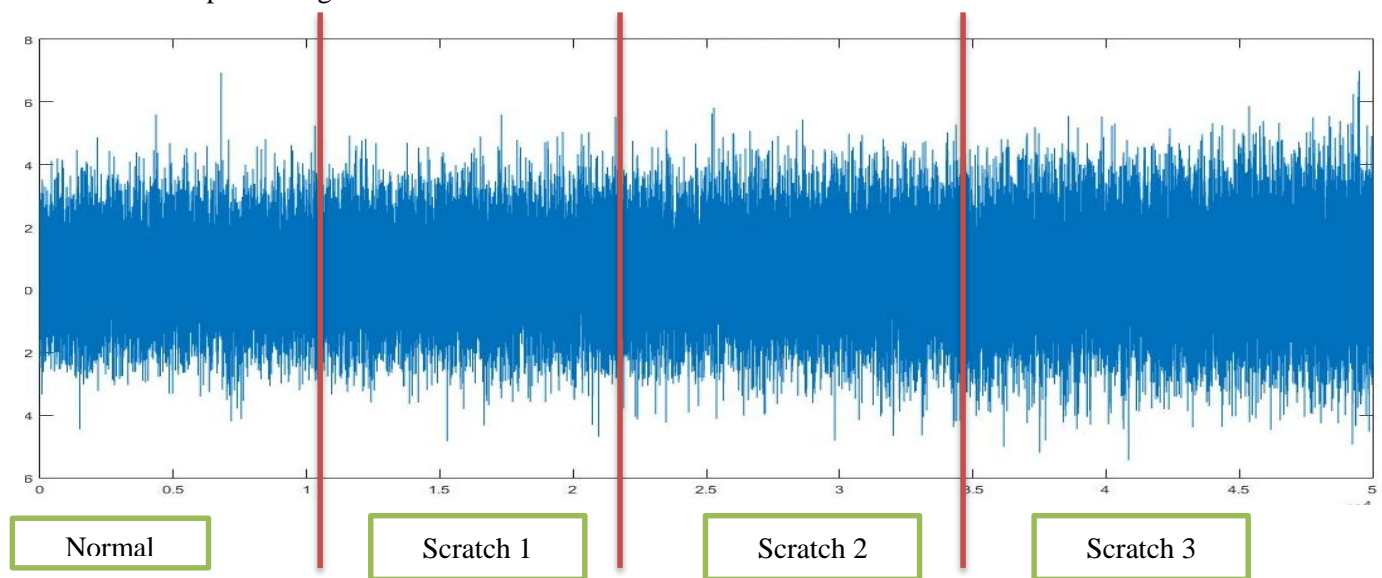


Figure 1.19.b: Graphical Representation of the Accelerometer Signal (Normal & Fault Mode)

Total data used in this research is 86 data observation. Where 70 % of them are used for training and the 30% remaining is used for testing and checking.

1.7 Determination of the Representation Space

1.7.1 Data Pre-processing

In general, the data acquired from the machine are noisy and redundant. Therefore, this data cannot be applied directly by a diagnosis or prognostic model. Further, this original data could be hidden inside many of the relevant information which can denote to the machine fault or machine degradation. Wherefore a set of parameters that plain of a relevant information should be extracted from this original data as indicators for this fault or this degradation. Usually, the strength of the diagnosis or prognostic model depends on the quality of the extracted and selected features. In addition, it is very important to identify the features that reflect the type of a fault and the progression of the failure of the machine [45][46].

A pattern is an observation made about the process. It is characterized by a set of d parameters (or features), and represented by a point in the dimension space d , defined by the different parameters called representation space. Since the parameters are often real numbers, a form ' i ' can be defined by a vector $X_i = [x_{i1}, x_{i2}, \dots, x_{id}]$ called form vector. If we place the problem in the context of the diagnosis, the parameters of the pattern vector reflect the state of the studied system. They come from analyzes performed on the signals measured by the sensors installed on the system (vibrations, speed, currents or even voltages for example). The typical patterns (or prototypes) are representative points of this space, and the problem of recognition consists in associating an observed pattern with a known standard pattern.

Due to disturbances (measurement noise, sensor accuracy ...), a new observation will rarely be identical to one of the prototypes. Thus, in order to express the influence of noise, the classes ($\omega_1, \omega_2, \dots, \omega_c, \dots, \omega_M$) correspond to zones in space, grouping the similar patterns (Figure 1.20).

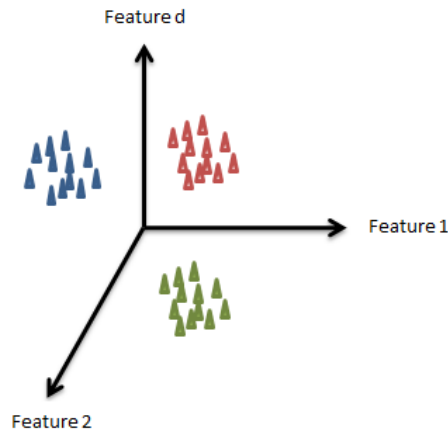


Figure: 1.20: Class Concept in Pattern Recognition.

The principle of recognition is to know which class, among M known classes, to associate a new form, $X_i = [x_{i1}, x_{i2}, \dots, x_{id}]$ observed. In terms of diagnosis, the classes correspond to the known modes of operation. They are our initial data set, called learning set and noted \mathbf{Xa} . To classify a new observation is to identify one of these modes. The development of a diagnostic system based on neural network takes place in three phases: a perception phase, an analysis phase and exploitation phase (Figure 1.21). The perception phase is the main source of information about the system. It is not only reserved for pattern recognition because it is common to other diagnostic approaches. It consists of two stages. A data acquisition step which consists in determining the hardware configuration (the type, the number of sensors to be used and the sampling rate, etc.) that they are necessary for the collection of signals on the studied system. The acquired signals must provide useful information in order to judge the operating state in which the system is located. This first step is followed by a signal preprocessing phase (filtering, de-noising, etc.). The analysis phase is to study the information provided by the sensors installed on the system. If the information is in the form of signals, then it is necessary to extract features (or parameters) digital. These parameters, which moreover constitute the pattern vector, must be able to describe the behavior of the

system. From this analysis phase, it must also get the precise definition of classes that represent the different operating modes. We then have a set of N observations distributed in M classes. This is the learning set. The observations of a class then represent the prototypes of this class.

A classification procedure is then applied to the learning set to establish boundaries between the different classes. This procedure will define a rule to assign or not a new observation to one of the known classes during the operation phase.

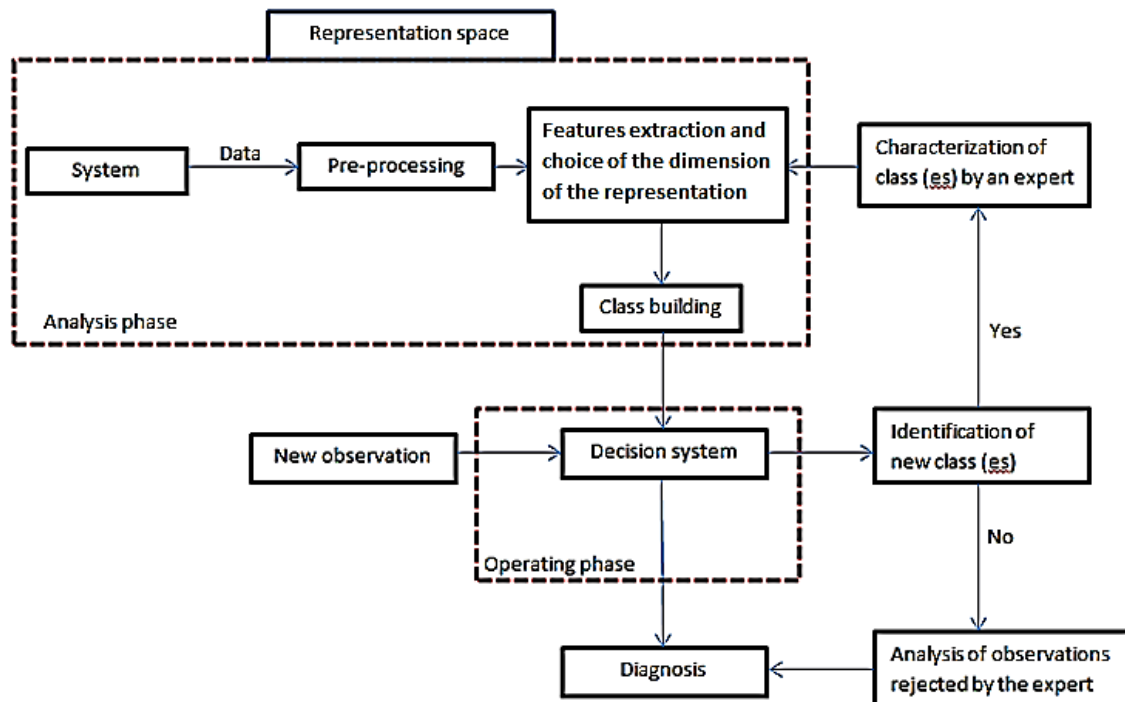


Figure 1.21: Schematic Diagram of the Diagnostic Procedure

The analysis phase is heavy in terms of calculation and often requires all the knowledge of the system studied to find the appropriate parameters for the appropriate treatment methods.

The operating phase (decision phase) allows associating a new unknown observation X , collected on the system to one of the classes defined during the classification phase by applying the associated decision rule. The proper use of the decision-making system depends on the relevance of the form vector and the performance of the decision rule [47].

1.7.2 Data Reduction

Fault diagnosis and failure prognostic of critical dynamic systems, such as aircraft and industrial processes, rely on degradation or fatigue models and measurements typically acquired online in real-time. Such sensor data must be pre-processed in order to remove artifacts and improve the signal-to-noise ratio.

Furthermore, they must be processed appropriately so that useful information in compact form can be extracted and used to detect incipient failures [48]. In the chemical process field, many analytical or measuring instruments can easily acquire values of many process variables in a very short period of time. In this way, one has to face multidimensionality problems. The multidimensionality complicates the data interpretation, increases the complexity of the fault diagnosis and detection in real time. Therefore, the first step in chemical process faults diagnosis and detection is to reduce the data dimensionality [49]. This step consists of constructing the pattern vector from the measurements made on the physical system or from the information collected during an observation of a phenomenon on the system. As mentioned before these measures are not all as informative, they may be noise, may be insignificant, correlated or redundant, aberrant or simply unworkable. Therefore, a parameter or feature generation step is required, i.e., producing a set of d parameters from the acquired signals using signal processing or data analysis techniques. The fundamental objective is to accentuate the important information of the acquired signal. This involves a transformation of the vector representing the time domain signal into a domain where the information contained in the signal will be better represented. These parameters are chosen to optimize the discrimination of the operating modes. The intervention of an expert to monitor the process is often very useful to guide this process.

The resulting observations are thus grouped in a numerical table of dimension ($N \times d$) which corresponds to a set of N vector forms characterized by values of d quantitative parameters. N forms (X_1, \dots, X_N) collected on the system are the training set. The performance of the diagnostic system will depend on the relevance of the calculated parameters. It is therefore preferable to have parameters varying, significantly, according to the different operating modes of the system.

In fact, the main objectives of reducing the dimension of data are:

- Facilitate visualization and understanding of the data.
- Reduce the storage space required.
- Reducing the training time.
- Identify relevant factors.

The reduction of the dimension of the pattern vector consists in looking for a subset of d' features ($d' < d$), which preserves as best as possible the separation of the classes of the initial learning set. This reduction of the representation space can be performed either by feature extraction methods [50].

Extraction consists of defining new features from the initial features. From the diagnostic point of view, the extraction methods do not decrease the number of features to be calculated, which remain at the number of d (the new features are linear combinations of the old ones), but on the other hand, the class representation space is of smaller size, which has the effect of speeding up the decision phase.

1.7.2.1 Data Reduction Based on Feature Extraction

Reducing the dimensionality of the data (also called feature extraction) is done by the construction of new features obtained by combining the initial characteristics. A data transformation risks losing the semantics of the initial set of characteristics and therefore the use of this family of methods is applicable only in the case where the semantics no longer intervenes in the steps that follow the reduction. So the parameter extraction aims to transform the data from the temporal space to another space where the parameter X is represented by a tell function $Y = f(X)$ (projection space) (Figure 1.22)

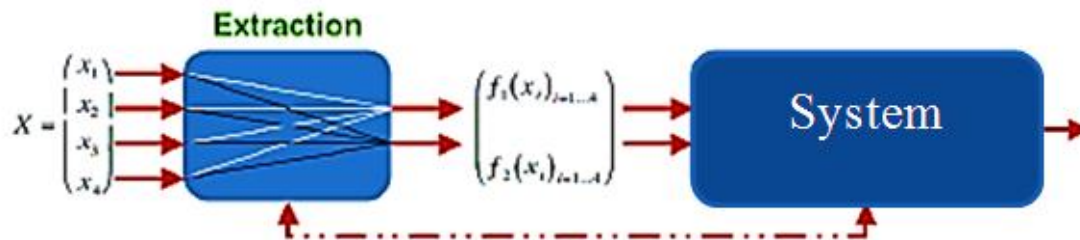


Figure 1.22: Principle of Feature Extraction

The extraction of signal parameters is an essential step before classification. It is necessary to extract the relevant, discriminant and most adapted parameters to the signal. The parameters related to a random signal are generally statistical or frequency. The statistical parameters are calculated from the probability density. The frequency parameters depend on the power spectral density. Other important parameters can be extracted from wavelet decomposition [51].

Signal processing based on features extraction combines a set of techniques for creating, analyzing and transforming input signals to extract fault-indicating parameters [52] [53]. As shown in Figure 1.23, it is possible to classify features extraction techniques into Time Analysis, Frequency Analysis, and Time-Frequency Analysis.

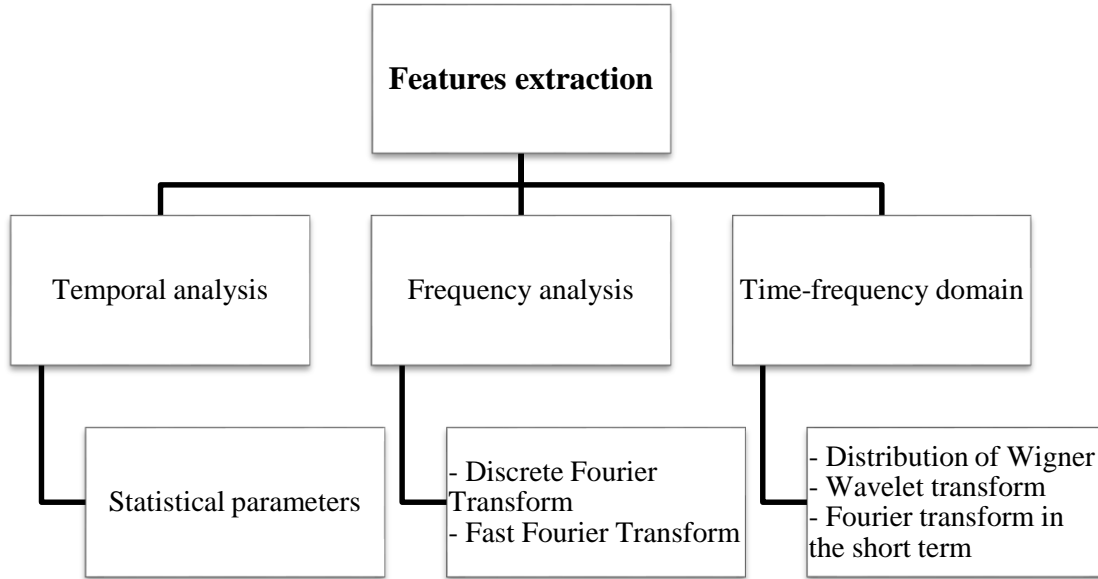


Figure: 1.23: Feature Extraction Techniques.

1- Temporal analysis

The temporal analysis makes it possible to extract fault-indicating parameters from raw data of the sensor. The parameters described here are called "statistical parameters" because they are based on analysis of the temporal characteristics of the recorded signal. For example:

- The mean value (eq.1.4): The average value denoted (μ) of a signal on a data sample window is a significant parameter for almost every type of sensor. It is defined by:

$$\mu = \frac{1}{n} \sum_{i=0}^{n-1} x_i \quad (1.4)$$

- The Envelope (env) Analysis: Fault diagnosis at an early stage, it can be determined reliably and quickly the shock repetition frequencies. We can look for the average and the variance of the envelope (eq.1.5) [54].

$$\text{env}(t) = \sqrt{|X(t)|^2 + |\tilde{X}(t)|^2} \quad (1.5)$$

Variance (σ^2) (eq. 1.6): The variance is a measure used to characterize the dispersion of a distribution or sample. It is defined by:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=0}^{n-1} (x_i - \mu)^2 \quad (1.6)$$

This parameter is often used as a metric base for classifiers such as dynamic Bayesian networks and neural networks.

- The RMS value (eq. 1.7): the RMS (Root Mean Square) value of a signal is the square root of the second order moment (or variance) of the signal:

$$RMS = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (1.7)$$

This is one of the most used parameters in time analysis. An increase in the value of the RMS will indicate a deterioration of the state of health of the system.

- Kurtosis: The Kurtosis noted S_{kurt} (eq. 1.8) represents the static moment of order 4. It measures the degree of crushing of the distribution of the recorded vibratory signal and is defined as being the ratio between the four-centered moment centered and the square of the variance

$$S_{kurt} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{(\sigma^2)^2} \quad (1.8)$$

A system in good condition generates a vibratory signal with a Kurtosis close to 3. For a degraded system, the amplitude of the signal is modified and Kurtosis becomes greater than or equal to 4.

- The crest factor (Fc) (eq. 1.9): the crest factor noted Fc is a characteristic measurement of a vibratory signal. This is the ratio between the amplitude of the peak of the signal and the RMS value of the signal. It is defined by:

$$F_C = \frac{|X|_{peak}}{X_{RMS}} \quad (1.9)$$

A system in good condition generates a vibration signal of low amplitude, both peak value and effective value. The crest factor remains low (between 2 and 6). A localized fault generates a vibration of high peak amplitude and low effective amplitude, so a large crest factor (> 6). The major defect of this parameter and Kurtosis is that they indicate about the same values in the new state and end of life of the system [55].

- Skewness: The Skewness noted S_{kew} (eq. 1.10) represents the static moment of order 3 centered on the cube of the standard deviation. It measures the symmetry of the distribution, or more precisely the lack of symmetry. A distribution is symmetrical if it has the same pace on both sides of the signal. It is defined as follows:

$$S_{kew} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{(\sigma)^3} \quad (1.10)$$

If the standard deviation σ is equal to 0, the distribution is symmetrical. If σ is smaller than 0, the distribution is asymmetric to the left. If σ is greater than 0, the distribution is asymmetric on the right.

The parameters described above all have the same drawback; they indicate the degradation of the system but fail to identify the fault responsible for this degradation [56].

2- Frequency analysis

Frequency spectrum analysis of a signal is the most commonly used technique for identifying faults in a system. This technique is based on the fact that a localized fault generates a periodic signal with a unique characteristic frequency. In contrast to time analysis, frequency analysis identifies the fault present in the system by identifying its characteristic frequency. This technique is generally applied during the steady state of the system [57]. A classic among the techniques used in the frequency domain is Fast Fourier Transform (FFT) Fourier transform spectral analysis. The spectrogram technique makes it possible to perform frequency analysis of the signals in the dynamic operating mode of the system. This technique consists in performing a repetitive calculation of the FFT on a sliding time window, which makes this technique sensitive to the length of the window, the type of windowing, the total duration of supervision and the sliding step of the window. Although this technique makes it possible to analyze signals in dynamic mode, the speed of the dynamic regime, which is in the case of asynchronous machines of the order of 150ms, significantly reduces the efficiency of this technique.

3- Time-frequency analysis

The time-frequency analysis of the signals deals with both the time domain and the frequency domain. Non-stationary signals are better represented by a time-frequency distribution, which aims to show the distribution of the signal energy over the two-dimensional time-frequency space. The most widely used techniques for time-frequency analysis are the Short-Time Fourier Transform (STFT), the Wigner-Ville distribution and the wavelet transforms [58].

1.7.2.2 Data Reduction Based on Features Selection

1.7.2.2.1 Introduction

Automatic classification or clustering technologies are some of the most used methods of data analysis and data mining. Despite their undisputed success in exploratory data analysis, clustering techniques have to adapt to ever larger volumes of data. Indeed, as storage technologies have evolved, the volume of available data has gradually exploded in numbers of individuals but also in a number of descriptors. So we are very often faced with the problem of the curse of dimensionality.

There are generally two types of approaches that can be used continuously: feature extraction (consists in constructing new attributes from the set of original variables) then feature selection (allows to keep only a relevant subset) variables). In this part, we are interested in the feature selection approach that allows us to select the most representative variables of the problem [59].

1.7.2.2.2 Problematic

Feature selection plays a very important role in classification when a large number of features are available, some of which may be insignificant, correlated or irrelevant [60]. It consists of selecting a subset of features. The selection also facilitates the learning step and reduces the complexity of the algorithms as well as the calculation times [61]. Unsupervised selection of features remains a difficult and time-consuming problem. When dealing with large databases, which cannot be manually tagged, it is desirable to have quick and efficient selection methods.

The selection of attributes or features is a research topic that has been active for several decades and is currently being developed in various applications [62]. It is an important step in the preprocessing of large data routed to supervised or unsupervised classification. Indeed the appearance of large databases in the field of learning and data mining systems "Data Mining" required a reduction in size, before starting the task of data classification.

Feature selection is a process of looking across the set of available explanatory variables for an optimal set of the most important characteristics of a given system [63].

1.7.2.2.3 Feature Selection Definition

Feature selection is a method of choosing an optimal subset of relevant variables from an original set of variables, based on a certain performance criterion. In fact, the choice of an optimal set of descriptors does not necessarily mean the selection of a set composed only of the variables deemed relevant and useful. It may contain irrelevant but better-performing features taken with other variables. Thus the feature selection procedure tries to select the smallest subset according to two main criteria:

- The accuracy of the classification does not deteriorate
- Class distribution is close to the original distribution.

The notion of relevance of a subset of features always depends on the objectives and In general, the problem of Features selection can be defined by:

Let $F = \{f_1; f_2...; f_N\}$ a set of features of size N or N represents the total number of features studied. Let Ev be a function that evaluates a subset of features. We assume that the largest value of Ev (eq.1.11) is obtained for the best subset of features. The objective of the selection is to find a subset $F' (F' \subseteq F)$ of size $N' (N' \leq N)$ such that:

$$Ev(F') = \max_{Z \subseteq F} Ev(Z) \quad (1.11)$$

Or $|Z| = N'$ and N' is either a number to be warned by the user or is controlled by one of the subset generation methods.

Feature selection consists in choosing relevant features from the measurement space (Figure 1.24) [64]

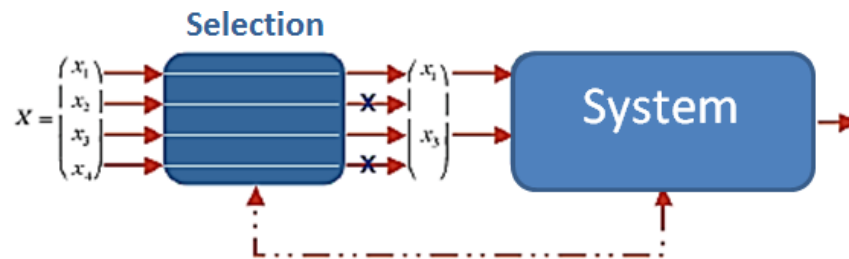


Figure: 1.24: Principle of Feature Selection.

Example 1.1:

If we have a vector P contain 10 elements a_i of features $P = [a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}]$, So $F = 10$. According to the selection if we obtain a vector P' of 6 elements thus: $P' = [a_1, a_4, a_5, a_7, a_9, a_{10}]$, $F' = 6$ and $Z = 6$

A general procedure proposed in 1997 by Dash et al. [63][64] for a feature selection method is illustrated in Figure 1.25.

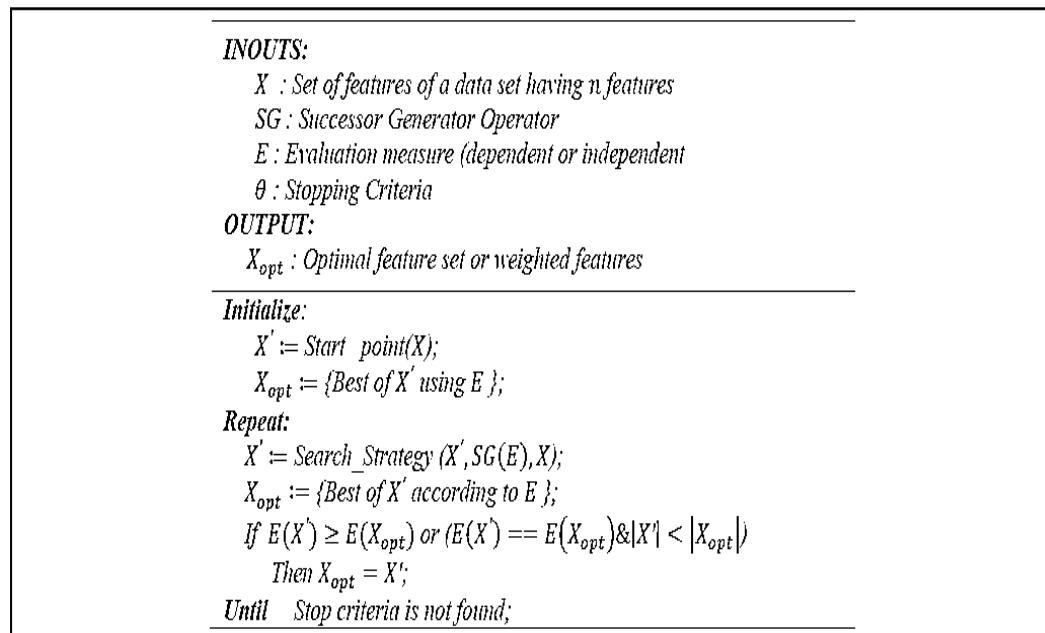


Figure 1.25: General Algorithm for Feature Selection [65]

1.8 Maintenance of Complex Systems

1.8.1 Introduction

The effectiveness of the maintenance of industrial systems is a major economic issue for commercial exploitation. Maintenance should improve the reliability, safety, and quality of industrial system at a lower cost. The main difficulties and sources of inefficiency are in the choice of maintenance actions. A maintenance action consists in replacing the failed equipment that is no longer capable of performing its function. A wrong choice of actions can lead to unsatisfactory maintenance and additional costs due to immobilization of the system. Optimizing maintenance involves reducing the downtime of the complex system by minimizing the duration of interventions and the number of maintenance actions. It is very difficult to determine a maintenance action for a distributed system. A distributed system is a complex system that results from an assembly of completely heterogeneous components (software, hardware). It is necessary to gradually monitor each component to be able to make a maintenance action decision for the overall system. Using new embedded technologies, it is possible to set up a supervisory system to monitor system components and detect online problems or failures that may occur in the system. It is then necessary to provide an online maintenance diagnosis that analyzes the different sources of observation and that identifies the failed equipment to replace. In order to improve the preventive maintenance of the distributed system, a prognostic function is integrated into this supervision system making it possible to program future maintenance phases.

1.8.2 Definition

The formalized concept of maintenance was born in the industry in the late 1970s. A definition of maintenance is given by the standard AFNOR NFX 13-306 [66]:

Definition 1: Maintenance is the set of actions that can maintain or restore a property in a specified state or able to provide a specific service.

According to this definition, maintenance maintains or restores a system to a previously specified state so that the system is able to provide the functions for which it was designed. The analysis of different forms of maintenance is based on three concepts.

- The events that cause the maintenance action: reference to a schedule, diagnostic result, sensor information, wear measurement, the occurrence of a failure, etc ...
- The maintenance methods associated with them: systematic preventive maintenance, conditional preventive maintenance, corrective maintenance.
- Maintenance operations: inspection, control, troubleshooting, repair, etc ...

These maintenance operations are performed on the equipment of the complex system whether hardware or software. Maintenance activities, in the sense of troubleshooting equipment, have always existed. They

essentially consist of repairing equipment when it is already out of order. Failed equipment is no longer able to perform the functions for which it was designed, it is said to be defective.

Definition 2 (Failure): A failure is a cessation of the ability of an entity to perform one or more required functions.

Two forms of failure can be identified depending on the degree of system degradation: partial failure and complete failure. Their definition is given below.

Definition 3 (Partial failure): A partial failure is a degradation of the ability of a system to perform required functions.

According to Zwingelstein et al. [67], a partial failure results from deviations of one or more system characteristics beyond the specified limits, so that the required functions do not completely disappear. When a failure causes the system functions to completely disappear, it is a complete failure.

Definition 4 (Complete Failure): A complete failure is represented by a cessation of the ability of a system to perform all the required functions.

Faulty equipment requires maintenance to be repaired. Troubleshooting activities do not include a preventive aspect. It is much more interesting to prevent a failure before it causes a failure of system equipment. Maintenance can also help rebuild and improve the system. For this, it must take into account many constraints such as quality, safety, environment, cost, etc.

1.8.3 New Evolutions of the Maintenance Function

Given the ever-increasing demands, maintenance costs have increased rapidly in recent years. For example, it is estimated that maintenance costs in the United States were \$ 200 billion in 1979 and that they grew in the range of 10% to 15% in the years that followed [68]. An important part of this cost of maintenance could, however, be avoided: a bad planning results in wastage in maintenance over time, and that possibly on equipment which does not have a big role in the continuity of the production. This increased cost alone does not justify the need to challenge traditional approaches to maintenance. Firstly, production systems are constantly evolving and new production techniques have emerged; in particular because of automation (machines can ensure production without human intervention).

Secondly, companies are more interested in rapidly adapting the quantity and quality of production according to the variation in customer demand, which requires a high level of flexibility for industrial equipment. Therefore, even if the maintenance activity is now considered a separate activity, companies are more reluctant to outsource to benefit from the strong skills trades service providers [69]. This evolution is largely due to the development of science, information and communication technologies. But before the development of maintenance architectures to reduce the distance between actors, it is the maintenance strategies themselves that evolve. Indeed, maintainers today want to go beyond static maintenance (without

anticipating the evolution of the state of equipment) and implement more "dynamic" maintenance strategies.

1.8.4 Navigating towards an Anticipation of the Phenomena of Failure.

Before the 1960s, the main mission of the maintenance department of a company was to intervene on broken equipment in order to repair them as soon as possible. This type of maintenance, called corrective maintenance, was then gradually completed by a more anticipatory approach to the phenomena of failure that is to say by a maintenance performed before the failure occurred. This second form of maintenance, called preventive, was initially implemented during the development of the Boeing 747 in 1960 [68]. These two main types of maintenance, corrective and preventive, have some variants explained below. Figure 1.26 gives an overall articulation [27].

There are two main maintenance families: Corrective Maintenance and Preventive Maintenance. The corrective maintenance is the one that the system undergoes when the failure is already present and that it must be repaired. Preventive maintenance is the anticipation of fault to prevent failures. A classification of maintenance strategies can be found in [70][71].

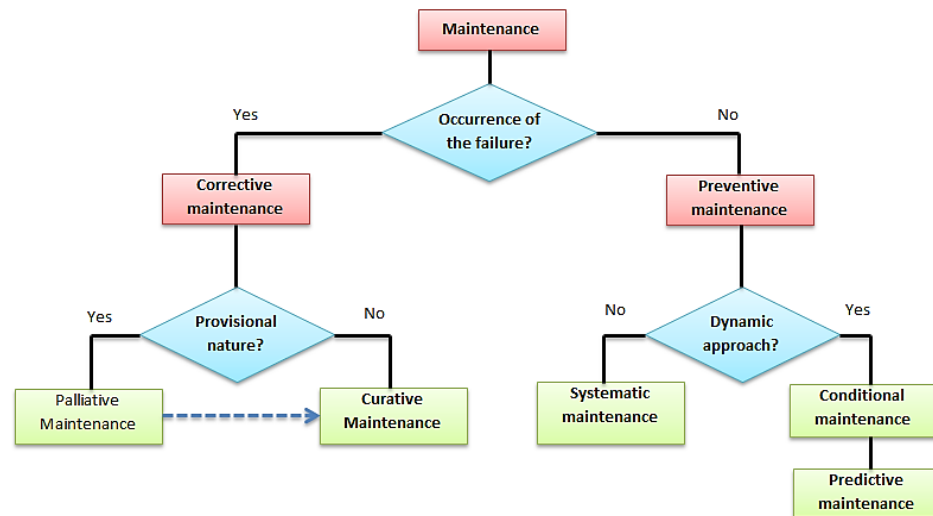


Figure 1.26: Maintenance Forms [27]

1.8.4.1 Corrective Maintenance

AFNOR [72] defines corrective maintenance as maintenance performed after detection of a failure and intended to return a machine to a state in which it can perform a required function. Corrective maintenance is generally adopted for equipment for which:

- The consequences of the breakdown are not critical,
- The repair is easy and does not require a lot of time,

- Investment costs are low.

Two forms of corrective maintenance can be distinguished. When the maintenance intervention is temporary, it is called palliative maintenance. If the work is definitive, it is called curative maintenance [73][74].

1- Curative Maintenance

This type of maintenance makes it possible to definitively restore the system after the occurrence of a failure. This system repair is a durable repair. The repaired equipment must perform the functions for which it was designed. A repair is a definitive curative maintenance operation that can be decided either immediately following a failure or after a repair (see next paragraph). It, therefore, causes an unavailability of the system [71].

2- Palliative Maintenance

Palliative Maintenance is temporary, provisional. It is mainly made up of operations that will, however, have to be followed by curative operations (repairs). Troubleshooting is a palliative maintenance operation that is intended to return the system to a provisional state of operation so that it can perform some of the required functions. Troubleshooting operations are often short-lived and can be numerous. Because they happen often, they are also very expensive [71].

1.8.4.2 Preventive Maintenance

This type of maintenance performed according to predetermined criteria aims to prevent failures during operation of the system [75]. The preventive aspect is important for reasons of operational safety but also for economic and sometimes practical reasons (the equipment is available for maintenance only at certain times). The purpose of preventive maintenance is to eliminate the causes of serious accidents by reducing the likelihood of system failure or degradation of system equipment. It aims to reduce the costs of breakdowns and corrective maintenance by minimizing or avoiding costly repairs and downtime with constant and preventive maintenance. There are three special forms of preventive maintenance: systematic preventive maintenance, conditional preventive maintenance and predictive preventive maintenance [76].

1- Systematic Preventive Maintenance

Systematic maintenance is developed in relation to a schedule established according to the operating time or the number of usage units [77]. Even if time is the most common unit of use, other units can be retained such as the number of flights for an airplane, the distance traveled, the number of cycles carried out, etc ... It consists of systematically replacing a number of previously defined equipment even if no failure has occurred. This is a scheduled maintenance. The frequency of maintenance operations is determined from commissioning and is essentially based on reliability data. This form of maintenance requires knowledge of the behavior of the equipment, the degradation modes (equipment wear) and the average or mean time of good operation between two failures of the system (MTBF) [77][76].

Systematic preventive maintenance ensures the periodic replacement of equipment, some parts of which are unusually worn. It also allows the replacement of equipment whose failure may cause serious accidents or equipment with a high cost of failure. This systematic method is quite expensive but it ensures a high level of security by setting a periodicity of the visit which reduces the risk of having a failure before the intervention.

2- Conditional Preventive Maintenance

Conditional preventive maintenance, also known as predictive maintenance, is subordinated to a predetermined type of event (diagnostic result, sensor data, wears measurement, etc.) indicative of the operating state of the system. It depends on experience but also uses real-time data that it analyzes to determine or predict a failure [78]. A deep knowledge of the equipment of the system makes it possible to be able to predict the failures by observing a certain number of parameters precursors of failure, as for example:

- a. Wear, visible in particular by dust, debris,
- b. The oxidation of parts,
- c. Loose electrical, mechanical or hydraulic connections,
- d. Abnormal, unusual vibrations,
- e. Leaks of fluids, compressed air,
- f. Unusual warm-up,
- g. The degraded results: the drift of the specifications of the parts, needs for frequent adjustments...

The measurements and parameters monitored are carefully determined to be representative of the operating state of the system. Whatever the technique used, the data collected or measured are always compared to a reference. The crossing of a predetermined threshold triggers an event (alarm) which makes it possible to decide on a maintenance operation on the system before the degradation causes a failure. This type of maintenance can be applied in the case where the MTBF of the system is not known. If the MTBF is known, it allows adjusting online time remaining until the next maintenance visit. This time depends directly on the monitoring of the precursor parameters of failure [79].

3- Predictive Preventive Maintenance

Predictive maintenance, also known as proactive maintenance, is also performed following an analysis of the monitored evolution of the precursor parameters of failure that qualify the operating state of the system. Proactive maintenance is a form of predictive maintenance that consists of determining the causes of failure and early wear of system equipment. Predictive maintenance makes it possible to anticipate and predict at best the moment when the maintenance operation will have to be carried out.

This form of maintenance reduces the number of unforeseen failures, and thus the unavailability of the system. It allows planning of maintenance operations in order to use the equipment to the maximum of their possibilities. By monitoring equipment, it is possible to correct driving errors or anomalies that can lead to

more serious failures later and improve safety by avoiding critical accidents. On the other hand, this form of maintenance requires setting up monitoring and measurement techniques that can be very expensive [80].

4- Operations

Monitoring operations (inspections, visits, checks) are necessary to control the evolution of the operating state of the system. They are performed continuously or at predetermined intervals calculated on the operating time or the number of using units.

- Inspections consist of periodically recording faults and performing simple adjustments that do not require system shutdown.
- Visits are surveillance operations which, as part of routine preventive maintenance, are carried out according to a specified periodicity. They correspond to a list of defined operations that can lead to system unavailability.
- Controls are compliance checks against pre-established data followed by a decision.

The revision corresponds to all the preventive maintenance operations carried out to avoid any major or critical system failure for a specified time or for a number of uses [75].

1.8.4.2.1 Implementation of a Conditional Preventive Maintenance (CBM)

The need for predictive maintenance has become unanimous and is of growing interest in the scientific community. A large number of works have recently emerged to propose comprehensive forecasting architectures that integrate traditional maintenance activities such as monitoring, diagnostics, prognostics and maintenance logistics [81][82][83][84][85].

In these works, the differences appear only at a relatively small level of detail and focus on the form of software architectures to implement (local, distributed, modular). , etc... Overall, the architecture proposed in 2009 by Medjaher et al. [86] and distributed by the MIMOSA group [m] is unifying. This is OSA / CBM (Open System Architecture for Condition Based Maintenance). This architecture consists of 7 functional layers that can be considered as sequential (Figure 1.27).

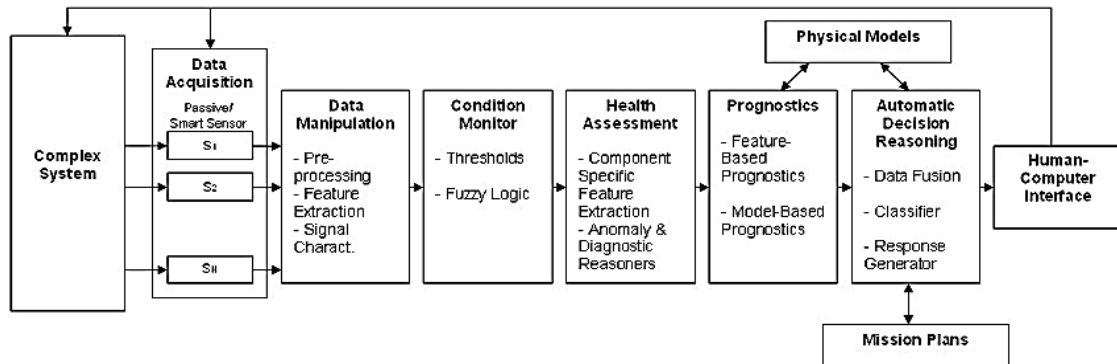


Figure 1.27: CBM Architecture [81]

- **Layer 1** Sensor module: This module provides the system with digital data from sensors or transducers.
- **Layer 2** Signal Processing Module: This module receives the data from sensors or transducers or other signal processors and performs signal transformations and extractions of features or descriptors.
- **Layer 3** Monitoring Module. The monitoring module compares the online data with some expected or known values; it must also be able to generate alerts according to previously set thresholds.
- **Layer 4** diagnostic module. This module determines whether the state of the system, subsystem or monitored component is degraded or not and suggests the likely failures.
- **Layer 5** prognostics module. This unit predicts the future state of the system, subsystem or component monitored. The module is based on the data from the previous modules.
- **Layer 6** Decision module: Its main function is to recommend maintenance actions or other alternatives to continue to operate the system until the completion of his mission.
- **Layer 7** Presentation module. This module receives information from all the previous modules. It can be built as an HMI (Human Machine Interface).

In summary, the anticipation of the failures necessary for a predictive maintenance strategy can only be achieved if the degradation phenomena are correctly understood (data acquisition, extraction of descriptors, detection/monitoring, diagnosis). Also, the prognostic is not an end in itself. On the other hand, it must make it possible to implement adequate reaction policies (decision support).

1.9 Conclusion

This chapter presents statistical studies on industrial accidents, in particular accidents related to chemical industries due to the extent of their moral and material gravity, whether at the level of workers in the industry or the society as a whole.

The distillation column is one of the most widely used equipment in the chemical industry. In this chapter, we have described the distillation column used in the industrial plants and explained the related methods of operation. This chapter also illustrates the most frequent accidents that occur in the distillation column.

To open the horizons on the development of effective methods able to ensure effective monitoring and prevent failure or possible accidents in the continuous distillation process, it was necessary to explain the nature of the data can be extracted from the distillation column system.

To achieve a high-precision diagnosis and prognostic of this catastrophic accidents occur in the distillation column, the pre-processing of these data (filtering, relevant features extraction, and features selection) should be thoroughly studied. The following chapter presents the main concepts of faults diagnosing, and their associated methods.



Normandie Université

Chapter 2

Industrial Diagnosis Methods Applied on Chemical Processes (nonlinear systems)-Application to Distillation Column

2.1 Introduction

The chemical industry is one of the pillars of the global economy, but in recent years it has faced an unflattering image of a dangerous and polluting industry. The history of hazardous chemical accidents shows that this industry remains one of the major sources of serious incidents that are relatively more likely than previously thought. Some recent examples of such accidents around the world that have caused major loss of life and major damage to property and the environment include the explosion of the four reactors at the Fukushima-Daiichi nuclear power plant in the North East of Japan in 2011, the destruction of the oxidation unit of Flixborough FP Less plant. (1996) [16], the toxic cloud of dioxin dispersed in the municipality of Seveso, the 4000 victims and 200000 wounded in the disaster of Bhopal or the explosion of the AZF plant, ... J.P. Gupta. (2005) [17] other major accidents can be cited such as Mississauga in Ontario, St-Basile-le-Grand in Quebec, Chernobyl in Ukraine, Sandoz in Basel, the Feyzin refinery in Lyon, Metz silos and Blaye, etc. so many images that remain engraved in the memory of those who have lived but also all security actors who still seek to draw the lessons and avoid the mistakes of the past J.L. Gustin. (2002) [18] T.A. Kletz (2006) [87].

Many industrial accidents have been caused by reactions whose implementation has not been controlled: thermal runaway, uncontrolled secondary reaction ... These consequences are often important insofar as an unwanted chemical reaction is likely, following the reagents used, to give substance to both an explosion and the emission of toxic or flammable products in the environment. The main causes and reasons identified for the development of uncontrolled reactions in the industry are attributable to a fault in the design or operation of the facilities and manufacturing processes. These accidents demonstrated how the incidents caused by this type of industry could be impressive and destructive. Faced with these events, it became necessary to rest the question of chemical risk by challenging industrial practices and identifying new technological issues (diagnosis, operating safety, surveillance ...). This development of safety aspects coincides with a change in the social context and the emergence of new themes such as ecology and environmental protection, which are today part of sustainable development [87]. At this time, this awareness of the public leads to less tolerance of the impact of industries on man and his environment (effluent discharge, waste, noise pollution, etc.). The phenomena involved in the chemical process must be characterized in order to ensure that it is best suited, in terms of the design of the facility that the choice of procedure. In a chemical process, non-linearity and the sensitivity of certain parameters influence the progress of the chemical reaction. A small variation of these parameters can change enormously during the reaction and could even cause an accident [88].

In the case of a chemical reactor, the main reasons that lead to its monitoring are, firstly, to maintain its optimal driving (online operation), and secondly, its maintenance. The driving support for a process gives the operator the tools to share decision-making to make it work the best (maximum production, security, non-degradation of equipment) [26].

This driver assistance requires monitoring process to detect any malfunctions and identify the best possible causes. Maintenance which aims to replace or repair worn or defective equipment is performed

most often offline. It is important to note that if the operation and maintenance are operations that are part of the time in different ways, they involve monitoring must be online for the two goals In this light concerning the operational safety, it is justified to be able to determine in real time the occurrence of a malfunction during the implementation of a chemical process in order to be able to effectively remedy the problem of detection of chemical failures [89]. In the field of chemical process surveillance, the industrial tools used are mainly based on hardware redundancy and thresholding metrics for the surveillance of the operating parameters and the conduct of the process. This aspect is known in the industry under the name "alarm management". The issues of detection and localization of faults and means of reconfiguration (aspect FDI (Fault Detection and Isolation) and FTC (Fault Tolerant Control)) remain open [14].

In recent years, alarm filtering and detecting problems and real-time location of faults are addressed in the framework of European projects Ould Bouamama B. (2002) [90, p. 1]E. Craye & al. (1998) [90] or by the industrial toolbox implementation processes applied in pilot Bouamama & B. Ould al. (2006) [91]. However, these systems relate to energy processes: the phenomena of material processing are not considered. In this context, many approaches have been developed, for fault detection and diagnosis by the different auto research communities, chemical engineering (INERIS: National Institute for Industrial Environment and Risks in France [87]) and artificial intelligence (Chemical Hazards and Processes Laboratory, INSA Rouen), interested in the field of supervision, in particular, the diagnosis, the main purpose is the safety of chemical processes (avoid the risk of thermal runaway, the hazards of chemical reactions used and the impact of the main drift effects on chemical reactions considered) [92].

2.2 Diagnosis of Dynamic Systems

The objective of diagnosis is to determine at every moment the mode of operation of the system by its outward manifestations. It is based on a priori knowledge of the operating modes and instant knowledge embodied by a new case of system status. Its general principle is to compare the data collected during the actual operation of the system with the knowledge that one has of its normal or faulty operation. If the mode of operation is identified as faulty mode, the diagnostic system can locate the cause. The function of the diagnosis is therefore to seek causality linking the symptom, the fault, and its origin. We can also say, that the diagnosis is to locate faulty elements and identify the root causes of the problem, this by establishing a causal link between the symptoms and the offending items to replace. The following phase is the decision. Its role is to identify and initiate actions to bring the best the system in (to) a normal state. These actions can be of emergency orders or judgments launches repairs or preventive operations. If we want to avoid a loss of production, this decision may be a reconfiguration of the process [93].

2.3 Characteristics of a Diagnostic System

Based on Venkatasubramanian in 2003 [94], the set of desired features a diagnostic system should possess is:

- Rapid detection.

- Insulation: it is the ability to differentiate faults.
- Robustness vis-à-vis for some noises and uncertainties.
- Modeling conditions: for quick and easy deployment of real-time diagnostic classifiers, the modeling effort should be as minimal as possible.
- IT Implementation Facility (low complexity in the algorithms and their implementation) and storage capacity.
- Identification of multiple faults: for large processes, combinatorial enumeration of multiple faults is too large and they cannot be explored comprehensively.
- Identification novelty: one refers to the ability to decide if the process is a normal or abnormal condition. In the event of a fault must be identified if it is a known fault or a new default.
- Estimation of default classification error (diagnostic) for reliability.
- Adaptability-the diagnostic system should be adaptable to changing conditions the process (disturbances, environmental changes).
- Facility explanation of the fault origin and the spread of it. This is important for making a decision online.

2.4 The Different Diagnosis Steps.

The diagnosis is divided into several sub-tasks to provide a set of Boolean decisions on the presence of faults: detection of an anomaly, localization of the sub-systems that are defective, and characterization of a fault or failure (Figure.2.1) [47].

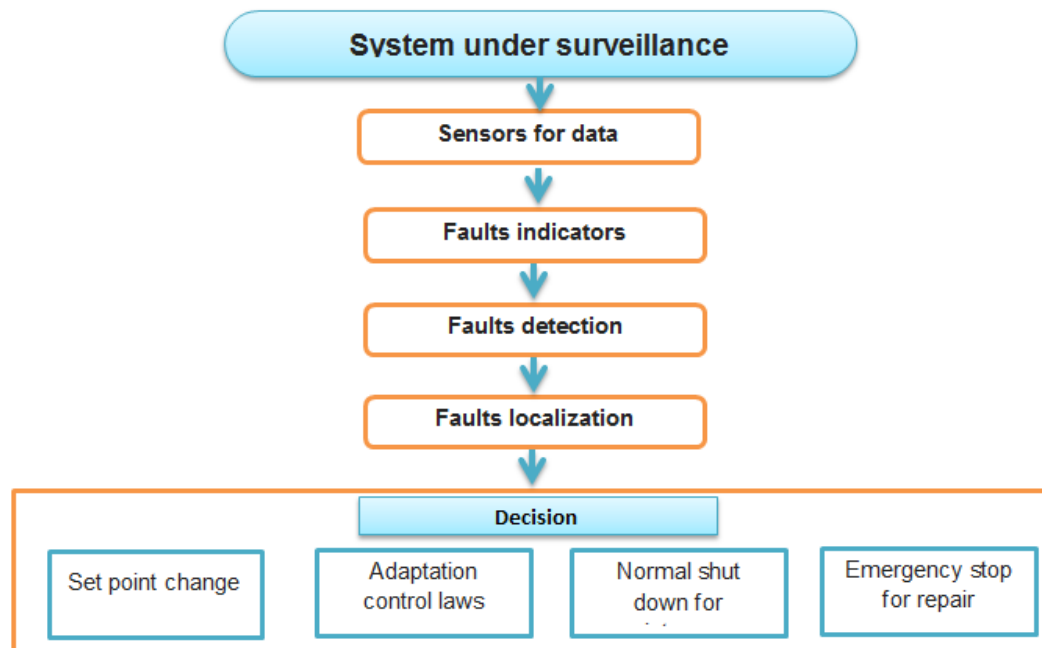


Figure 2.1: Diagnosis Steps

1- Data Acquisition Step.

The diagnostic procedure requires the availability of information on the functioning of the system to be monitored. This information is collected during a data acquisition phase followed by a validation.

This step involves the use of suitable sensors for measuring the variables of the process.

2- Faults Indicators Development Step

From measurements and observations conducted by the operators in charge of the installation, it is to build indicators to highlight any faults that may occur in the system. In the field of diagnostics, fault indicators are commonly referred to as residues or symptoms.

3- Detection Step

This step should allow the system to decide whether or not they are in a normal operating state. It is not enough to test the non-nullity of residues in determining the occurrence of a default because, in practice, the measured values are still marred by noise and the system monitor is always subject to disruption. Therefore, this step is usually to call statistical tests or, more simply, is performed using a threshold.

4- Localization Step

This is from the statistically non-zero residues, locate the fault, that is to say, or to determine the faulty elements. The location procedure requires the use of a set (or vector) residues, which must have properties to uniquely characterize each fault. To do this, two methods may be used:

- Construction of structured residues.
- Construction directional residues.

5- Decision Making Step

It is to decide what to do to maintain the desired performance of the system under surveillance. This decision should allow generating, possibly under the control of a human operator, corrective actions necessary to return to normal operation of the installation.

In short, regardless of the method used, the diagnostic procedure comprises two main steps, a residue generation step, and a tailings evaluation stage.

2.5. Classification of Diagnostic Methods

Studies conducted over the past fifteen years have led to the development of methodologies for the development of procedures tailored to the functional chemical facilities and security constraints (tubular reactors, continuous reactors, batch, and semi-batch ... etc.). However, reactors, even if they provide the characteristics of flexibility and versatility required, present a number of technological limitations. In particular, poor drainage conditions the heat from the chemical reactions are a serious security problem. So various factors such as system complexity, high dimension, non-linearity of the processes have often made very difficult accident risk detection (very) serious or major: fire and/or explosion, poisoning, chemical burns, pollution,...[95]. All these factors clearly demonstrate the need, always present, to develop techniques for the prevention of chemical risk and explain the growing interest in research on security

issues. This has contributed to the development and improvement of many methods and tools that are now recognized and widely used in the industrial world Coker [96].

Diagnostic methods differ according to different criteria: the dynamic process (discrete, continuous or hybrid), complexity, online diagnostic implementing and / or offline, the nature of information (qualitative and / or quantitative) , depth (structural, functional and / or temporal), distribution (centralized, decentralized or distributed) ... in this context, several classifications are proposed in the literature [Franck and Köppen-Seliger, 1997 [97]; Isermann, 2006 [98]; Dash and Venkatasubramanian, 2000 [99]; Venkatasubramanian et al, 2003 [94]]. These classifications are influenced by terminologies and specific contexts of each community and are not always consistent.

The objective is to determine the most appropriate method for the resolution of our problem detection and diagnosis, as well as position the class of diagnostic methods that interests us from the different methods of literature. We provide a non-exhaustive classification of diagnostic methods into two large families (Figure 2.2):

- 1- Methods without a mathematical model that does not require increased knowledge of the physical system, but the use of superficial knowledge.
- 2- The methods based on models that require the thorough knowledge of the physical system.

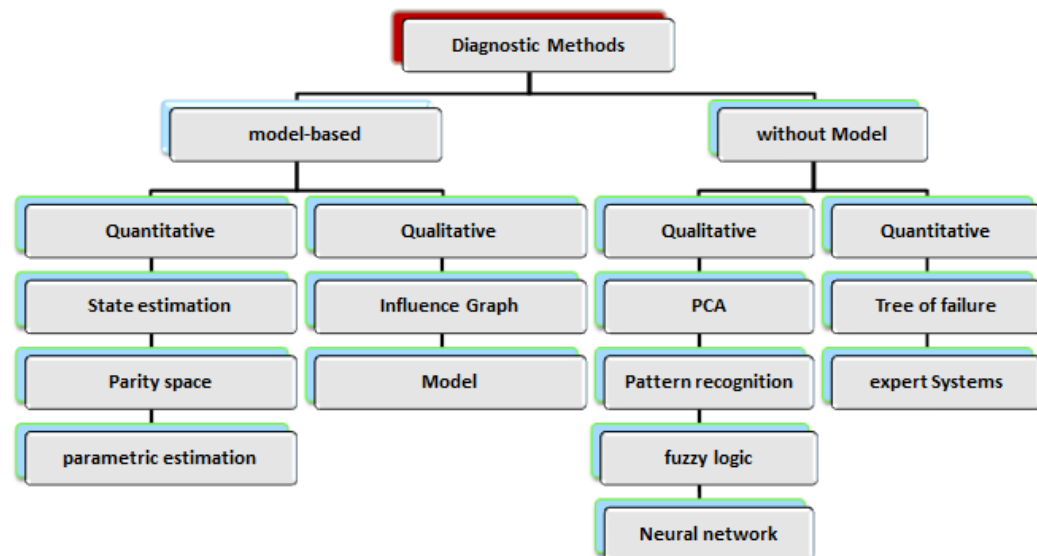


Figure 2.2: General Classification of Diagnostic Methods.

2.5.1 Model-based Methods

The use of models for diagnosis dates back to the early 1970s. Since many studies have been proposed by Willsky in 1976 [100], Chow et al. in 1984 [101], Basseville in 1988 [102], Patton et al. in 1989 [103];

Evsukoff et al. in 1997 [104]; Isermann and Ball in 1997 [29] and Fussel and Isermann in 1998 [105]. A comprehensive study of the model-based methods can be found in the study of Frank et al. in 1996 [106] or in recent books like the book of Patan in 2008 [107]; and the book of Chiang et al in 2000 [32]. These methods are alternatives to physical or hardware redundancy. The general structure of most of these methods is based on the idea of analytical redundancy [101]. The principle of model-based methods is to identify the gap between the real system and its model. These methods rely on explicit behavioral patterns of the system being diagnosed.

2.5.2 Model Free Methods

For industrial applications, the design of a mathematical model is difficult or impossible to obtain, because of the numerous reconfigurations involved in the production process or the complexity of the phenomena involved. In this case, it uses methods that do not require detailed knowledge of the process. Two classes in this type of approaches are possible [108]:

- Quantitative methods called also knowledge-based methods
- Qualitative methods or methods based on data processing.

2.5.2.1 Qualitative Methods

Qualitative methods consist of operating a symbolic knowledge base and require the existence of a wide range of historical data corresponding to the various modes of installation.

1- Principal Component Analysis: principal component analysis (PCA) is a multivariate statistical technique, capable of compressing the data and reducing their size. It can be seen as a technical linear orthogonal projection which projects the cases represented in a multidimensional space of dimension n (n is the number of observed variables) in a subspace of dimension $q < n$, maximizing the variance of the projections (or minimizing the squared estimation error). This method is successfully used in diagnostic studies [109] [110].

2- Pattern Recognition: the goal of a pattern recognition method is automatic object classification following his likeness with respect to a reference object. In a diagnostic problem, a class is formed by the set of observations characterizing a situation or process operation: for example, the class C1 may be related to the normal functioning of the method, Class C2 for the gradient operation and class C3 for the failed operation. The diagnosis is to combine new observations to a class. The diagnostic problem is equivalent to the search for boundaries between classes that minimize misclassification. The calculation of the distance (Euclidean distance) can be chosen as a decision criterion for assigning a form to a class and determine what confidence is affected by the decision.

3- Spectral Analysis: Under certain normal conditions, certain measures have a typical frequency spectrum; any deviation of the frequency characteristics of a signal is connected to an anomaly. This

method is useful for analyzing signals which show oscillations with long periods (flow rates, pressures ...). The application of a decision procedure to detect and locate the faulty component in the system. Among the decision-making procedures applied to a sample of measurements include the empirical test threshold crossing, variance test, the test of the average [102][111].

4- Artificial Neural Networks (ANN): ANN is a computer system composed of a number of interconnected elementary processors (or nodes) that dynamically process the information that arrives from external signals. In general, the use of the ANN is done in two phases. First, the synthesis of the network includes several steps: the choice of network type, the type of neurons, the number of layers and the learning methods. The learning then allows, on the basis of optimizing a criterion, to reproduce the behavior of the system to be modeled. It consists in the search for weight parameters and can be done in two ways: supervised (the network uses the input and output data of the system to model) and unsupervised (only system input data is supplied and learning is done by comparing examples). When the learning outcomes achieved by the ANN are satisfactory, it can be used for generalization. This is the second phase where new examples that have not been used for learning, are presented to ANN to assess its ability to predict the behavior of the modeled system as well. Their low sensitivity to measurement noise, their ability to solve nonlinear problems and multivariate, store compactly knowledge, to "learn" online and in real time, are properties that make the use of ANN common. Their use can then be done at three levels:

- As a model system to watch in a normal state and generate an error residue between observations and predictions.
- As residues assessment system for the diagnosis.
- AS detection system in one step (as a classifier), or in two steps (for the generation of residues and diagnosis).

Several works have been done in the literature to develop methods for diagnosing chemical processes, based on the different information available to describe the behavior of systems. In 1990 E.M. Assaf et al. [112] and in 1984 J.A Barton [113] presented the results given by the neural model for the possible thermal runaway situations strongly exothermic process. The objective is to establish a reliable alarm algorithm for the detection and early prevention of this situation. So, the results show that the neural model is representative NARMAX for the dynamic behavior of the nonlinear chemical process. Experiments were performed in an exchanger-reactor and experimental data were used to define and validate the model. Then, the test of the cumulative sum is applied to indicate any drift of the normal behavior of the process. The abnormal behavior of the process due to two faults in its control parameters was examined. Indeed, each fault was caused by a sudden change in the flow of cooling. Note that the most common faults are due to unsatisfactory heat cooling, poor knowledge of operating conditions, the presence of impurities, failure of auxiliary equipment (controller, actuator, sensor,...) And the appearance of adverse events. In 1996, Gustin has classified causes thermal runaway [114] where the loss of the cooling capacity is the most important cause of the onset of thermal runaway.

Fault diagnosis with the ANN to radial basis functions with one hidden layer may be applied to develop a non-linear representation of the polymerization reactor based on a repeating structure [115]. In the first case study of a batch reactor, a temperature sensor fault and a jacket fouling were studied and classified

successfully using the model of neural processes and neural radial basis function (RBF) Classifier. The ability of ANNs to extract information directly from the data available on the process can be the most important reason to apply this approach in the industry. They are useful in problems such as sensor data analysis, which is beyond the scope of the conventional techniques algorithmic expert system. Multiple fault diagnostic possibilities (occurring two, three at a time) are also studied. The neural network can diagnose many double faults compared to triple faults.

In industrial processes, ANNs were applied for the detection and fault diagnosis. For example, in 1998, A. F. Cubillos et al. [116] have described an adaptive hybrid system established on prior knowledge and neural networks to model the process control strategies and uncertain parameters in a strongly non-linear continuous reactor. In 2004 Power & Y. al. [117] described a supervisory framework in two steps fault diagnosis using neural networks. Based on this framework, a failure detection system has been implemented to identify the exact location of the faults and to diagnose. ANNs have made rapid developments in fault diagnosis of chemical processes and related areas. The process fault detection by artificial intelligence techniques was studied in 1990 by V. Venkatasubramanian et al. [118], in 1988 by JC Hoskins et al. [119] and in 1994 by K. Watanabe et al. [120]. In 1988 Hugo P. et al. [121] concluded that with regard to the safe operating problem in the reaction, the reactors must reach the desired temperature, which may end with secondary reactions or possible runaway.

Neural network-based methods have received a lot of attention because of their fast and robust execution, their capacity for structure recognition and association. The fault diagnosis problem can be interpreted as a structure recognition task. While, limitations lie in the need for an important database and the presence of a real process, which is not always available especially in the case of learning the failure modes.

As stated previously, the ANN can be used to diagnose failures. Their low sensitivity to measurement noise, their ability to solve nonlinear and multivariate problems, store compactly knowledge, learn online and in real time, are indeed many properties that make them attractive for this use.

5- Fuzzy Logic (FL): fuzzy logic is a mathematical theory introduced in 1965 by L. Zadeh [122], which takes into account uncertainty and allows merging of information. The idea of the fuzzy approach is to build a device, called fuzzy inference system able to imitate the decisions of a human operator from verbal rules translating his knowledge of a given process. Find a mathematical relationship between a failure and its symptoms is often difficult. However, based on their experience, human operators are able to determine the faulty component that is causing the symptoms observed. This kind of knowledge can be expressed using rules of the form: IF condition THEN conclusion. Where the condition part includes symptoms observed and the conclusion part includes the faulty component. Thus the diagnostic problem is considered as a classification problem. The vector of the symptoms of the classifier, developed from the measured values on the system, can be seen as a form, it is to rank among all the shapes corresponding to normal operation or not.

2.5.2.2 Quantitative Methods

Quantitative methods or knowledge base are implemented when the majority of measures are unavailable and when building the model is difficult. They can be used to identify the causes of failures of an industrial process. Its functional and structural analyzes are based on the experience and knowledge of the operator.

1- Tree of Failure: the fault tree first appeared in 1988 by Villemeur [123], is one of the major tools for analyzing technological risks.

The objective of this approach is to identify the various possible combinations of events that cause the creation of a single adverse event. The graphic representation is composed of a tree structure allowing treatment of both qualitative and quantitative. The fault tree is made of several layers where the root corresponds to the adverse event. The levels are ordered successively such that each event is generated from the lower level events through logical operators (AND, OR). The decomposition stops at the levels of elementary events, characterized by the fact that they are independent and not be broken down into simpler components.

2- Expert System (The case-based reasoning): The case-based reasoning is modeling expertise and reasoning skills of qualified specialists in the emerging field. This reasoning is qualified to solve a problem by relying on past experiences. Knowledge is stored as appropriate. A case is a contextual piece of knowledge, representing an experience that can be used to achieve the goals of the reasoning engine. Thus, a case can be seen as a proven situation in the past, associated with the result of some relevant action. The reasoning from cases is reasoning by analogy. The attributes of a situation are used as an index in the case of the library to get the best, according to some similarity criteria, and thus to determine the solution. Expert systems based on the use:

- A knowledge base that contains the expertise of the specialist described as rules whose structure is as follows: IF <condition> THEN <conclusion>.
- A fact base that contains the basic information needed to establish a diagnosis.
- An inference engine that mimics the expert's reasoning process.

The difficulty here is to clearly define the cases in other words, to identify those useful and necessary for the description of a situation. Their determination for dynamic systems is far from obvious [124].

2.5.3 Comparative Analysis of Diagnostic Methods on Chemical Processes

The comparison between the methods of diagnosis is a very difficult task. Indeed, the decision of selecting a fault detection method depends on several factors: a priori knowledge available on the system, the presence or absence of a mathematical process model to diagnose, type of faults to be detected, the presence or absence unknown inputs (noise, uncertainty), nonlinearity, system is closed or open loop ...

For example, for the electrical and mechanical systems, it is easy to design a mathematical model. The choice of model-based diagnostic methods will be preferred. By cons, for chemical industrial processes, modeling is difficult to achieve because even if a mathematical model is obtained, it is complex. In this case, the reference model without flaw detection methods can be applied. In 2006 Isermann et al. [98] proposed a comparative study of diagnostic methods by analytical redundancy. It should be noted that many methods are born from the coupling of several diagnostic strategies which aims to combine the advantages of each strategy. Some combined approaches are presented in [Isermann, 2006] [98].

Table 2.1 provides a synthetic comparative analysis of approaches for detection and diagnosis that we have just explained.



With/without a model	Method	Advantage	Disadvantage
Model-Based	Influence graph	<ul style="list-style-type: none"> Knowledge of behavior (models) unnecessarily default location Explanatory capacity Maximum operating structure Unique graphic language applicable to many systems Based solely on the course of causal paths 	<ul style="list-style-type: none"> Generating a large number of assumptions which may lead to misdiagnosis Construction of the structure can cause loss of information
	The parametric estimation	<ul style="list-style-type: none"> Well adapted to multiplicative faults (affecting parameters) Provide information on the extent of deviations Application Process low number of variables where specific templates can be defined. 	<ul style="list-style-type: none"> Localization task difficult diagnosis Relations between mathematics and physical parameters not always reversible in a unitary manner Working hard for complex installations due to a large number of variables involved Need for a permanently excited physical system: Problem in the case of hazardous processes or operating in a stationary mode
	The parity space	<ul style="list-style-type: none"> Knowledge of the system decoupled from the knowledge of diagnosis Very general method Lower cost of development Structure interesting for complex processes Facilitates analysis: insulation failure 	<ul style="list-style-type: none"> Unsuitable for non-linearity and non-additive failures Consideration of additive uncertainties
	The state estimation	<ul style="list-style-type: none"> Simple calculation stronger method for measuring noise Applicable to linear and non-linear systems very common Methods Ability to decouple the responses 	<ul style="list-style-type: none"> No detection warranty if the fault type has not been modeled Diagnostic error due to disturbance Need to have a precise and complete model Adaptability to difficult processes of change and lack of general method due to the local nature of the model (applied to the system studied)



		of unknown inputs	<ul style="list-style-type: none"> Poorly adapted to the complex process
Model Free	Expert systems	<ul style="list-style-type: none"> Low calculation time No complicated reasoning (no intermediate calculation), direct generation rules by a human expert who results comprehensible to the operator. 	<ul style="list-style-type: none"> Depends on the expertise done on the system Uncertain data: difficulties in the analysis of a set of uncorrelated data incomplete and ambiguous Deep Lack of knowledge: no explanations on the conclusions adopted. Inconsistency rules: add or delete rule impact on other rules difficult to detect Difficulties in acquiring expertise Lack of generic because the rules depend on the system architecture The problem of the evolution of the system: adding / component change involves new expertise Robustness: fixed rules and not robust to unrecognized situations
	Tree of failures	<ul style="list-style-type: none"> Very effective for analyzing and solving diagnostic problems failures in industrial processes. Identification of common failure modes that could affect the system Very powerful approach for analyzing single failures of elements leading to global failure No treatment necessary 	<ul style="list-style-type: none"> Multiple faults handling disability. Axes by ray constructed dependent on its creator: error sensitivity No formal method to verify the accuracy of the developed tree. Request a long experience Any Change or development of the system requires rewriting of the table Poorly adapted to dynamic systems, highly time-dependent, because of the high number of variables and process
	PCA	<ul style="list-style-type: none"> Handling noise and correlation to extract information efficiently A powerful tool capable of reducing the size of data so that the information is retained Facilitates data analysis 	<ul style="list-style-type: none"> No property signature making it difficult insulation faults Invariant representation in time Periodic Update
	Fuzzy logic	<ul style="list-style-type: none"> Approximation of the behavior of a complex system 	<ul style="list-style-type: none"> Do not provide a better understanding of the relationship between variables



Normandie Université

		<ul style="list-style-type: none"> • Output variable directly related to the input variables • Description of the known physical structure of a system 	<ul style="list-style-type: none"> • Sets with no clear semantic meaning
	Pattern recognition	<ul style="list-style-type: none"> • Capable of processing data that is both uncertain and imprecise • Simple to apply • Low computing time for the classification of a new observation and independent of the size of the training set 	<ul style="list-style-type: none"> • Any comments must belong to a defined class • The characteristics of some modes remain unknown • No general rules for the choice of the space of representation which is the diagnostic success factor • Number of classes supposed known at the outset supposedly exhaustive knowledge
	Neural network	<ul style="list-style-type: none"> • Ability to solve nonlinear and multivariate problems • Ability to learn online and in real time • Low sensitivity to measurement noise, 	<ul style="list-style-type: none"> • Setting parameters of the delicate network • Time Significant calculation during the learning phase

Table 2.1 Comparative Analysis of Synthetic Fault Detection and Diagnosis Approaches



Plant operation today is becoming more complex as plants are often operated at extreme pressures and temperatures to achieve optimal performance. These extreme conditions may cause equipment failures and deviation in a process that may lead to catastrophic accidents. Although these plants are equipped with automated control, still the role played by the computer is limited and highly depended on human operators to maintain process plant integrity.

In some modern high technology process plant, there may be hundreds or thousands of process variables that need to be observed. As a human with limited capability, they are out of hand when handling these problems. Industrial statistics show that although major catastrophe may be infrequent, minor accidents are very common. These minor accidents cost the society billions of dollars every year [94]. Thus, to prevent this from happening, a quick detection and diagnosis of process fault are needed. An automated process fault detection and diagnosis (PFD&D) can provide a good solution to better safety in chemical process plant [125].

It is clear that a large number of developments concerned approaches based on quantitative models of the process using the technique of calculating the residuals [125] [1] [126].

However various factors such as the complexity of the system [127], the higher dimension, the non-linearity of the process have often made it very difficult to develop an accurate mathematical model [128]. This difficulty limits the application of this approach to real industrial processes. In particular, a too simple model does not generate residues only representative failures, but they often include model error, or normal drift parameters [128] [129]. The difficult to develop a mathematical model to diagnose and detect a fault in the chemical process has made use of the quantitative and qualitative methods without a model as the better solution in FDI domain such as the expert system, the tree of failures, the fuzzy logic, the pattern recognition and the neural network [108]. We noticed that if we want to dive over in the heart of the problem, we find that many of the methods without model suffer from many disadvantages in comparison with their advantage. The method depend on the human expertise, suffer from the difficulties in the analysis of a set of uncorrelated data incomplete and ambiguous in addition to the deep lack of knowledge no explanations on the conclusions adopted, the difficulties in acquiring expertise, the lack of generic due the rules that depend on the system architecture, the problem of evolution of the system (adding / component change involves new expertise) and finally the non-robustness fixed rules and not robust to unrecognized situations that all disadvantage reduce dependence on expert system on the fault diagnosis and detection (Table 2.1)

For the tree of failures, many of the weaknesses prevent the use of this method such as The absent of formal method to verify the accuracy of the developed tree, the need to request a long experience, also this method suffer from a poorly adapted to dynamic systems, highly time-dependent, because of the high number of variables and process. (Table 2.1) On the other side if we look to the fuzzy logic we find that this method has some advantages when we use it on industrial faults diagnosis and detection or this method has a strong approximation of the behavior of a complex system, and their output variable directly related to their input variables and it's has the ability to description of the known physical structure of a system (Table 2.1) Although the advantages of fuzzy logic, but it remain few in comparison with the advantages of the neural network [130] [131], the benefits of ANN allow it to be applied in a variety of areas.



Normandie Université

These areas are automotive, aerospace, business, banking, voice recognition and robotics. In the mid-1980s, researchers in the field of process safety proposed a potential approach to detect and diagnose faults using the ANN. The first researchers to apply ANN in PFD & D, is done in 1989 by Venkatasubramanian and Chan [132]. After that, many works have studied the benefits of the implementation of the neural network in this area, such as the work of Li et al in 1990 [133], Becraft and Lee in 1993 [134] and Leung and Ramagnoli in 2000 [135]. ANN has also been used to perform several functions that could also be performed by integrating other fault-detection techniques such as the knowledge-based expert system [9].

Some of the benefits of using the ANN on chemical engineering are as follows [136] [126]:

- Ability to solve nonlinear and multivariate problems
- The potential for online use – ANNs may take a long time to train, but once trained, they can calculate results from a given input very quickly. Since a trained ANN may take less than a second to calculate results, it has the potential to be used online in a control system.
- Adaptive behavior – ANNs have the ability to adapt or learn, in response to their environment. They learn through training, where when given the input-output patterns, they adjust themselves to minimize the error.
- Pattern recognition properties – ANNs perform multivariable pattern recognition very well and this is where ANNs will probably find the most useful especially in process control and fault diagnosis.
- Filtering capacity: low sensitivity to noise and incomplete information – ANNs can deal with the imperfect world, generalize and draw substance conclusions more effectively than the less flexible empirical models.
- Automated abstraction – ANNs can ascertain the essentials of relationships and can do so automatically. They do not need the domain expert that knowledge-based required. Instead, through training with direct (and sometimes imprecise) numerical data, ANNs can automatically determine cause-effect relations.
- Ability to learn online and in real time

Their low sensitivity to measurement noise, their ability to solve nonlinear problems and multivariate, store compactly knowledge, to "learn" online and in real time, are properties that make use of attractive ANN. However, the major disadvantage is how to determine a methodology to master the inherent problems, which are mainly the selection of the structure, network size and learning algorithms for a specific problem [47].

In recent years, many studies have conducted process fault detection and diagnosis using a variety of methods such as knowledge-based expert system (KBES), artificial neural network (ANN) and mathematical modeling. Using KBES in PFD&D has the advantage of having an overview of problem-solving in a chemical plant. The tedious nature of knowledge acquisition represents its own limitations, the system's inability to dynamically learn or improve its performance and the unpredictability of the system outside its area of expertise. Fault detection and diagnosis based on knowledge has also been introduced by the combination of heuristic



knowledge (rules of operator knowledge) and knowledge of the procedure (mathematical models, Kalman filtering algorithms, and procedures of signal processing) [125]. However, the technique could be very complex when it comes to a non-linear process [126].

As results of these limitations, for PFD&D of a non-linear and complex system such as the chemical industrial process, ANN can provide a better solution especially when it combines with fuzzy logic because of its utility in the representation of the data input-output, data classification, and pattern recognition [137]. According to [3], there is no single method that can handle the entire system failure. Thus, the combination of ANN and fuzzy logic is very convenient because it combines both advantages and makes the whole system more robust. ANN can work simultaneously on qualitative and quantitative data and in the case of the difficulties to build a mathematical model of the system; ANN demonstrates that is very useful in this case while fuzzy logic has the ability to mimic the ability to detect, generalize, process, operate and learning of the human operator [138] [131].

2.6. Artificial Neural Network and Industrial Diagnostics.

Artificial neural networks (ANNs) are parallel and distributed systems of information processing inspired by the functioning of the human brain. Following we will describe the operating principle of ANNs, their different architectures, the available learning algorithm families, and a summary of the main advantages of neural networks as an information processing system. ANNs are a new approach to information processing. They offer compact and fast solutions for a wide range of problems, especially those with real-time constraints such as most current space applications. This is truer with the use of emulations and hardware implementations. They can provide an interesting solution for monitoring problems of industrial equipment. Important properties of neural networks include their fault tolerance which measures their ability to perform the task they are asked for in the presence of erroneous information and to maintain their computability even if part of the network is damaged [139].

2.6.1.2 Artificial Neural Network

Definition: *Kohonen* proposes the following definition: "The ANNs are massively connected networks in parallel to simple elements (usually Adaptive) and their hierarchical organization. They are supposed to interact with the objects of the Real world in the same way as the biological nervous systems." According to this definition, we can say that an ANN performs one or more algebraic functions of its entries. By the composition of the functions performed by each of the neurons, we can model an ANN using a graph-oriented by the interconnection of simple elements (neurons) and the exchange of information via connections. The calculation will be done in a distributed, parallel and cooperative way. Figure 2.3 illustrates an ANN with four inputs is an output.

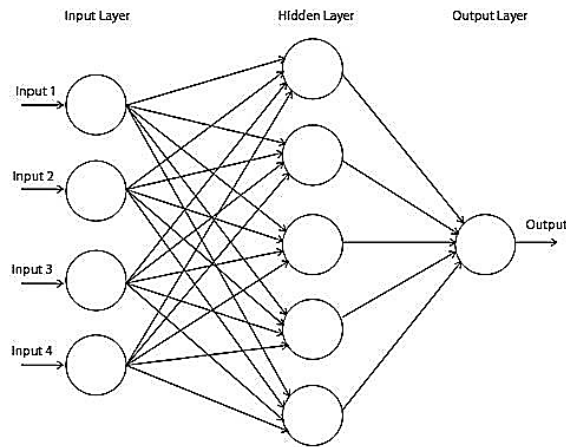


Figure 2.3: Example of an Artificial Neural Network

The essential characteristics of a neural network are its architecture (topology): type of interconnection, choice of the transfer function and its mode of learning, ie how to estimate or learn the weights and especially the tool representation of knowledge. It is a distributed representation, where each neuron participates, which leads us to observe that the connections between the neurons that make up the network describe the "topology" of the model [140].

2.6.2 Neural Architectures

2.6.2.1 Non-looped Neural Networks "Feedforward"

A non-looped neuron network is presented by a set of neurons connected to each other such as the information flowing from the inputs to the outputs without going back. The calculation of Y (output) is done by propagating calculations from left to right, possibly with linear direct connections: $Y = a \times x + fw(x)$

This type of network includes two groups of architectures: single-layer networks and multi-layer networks. They differ by the existence or not of intermediate neurons called neurons hidden between the input units and the output units called source nodes or input nodes and output nodes respectively [141].

2.6.2.1.1 Non-looped Mono-layer Networks

This type of network has an input layer receiving the stimuli to be processed via the source nodes. This layer is projected into an output layer composed of neurons (computing nodes) transmitting the results of the treatment to the external environment.



2.6.2.1.2 Non-looped Multi-layer Networks

This type of proactive network is characterized by the presence of one or more hidden layers, whose corresponding compute nodes are called hidden nodes or hidden units. Hidden layers intervene between the input of the network and its output. Their role is to pretreat the input signals, received by the input layer from the external medium, and transmits the corresponding results to the output layer or will determine the final response of the network before it is transmitted to the outside environment. This pretreatment role verified by adding one or more hidden layers, the network is able to extract more statistical properties than those extracted from a similar network with fewer hidden layers. This is useful for performing more complex functions than simple linear separations. In this type of network, the inputs of the neurons of a particular layer come only from the outputs of the previous adjacent layer. The most frequently used networks in this category are Multi-layers Perceptron (MLPs) [142].

2.6.2.2 Looped Neural Networks (Recurrent)

A discrete-time looped neural network produces one or more equations with non-linear differences, by the composition of the functions performed by each of the neurons and delays associated with each of the connections [143]. These networks are characterized by the presence of at least one feedback loop at the level of the neurons or between the layers, and taking into account the temporal aspect of the phenomenon. But it is harder to implement meter models.

2.6.4 Most Used Neural Networks

Today, the number of possible neural network types is quite high. We present a summary diagram (Figure 2.4) of what is the most used ANNs in the diagnosis field.

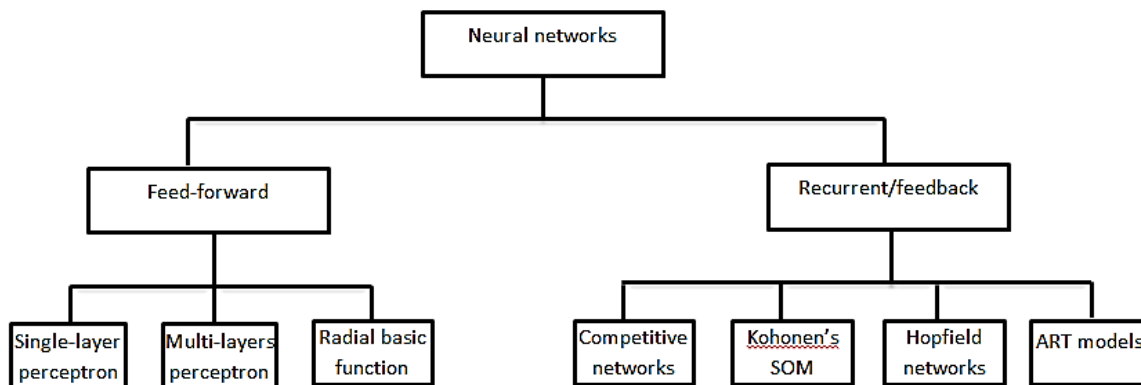


Figure 2.4: Some Usual ANNs

2.6.4.1 Simple Perceptron

The perceptron is the first model of neural network invented in 1957 by Frank Rosenblatt [144]. The goal of the perceptron is to associate input forms with answers. The perceptron



consists of two layers: the retina and the output layer that gives the response corresponding to the stimulation present at the input. The cells of the first layer answer yes/no. The answer "yes" corresponds to a value "1" and the answer "no" corresponds to a value "0" at the exit of the neuron. The input cells are connected to the output cells through synapses of varying intensity. The learning of the perceptron is done by modifying the intensity of these synapses. The output cells evaluate the intensity of the stimulation from the cells of the retina by summing the intensities of the active cells.

The perceptron must find all the values to be given to the synapses so that the input configurations are translated into desired responses. For this, we use the Windrow-Hoff learning rule. To learn, the perceptron (Figure 2.5) must know that he has made a mistake and must know the answer he should have given. As a result, we speak of supervised learning. The learning rule is local in the sense that each output cell learns without needing to know the response of the other cells. The cell only changes the intensity of its synapses (learns) when it is wrong. Minsky has shown that a simple form (the xor) cannot be learned by a perceptron neuron. A neuron can only separate two regions separated by a hyperplane. With several neurons, it's better now, but it's clear that a single layer of perceptron cannot learn complex shapes [145].

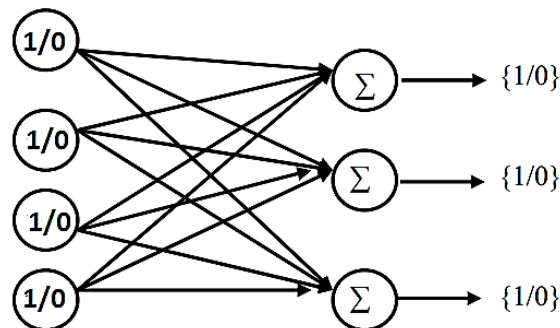


Figure 2.5: General Scheme of Simple Perceptron.

2.6.4.2 Multi-layer Perceptron

The Multi-layer Perceptron (MLP) is an organized network of artificial neurons organized in layers and where information moves in one direction, from the input layer to the output layer. Figure 2.6 gives an example of a network containing an input layer, two hidden layers, and an output layer. The input layer always represents a virtual layer associated with the inputs of the system. It contains no neurons. The following layers are layers of neurons. In the illustrated example, there are 3 inputs, 4 neurons on the first hidden layer, three neurons on the second and four neurons on the output layer. The outputs of the neurons of the last layer always correspond to the outputs of the system. In the general case, a Multi-Layer Perceptron may have any number of layers and number of neurons (or input) by one layer also.



Normandie Université

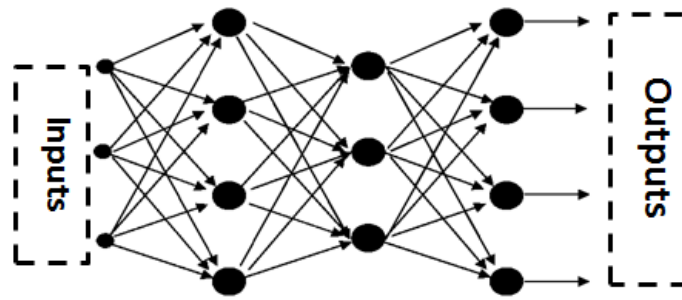


Figure 2.6: Example of Multi-layer Perceptron Network.

The creation of a multi-layer perceptron to solve a given problem therefore passes through the inference in the best possible application as defined by a set of learning data consisting of pairs of input vectors and desired outputs. This inference can be made, among others, by the said retro propagation algorithm [145].

2.6.4.3 Retro Propagation Network

One of Perceptron's disadvantages is that it minimizes an all-or-nothing error because of its activation function. It does not take into account the notion of distance. Because of this, it is very weak. The Widrow-Hoff learning rule (Delta rule) [146] no longer works in all or nothing but minimizes a quadratic error function, which is more robust. Unfortunately, this rule can only be applied to single-layer networks of adaptive weights. It is therefore by extending the Widrow-Hoff rule that several teams of researchers have developed a learning algorithm called the backpropagation gradient error, which was then generalized by the Rumelhart team in 1986 [147]. This algorithm provides a way to modify the weights of the connections of all layers of a Multi-layer Perceptron (MLP) [34] (Figure 2.7).

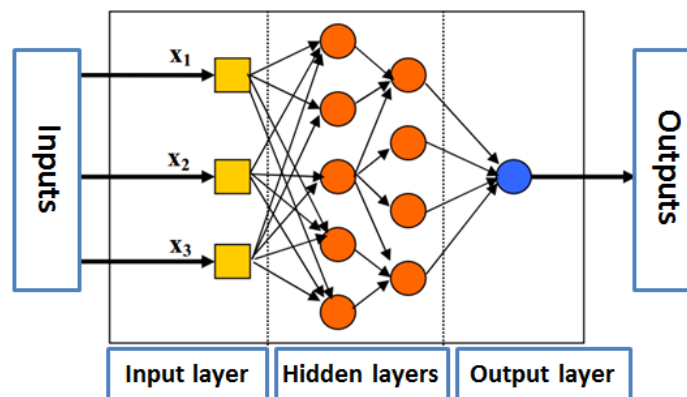


Figure 2.7: Non-looped Neural Network



Normandie Université

This algorithm is made in order to answer the question: how to pass on, on each of the connections, the error signal which can only be measured on the output layer after going through several non-linear steps?

For this algorithm as well as one is able to propagate a signal from the input cells to the output layer, one can by following the reverse way, retro propagate the calculated error output to the inner layers. The algorithm of retro propagation of the gradient of the error made it possible to exceed the limits of the simple Perceptron. It is able to solve a large number of classification and pattern recognition issues and has resulted in many applications. This Algorithm nevertheless suffers from numerous faults, including [143]:

- The calculation time: the learning time is very long;
- A high sensitivity to the initial conditions, i.e. the way in which the weights of the connections are initialized;
- Many problems are caused by the geometry of the error function: local minima. This problem is partly solved with the stochastic gradient, but it still remains;
- The problem of dimensioning the network: The retro propagation learns a base of learning on a network whose structure is fixed a priori. The structure is defined by the number of hidden layers, the number of neurons per layer, and the topology of the connections. A poor choice of structure can significantly degrade network performance.

2.6.4.4 Radial Basis Function Neural Network (RBF)

Networks with radial basis functions (RBF) or more simply radial-based networks have been proposed by J. Moody and C. Darken. There is an organization with an input layer, a hidden layer, and an output layer. Each hidden neuron only reacts to a small part of the input space (its area of influence). For a network having n inputs and m hidden units, activation of hidden neurons is given by a Gaussian function of the type (eq. 2.1) (the input and activation functions are combined):

$$a_i = \exp\left(-\frac{1}{2} \sum_{k=1}^n \frac{(e_k - c_{k,i})^2}{\sigma_{k,i}^2}\right) = \prod_{k=1}^n \exp\left(-\frac{1}{2} \frac{(e_k - c_{k,i})^2}{\sigma_{k,i}^2}\right) \quad (2.1)$$

Where i denotes the index of the neuron, k traverses the set of entries noted e_k , and $c_{k,i}$ and $\sigma_{k,i}^2$ are parameters called respectively centers and variances of Gaussians. Figure 2.8 shows the shape of this activation function for a neuron that has only one input.

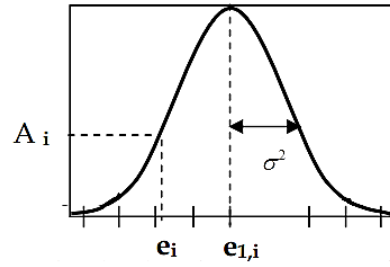


Figure 2.8: Activation Function of a Hidden Neuron with a Single Input

Each of these neurons is thus activated significantly only for input values relatively close to the Gaussian centers. Connections from input neurons are unweighted. The activation of an index output neuron i is given by (eq.2.2):

$$\sigma_i = \frac{\sum_{j=1}^m w_{ij} a_j}{\sum_{j=1}^m a_j} \quad (2.2)$$

Where j runs through all the clues hidden neurons, neurons of this type thus realize a weighted sum of the activation values of the hidden neurons. The term $\sum_{i=1}^m a_i$ called normalization factor is not mandatory. We talk about standard network when used. Learning is done in these networks by changing the weights of the connections between the hidden neurons, the output neurons, the centers and the variances of the Gaussians. As before, a gradient descent is performed with the aim of minimizing the quadratic error, the expression of which is given by equation (eq.2.3).

$$q = \frac{1}{2} \sum_i [a_i - s_i]^2 \quad (2.3)$$

This model, however, suffers from a disadvantage compared to multi-layer networks since unlike them, its approximation domain (i.e. domain in which it realizes a satisfactory approximation) is strictly limited. It confines itself to the areas of influence of hidden neurons, without which the network is unable to extrapolate. When the size of the input domain is very important, the number of necessary neurons can become considerable and the use of multi-layer networks may be more appropriate [148].

2.7. Fuzzy Logic and Industrial Diagnostics.

2.7.1 History of Fuzzy Logic:

Initially, the theory of fuzzy logic was affirmed as an operational technique. Used alongside other advanced control techniques, it makes a discrete but appreciated input into industrial control



automation. The theoretical foundations of fuzzy logic were established in the early 1965s by Professor Lotfi Zadeh of the Berkeley University of California [122], this technique combines the notions of "fuzzy subset" and "theory of possibilities". In 1970, it was the first application of fuzzy logic in expert systems of decision support in medicine, and in 1975 Mamdani [149] performed a fuzzy regulation of a steam boiler. The Japanese [150], in 1985, was the first to use fuzzy logic in Fuzzy Logic Inside consumer products. The fuzzy set theory has also given rise to an original treatment of uncertainty, based on the idea of order, which formalizes the treatment of ignorance, and allows the formalization of advanced information systems. Fuzzy sets also have an impact on automatic classification techniques and have contributed to some renewal of existing approaches to decision support [139].

2.7.2 The Requirements of Fuzzy Logic

Fuzzy logic is called linguistic logic because its values of truth are words of the current language; it is not an imprecise logic but a logic that adapts itself to the human being. Fuzzy logic is a tool for integrating human knowledge into practical algorithms. It has the advantage of combining digital and symbolic data processing.

Fuzzy logic methods can be used to design intelligent diagnostic systems [151] based on knowledge expressed in natural language. It is an approach modeled on human reasoning rather than rigid calculations. Indeed, the reasoning mode in fuzzy logic is more intuitive than classical logic. It allows designers to better understand natural phenomena, inaccurate and difficult to model by relying on the definition of rules and membership functions in sets called "fuzzy sets" [122] [152]. A field of application of the fuzzy logic which becomes frequent is that of the regulation and control of industrial regulations. This method makes it possible to obtain a control law which is often effective, without having to resort to significant theoretical developments. It presents the interest of taking into account the experiences acquired by the users and operators of the process to be controlled. The basic elements of fuzzy logic are [153] [152]:

- Linguistic variables
- Membership functions
- Deductions to inferences: Decision-making from a rule base *If ... Then*

2.7.3 Fuzzy Sets

2.7.3.1 Introduction

Fuzzy sets theory is a mathematical theory, it was introduced by Lotfi Zadeh in 1965, which showed that this theory is a special case of the theory of classical subsets where the membership functions considered take binary values ($\{0,1\}$) [150]. The notion of fuzzy set aims to allow the idea of a partial membership of an element to a set or a class that is to say to allow an element to belong more or less strongly to this class [154]. This concept allows the use of data categories with poorly defined boundaries, intermediate situations between everything and nothing, the gradual transition from one property to another, etc. Let U be the set of values of the variable x ,



called the universe of discourse; a subset A of U and a function $\mu_A(x)$ between 0 and 1. This function $\mu_A(x)$ quantifies the degree to which each element x of U belongs to A .

U : the universe of discourse.

A : subset of U .

- If $\mu_A(x) = 1$ x completely belongs to the subset A
- If $\mu_A(x) = 0$ x does not belong to subset A
- If $0 < \mu_A(x) < 1$ x partially belongs to subset A

The subset A is therefore a fuzzy set and $\mu_A(x)$ is called the membership function. Figure 2.9 provides a better understanding of the notion of a fuzzy set compared to a classical set.

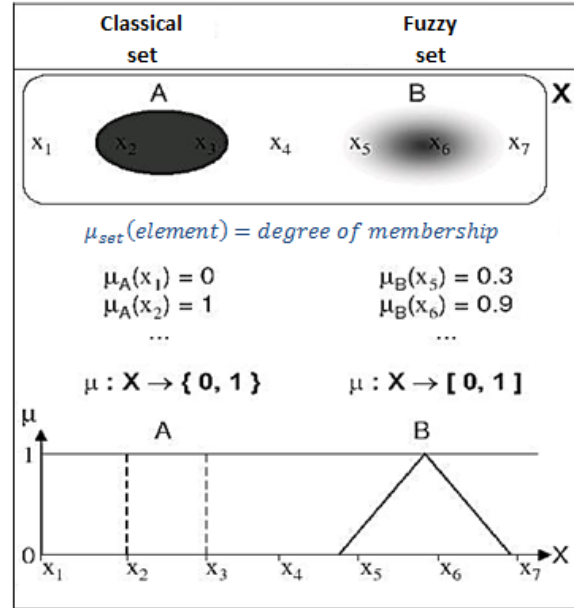


Figure 2.9: Comparison between Classical and Fuzzy Set.

Equation 2.4 defines a fuzzy set completely described by its membership function $\mu_A(x)$.

$$A = \{(x, \mu_A(x)) / x \in U\} \quad (2.4)$$

In the case of a discrete set $U = \{x_i, i = 1, 2, \dots, n\}$, a fuzzy set A (eq.2.5) can be defined by a list of ordered pairs: degree of membership / element of the set:

$$A = \{\mu_A(x_1)/x_1, \mu_A(x_2)/x_2, \dots, \mu_A(x_n)/x_n\} \quad (2.5)$$

We often use vector formalization, more convenient for programming (eq.2.6):

$$x = [x_1, x_2, \dots, x_n]^T, \mu = [\mu_A(x_1), \mu_A(x_2), \dots, \mu_A(x_n)]^T \quad (2.6)$$



2.7.4 Membership Function

The membership function measures the degree to which an element x belongs to a fuzzy set A . They are either uniformly distributed or random [153] [155]. Membership functions can have several forms (Figure 2.10):

- Triangular.
- Bell-shaped.
- Monotone (increasing or decreasing).
- Trapezoidal.
- ...

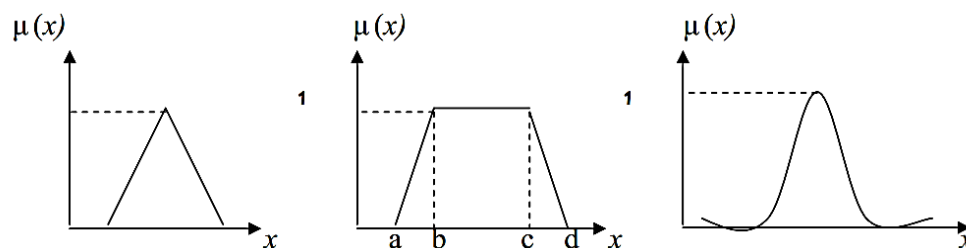


Figure 2.10: Different Form of Membership Function.

The shape of the membership function is to choose according to the processed application. However, for a fuzzy set, what matters is less the precise value of the degrees of membership of the elements that support the scheduling of these membership degrees between them. The most used membership functions are in trapezoid or triangle form, and allow to respect this constraint while keeping a very simple analytical form. In some applications, where we must derive the membership function, we will instead choose S-functions or Gaussian-type functions, continuously differentiable on their support.

2.7.6 Fuzzy Inference System [156]

The concept of the fuzzy rule defines a fuzzy expert system as an extension of a conventional expert system, manipulating the fuzzy proposition. Thus a fuzzy inference system (FIS) is formed of three blocks as shown in Figure 2.11. The first fuzzification block transforms numeric values into degrees of membership in the various fuzzy sets of the partition. The second block is the inference engine, made up of the set of rules. Finally, the block of defuzzification allows, if necessary, to infer a net value, from the result of the aggregation of the rules.

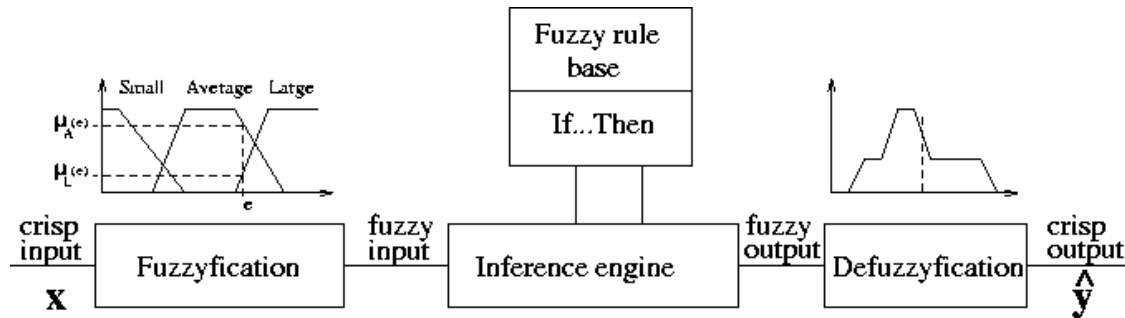


Figure 2.11: Fuzzy Inference System

Fuzzy systems based on "if ... then" rules have antecedents and consequents that are symbolically specified. In the context of system modeling and control, the exploitation of such knowledge generally requires the establishment of numerical/symbolic (N / S) and symbolic/numerical (S / N) interfaces. These are essential gateways for establishing a link between the set of rules (rule base) that interface the fuzzy system and the process, on which only measurements and numerical actions are possible. Conventionally, the internal operation of fuzzy systems is based on a structure, as shown in Figure 2.11, which includes:

- A rule base containing a number of rules *if ... Then* of the control strategy of the expert; and a database that includes all the definitions used in the fuzzy control (universe of discourse, fuzzy partitions, choice of operators ...).
- A decision unit (Inference) that transforms interference operations into rules from a knowledge base (provided by the expert) and the fuzzy subset corresponding to fuzzification of the measurement vector. In general, several fuzzy variable values, suitably defined by membership functions, are linked together by rules, in order to draw conclusions. We then speak of fuzzy deductions. In this context, we can distinguish two kinds of inference rules:

- 1- *Inference with several rules*
- 2- *Inference with a single rule.*

In the case of an inference with multiple rules, they are expressed in the general form:

If condition 1, then operation 1, OR
If condition 2, then operation 2, OR
If condition 3, then operation 3, OR

 \vdots
If condition m, then operation m,

Conditions may depend on one or more variables.

In the second case, the variables are linked together by fuzzy operators AND and OR. Each variable is assigned membership functions, taking into account the fuzzy sets formed by these variables.



- A fuzzification interface that transforms crisp entries into verification degrees of linguistic values.
- Finally, a defuzzification interface, with an optional post-processing information which converts the fuzzy inference results in a crisp output.

2.7.6.1 Fuzzification (Fuzzy Quantization)

The fuzzification is the step of passing the digital domain into the symbolic domain (blurred). This step is necessary as long as we want to manipulate, using the fuzzy set theory, measurable physical quantities (precise or imprecise). Depending on the application, fuzzification can be done in different ways. This can be done by transforming numerical data into linguistic values to give a subjective view of the state of the observed system. It is necessary to transform the non-fuzzy variables from the outside world to fuzzy sets. To do this, use an operator called fuzzification which associates a measure of a particular membership function. Fuzzy logic systems deal with fuzzy input variables and provide results on output variables that are themselves fuzzy [157].

How to fuzzify?

For fuzzifier must be given:

- The universe of discourse, i.e.: Range of possible variations of the input considered.
- A fuzzy classroom partition of this universe.
- The membership functions of each of these classes.

2.7.6.4 Defuzzification

The inverse operation of moving from a linguistic variable representation to a physically applicable numerical variable is called defuzzification. There are several methods to obtain an accurate value from a fuzzy set as input. It cites as examples the average of the maximum and the center of gravity. Defuzzification, also called the combination of rules, is necessary when several rules of inference are validated because one finds oneself in this case with several fuzzy sets of exit, one must, therefore, apply a technique to find an exit value [158].

1. Center of Gravity (CoG) Method [158] [159]:

In fuzzy control, CoG defuzzification (Figure 2.12) is the most used. The output value is given in the continuous case (eq.2.7):

$$x_0 = \frac{\int_{x_1}^{x_2} x\mu(x)dx}{\int_{x_1}^{x_2} \mu(x)dx} \quad (2.7)$$

In the discrete case this output value is defined as the following equation (eq.2.8):



Normandie Université

$$x_0 = \frac{\sum_{i=1}^n \mu_i x_i}{\sum_{i=1}^n \mu_i} \quad (2.8)$$

With:

n : Level of discretization

x_i : i th exit.

μ_i : Membership value of the i th output.

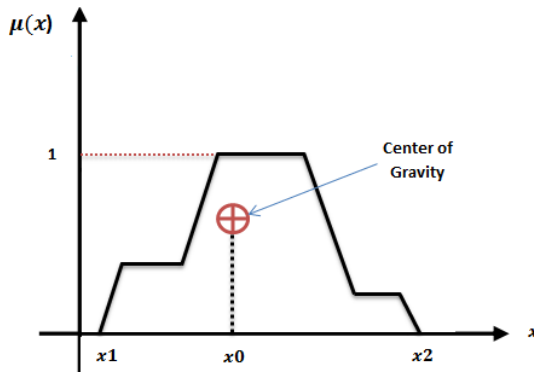


Figure 2.12: Defuzzification Using Center of Gravity Method

2. Method Mean of Maximum (MoM) [158][159]:

This is the average of the most likely output values. MoM defuzzification (Figure 2.13) is rather used when it comes to discriminate an output value (eg recognition). In the discrete case the output value is given by (eq.2.9):

$$x_0 = \sum_{i=1}^L \frac{r_i}{l} \quad (2.9)$$

L : Number of quantified values r_i where the membership is max.

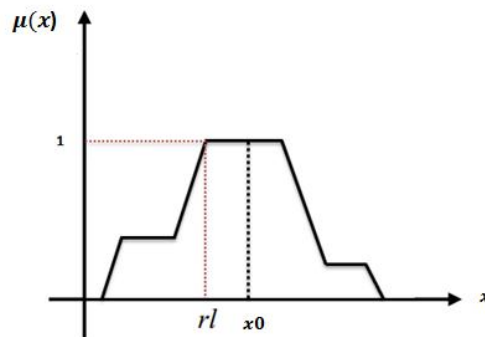


Figure 2.13: Defuzzification Using Mean of Maximums Method



2.8 Faults Diagnosis-Application to a Distillation Column

2.8.1 Introduction

These days, with advances in plant operation, more complexity is operated at extreme conditions to realize best performance. Extreme conditions such as pressure and temperature could cause failures in instrumentation and some deviations that may cause harmful accidents, as since explosion of the four reactors at the Daiichi nuclear plant Fukushima- Japan in 2011 [108], the toxic cloud dispersed dioxin in the town of Seveso, 4,000 injured and 200,000 victims of the disaster in Bhopal or the explosion of the AZF factory, Gupta in 2005, etc. Although the plants are equipped with automatic systems, computer simulation and analysis is still limited to maintain process plant integrity and extremely relied on human operators. For instance, humans can't detect hidden faults or predict future problems. Previous Industrial statistics have shown that major catastrophes may be infrequent, and minor accidents are very common. The yearly costs of these accidents are beyond billions of dollars [160]. Therefore, it's necessary to build methods for the detection and diagnosis of faults. This could provide solutions for the chemical process plant safety [160]. Diagnostic methods differ according to different criteria: the dynamic process, complexity, online diagnostic implementing, the nature of information, depth, distribution... in this context, several classifications are also proposed in the literature. These classifications are influenced by terminologies and specific contexts of each community and are not always consistent.

Refer to state of arts many researchers said that the use of particular techniques without a model, such as the expert system, the tree of failures, the fuzzy logic, the pattern recognition and the neural network, is more effective than the methods with a model, especially for the diagnosis and prognostics of faults in real time on a distillation process [126] [16]. In chemical engineering applications, the fuzzy logic and the ANN technique has a lot of advantages over the other techniques [108]. The distillation plays a major role in the operations unit; in fact, it is the method that is most used in terms of separation, rather than the liquid-liquid (L-L) extraction or even other absorption columns. Its use is less complex than the other unit operations. In addition, it allows for a better separation of the L-L extraction process. In fact, it depends on the nature of the product to be separated. For example, to separate the various components of fuel oil, we may not use the L-L extraction process but the fractional distillation procedure which meets the requirements of separation.

The purpose of this type of distillation is to separate between two chemical compounds by varying the difference in boiling temperatures. One of the compounds being more volatile, it will vaporize until column head where it will be condensed with a total condenser. It will, therefore, make an exchange between the ascending vapor and descending liquid. This is called the enrichment of steam. Part of the condensate will be separated at the top of the column, this is called the distillate. Another part will, in turn, be separated into bottoms, it is called the residue.



Normandie Université

Signals acquisition is done by a supervision system that consists of software ETP 200. It allows monitoring of changes in parameters such as differential pressure or temperature at a given point of the distillation column. The signals obtained during each acquisition represent:

1. *S1: Pre-Heated Temperature,*
2. *S2: The Timer: Reflux Rate,*
3. *S3: Preheated Power,*
4. *S4: Loss of Charge,*
5. *S5: Heating Power,*
6. *S6: Feed Flow Rate,*
7. *S7: Boiler Temperature,*
8. *S8: TIC2: Column Head Temperature.*

While accidents automated continuous distillation that occurs in industry cause variations of the parameters especially:

- 1- *The Reflux Rate.*
- 2- *The Heating Power.*
- 3- *The Preheating Power*
- 4- *The Feed Rate.*

This is clearly illustrated in Figure 2.14. For more details see chapter 1.

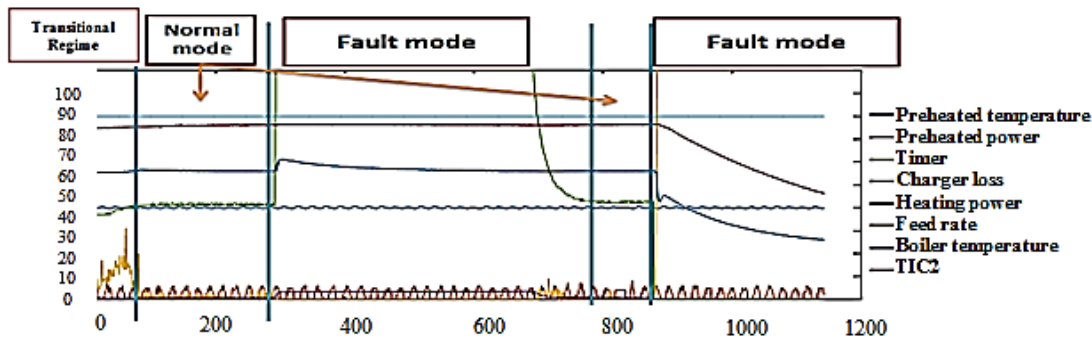


Figure 2.14: Graphical representation of the signals (normal & fault Mode)

So far in the previous sections of this chapter, we have presented the main framework for process fault diagnosis. In this section, we provide a real application for a diagnosis purpose. The proposed methodology is designed to classify 8 kinds of faults (E1...E8). For more details see in chapter 1

- *E1: Accident Reflux Ratio 0% (Timer);*
- *E2: Accident Reflux Ratio 100% (Timer);*
- *E3: Accident Heating Power 100%;*
- *E4: Accident Heating Power 0%;*



- ***E5: Accident Preheated Power 0%;***
- ***E6: Accident Preheated Power 95%;***
- ***E7: Accident Feed Rate 100%;***
- ***E8: Accident Feed Rate 0%.***

Normal mode signals (Table 1.1) are characterized by the following parameters: Feed flow rate set at 80% of its capacity, Pressure drop at 0.7 mbar, Temperature of the preheated to 40 ° C, Temperature of the boiler at 76 ° C, Column head temperature at 56 ° C.

2.8.2 Distillation Column Faults Diagnosis Based on Fuzzy Logic

Refer to the state of arts fuzzy techniques have received a lot of attention because of their fast and robust implementation, their ability to integrate prior knowledge, their performance in reproducing non-linear mappings and their generalization capabilities. Thus, fuzzy logic techniques are now being studied in the FDI research community as a powerful modeling and decision tool, with neural networks and other more traditional techniques such as nonlinear and robust observers, parity methods. To work around this problem of precision modeling, more models based on qualitative approaches can be used. Alternatively, fuzzy logic rules can be developed to help or replace the use of a model for diagnosis. The key advantage of fuzzy logic is that it allows the behavior of the system to be described by "if-then" relationships. The main trend in the development of fuzzy FDI systems has been to generate residuals using either parameter estimation or observers and to allocate decision making to a fuzzy logic inference engine.

The role of the fuzzy logic system, in this case, is to classify between normal mode and any accident mode generated on the output of the distillation column system.

2.8.2.1 Input Status Words

- *Low*
- *Normal*
- *High*

2.8.2.2 Output Action Words

- *Normal*
- *E1*
- *E2*
- *E3*
- *E4*
- *E5*
- *E6*
- *E7*
- *E8*

2.8.2.3 Defining of the Inference Rules

We call inference rules, the set of different rules that connect the fuzzy input variables of a system to the fuzzy output variables of this system. In our application these rules are in the form:

1. *If (S1 is Normal) and (S2 is Normal) and (S3 is Normal) and (S4 is Normal) and (S5 is Normal) and (S6 is Normal) and (S7 is Normal) and (S8 is Normal) then (system is Normal) (1)*
2. *If (S1 is Normal) and (S2 is Normal) and (S3 is L) and (S4 is Normal) and (S5 is L) and (S6 is Normal) and (S7 is Normal) and (S8 is Normal) then (system is E1) (1)*
3. *If (S1 is Normal) and (S2 is Normal) and (S3 is H) and (S4 is Normal) and (S5 is H) and (S6 is Normal) and (S7 is Normal) and (S8 is H) then (system is E2) (1)*
4. *If (S1 is Normal) and (S2 is Normal) and (S3 is L) and (S4 is H) and (S5 is H) and (S6 is Normal) and (S7 is Normal) and (S8 is H) then (system is E3) (1)*
5. *If (S1 is Normal) and (S2 is Normal) and (S3 is H) and (S4 is L) and (S5 is L) and (S6 is Normal) and (S7 is H) and (S8 is L) then (system is E4) (1)*
6. *If (S1 is L) and (S2 is L) and (S3 is Normal) and (S4 is Normal) and (S5 is Normal) and (S6 is Normal) and (S7 is Normal) and (S8 is Normal) then (system is E5) (1)*
7. *If (S1 is H) and (S2 is H) and (S3 is Normal) and (S4 is H) and (S5 is Normal) and (S6 is Normal) and (S7 is Normal) and (S8 is H) then (system is E6) (1)*
8. *If (S1 is Normal) and (S2 is Normal) and (S3 is Normal) and (S4 is Normal) and (S5 is Normal) and (S6 is H) and (S7 is Normal) and (S8 is H) then (system is E7) (1)*
9. *If (S1 is Normal) and (S2 is Normal) and (S3 is H) and (S4 is Normal) and (S5 is Normal) and (S6 is L) and (S7 is Normal) and (S8 is Normal) then (system is E8) (1)*

In terms of artificial intelligence, these rules actually summarize the experience of the expert and they are generally not definable in a unique way since each individual creates his own rules.

2.8.2.4 Fuzzification of Inputs and Outputs:

Define the membership functions for inputs and output variables. On the fuzzification step, after many tests, we choose the triangle function as the best membership function can be used for the faults diagnosis on distillation column data. When the system becomes infected with an accident, this means that a change will occur on all the signals acquired from it or on a part of them as explained in chapter 1. This change in signal means that it has deviated from its normal course to another path that may be lower or higher than normal. Therefore we divide the range of the signal into three parts: normal, lower than normal and higher than normal.

The following figures (Figure 2.15, Figure 2.16, Figure 2.17, Figure 2.18, Figure 2.19 Figure 2.20, Figure 2.21 and Figure 2.22) illustrates the degree of membership applied to each input signal S_i

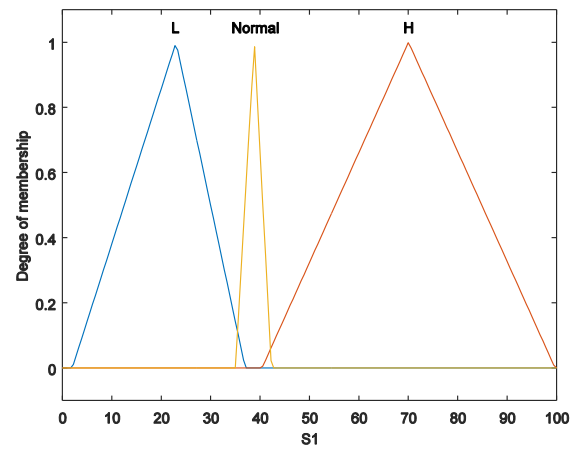


Figure 2.15: Fuzzification of Pre-heated Temperature (S1)

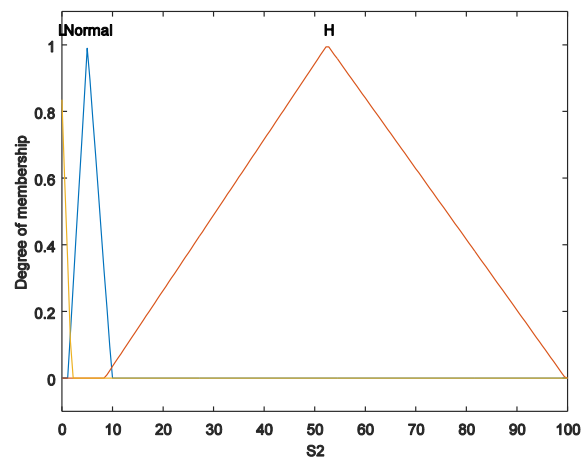


Figure 2.16: Fuzzification of the Timer: Reflux Rate (S2)

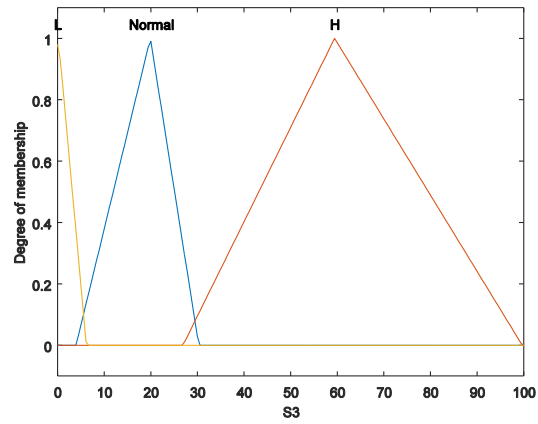


Figure 2.17: Fuzzification of Pre-heating Power (S3)

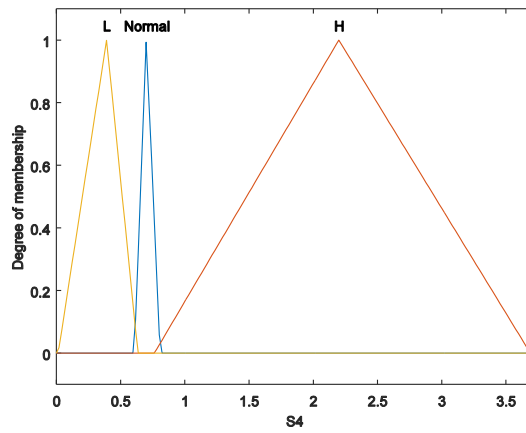


Figure 2.18: Fuzzification of the Loss of Charge (S4)

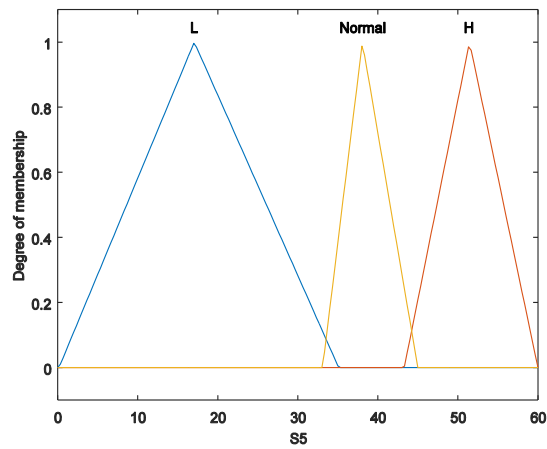


Figure 2.19: Fuzzification of the Heating Power (S5)

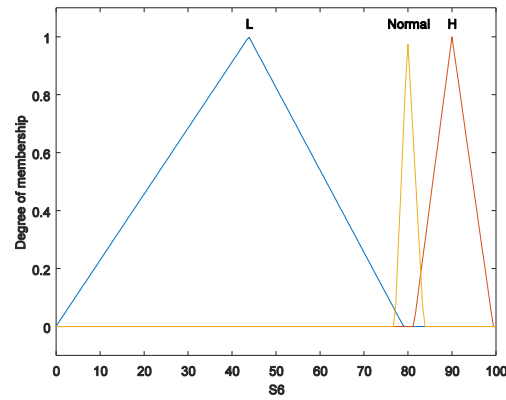


Figure 2.20: Fuzzification of the Feed Flow Rate (S6)

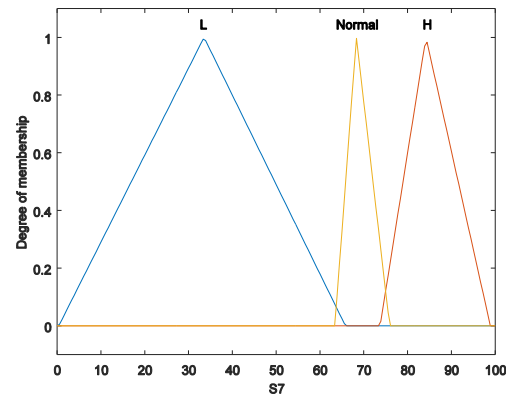


Figure 2.21: Fuzzification of the Boiler Temperature (S7)

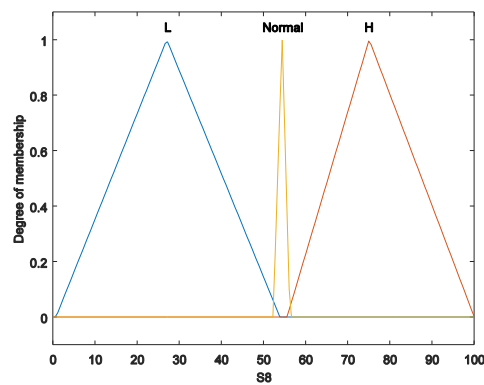


Figure 2.22: Fuzzification of the Column Head Temperature-TIC2 (S8)

The following figure (Figure 2.23) illustrates the degree of membership applied to each accident output E_i

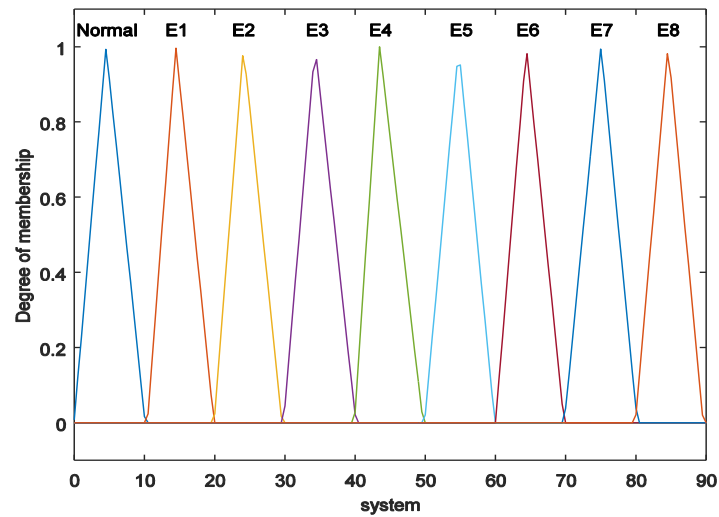


Figure 2.23: Fuzzification of the System Output

On the output of the system, we have 9 outputs where 1 is normal and the 8 others are the accidents modes (E1 to E8) (Figure 2.30), the table below (Table 2.2) represents the range of each output mode.

Mode	Range
Normal	[0 10]
E1	[11 20]
E2	[21 30]
E3	[31 40]
E4	[41 50]
E5	[51 60]
E6	[61 70]
E7	[71 80]
E8	[81 90]

Table 2.2: The Range of Each Output Mode

The Figure 2.24 represents the whole system

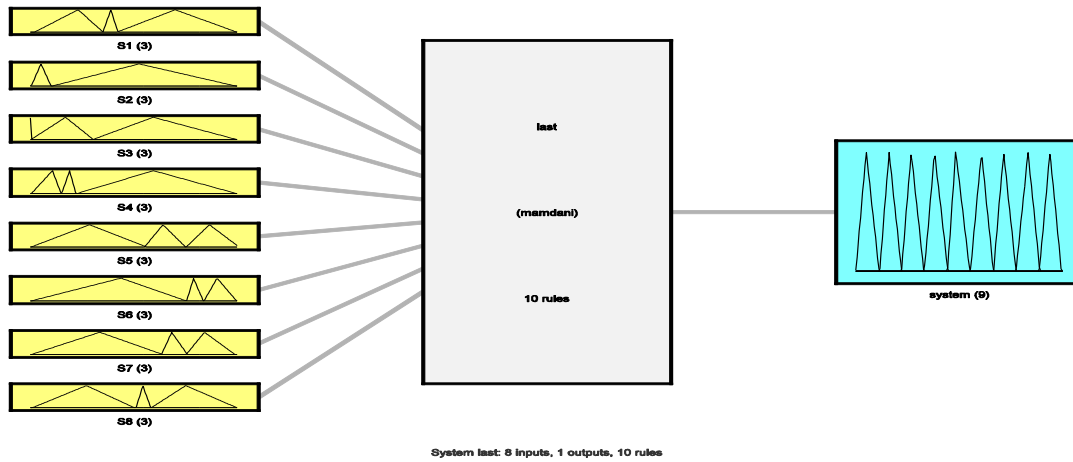


Figure 2.24: Fuzzy Logic Applied to Distillation Column Data

Figure 2.25 represents the rules viewers of the inputs and outputs

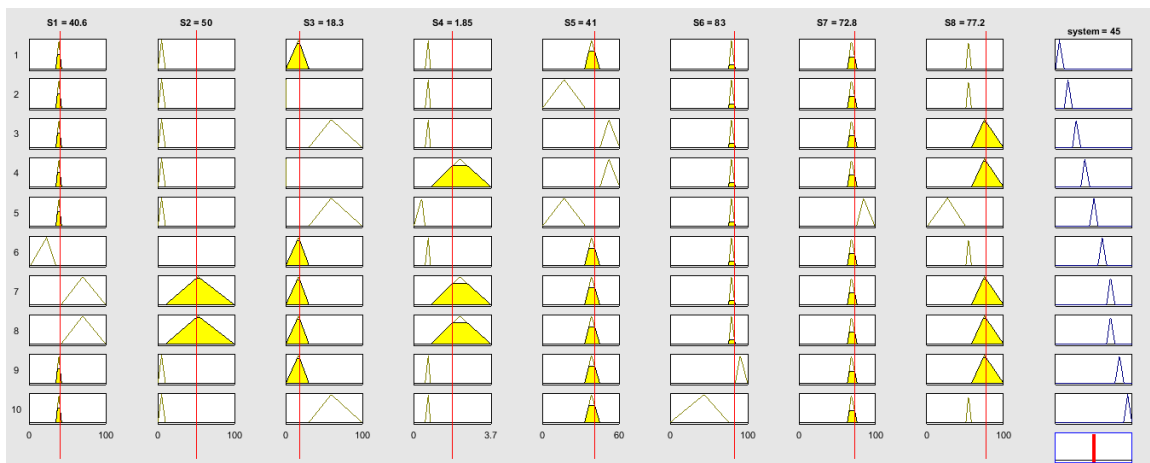


Figure 2.25: Rule Viewer of the Whole System

The Rule Viewer is a displaying of all parts of the fuzzy inference process from inputs to outputs. In the Figure 2.25, each row of graphs represents one rule and each column $C_i, i = 1, \dots, 8$ represents to either an input variable $S_i, i = 1, \dots, 8$. The last column represents an output variable that means the decision result of the fuzzy logic algorithm.

2.8.2.5 Defuzzification

At the end of inference, the output fuzzy set is determined but it is not used directly to provide accurate information to the operator or control an actuator.

It is necessary to pass from the "fuzzy world" to the "real world", it is the defuzzification. There are several methods of defuzzification: the center of gravity of the surface, the bisector of the surface, the average of the maxima, the smaller of the maxima in absolute value, the greater of the maxima in absolute value. The most often encountered is the method of calculating the "center of gravity" of the fuzzy set (eq.2.10).

$$x_0 = \frac{\sum_{i=1}^n \mu_i x_i}{\sum_{i=1}^n \mu_i} \quad (2.10)$$

With:

n : Level of discretization

x_i : i th exit.

μ_i : Membership value of the i th output.

2.8.2.6 Results

The method of fuzzy logic is applied on 50 data observations from each mode, the total number of observations is 50. Figure 2.26 illustrates the results of the classification between the modes or faults types that may occur in distillation column system, the blue curve represents the real output and the green one is the desired output and the red points represent the observations that have been incorrectly classified. As you see in Figure 2.26 the vast majority of these observations are correctly classified that means fuzzy logic allows diagnosing the mode where the fault is originated and classify the type of faults occurred during the distillation process.

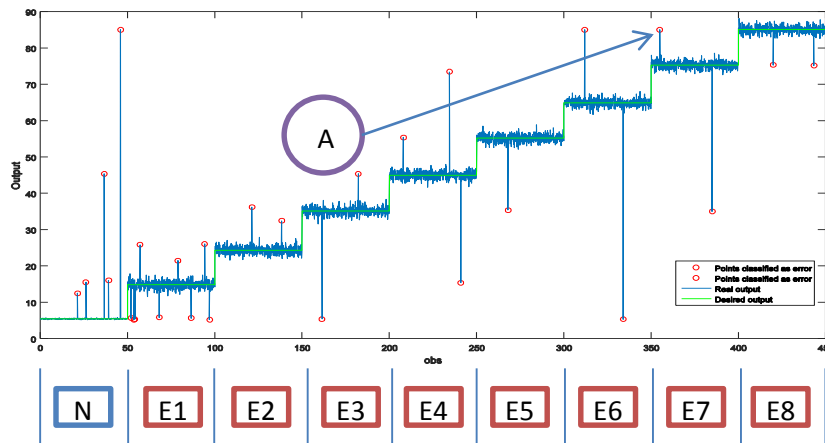


Figure 2.26: Outputs of Fuzzy Logic when it applied to the Distillation Column Data

The figure 2.27 represents the absolute difference (eq.2.11) between the real output and the desired output on each mode.

$$difference_i = |R_i - D_i| \quad (2.11)$$



Where R_i is the real output of the mode i and D_i is the desired output of the same mode i .

The results shown in the Figure 2.27 allow that this difference is very small except in the case in which an observation is incorrectly classified. For example, the observation A in figure 2.26 should be classified as E6 but it's classified as E8.

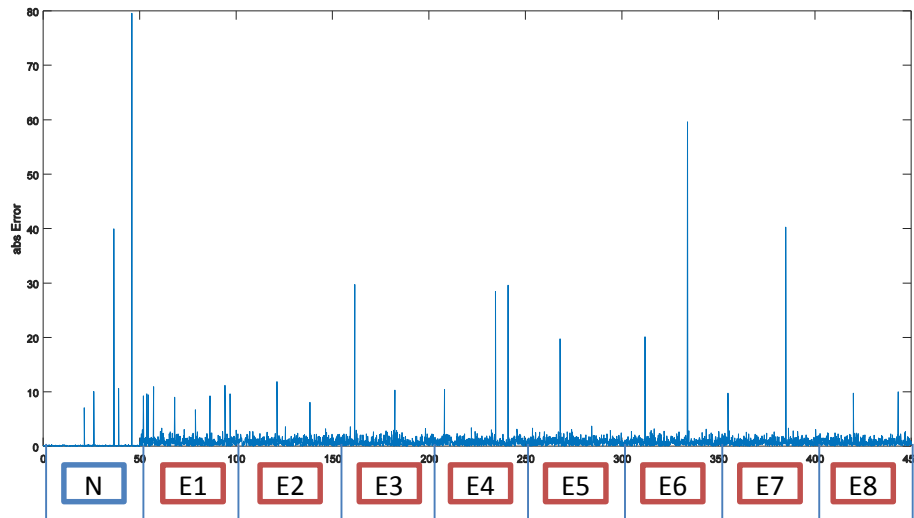


Figure 2.27: Difference between Real and Desired Outputs

The figure 2.28 represents the percentage of the observation that correctly classified in terms of the others that incorrectly or wrongly classified, as you see we can say that 94% of the data are correctly classified.

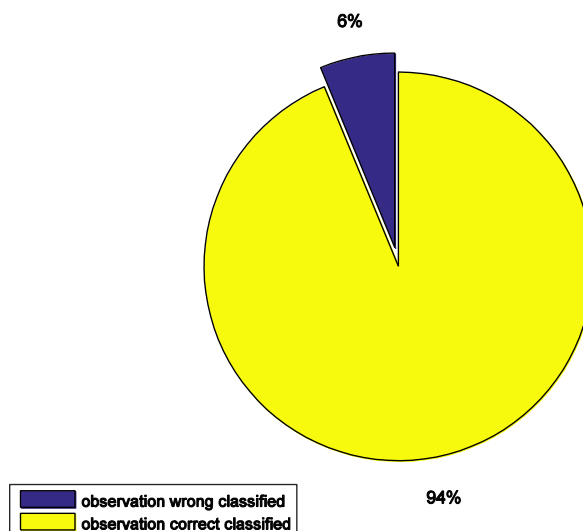


Figure 2.28: Percentage of the Observation that Correctly Classified in Terms of the Observations that Incorrectly Classified



In general, one can notice clearly the difference between normal and faulty modes, also the difference between the different cases of faults. 6% of error percentage is considered good but not very excellent. We have seen that the designer of a fuzzy system has to make a lot of choices. These choices are mainly based on the expert's advice or on the statistical analysis of past data, in particular, to define the membership functions and the decision matrix.

Thus, all the power of fuzzy logic is to make possible the establishment of inference systems whose decisions are seamless, flexible and non-linear, closer to human behavior than is conventional logic. In addition, the rules of the decision matrix are expressed in natural language. This has many advantages, such as including the knowledge of a non-computer expert at the heart of a decision-making system or modeling more finely some aspects of natural language.

2.8.3 Distillation Column Faults Diagnosis Based on Artificial Neural Network

Normally, in the domain of faults diagnosis, artificial neural networks can be directly applied to the original signals acquired from the system, but their effectiveness is very limited and their results are unsatisfactory. In addition, in general, the original signals are large and this leads to the inaccurate diagnosis. Therefore, the extraction of signal parameters or features is an essential step before classification. This is because it is harmful and necessary to extract the relevant, discriminant and most adapted parameters to the signal.

In general, the illustration of a diagnosis system based on ANN is as follows (Figure 2.29)

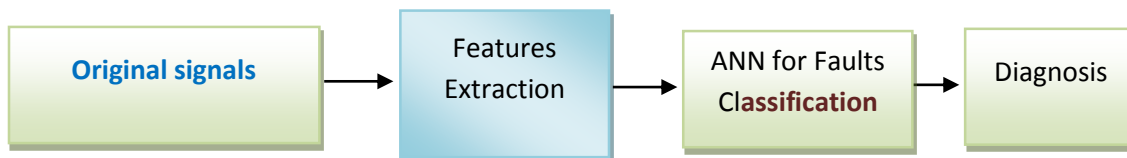


Figure 2.29: General Diagnosis System Based on ANN

2.8.3.1 Features Extraction Applied to the Distillation Column Data

The following table (Table 2.3) presents the results of features extraction when it's applied to a normal real data acquired from distillation column system.

Signal	Mean	Max.	Min.	Variance	Frequency peak
Preheated temperature (°C)	40	40.6	39.6	0.006	0.0006
Timer %	0.5	7	0	1.6	0.001
Charge loss (mbar)	0.7	0.7	0.7	0	0.0001
Pre-heating power	2.8	7	0	6.7	0.0003
Heating power	42.7	42.9	42.3	0.009	0.0001
Flow %	80	80	80	0	0.0001
Boiler temperature (°C)	76	76.1	75.8	0.009	0.0001
TIC2 (°C)	56	56.1	55.9	0.001	0.0001

Table 2.3: Statistical Analysis of Distillation Column Data in Normal Mode

2.8.2.2 Discussion

The statistical analysis of normal real signals acquired from distillation column signal table 2.3 , proves the need to skip the classical feature extraction step (time and frequency domain features), because it was noticed that most time parameters are equal and the signals are deterministic with poor frequencies (Figure 2.30 and Table 2.3) and thus couldn't provide significant results [3].

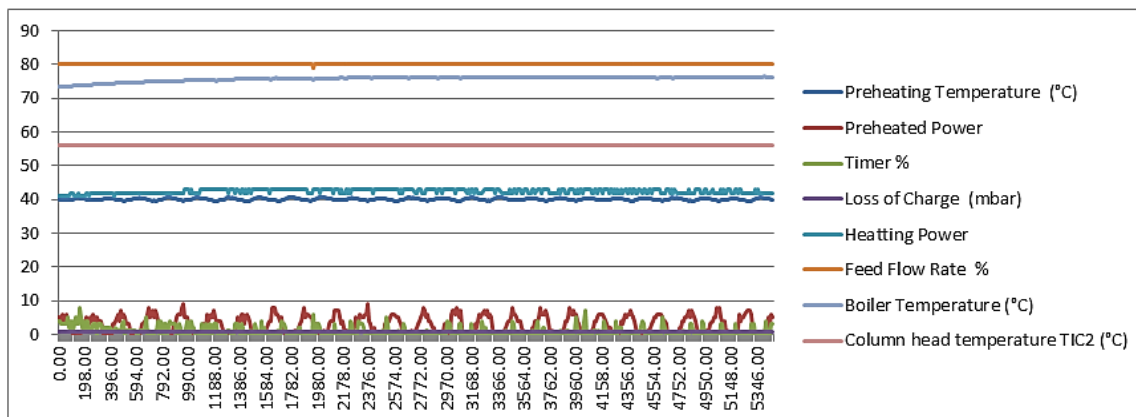


Figure 2.30: Graphical Representation of the Signals (Normal Mode)

the skipping of the features extraction step put the application of ANN in the distillation column data in an embarrassing position, therefore it is necessary to search another solution can minimize the percentage of incorrect fault classification obtained when we apply the fuzzy logic on the real signals (Figure 2.30) and Solve the problem of impossibility of application of features extraction on the original signals (Table 2.3).



2.8.4 Modified Fuzzy c-Means Combined with Neural Network Based Fault Diagnosis as New Approach for a Distillation Column

The previous unsatisfactory results obtained when we apply the fuzzy logic or ANN individually on the data it encourages us to propose an approach more efficient in real-time analysis of distillation column system. It proposes a methodology that combines fuzzy mean clustering and neural network for diagnosis, detection, and classification of many faults. Moreover, a modified FCM method (MFCM) is presented in place of a feature extraction and selection approach. MFCM is a clustering method that allows calculating the degree of variation between normal and abnormal modes. The output of the MFCM is considered as inputs for the neural network classifier. This methodology is tested to a real experimental data obtained from a distillation column, after a pre-processing step including filtering and smoothing of the signals. A database with normal and faulty observations is analyzed. The database is composed of eight different types of faults (E1 to E8) that may occur during the automated distillation process in the chemical industry.

2.8.4.1 Fuzzy c-Means Clustering (FCM)

Fuzzy c-means (FCM) is a method of clustering that allows one piece of data to belong to two or more clusters. This method is proposed by Bezdek et al. in 1981 [161] it is frequently used in pattern recognition. The algorithm (Algorithm 2.3) works by assigning membership to each data point corresponding to each cluster center on the basis of the distance between the cluster center and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, the summation of membership of each data point should be equal to one. After each iteration, membership and cluster centers are updated and minimized.

We consider a space X (eq.2.12) with (n) points and has (p) dimension:

$$X : \begin{bmatrix} x_1^1 & \dots & x_1^j & \dots & x_1^p \\ \dots & \dots & \dots & \dots & \dots \\ x_i^1 & \dots & x_i^j & \dots & x_i^p \\ \dots & \dots & \dots & \dots & \dots \\ x_n^1 & \dots & x_n^j & \dots & x_n^p \end{bmatrix} \quad (2.12)$$

It is assumed that the N points can be regrouped into c clusters $c < n$. The clusters are described by the following set of centers V_i (eq.2.13):

$$V_i = [v_i^1, v_i^2, \dots, v_i^j, \dots, v_i^p], 1 \leq i \leq c \quad (2.13)$$

Consider the following proximity matrix U_{ik} (eq.2.14) that represents the membership degree of the point X_k in the center V_i



Normandie Université

$$U = \begin{bmatrix} u_{11} & \dots & u_{1k} & \dots & u_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ u_{ij} & \dots & u_{ik} & \dots & u_{in} \\ \dots & \dots & \dots & \dots & \dots \\ u_{c1} & \dots & u_{ck} & \dots & u_{cn} \end{bmatrix}, k = 1, \dots, n \text{ \& } i = 1, \dots, c \quad (2.14)$$

FCM Algorithm

The FCM algorithm is then described by the following steps:

- 1- Iteration membership (eq.2.15)

$$\mu_{kl} = \frac{1}{\sum_{i=1}^n (d_{kl}/d_{ki})^{2/m-1}} \quad (2.15)$$

- 2- Cluster center selection (eq.2.16)

$$V_l = \frac{\sum_{k=1}^n (\mu_{kl})^m x_k}{\sum_{k=1}^n (\mu_{kl})^m}, k = 1, 2, \dots, c \quad (2.16)$$

***n** is data points number*

***V_l** is the **lth** cluster center*

***m** is the index of fuzziness **m** ∈ [1, ∞]*

***c** is the nb of cluster centers*

***μ_{kl}** is the membership of **kth** & **lth** cluster center*

***d_{kl}** is the Euclidean distance between*

***kth** & **lth** cluster center*

- 3- Minimization of the following equation (eq.2.17):

$$J(U, V) = \sum_{k=1}^n \sum_{l=1}^c (\mu_{kl})^m \|x_k - v_l\|^2, \quad (2.17)$$

***||x_k - v_l||** is the is the Euclidean distance*

*between **kth** & **lth** cluster centers.*

Algorithm 2.3: FCM Algorithm

2.8.4.2 Feed Forward ANN

Feedforward neural networks (FFNN) are the most popular and most widely used models in many practical applications. This network spreads the network input to the following layers and never goes backward. This type of network is used in this study. In addition to the neural network type, it is necessary to choose an error function and an activation function for the neurons. These choices are often guided by the type of data processed. In this study, at the level of the hidden layer, the activation function used is a logistic function defined by (eq.2.18):



Normandie Université

$$b_i^{(j)}(x) = 1 / (1 + \exp(-x)) \quad (2.18)$$

The linear transfer function is used at the level of the output layer. For the error function, a simple choice is used (eq.2.19):

$$E = \frac{1}{2} \|y - t\|_2^2 = \frac{1}{2} \sum_{j=1}^n (y_j - t_j)^2 \quad (2.19)$$

This equation (eq.2.19) is half the square of the Euclidean distance between the network output (y) and the target (t). It is now necessary to minimize the average error data by the function E on all data input (eq.2.20):

$$E_{average} = \frac{1}{N} \sum_{i=1}^N E_i \quad (2.20)$$

Where N is the number of training data given to the neural network and E_i represents the i^{th} learning error [3].

2.8.4.3 Proposed Methodology

For fault detection, we decided to work with artificial neural networks. The statistical analysis of real signals proves the need to skip the classical feature extraction step (time and frequency domain features) because it was noticed that most time parameters are equal and the signals are deterministic with poor frequencies (Figure 2.30 and Table 2.3) and thus couldn't provide significant results. Therefore, we find the solution to this problem is to use the output of modified FCM as input layer for neural network classifier.

Figure 2.31 presents the block diagram of our proposed methodology. The first step consists of database development. This step was detailed in the previous section. In the second step, a modified FCM clustering algorithm (MFCM) is proposed after a pre-processing step including filtering and smoothing of the signals. MFCM tries to put each of the data points to one of the clusters. What makes MFCM different is that it calculates the degree of variation between normal and abnormal modes followed by the calculation of the Euclidean distance between two cluster centers.

MFCM divides the data into S segments and for each segment, it gives a vector of Euclidean distances ϑ (Figure 2.32 and Figure 2.33).

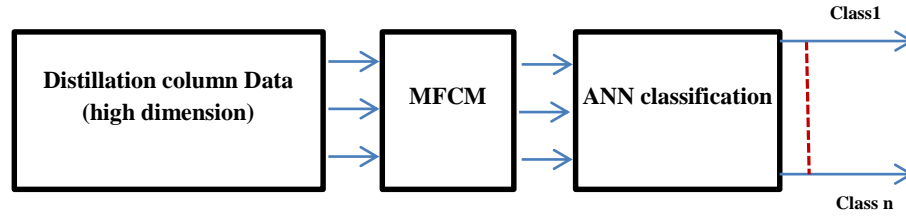


Figure 2.31: Block Diagram of Our Proposed Methodology

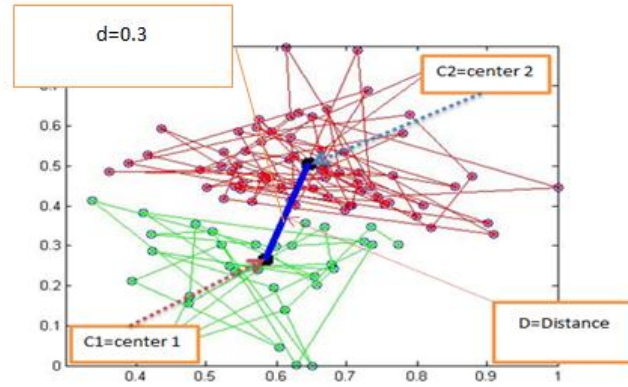


Figure 2.32: Result of Clustering in the Same Mode (No-Fault)

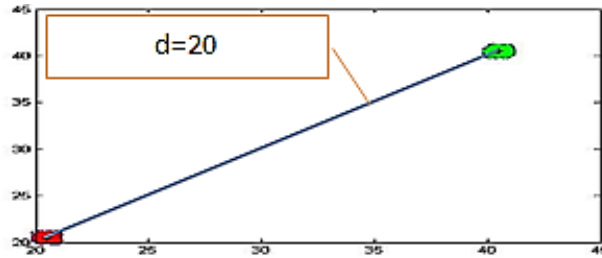


Figure 2.33: Distance between Two Clusters in Different Modes (Fault Mode)

The vector θ contains 8 elements $\theta=[d1, d2, d3, d4, d5, d6, d7, d8]$. In the end, each element represents a vector of the degree of variation in distance of a signal. This could provide information if we stay in the same mode (normal-normal) or pass to another mode (normal-fault).

MFCM is detailed in the next subsection. In our case, we have two clusters groups so we have two centers $c1$ and $c2$ (eq.2.21)

$$C1 : (x_1, y_1) \text{ And } C2 : (x_2, y_{21}) \quad (2.21)$$

The distance between the two centers is calculated as (eq.2.22):

$$d = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \quad (2.22)$$



After extracting the vectors of Euclidean distances, we then used neural networks classifier because of their usefulness for the problem of fault diagnosis. The neural network classifier is based on the extracted Euclidean distance vectors, which are considered as the feature set. The aim of using MFCM is to decrease the calculation time and increase the performance of the ANN classifier. The architecture of the ANN model uses the feed forward neural network as a training scheme. Finally, we make a real application on faulty conditions simulated using the column system discussed previously. The system consists of eight different fault conditions. The Modified FCM algorithm (MFCM) is summarized in Algorithm 2.4. First of all, the algorithm starts with initial positions of the cluster centers. It then calculates and optimizes the data points. More the data is near to the cluster center more its membership towards the particular cluster center. Thus, a matrix U is formed. This matrix represents the final clustered groups.

2.8.4.5 Proposed MFCM Algorithm

MFCM Algorithm

- 1) E is the vector of Euclidean distances
- 2) Start $E = [\emptyset]$
- 3) X is the data (eq.2.23)

$$X : \begin{bmatrix} x_1^1 & \dots & x_1^j & \dots & x_1^p \\ \dots & \dots & \dots & \dots & \dots \\ x_i^1 & \dots & x_i^j & \dots & x_i^p \\ \dots & \dots & \dots & \dots & \dots \\ x_n^1 & \dots & x_n^j & \dots & x_n^p \end{bmatrix} \quad (2.23)$$

- 4) Apply moving window after the dividing of the data X to S segments $S=[S_1, S_2, \dots, S_n]$
- 5) For each segment S_i calculate the position of the two centers (eq.2.24)

$$V_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}, \quad i = [1,2] \quad (2.24)$$

- 6) calculate the matrix : $U^{(l+1)}$ depending to formula (2.14)
- 7) if $\|U^{(l+1)} - U^{(l)}\| < \varepsilon$ for each S_i
stop
else
l=l+1 and return to 4)
- 8) calculate the Euclidean distance (eq.2.25) between the calculated centers in S_i ,
 $E = [d_1 \ d_2 \ \dots \ d_n]$ (2.25)
- 9) $S_i = S_{i+1}, \quad i < n$
- 10) End

Algorithm 2.4: MFCM Algorithm

2.8.4.6 Classification with Multi-layer Feed Forward ANN

In this research, we use the multi-layer feed forward neural network as the first classifier. Our ANN (Figure 2.34) is formed of 3 layers.

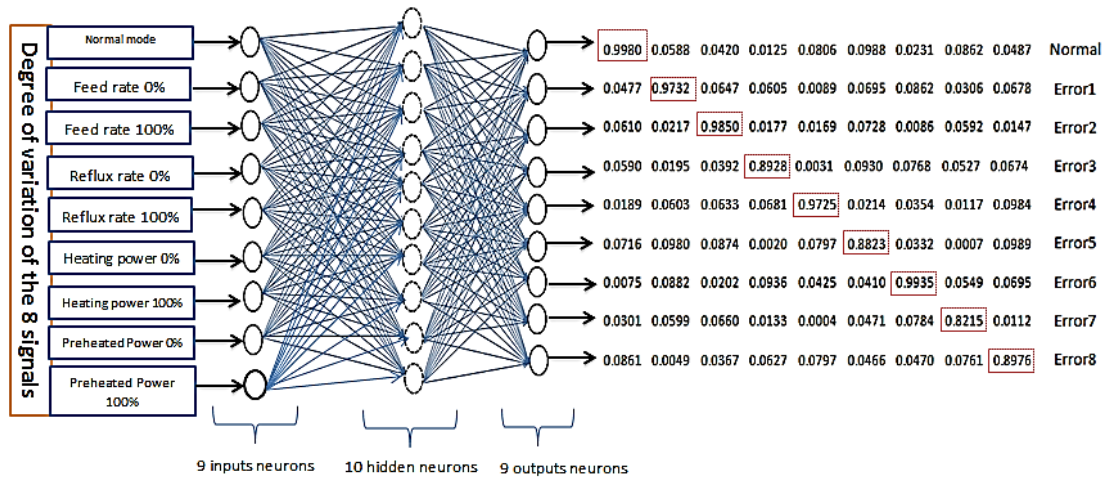


Figure 2.34: Neural Network Proposed Architecture

The first one is the inputs layer that contains 8 inputs representing the degree of variation of 1- the Timer: reflux rate, 2- the Heating power, 3- the Feed flow rate, 4- the preheated Power, 5- the Pressure drop, 6- the preheated temperature, 7- the temperature boiler, and 8- the TIC2 Column head temperature. The second is a hidden layer having 10 neurons with logistic learning function. And the third one is the output layer containing 9 linear neurons (normal or Error1 or Error2 ... Error8). The network is able to find the solution after 61 epochs (Figure 2.35 and Figure 2.36) with learning rate $\mu=0.001$ and the method generates a warning message indicating the state of the system (normal or any Error)

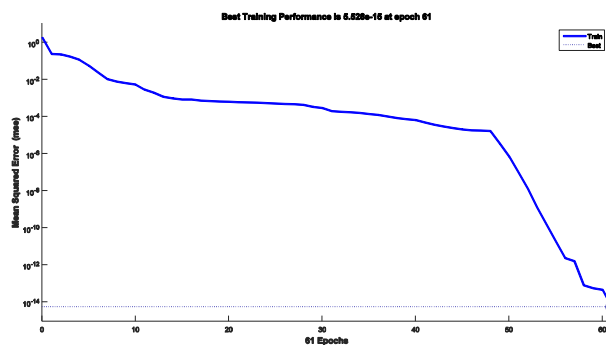


Figure 2.35: ANN Training Performance

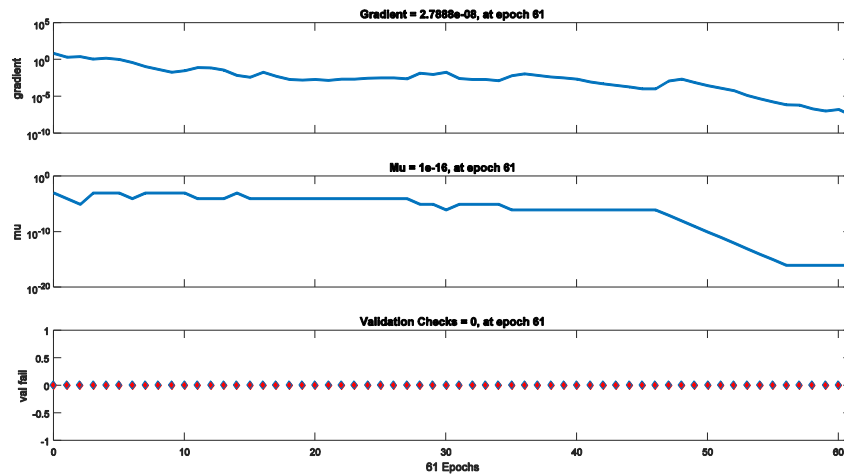


Figure 2.36: ANN Training State

2.8.4.7 Results and Discussion

So far in the previous section, we have presented the main framework for process fault diagnosis. In this section, we provide a real application for a diagnostic purpose. The proposed methodology is designed to detect simultaneously 8 kinds of faults (E1...E8).

- E1: Accident Reflux Ratio 0% (Timer);
- E2: Accident Reflux Ratio 100% (Timer);
- E3: Accident Heating Power 100%;
- E4: Accident Heating Power 0%;
- E5: Accident Preheated Power 0%;
- E6: Accident Preheated Power 95%;
- E7: Accident Feed Rate 100%;
- E8: Accident Feed Rate 0%.

Previously, all researches applied to diagnose and detect the faults on distillation column signals are completely dependent on the modeling techniques to calculate the residual between the real signal and the model. The black box ANN is the most used and is the best technique in this domain[130]. Most of the studies on the detection of the changes in distillation column were carried out by Chetouani et al. [162][160][130] with the most recent study in 2014 where they applied the Bayes decision theory combined with neural adaptive black-box identification for modeling such system [162]. Their works provided an efficient fault detection method when tested on a real distillation process. However, such work models and process each signal independently which needs long time simulation. Further were, the work did not take into account how a fault detected in a signal affects the other signals and the whole system. In some cases, a false alarm may occur because of unexpected faults that may affect the other signals. Then, in this case, the model will fail to detect the true fault. For example, when the Timer is blocked (0%),



this causes the overhead temperature (TIC2) to decrease. The method succeeds in this case to detect the fault. On the other hand, if the boiler has a failed operation status (0% power heating), TIC2 also decreases in this case. This will create a false alarm, not a real fault detected in the system. However, our proposed method is a full scan system i.e., it takes all eight signals at the same time. It will solve the problem of false alarm by analyzing the fault and its influence on all signals. The algorithm makes the analysis in one iteration step and thus needs less time to perform the diagnosis operation.

The results allow diagnosing the mode where the fault is originated and classify the type of faults occurred during the distillation process. The results of the MFCM are presented in figure 2.37. In general, one can notice clearly the difference between normal and faulty modes, also the difference between the different cases of faults. In normal mode (normal case), there are no variations with the absence of peaks. However, for example, in E2 case (accident reflux ratio timer 100%), we notice that the degree of variation for parameters number 3 (timer), 4 (charge loss) and 5 (heating power) increases significantly. Therefore, the degree of variation of the eight parameters (shown in Figure 2.37) could be considered as an indicator of fault, thus could differentiate between normal and abnormal cases and between the eight different types of faults

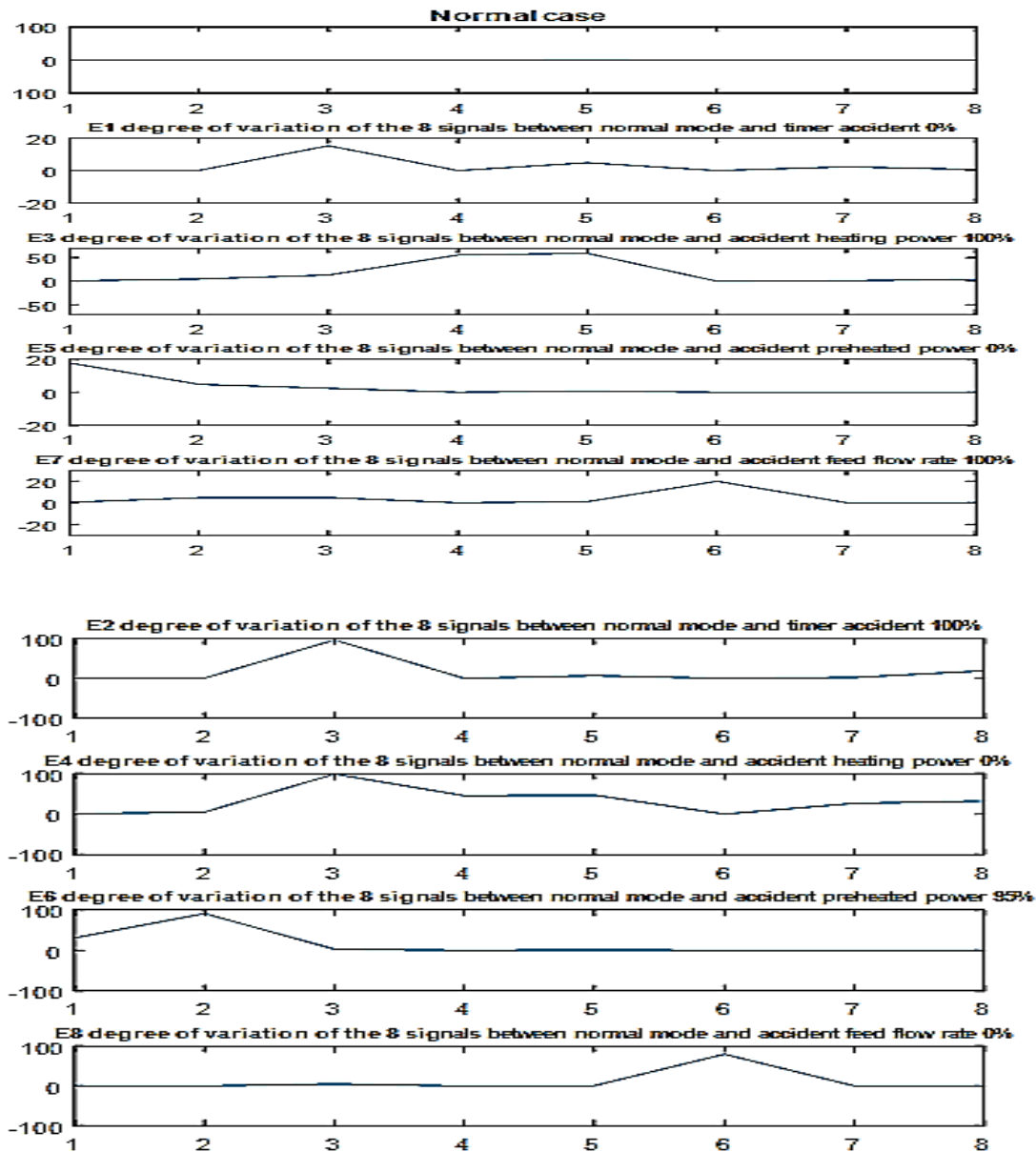


Figure 2.37: Degree of Variation in Processed Signals for Normal and Eight Error Modes. 1- Preheated Temp 2-Preheated Power 3-Timer 4-Charge Loss 5-Heating Power 6-Feed Rate 7- Boiler Temp 8- TIC2.

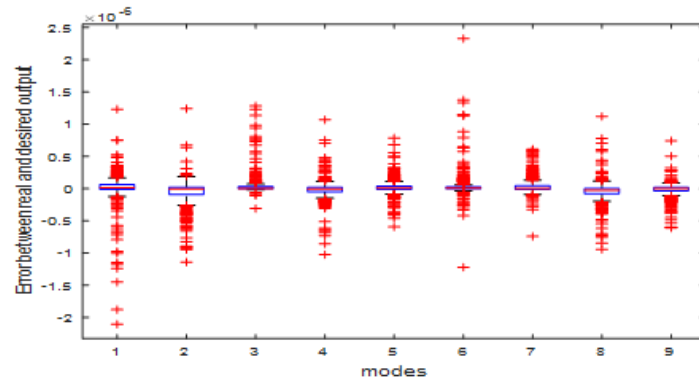
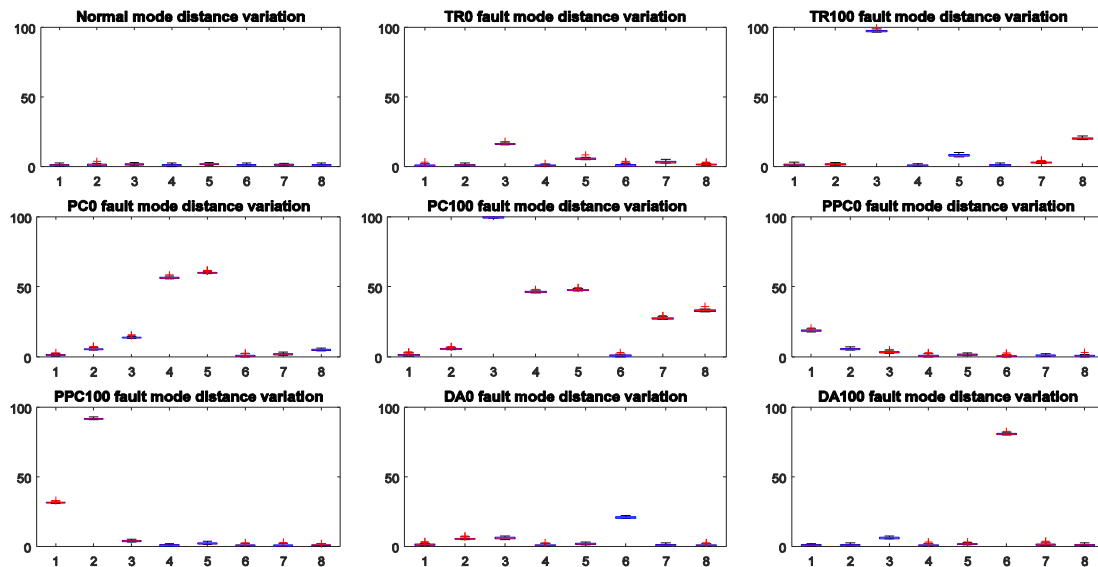


Figure 2.38: Error Boxplot- (Error between Real and Desired Output)

The method is applied to 50 data observations from each mode, the total numbers of observations are 450. 70% are used for training and 30% for testing. Figure 2.39 illustrates the boxplot which shows the difference between the medians of the degree of variation in distance. One can notice clearly the difference between the different modes. Figure 2.38 shows that the error between the real and desired output is negligible (in order of 10^{-6}). This proves that the proposed methodology is successful in detecting faults introduced within the distillation process plant model. The target or output vector is a 9-element vector with a value near to '1' in the position of the fault it represents and a value near to '0' everywhere else (Figure 2.34). In other words, the output value is set to '0' to indicate no fault and '1' to indicate the presence of a fault.



1-Preheated Temperature, 2-Preheated Power, 3-Timer, 4- Loss of Charge, 5-Heating Power, 6-Feed Rate, 7-Boiler Temperature, 8-TIC2

Figure 2.39: Boxplot Showing the Difference between the Median of the Degree of Variation



On the testing step (15 observations from each mode) only 2 observations are incorrectly classified (Figure 2.40) so the percentage of error is $2/135$ so it's equal to 1.4815%.

As you see the percentage of error percentage of the observation that correctly classified was reduced from 6% (Figure 2.28) in the case of faults diagnosis based on fuzzy logic to 1.48% in the case of faults diagnosis based on Modified fuzzy c-means combined with Neural Network, so our proposed method reduces the error rate for 81.5 %. This reducing of error it was due because we have benefited from the combination of the advantages of fuzzy logic with the advantages of ANN in the same algorithm. In addition in our proposed methodology, we have eliminated the role of the user or expertise in determining the number and the type of the membership function used and the intersection level between the inputs in the fuzzification step. Therefore we reduce the source of errors.

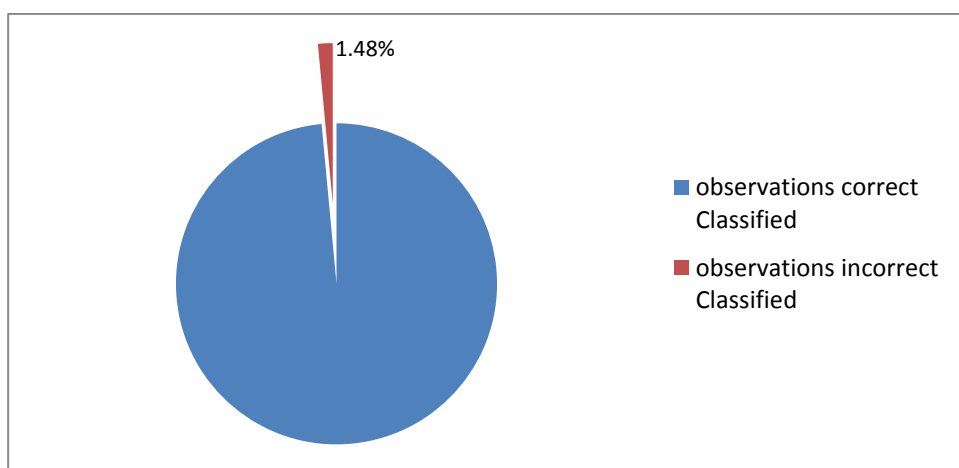


Figure 2.40 Percentage of the Observation That Correctly Classified In Terms of the Observations That Incorrectly Classified

The results showed also proved that the proposed methodology for MFCM, neural network classification, succeed to diagnose and detect in real time the simultaneous faults in real experimental data obtained from the distillation column during the automated continuous distillation process.

The results also confirmed the ability to classify between normal and eight abnormal classes of faults with very low classification error.

2.9 Conclusion

This chapter provided a brief overview of the various diagnostic techniques conventionally used as methods of detection and fault isolation of chemical processes.

A very large number of developments concerned approaches based on quantitative models in the process using the residue calculation technique. However, various factors such as the complexity of the system, the nonlinearity of the processes have often made it very difficult to develop a precise mathematical model to ensure a reliable and robust diagnostic chemical reactor.



In particular, a too simple model not generates residues only representative failures, but they often incorporate modeling errors or drifts of normal parameters. With presence of multi-physical character and a strong coupling of several energy in chemical processes it has been shown that the neural network (by its causal properties, structural and functional) can be an alternative for generating resistant and sensitive indicators several types of faults (sensor, actuators, and measurement process) with knowing that we cannot ignore the merits of the fuzzy logic method that is used heavily in this domain, especially when it coupling with the neural network method under the title neuro-fuzzy faults diagnosis and detection system for chemical process.

The diagnosis is a main component of the monitoring module. It consists in determining at each instant the mode of operation wherein the system is located. It is based on a priori knowledge of the operating modes and instant knowledge materially with a new observation system status. There are several approaches to make the diagnosis; choice of approach is related to the mode of representation of knowledge.

This chapter is mainly intended to present some reminders on basic concepts in diagnosis and the observability of nonlinear systems. The first part proposes a state of art of the different types of faults and their influence on the process to diagnose. On the second side, the different diagnostic methods are presented, they will be grouped into two main categories: methods using mathematical models and the other without a model. This chapter is ended by an application of distillation column to fuzzy logic and ANN separately the results of fuzzy logic when it is applied to the data shown that we can notice clearly the difference between normal and faulty modes, also the difference between the different cases of faults. 6% of error percentage is considered good but not very excellent. We have seen that the designer of a fuzzy system has to make a lot of choices. These choices are mainly based on the expert's advice or on the statistical analysis of past data, in particular, to define the membership functions and the decision matrix.

The nature of the data acquired from the distillation column shown that it is impossible to apply ANN individually to this data. This impossibility and the percentage of error that equal to 6% that is quite large in the case of the application of fuzzy logic on the data were the reasons we were forced to propose a new methodology that combines the benefits of fuzzy logic and ANN in the same algorithm to perform the results and minimize the percentage of error. The results proved that the proposed methodology for MFCM, neural network classification, succeed to diagnose and detect in real time the faults in real experimental data obtained from the distillation column during the automated continuous distillation process and the proposed methodology also succeed in the minimization of the percentage of error from 6% to 1.4815%.

The results also confirmed the ability to classify between normal and eight abnormal classes of faults. As perspective, we plan to propose a prognostics system for the early detection of faults in distillation column process to prevent damage or catastrophic accidents. Due to the importance of identifying the industrial system future status and estimate its remaining useful life (RUL), the following chapter will open the horizons around the faults prognostic and its importance in developing the maintenance strategy.



Normandie Université

Chapter 3

Industrial Prognostics Methods Applied on Chemical Processes (nonlinear systems)- Application to Distillation Column



3.1 Introduction to Industrial Prognostics

One of the most important issues facing the industry today is the operational safety of industrial systems and the search for increased availability at lower costs. Also, the maintenance activity is taking a growing share in companies and tends to evolve for needs of responsiveness and cost in particular. A particular evolution concerns the way to apprehend the phenomena of failure: little by little the industrialists tend, not only to anticipate them in order to resort to preventive actions but in addition to do it in the most just possible way for a purpose reducing costs and risks. This evolution has given a growing part to the prognostic process which is today considered as one of the main levers of action in the search for a global performance. For manufacturers, it is therefore imperative to understand the state of gravity of a fault and to predict the optimal time to stop a machine and intervene. Rather than understand a posteriori a phenomenon that has just manifested (failure ...), it is appropriate to "anticipate" the occurrence in order to resort to protective actions accordingly. This is what can be understood under the label "fault prognostics". As a result, the traditional concepts of preventive and corrective maintenance are gradually being complemented by more responsive and proactive consideration of failures [163]. Thus, the topic of Prognostics and Health Management (PHM) becomes a frame of work of foreground and the potential advantages of the implementation of the prognostics in industrial environments, related to the safety of the work, the economic aspects and the human resources pushed the scientists to s' to interest [164] [165]. Today the "Prognostics and Health Management (PHM)" is considered a key process in maintenance strategies.

3.2 Emergence of the Prognostics in the Maintenance Activity

Conditional maintenance is defined according to the standard [166], as a preventive maintenance based on a monitoring of the operation of the property and/or significant parameters of this operation integrating the resulting actions. This maintenance strategy is based on real-time data analysis of industrial equipment (ex. vibration, temperature, etc.). It aims to detect anomalies in the operation of industrial machines: the discovery of changes in their characteristics foreshadows in the short term a future failure. Conditional maintenance makes it possible to better take into account the conditions of use of equipment than traditional systematic maintenance. This being the case, it does not make it possible to dimension with certainty the maintenance policies: the date of occurrence of the failure remains uncertain. Predictive maintenance aims to overcome this lack of knowledge. It is defined as follows [166]:

Predictive maintenance is a conditionally performed maintenance based on the extrapolated forecasts of the analysis and evaluation of significant parameters of the degradation of the property. The idea is to project the current state of the property in the future, to estimate the operating time before the failure, and thus to better size the maintenance policies. Predictive maintenance is thus more dynamic. It takes into account the current conditions of the equipment and tries to predict the evolution in the time of the state of the property.



Predictive maintenance can detect anomalies on machines before they become too serious. The strength of predictive maintenance is therefore to anticipate breakdowns. This avoids any expensive shutdown of the production line.

If predictive maintenance emerges, it is now possible to pick up weak signals on the machines. It then remains to trace the data and analyze them. This analyzes help increase customer satisfaction and save money.

According to a McKinsey study, predictive maintenance will save businesses \$ 630 billion by 2025. These savings will be made possible by several factors. First, a reduction in maintenance costs of 10 to 40% then, reducing the number of breakdowns by half. Finally, decreasing the amount invested in the new machinery of 3 to 5% by increasing the duration of lives of existing machines.

It is, therefore, a bright future promised by the famous consulting firm to the world of industry. But if predictive maintenance brings a break from what is done today, it is first necessary to study the most widespread maintenance strategies today.

Maintenance interventions being planned with greater precision and accuracy, the forecasting maintenance must allow making substantial savings and has been the object of a growing attention in recent years. The expected benefits are indeed numerous like:

- Reduction of the number of breakdowns,
- Reliability of productions,
- Improved staff safety and company image,
- Reduced downtime of equipment (expensive),
- Increased business performance.
- ...

The following section outlines the strategy and highlights the importance of the prognostics in this approach.

3.2.1 Predictive Maintenance and Prognostics

3.2.1.1 Diagnostic vs Prognostic

The maintenance activity traditionally uses various business processes aimed first, to "perceive" certain phenomena (detection), then to "understand" (diagnosis), and finally, to "act" accordingly (choice of 'actions of control). Also, rather than understand a posteriori phenomenon that has just manifested (failure), it may be opportune to "anticipate" the appearance in order to resort to protective actions accordingly: this is what can be heard under the label "prognostics of failures". The relative positioning of these processes of "detection", "diagnosis" and "prognostic" is schematized on (Figure 3.1.a). From a phenomenological point of view, their complementarity can be explained as follows [167] (Figure 3.1.b):



- The detection aims to identify the mode of operation of the system and its state,
- When a failure has occurred, the diagnostic can isolate and identify the component that has stopped working (backward propagation: from effects to the causes),
- The prognostic is aimed to predict the future states of the system (forward propagation: from causes to effects).

Thus, the main industrial maintenance activities listed here complement each other perfectly (this point is slightly developed later). However, they naturally rely on an understanding of failure phenomena. This can be tricky when the dynamics of the equipment are marked or when their conditions of use are variable: the modeling of the phenomena can sometimes be difficult to ensure by an expert and it is advisable to instrument the systems to collect field data allowing in the long term to deploy the different maintenance processes. Also, it appears that the prognostic process, although essential for deploying a predictive maintenance strategy (from causes to effects), should not be considered in isolation, but should be seen as part of a more global complex process.

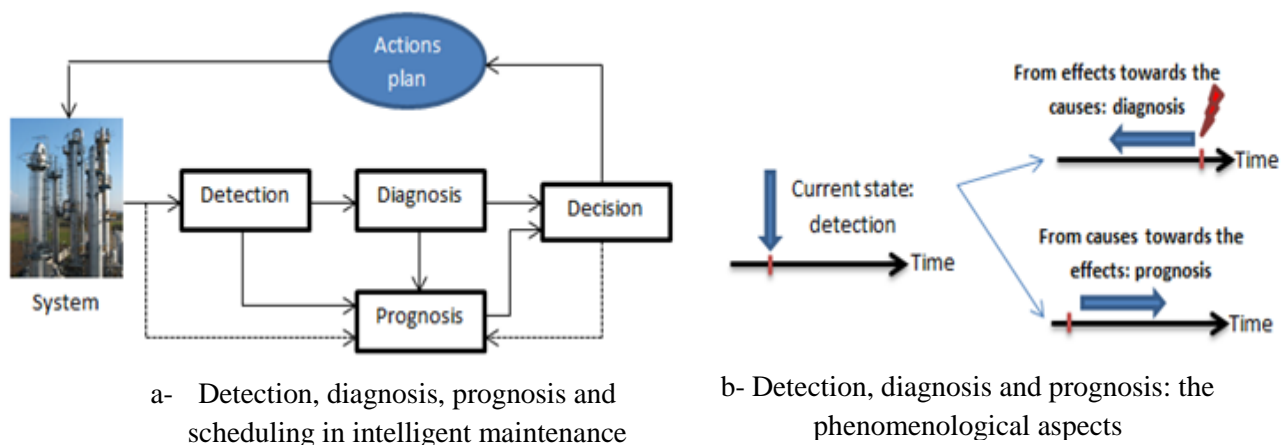


Figure.3.1: Complementarity of Detection, Diagnosis and Prognostics Activities

3.2.1.2 Concept of Prognostics

It has been a long time since man seeks to anticipate phenomena in order to see his projects succeed without risk of failure. The term "prognostics" comes from the Greek "progignoskein" meaning "know in advance". It was then mainly used in medicine where it refers to a prediction made following a diagnosis: the medical prognosis is on the one hand on the evaluation of the degree of severity of pathology, and on the other hand, on the estimate of the subsequent course of the disease. Recently, this term has been transposed to the industrial world. The patient is replaced by a machine or an industrial facility and the objective is to predict the future state of operation of the equipment concerned. Although the "prognostics" is the subject of an international standard (ISO 2004), there are some differences in the interpretation of this concept. The term is also absent from the list of keywords of companies IFAC (International Federation of



Automatic Control) and IEEE (Institute of Electrical & Electronics Engineering), referring organizations in the scientific community.

In this context, a first problem that comes to us is to determine the general framework of the concept of prognostics, while starting with the definition of the concept. The process is difficult because the terms "prediction", "forecast" and "prognostics", distinct in French literature have the same equivalent in the English translation: "forecasting".

The forecast is defined as the estimate of future conditions of the phenomena for a given period from past and present observations. Its overall objective is to provide the best estimates of what can happen at a given point at a specified future date [168]. The prediction is estimating future conditions without reference to a specific time [169].

Depending on requirements and maintenance constraints, prognostics was associated with the following terms:

- ✓ In 2003 Luo et al. [170] and In 2002 Yan et al. [171] assimilate the prognostics by RUL (Remaining Useful Life) or the remaining time of functioning, the remaining time before observing a failure, the remaining useful life, the life remaining or residual life. The term RUL will be extended thereafter and will be implicitly associated with the "prognostics". The definition of the most widely used prognostic process in this acceptance is the following: "a process that aims to predict the number of hours remaining before failure relative to the current time and operating history" [165]. In this context, the definition of failure is crucial to the interpretation of RUL.
- ✓ Another analogy between prognostics and existing terms, this time in a probabilistic manner, attempts to predict the chance that a machine works without fault or failure until a certain date. In the general context of maintenance, this "probabilistic value prognostics" is more an interesting indication that the fault or failure can have catastrophic consequences (ex. nuclear power plants). However, a small number of papers stand this association [172] [173].

The different interpretations assigned to the prognostics in the literature:

- ✓ "After detecting the degradation of a component or subsystem, the role of the prognostics is to predict the future evolution of production system performance, taking into account maintenance work planned and possibly operational conditions or changing environmental ".
- ✓ "The aim of the prognostics (predictive diagnostics) is to identify the causes and locate the bodies led to a particular degradation" [174]. In this definition, the prognostic is considered a predictive diagnosis, which is exerted on the degradation and not the fault like the classic diagnosis. Moreover, it is not adapted to the context of predictive maintenance because it does not include the proactive dimension of the approach.



- ✓ "Prognostic is the ability to predict the future condition of a machine based on the current diagnostic state of the machinery and its available operating and failure history data" [175].
- ✓ "Prognostic is the ability to perform a reliable and sufficiently accurate prediction of the remaining useful life of equipment in service. The primary function of the prognostic is to project the current health state of equipment into the future taking into account estimates of future usage profiles" [81].
- ✓ "Diagnosis and prognostics are processes of assessment of a system's health. Diagnosis is an assessment of the current (and past) health of a system based on observed symptoms, and prognostics is an assessment of the future health"[176].
- ✓ "In the industrial and manufacturing areas, prognostics is interpreted to answer the question: what is the remaining useful lifetime of a machine or a component once an impending failure condition is detected and identified" [176].

The prognostic is normally intuitive and based on experience. The prognostic is generally effective for faults and failure modes with known, age-related, or progressive deterioration characteristics, the simplest being linear. A failure must be defined in terms of monitored parameters or descriptors. Surveillance data alone is insufficient to establish a prognostic.

The general conceptual basis of a prognostic process is (ISO, 2004):

- Define the limit point (usually zeroing),
- Establish the current gravity,
- Determine or estimate the behavior of the parameters and the expected speed of deterioration,
- Determine the estimated duration of operation before failure.

It is important to understand that the diagnosis is, by nature, retrospective and focused on data existing at a given moment. However, the prognostic is focused on the future and, therefore, must take into account the following aspects (**ISO 2004**):

- Existing failure modes and deterioration rates,
- The criteria for triggering future failure modes,
- The role of existing failure modes in triggering future failure modes,
- The influence between existing failure modes and future modes and their deterioration rates,
- The sensitivity to detection and modification of existing and future failure modes due to current monitoring techniques,
- Design and changes of monitoring strategies to fit all of the above,
- The impact of maintenance actions and / or operating conditions,
- The conditions or hypotheses in which the prediction remains valid.

Many definitions of the term prognostics have been proposed [176] [174] [175] and there exists no totally consensual. A striking feature, however, can be identified: the prognostic is often



likened to a prediction process (a future situation must be understood). As a result, two large acceptances prognostics can be considered unifying: the prognostics refer as appropriate a process to determine the remaining life of a system, that is to say, its RUL (Remaining Useful Life) [165] (Figure 3.2.a), or the probability that the system will work for a certain time [172] (Figure 3.2.b).

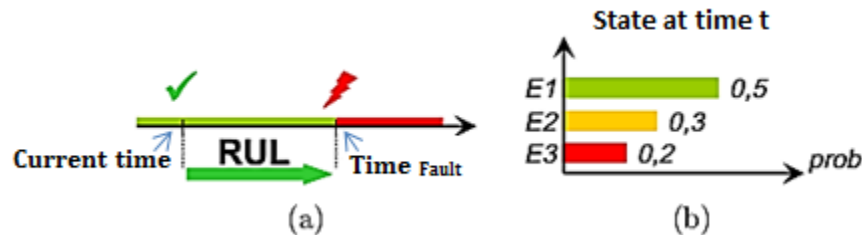


Figure: 3.2 (a) Prognostics as RUL estimation, (b) State probabilities

In conclusion, there are many definitions and interpretations of the prognostics in the literature. In all cases, the authors add comments and remarks to the definition given by ISO. On the other hand, all the definitions described above associate the prognostics with a prediction process. This obviously assumes that the current situation can be grasped (practically, the prognostic is the synthesis of a detection method and measured data of the system). In addition, these approaches are based on the notion of failure (or default), which implies that the prognostic is associated with a degree of acceptability (a system must perform a required function). We thus consider that the prognostic should be based on the evaluation criteria, the limits of which depend on the system itself and the execution objectives.

3.3. Contribution to the Formalization of the Prognostics and to the Development of an Application Framework

3.3.1 Formalization of the Prognostic Process

All proposed definitions equate the prognostic with a "forecasting process": a future situation must be identified. In addition, these interpretations of the prognostic are based on the notion of failure (or default), which implies that the "predicted" situation is associated with a degree of acceptability (a system must perform a required function) (ISO, 2004). We thus consider that the prognostic should be based on the evaluation criteria, the limits of which depend on the system itself and the execution objectives. From this point of view, in this section, the prognostic is associated with the notions of forecasting and evaluation. This obviously implies that the current situation can be seized (virtually, that is the synthesis of a detection method and measured data of the system). Moreover, it is not so much the concept of failure in the sense of total loss of the ability to accomplish a mission that is relevant, but the performance loss.

Accordingly, the definition that we consider is that proposed by Dragomir [177] **"Prognostic could be split into 2 sub-activities: a first one to predict the evolution of a situation at a**



given time (forecasting process), and a second one to assess this predicted situation with regards to a referential".

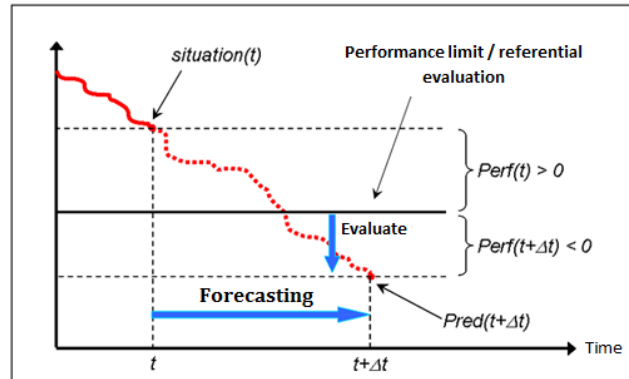


Figure 3.3: The Concept of Prognostic (Dragomir et al., 2007)

Consider Figure 3.3 to illustrate this propositional defining: the situation "predicted" $t + \Delta t$ is considered due to the degradation limit considered the objective viewpoint of performance. Otherwise, if this threshold does not exist, it would be impossible to conclude on the predicted position and therefore impossible to assess the severity of this situation. The maintenance process could then be affected.

Thus, the prognostic could be divided into two sub-activities (Figure 3.4):

- First to predict the evolution of a situation at some point,
- A second to assess the situation in relation to decision-making framework.

The prediction step should determine the future state of the closest way possible system of reality. At the evaluation level, the expected values must be estimated quantitatively and qualitatively: references, RUL, confidence, and precision.

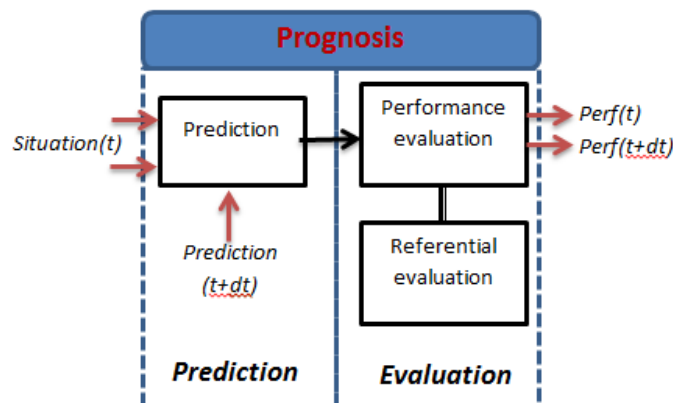


Figure 3.4: Prognostic- Process for Prediction and Assessment (Dragomir Et al., 2007)



In this new sense of prognostic, the complementary aspects of the detection, diagnosis and prognostic can be explained as follows:

- Detection is to identify the mode of operation of the system following the monitoring process, i.e., by identifying its current state,
- Assuming that a fault has occurred, the diagnostic can identify the component that has stopped working (from effects to causes: backpropagation),
- The prognostic deals with the prediction of the future states of the system (from causes to effects: propagation) in two stages: first the situation is foreseen in time and secondly the situation is evaluated by the use of evaluation criteria.

The prognostic is essential as corresponding to a process of anticipation of system failure. Integrated into maintenance strategies it should allow to optimize strategies. The implementation, the cost of the applicability and effectiveness are variable with a dynamic very difficult to "evaluate", especially for industrial land facing the daily stress. In this context, the choice of tool to implement the outcome of the activity is a critical step in terms of overall performance objectives. The next sections will discuss this process.

3.3.2. Prognostic Indicators

We already presented the prognostic can be divided into two processes: the prediction and evaluation [177]. The main prediction goal is to provide useful information to act accordingly, that is to say, to select maintenance actions. The first set of interesting metric is related to the measures and risks involved in the monitored system. This kind of action is called *prognostic measures* [178].

In the same vein, considering that the prediction is, in essence, an uncertain process, it is useful to be able to judge its "quality" in order to imagine the most appropriate actions. Thus, various indicators can be used, which will be called this time, performance measures for the prognostic process [178].

3.4.2.1. The Prognostic Measures

The result of the prediction is the future estimated value of the process (\hat{y}). Analyzing this value, the uncertainties inherent in the prediction process must be taken into consideration (very little information on the phenomenon studied, difficulties in formalization). These are very important sources of errors (e) that make the difference between the actual measured future value of the system (y) and that predicted earlier ($e = y - \hat{y}$).

In particular, the error can be considered:

- Compared to the moment of detection of a T fault.
- Compared to the moment of appearance of the failure (100% degradation).



The relevant evaluation of the chosen prognostic method depends strongly on the metric used for the measurement of the prediction errors. The choice of error measures for the comparison of prediction methods has been much discussed starting with the 1980s. (De Gooijer et al., 2006) [179] identifies different indicators of prediction error measures.

Generally, the error is measured from two following methods:

- MAPE (Mean Absolute Percentage Error) method
- RMSE (Root Mean Square Error) method.

As mentioned above, the main predictive metric sought is the remaining time before failure (Time To Failure: TTF or remaining useful life RUL). In addition, a confidence measure can be constructed to indicate the degree of certainty of the predicted failure time. By extension, and considering that users may be interested in evaluating the system against performance limits, the RUL and trust can be generalized: in Figure 3.5, TT_{xx} is the time remaining to exceed the performance limit $Perf/xx$, and $Conf / xxT$ is the confidence with which the indication $TT_{xx} / T > T$ can be taken.

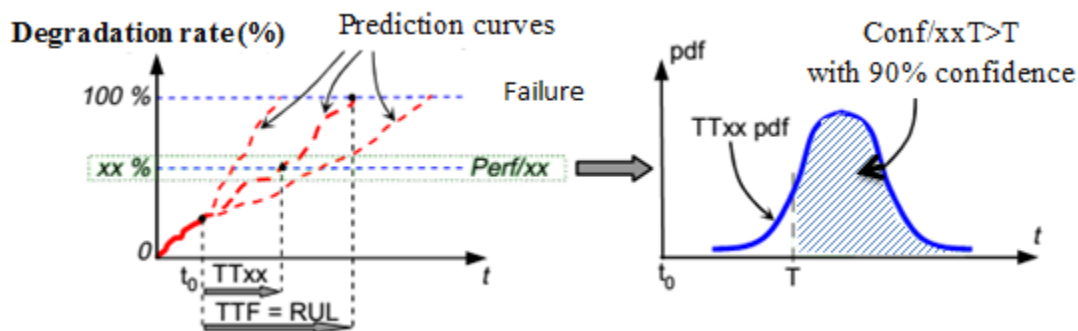


Figure: 3.5 prognostic measures: RUL and confidence

3.4.2.2. Performance Measures for the Prognostic Process [180]

- *Accuracy*

Accuracy measures the proximity of the expected value to the actual value [28]. It has an exponential shape and its value is even greater than the FTT of the prediction error is small. The calculation of this metric represents a critical point in the prognostic process. The question that arises first is whether the prediction is "pretty good". The answer depends heavily on the rigidity of the evaluation criteria imposed. It should be noted that the calculation of this quantity is based on the existence of historical data on several components that have failed as a result of solicitations sustained throughout a known period of time, which is not always possible (unique material).

Assuming this, for experiment i, the real-time t_0 and the expected time of failure t_p the accuracy is defined as the equation 3.1:



Normandie Université

$$Accuracy(t_p) = \frac{1}{N} \sum_{i=1}^N e^{\frac{D_i}{D_0}} \quad (3.1)$$

- $D_i = |t_p(i) - t_0(i)|$ is the distance between the actual and expected moments of failure,
- D_0 is a normalization factor, a constant whose value is based on the importance of real value in the application,
- N is the number of experiments.

- **precision**

Precision is a measure of prediction dispersion that evaluates how predicted values are grouped around the interval in which the failure occurs [181]. Accuracy strongly depends on the level of confidence and the distribution of predictions. The precision formula at a specific future time t_{pf} with respect to the current moment t_{af} is as follows (eq.3.2):

$$Precision(t_p) = \left(\frac{1}{N} \sum_{i=1}^N e^{\frac{-R_i}{R_0}} \right) e^{\frac{\sigma^2}{\sigma_0^2}} \quad (3.2)$$

Where:

$$E_i = t_{pf}(i) - t_{af}(i)$$

$$\bar{E} = \frac{1}{N} \sum_{i=1}^N E_i$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (E_i - \bar{E})^2$$

σ_0^2 and R_0 are normalization factors, R_i is the prediction confidence interval for the experiment i [181].

Similarly, an exponential function is used here to define the relationship between the standard deviation of the prediction, the confidence interval, and accuracy. The precision has a value between 1 and 0 (1 indicating highest precision and 0 the lowest).

The complementarity of the accuracy and precision is illustrated in Figure 3.6.

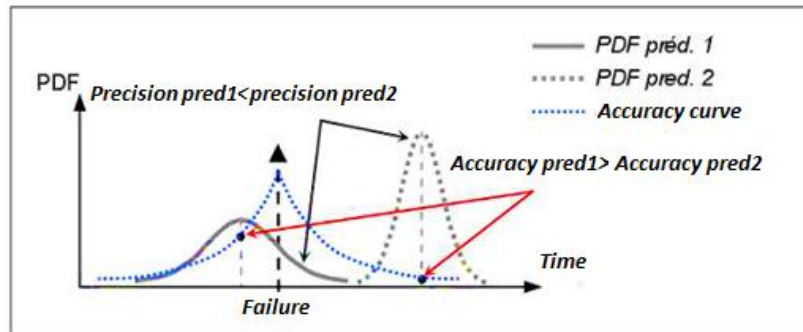


Figure 3.6: Accuracy and Precision

- **Timeliness**

The performance measures for the prognostic should take into account two outcomes: the expected time to failure - RUL (the expected value based on historical data) and the confidence interval. The "timeliness" is the relative position of the PDF of the prognostic model with respect to the occurrence of the failure event. This measure evolves as the data become available and makes it possible to judge the appropriate time to take preventive actions (see Figure 3.7).

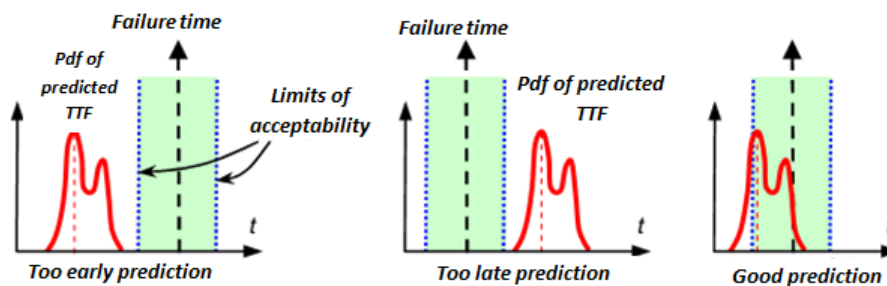


Figure 3.7: Timeliness from (Goebel et al., 2005)

In 2005 Goebel et al. defined the limits at the earliest and at the latest beyond which the predicted value will be considered unacceptable from a performance point of view. These two limits are the consequence of the fact that the prediction error is not systematically centered with respect to zero (where the error is defined as the difference between the actual remaining life and the estimated remaining life) [182]. For example, if the prediction is too early, the resulting alarm requesting too early intervention to check the potential occurrence of a failure to monitor the various process variables and to perform corrective action. In the other case, if the failure is expected too late, this error reduces the time available to assess the situation and react accordingly. The situation is completely degraded when the failure occurs before a prediction is made. Therefore, it is in most situations preferable to have a positive bias of errors (early forecast), rather than negative (late forecasts). Of course, acceptability limits must be set on how a forecast can be considered as too early, too late or acceptable. Any prediction outside of limits will be considered inappropriate. Other prognostic performance measures were proposed and



detailed in 2006 by Vachtsevanos et al. [183] such as the similarity or sensitivity. In 2005 Létourneau et al. [184] used a reward function to penalize the positive or negative prediction errors.

3.4.2.3 Prognostics Modeling Methods for RUL Estimation

The construction of an effective model that able to predict the evolution of the degrading features and able to calculate the RUL of the system represents the primary objective of a prognostics algorithm. In point of fact, the RUL estimation is a very difficult task. Therefore, the construction of an accurate prognostic model especially that is based on data-driven approach has been the focus of researchers in recent years. Depending on these researches we can classify data-driven RUL estimation into two categories:

- 1- Univariate degradation based modeling method
- 2- Multivariate degradation based modeling method.

1- Univariate Degradation-based Modeling Method

Usually, when the prognostic modeling depends on a prediction of continuous degrading curve accompanied with previously determination of the criteria of failure this is known as univariate degradation calculation, in this case, RUL is estimated at the intersection point between this degradation curve and a failure threshold (FT) that previously determined (Figure 3.8). In this case, a time series observations calculated during regression is used training inputs of data-driven approach [185][81]. The calculation of the FT is considered the main difficulty in this method, that's because it can be restricting the efficacy of univariate degradation based approach. In general, all studies based on univariate degradation based approach resort to the estimation of FT [186][187] which are often uncertain and can lead to inaccurate prognostics.

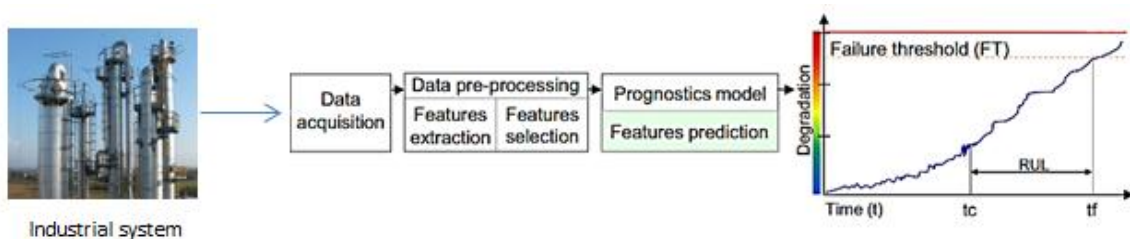


Figure 3.8: Univariate Degradation Based Modeling Method

2- Multivariate Degradation- Based Modeling Method

To avoid the limitation of the prognostic based on the univariate degradation based approach, a health assessment model and a prediction model are used in same time to estimate the RUL [177] this is known multivariate degradation based modeling [188]. The main objective of the multivariate degradation based modeling is to create integration between a prediction model and a classification model for accurate calculation of FT. Therefore, to determine the system degradation level the RUL is calculated basically on the estimation of discrete states calculated



simultaneously and continuous predictions of a set of features v and the FT, in this case, calculated dynamically rather than in univariate degradation approach [189] (see Figure 3.9).

The Multivariate degradation is the newest method in the domain of data-driven prognostic and it's proved their worth on the detection of system degradation and RUL calculation. However, the number of degradation states is considered as the main limitation of this approach that's because it can be different from a system to another. At the end, many applications affirm that the multivariate degradation has achieved very good and effective results compared to those achieved by the univariate degradation based approach.

Consequently, in this work, an enhanced multivariate degradation based modeling is used as a step from the RUL calculation methodology.

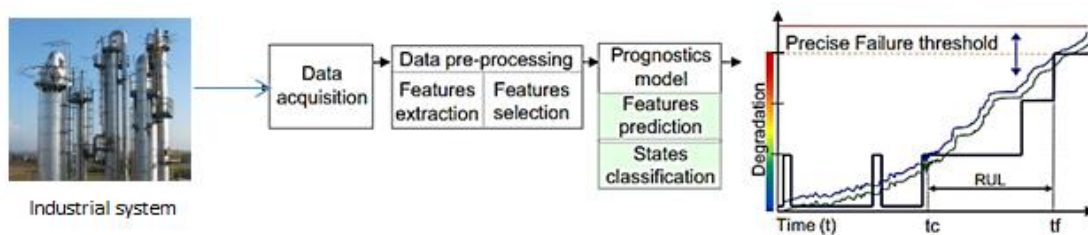


Figure 3.9: Multivariate Degradation Based Modeling Method

3.5. Prognostic Approaches

The field concerning the prognostic of approaches is very broad; the purpose of this section is not to make an exhaustive synthesis of the existing, but to show the wealth of opportunities. Many tools and methods of failure prognostic have been proposed during the last decade. Prognostic methods generally differ by the type of application considered, while the tools used depend mainly on the nature of the data [190] [191] and knowledge available to build a real system behavior model including the phenomenon of degradation. Also, these methods and tools can be grouped into a limited number of approaches.

3.5.1. Classification of Prognostic Approaches

3.5.1.1. Emerging Classifications in the Literature

The first classification of prognostic approaches has been proposed by (Lebold et al., 2001). In their paper, the authors suggest a pyramidal three-level classification of prognostic approaches [81] (Figure 3.10).



Normandie Université

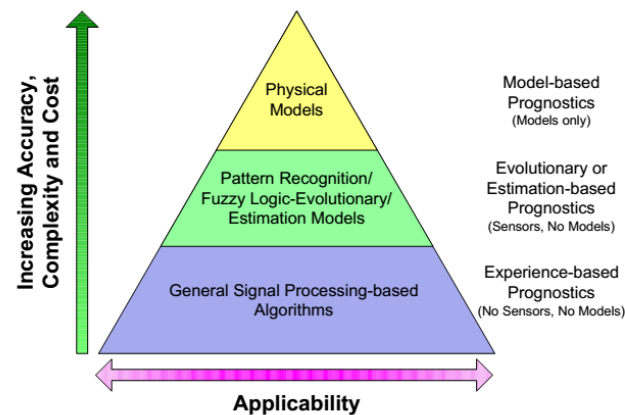


Figure 3.10: Taxonomy of Prognostic Approaches [192]

Similarly to the acceptance of the "scientific community", prognostic methods can be associated with one or more of the following three approaches:

- ❖ ***Prognostics based on the physical model*** - use the causal relations derived from the laws of physics for the mathematical representation of the mechanism of degradation,
- ❖ ***Prognostic data-driven*** - is based on the following assumptions: (1) the statistical characteristics of the data are relatively unchanged unless a malfunction occurs and (2) it implies being able to learn (by the examples) and to capture the subtle relationships between data,
- ❖ ***Prognostic based on experience*** - is based on the formalization of the physical mechanisms of deterioration of components with stochastic models initiated by a priori knowledge and expert judgment.

The complementary approaches (Table 3.1) that emerge from the scientific literature are those of Byington et al. in 2004 [193] and Muller et al. in 2008 [194]. To establish a point of view referring to these works, in the following, we retain the characteristics distinguishing them from "classical" approaches.

Approach	Techniques
<ul style="list-style-type: none"> • Prediction based on the model 	<ul style="list-style-type: none"> • Parity Space • Observers • Parametric Estimation
<ul style="list-style-type: none"> • Prognostic data-driven: • Prognostic for trend analysis • Machine learning Prognostic • Prediction based on state estimators 	<ul style="list-style-type: none"> • Techniques Of AI ("Black Box") • Statistical Techniques
<ul style="list-style-type: none"> • Prediction based on experience 	<ul style="list-style-type: none"> • ALM (Accelerated Life Model) • PHM (Proportional Hazard Model) • Monte Carlo

Table 3.1: Emerging Approaches and Prognostic Techniques in the Literature

In 2004, Byington et al. lists the three categories and prognostic methodologies previously mentioned and adds a more in-depth classification of the methods according to the type of model they use in the data-driven prognostic [193]

- ***Prognostic for trend analysis*** (Evolutionary / Feature-based predictions) based on the exploitation of statistical models.
- ***Machine learning Prognostic*** (Machine learning / artificial intelligence (AI) - based prognostic) using black box techniques derived from artificial intelligence.
- ***Prognostic based on state estimators*** (State Estimator Prognostic). These methods are used when a diagnosis by pattern recognition is implemented beforehand. The mechanism is to predict the evolution of the trajectory of the form via a Kalman filter.

(Muller, 2005) offers a prognostic methodology based on:

- coupling probabilistic / event approaches incorporating probabilistic information and events,
- modeling of degradation, the impact of the degradation on the performance and impact of the maintenance action on the performance,
- evaluation of the expected performance of a system based on a predefined maintenance strategy [195].

In 2006, Jardin et al. [165] proposed a new taxonomy of prognostic methods, distinguishing two main categories of methods. The first group includes methods for estimating the future state of the component, subsystem, or system (estimation of the RUL or the TTF), and the second category concerns methods for determining the RUL while integrating the context of system operation (maintenance actions and operating conditions).



We note that the academic vision of the prognostic is restricted to its application on critical components and the approaches they are focused on the analysis of an "elementary component". The modeling of complex systems is almost non-existent.

3.5.1.2. Discussion

Without pretending that our existing synthesis is complete, we can summarize the strengths and weaknesses identified as follows. For prognostic trend analysis (Evolutionary / Feature-based prognostic) of (Byington et al., 2004) [193], subcategory prognostic data-driven and based on the use of statistical models, it appears that the disadvantaged are related:

- ✓ *Lack of responsiveness to the occurrence of the fault,*
- ✓ *The possible errors in the forecasts in case of noise or an insufficient number of measures.*

Thus, to be efficient, the methods belonging to this category of prognostic must imperatively correct the parameters of the models in case of modification of the operating conditions of the component.

In the same context of the data-driven prognostic, we have identified in the the proposal the prognostic by machine learning (Machine learning / AI-based prognostic). They use black box type models derived from artificial intelligence. Their effectiveness is conditioned by the existence of the data sample of degradation scenarios. The absence of a scenario in the history makes the model ineffective in case of occurrence of this scenario. The problem is to overcome the inability to consider a new mechanism. In this case, the prognostic is random [193]. Similar remarks can be made about the prognostic based on a State Estimator Prognostic. These methods are based on state estimators and are used when pattern recognition is implemented beforehand. The mechanism consists of predicting the evolution of the trajectory of the operation of a system by means of a Kalman filter. Without the existence of the diagnosis upstream they are impractical. From an application point of view, the information required to deploy the prognostic approaches is of various kinds: engineering models, data, failure history, system demands, operating conditions, etc. In 2001 Lebold et al. [81] a generalization of what could be the set of inputs and outputs of a model of prognostics.

Approach	Model-based	Data-driven	Experience-based
Information			
System model	Necessary	Useful	Not necessary
Failure History	Useful	Not necessary	Necessary
Past conditions	Necessary	Not necessary	Useful
Current conditions	Necessary	Necessary	Useful
Failure Recognition Methods	Necessary	Necessary	Not necessary
Maintenance history	Useful	Not necessary	Useful
General	Sensors and model	Sensors, no model	No sensor, no model

Table 3.2: The Information Required Deploying Prognostic Approaches

Our analysis of prognostic approaches identified in the literature also unveiled some terms in the proposal of (Muller, 2008). His vision has brought novelties through the modeling of degradation, the impact of degradation on performance and the impact of maintenance action on performance. He also integrated probabilistic and event information into the prognostic process and evaluated the expected performance of a system in the future based on a predefined maintenance strategy. However, multiple difficulties related to multiple degradation modes, the quantization of a priori conditional probability tables and the modeling of maintenance efficiency persist.

3.5.2. Prognostic Based on the Physical Model

The implementation of this approach is generally based on an available mathematical representation of the degradation mechanism. The causal relationships derived from the laws of physics are used to model the interactions between the entities of the system. Thus, the residues are used as mathematical instruments. Large residues denote the presence of malfunctions, and small residues the presence of normal disturbances such as noise or modeling errors. The use of model-based prognostic methods involves specific knowledge related to the failure as well as a strong control of the mode of operation of the analyzed system.

3.5.2.1. The principle of Operation, Advantages, and Disadvantages

The premise principle based on physical models is to determine the current level of degradation of the system (through monitoring and diagnostic processes) and to evaluate the RUL using the evolution curve of the system degradation depending on the solicitation of the system.

The main advantages of this approach are:

- ✓ Flexibility: If there is any changing of the properties of the system or in their degradation form, the model can be adjusted to reflect this change.
- ✓ The knowledge of the structure of the model makes it possible to find a link between the variation of indicators and a modification of a parameter, which is not available in the case of "data-driven" approaches [196].

The main disadvantages of prognostic based on physical models are:

- ✓ Model development is extremely expensive. It requires a high level of qualification and experience of the developers,
- ✓ It is difficult to build a global model for complex systems: complexity to modeling all the interactions between the different mechanisms, computational difficulties associated with solving a system of differential equations
- ✓ It may be impossible to generalize an approach based on an oriented physical.
- ✓ It is necessary to know the mechanisms of degradation and influencing factors on them.



3.5.2.2. Techniques and Tools

The techniques belong to this prognostic approach are based on the following tools: parity space, observers (Kalman filters) and parametric estimation.

❖ *The parity space*

It is an analytical redundancy relationship represented by an equation in which all the variables are known. The generation of such relationships can generate residues. A residue is a timing signal based on inputs and outputs of the process, independent (if possible) of the operating point of the presence of faults. In the absence of this, the residue is statistically zero. When a fault occurs, its amplitude changes significantly [197].

❖ *Observers*

The generation of residuals using a state estimate consists of reconstructing the state or, more generally, outputting the process using observers and using the estimation error as a residual. This method has grown significantly since it leads to the design of flexible residual generators [197].

❖ *Parametric Estimation*

The parametric estimation approach considers that the influence of faults is reflected in the parameters and not only in the variables of the physical system, as in the case of observers. The principle of this method consists in continuously estimating the parameters of the process using the input/output measurements and then evaluating the distance between them and the reference values of the normal state of the process. Parametric estimation has the advantage of providing information on the importance of deviations. However, one of the major disadvantages of the method is represented by the need to have a constantly excited physical system. This, therefore, poses practical problems in the case of hazardous, expensive or stationary mode. In addition, the relationships between mathematical and physical parameters are not unitarily invertible, which complicates the task of residue-based diagnostics [170].

3.5.2.3. Applications

Before going further into the problem of prognostic, we will draw the brief state of the art applications that used the prognostic based on models. In 2004, Rafiq et al. [198] investigated the deterioration of bridges due to a process of chlorine induction. A Bayesian approach has been implemented to update the stochastic model. In 2003, Luo et al. [170] introduced the concept of an integrated prognostic process based on data generated following simulations on a model of the system, in nominal mode and in degraded mode. Thus, he has developed a generic prognostic methodology based on the model that respects the operating principle described in Figure 3.11.

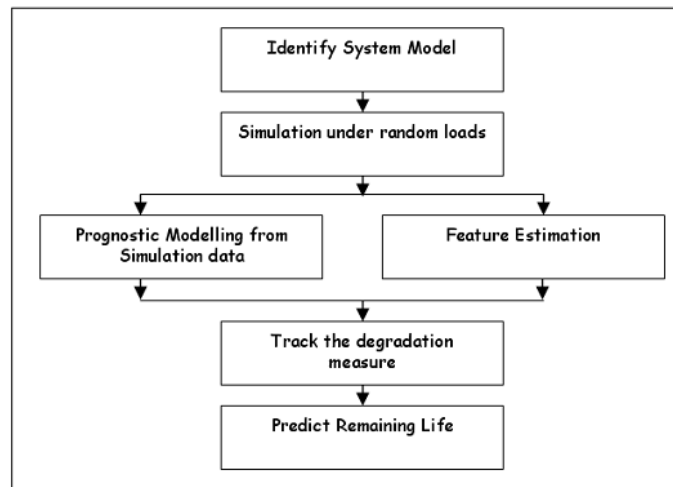


Figure 3.11: Prognostic Based on the Model (Luo et al., 2003)

3.5.3. Data-Driven Prognostics

The prognostic guided by the data is based on the following observation: measurements (input/output) are often the strongest and safest source of information to understand degradation phenomena ... Its strength lies in the ability to learn (through examples) and to capture the subtle relationships between data, even if these relationships are unknown or difficult to describe [199].

3.5.3.1. The principle of Operation, Advantages, and Disadvantages

The data-driven prognostic exploits the indicators of degradation or maintenance interventions delivered by the monitoring and decision support processes respectively (ex. calorimetric calibration data, spectrometric data, power, vibration and signal, temperature pressure, oil debris, acoustic voltages of currents). The upstream diagnosis determines the success of the prognostic by its ability to provide a reliable and accurate estimate of the current state of health of the system and an update of the parameters of the deterioration processes. This type of prognostic is based on the assumption that the statistical characteristics of the data are relatively unchanged unless a malfunction occurs in the system. The ability to adapt to any type of application with sufficient data in quantity and quality is a strong point for this prognostic approach. At the same time, the implementation of a "data-driven" approach is relatively straightforward as it does not require formal knowledge of the mechanisms of degradation. The ability to transform noisy data into information relevant for diagnostic/prognostic decisions is another advantage that can be highlighted.

The main disadvantage of these approaches is that their effectiveness is greatly dependent on the quantity and quality of operating system data [199].



3.5.3.2. Techniques and Tools

In the literature are distinguished from the techniques of artificial intelligence (AI) and statistical techniques.

➤ **black box Techniques**

When the only information available on the system is the measurable variables and physical redundancy cannot be used, the usual technique is to learn the behavior of the system using the data history: these are the learning data. It is assumed that the same cause will always have the same effects. These are systems of the "black box" type that has the main advantage of using "blind" data, without any physical consideration. Their strength lies in the ability to learn and to capture the subtle relationships between data, even if these relationships are unknown or difficult to describe [200].

1- Techniques of Artificial Intelligence (AI)

Neural networks (NNs) and wavelet networks (multilayer perceptron, probabilistic neural networks, wavelet networks with self-organization, etc.) are the main classes of tools of this type. The main disadvantage of neural networks is the acquisition and coverage of training data.

2- Statistical Techniques

The multivariate statistical techniques are powerful tools that can compress data and reduce their dimensionality so that essential information is maintained. They can also manipulate noise and correlation to extract information efficiently. The main function of this type of technique is, using a mathematical procedure, to transform a number of correlated variables into a smaller set of uncorrelated variables.

- ***Principal Component Analyses (PCA)***

The Principal Component Analyses (PCA) is a multi-variable statistical technique. The PCA is essentially based on an orthogonal decomposition of the covariance matrix of process variables along the directions that explain the maximum variation of the data. The first axis contains the largest variation. The second axis will contain the second largest orthogonal variation at the first. The main purpose of the PCA is to find a set of factors (components) of smaller size than the original set of data that can describe correctly the main trends. PCA is a procedure that only considers process variables. Sometimes an additional set of data is available, ex. product quality variables. It is desirable to include all available data for process monitoring and to use process variables in this way to predict and detect changes in product quality variables. For this, the Partial Least Squares (PLS) method can be used. This method models the relationship between two blocks of data while compressing them simultaneously. It is used to extract the latent variables that explain the variation in process data. An important limitation of PCA-based monitoring is that the representation obtained is invariant over time, while most real processes



evolve over time. Therefore, the representation obtained from the PCA also needs to be updated periodically. Another disadvantage is that it does not have signature properties for diagnostics, which makes fault isolation difficult [201].

- ***Trends Analysis and Qualitative Representation.***

A general signal processing aims analysis and qualitative representation of the process trends (series of episodes with representation). In this formalism, each episode is represented by its initial slope, the final slope (at every critical point) and a line segment connecting the two critical points, or a qualitative description of signals (TDL - Trend Description Language) using primitives, episodes, trends, and profiles. These trends can be used for the identification of abnormal situations in the process. Thus, a proper analysis of the trends of the process can help in the early detection of failure [202].

3.5.4. Prognostics Based on Experience

The prognostic based on experience is based on the formalization of the physical mechanisms of deterioration of components with stochastic models (reliability of law, Markov processes or non-Markov) initiated by knowledge a priori and expert judgment.

3.5.4.1. The principle of Operation, Advantages, and Disadvantages

The main ***advantage*** of this method is that it does not require advanced knowledge of the physical mechanism (s) of degradation. Also, it is relatively simple to implement and inexpensive.

The main ***disadvantages*** are the following:

- The lack of responsiveness to the change in the behavior of a system or the environment,
- The applications focus on critical components, treated individually and therefore the development of "system" oriented approaches is rare,
- There is often a discrepancy between the models (two-state mono-component system) developed and the industrial reality (multi-component multi-state system). The origin of this shift often comes from the inability of the methods to perform the calculations generated by a complex system [203].

1.5.4.2. Techniques and Tools

The use of an evolutionary reliability model of Accelerated Life Model (ALM) type, Proportional Hazard Model (PHM), or the implementation of a Bayesian approach that updates the parameters of the degradation law with all new information available can represent a solution. The Monte Carlo simulation is another preferred alternative but it is itself confronted with another problem: the explosion of simulation times.



In conclusion, each of these approaches has its own advantages and disadvantages, and therefore they are often used with the association for specific applications [204].

3.6. Towards Predictive Systems

3.6.1. Typology of Prediction Systems

The prediction is the process of estimating unknown situations in the future. The fields of application are therefore broad: floods occurring over a given period, planning of demand in manufacturing companies, weather forecast, prediction in the financial world, etc. The prediction plays an essential role in making a decision; it concerns the security or the company's capital. Traditionally, the prediction is the estimation of a value in the future by analyzing data from the past, or more informally, by expertise. Two families of methods [205] exist to perform the prediction task.

- **Qualitative:** This type of technique uses experience and judgment to establish future behaviors.
- **Quantitative:** This type of technique uses historical data to build relationships and trends that can be projected in the future.

To select the correct method, it is mandatory to consider the context in which the prediction tool is applied. The following sections deal with the criteria for choosing a predictive system, namely the industrial constraints and the performance of a predictive model.

3.6.2. Criteria for Choosing a Predictive System

The key points for choosing a tool to highlight the advantages and disadvantages of the different monitoring methods encountered are those shown in Figure 3.12.

Each of the potential techniques for prognostic application has its own advantages and disadvantages. As a result, they are often used in combination in many applications. However, it should be noted, that many approaches only focus on the analysis "of an elementary component", and the modeling of complex systems on a global level is almost non-existent [206].

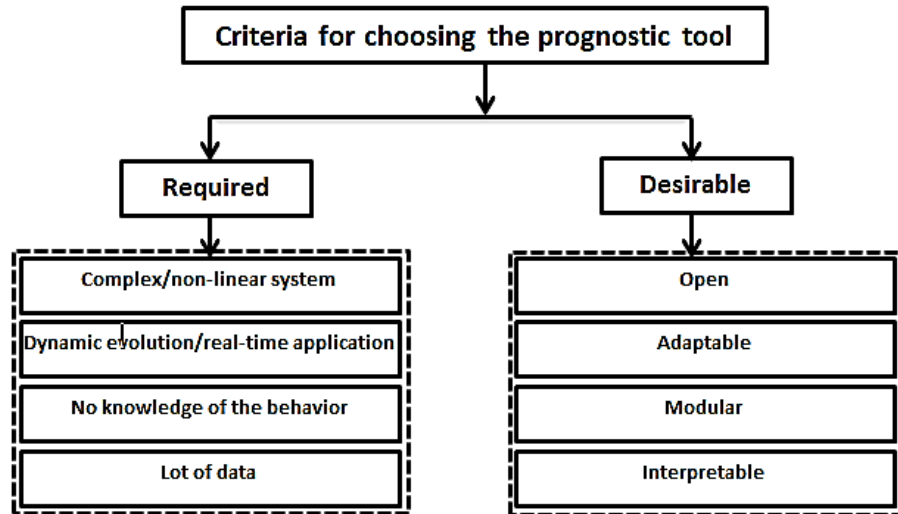


Figure 3.12: The Criteria for Choosing a Prognostic Tool

1. *Industrial Constraints*

The evolution of technology and its incorporation into industrial equipment has made them more complex. The goal is to make equipment efficient and reliable. These devices have become more and more complex, and they evolve in a dynamic and non-linear environmental context. Moreover, the operators have difficulty understanding their operation and controlling them. Indeed this complex and evolving context complicates the task of prediction. However, in the industrial world, the data acquisition system is generally operational and makes it possible to measure different parameters with precision, as well as to collect and distribute a large volume of information in real time. In the end, the industrial world has the following specificities:

- ✓ ***Complex equipment,***
- ✓ ***Dynamic environment / non-linear,***
- ✓ ***No knowledge of the behavior,***
- ✓ ***Many data.***

2. *Desirable features*

An important point in choosing a prediction tool is the possibility of integrating it into our prognostic approach, which consists first of all in detecting an anomaly monitored by a physical parameter or a degradation index and then in predicting its future evolution.

- ✓ ***Real-time application:*** the predictive system must be used and evolve with the real equipment,
- ✓ ***Flexibility:*** the prediction tool must be able to adapt to several applications and not be limited to particular environmental conditions. In addition, this prediction system must



react according to the different behaviors of the system (non-linear, dynamic, non-stationary, etc.).

- ✓ **Interpretability:** the prediction tool must maintain a minimum level of transparency so that an expert can intervene to modify the parameters,
- ✓ **Open:** An open system is a scalable system to which parameters can be added during execution.

3.6.3. Restrictive Industrial Characteristics of Prognostic Tools

From the industrial point of view, two types of conditions can be considered:

- Some of them are strictly "necessary" because a prognostic technique must adapt to real systems,
- Others may be considered desirable as they have "translated" the identified characteristics.

In this context, a prognostic tool must be able to capture the dynamic behavior of the system by providing, if possible, some related indicators. Also, it is particularly interesting to implement a technique that evolves in real time with the system, reducing quantitatively improving its sensitivity and accuracy. These are very hard objectives to achieve because the real processes are complex which prevents their characterization with simple models. Moreover, real systems cannot be considered as static systems. Thus, the tools of prognostic should evolve to be appropriate. At another level, the features that can be expected are: adaptability/flexibility (transferring technology from an application to another), modularity/integrability (globally integrated local prediction), and accessibility (promote knowledge of the system).

In practice, the need for robustness defined here as the ability of the system to detect faults independently of modeling errors, sensitivity (the ability of the system to detect faults of a certain magnitude), detectability (system suitability prognostic to be able to detect the presence of a failure on the process), isolability (the ability of the prognostic tool to go directly back to the origin of the fault) and reliability for the processes with a dynamic evolution (in real time) is imperative [206].

3.7. Selection Guide for a Family and a Prognostic Tool

The literature review revealed the existence of a classification of prognostic approaches in the sense of the "scientific community" as follows: model-based prognostic, data-driven prognostic, prognostic based on the experience. A brief description of the three approaches has been given. In addition, the advantages and disadvantages of each of the tool classes have been identified. In the scientific literature, two other different "prognostic" approaches, from the conceptual point of view of "generally" accepted, have been identified. This is the proposal of (Byington et al., 2004) [193] and that of (Muller et al., 2007) [163].



One of our most important goals is to adopt the different existing classifications in order to propose a framework allowing the identification of a type of tool adapted to the specifications of the industrial context (Figure 3.13) [206].

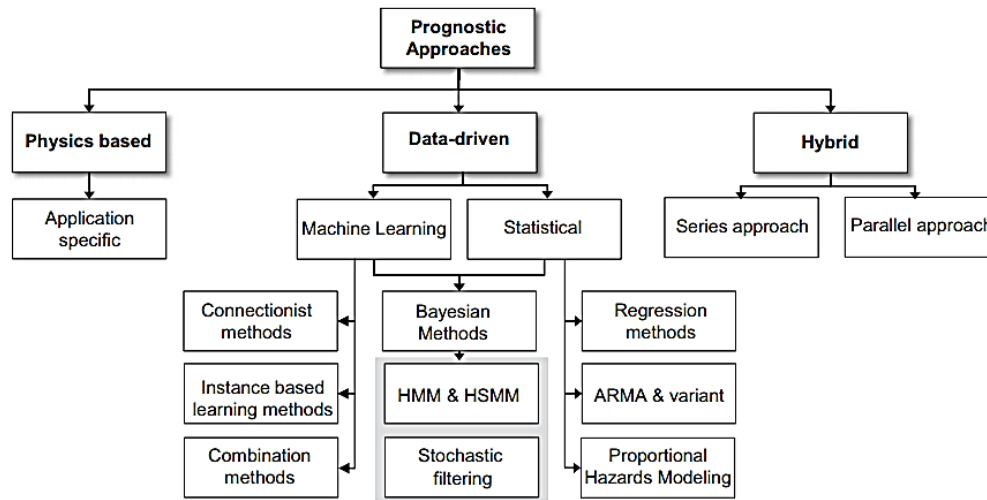


Figure 3.13: Other Classification of Prognostic Techniques [191]

3.7.1. Data-driven Tools

In the section "Data-Driven Approaches", we identify in the literature two distinct categories of support tools: techniques based on artificial intelligence and statistical techniques. We consider that a distinction is needed at the level of techniques based on artificial intelligence. Thus, only artificial neural networks (ANNs) or "black box" tools are considered to be data-based tools. The others, the fuzzy rules-based systems, the decision trees, the graphical models, require additional meanings and information related to the context of the prognostic for execution and thus surpasses the notion of data. ANNs are tools adaptable to nonlinearities of multi-variable systems, with low sensitivity to measurement noises and changes in effective modes (systems with multiple configurations). The simple assessment function gives them online and real-time learning capability. Even though the inputs (historical data) are decoupled from the system structure, the prerequisites for modeling are minimal because the ANNs can capture the hidden input/output relationships. The difficulty of detecting multiple faults due to the inability to generalize and think in unknown space through learning, unable to explain how the decision was made and the lack of guarantee on the convergence of the learning data are the main disadvantages of these tools. In this category of data-oriented methods, we also consider statistical techniques. In the modeling process, the data is used only as quantitative measurements. These techniques are based on the assumption of rapid changes in the characteristics or parameters of the models compared to dynamics considered as slow (quasi-stationary processes). They are used for detecting gradual changes with low detection thresholds. The information provided by the increasing number of sensors installed on the processes makes it



very difficult to analyze the results. On the other hand, since the calculation time for the filtering is too long, the representation used is redundant and only the additive faults are detected.

3.7.2. Information-Oriented Tools

Similarly, according to literature, the methods based on information in this proposal are based on two categories of techniques: *quantitative and qualitative*. We consider the category of methods based on qualitative information as potential candidates reasoning tools from cases and pattern recognition.

Their strengths are:

- ▶ Knowledge acquisition efforts are reduced,
- ▶ They are relatively easy to maintain,
- ▶ They allow the use of existing data in databases,
- ▶ They can adapt to environmental changes.

On the other hand :

- ▶ The precise definition of the classes between which will take place the decision,
- ▶ There is no systematic method for choosing the parameters,
- ▶ The number of attributes sets the dimension of the representation space and, therefore,
- ▶ The amount of calculations to be carried out is a strong constraint for real-time processing,
- ▶ The existing information on the various modes of operation of a system is always incomplete,
- ▶ The difficulty lies precisely in this case structure and the information it must contain. Indeed, the extraction of knowledge and their representations are essential in this type of application.

Quantitative techniques (parity space, parametric estimation, observers) are in fact statistical techniques, usually used in methods based on "classical" models (use thresholds to detect the presence of faults) technical lies in their ability to take into account in the design of the system architecture, multiple faults and to provide information on the importance of deviations. However, these techniques require having a permanently excited physical system, which limits the industrial applicability. The processes are usually described in the stationary mode and are very expensive, which is impossible for dangerous systems. Usually, their offline operation prevents real-time processing. Disturbance modeling that can lead to errors in the model because the adaptability of these approaches to process changes does not exist.

3.7.3. Knowledge-Oriented Tools

Knowledge provides users with additional information related to the remaining life until the failure of the process being studied. These methods include fuzzy rules-based systems (FS), decision trees (DT) and expert systems (ES). Modularity is a general feature of these methods and



partially explains that they are often used in combination with other techniques. One particular technique that can be cited here is that of the *neuro-fuzzy systems*. Approaches based on knowledge have the ability to model the systems identified with minimal distortion of reality. This is because surveillance data is context-related. The language used is very close to the normal language and a physical understanding of the system is easier. Fuzzy systems require knowledge expressed by rules. These are recommended causal methods if we do not have quantitative information but dependence rules describing the propagation of faults. A fuzzy system can be automatically adjusted and mathematical models are not required. The previous information on the rules can be used because the interpretation and implementation are simple. FS is also the only framework in which inaccuracies and uncertainties can be addressed and which also allows the treatment of certain incompleteness.

On the other hand:

- ▶ The rules must be available,
- ▶ They do not have the ability to learn,
- ▶ Adaptation to environmental changes is difficult,
- ▶ No formal method for adjusting the initially created knowledge base is available.

Expert systems are transparent methods that reasoning under uncertainty (fuzzy) that by cons, explicit decisions. The ease of development of this type of tool is a consequence of the fact that they do not require many details related to the system. Knowledge-based on which they are highly specialized, require an abundance of experience and are difficult to actualize

For decision trees (DT) can learn the cases not encountered during use. On the other hand, the problem is complicated when temporal constraints are added, which are not frequently explained. They, therefore, require a very good expertise of the system and these malfunctions.

3.8. Investigation Track to Choice of a Prognostic Tool

Table 3.3 reveals the strengths of current approaches and areas for improvement, in accordance with the criteria necessary and/or desirable for the industry.

The objective, let us remember once again, is not here to "judge" prognostic techniques, to the decree which ones are "good" and which ones are "bad", but to identify the strong points, weaknesses and if possible, the tracks could improve the contribution of these different tools.

			Support tools for the prognosis						
			Data-driven methods		Information-oriented methods		Knowledge-oriented methods		
Choice criteria			Black box tools	Statistical tools	Qualitative tools	Quantitative tools	Decision tree, Petri networks	Fuzzy systems	Expert Systems
	Required	Complex system / non-linear	+++	+--	+--	+--	+--	+++	+--
		Dynamic evolution / real-time application	+++	+--	+--	+--	+--	+++	---
		No knowledge of the behavior	Without influence				Unusable		
	desirable	Lot of data	+++	++-	+--	+--	++-	Without influence	
		Open	+++	+++	+++	---	+--	++-	+--
		Adaptable	+++	+--	++-	---	+--	+++	---
		Modular	+++	+++	+--	+++	+--	+++	+--
		Interpretable	---	+++	+++	+++	+++	+++	+++

Table 3.3: Prognostic Support Tools (Vasile et al., 2008)

In our case, the decision problem is presented as a matrix where each column corresponds to one of the prognostic techniques available to the decision maker and each line to a possible state / environmental criterion (Table 3.3). At the intersection of the rows and columns, each box expresses the adequacy of the technique to the state of the environment j

We see that in this abstract vision of the decision:

- Possible alternatives sets (actions and decisions) are finished and known,
- The different alternatives are mutually incompatible,
- The set of possible states of the environment is finite and known,
- The consequences of each alternative for each state of the environment are known.

Once these conditions are satisfied, deciding would then be to evaluate the different alternatives, i.e. to value the consequences of each of them in terms of usefulness for the decision-maker and to select the best one that is, the one that best satisfies the satisfaction criterion used. However, to decide is to choose between several mutually incompatible options.

In concrete terms, the decision constitutes a complete process of information search, evaluation, and selection in order to act on a specific system in order to achieve one or more objectives.

This process includes:

- ✓ A phase of identification of the different possible actions on the system (good correlation (+++), correlation with limits (++ -), weak correlation (+ -), without influence / undesirable (---), not the case / unusable),
- ✓ A phase of evaluation of their respective consequences,
- ✓ A selection phase of one (or more in case of a program of actions) of them depending on preference criteria,
- ✓ A phase of implementation of the decision.

The valuation effects of establishing a ranking between them depending on the preferences and objectives of the decision maker. This ranking is, in our case, expressed qualitatively, in terms of "linguistic variables" (for example good correlation, correlation with limits, no influence / undesirable, etc.). It can also be a quantitative evaluation by means of an appropriate quantitative metric or in monetary terms.

3.9. Neuro-Fuzzy Systems for Time Series Prediction

Since the 1980s, adaptive networks have been used for time series prediction [7]. The goal of introducing such systems is to streamline the decision-making process of decision-makers to achieve better results. Historically, neural networks (ANNs) have been used to predict time series. Then, the idea came to combine them with the principle of fuzzy logic. In the next section, we present a history of using ANNs in prediction by showing the advantages of combining them with fuzzy logic-based systems.

3.9.1. Neural Networks and Prediction

Neural networks are a special case of adaptive networks. They have had great success because of their characteristics: they can "model" and "reproduce" nonlinear phenomena without prior knowledge and are able to grasp the hidden relations between inputs and outputs. From an operational point of view, they are fast systems.

The idea of using ANNs for prediction dates back to 1964: Hu used the Widrow Adaptive Network (WAN) to make climate predictions. The lack of learning algorithms limited the continuation of this type of study. Since the 1980s, research in the field has been revived. We will retrace the evolution below (Figure 3.14) [207].

In 1987, Lapedes and Farber [208] performed the first work showing the possibility of identifying and predicting deterministic chaotic time series using multilayer perceptron. This article has launched several applications on real data. H.White in 1988 [209] studied the case of the return of stock forecast for IBM. This effort was followed by Sharda and Patil (1990) [210], which led to prediction competitions between neural networks and traditional techniques: out of 75 series tested, the neural networks were more efficient for 39 series.

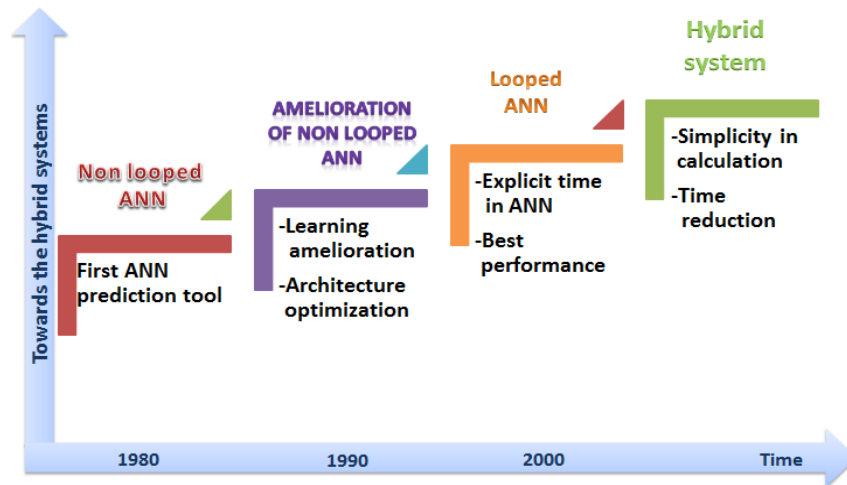


Figure 3.14: Towards Hybrid Systems

We propose below a summary on the use of ANNs for the prediction.

- **Phase 1:** Non-looped networks, one of the first ANN applications for prediction dates back to 1987. Lapedes and Farber in 1988 [211] built an ANN to approximate a chaotic signal [212]. As a result, the non-looped ANNs associated with the backpropagation algorithm (introduced at this time) performed better than the classical self-regression models for the prediction of nonlinear time series [192].
- **Phase 2:** Improved learning and parameterization of ANNs, many factors affect the performance of ANNs (number of inputs and outputs, number of layers, activation functions, choice of test base, learning algorithms) and defining an appropriate ANN to a given problem is not an easy task. Also, since the 90's, many developments are carried out to improve the accuracy of the predictions made by the ANNs while decreasing the complexity of the models and the calculation time. This work aims at proposing a "guide" for the optimization of architectures of ANNs and learning algorithms [213].
- **Phase 3:** Recurrent networks, in order to explicitly take time into account, recurrent network architectures have been developed and compared to other nonlinear time series prediction techniques. The results show that these ANNs perform better than conventional methods and even more so than unscrambled networks.
- **Phase 4:** Towards hybrid systems, it appears that ANNs have been used successfully to support the prediction activity. However, some authors remain skeptical: first, the optimization of an ANNs is more an art than a science, then the ANNs are black boxes and it is not possible to explain and analyze the relations between inputs and outputs. Thus, work emphasizes the interest of hybrid systems to overcome this weakness of ANNs (preserving their learning capacity or even reducing the complexity of models). For this purpose, research is moving towards the combination of ANNs with other AI principles including fuzzy logic [192].



3.9.2. Towards a Hybrid Neuro-fuzzy Prognostic System

The concept of fuzzy modeling has its origins in the fuzzy set theory proposed in 1965 by Zadeh [122] as a way of dealing with uncertainty, based on the idea of defining sets that can contain elements in a gradual manner. This theory introduces a way of formalizing human reasoning methods by using rule bases and linguistic variables for the representation of knowledge [214]. Most applications developed in the 80s-90s were based on a "knowledge-based" approach based on the expertise of an operator for a given problem and of limited complexity. When one wanted to move on to more complex problems, it was difficult to write (even for an expert) large rule bases and the knowledge-based approach was no longer appropriate. In order to cope with this problem, one can take advantage of a type of knowledge [215] which constitutes the many input-output data on the process. Thus, the solution consists in using the properties of the ANNs in order to learn from the data (inputs/outputs) the fuzzy structure and to adapt the parameters accordingly.

The joint use of neural networks and fuzzy logic allows getting the benefits of the two methods: the learning capabilities of the first and the readability and flexibility of the second. To summarize the contribution of neuro-fuzzy, Table 3.4 groups together the advantages and disadvantages of neural networks and fuzzy logic.

Neural Networks	Fuzzy logic
Advantages	
Data-driven	No mathematical model
	Interpretability
Learning algorithm	
Disadvantages	
Black box model	Fitting adjustment
	Need to own the rules

Table 3.4: Advantages and Disadvantages of ANNs and FL

The fuzzy rules encoded in a neuro-fuzzy system represent inaccurate samples and can be seen as imprecise prototypes of the training data. A neuro-fuzzy (NF) system should not be seen as an expert system (fuzzy), and it has nothing to do with fuzzy logic in the strict sense of the term. It can also be noted that neuro-fuzzy systems can be used as universal approximators [8].

A definition of neuro-fuzzy systems is given by Nauck et al. in 1997 [216] and updated by Palluat et al. in 2006 [217] according to which:

The neuro-fuzzy systems are fuzzy systems formed by a learning algorithm inspired by the neural network theory. The learning technique operates according to local information and produces only local changes in the original fuzzy system.

So, the interest is to build a predictive system relies on the integration of neural networks and fuzzy inference systems (FIS) because of their complementarity. The FIS exploit linguistic rules



if-then; in order to translate knowledge about the dynamics of a system. Given a situation characterized at the date "t" (input), they thus make it possible to predict the evolution at "t + r" (output). However, a FIS is not able to learn: the rules must be formulated which is sometimes difficult ... So, one solution is to use the properties of the ANNs to "learn" the structure and fuzzy and adjust the parameters accordingly. In general, neuro-fuzzy networks replace the different hidden layers of neural networks by fuzzy rules (i.e. linguistic rules). They then use learning algorithms to define and optimize these parameters.

In addition, the rules of an NF system are transparent, allowing validation and manipulation by an expert [218]. As results, NF systems are very promising in cases where the available data are limited [9]. Finally, NF systems are suitable tools to support the prediction activity of the prognostic process. By being at the intersection of neural networks and fuzzy logic, neuro-fuzzy networks benefit from both methods. Neural networks are already a powerful tool. A neuro-fuzzy network also makes it possible to automatically determine the system parameters.

The Neuro-Fuzzy networks were chosen because with a dual interest: industrial and scientific (Table 3.5).

Industrial advantages	Scientific advantages
✓ Complexities of real systems	✓ Scientific opening
✓ Real system = dynamic systems	✓ Easily adaptable
✓ Real-time	✓ Capabilities of interpretable
✓ Model from data	✓ ...

Table 3.5: Industrial and Scientific Benefits of Neuro-Fuzzy Networks

3.9.3 ANFIS (Adaptive-network-based Fuzzy Inference System)

ANFIS represents a fuzzy inference system implemented as part of adaptive networks. It uses the hybrid learning procedure. This architecture (Figure 3.15) refines the fuzzy rules obtained by human experts to describe the input-output behavior of a complex system. This model gives very good results in trajectory tracking, nonlinear approximation, dynamic control and signal processing. Therefore ANFIS is the most used neuro-fuzzy architecture.

Our study will focus on ANFIS (Adaptive Neuro-Fuzzy Inference System). The following is dedicated to the presentation of this neuro-fuzzy architecture. An explanation by the example of the operation of this architecture will be made to better understand the mechanisms mentioned above. In addition, a comparative study between artificial neural network (ANN) and ANFIS method [4] confirmed that the mean percentage error generated by ANFIS algorithm is much less than generated by ANNs [11] and also ANFIS have better performance, and a faster learning process than ANNs and other conventional approaches [178]. Recently researches have shown



also that the use of the largest number of inputs in ANFIS method increases the accuracy of the forecasting data [11].

3.9.3.1 ANFIS a Hybrid Tool for Prognostic

Real systems are complex and generally non-stationary behavior and nonlinear making a modeling step difficult. Yet implementing a predictive tool must accommodate this. Also, various artificial intelligence techniques were tested on prediction problems and have shown better performance than those of "conventional" methods [212] [192] [219]. It is clear from this work that the neuro-fuzzy networks are particularly suitable. In this set, our work deals more specifically with the ANFIS (Adaptive Neuro-Fuzzy Inference System) proposed by Jang in 1993 [220].

The ANFIS is an adaptive array class. It can be seen as a feed-forward neural network for which each layer is a component of a neuro-fuzzy system and, as such, it is an "approximator" universal. It is thus used in various applications of predictions. ANFIS system performs a linear approximation of the output variable by decomposing the input space into different fuzzy spaces. Consider Figure 3.15 to describe the architecture of a system ANFIS and briefly explain the inference mechanism of such a system. Jang et al. in 1995 and Yam et al. in 2001 [192] proposed a more detailed view. Adaptive neuro-fuzzy inference system is a type of hybrid system that combines the evident knowledge of Takagi–Sugeno (TS) fuzzy inference system and the supervised learning potential of the multilayer feedforward neural network in one approach called ANFIS. It is a very robust technique that aims to achieve the nonlinear and complex relationship between input and output data [220], it is much simpler, suppose that we have two inputs x and y , and one output f . for “If-Then” of Takagi–Sugeno (TS) model, two rules are used as follows (eq.3.3) and (eq.3.4):

$$\text{rule1: if } x \text{ is } A1 \text{ and } y \text{ is } B1 \text{ Then } f1 = p_1x + q_1y + r_1 \quad (3.3)$$

$$\text{rule2: if } x \text{ is } A2 \text{ and } y \text{ is } B2 \text{ Then } f2 = p_2x + q_2y + r_2 \quad (3.4)$$

Where $A1$, $A2$, and $B1$, $B2$ are the membership functions of the two input x and y respectively and the $p1$, $q1$, $r1$ and $p2$, $q2$, $r2$ are linear parameters of the output of Takagi–Sugeno fuzzy inference model.

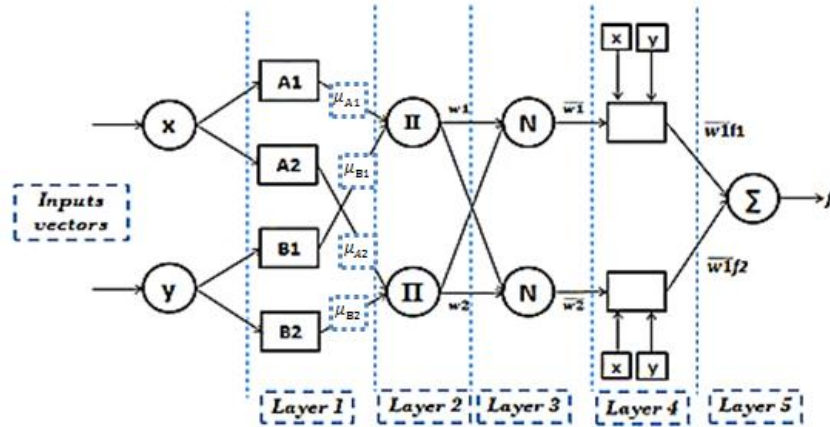


Figure 3.15: ANFIS Architecture

Figure 3.15 shows the most known architecture of ANFIS. Usually, it has two inputs and one output. As standard, ANFIS architecture has five layers. Where the first and fourth layers (square nodes) formed from an adaptive node, and fixed nodes (circle nodes) constitute the essential ingredient of the other layers. Explanation of each layer is described in the following paragraphs:

Layer 1: Each node of this layer is an adaptive network, $A1-A2$, and $B1-B2$ are four fuzzy parameters described by the type of membership function like Gaussian-shaped, triangle-shaped etc. [11] The output of each node in this layer represents the degree of membership value o_{1i} for the fuzzy set A_i and B_i , respectively (eq.3.5)

$$o_{1i} = \mu_{A_i}(x), i = 1, 2$$

$$o_{1i} = \mu_{B_{i-2}}(y), i = 3, 4 \quad (3.5)$$

Layer 2: Every node in this layer is fixed or non-adaptive, and represents the firing strength for each rule " w_i ". The T-norm, **AND** operator is applied to obtain the output as shown in (eq.3.6)

$$o_{2i} = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), i = 1, 2 \quad (3.6)$$

Layer 3: The main role of this layer is the normalization of the firing strength. The nodes in this layer are fixed, circle and labeled by N symbol. Each node calculates the ratio between the i -th rules firing strength and the sum of all firing strengths (eq.3.7):

$$o_{3i} = \bar{w}_i = \frac{w_i}{\sum_i w_i}, i = 1, 2 \quad (3.7)$$

Layer 4: The nodes in this layer are adaptive with a node function (eq.3.8):

$$o_{4i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), i = 1, 2 \quad (3.8)$$

Where, \bar{w}_i is the normalized firing strength calculated on layer 3, (p_i, q_i, r_i) are the



consequent parameters of this layer and \mathbf{x} , \mathbf{y} are the inputs vectors.

Layer 5: This layer formed from a circle, fixed, and (the) single node labeled as Σ . It represents the sum of all overall output of the layer 4 (eq.3.9).

$$o_{5i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, i = 1, 2 \quad (3.9)$$

The gradient descent algorithm [221] combined with the least squares method to form hybrid learning algorithm proposed by Jang (1993) [220], used to learn the ANFIS algorithm or to update the nonlinear premises parameters in layer 1 and the linear consequent parameters in layer 4. Forward path and backward path are the two ways of the hybrid learning algorithm. Firstly, on the forward path, the premises parameters in layer 1 are fixed and a recursive least square estimator (RLSE) method was applied to update the consequent parameters in the layer 4. The linearity of the consequent parameters is the reason to use the RSLE method which aims to accelerate the convergence rate in hybrid learning process. While the algorithm is in the backward way the consequent parameters are fixed and the gradient descent algorithm runs to update the premises parameter in the layer 1, and an error generated that represents the difference between the desired output and the actual output, is propagated back to the first layer [220].

3.13 Frame of ANFIS in the Prognostics Domain

As long as ANFIS is a combination of data-driven method and knowledge-oriented method, therefore, when focusing on the process of Prognostics based on ANFIS or any others data-driven method data, we can highlight a stream that passes from multidimensional data to the RUL of the system. Based on this, it is essential for Prognostics to rely on the following steps:

- 1- Data acquisition,
- 2- Data processing: feature extraction and selection
- 3- Prognostics modeling: RUL estimation.

3.13.1 Data Acquisition

In the implementation of PHM program, Data acquisition or data collection is represents an important step for industrial system diagnostic and prognostic. In general, the data acquired for PHM can be classified into two groups [165]:

1. Event Data
 2. Condition Monitoring Data
- 1- Event data: it represents the machine archived files that include all preventive and corrective maintenance done on this machine; this type of data provides a clear image of the machine health status from the moment of installation to the moment of the last measure.



- 2- Condition monitoring (CM) data: it represents the data acquired or collected during an experiment applied on the machine for identifying their physical parameters of health conditions, where the collection of these parameters aims to detect the changes on this machine that can lead to a failure of the machine. These parameters can be temperature, vibration, acoustic, pressure, etc. the collection of this data is done via a set of sensors that are built into the device or added as an external monitoring system, this set can include many types of sensors like an accelerometer for vibration recording or a barometer for pressure registration etc. For more details please see the section of data Collection for faults prognostic in chapter 1.

3.13.2 Data Pre-processing

As mentioned in chapter 1, a section of data pre-processing, the data acquired from the Physical system are noisy and redundant. Therefore, this data cannot be applied directly by a prognostic model. Further, this original data could be hidden inside many of the relevant information which can denote to the system degradation. Wherefore a set of parameters that plain of a relevant information should be extracted from this original data as indicators for this degradation. Usually, the strength of the prognostic model depends on the quality of the extracted and selected features. In addition, it is very important to identify the features that reflect the progression of the failure in the system and can be used to construct the prognostics model.

3.14 New Prognostic Approach for Preventive and Predictive Maintenance- Application to a Distillation Column

3.14.1 Introduction

Nowadays, the research activities concerning the increase in the availability of the industrial systems at less cost and in best performance represent one of the most important things in the industrial safety domain [1]. Temperature, pressure and other extreme conditions may cause malfunctions in devices and a little deviation could cause unexpected accidents, i.e. the explosion of the four reactors at the Daiichi nuclear plant Fukushima-Japan in 2011 [16]. The role of computer simulation and analysis is still restricted in the ability to preserve process plant safety and highly dependent on human operators. Humans could not be able to discover the hidden faults or predict future failures. Industrial statistics showed that the major disasters may be rare, but the minor accidents are very frequent with an annual cost that exceed billions of dollars [1], [3]. Therefore, it's necessary for the industrialists to surround the severity status of the fault and to predict the ideal moment to intervene and stop the instrument. This is known as the prognostic process [27].

In the previous works [164] [165], authors have proved that the prognostic topic represents a work main frame that ensures the safety for industrial environment, and is considered as a key process in maintenance strategies. In 2007, Dragomir et al. suggested that a prognostic process



has two main activities: It can be used to predict the evolution of a situation at a given time as well as to assess this predicted situation with regards to a referential [86].

In the present study, we propose an approach that can be used in real time analysis of distillation column system. The adaptive neuro-fuzzy inference system (ANFIS) is chosen as the hybrid system since it combines the advantages of fuzzy logic and ANNs in the same algorithm. This methodology is then tested on a real experimental data obtained from a distillation column, after a pre-processing step including filtering and smoothing of the signals. A database with normal and faulty (degraded) observations is analyzed. The database is composed of eight different types of faults that may occur during the automated distillation process in the chemical industry. Depending on the state of art we note that the recurrent ANN is the first competitor for the ANFIS in terms of accuracy of prediction, Therefore, our work must include a comparison study between the recurrent ANN and the ANFIS to choose the best of them in terms of the highest ability to predict the data acquired from the distillation column system.

3.14.2 Data Acquisition and Measurements

The distillation process is the most method used in terms of separation because it's simple in terms of the other unit operations; therefore it plays an important role in the operations unit, especially for the liquid-liquid extraction and the fractional distillation that largely used to separate the fuel oil components.

The methodology in this part of research is applied on a real experimental data obtained from a distillation column installed at IUT of Rouen-France. A database with normal and 7 types of degradations is analyzed. The database is composed of 50 observations of each type when increasing or decreasing faults over time (degradation of the system) that may occur during the automated distillation process in the chemical industry.

The signals obtained from every observation are:

- 1- S1: Timer: Reflux Rate,
- 2- S2: Heating Power,
- 3- S3: Feed Flow Rate,
- 4- S4: Preheated Power,
- 5- S5: Loss of Charge,
- 6- S6: Pre-Heated Temperature,
- 7- S7: Boiler Temperature,
- 8- S8: TIC2: Column Head Temperature.

While an accident occurs in the automated distillation process, it causes a cumulative increasing or decreasing over time on the following parameters:

- 1- The Reflux Rate.
- 2- The Heating Power.
- 3- The Preheating Power

4- The Feed Rate.

This is clearly illustrated in Figure 3.17 and this variation represents the data failure of the system. For more details see chapter 1.

For example, if the system was harmed by a cumulative degradation of reflux rate (green signal in Figure 3.16), in this case, we can see that the temperature at the top of the column TIC2 (brick signal in Figure 3.16) has increased very rapidly from 55 °C to 75 °C. which means this temperature (75°C) is very near to the boiling temperature of ethanol, which is 78.6 °C. This is explained by the fact that there is no liquid falling back into the column. Indeed, to compensate for the decreasing in the pressure drop, the boiler heater must send more steam into the column, so the heating increases. In the column, the vapor remains only, so the TIC2 temperature increases strongly. Because of this, the liquid level in the boiler decreases rapidly and the level of the distillate increases. Therefore should interrupt the system because there was a real danger.

In summary, this type of accident can have serious consequences. In fact, if the liquid is completely empty in the boiler, the heater can be seriously damaged and the boiler may explode.

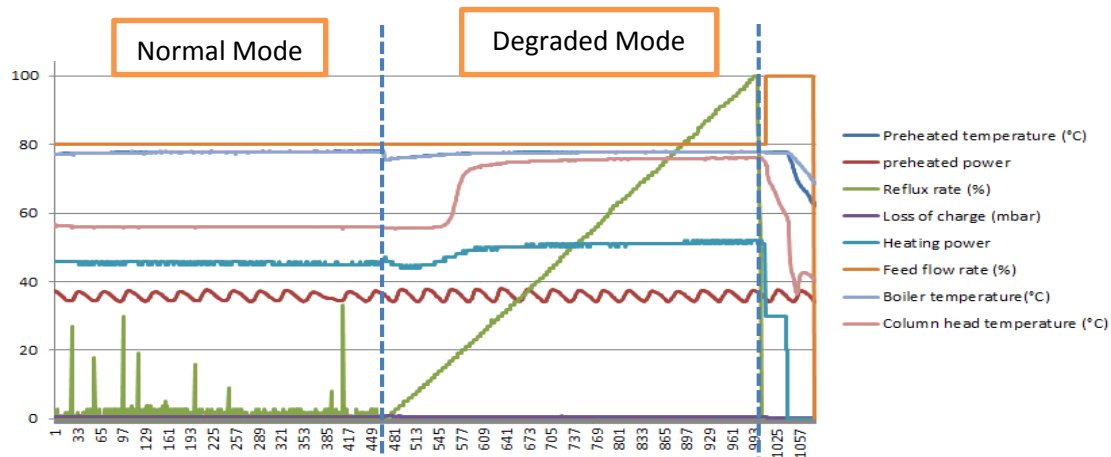


Figure 3.16: Graphical Representation of the Signals (Normal & Degraded Modes)

A database with normal and faulty observations is analyzed. The database is composed of 42 observations; each observation has 1507 points with the sampling frequency, $SF=1\text{sample}/10\text{sec}$.

On the other hand, the aging of the distillation column components should be considered, especially the aging of the metering pump, because it controls the input flow rate. When we have a problem with the pump Impeller, the flow rate will decrease in this case. The most serious consequence would be when the Timer remains open and thus all the liquid in the boiler turns into vapor. This vapor will condensate when it in contact with the condenser which installed at the top of the column. The result is a liquid stocked at the distillate reservoir.

The result of all this it will be a rapid drop in the level of the liquid that is found in the boiler, which means the heater is no longer in contact with the liquid and this could cause a break in the boiler. To simulate the pump aging, we scratch multi times the Impeller of the pump and measure



each time the generated vibration via an accelerometer fixed on the pump chassis and a data acquisition system connected to Labview software. Figure 3.17 shows a signal representative of the acceleration dimension z , where the x -axis is the sample index and y -axis is the amplitude of z acceleration signal. This procedure (introducing more scratches) is repeated many times. A database with normal and faulty observations is analyzed. This database is composed of 10 observations, each observation is 5×10^4 points with the sampling frequency, $SF=500$ samples/sec. A pre-processing step including filtering, normalization, and smoothing is applied to the data before processing.

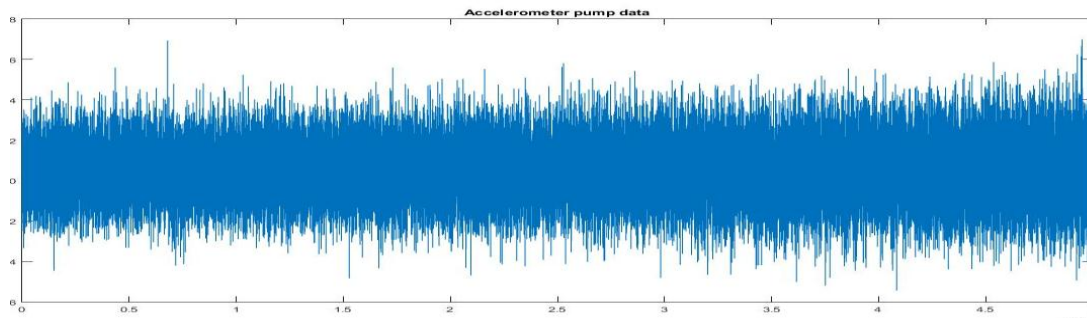


Figure 3.17: Graphical Representation of Accelerometer Signal (Normal & Fault Mode)

So far in the previous sections of this chapter, we have presented the main framework for process fault prognostic. In this section, we will extend a real application for a prognostic purpose. The designed methodology is proposed to predict seven types of degradation that can be maybe occurred in the distillation column (D1...D7) and metric pump impeller degradation D8.

- D1: Increasing Degradation Reflux Ratio From 0 To 100% (Timer);
- D2: Increasing Degradation Heating Power 0 To 100%;
- D3: Decreasing Degradation Heating Power 100 To 0%;
- D4: Increasing Degradation Preheated Power 0 To 85 %;
- D5: Decreasing Degradation Preheated Power 100 To 0%;
- D6: Increasing Degradation Feed Rate 0 To 100%.
- D7: Decreasing Degradation Feed Rate 100 To 0%;
- D8: Pump Impeller Degradation.

Total data used in this research is 66 data observation. Where 70 % of them are used for training and the 30% remaining is used for testing and checking.

For more information see the section (Data Collection for faults prognostic) in chapter 1

3.14.3 Time Series Calculation

The failure occurrence it may be at any time without indication, so it must be a reliable method able to save the system, and gives the quick and accurate prediction (short, mid and long-term), to prevent any unexpected event. To get good results and to protect the system, ANFIS should be properly and strongly trained, [178], therefore a delayed matrix should be created and

used as input of ANFIS as shown in Figure 3.18.

As a description of Figure 3.18, the x -axis is the sample index and y -axis is the signal amplitude. The red dashed lines represent the time series generated as the input of ANFIS, and the green circles are the predicted points.

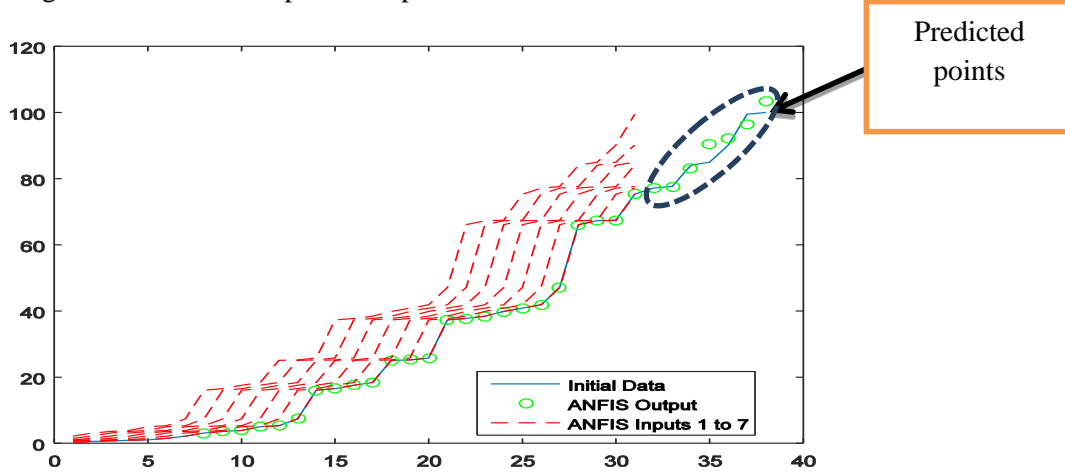


Figure 3.18: Graphical Representation of Time Series Creation

Let's consider a space X with (p):

$$X : [x_1, \dots, x_r, \dots, x_p]$$

Prognostic is a process based on the collecting data of the past $x(t-k)$ and present states $x(t)$ of the system, to predict the future ones $x(t+k)$.

X is considered as the signal to be predicted; from this X we should create a time series matrix (shifted matrix) to train the prediction algorithm as inputs of ANFIS algorithm.

Let's consider a delay vector as follow (eq. 3.10): Delay = $[1, 2, 3, \dots, n]$ $n < P$

Where n : is the numbers of inputs vectors and it determines the type of the prediction mode (short, mid or long-term) when n grows that means we're heading to long-term prediction.

$$\text{Inputs : } \begin{bmatrix} x(t) & x(t+1) & \dots & x(t+(P-(n+1))) \\ x(t-1) & x(t) & \dots & x(t+(P-(n+2))) \\ \dots & \dots & \dots & \dots \\ x(t-(n-1)) & x(t-n) & \dots & x(t-(P-(n+1))) \end{bmatrix} \quad (3.10)$$

Target : $[x(t+1), \quad x(t+2), \quad \dots \quad x(t+(p-n))]$

3.14.4 Fuzzy C-Means Clustering (FCM) for Health Assessment

In this study, FCM is used as unsupervised classification phase for lifetime assessment. The main role of the unsupervised classification is to dividing the degradation curve into many classes. Then, these classes are used to estimate discrete states and to determine the time to failure



(TTF) or the remaining useful life (RUL) of this system. When the temporal prediction achieved by the predictor, the classifier (FCM) estimate the future status of the system.

To handle unlabeled data and due to the using of real information, FCM approach represents the best technique able to resolve this problem, [221], [222]. Fuzzy C-means (FCM) is a method of clustering that allows one piece of data to belong to two or more clusters. Whereas, FCM is frequently used in pattern recognition, [223]. This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of the distance between the cluster center and the data point. More the data is near to the cluster center more the membership towards to the particular cluster center. Clearly, the summation of membership of each data point should be equal to one. After each iteration, the membership and the cluster centers are updated and minimized [3]. For more information see the section 2.8.4.1 (Fuzzy c-Means clustering (FCM)) in chapter 2

3.14.5 Proposed Methodology

Previously, ANFIS was used to calculate the RUL of the system depends on a prediction step followed by the location identification of the predicted value. In our study the database can be divided into two types:

- 1- Data acquired from distillation column via ATP200, this type of data is deterministic and has a low sampling frequency (1 sample/10sec) (see normal mode in Figure 3.16 and Table1.1). It's clear to find the difference between normal and degradation mode (Figure 3.16).
- 2- Data acquired from the pump via accelerometer, this type of data is rich in information (500 samples /sec). In this type of data, it's very difficult and complex to differentiate between normal mode and degradation mode (Figure 3.17). Usually, features extraction technics are applied to the originally acquired signal to find the better parameter (mean, variance, kurtosis,...) who can represent the system degradation [224]. In our case any parameter could not achieve the desired goal, for this reason, it is necessary to think about another way able to discover this small variation in the signal over time and put a strategy for the system path or to find the better curve that can represent the system degradation. Then use this degradation curve to calculate the RUL of the studied system certainly after passing several steps will be detailed later.

As a description of Figure 3.19, the *x-axis* is the sample index and the *y-axis* is the amplitude of degradation.

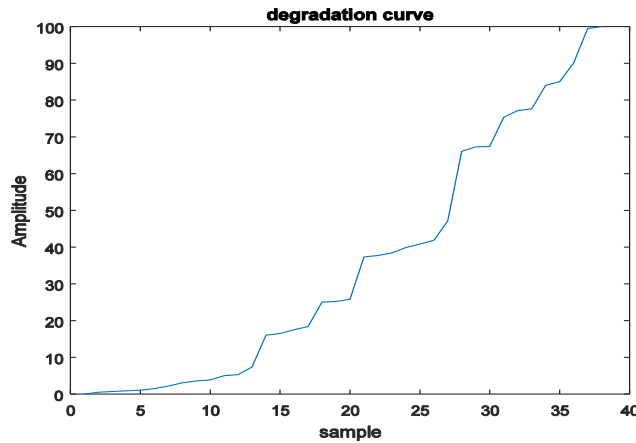


Figure 3.19: Graphical Representation of System Degradation Mode: Case of Aging of the Pump Impeller

The proposed methodology can be summarized by the following four steps:

Step1: Training of ANFIS, Max of Kurtosis Calculation

- *Train ANFIS using normal data (this produce ANFIS 'A1').*
- *Apply ANFIS on whole signal, this produce error ϵ_1 .*
- *Estimate the max of kurtosis vector ($mk1$) from this error ϵ_1 .*

Step2: Degradation Curve Calculation

- *Apply ANFIS (A1) on real degraded data, this produce error ϵ_2 .*
- *Estimate the kurtosis vector ($k2$)*
- *Calculating the degradation curve via ($mk1 - k2$).*

Step3: Modeling and Clustering

- *Modeling the degradation curve by: $f(x) = px^2 + qx + r$.*
- *Clustering this model via FCM algorithm, produce 3 classes.*

Step4: For Real-Time Application, For New Data:

- *Apply ANFIS (A1) on large window of unknown data.*
- *Calculating the error vector ϵ' .*
- *Estimate the kurtosis vector k' from ϵ' then calculate ($mk1 - k'$)*
- *Prediction of new kurtosis values from $mk1 - k'$.*
- *Project predicted kurtosis values to the clustered classes, calculate the RUL.*

These steps can be represented by a simple block diagram shown in Figure 3.20, Figure 3.22, Figure 3.23 and Figure 3.25. In the first step (Figure 3.20), a normal signal is divided into N segments (S) where each segment has n samples ($n=500$). $[(S_j), j = 1, \dots, N,$



N : number of segments]), then a time series matrix is generated from each segment as an input to ANFIS algorithm, where 70% of this matrix is used to train ANFIS, and the remaining amount (30%) is used as testing step for ANFIS validation as shown in Figure 3.21a. Accordingly a vector (n elements) of residual error (ε_j) is generated between the target and the actual output of ANFIS. Then the kurtosis value (k_j) (eq.3.11) of this error is calculated. When we are on the segment N , the kurtosis (k_N) is calculated, in the end, we have got a kurtosis vector (K_1) consists of N elements. Then calculate the maximum (mK_1) (eq.3.12) from this kurtosis vector (K_1). (mK_1) is calculated to be used as a normal reference. This step was produced an ANFIS(A_1), where the premise and the consequence parameters of this ANFIS updated based on normal data.

$$k_j = n \frac{\sum_{i=1}^n (X_i - \bar{X}_{avr})^4}{(\sum_{i=1}^n (X_i - \bar{X}_{avr})^2)^2}, \quad j = 1, \dots, N \quad (3.11)$$

$$mK_1 = \max[k_j, k_{j+1}, \dots, k_N] \quad (3.12)$$

Where:

n : Number of samples for each segment.

N : Number of segments

To calculate the kurtosis, you first need to calculate each observation's deviation from the mean (the difference between each value and arithmetic average of all values). The deviation from the mean for i^{th} observation equals: $(X_i - \bar{X}_{avr})$. After many tests, the kurtosis parameter is chosen because it gives a convincing and very clear difference between failure modes more than the other parameters (time and frequency parameters) such as the mean, variance, skewness, median frequency,... etc. In the case of normal data, the error is always small (samples index from 0 to 650 as shown in Figure 3.21.a and Figure 3.21.b). If the error increases between the ANFIS output (red line in Figure 3.21.a) and the real checked signal (black line in Figure 3.22.a), this is an indication that the signal starting deviation from the normal condition (samples index from 651 to 920 in Figure 3.21.a and Figure 3.21.b).

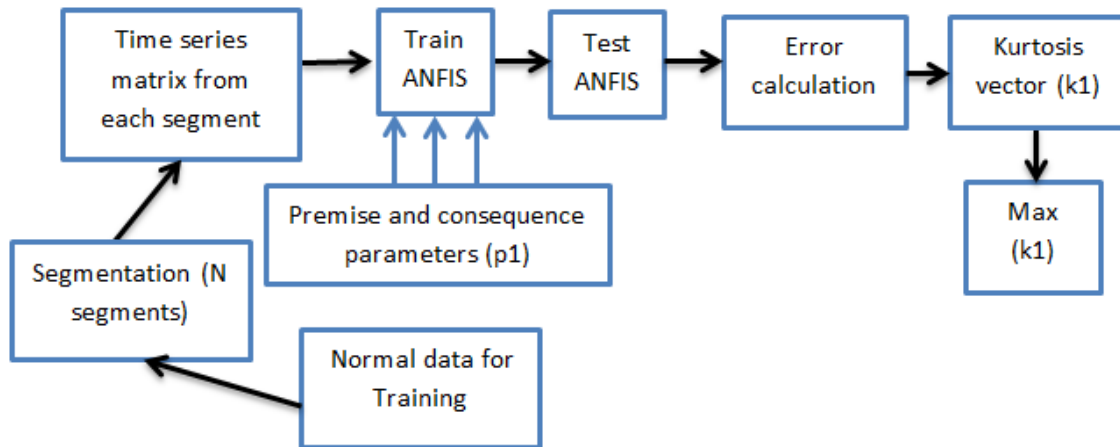


Figure 3.20: Training of ANFIS

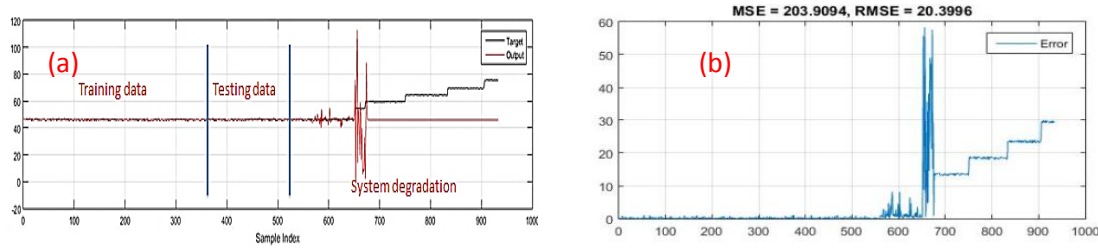


Figure 3.21.a: Graphical Representation of ANFIS Training, Testing and Checking

Figure 3.21.b: Graphical Representation of the Error between Target and ANFIS Output

In the Figure 2.21. a the *x-axis* is the sample index and *y-axis* is the signal amplitude. For the Figure 3.21.b, the *x-axis* is the sample index and the *y-axis* is the error amplitude.

In step 2 (Figure 3.22), after training of ANFIS (A_1) with the same parameters (premise and consequence) calculated before in step 1, a segmentation procedure is applied on a real historical degraded signal (M segments each segment has a length of m samples with no overlap, $m=500$). For each segment, we applied ANFIS (A_1) and a residual error (ε_j) is produced as resua It of this process, then the kurtosis value (k_j) (eq.3.11) is calculated from this error. When we are on the segment M, the kurtosis (k_M) is calculated, at the end, we have got a kurtosis vector (K_2) which consists of M elements.

$$K2 = [k_j, k_{j+1}, \dots, k_M] \quad (3.13)$$

At the end of this part the absolute vector of the difference (diff) (eq.3.14) between the max of statistical kurtosis vector (mk_1) (eq.3.12) and statistical kurtosis vector K_2 , (eq.3.13) is calculated. Each point of the vector (diff) is the difference between the max of the kurtosis vector calculated from the normal data and the kurtosis vector calculated from the degraded data. The value of each point gives an indication how far their related data from the normal reference (mk_1). The result was a reliable curve represents the degradation profile of the system as shown on Figure 3.22.

$$diff = |mk1 - k2| \quad (3.14)$$

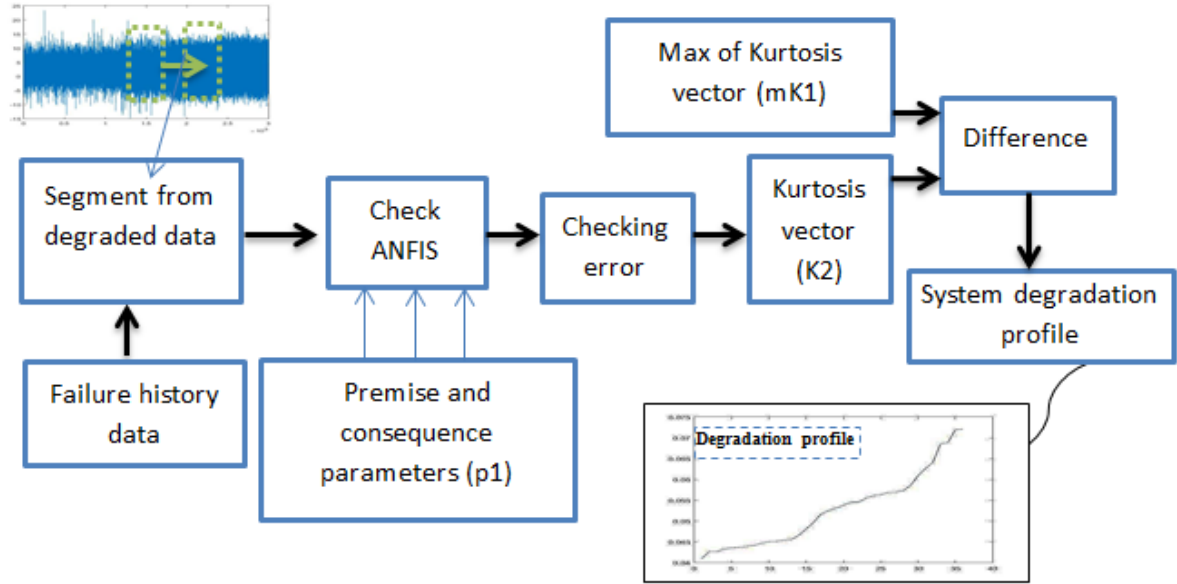


Figure 3.22: Designing of Degradation Curve

In step 3 (Figure 3.23), after obtaining the degradation profile that represents the functional mode of the system, a fitting curve $f(x)$, (eq.3.15) will be estimated or modeling this degradation profile.

$$f(x) = px^2 + qx + r \quad (3.15)$$

For higher accuracy, five fitting curves are calculated from five real acquired data. The best fitting curve is the mean of these five fitting curves, Figure 3.24). Many functioning modes should be taking into consideration. In this study, three stages are proposed: initiation class, progression class and critical class. Therefore, FCM algorithm is used to partitioning this model of degradation into n classes determined by the user ($n=3$ in our study). FCM returns the coordinates for each cluster center $(c_i, i = 1,2,3)$ and the membership function (matrix U) that contains the membership grade of each data point in each cluster.

$$C = [c_1, c_2, \dots, c_n], n = 3 \quad (3.16)$$

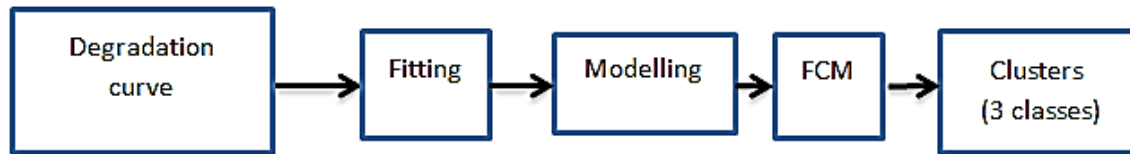


Figure 3.23: Fitting and Clustering of Degradation Model

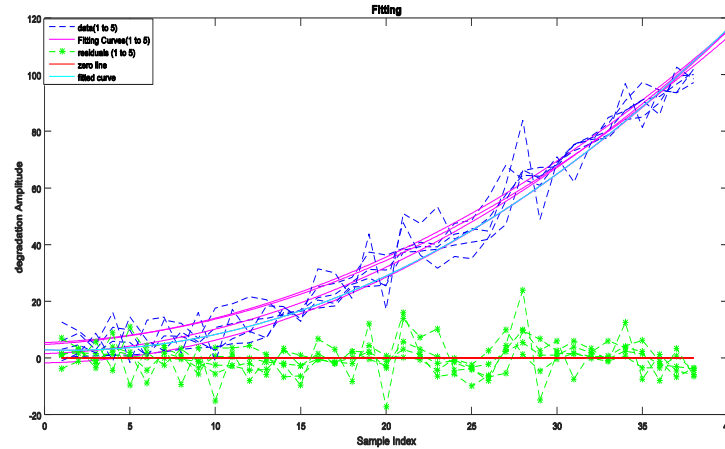


Figure 3.24: Graphical Representation of Fitting Curves to 5 Real Observations Degraded Curves

To obtain an accurate degradation curve, many observations are executed (5 observations), In Figure 3.24 the dashed blue lines are the real five degradation curves, the violets lines are the fitting curves of the degradation curves, the sky blue line is the average of the all fitting curves, the green star's lines are the residual between the real degradation curve and their fitting curve (real degradation curve-fit curve), and the red line is the zero line.

In step 4 (Figure 3.25), for the real application, when it is in the online mode, an unknown new large signal is acquired then segmented (L segments). ANFIS (A_1) is applied again on each segment and a new kurtosis value calculated. At the end of the segmentation process, we obtain a vector contains L kurtosis values which called (K'). Then, the absolute of the difference ($|mK_1 - K'|$) is calculated as an indication of the data degradation state.

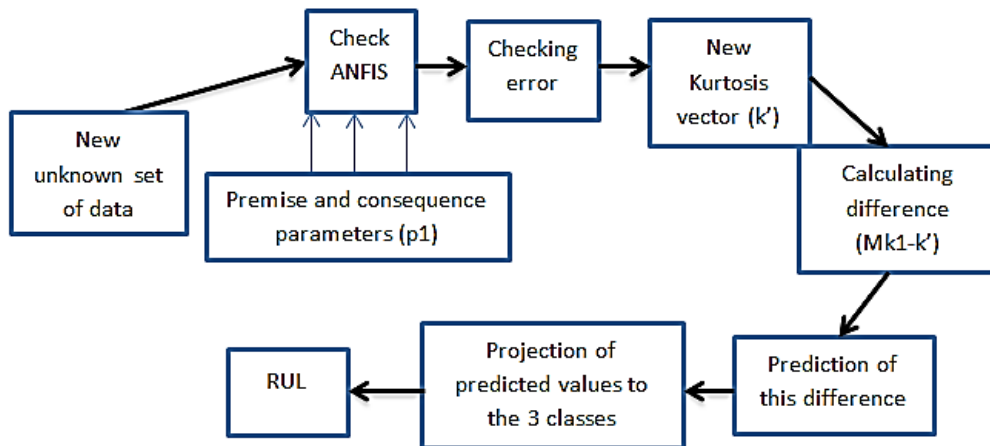


Figure 3.25: Testing, Prediction, and Clustering of New Set of Data

Using this vector of kurtosis, a new ANFIS is applied to predict the new kurtosis values; this gives an indication about the future distribution of this vector.



When the predicted vector calculated using ANFIS, FCM with the same coordinates of centers (C) calculated before (eq.3.16) and the same membership function (U) is used to project the temporally predicted values of the vector P to one of the three classes (initiation, progression or critical). This based on fuzzy partitioning of multidimensional data, that means the membership degree is calculated ($\lambda_i, i = 1, \dots, n$) (eq.2.14) of each predicted point in each class. Then (λ) is used to calculate the gravity center (GC), (eq.3.17). Finally, the RUL calculated using this formula ($RUL = th - GC, th = \text{threshold of 100\% system degradation}$). For an accurate prediction results, if the RMSE between the predicted point and the profile of degradation is very large, the prediction process should be refused and repeated again. Finally, the user will get a clear vision, able to expresses for him/her how the system is far from the critical or failure threshold [222] .

$$GC = 100 * \frac{\sum_{i=1}^n \lambda_i c_i}{\sum_{i=1}^n c_i}, i = 1, \dots, n \quad (3.17)$$

3.14.6 Results and Discussions

In this section, we will extend a real application for a prognostic purpose. The designed methodology is proposed to predict seven types of degradation that can be may occur in the distillation column (D1...D7) and metric pump impeller degradation D8.

Previously, all prognoses researches applied to distillation column system have objectives other than what we aspire to it in this research. Sahraie et al, in 2014 [225], use ANFIS as a prediction algorithm to estimate the composition of the distillate product. Sivakumar et al, in 2010 [226], use ANFIS as an estimator to control the composition of the distillate, control of the composite of bottom products, control of liquid retention in reflux cylinder and control of liquid retention at the base of the distillation column.

This study is the first one designed to predict the degradation of the distillation process and distillation column system. Suddenly faults on reactors may be disastrous, deadly and very expensive, especially when we have no knowledge about the future operations of the system, for this reason this work provided an efficient fault prognostic method when tested on a real distillation process.

Previous studies relied on the future extraction techniques (time and frequency parameters) to find the better parameter that can represent the system degradation, but usually it very clear and very easy to find out. In this work, when we apply the features extraction technique to the degraded signals, acquired from the metering pump via accelerometer shown in Figure 3.17, the obtained results were not satisfied, and they could not achieve the desired goal. Therefore it was necessary to think carefully about a new method able to solve this problem. The biggest challenge in this study is the using of ANFIS firstly as an unusual known way, where it has the same functionality of features extraction technique and secondly it used as usual known that is the approximation of nonlinear functions.

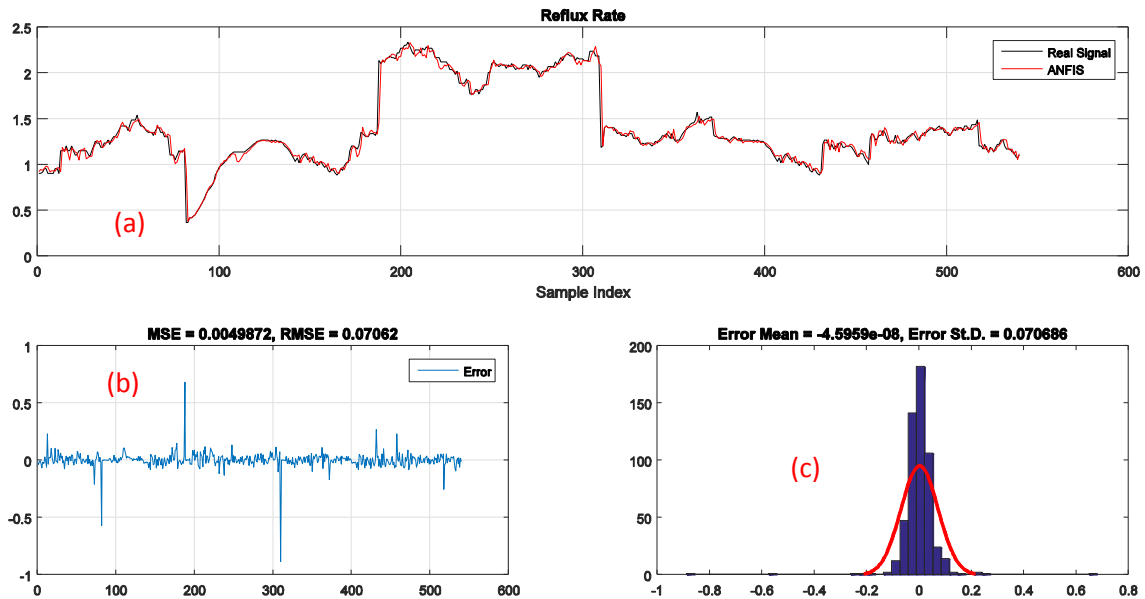


Figure 3.26.a: Graphical Representation of Results-Using ANFIS as A Predictor of Distillation Column Data (Reflux Rate Signal)

Figure 3.26.b: Graphical Representation of Absolute Value Of Error Calculated In Each Iteration

Figure 3.26.c: Graphical Representation of Error Absolute Value Distribution

The black curve in Figure 3.26.a is the real signal and the red one is the output of ANFIS, the *x-axis* is the sample index and the *y-axis* is the amplitude of reflux rate signal.

Degradation type	Average RMSE	Average Error St.D	Average ME
S1	0.17	0.23	1.5×10^{-8}
S2	0.016	0.01	2.7×10^{-8}
S3	0.21	0.35	3.3275×10^{-8}
S4	0.01	0.03	2.4×10^{-8}
S5	0.07	0.055	0.74×10^{-8}
S6	0.117	0.65	1.2×10^{-8}
S7	0.041	0.4	0.9×10^{-8}
S8	0.093	0.19	1.78×10^{-8}

Table 3.6: Performance of ANFIS when it applied on distillation column signals

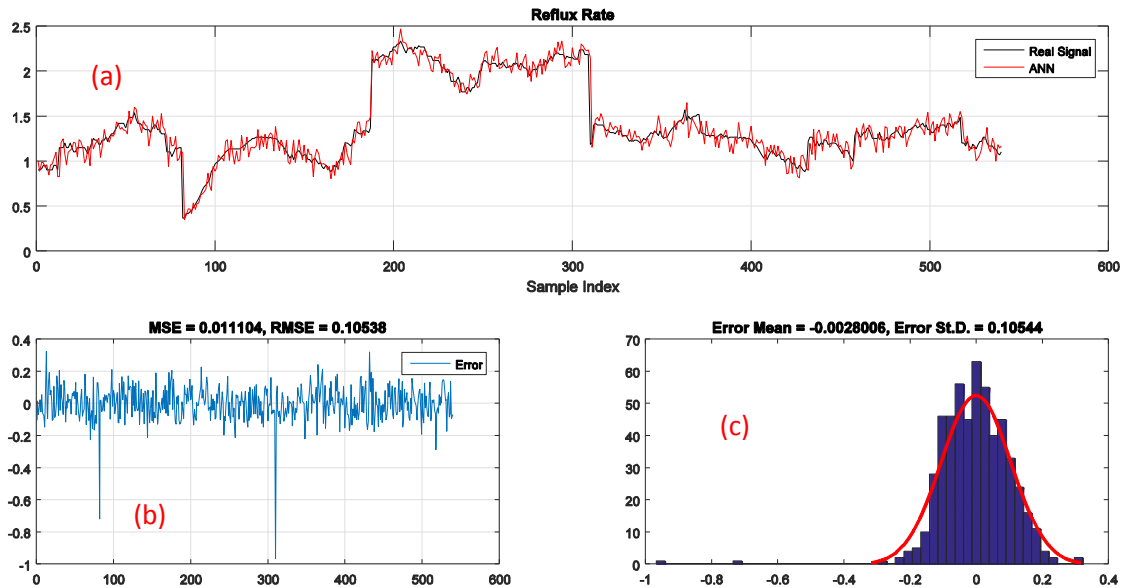


Figure 3.27.a. Graphical Representation of the Results - Using ANN as a Predictor of Distillation Column Data (Reflux Rate Signal)

Figure 3.27.b. Graphical Representation of Absolute Value of Error (Calculated Each Iteration)

Figure 3.27.c. Graphical Representation of Error Absolute Value Distribution

The results are shown on Figures 26.a.b.c, and the Table 3.6 confirms that ANFIS has the ability with a great confidence to predict the distillation column data ($MSE \cong 0.004$, $RMSE \cong 0.07$, Error mean $\cong -4.5 \times 10^{-8}$, Error standard deviation $\cong 0.07$). Referring to the Tables 3.6 and 3.7 it is clear that ANFIS has better performance and higher predictability results when it compared with ANN (Figure 3.27 a.b.c) for most of the tested signals.

Degradation type	Average RMSE	Average Error St.D	Average ME
S1	0.3060	0.5290	0.0026
S2	0.0288	0.0230	0.0049
S3	0.3780	0.8050	0.0016
S4	0.0180	0.0690	0.0038
S5	0.1260	0.1265	0.0035
S6	0.2106	1.4950	0.0020
S7	0.0738	0.9200	0.0033
S8	0.1674	0.4370	0.0037

Table 3.7: Performance of ANN When it applied on Whole Signals Extracted from Distillation Column

That indicates that ANFIS can be used with higher confidence to prognostic the degradations that may occur in the distillation column system and we can rely on it to calculate the RUL of the system.

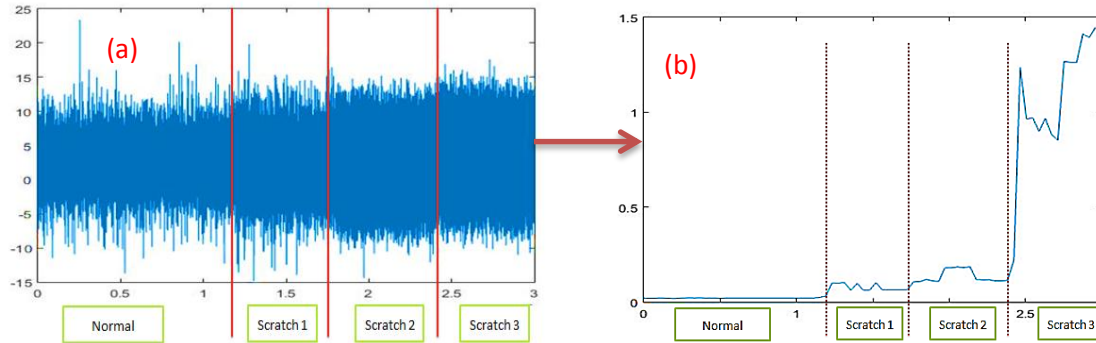


Figure 3.28.a. Vibration Signal extracted from Degraded Impeller Pump
Figure 3.28.b. Calculation of System Degradation Curve Using ANFIS

Figure 3.28.a shows the vibration signal acquired from the metric pump in two cases, pump without any malfunction (normal data) and pump with impeller degraded over time (degraded data). Figure 3.28.b explains the results when ANFIS used as a technique can capture the degradation curve of the pump. The *y-axis* in Figure 3.28.b represents the kurtosis amplitude.

It can be shown that the result of the step2 of our proposed methodology Figures 3.28.a.b represent a good indicator or feature that can illustrate the degradation occurs on vibration signal acquired from the pump has an impeller degraded over time, all this after the successfully of this strategy on discovering and detecting the smallest variation in the signal (Figures 3.28.a).

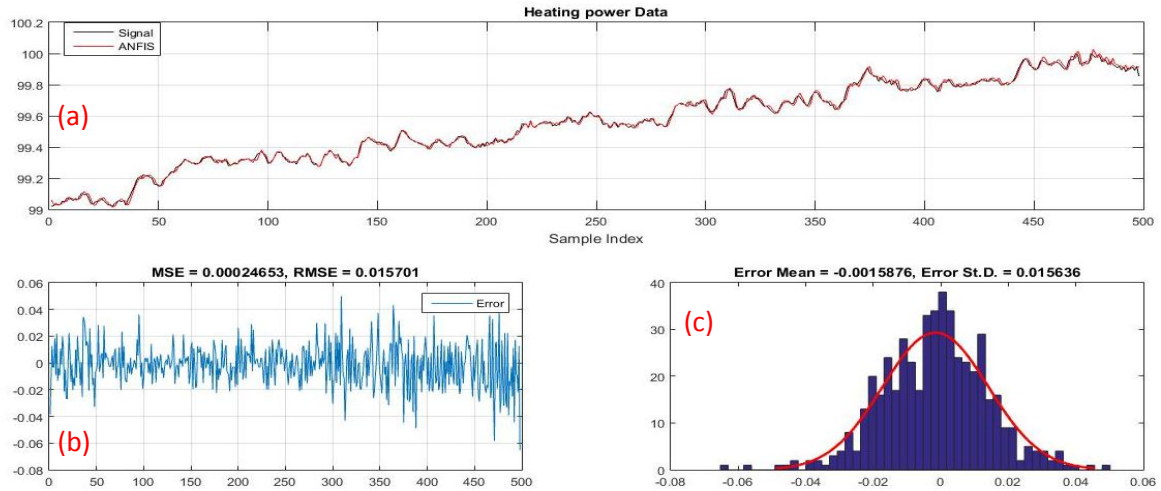


Figure 3.29.a. Graphical Representation of Results-Using ANFIS as a Predictor of the Degradation Curve D3

Figure 3.29.b. Graphical Representation of Absolute Value of Error Calculated Each Iteration

Figure 3.29.c. Graphical Representation of Error Distribution

In Figure 3.29.a, the black line is the real degradation curve calculated from degraded heating power signal (D2) that represents one from eight degradations types (D1...D8). Referring to the Figure 29.b, D2 has $RMSE \cong 0.01$, Error standard deviation $\cong 0.01$ and maximum error amplitude (MEA), ($MEA \cong 0.05$). The smallest in RMSE, in the standard deviation of residual error and in MEA affirm that ANFIS has the ability with great confidence to predict the D2.

For all other degradation modes, you can find all the results in Table 3.8 (column 3, column 4 and column 5).

Degradation type	Best MF	Average RMSE	Average Error St.D	Average MEA	Average MRES	Average SRES	RUL percentage
D1	Trapezoidal-Shaped	0.17	0.23	0.63	-9.0688e-15	1.5	86.2
D2	Pi-Shaped	0.016	0.01	0.07	-1.5360e-15	2.7	45.2
D3	Trapezoidal-Shaped	0.21	0.35	0.91	-1.1084e-14	3.3275	28.8
D4	Pi-Shaped	0.01	0.03	0.06	3.1974e-14	2.4	2.45
D5	Trapezoidal-Shaped	0.07	0.055	0.11	-3.3718e-15	0.74	83.6
D6	Pi-Shaped	0.117	0.65	0.85	1.0936e-14	1.2	47.2
D7	Trapezoidal-Shaped	0.041	0.4	0.32	1.0499e-14	0.9	73.8
D8	Trapezoidal-Shaped	0.093	0.19	0.42	1.0097e-14	1.78	71.5

Table 3.8: Statistical Characteristics of Eight Degradation Modes

In Table 3.8, the column 2 represents the best membership function for each degradation type used in layer1 of ANFIS architecture, the columns 3, 4 and 5 represent the average RMSE, average of standard deviation of the residual error and the average of MEA respectively for 10 observations from each degradation mode (D1 to D8).

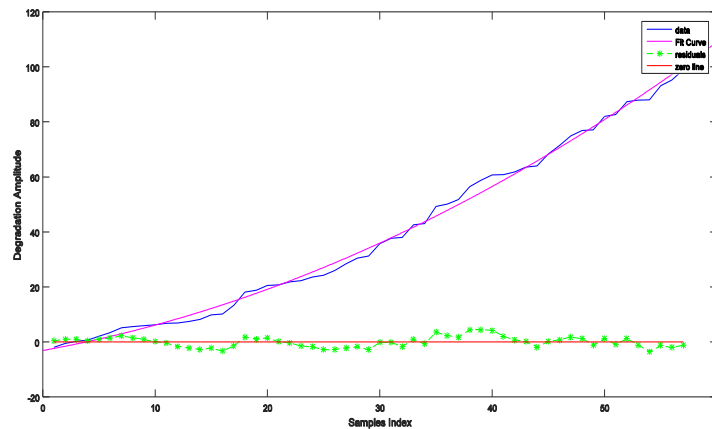


Figure 3.30: Fitting of Degradation Curve (D4)

Figure 3.30 represents the result of the fitting of degradation curve (D5), the blue line is the real degradation curve (RDC), the purple line is the calculated fitting curve (CFC), the green line is the residual (RES) between the RDC and CFC, ($RES = RDC - CFC$) and the red line is the zero line.



In Table 3.8, column 6 represents the average mean of 10 RES (MRES) calculated from 10 observations from every degradation mode. Column 7 represents the average standard deviation of 10 RES calculated from 10 observations from every degradation mode (SRES). As mentioned before the best-chosen fitting curve is the average of all fittings curves calculated from 10 observations.

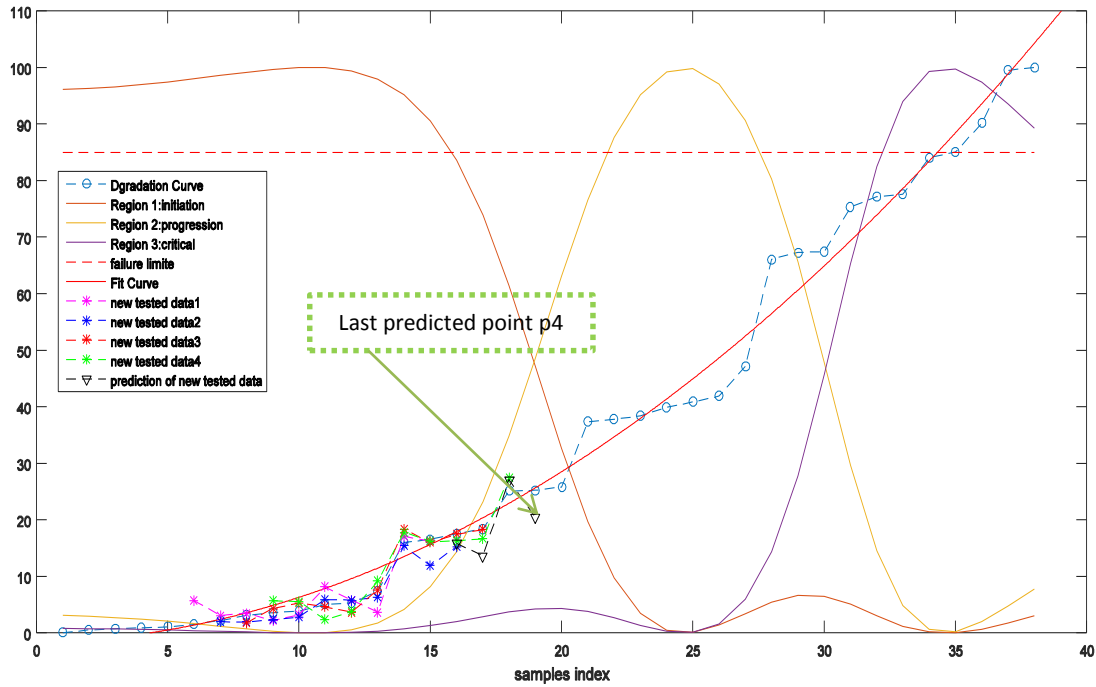


Figure 3.31: RUL Calculation (D8)

Figure 3.31 shows the results of the last step of the proposed methodology. When we apply it to the vibration signal acquired from the pump that has an impeller degraded over time. The blue circles dashed line represents the real degradation curve, the brick, yellow and purple lines are the plotting of the membership functions used in the fuzzy algorithm. The dashed red line is the failure limit chosen by the user, the red line is the fit curve of the degradation curve, the stars purple, blue, red and green lines are the new degraded data, chosen for the testing of this methodology. The black triangles line is the predicted values calculated for each new tested data, the result shown that FCM algorithm can successfully cluster the fitting curve (calculated before) on three groups, for example, if we take the last predicted point (p4) we see referring to UP matrix (eq.3.18), that the probability to classify p4 (Pp_i/c_j ; $j = 1,2,3$) on class1, class2 or class 3 are 0.47, 0.48 and 0.042 respectively, that indicate that p4 classified on the class2 ($Pp4/c2 > Pp4/c1 > Pp4/c3$), the $CG_4 = 28.4465$ and $RUL = 71.5\%$ that means the remaining Useful life of the pump impeller is equal 71.5%.

The machine should be stopped if the probability of the predicted value classified on the 3'rd class (critical class) or the value overtakes the failure threshold.



Normandie Université

$$Up = \begin{bmatrix} \text{class1} & p1 & p2 & p3 & p4 \\ \text{class2} & 0.8361 & 0.7390 & 0.6149 & 0.4721 \\ \text{class3} & 0.1438 & 0.2323 & 0.3481 & 0.4854 \\ \text{class3} & 0.0202 & 0.0287 & 0.0370 & 0.0425 \end{bmatrix} \quad (3.18)$$

The last column in Table 3.8 represents the RUL percentage on each degradation mode

One may say, why do not simply calculate the RUL via the coordinates of the predicted point, ($Rul = 100 - amplitude_{(predicted\ point)}$, $\rightarrow RUL = 100 - 22.22 = 77.78$), without recourse to clustering technique which is much simpler and easier.

However using this method we are considering that the predicted point belongs equally to the all classes (3 classes), in this case, there is no effect of the class weight on the RUL calculated. Also, we will not have any information about the degradation stage. But using our approach, we take into consideration the weight of each class in the predicted point, therefore the RUL will be dependent on the combinations of this weight and the class with the highest weight will influence the RUL results, and the results showed illustrate this issue very well.

On the other hand, all previous work looks to the RUL estimation as a probabilistic problem [221] [4] [227]. All these combined reasons were the main incentive to use this method.

Conclusion

An accurate prognostic of the remaining useful life of a system is critical to reducing the maintenance cost, ameliorate system reliability and improve the safety level that represents the first important term in the world. The subject becomes even more important when it comes to equipment that has a high risk and is very costly materially and humanly such as the reactors.

This section presented a methodology that aims to realize more accurate RUL prediction of equipment deterioration over time. In this work, the proposed ANFIS model has two different functions, the primary one is the creation of a model can be presenting the system deterioration which the features extraction techniques were unable to perform it also a strategy for early detection of faults. Secondly, the ANFIS model, take the previous inspection points of degradation values that they have a correlation with the age of this system as inputs, and the life percentage as the output.

The results shown in this research demonstrated that the proposed methodology, succeed to create a new technique that is very effective in determining the path of deterioration of the distillation column device and also predicts the future path of this device by determining their RUL. Faults prognostic is applied in real experimental data obtained from the distillation column during the automated continuous distillation process.

The results also confirmed the ability to classify between initiation, progression, and critical classes of the system lifetime percentage and an early detection of failure occur in distillation column process to prevent damage or catastrophic accidents.



3.15 Parzen Window Distributions as New Membership Function for ANFIS Algorithm- Application to a Distillation Column Faults Prediction.

3.15.1 Introduction

The deviation of the distillation column data from its normal path is a sign that the device is heading towards a malfunction or a problem; this is a dangerous indicator of system failure and damage in the distillation process. In such a case, losses will not only be material but there may also be casualties due to the significantly high pressure and the flammable materials in the device, possibly causing an unexpected explosion in the distillation column. Consequently, the data of this device must be carefully tracked in order to understand the future direction of data. This is known as the predictive or forecasting process [16]. The ability to construct a reliable and effective prediction system puts the device under cautious monitoring and supports industrial security, which is the most important issue for industrialists around the world [162]. When choosing the best forecasting techniques, one should take into account several factors such as real-time, desirable degrees of accuracy, time analysis, historical data relevance and availability, cost, complexity, operating conditions, heat transfer coefficient, reaction rate, reaction enthalpy, activation energy, and their unpredictable variations, information volume and nature, etc. [228][229][3].

The most important factor that affects the prediction accuracy of ANFIS is the type of Membership Function (MF) used on the first layer of ANFIS architecture (Figure 3.15). The execution time is also important for real-time processing [11]. In each instance, ANFIS uses the gradient descent algorithm to update the premises parameters of the MF. For example, if we have only two inputs, x and y , and each input has two MFs, supposing we use the Gaussian function as MF, which has two parameters (the mean (μ) and the variance (σ)), this means in total we have eight parameters that must be updated at every iteration. Therefore, if the number of the inputs is high, then the number of the parameters and execution time is also substantial. In the case that the Gaussian shape is applied as the MF, we have eight parameters while in the case of using the Pi-shape as the MF, the number of parameters will be 20. Therefore, a solution is needed to reduce the vast amount of parameters in order to minimize the calculation cost and keep real time processing. On the other hand, previous research has clarified that there is no a priori ideal shape that can be used as a standard for all applications. In addition, the quality, the nature of the data and its application determine the best MF that should be selected (Rui et al., 1995) [230][5].

In a comparative study in 2016, Ardhian et al. [11] suggested that the trapezoidal shape is the best MF that can be used for load forecasting. In 2011, Mayilvaganan et al. [12] also proved that the Gaussian shape showed significant results for the prediction of the groundwater level of a watershed. Additionally, in 2012, Singh et al. demonstrated that the Gaussian and the bell-shape are the best MFs for the estimation of the elastic constant of

rocks [13].

In the absence of any study applied to data extracted from the distillation column in search of the best forecasting technique, it is clearly necessary to do a comparative study between different types of membership functions in order to determine the optimal MF for forecasting the distillation process data.

A meaningful consideration for the best method in determining a new MF with a low number of parameters is the first incentive in proposing the Parzen windows distribution as a new membership function to be used on the first layer of ANFIS algorithm. As it is well known, only the *standard deviation* (h) can modify the Parzen shape form, and in this case, we have only one parameter that should be updated per iteration.

In this study of this section, we aim to determine the most advantageous type of MF which has the smallest root mean square errors (RMSE) among actual and forecasting data, when considering the execution time.

3.15.2 Proposing Parzen Windows as a New Membership Function for ANFIS Algorithm

3.15.2.1 Membership Function (MF)

As aforementioned, the premise parameters of the ANFIS algorithm are completely related to the membership functions (MFs) chosen at the beginning of the learning process [11].

If we consider X as an ensemble of x elements, then a fuzzy set A (eq.3.19) in X is defined as:

$$A = \{x, \mu_A(x) \mid x \in X\} \quad (3.19)$$

$\mu_A(x)$ is known as the degree of membership of x in A ($\mu \in [0, 1]$). This degree of membership is modified according to the form of the shape used. There are different types of membership functions such as the Gaussian shape (gaussmf), generalized bell shape (gbellmf), triangle shape (trimf), and trapezoid shape (Trapmf), etc.

Below is a brief explanation of each MF type used in this study:

1- *Triangular-Shaped Membership Function (Trimf)*

Triangular membership function (Trimf) (eq.3.20) is formed of three straight lines constituting a triangle. This MF is very simple and linear, it includes three parameters, a , b and c , which represent the 3 points of the triangle.



Normandie Université

$$f(x; a, b, c) = \left\{ \begin{array}{ll} 0, & x \ll a \\ \frac{x-a}{b-a} & a \ll x \ll b \\ \frac{b-x}{c-b} & b \ll x \ll c \\ 0, & x \gg c \end{array} \right\} \quad (3.20)$$

2- Trapezoidal-Shaped Membership Function (Trapmf)

The *trapezoidal* membership function (Trapmf) (eq.3.21) is formed of a truncated triangle curve with a flat top. It includes four parameters, a, b, c, and d, that form the shape of a trapezoid.

$$f(x; a, b, c, d) = \left\{ \begin{array}{ll} 0, & x \ll a \\ \frac{x-a}{b-a} & a \ll x \ll b \\ 1 & b \ll x \ll c \\ \frac{d-x}{d-c} & c \ll x < d \\ 0, & x \gg d \end{array} \right\} \quad (3.21)$$

The triangular and trapezoidal MFs, both formed of straight lines, have the advantage of simplicity.

3- Gaussian Membership Function (Gaussmf):

The *Gaussian* membership function (Gaussmf): (eq.3.22) has two parameters, the mean (a) and the variance (c). This MF can achieve smoothness, contrary to the triangle and trapezoid functions.

$$f(x; a, c) = e^{\frac{-(x-a)^2}{2c^2}} \quad (3.22)$$

4- Generalized Bell-Shaped Membership Function (Gbellmf)

$$f(x; a, c) = \frac{1}{1 + e^{-a(x-c)}} \quad (3.23)$$

The *Generalized Bell-Shaped Membership Function* (Gbellmf) (eq.3.23) depends on three parameters, a, b, and c, where b is generally positive. The parameter c is located at the center of the curve.



The smooth Gaussian and bell membership MFs are unable to achieve asymmetric membership functions, which is an important characteristic in certain applications.

5- Difference Between Two Sigmoidal Membership Functions (Dsigmf)

The Dsigmf (eq.3.24) is the difference between two sigmoidal membership functions and it has four parameters, a_1 , c_1 , a_2 , and c_2 . A_1 and c_1 are the variances and the mean of the first sigmoidal MF whereas a_2 and c_2 are the mean and the variance of the second sigmoidal MF

$$f(x, a_1, c_1, a_2, c_2) = \left(\frac{1}{1 + e^{-a_1(x-c_1)}} - \frac{1}{1 + e^{-a_2(x-c_2)}} \right) \quad (2.24)$$

6- Phi-Shaped Membership Function (Pimf)

The *Phi-shaped MF* (Pimf) (eq.3.25), which is formed of four parameters, is zero on both extremes with a rise in the middle.

$$f(x; a, b, c, d) = \begin{cases} 0, & x \ll a \\ 2 \left(\frac{x-a}{b-a} \right)^2 & a \ll x \ll \frac{a+b}{2} \\ 1 - 2 \left(\frac{x-b}{b-a} \right)^2 & \frac{a+b}{2} \ll x \ll b \\ 1 & b \ll x \ll c \\ 1 - 2 \left(\frac{x-c}{d-c} \right)^2 & c \ll x \ll \frac{c+d}{2} \\ 2 \left(\frac{x-d}{d-c} \right)^2 & \frac{c+d}{2} \ll x \ll d \\ 0, & x \gg d \end{cases} \quad (3.25)$$

7- Parzen Windows as a Membership Function

The Parzen window distribution approach was created by Emanuel Parzen in 1962. It defines an unknown probability density $p(x)$ based on a set of observations. This approach provides an accurate mathematical analysis and is used in different domains and applications, such as pattern recognition and classification [231], and image processing [232], etc. Parzen window density estimation is a data-interpolation method [231]. Let us consider that X is a random sample, and $P(X)$ (eq.3.26) is the PDF of this sample estimated by Parzen window.

The general expression of non-parametric density estimation is:

$$p(x) \cong \frac{k}{VN} \text{ where } \begin{cases} N & \text{volume surrounding } x \\ V & \text{total \# examples} \\ K & \text{\# examples inside } V \end{cases} \quad (3.26)$$

When we fix V and determine K from the data, this leads to **kernel density estimation** (KDE). If we assume that the region R that encloses the K examples is a hypercube with sides of length h centered at x , then its volume is given by $V=h.D$, where D is the number of dimensions.

It is necessary to introduce (eq.3.27):

$$\phi\left(\frac{x_i-x}{h}\right) = \begin{cases} 1 & \left|\frac{x_{ik}-x_k}{h}\right| \leq 1/2, \quad k = 1,2 \\ 0 & \text{otherwise} \end{cases} \quad (3.27)$$

This equation indicates the location of (x_i) whether it is inside the square (centered at x , with a width h) or not.

The total number of k samples falling within the region R , out of n , is given by (eq.3.28):

$$k = \sum_{i=1}^n \phi\left(\frac{x_i-x}{h}\right) \quad (3.28)$$

The Parzen probability density estimation formula (for 2-D) is given by (eq.3.29):

$$p(x) = \frac{k/n}{V} = \frac{1}{n} \sum_{i=1}^n \frac{1}{h^2} \phi\left(\frac{x_i-x}{h}\right) \quad (3.29)$$

The equation $\phi\left(\frac{x_i-x}{h}\right)$ is known as a window function. We can generalize the idea and allow the use of other window functions in order to yield other Parzen window density estimation methods. For example, if the Gaussian function is used, then for (1-D) we have (eq.3.30 and eq.3.31)

$$p(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i-x)^2}{2\sigma^2}\right) \quad (3.30)$$

$$p(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{(h\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2\right) \quad (3.31)$$

- In both equations, h is the **standard deviation** of the Gaussian PDF along each dimension.
- For a large sample size, the Parzen window estimate comes quite close to the Gaussian PDF.

The final Parzen equation used as MF for ANFIS is (eq.3.32):

$$f(x; h) = \exp\left(-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2\right) \quad (3.32)$$

Parzen window has many advantages. Firstly, it can be used for bimodal, unimodal and normal mixtures. It also does not need an assumption about the early time distribution. In addition, it has the ability to converge towards an arbitrarily complicated target density with sufficient samples. In theory, it converges when the number of samples reaches infinity and it can be applied to the data that is taken from any distribution [233].



Normandie Université

Usually, In Parzen distribution, the size of the chosen windows totally depends on the standard deviation (h), therefore, the only disadvantage of Parzen window is the difficulty to find the best value of h which is usually determined by trial without any general concept or rule. In our work, the ANFIS method eliminates this challenge because it uses the gradient descent algorithm to find the appropriate size of (h) window. Here we guarantee that ANFIS converges faster to the desired goal in the least amount of time, although it is not certain that results will always be better. The selection of the best MF completely depends on the application and the nature of the processed data.

3.15.3 Proposed Methodology

In this study, ANFIS was used as a predictor for the signals acquired from the distillation column system to estimate the future state of this system. This is shown in the following flowchart (Figure 3.32):

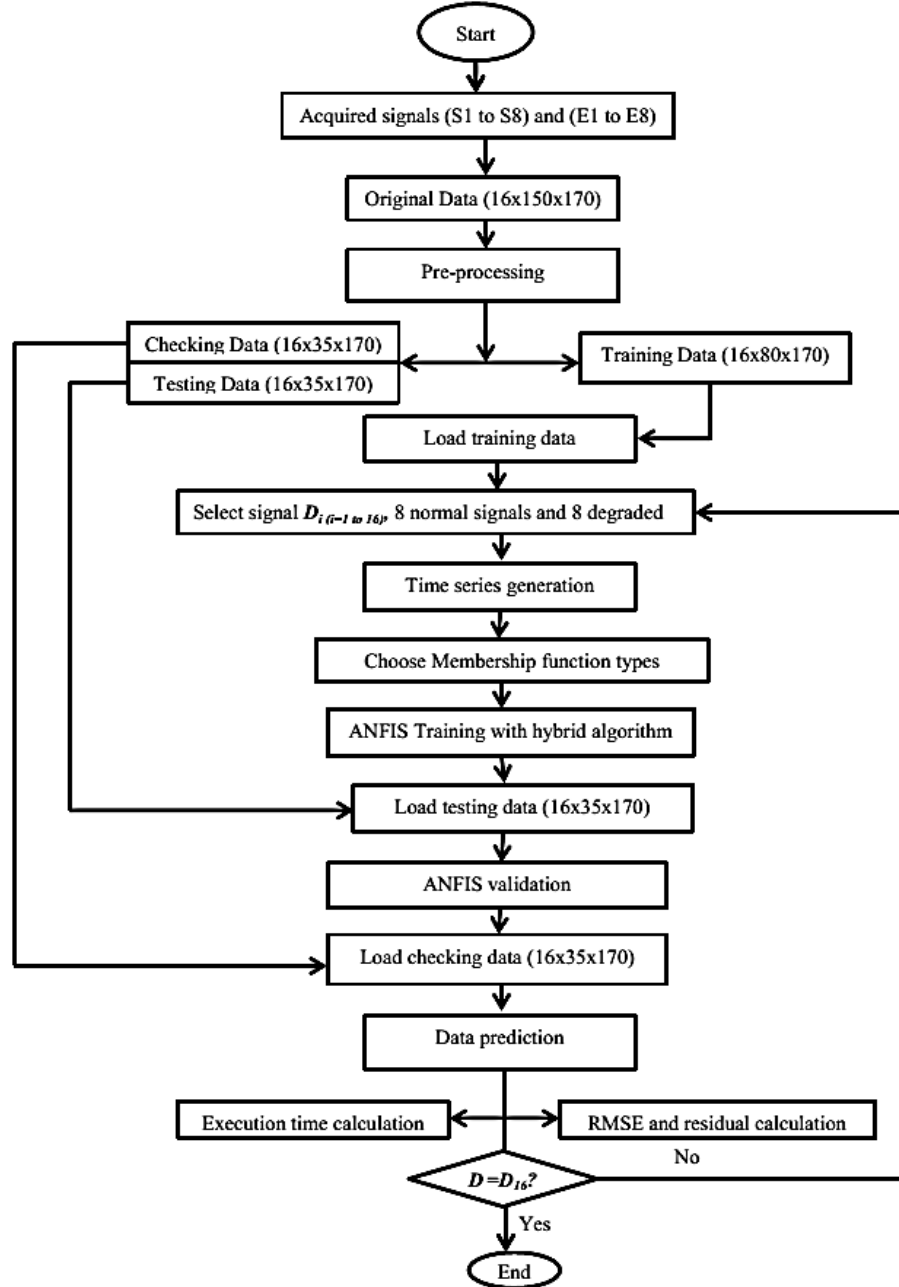


Figure 3.32: Flow Chart of Our Proposed Methodology

1- ANFIS Model

In this study and after many attempts to find the best ANFIS model, ANFIS has six-time series (Ts) vectors as inputs and one predicted vector as an output. Each input has two MFs formed of 24 rules of ANFIS shown inputs. These are the previous and the present distillation data; the output is the predicted distillation data.



The ANFIS model in this study consists of seven steps:

- Step 1: Pre-processing data such as filtering and smoothing
- Step 2: Load training, testing and checking data (150 observations, each observation consists of 170 samples).
- Step 3: Training ANFIS with 80 observations.
- Step 4: Testing ANFIS with 35 observations (validation step).
- Step 5: Checking ANFIS with 35 observations.
- Step 6: RMSE and execution time calculation
- Step 7: Residual calculation

Figure 3.33 represents the ANFIS architecture that was used in this study. Layer 0 is the six-time series vectors inputs that were calculated from the original signals acquired from the distillation column; from each acquired signal we calculate 6 times series vectors. Layer 1 is the membership functions that were applied (inputs MFs), layer 2 is the prod layer that produces outputs that have 64 rules, layer 3 is the normalization layer, and layer 4 calculated the 64 output functions (f_1, \dots, f_{64}) as the following:

Rule 1: if $TS1$ is $A1$ and $TS2$ is $A3$ and $TS3$ is $A5$ and $TS4$ is $A7$ and $TS5$ is $A9$ and $TS6$ is $A11$ then $f1 = TS1.A1 + TS2.A3 + TS3.A5 + TS4.A7 + TS5.A9 + TS6.A11$

Rule 2: if $TS1$ is $A2$ and $TS2$ is $A3$ and $TS3$ is $A5$ and $TS4$ is $A7$ and $TS5$ is $A9$ and $TS6$ is $A11$ then $f2 = TS1.A2 + TS2.A3 + TS3.A5 + TS4.A7 + TS5.A9 + TS6.A11$

Rule 3: if $TS1$ is $A1$ and $TS2$ is $A4$ and $TS3$ is $A5$ and $TS4$ is $A7$ and $TS5$ is $A9$ and $TS6$ is $A11$ then $f3 = TS1.A1 + TS2.A3 + TS3.A5 + TS4.A7 + TS5.A9 + TS6.A11$

⋮

Rule 64: if $TS1$ is $A2$ and $TS2$ is $A4$ and $TS3$ is $A6$ and $TS4$ is $A8$ and $TS5$ is $A10$ and $TS6$ is $A12$ then $f64 = TS1.A2 + TS2.A4 + TS3.A6 + TS4.A8 + TS5.A10 + TS6.A12$

Layer 5 is the ANFIS output that produces the summation of the f_i calculated in layer 4 where $i=1$ to 64.

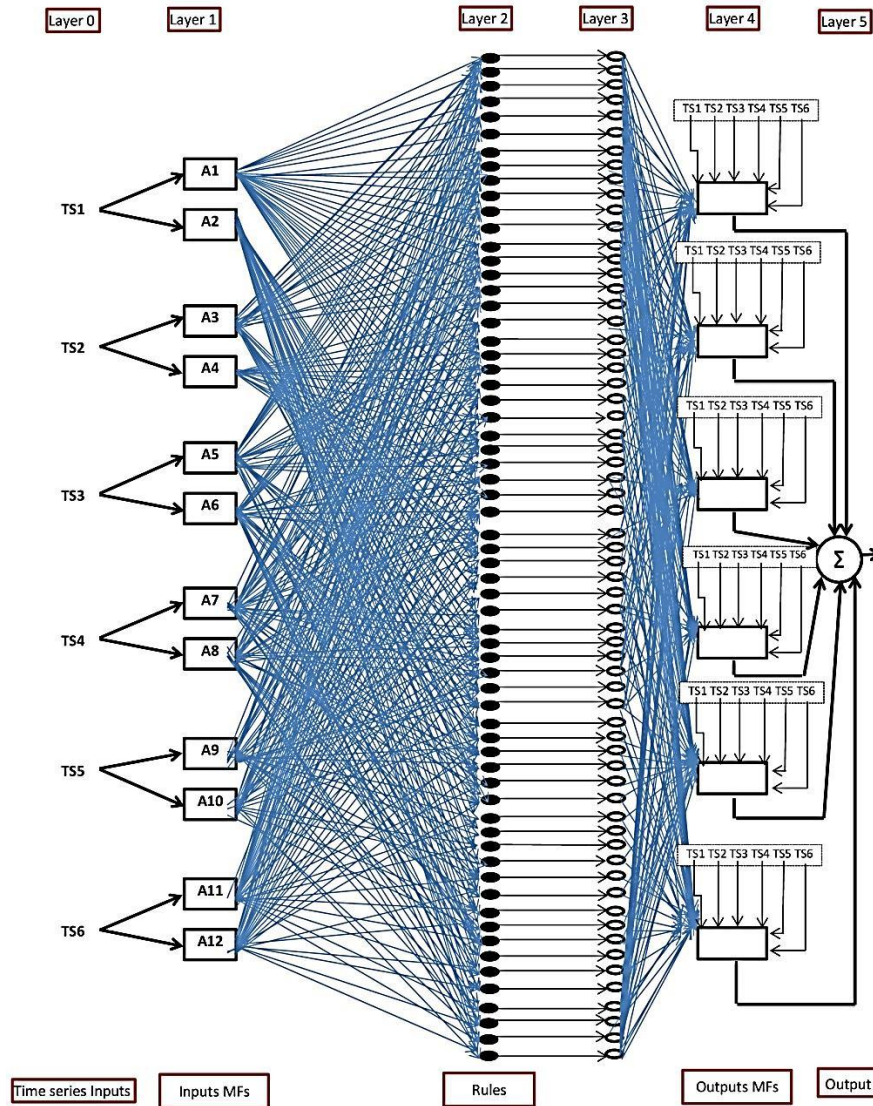


Figure 3.33: ANFIS Architecture that was used in this Study

2- Predictions residual

The difference between the target (T), that represents the real signal and the output of ANFIS (O), is called the residual (R), $R = |T - O|$. As long as the residual remains small, then this means that the prediction values are within the normal range. When the residual increases, this is an indication that the signal is starting to deviate from the normal state and it is heading towards an accident state. This is a relevant indicator of where the system is heading in the future.

3.15.4 Results and Discussion

In this section, we will demonstrate the real application for the purpose of prognostic. The methodology is proposed in order to predict eight normal signals (S1...S8) and eight degraded signals over time (E1...E8) (Table 3.9) as extracted from a distillation column, where each degraded signal Ei represents the deviation of this signal from the normal mode due to the failure of any components in the system or due bad control. For example, the signal E3 (decreasing degradation heating power from 100 to 0%) means that the boiler is failing from the other side; E2 means that there are no controls on the boiler (thermostat failure). It is important to note here that there is no relationship between Si and Ei in terms of i index.

- S1: Reflux rate (Timer);	-E1: increasing degradation reflux ratio from 0 to 100% (Timer);
- S2: Heating power;	-E2: increasing degradation heating power from 0 to 100%;
- S3: Feed flow rate;	-E3: decreasing degradation heating power from 100 to 0%;
- S4: Preheated Power;	-E4: increasing degradation preheated power from 0 to 85 %;
- S5: Loss of charge;	-E5: decreasing degradation preheated power from 100 to 0%;
- S6: Pre-heated temperature;	-E6: increasing degradation feed rate from 0 to 100%.
- S7: Boiler temperature;	-E7: decreasing degradation feed rate from 100 to 0%;
- S8:TIC2: Column head temperature;	-E8: pump impeller degradation.

Table 3.9: Normal and Degraded Signals

Previously, research on forecasting was applied to the distillation column system, but it was conducted for objectives other than those set out in this research. For instance, in 2014, *Sahraie et al* [234] used ANFIS to estimate the composition of the distillate product. In 2010, *Sivakumar et al* [226] used ANFIS to control the composition of the inlets distillation products. However, the risk that sudden faults in the reactors could be disastrous, deadly and very expensive remains. Therefore, it is necessary to have previous knowledge about the future state of the system and be sure that the system is still working within the normal range. As aforementioned, what distinguishes one method from another is its ability to achieve quality results within a very short time. Time is especially important when the application must be in real time, as is the case in this study, where ANFIS processes data for one of the most dangerous systems: the chemical reactor.

Despite the high risk of catastrophic problems in the system, it is not equipped with any prediction algorithm that can provide insight into potentially serious faults that may be fatal. Hence, this demonstrates the importance of strengthening the system by implementing this algorithm on a computer that processes the data acquired from the distillation column. Depending on the Matlab real-time application software, the algorithm receives the data in real time then activates an actuator (alarm, warning lamp, etc.) or triggers a control mechanism on the system (increases the feed flow rate, turns off the heater, etc.) whenever any of the input signals deviate from their normal mode to a faulty mode, depending on the predicted data.

The proposition to use Parzen window distribution as a membership function for the ANFIS method was based on a well thought out plan; the goal is to significantly reduce the numbers of



ANFIS parameters closer to real-time application. As demonstrated earlier, if the Gaussian shape as MF is applied, in every instance, ANFIS should update 8 parameters. Knowing that the Gaussian shape is controlled by 2 parameters only (means and variance), how will the case be when we use the trapezoidal-shape that is controlled by 4 parameters or when we use the Pi-shaped which has 5 parameters. When we use Parzen windows as MF, we have only 1 parameter (h) that has the right to control the shape. Therefore, when we use Parzen window, if we have the same number of inputs and rules, the number of parameters will be reduced by half, in comparison to using the Gaussian shape. As long as the number of parameters is reduced, the consumption time will be automatically reduced as well.

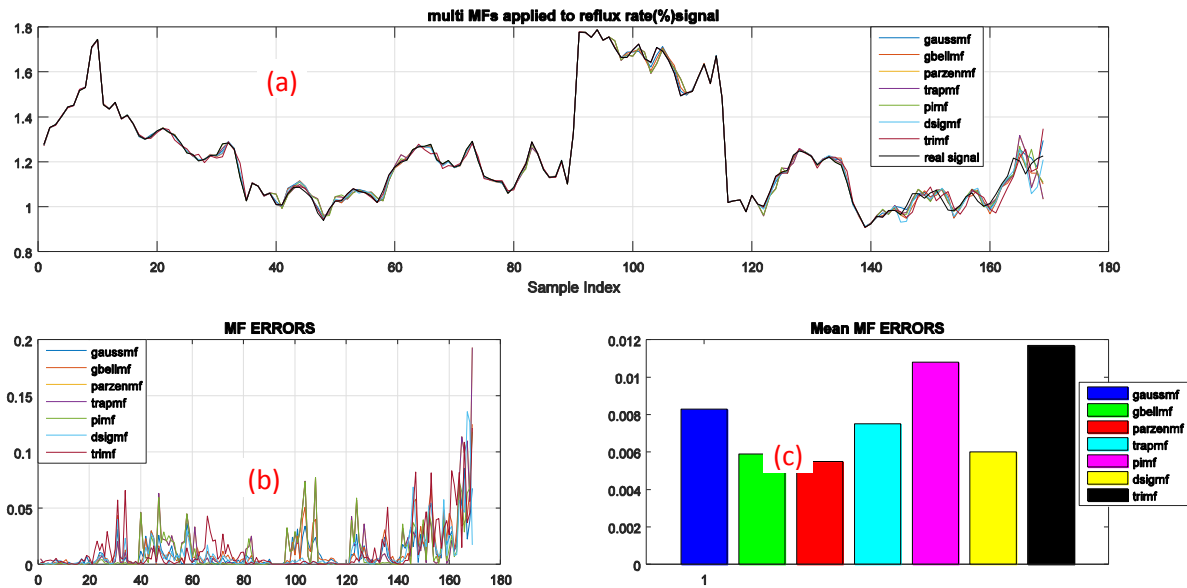


Figure 3.34.a: Graphical Representation of ANFIS Results Applied to Distillation Column Data (Reflux Rate Signal) with Different MFs of ANFIS

Figure 3.34.b: Graphical Representation of Residual Errors for Different MFs

Figure 3.34.c: Graphical Representation of Mean Errors for Different MFs

The black curve in Figure 3.34.a is the real signal and the other colored curves are the outputs of ANFIS, with different types of shapes used as MFs of ANFIS. The x-axis is the sample index and the y-axis is the reflux rate signal amplitude (S4).

Figure 3.34.b is a graphical representation of the residual errors between the real signal and the output of ANFIS for different types of MFs. The x-axis represents the sample index and the y-axis represents the error amplitude.

The results shown in Figure 3.34.a and Figure 3.34.b demonstrate that ANFIS has the ability to predict the data of the distillation column with a higher degree of accuracy. Comparing the results between the different types of MFs, Figure 3.34.c shows that Parzen window has the smallest RMSE (0.005) when ANFIS is applied to the reflux rate signal. As mentioned before,



our study is applied to 8 kinds of normal signals (S1...S8) various times (90 observations), and to 8 degraded signals (60 observations). For the normal signals, all the results are presented in Table 3.10.



Figure 3.35: Graphical Representation of Consumption Time in Seconds for Different MFs (ANFIS Applied to Normal Reflux Rate Signal)

Figure 3.35 is a bar representation of the consumption time of ANFIS for different types of MFs. The results have shown that Parzen window has the shortest time (14 sec) among the other MFs when ANFIS is applied to the normal reflux rate signal. All results of normal signals are listed in Table 3.10.



MF signal	Gaussian			Bell-shape			Trapezoidal		
	Average RMSE	Average Error St.D	Average Execution time	Average RMSE	Average Error St.D	Execution time	Average RMSE	Average Error St.D	Average Execution time
S1	0.0162	0.0395	15.117	0.0220	0.0474	20.510	0.0222	0.0464	27.832
S2	0.0137	0.0849	12.9889	0.0101	0.0283	20.0159	0.0122	0.0395	30.4860
S3	0.0235	0.1015	13.8395	0.0227	0.0536	19.9751	0.0233	0.0697	27.6670
S4	0.0120	0.0523	14.6847	0.0133	0.0632	19.9014	0.0127	0.0872	26.9641
S5	0.0099	0.0308	15.0306	0.0095	0.0714	19.1836	0.0092	0.0741	25.6672
S6	0.0121	0.0415	13.4683	0.0130	0.0457	23.8216	0.0134	0.0430	29.1184
S7	0.0091	0.0928	15.9022	0.0098	0.0455	17.7548	0.0089	0.0225	27.0609
S8	0.0117	0.0310	15.0399	0.0138	0.0394	20.8412	0.0136	0.0631	29.9868
MF signal	Parzen			Pi-shaped			Triangle-shape		
	Average RMSE	Average Error St.D	Average Execution time	Average RMSE	Average Error St.D	Execution time	Average RMSE	Average Error St.D	Average Execution time
S1	0.0222	0.0264	12.321	0.0249	0.0678	37.487	0.0232	0.0441	23.651
S2	0.0148	0.0435	10.8700	0.0125	0.0649	37.8790	0.0138	0.0884	21.0721
S3	0.0136	0.0749	13.3244	0.0130	0.0436	40.0144	0.0198	0.0305	25.8612
S4	0.0082	0.0553	12.0383	0.0096	0.0768	38.6691	0.0141	0.0684	27.6624
S5	0.0090	0.0256	8.8189	0.0143	0.0485	41.0849	0.0123	0.0814	24.1958
S6	0.0111	0.0594	12.6473	0.0248	0.0633	37.9413	0.0081	0.0120	23.5139
S7	0.0148	0.0466	13.9371	0.0113	0.0656	38.3064	0.0106	0.0569	25.5321
S8	0.0115	0.0548	14.4828	0.0131	0.1180	38.2318	0.0095	0.0623	20.2782
MF Signal	Dsigmf								
	Average RMSE	Average Error St.D	Average Execution time						
S1	0.0249	0.0551	33.225						
S2	0.0104	0.0439	37.3589						
S3	0.0154	0.0518	31.9551						
S4	0.0120	0.0509	30.6536						
S5	0.0146	0.0519	35.6401						
S6	0.0098	0.0423	34.5521						
S7	0.0122	0.0728	32.6258						
S8	0.0141	0.0529	35.0376						

Table 3.10: Statistical Characteristics of MFs of Acquired Signals in Normal Mode

Table 3.10 presents the variation of RMSE, the standard deviation and the execution time when ANFIS is applied to 90 observations of acquired normal signals with different types of MFs. The Parzen window proved its worth when chosen as the best MF (smallest RMSE) for three out of eight kinds of signals in normal mode (S1, S4, and S5). The results in Table 3.10 also affirm that Parzen window requires the shortest execution time for all signals in normal mode when compared to all other MFs.

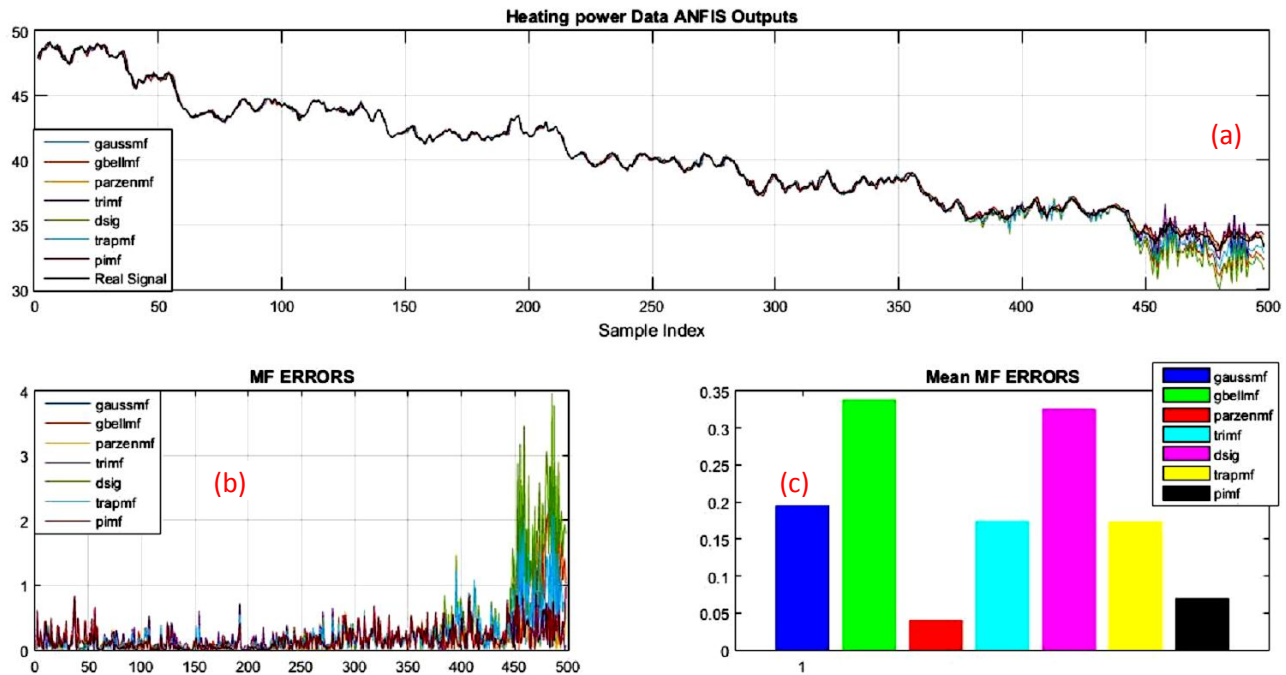


Figure 3.36.a: Graphical Representation of ANFIS Results Applied To Distillation Column Data (Heating Power Degraded Signal) with Different MFs of ANFIS

Figure 3.36.b: Graphical Representation of Residual Errors for Different MFs

Figure 3.36.c: Graphical Representation of Mean Errors for Different MFs

In this phase, ANFIS with different types of MFs is applied to degraded signals acquired from the distillation column. The black curve in Figure 3.36.a is the real signal and the other colored curves are the outputs of ANFIS, with different types of shapes used as MFs. The x-axis represents the sample index and the y-axis represents the degraded heating power signal amplitude (E4). Figure 3.36.b is a graphical representation of the residual errors between the real signal and the output of ANFIS for different types of MFs. The x-axis represents the sample index and the y-axis represents the error amplitude.

The results are shown in Figures 3.36.a and Figure 3.36.b demonstrate that ANFIS has a higher ability to predict the data of the power heating degraded signal that is acquired from the distillation column with higher accuracy. A comparison of the results between different types of MFs, shown in Figure 3.36.c, affirms that Parzen window has the smallest RMSE (0.4) when ANFIS is applied to the degraded heating power signal. All the results of degraded signals are presented in Table 3.11.

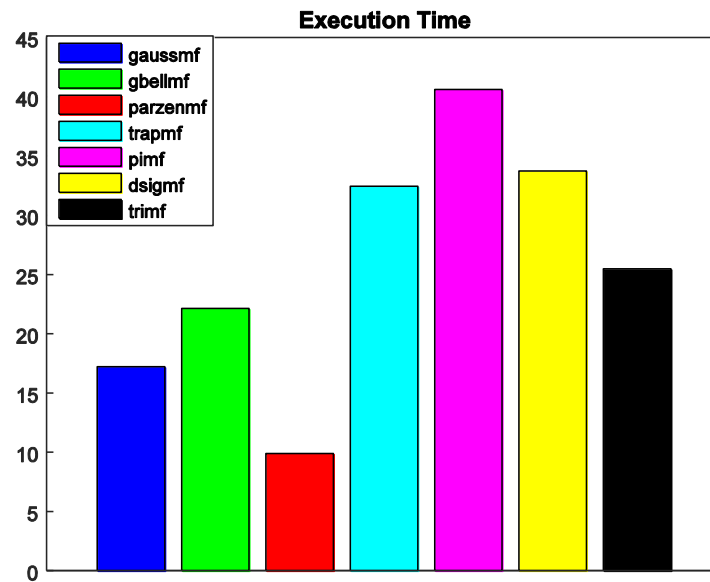


Figure 3.37: Graphical Representation of Consumption Time in Seconds for Different MFs (ANFIS Applied to Degraded Heating Power Signal)

Figure 3.37 is a bar representation of the time consumption of ANFIS for different types of MFs. The results show that Parzen window has the smallest time (19.2 sec) among the other MFs when ANFIS is applied to the degraded heating power signal. All the results of degraded signals can be seen in Table 3.11.



MF signal	Gaussian			Bell-shape			Trapezoidal		
	Average RMSE	Average Error St.D	Average Execution time	Average RMSE	Average Error St.D	Execution time	Average RMSE	Average Error St.D	Average Execution time
E1	2.0044	1.3851	13.4914	1.3806	1.2045	23.1888	1.0256	0.7589	26.5714
E2	2.4660	2.0338	15.0845	1.7744	2.3341	26.1951	1.3414	0.9934	26.6337
E3	2.0097	0.7695	13.6155	1.5764	1.2566	23.2522	0.6652	0.6586	24.8378
E4	1.6867	1.4606	16.3652	1.9909	1.4516	22.4751	0.7020	0.8897	28.2194
E5	0.9554	0.7837	15.1001	0.5308	2.8527	25.4225	0.6256	0.7703	26.8085
E6	1.6429	1.1095	12.6073	1.1804	2.4551	18.9890	1.1180	0.7933	29.0016
E7	0.6998	0.7594	16.6406	2.2485	1.8054	23.2660	0.9980	0.8170	28.3545
E8	2.3637	1.2657	14.2806	1.7637	0.6839	25.7442	1.0067	0.9092	28.7257
MF signal	Parzen			Pi-shaped			Triangle-shape		
	Average RMSE	Average Error St.D	Average Execution time	Average RMSE	Average Error St.D	Execution time	Average RMSE	Average Error St.D	Average Execution time
E1	0.9286	0.4324	6.9647	0.9854	0.6519	33.0954	2.3717	0.5549	20.1457
E2	1.3058	0.5442	7.1437	0.7920	0.3647	35.9886	2.9292	0.4124	21.9674
E3	0.4053	0.5009	5.4380	0.9157	0.5068	30.3677	2.8574	0.5450	20.3074
E4	1.0014	0.6713	8.7098	1.0860	1.2451	39.4668	1.5101	0.9993	24.2025
E5	0.2725	0.7782	6.3140	1.0864	1.0315	32.4521	3.1522	0.8685	16.6878
E6	0.6483	0.8076	4.6545	1.1910	0.8590	37.7203	2.5736	0.8928	16.6014
E7	1.1095	0.8235	6.9108	0.9868	0.9794	32.7332	3.4891	0.8112	19.7368
E8	0.8372	0.5099	6.6788	1.0680	0.7482	32.2240	1.6825	0.8232	20.1531
MF Signal	Dsigmf								
	Average RMSE	Average Error St.D	Average Execution time						
E1	0.9842	0.5410	24.2345						
E2	0.6252	0.6639	18.5393						
E3	1.2711	0.8525	24.8805						
E4	0.3609	0.2426	25.2238						
E5	0.5968	0.6030	23.2395						
E6	2.2000	0.9078	23.0783						
E7	2.2963	1.0607	23.5643						
E8	1.0503	0.5497	24.3967						

Table 3.11: Statistical Characteristics of MFs of Acquired Signals in Abnormal Mode

Table 3.11 presents the variation of RMSE, the standard deviation and the execution time when ANFIS is applied to 60 observations (each observation represents 8 signals) of degraded acquired signals with different types of MFs. In this table as well, it is observable that Parzen window has proven to be the best MF (smallest RMSE) for 5 kinds of signals out 8 degraded signals (E1, E3, E4, E5, and E6). Again, Parzen window requires the shortest execution time for all signals in degraded mode when compared with all other MFs.



Parzen MF was not the best MF nor did it have the smallest RMSE for all signals in normal and degraded mode, and this affirms, as seen in Table 3.10 and Table 3.11, that results are really congruent with those of the previous research (Rui et al., 1995) asserting that only the comparative study is able to determine the best MF for each acquired signal.

3.15.5 Conclusion

The complementarity of the traditional preventive and curative maintenance by a more reactive and proactive orientation maintenance in operational condition of the industrial systems at lower cost offering a better performance and competitiveness to the companies.

The prognostic methods presented above do not constitute a complete state of the art of the existing methods. Only, the described techniques are the most known and the most used. The prognostic, when considered as a key process in maintenance strategies, such as CBM, proves to be a very promising activity. Prognostic activity is supported by model-based approaches, data-driven approaches, and experiential approaches to the diversity of methods and tools available. The knowing that no prognostic approach is universal and that the choice of an appropriate technique depends on conventional constraints limiting the applicability of tools: availability of data and/or knowledge and/or experiences, dynamics and complexity of real systems, constraints implementation (accuracy, calculation time, etc.), measurement possibilities (sensors, system, etc.).

Forecasting the state of a system has an important role at various levels, including the minimization of the cost of maintenance, the reduction or prevention of sudden accidents, the increase of safety and performance, all of which are considered as the primary objectives of industrialists and civil society actors alike. The relevance of forecasting and its accuracy increases when discussing a system whose failure leads to human and material disasters, such as the case of reactors. Although there are many forecasting techniques, this study aims to use ANFIS as a prediction method applied to real data that is acquired from a distillation column. The type of MF chosen on layer 1 of ANFIS can improve the performance of this method in terms of reducing the execution time and minimizing the RMSE. Therefore, in this research we proposed Parzen window as a new membership function of ANFIS algorithm, then we moved onto a comparative study that aims to choose the best MF used by ANFIS when it is applied to the normal and degraded data of the distillation column.

The results obtained in this study demonstrated that Parzen window proved its valuable capacity as a new membership function of ANFIS algorithm when it is applied to the distillation column data and it also proved to be successful in reducing the execution time of ANFIS.

The results have also shown that Parzen MF is chosen as the best MF for three out of eight types of normal signals and for five out of eight degraded signals.



Normandie Université

Finally, to prevent the limitation of ANFIS, reaching a compromise must be found between simplicity and generalization. From our point of view, the prediction is still an art and many of the techniques that we have learned out of experience can prove to be very helpful to others



Normandie Université

Chapter 4

Soft Computing Fault Prediction and Diagnosis Algorithm- Application to a Distillation Column

4.1 Introduction

The reliability and dependability of an industrial system, at a lower cost, are the main objectives of industrial enterprises to remain competitive in an ever-growing market. This interest is fueled by the fact that an unplanned shutdown can have very serious economic consequences in key sectors.

As mentioned before the high costs of maintenance highlight the increasing importance of monitoring the state of the systems in the industry through maintenance. The objectives mentioned above can be achieved through the implementation of an adequate maintenance strategy. As a result, there is an urgent need to continually develop and improve intelligent maintenance strategies to identify service requirements, optimize maintenance actions, and prevent unplanned downtime [235]

An effective health monitoring technique must be adapted to determine the state of the system at all times. A diagnostic method determines the current state of the system and identifies the probable causes (interaction with the environment, faults, etc.) that can lead to this state by reasoning on the observations. A prognostic method uses the current mission plan and knowledge of system degradation to anticipate abnormal behaviors or faults and thus predict future states.

The prognostic is generally associated with the end-of-life (EOL) prediction of a system in service, when the system is no longer operational, or its residual life (RUL). The diagnosis and the prognostic thus make it possible to have a report on the current health of the system as well as a prediction of the evolution of its state in the future. This information is used to reconfigure the system and update the mission and maintenance plans, hence the importance of diagnostic quality and prognosis [236].

This chapter will present a new proposed approach which can be applied to obtain real-time monitoring system which relies on the fault prediction module to reach the diagnosis module in contrast to the previous strategies; this means this method predict the future state of the system then diagnosis what is the probable fault source.

The Adaptive Neuro-Fuzzy Inference System (ANFIS), as a hybrid system, has been selected for the step related to prediction since it combines the advantages of fuzzy logic and ANNs in one simultaneous algorithm. In this research, we tested this methodology with real experimental data that was obtained from a real distillation column. This experimental data was acquired, however, after pre-processing was conducted in order to filter and smooth the signals. This resulted in the analysis of a database with different types of faults that could potentially occur during the automated distillation process.

4.2 problem formulation

If we go back to the maintenance strategy detailed in chapter 1 you can see that all the previous researchers rely on the diagnosis module to build or to reach a fault predictive module, where the diagnosis module ensures the failure data that represent the inputs for this predictive module as you see in Figure 4.1. One of the main disadvantages of this method is that it depends on previous failures; this means that we are supposed to incur the consequences of these faults, which are sometimes disastrous.

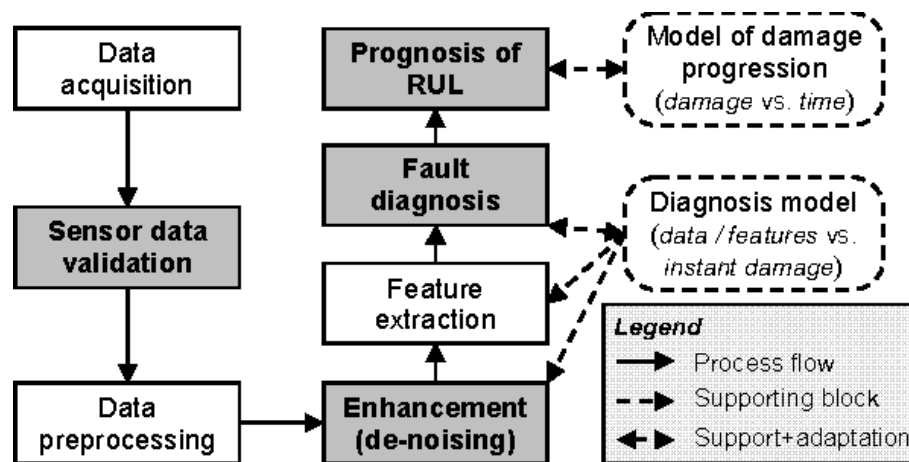


Figure 4.1: From Fault Diagnosis to Fault Prognostic

This lays out the undisputable justification for industrialists' need to examine levels of severity of faults within their systems and predict the optimal moment for intervention, and even stoppage of the instrument. This is better known as the prediction and diagnosis process [27].

Conditional preventive maintenance requires a predictive approach. It provides the maintenance personnel an indication of the future state of the system and, ideally, gives sufficient time for staff, equipment, and spare parts are organized, minimizing downtime and maintenance costs [237][238] Figure 4.2 shows the steps of a process for dealing with a possible fault at a system level. This system is considered to be in working order at the beginning, then, after a while, an incipient fault develops in the system.

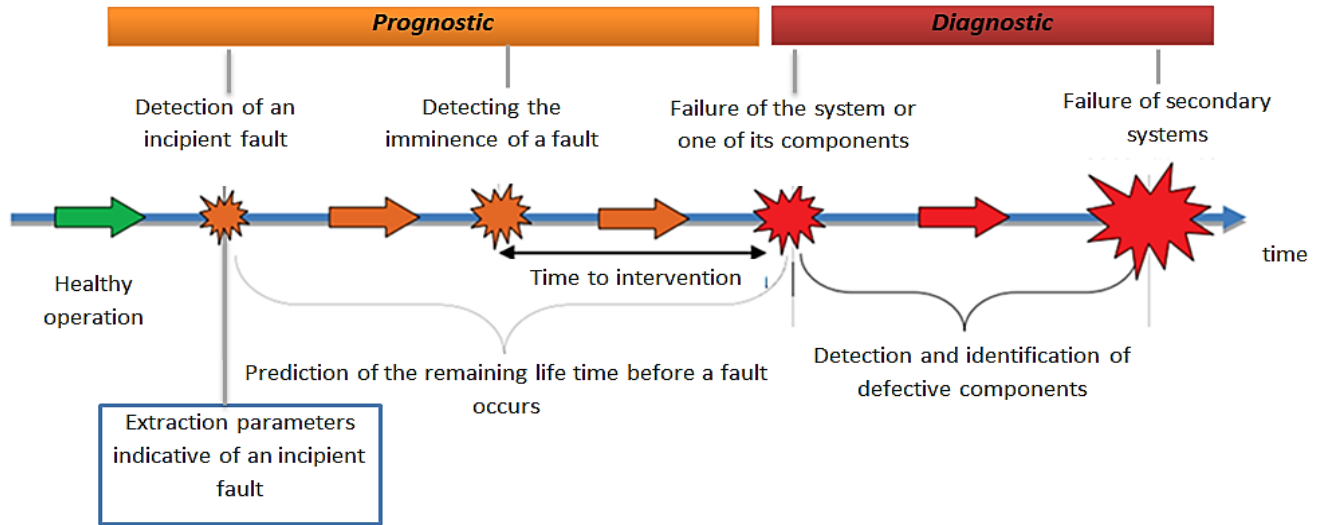


Figure 4.2: From fault prognostic to fault diagnostic

Over time, the severity of the fault increases until the system is completely degraded. If the system is allowed to continue functioning, there is a possibility to see other faults.

The diagnostic process typically occurs when a fault occurs and/or in the interval between system failure and failure of secondary systems.

However, if an incipient fault can be detected at an early stage, then maintenance operations may be delayed until the state of the system changes to a more degraded state. This interval, between the detection of an incipient fault and its appearance, defines the (temporal) domain of the prognosis. Providing a sufficient interval, commonly referred to as the remaining lifetime, between the detection of the incipient fault and the system failure, allows the system to be operated better and the maintenance operations to be reduced further.

To take advantage of the benefits of prognosis, maintenance personnel must:

- Have techniques to detect and identify an emerging defect;
- Have a reliable estimate of the time remaining before the intervention, i.e. the time remaining before the appearance of a fault.

4.3 Data acquisition

4.3.1 Signal acquisition

The signal acquisition allows the monitoring of changes in parameters such as differential pressure or temperature at a given point of the distillation column. The signals obtained during each acquisition represent (Table 4.1):

<i>D1: Feed flow rate,</i>	<i>D5: Heating power,</i>
<i>D2: Loss of charge,</i>	<i>D6: Pre-heated temperature,</i>
<i>D3:Timer: reflux rate</i>	<i>D7: TIC2: Column head temperature.</i>
<i>D4: Preheated Power</i>	<i>D8: Boiler temperature</i>

Table 4.1: The Acquired Signals

Normal mode signals (Table 4.2) are characterized by the following parameters: feed flow rate set at 80% of its capacity, pressure drop at 0.7 mbar, preheated temperature of 40°C, boiler temperature at 76°C, column head temperature at 56°C.

Signal	Min	Mean	Max	Variance	Frequency peak
Preheated Temp (°C)	39.6	40	40.6	0.006	0.0006
Timer %	0	0.5	7	1.6	0.001
Charge loss (mbar)	0.7	0.7	0.7	0	0.0001
Heating power	42.3	42.7	42.9	0.009	0.0001
Flow %	80	80	80	0	0.0001
Temp heater (°C)	75.8	76	76.1	0.009	0.0001
Head temp TIC2 (°C)	55.9	56	56.1	0.001	0.0001

Table 4.2: Statistical Characteristics of Acquired Signals in Normal Mode

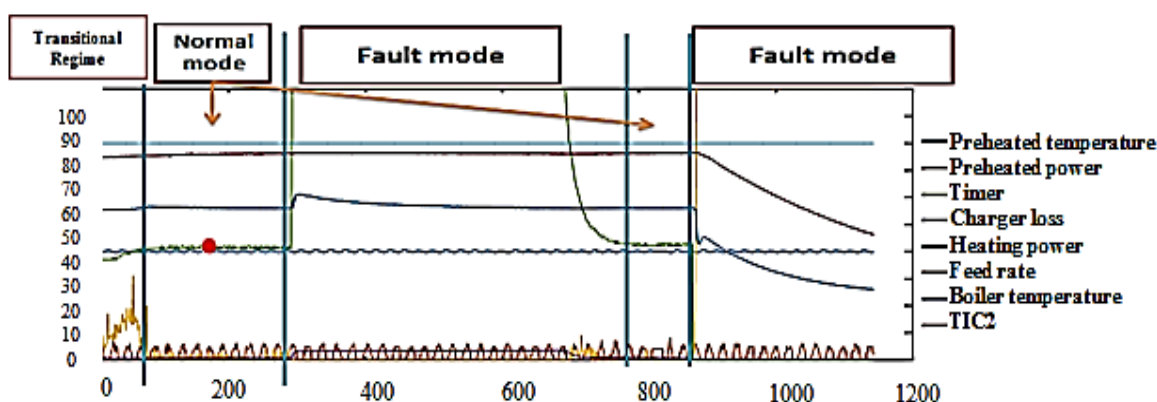


Figure 4.3: Graphical Representation of Signals (Normal & Faults Mode)

While accidents automated the continuous distillation that occurs in the industry, the cause was variations of the parameters. The most common variation that was observed, as illustrated in Figure 4.3, where: 1- the reflux rate, 2- the heating power, 3- the preheating power and 4- the feed rate.



A database with normal and faulty observations is analyzed. The database is composed of 50 observations; each observation has 2507 points with the sampling frequency, $SF=1\text{sample}/11\text{sec}$.

4.3.2 Data Pre-processing

In general, the data acquired from the machine are noisy and redundant. Therefore, this data cannot be applied directly by a diagnosis or prognostic model. Further, this original data could be hidden inside many of the relevant information which can denote to the machine fault or machine degradation. Wherefore a set of parameters that plain of a relevant information should be extracted from this original data as indicators for this fault or this degradation. Usually, the strength of the diagnosis or prognostic model depends on the quality of the extracted and selected features. In addition, it is very important to identify the features that reflect the type of a fault and the progression of the failure of the machine [45][46].

A pattern is an observation made about the process. It is characterized by a set of d parameters (or features), and represented by a point in the dimension space d , defined by the different parameters called representation space. Since the parameters are often real numbers, a form ' i ' can be defined by a vector $X_i = [xi1, xi2, \dots, xid]$ called form vector.

If we place the problem in the context of the diagnosis, the parameters of the pattern vector reflect the state of the studied system. They come from analyzes performed on the signals measured by the sensors installed on the system (vibrations, speed, currents or even voltages for example).

The typical patterns (or prototypes) are representative points of this space, and the problem of recognition consists in associating an observed pattern with a known standard pattern.

Due to disturbances (measurement noise, sensor accuracy ...), a new observation will rarely be identical to one of the prototypes. Thus, in order to express the influence of noise, the classes $(\omega_1, \omega_2, \dots, \omega_c, \dots, \omega_M)$ correspond to zones in space, grouping the similar patterns.

The principle of recognition is to know which class, among M known classes, to associate a new form, $X_i = [xi1, xi2, \dots, xid]$ observed.

In terms of diagnosis, the classes correspond to the known modes of operation. They are our initial data set, called learning set and noted \mathbf{Xa} . To classify a new observation is to identify one of these modes.

The development of a diagnostic system based on the neural network takes place in three phases: a perception phase, an analysis phase, and exploitation phase.

The perception phase is the main source of information about the system. It is not only reserved for pattern recognition because it is common to other diagnostic approaches. It consists of two stages. A data acquisition step which consists in determining the hardware configuration (the type, the number of sensors to be used and the sampling rate, etc.) that they are necessary for the collection of signals on the studied system. The acquired signals must provide useful



information in order to judge the operating state in which the system is located. This first step is followed by a signal preprocessing phase (filtering, de-noising, etc.).

The analysis phase is to study the information provided by the sensors installed on the system. If the information is in the form of signals, then it is necessary to extract features (or parameters). These parameters, which moreover constitute the pattern vector, must be able to describe the behavior of the system.

In the exploitation phase, the diagnostic system based on ANN can be commissioned. It makes it possible to classify each new observation collected on the system in one of the known classes, by applying the decision rule developed in the analysis phase. The determination of this class makes it possible to know the current mode of operation of the system.

4.3.3 Signal processing

Signal processing combines a set of techniques for creating, analyzing and transforming input signals to extract fault-indicating parameters [239]. Signal processing techniques classified into Time Analysis, Frequency Analysis, and Time-Frequency Analysis.

1- Temporal Features

The temporal analysis takes into account:

- The time required to acquire the information,
- The delay of diagnosis and treatment of information,
- The decision period,
- The reaction time of the system to the resulting action

a- Direct Measures:

- $\text{Min}(X)$: It represents the minimum of X.
- $\text{Max}(X)$: It represents the maximum of X.

b- Statistical measures:

They represent these values:

- Mean value(μ) (eq.4.1): is the statistician's jargon for the average value of the signal X.

$$\mu = \frac{1}{n} \sum_{i=0}^{n-1} x_i \quad (4.1)$$

- Quadratic mean value (\bar{X}) (eq.4.2): The quadratic mean of a set of numbers is the square root of the arithmetic mean of the squares of these numbers.

$$\bar{X} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} x_i^2} \quad (4.2)$$

- Variance: In probability theory and in statistics, the variance (eq.4.3) is the expectation of the quadratic deviation of a random variable from its mean.



Normandie Université

$$\sigma^2 = \frac{1}{n-1} \sqrt{\sum_{i=0}^{n-1} (x_i - \mu)^2} \quad (4.3)$$

- **Standard Deviation:** $SD = \sigma = \sqrt{\sigma^2}$

c- Power Measurements

- Power of the signal (eq.4.4):

$$P = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{i=-n}^n |x(i)|^2 \quad (4.4)$$

- Root Mean Square (eq. 4.5):

$$RMS = \sqrt{P} \quad (4.5)$$

- d- The Crest Factor:** The crest factor (eq. 4.6) is a characteristic measure of a signal. This is the ratio between the amplitude of the peak of the signal and the RMS value of the signal. This factor is independent of the operating conditions. It decreases when faults develop. It is commonly correlated with the Peak-to-Average Power Ratio (PAPR) which indicates a ratio between peak power and average power [240]:

$$C = \frac{|X|_{peak}}{X_{RMS}} \quad (4.6)$$

- e- The Envelope Analysis:** Fault diagnosis at an early stage, it can be determined reliably and quickly the shock repetition frequencies. We can look for the average and the variance of the envelope (eq.4.7) [54].

$$\text{env}(t) = \sqrt{|X(t)|^2 + |\tilde{X}(t)|^2} \quad (4.7)$$

Where, $\tilde{X}(t) = \frac{1}{\pi t} \times X(t)$

1- Frequency Features

Frequency features extraction is based on the Fourier transform. Frequency analysis makes it possible to locate faults and to make a reliable diagnosis. Likewise, it does not require any additional measurements. We define the spectral moment of a signal by the following formula (eq.4.8)

$$Mr^* = 2 \int_0^\infty (f - MPF)^r S_x(f) df \quad (4.8)$$

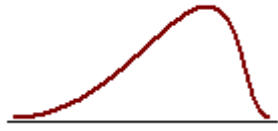


Where $S_x(f)$ it is the spectral density of the signal.

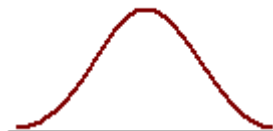
Frequency features include:

- a- **Signal Energy:** It represents the distribution of the energy **M0** (moment of order 0) of the signal on the frequency axis.
- b- **Mean Power Frequency (MPF):** $MPF = \frac{M_1}{M_0}$
- c- **Skewness (CD):** The Skewness coefficient (eq.4.9) measures the degree of asymmetry of the distribution. It is defined as the third order moment centered on the standard deviation cube. If CD is equal to 0, the distribution is symmetrical. If CD is smaller than 0, the distribution is asymmetric to the left. If CD is greater than 0, the distribution is asymmetric on the right [241].

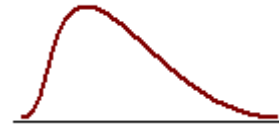
$$CD = \frac{M_3^*}{\sqrt{M_2^{3*}}} \quad (4.9)$$



Skewness <0



Skewness =0



Skewness >0

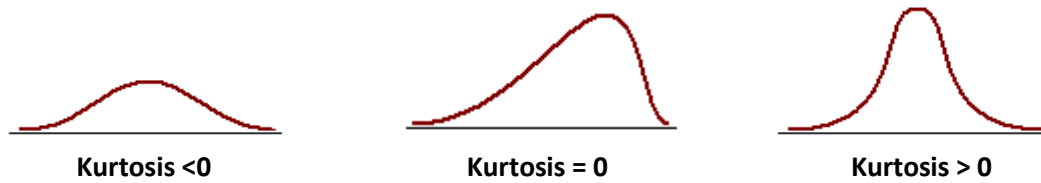
- d- **Median Frequency (MF):** It is the frequency that divides the surface into two equal parts. In other words, the surface before this frequency is equal to that which is after it (eq.4.10) [242].

$$\int_0^{MF} s_x(f) df = \int_{MF}^{F^{max}} s_x(f) df \quad (4.10)$$

- e- **Kurtosis Value (CA):** The Kurtosis coefficient (eq.4.11) measures the degree of crushing of the distribution. It is classically defined as the ratio between the four-centered moment centered and the square of the variance. When it is positive, it indicates that the distribution is "pointed". When it is negative, it indicates that the distribution is relatively "overwritten" [241].

It has this formula:

$$CA = \frac{M_4^*}{M_2^{2*}} \quad (4.11)$$



f- **Frequency Peak:** This is the frequency that corresponds to the maximum of energy [243].

g- **Relative Energy per Frequency Band / Deciles:** We have seen that the median divides the distribution of the spectral density into two parts. The division of this distribution into four, ten, one hundred, or any number of parts can be generalized. The values thus obtained are called quartiles, deciles, percentiles (or percentiles), or quantiles. The energy of each interval is given by the following formula (eq.4.12) [244]:

$$W_n = \frac{\int_{f_{n-1}}^{f_n} S_x(f) df}{M_0} \quad (4.12)$$

$$f_n = \frac{n}{N} f_{\max} \quad 1 < n < N$$

With N is the number of intervals. These parameters represent the spectral variance. The frequency axis is distributed in ten equal intervals.

h- **Percentile:** For this parameter, the corresponding frequencies are removed at equal energy bands. It has this formula [244] (eq.4.13).

$$\int_{fp-1}^{fp} S_x(f) df = k \int_0^{F_{\max}} S_x(f) df \quad (4.13)$$

$$0 < k \leq 1$$

We distinguish:

- Decile: $k = 0.1$ so each part represents 10% of the total area of the power spectral density
- Median: $k = 0.5$ as it has already been said divides the total area into 2 equal parts
- Quartile: $k = 0.25$ divides the total area into 4 equal parts.

i- **H / L (High / Low) ratio:** It is a relationship between two bands of energy extracted after studies a priori of the shapes of the spectral density in the situations of interest. In general, one band increases when the other decreases or vice versa. This gives a good difference and therefore a good discriminating parameter .H and L (eq.4.14) are chosen beforehand and this report will then be of the form:

$$H = [H1, H2] \quad L = [L1, L2]$$



$$\frac{H}{L} = \frac{\int_{fL1}^{fH2} S_x(f) df}{\int_{fL1}^{fH2} S_x(f) df} \quad (4.14)$$

- j- Spectral Entropy:** Entropy measures the amount of average information contained in a signal; it is significant of the spectral variance (eq.4.15) [245]:

$$H = - \int_0^{f \max} S_x(f) \ln[S_x(f)] df \quad (4.15)$$

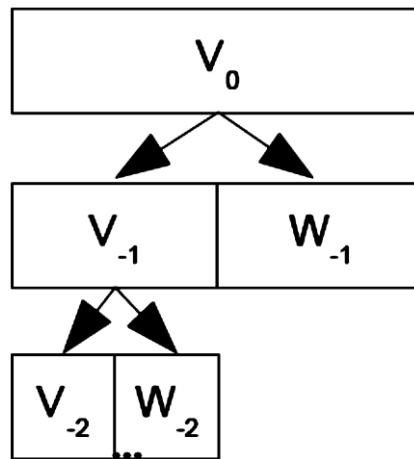
- k- Cepstre:** The cepstrum (eq.4.16) of a signal $x(t)$ is a transformation of this signal from the time domain to another domain analogous to the time domain. This is the inverse Fourier transform applied to the logarithm of the Fourier transform of the signal [246].

$$C_x(\tau) = FT^{-1}\{\ln(X(f))\} \quad (4.16)$$

The cepstrum makes it possible to highlight the periodic components of a spectrum and makes it possible to locate and determine the origin of the faults inducing periodic shocks.

- 2- Time-Frequency Features:** The Fourier transform is used in the case of periodic signals or stationary random signals. For its signals, the same frequencies exist on the entire time domain. But in the case of non-stationary and irregular signals the Fourier analysis cannot give good results as at each time we will have a frequency. That's why you have to go to Time-Frequency or Time-Scale analysis.

- a- Wavelet:** Wavelets generalize Fourier analysis based on time-scale analysis.



The continuous wavelet transform (eq.4.17) of a signal $x(t)$ is:

$$T_x^\psi(a, b) = \int_{-\infty}^{+\infty} x(t) \psi_{ab}(t) dt \quad (4.17)$$

This expression can also be interpreted as a projection of the signal on a family of analytic functions ψ_{ab} (eq.4.18) constructed from an analyzing "mother" function ψ according to the following equation, which is localized in time and frequency:

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (4.18)$$

This equation corresponds to translating a reference wavelet after having dilated it ($a > 1$) or compressed ($a < 1$) (it is the parameter of scale, b parameter of translation). It is possible to reconstruct the signal $x(t)$ from its decomposition into wavelet coefficients, which makes it possible to say that the wavelet transform is invertible and reconstructed (eq.4.19):

$$x(t) = A \iint T_x^\psi(a, b) \psi\left(\frac{t-b}{a}\right) da db \quad (4.19)$$

This parameters presented above can be extracted manually "classical calculation" or depending on other methods known and has a high credibility. Among the different methods for extracting parameters, we distinguish the Principal Component Analysis (PCA) or the Karhunen - Loève transform [247][248].

4.4 Materials and Methods

4.4.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

In order to understand the selection of the methods and techniques applied in our methodology, it is essential to understand the role of each technique used. The grounds for application of the Adaptive Neuro-Fuzzy Inference System is that it is a hybrid system that combines the evident knowledge of the Takagi–Sugeno fuzzy inference system with the supervised learning potential of the multilayer feed-forward neural network in one approach. ANFIS is known to be a vigorous robust technique that aims to realize the nonlinear and complex relationship between input and output data [220] and it is much simpler.

4.4.2 ANFIS Accuracy

Recent studies and experience of system operators indicate that the cost function in the prediction problem is clearly non-linear and that large errors can have disastrous consequences. For this reason, measures based on Root Mean Square Error (RMSE) (eq.4.20) are sometimes suggested because they penalize large errors so can be considered better suited (De Gooijer et al., 2006) [5].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{Y}_i)^2} \quad (4.20)$$

Where $(y_i - \hat{Y}_i)^2$ is the difference between the measured actual value of the system (y) and the previously estimated \hat{Y} , whenever the RMSE has a small value, this means that the predicted value will be more accurate and more acceptable.

4.4.3 ANN Classifier

Neural networks, used for fault diagnosis, can be classified according to (i) the network architecture, and (ii) the learning method [249]. There have been a number of papers that address the problem of fault diagnosis by neural networks. In chemical engineering, Watanabe et al. in 1989 [250], Ungar et al. in 1990 [250] and Hoskins et al. in 91[251] were among the first researchers to demonstrate the usefulness of neural networks for diagnosis. Their strength lies in the ability to learn and capture the relationships between neural network input and output even if these relationships are unknown or difficult to describe. However, their main disadvantage lies in the acquisition and availability of learning data.

The strength in this study is represented by the using of ANFIS combined with ANN in the same algorithm. Feedforward neural networks (FFNN) [252] are the most popular and most widely used models among the many practical applications. This network propagates the input of the network to the following layers without ever going back. This type of network was selected for our research and it's used for the rest of this study. After selecting the neural network type, it is also inevitable to choose an activation function and an error function for their neurons. These options are often guided by the type of data being processed. In this research, a logistic function defined (eq.4.21) is used as an activation function at the level of the hidden layer of the ANN.

$$b_i^{(j)}(x) = 1 / (1 + \exp(-x)) \quad (4.21)$$

For the output layer, the linear transfer function is used. And a simple choice is used for the error function (eq.4.22)

$$E = \frac{1}{2} \|y - t\|_2^2 = \frac{1}{2} \sum_{j=1}^n (y_j - t_j)^2 \quad (4.22)$$

This equation is half the square of the Euclidean distance between the network output (y) and the target (t). Then the minimizing of the average error data is a necessary step taken on all data input by the function E (eq.4.23):

$$E_{average} = \frac{1}{N} \sum_{i=1}^N E_i \quad (4.23)$$

Where N is the number of training data used as input of ANN, and E_i represents the i^{th} learning error.

4.4.4 Time Series Creation

A failure can potentially occur at any given time and can occur without offering prior indication of an inbound failure. For this reason, it is necessary to establish a reliable method that has the capacity to save the system in a, in addition, provide time-sensitive and accurate predictions



in the short, medium and long-term so that unexpected events can be prevented. It is important to note that the proper and effective training of ANFIS should be completed in order to acquire quality results that will lead to durable protection of the system [178]. This can be done by creating a delayed matrix and applying it as the input of ANFIS; this process is shown in Figure 4.4.

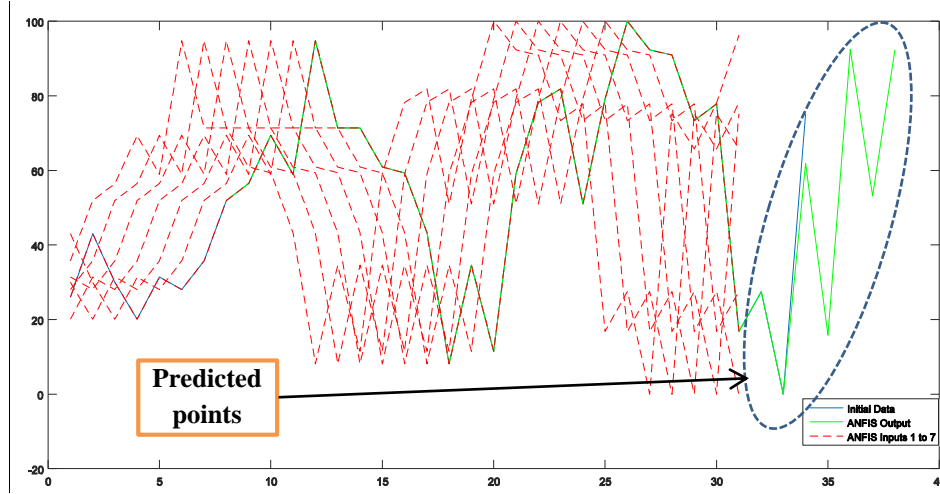


Figure 4.4: Graphical Representation of Time Series Creation

Figure 4.4 demonstrates that the x-axis is the sample index and y-axis is the signal amplitude. The red dashed lines represent the time series generated as the input of ANFIS, and the green circles are the predicted points

Let's consider a space X with (p) points (eq.4.24):

$$X : [x_1, \dots, x_r, \dots, x_p] \quad (4.24)$$

Prediction is a process based on collecting data of the past $x(t-k)$ and present states $x(t)$ of the system, to predict future ones $x(t+k)$. X is considered as the signal to be predicted; from this X we should create a time series matrix (shifted matrix) to train the prediction algorithm as inputs of the ANFIS algorithm (eq.4.25).

$$\text{Inputs : } \begin{bmatrix} x(t) & x(t+1) & \dots & x(t+(P-(n+1))) \\ x(t-1) & x(t) & \dots & x(t+(P-(n+2))) \\ \dots & \dots & \dots & \dots \\ x(t-(n-1)) & x(t-n) & \dots & x(t-(P-(n+1))) \end{bmatrix} \quad (4.25)$$

$\downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$
 Target : $[x(t+1), \quad x(t+2), \quad \dots \quad x(t+(p-n))]$

Let's consider a delay vector as follows: Delay= [1, 2, 3, ..., n] $n < P$ Where n is the number of



input vectors and it determines the type of the prediction mode (short, mid or long-term); when n grows that means we are heading towards a long-term prediction.

4.5 Proposed Methodology

For the purposes of this study, Artificial Neural Networks was the selected method. The classification method presented here are usually associated with pattern recognition. As shown in Figure 4.5, a complete pattern recognition system is associated with several modules. By disregarding the acquisition module and the preprocessing module, a classification-based diagnostic system is associated with three modules: constructing the parameter vector, retrieving/selecting parameters, and classifying.

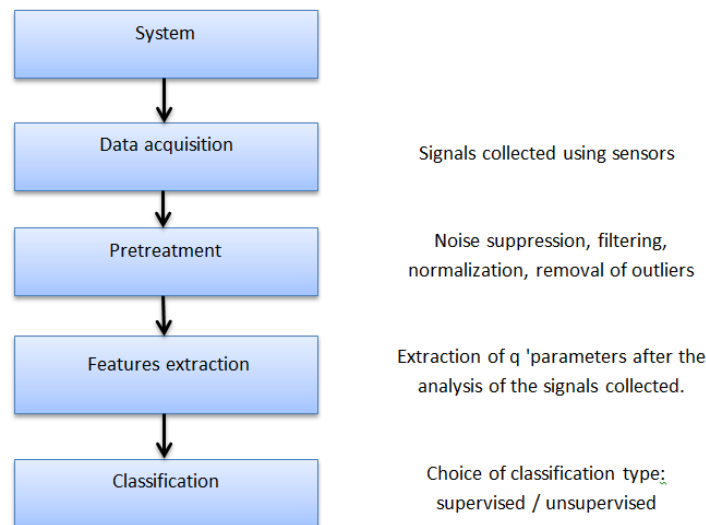


Figure 4.5: classification system steps

The statistical analysis of real signals proves that the acquired signals are deterministic and they are signals with very poor frequencies and the most time parameters are equal (normal mode in Figure 4.3 and Table 4.2) therefore, in the classification system it is necessary to skip the classical feature extraction step as shown in Figure 4.6 (time and frequency domain features) and thus couldn't provide significant results.

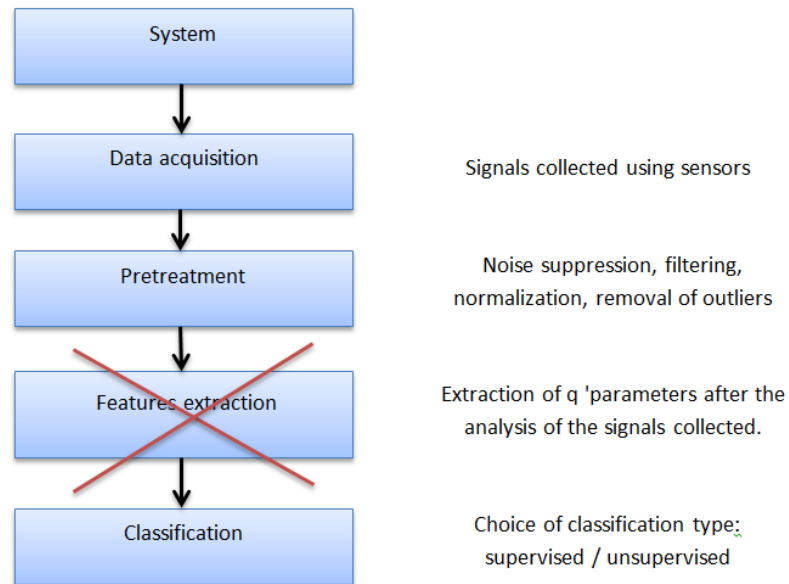


Figure 4.6: Cancellation of feature extraction step

Therefore, we found that the solution to this problem is to apply the features extraction step on the error calculated at the output of the ANFIS algorithm and subsequently these calculated features can be used as the input layer for the neural network classifier. In addition, the classification of faults depends on the training of the classifier with all types of faults possible and this is a problem in itself because the fault may have an infinite form so it is very difficult to compute and understand this countless number of faults, the solution of this problem it to diagnose the origin of fault depending only on normal data.

Our proposed methodology can be summarized by the following three steps:

Step1: Training ANFIS

For each normal acquired signal

- *Training ANFIS (A_1, A_2, \dots, A_n) using 70% of normal data (D_1, D_2, \dots, D_n), where $n=8$ in our case.*
- *Validating ANFISs (A_1, A_2, \dots, A_n) with 30% from the same normal data (D_1, D_2, \dots, D_n).*
- *Apply ANFISs (A_i) where ($i=1,2,\dots,8$) on the whole normal signals, this produces normal RMSE errors (ε_i , $i=1,2,\dots,8$).*
- *Calculate the predicted data (P_i)*
- *Error signals calculated from normal mode*

Step2: Training ANN

- *Extract the features vector (NF_i) from each normal error ε_i*
- *Training each ANN_i ($i=1,\dots,8$) using 70% of (NF_i)*



- Validation of each ANN_i by 30% of extracted features (NF_i).
- Thresholding (if the output of any $ANN < 0.5$ put the output=0, else output=1)

Step3: For a real-time application,

- Testing ANFIS (A_i) on a window of unknown data (D'_i).
- Predict the future estimated data signal (P'_i)
- Calculating the error vector ϵ'_i from the difference between the target and the actual output
- Extract the vectors of features (UF) from (ϵ'_i)
- Testing the ANN_i using extracted vector of features
- Diagnose the state of the system

These 3 steps can be represented by a simple block diagram shown in Figure 4.7.

In the first step, with normal data observations (40 observations), each observation has 8 signals (D_1, D_2, \dots, D_n) (see Table 4.1), a time series matrix is generated from each signal D_i as an input for the ANFIS (A_i) algorithm ($i=1, \dots, 8$), where 70% of this time series matrix is used to train ANFIS, and the remaining amount (30%) is used for ANFIS validation. Accordingly for each ANFIS (A_i) a vector of predicted data (P_i) is calculated on the output of this ANFIS (A_i) and a vector of residual error (ϵ_i) is generated between the target and the actual output. At the end of this step a database of 8 RMSE vectors is created from the normal data.

In step 2, the generated error in steps 1 represents random signals that contrast the original signals that are deterministic.

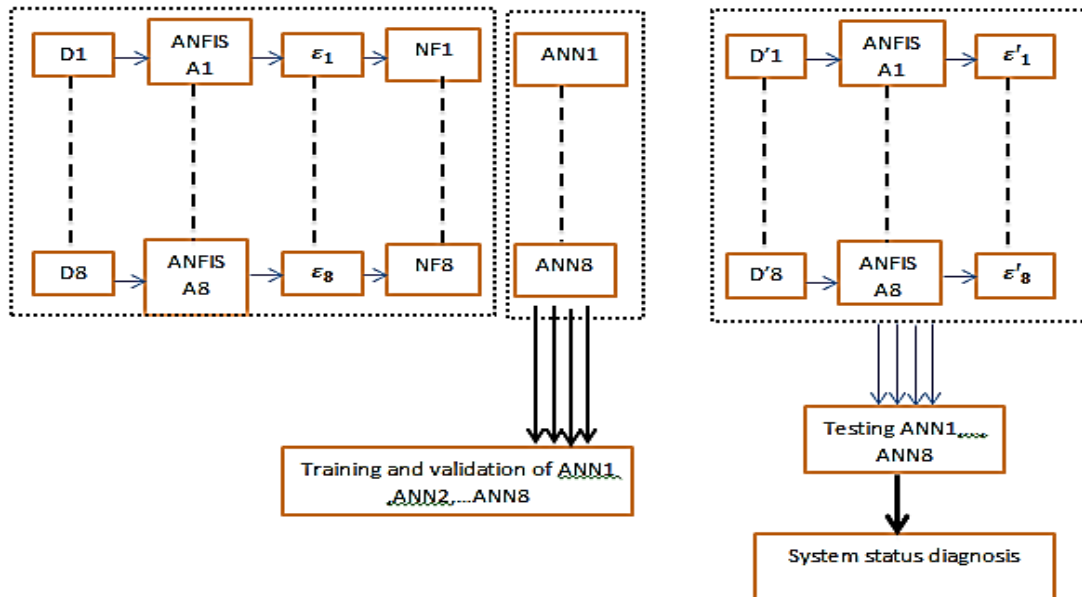


Figure 4.7: Block Diagram of Proposed Methodology



Normandie Université

Therefore, we used each generated errors signals (ϵ_i) as an input of each Feed Forward Artificial Neural Networks ($FFANN_i$) after a features extraction step. A database of vectors of features (NF_i) is created. where 70% of this database is used to train each ($FFANN_i$) and 30% of them used to validate each ($FFANN_i$). Each feature vector (NF_i) contain 34 features.

The extracted features from each signal are:

1. Time domain features: minimum, maximum, mean, variance, quadratic mean value, standard deviation, RMS, crest factor, envelope;
2. Frequency domain: the power of signal, mean frequency (MPF), skewness, kurtosis, median frequency, frequency peak, relative energy in frequency bands, percentiles, H/L (high/low) ratio, spectral entropy;
3. Time-frequency domain: wavelets [248].

In step 3, after training each ANFIS (A_i), with the same parameters (premise and consequence) calculated before in step 1, and after the training of ($FFANN_i$) is step 2 a new windows of unknown acquired data (D'_1, D'_2, \dots, D'_n) is applied for testing each ANFIS (A_i). A residual error (ϵ'_i) is produced as a result of this process, and a prediction data (P'_i) is calculated. Then an ensemble of vectors of features is calculated from the residual error (ϵ'_i), this ensemble is used to test each ($FFANN_i$) created before. As long as each ANFIS (A_i) is trained by a normal data in step 1, the predicted values (P'_i) should also be in the normal range, but this case is true only when the tested data is also normal. When the data starts to deviate from the normal range, the difference between the target of ANFIS (A_i) and the output will increase; this demonstrates that the system has begun to deviate from its normal course. At this point, ANFIS has the ability to predict that the system seems to be deviating from its normal path and is heading towards a problem. This assists ANN to diagnose the future state of the system.

In this study, we use 8 multi-layer feed forward neural networks as the first classifiers. Each ANN (Figure 4.8) is formed of 3 layers. The first is the inputs layer that contains 2 inputs representing the extracted features vector from the error calculated at the output of ANFIS and the desired output that it is zero (target=0) in our study. the respectively input of the 8 ANFIS is the acquired signals D1-the Timer: reflux rate, D2- the Heating power, D3- the Feed flow rate, D4- the preheated Power, D5- the Pressure drop, D6- the preheated temperature, D7- the temperature boiler, and D8- the TIC2 Column head temperature. The second is a hidden layer having 10 neurons with a logistic learning function. The third is the output layer containing 1 linear neuron.



Normandie Université

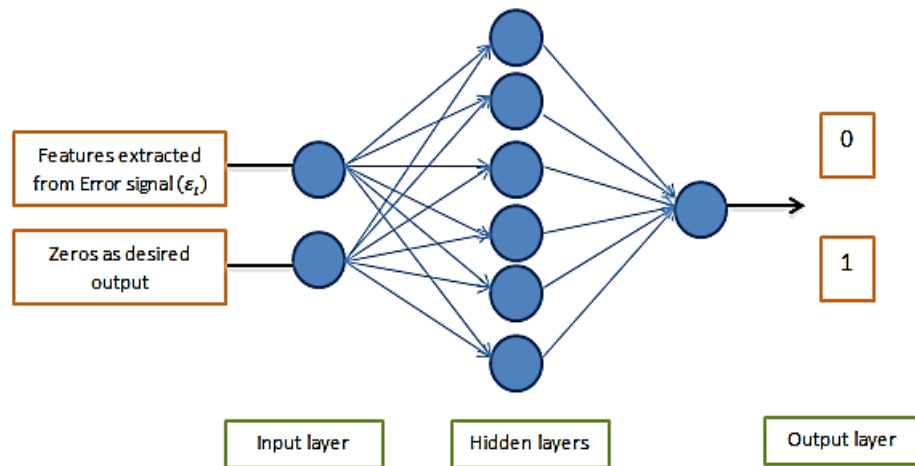


Figure 4.8: Neural Network Proposed Architecture

Approximately every network is able to find the solution after 75 epochs with a learning rate $\mu=0.0009$.

4.6 Results and Discussions

In the previous section, we have presented the main framework for the process of fault prediction and diagnosis. In this section, we provide the results of the real application of the process for diagnostic purposes. The proposed methodology is designed to detect 8 kinds of faults (E1...E8):

- E1: Increasing Degradation Reflux Ratio From 0 To 100% (Timer);
- E2: Decreasing Degradation Reflux Ratio From 10 To 0% (Timer);
- E3: Increasing Degradation Heating Power 0 To 100%;
- E4: Decreasing Degradation Heating Power 100 To 0%;
- E5: Increasing Degradation Preheated Power 0 To 85 %;
- E6: Decreasing Degradation Preheated Power 100 To 0%;
- E7: Increasing Degradation Feed Rate 0 To 100%.
- E8: Decreasing Degradation Feed Rate 100 To 0%;

In previous studies, all applied research on the prediction; diagnosis and detection of faults in distillation columns have relied on the use of modeling techniques in the calculation of residual between the real signal and the model. The black box ANN is the best and most commonly used technique in this domain[130]. Most studies on the detection of changes in distillation columns were carried out by Chetouani et al. [162][160] with the most recent study completed in 2014. This 2014 study, researchers applied the Bayes decision theory and combined it with the neural adaptive black-box identification for modeling such systems [162].

The research conducted in this study provided an efficient fault prediction and diagnosis method when tested on real distillation processes. However, the weakness in such work models



and processes is that they each signal independently of each other which require long periods of time for simulation. Furthermore, their research did not take into consideration how a fault detected in one signal affects other signals and the whole system. In some cases, a false alarm may occur because of unexpected faults that may affect the other signals. Then, in this case, the model will fail to detect the true fault. For example, when the Timer is blocked (0%), this causes the overhead temperature (TIC2) to decrease. The method succeeds in this case to detect the fault. On the other hand, if the boiler has a failed operation status (0% power heating), TIC2 also decreases. This will create a false alarm rather than real fault detection in the system. Unlike this study, our proposed method provides an overall scan of the complete system and it analyzes all eight signals simultaneously.

This approach will solve the problem of false alarms by analyzing the fault and its influence on all other signals. The algorithm makes the analysis in one iteration step and thus needs less time to perform the diagnosis operation. On the other hand, this methodology is designed to detect, in real time, the small variation in the system then diagnoses the type of fault that may potentially occur.

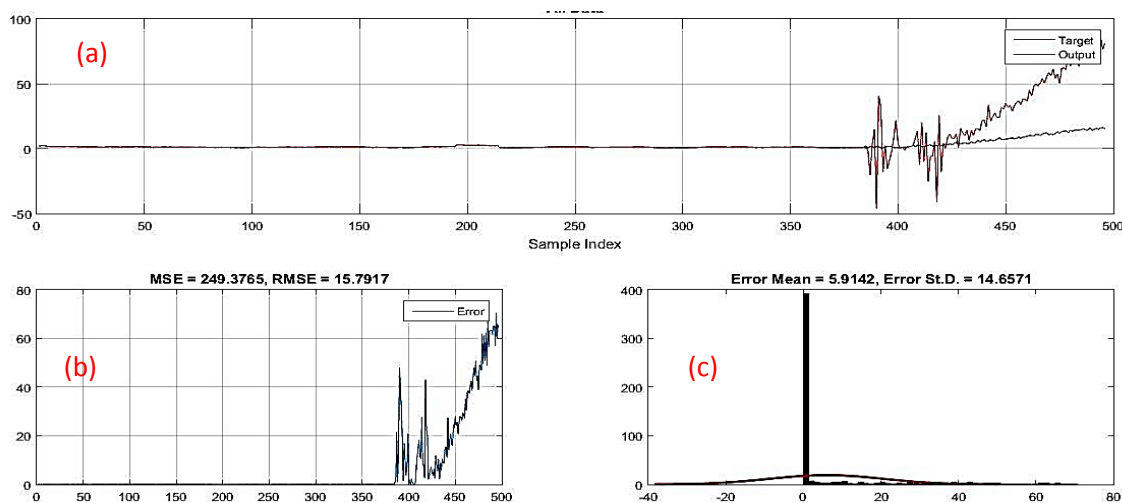


Figure 4.9.a: Graphical Representation of Results Using ANFIS as Predictor of Distillation Column Data (Reflux Rate Signal)

Figure 4.9.b: Graphical Representation of Absolute Value of the Error Calculated Each Iteration

Figure 4.9.c: Graphical Representation of Error Absolute Value Distributions

The black curve in Figure 4.9.a is the real signal and the red one is the output of ANFIS. The x-axis is the sample index and the y-axis is the amplitude of the reflux rate signal. The results are shown in Figure 4.9. a, Figure 4.9.b, Figure 4.9.c, confirm that ANFIS has the extensive ability to a high level of accuracy to predict the small variation occurrence in distillation column data. As can be observed when the signal (black curve in Figure 4.9.a) started to deviate from its natural course ($t=380$), it was evident that the difference (error) between the signal and the ANFIS output clearly increases (Figure 4.9.b). This is an indication that ANFIS has a significant ability to



detect a small undetected variation in the signals. The results also allow diagnosing the mode where the fault originates and they classify the type of faults that occur during the distillation process. The results of the errors calculated on the output of ANFIS are presented in Figure 4.10: In general, it is clearly demonstrated that these errors are random signals that have various variations contrary to the inputs signals shown in Figure 4.3, and this is an advantage in applying the features extraction techniques on this calculated signals (errors signals).

However, for example, in the E1 case (accident reflux ratio timer 100%), we notice that the parameter numbers 4 (timer), 5 (charge loss) and 7 (heating power), increase significantly. Therefore, these variations could be considered as fault indicators and could differentiate between normal and abnormal cases and between the eight different types of faults. Figure 4.12 shows that the error between the real and desired output is negligible (in order of 10^{-6}). This proves that the proposed methodology is successful in detecting faults introduced within the distillation process plant model.

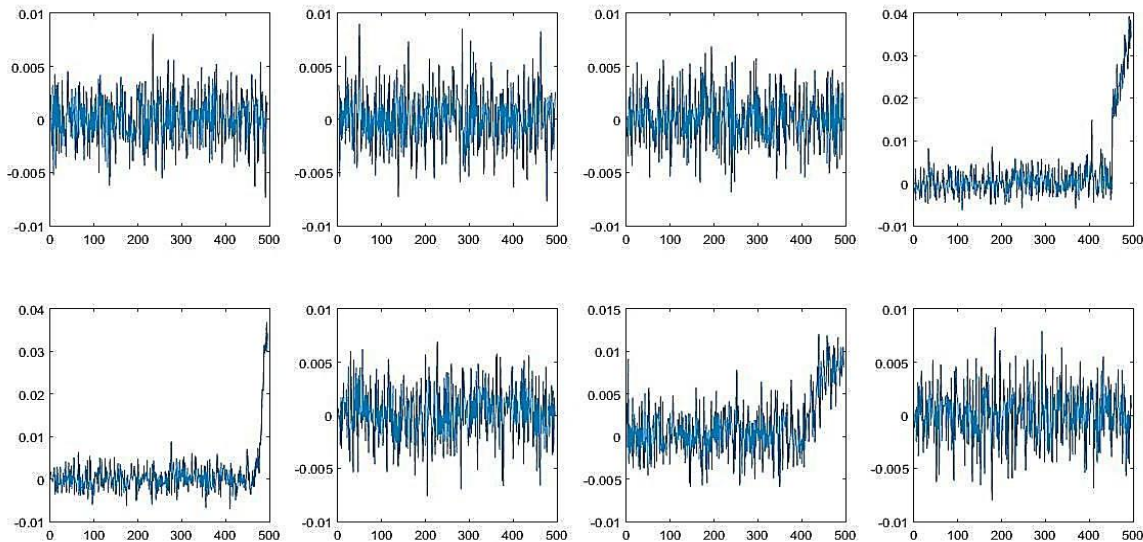


Figure 4.10: Graphical Representation of Absolute Value of Errors Calculated on ANFIS Outputs (Accident E1)

In the ANN testing step, as previously mentioned, the FFANN is learned through normal and fault data. For example, if the checked data is considered from the normal mode, therefore the outputs of FFANN after a thresholding step (if the output of FFANN < 0.5 put the output = 0, else output = 1) is a vector of 8 zeros, it means that no one from the predicted signal ($\mathbf{P}'_1, \dots, \mathbf{P}'_8$) will leave the normal mode. In the case of a degraded mode such as an increasing degraded reflux rate mode (E1) (Figure 4.11), the output of each FFANN variate according to the data as can be seen in Table 4.3.

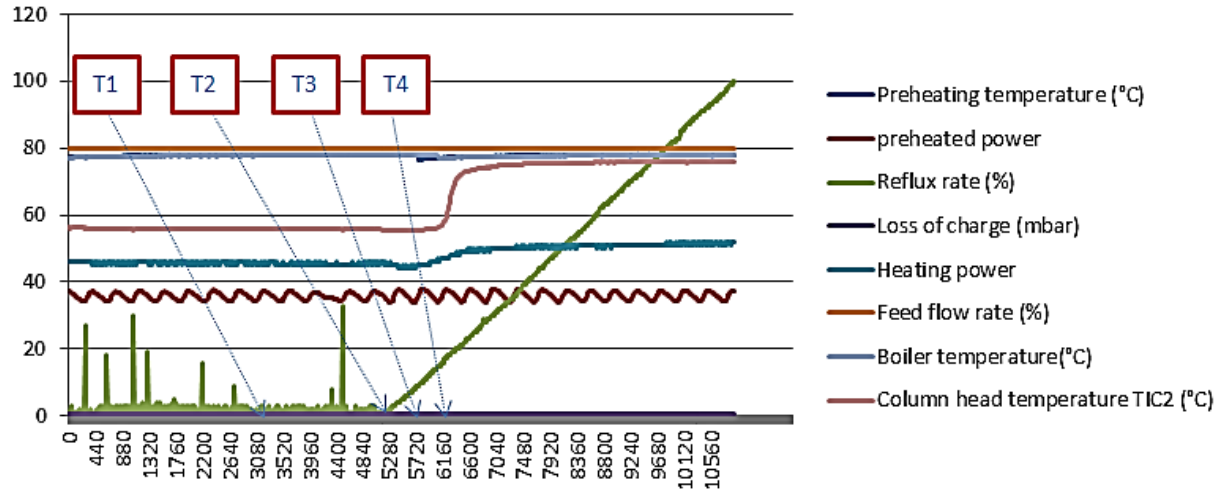


Figure 4.11: Graphical Representation of Increasing Degraded Reflux Rate Accident (E1)

	D1	D2	D3	D4	D5	D6	D7	D8
T1=3113	0	0	0	0	0	0	0	0
T2=5005	0	0	1	0	0	0	0	0
T3=5700	0	0	1	0	1	0	0	0
T4=6149	0	0	1	0	1	0	1	0

Table 4.3 Artificial Neural Network Output Results on Reflux Rate Accident E1

In Table 4.3 at the time $T1= 3113$, all the signals are at normal values, so the $FFANN_{1...8}$ classifies the data in the normal mode and gives a vector of 8 zeros on the output (Table 4.3, row 2). On $T2=5005$, the timer signal (D'_1) leaves the normal range and starts to increase, so the output of $FFANN_1$ is updated to 1 on (D'_1) and all others remain zeros (Table 4.3, row 3). After some time, at $T3=5700$, the increase of the timer signal affects the heating power signal (D'_2), therefore the $FFANN_2$ updates its output again (Table 4.3, row 4). After more time has passed ($T4=6149$), this accident has an effect on the TIC2 (D'_7) signal, and again the output of the $FFANN_7$ is updated appear as the row 5 in Table 4.3. In conclusion, the $FFANN$ updates its output concurrently with the variation of any of input signals.



D'_1	D'_2	D'_3	D'_4	D'_5	D'_6	D'_7	D'_8	
0	0	0	0	0	0	0	0	Normal
0	0	1	0	1	0	1	0	Accident E1
0	0	1	0	1	0	0	1	Accident E2
0	1	1	1	1	0	0	0	Accident E3
0	0	1	1	1	0	1	1	Accident E4
1	1	0	0	0	0	0	0	Accident E5
1	1	1	0	0	0	0	0	Accident E6
1	0	1	0	0	1	0	0	Accident E7
0	0	1	0	0	1	0	0	Accident E8

Table 4.4: The Result of Artificial Neural Network Output for All Accidents

Table 4.4 represents the output of the artificial network for each accident type, as you see in this table every accident has their own profile. The target or output vector is an 8 element vector with a '1' in the position of the fault it represents and '0' everywhere else (Figure 4.8). In other words, the output value is set to '0' to indicate no fault and '1' to indicate the presence of a fault.

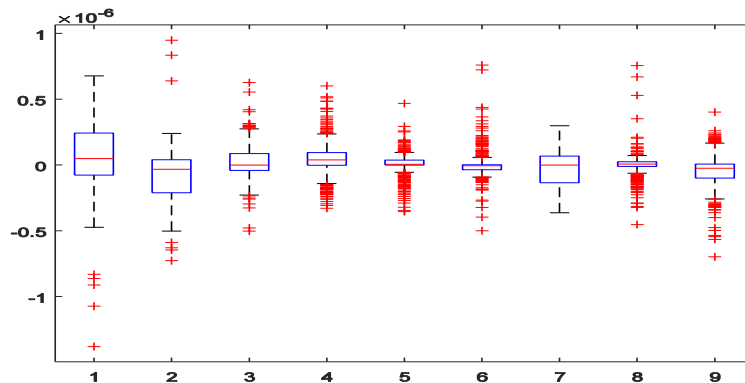


Figure 4.12: Error Boxplot- Error between Real and Desired Output

The results shown affirm the high-level capability of the proposed algorithm, to predict and diagnose all type of faults that may occur in the distillation process; in addition, the diagnosis of faults was only based on normal data without the need to train the algorithm any abnormal data. In this case, we have eliminated the problem of infinite forms of faults that are difficult to understand.

4.7 Conclusion

It is widely known in the industry that the accurate prediction and diagnosis of faults is fundamental in reducing maintenance costs. It is also a contributing factor to improving system



reliability as well as the level of safety that can affect all stakeholders of a system. This is one of the most relevant considerations of the industry at a global level.

The topic of fault prediction and diagnosis becomes even more pertinent when the discussion is linked to the strength and viability of equipment that has potential risks which can lead to various levels of material costs at the minimal level and even human costs at a catastrophic level. This was clearly reflected in the catastrophe of the reactors in Japan.

This chapter has presented and demonstrated the accuracy and strength of a methodology that aims at efficient and strategic fault diagnosis and prediction in equipment that increasingly deteriorates over time and is affected by several external factors. In this study, the proposed ANFIS that was used to detect distillation column deterioration, a step that features extraction techniques we unable to do, is also a strategic method that can be used for early, timely and detailed detection of faults. Furthermore, the FFANN model uses the data of previous inspection points of degradation values in order to complete a wider diagnosis of the state of the entire system.

This chapter has demonstrated results that prove the success and efficiency of the proposed methodology for ANFIS, neural network classification. Through its application on actual experimental data, the methodology was able to predict and diagnose, in real time, faults in the distillation column during the automated continuous distillation process. The results also confirmed that it is possible to classify between normal and eight abnormal classes of faults. Building on the results of our experiment that was presented in this work the possibility of prospective study emerges.



Normandie Université

General conclusion



This study is in the field of health systems management, basically aims to develop maintenance support tools. The work presented in this thesis deals with the integration of the diagnosis and the prognosis of failures in the dependability of industrial chemical reactors in process engineering. The diagnosis makes it possible to determine online the system components of that are at fault and need to be repaired. The prognosis makes it possible to anticipate the errors by providing information on the future state of the system and it's needed maintenance actions that should be applied later on. Even though these techniques seem to be correlated, they are usually studied separately. this is because the time scales manipulated by the two processes are very different. Therefore, in addition to the study of faults diagnosis and faults prognostics separately, this thesis is ended by a major challenge represented by their integration as a new maintenance strategy.

The objective of this work is to study is the development of a new diagnostic approach and a new prognostic approach for a high-level monitoring of a distillation column. All this followed by an integration of a new approach that combines the diagnostic and prognostic methods in a common algorithm for the goal of improving the maintenance strategy.

Chapter I provides an overview of general industrial accidents and accidents occurring in the chemical industry, especially in chemical reactors. This chapter focuses on the entire distillation process with an in-depth explanation of the continuous automatic distillation process. A detailed explanation of the data that reflects the errors that may occur in the distillation process is given to reach data collection for faults diagnosis and prognosis with clarification of the characteristics of the acquired signals. This chapter is ended by a pre-processing of the data used in this study followed by a state of art for data reduction including features extraction, features selection, and a framework of the maintenance strategy of complex systems.

The second chapter presents the qualitative and quantitative methods most widely used in the literature for the diagnosis of chemical processes and the different methods of risk analysis. This state of the art is the result of a bibliographic search of about a hundred well-studied articles on the evolution of the monitoring of the chemical processes and the different methods developed. A synthesis on methods of monitoring chemical reactors models free is developed in this section. Then we move to a state of the art that demonstrates the outcome work of detection and diagnosis of faults that may occur in the distillation column and clarify the weaknesses of the previous methods. The results of this biography prove that the selection of the neuro-fuzzy is a better technique for the diagnosis of faults that occur in a distillation column. This chapter is ended by an application of fuzzy logic and ANN separately on the data extracted from the distillation column. Depending on the results we obtained, we decided to propose a more efficient approach in real-time analysis of distillation column system. It proposes a methodology that combines fuzzy mean clustering and neural network for diagnosis, detection, and classification of many faults. Moreover, a modified FCM method (MFCM) is presented in place of a feature extraction and selection approach. MFCM is a clustering method that allows calculating the degree of variation between normal and abnormal modes. The output of the MFCM is considered as an input for the neural network classifier. This proposed methodology is then tested via real



experimental data obtained from a distillation column, after a pre-processing step including filtering and smoothing of the signals. A database with normal and faulty observations is analyzed. The database is composed of eight different types of faults that may occur during the automated distillation process in the chemical industry. The results of the proposed method confirm the ability to differentiate between normal and eight abnormal classes of faults.

The design of PHM has become an important element in the realization of an effective predictive maintenance strategy. Effective maintenance is understood to ensure the reliability of systems while reducing the costs incurred by interventions and shutdowns due to sudden failures. The third chapter is devoted to the positioning of prognostic activity in the context of industrial maintenance and to the study of potentially useful tools to support this process and gives an overview of prognostic methods. An overview of the classification of prognostics approaches applied to the chemical reactors and the distillation column are also discussed in this chapter including their pros and cons and many RUL estimated strategies are also reviewed. The investigation track in this chapter is aimed to choose the adaptive neuro-fuzzy inference system (ANFIS) approach as the best technique. ANFIS is able to calculate the RUL of the distillation column degradation (guide to choosing a prognostic tool). This chapter is ended by the development of a new prognostic strategy applied to a real experimental data acquired from distillation column and from a metric pump. This methodology is a new technique which is effective in determining the path of deterioration of the distillation column system and also predicts the future path (prognostic) of this system by determining their RUL. Also, this work presents a direct monitoring approach based on the technique of adaptive neuro-fuzzy inference system (ANFIS) combined with fuzzy C-means algorithm (FCM). The results of a comparative study between the results of our proposed methodology and other based on ANN are discussed at the end of this application. The Results demonstrate the validity of the proposed technique to achieve the needed objectives with a high-level of accuracy, especially in determining a more accurate Remaining Useful Life (RUL) when it is applied on the automated distillation process in the chemical industry.

To improve the performance of ANFIS algorithm, Parzen windows distribution is proposed as a new membership function for ANFIS algorithm. The aim of this proposal is to reduce the consumption of time and make the processing closer to a real-time application or minimizing the root means square error (RMSE) between the real and the predictive data. The methodology is tested on real experimental data obtained from a distillation column aiming to predict the failure that may occur during the automated continuous distillation process. A comparative study was needed to choose the better membership function that can be used for ANFIS algorithm when ANFIS applied to distillation column data. The results obtained in this research demonstrated that Parzen window proved its worth as a new membership function of ANFIS algorithm when it is applied to the distillation column data and it also proved to be successful in reducing the execution time of ANFIS. The results have also shown that Parzen MF is chosen as the best MF for three over eight types of normal signals and for five over eight degraded signals.



The fourth and last chapter proposes a new integrated methodology that works for fault prognostics and diagnosis in the same time as a full scanning system for distillation column faults. The Adaptive Neuro-Fuzzy Inference System (ANFIS), as a hybrid system, has been selected for the step related to prediction since it combines the advantages of fuzzy logic and ANNs in one simultaneous algorithm. In our research, we tested this methodology with real experimental data that was obtained from a real distillation column. This resulted in the analysis of a database with different types of faults that could potentially occur during the automated distillation process. The results that were observed proved the validity and strength of this proposed technique. It was also demonstrated that the technique achieved a high level of accuracy, the objective of prediction and diagnosis especially when applied to the data obtained from automated distillation process in the chemical industry.

Throughout Chapters 2, 3 and 4, after the illustration of theoretical concepts, new methodologies for diagnosis and prognosis have been applied to a distillation system. The results of the diagnostic function illustrate how the diagnosis is robust to uncertain, missing, and false observations. The results of the prognostic function illustrate how the prognosis tips the current diagnosis, behavior model, and system degradation model to return a distribution belief over the remaining useful life (RUL).

It is difficult to develop an applicable diagnostic or prognostic algorithm with the same efficiency on all types of systems. Performance varies from one application to another. A compromise must be made between performance and applicability.

Throughout this manuscript, hypotheses about the problem or its resolution are posed. It would be interesting to question them for future developments.

This work, although defined in a health management context, is limited to monitoring the health of the systems. One of the prospects of this work is, therefore, the exploitation of the results of diagnostic and prognostic methods having forms of belief distributions. This exploitation can be directed towards several objectives:

- The use of diagnostic results and prognosis to improve health surveillance with respect to health management objectives of the system.
- The integration of an active diagnostic function (which makes it possible to perform actions in order to refine the diagnostic result) which could also be influenced by the prognostic results.
- Re-planning and reconfiguring the system autonomously, in order to ensure the mission of the system, to minimize its wear and/or its unavailability.

The prospects of this thesis work are therefore in five directions:

- The consideration in the model of an observable hybrid dynamic.
- The proposal of a prognostic method coupling the analytical resolution and the simulation of the degradation model for a more precise prognostic result.



Normandie Université

- The extension of the methodology to heterogeneous systems and the integration of mechanisms of neuro-fuzzy networks such as parallelism and synchronization.
- The extension of the methodology to repairable systems.
- Exploitation of the results of diagnostic and prognostic functions taking forms of belief distributions.



References:

- [1] M. R. Othman, M. W. Ali, and M. Z. Kamsah, "Process fault detection using hierarchical artificial neural network diagnostic strategy," *J. Teknol.*, vol. 46, no. 1, pp. 11–26, 2012.
- [2] Y. Chetouani, "Fault detection in a chemical reactor by using the standardized innovation," *Process Saf. Environ. Prot.*, vol. 84, no. 1, pp. 27–32, 2006.
- [3] A. Daher, Y. Chetouani, G. Hoblos, and M. Khalil, "Modified fuzzy c-means combined with neural network based fault diagnosis approach for a distillation column," presented at the Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International, 2016, pp. 185–190.
- [4] R. Khelif, S. Malinowski, B. Chebel-Morello, and N. Zerhouni, "RUL prediction based on a new similarity-instance based approach," presented at the Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium on, 2014, pp. 2463–2468.
- [5] O. E. Vasile-Dragomir, "Contribution au pronostic de défaillances par réseau neuro-flou: maîtrise de l'erreur de prédiction.," 2008.
- [6] T. Sutharssan, D. Montalvao, Y. K. Chen, W.-C. Wang, C. Pisac, and H. Elemara, "A review on prognostics and health monitoring of proton exchange membrane fuel cell," *Renew. Sustain. Energy Rev.*, 2016.
- [7] A. K. Palit and D. Popovic, "Transparent Fuzzy/Neuro-fuzzy Modelling," *Comput. Intell. Time Ser. Forecast. Theory Eng. Appl.*, pp. 275–303, 2005.
- [8] J.-S. Wang and C. G. Lee, "Efficient neuro-fuzzy control systems for autonomous underwater vehicle control," presented at the Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on, 2001, vol. 3, pp. 2986–2991.
- [9] C. Mahabir, F. Hicks, and A. R. Fayek, "Neuro-fuzzy river ice breakup forecasting system," *Cold Reg. Sci. Technol.*, vol. 46, no. 2, pp. 100–112, 2006.
- [10] C. Li and K.-H. Cheng, "Recurrent neuro-fuzzy hybrid-learning approach to accurate system modeling," *Fuzzy Sets Syst.*, vol. 158, no. 2, pp. 194–212, 2007.
- [11] A. M. Ardhian and S. P. Hadi, "Membership Function Comparative Study on Load Forecasting using ANFIS Framework."
- [12] M. K. Mayilvaganan and K. Naidu, "ANN and Fuzzy Logic Models for the Prediction of groundwater level of a watershed," *Int. J. Comput. Sci. Eng.*, vol. 3, no. 6, pp. 2523–2530, 2011.
- [13] R. Singh, A. Kainthola, and T. Singh, "Estimation of elastic constant of rocks using an ANFIS approach," *Appl. Soft Comput.*, vol. 12, no. 1, pp. 40–45, 2012.
- [14] R. El Harabi, "Supervision des processus chimiques à base de modèles Bond Graphs," 2011.
- [15] R. Glaubitz, *The Aria Database*. Robert Glaubitz, 1996.
- [16] F. Lees, *Lees' Loss prevention in the process industries: Hazard identification, assessment and control*. Butterworth-Heinemann, 2012.
- [17] J. Gupta, "Bhopal gas tragedy and its effects on process safety," *J. Loss Prev. Process Ind.*, vol. 18, no. 4, pp. 197–199, 2005.
- [18] J.-L. Gustin, "How the study of accident case histories can prevent runaway reaction accidents from recurring," *Process Saf. Environ. Prot.*, vol. 80, no. 1, pp. 16–24, 2002.
- [19] T. A. Kletz, "Accident investigation: keep asking 'why?'," *J. Hazard. Mater.*, vol. 130, no. 1, pp. 69–75, 2006.



- [20] D. A. Crowl and J. F. Louvar, *Chemical process safety: fundamentals with applications*. Pearson Education, 2001.
- [21] A. Neuberger, "ROLE OF RISK MANAGEMENT IN THE OIL AND GAS INDUSTRY: THE EFFECT OF DERIVATIVE CONTRACTS USED TO MANAGE OIL AND GAS PRICE RISK ON STOCK PRICE SENSITIVITY IN THE OIL AND GAS INDUSTRY DURING A PERIOD OF DECLINING OIL AND GAS PRICES," 2016.
- [22] S. Alexander *et al.*, "Dioxines dans l'environnement: quels risques pour la santé?," 2000.
- [23] A. LA UNE, "Risques industriels," 2008.
- [24] S. Imperatives, "Report of the World Commission on Environment and Development: Our Common Future," 2016.
- [25] M. Grunell, "The Netherlands: The representativeness of trade unions and employer associations in the chemical sector," *EurWORK*, 2014.
- [26] D. Fragkoulis, "Détection et localisation des défauts provenant des actionneurs et des capteurs: application sur un système non linéaire," 2008.
- [27] *Norme X60-319/NF EN 13306 Terminologie de la maintenance*. 2001.
- [28] J.-M. Chanut *et al.*, "L'industrie française au milieu du 19^e siècle (les enquêtes de la statistique générale en France)," *Rech. Hist. Sci. Soc.*, 2000.
- [29] R. Isermann and P. Balle, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control Eng. Pract.*, vol. 5, no. 5, pp. 709–719, 1997.
- [30] M. Blanke, W. C. Frei, F. Kraus, J. R. Patton, and M. Staroswiecki, "What is fault-tolerant control?," *IFAC Proc. Vol.*, vol. 33, no. 11, pp. 41–52, 2000.
- [31] G. Zwingelstein, "failures diagnosis. Theory and practice for industrial systems," 1995.
- [32] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2000.
- [33] K. Khedri, "Modélisation et simulation d'une application industrielle par le générateur du système expert G 2," *Mém. Magister Au Dép. Hygiène Sécurité Univ. Batna*, 2002.
- [34] R. Zemouri, "Contribution à la surveillance des systèmes de production à l'aide des réseaux de neurones dynamiques: Application à la e-maintenance," 2003.
- [35] R. Isermann, R. Schwarz, and S. Stolz, "Fault-tolerant drive-by-wire systems," *IEEE Control Syst.*, vol. 22, no. 5, pp. 64–81, 2002.
- [36] G. J. Ducard, *Fault-tolerant flight control and guidance systems: Practical methods for small unmanned aerial vehicles*. Springer Science & Business Media, 2009.
- [37] J. G. Stichlmair and J. R. Fair, *Distillation: principles and practices*. Wiley-VCH New York, 1998.
- [38] J. G. Speight, *The chemistry and technology of petroleum*. CRC press, 2014.
- [39] C. S. Robinson, *The elements of fractional distillation*. McGraw-Hill book company, inc., 1922.
- [40] L. Serafimov, T. Chelyuskina, and P. Mavletkulova, "Special cases of determination of the minimum reflux ratio for distillation of binary mixtures," *Theor. Found. Chem. Eng.*, vol. 47, no. 3, pp. 231–238, 2013.
- [41] Y. Chetouani, "MODEL-BASED ONLINE FAULT DETECTION AND DIAGNOSIS (FDD) STRATEGY FOR A CHEMICAL REACTOR."
- [42] F. Bolaers, J. Dron, and L. Rasolofondraibe, "Prédiction et suivi de l'évolution d'un écaillage de fatigue de roulement par analyse vibratoire," *XVIème Congrès Fr. Mécanique Nice Réf.*, vol. 649, 2003.



- [43] E. Ftoutou, *Détection et Diagnostic des Défauts de Moteurs par Analyse vibratoire*. Éditions universitaires européennes, 2015.
- [44] O. Djebili, "Contribution à la maintenance prédictive par analyse vibratoire des composants mécaniques tournants: Application aux butées à billes soumises à la fatigue de contact de roulement," 2013.
- [45] F. Camci, K. Medjaher, N. Zerhouni, and P. Nectoux, "Feature evaluation for effective bearing prognostics," *Qual. Reliab. Eng. Int.*, vol. 29, no. 4, pp. 477–486, 2013.
- [46] W. Yan, H. Qiu, and N. Iyer, "Feature extraction for bearing prognostics and health management (phm)-a survey (preprint)," AIR FORCE RESEARCH LAB WRIGHT-PATTERSON AFB OH MATERIALS AND MANUFACTURING DIRECTORATE, 2008.
- [47] G. Mohammed, "Diagnostic par Reconnaissance des formes: Application à la machine asynchrone."
- [48] E. Avci, "A new optimum feature extraction and classification method for speaker recognition: GWPNN," *Expert Syst. Appl.*, vol. 32, no. 2, pp. 485–498, 2007.
- [49] Z. Qunxiong and L. Chengfei, "Dimensionality Reduction with Input Training Neural Network and Its Application in Chemical Process Modelling1," *Chin. J. Chem. Eng.*, vol. 14, no. 5, pp. 597–603, 2006.
- [50] Y. Bennani, E. VIENNET, and S. GUÉRIF, "Réduction des dimensions des données en apprentissage artificiel," 2008.
- [51] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *J. Med. Syst.*, vol. 29, no. 6, pp. 647–660, 2005.
- [52] K. Chen, X. Li, F. Wang, T. Wang, and C. Wu, "Bearing fault diagnosis using wavelet analysis," presented at the Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2012 International Conference on, 2012, pp. 699–702.
- [53] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, 2013.
- [54] P. Borghesani, R. Ricci, S. Chatterton, and P. Pennacchi, "A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions," *Mech. Syst. Signal Process.*, vol. 38, no. 1, pp. 23–35, 2013.
- [55] K. Medjaher, F. Camci, and N. Zerhouni, "Feature extraction and evaluation for Health Assessment and Failure prognostics.," presented at the Proceedings of First European Conference of the Prognostics and Health Management Society, PHM-E'12., 2012, pp. 111–116.
- [56] N. Tandon and A. Choudhury, "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings," *Tribol. Int.*, vol. 32, no. 8, pp. 469–480, 1999.
- [57] G. Didier, "Modélisation et diagnostic de la machine asynchrone en présence de défaillances," *These Dr. L'université Henri Poincaré Nancy-I*, 2004.
- [58] W. Wang and P. McFadden, "Application of wavelets to gearbox vibration signals for fault detection," *J. Sound Vib.*, vol. 192, no. 5, pp. 927–939, 1996.
- [59] S. Guérif and Y. Bennani, "Sélection de Variables en Apprentissage Numérique Non Supervisé," *Univ. Paris*, 2007.
- [60] M. Campedel and E. Moulines, "Méthodologie de sélection de caractéristiques pour la classification d'images satellitaires.," presented at the CAP, 2005, pp. 107–108.



- [61] D. SERNANI, C. FRELICOT, and P. Courtellemont, "Une approche de type filtrage pour la sélection de variables. Application à la reconnaissance automatique de poissons," presented at the 20^e Colloque sur le traitement du signal et des images, FRA, 2005, 2005.
- [62] S. El Ferchichi, "Sélection et Extraction d'attributs pour les problèmes de classification," 2013.
- [63] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, no. 3, pp. 131–156, 1997.
- [64] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, no. Mar, pp. 1157–1182, 2003.
- [65] V. Kumar and S. Minz, "Feature selection," *SmartCR*, vol. 4, no. 3, pp. 211–229, 2014.
- [66] H. Mahfoud, "Optimisation des stratégies de maintenance d'un système complexe- Analyse bibliographique," presented at the Congres International du Génie Industriel et Management des Systemes (CIGIMS'15), 2015, p. 16.
- [67] G. Zwingelstein, *Diagnostic des défaillances: théorie et pratique pour les systèmes industriels*. Hermès, 1995.
- [68] M. Bengtsson, "Condition based maintenance system technology—Where is development heading," *Cond. BASED Maint. Syst. Investig. Tech. Const. Organ. Asp.*, vol. 55, 2004.
- [69] V. Tararykine, "Modélisation des flux d'information dans un système de e-maintenance.," 2005.
- [70] R. Kothamasu, S. H. Huang, and W. H. VerDuin, "System health monitoring and prognostics—a review of current paradigms and practices," *Int. J. Adv. Manuf. Technol.*, vol. 28, no. 9–10, pp. 1012–1024, 2006.
- [71] N. Ungureanu, M. Ungureanu, A. Cotetiu, B. Barisic, and S. Grozav, "Principles of the maintenance management," *Sci. Bull. Ser. C Fascicle Mech. Tribol. Mach. Manuf. Technol.*, vol. 24, p. 69, 2010.
- [72] N. AFNOR, "Terminologie de la maintenance," *NF-EN*, vol. 13306, pp. X60–319, 2001.
- [73] C. Sheut and L. Krajewski, "A decision model for corrective maintenance management," *Int. J. Prod. Res.*, vol. 32, no. 6, pp. 1365–1382, 1994.
- [74] R. Eisenmann and R. Eisenmann, "Applied condition monitoring," *Mach. Malfunct. Diagn. Correct.*, vol. 13, pp. 703–741, 1997.
- [75] W. W. Tiddens, A. J. J. Braaksma, and T. Tinga, "Framework for the selection of the optimal preventive maintenance approach," *Reliab. Eng. Syst. Saf.*, 2018.
- [76] S. Alaswad and Y. Xiang, "A review on condition-based maintenance optimization models for stochastically deteriorating system," *Reliab. Eng. Syst. Saf.*, vol. 157, pp. 54–63, 2017.
- [77] A. K. Jardine and A. H. Tsang, *Maintenance, replacement, and reliability: theory and applications*. CRC press, 2013.
- [78] E. Deloux, B. Castanier, and C. Bérenguer, "Predictive maintenance policy for a gradually deteriorating system subject to stress," *Reliab. Eng. Syst. Saf.*, vol. 94, no. 2, pp. 418–431, 2009.
- [79] I. Gertsbakh, *Reliability theory: with applications to preventive maintenance*. Springer, 2013.
- [80] A. Garg and S. Deshmukh, "Maintenance management: literature review and directions," *J. Qual. Maint. Eng.*, vol. 12, no. 3, pp. 205–238, 2006.
- [81] M. Lebold and M. Thurston, "Open standards for condition-based maintenance and prognostic systems," presented at the Maintenance and Reliability Conference (MARCON), 2001, pp. 6–9.



- [82] P. Wang and G. Vachtsevanos, "Fault prognostics using dynamic wavelet neural networks," *AI EDAM*, vol. 15, no. 04, pp. 349–365, 2001.
- [83] D. Djurdjanovic, J. Lee, and J. Ni, "Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction," *Adv. Eng. Inform.*, vol. 17, no. 3, pp. 109–125, 2003.
- [84] R. B. Chinnam and P. Baruah, "A neuro-fuzzy approach for estimating mean residual life in condition-based maintenance systems," *Int. J. Mater. Prod. Technol.*, vol. 20, no. 1–3, pp. 166–179, 2004.
- [85] R. Patrick-Aldaco, "AModel BASED FRAMEWORK FOR FAULT DIAGNOSIS AND PROGNOSIS OF DYNAMICAL SYSTEMS WITH AN APPLICATION TO HELICOPTER TRANSMISSIONS," 2007.
- [86] K. Medjaher, R. Gouriveau, and N. Zerhouni, "A procedure for failure prognostic in dynamic system.," presented at the 13th IFAC Symposium on Information Control Problems in Manufacturing, INCOM'09., 2009, pp. 1580–1585.
- [87] B. Giovannini, D. Gaston, and D. Gaston, "Etude bibliographique relative au développement de réactions chimiques non contrôlées dans l'industrie," *INIRIS CGR*, vol. 23414, 2000.
- [88] P. Trambouze and J.-P. Wauquier, *Le développement des procédés de raffinage et de pétrochimie*, vol. 3. Éditions Technip, 1975.
- [89] F. I. Khan and M. M. Haddara, "Risk-based maintenance (RBM): a quantitative approach for maintenance/inspection scheduling and planning," *J. Loss Prev. Process Ind.*, vol. 16, no. 6, pp. 561–573, 2003.
- [90] E. de Base pour l'énergétique, "Publications de Belkacem OULD BOUAMAMA1," *J. FI 1578*, vol. 17, no. 2–3, pp. 93–108.
- [91] B. O. Bouamama, M. Staroswiecki, and A. Samantaray, "Software for supervision system design in process engineering industry," in *Fault Detection, Supervision and Safety of Technical Processes 2006*, Elsevier, 2007, pp. 646–650.
- [92] J.-P. TAGUTCHOU, L. ESTEL, A. LEDOUX, and M. WANG, "influence of fly ashES on mass transfer."
- [93] H. Derbel, "Diagnostic à base de modèles des systèmes temporisés et d'une sous-classe de systèmes dynamiques hybrides," 2009.
- [94] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis: Part III: Process history based methods," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 327–346, 2003.
- [95] A. Laurent, "Sécurité des procédés chimiques," *Lavoisier Paris ISBN 2743006358*, vol. 21, p. 35, 2003.
- [96] A. K. Coker, *Modeling of chemical kinetics and reactor design*, vol. 1. Gulf Professional Publishing, 2001.
- [97] P. M. Frank and B. Köppen-Seliger, "New developments using AI in fault diagnosis," *Eng. Appl. Artif. Intell.*, vol. 10, no. 1, pp. 3–14, 1997.
- [98] R. Isermann, *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media, 2006.
- [99] S. Dash and V. Venkatasubramanian, "Challenges in the industrial applications of fault diagnostic systems," *Comput. Chem. Eng.*, vol. 24, no. 2, pp. 785–791, 2000.
- [100] A. S. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, no. 6, pp. 601–611, 1976.



- [101] E. Chow and A. Willsky, "Analytical redundancy and the design of robust failure detection systems," *IEEE Trans. Autom. Control*, vol. 29, no. 7, pp. 603–614, 1984.
- [102] M. Basseville, "Detecting changes in signals and systems—a survey," *Automatica*, vol. 24, no. 3, pp. 309–326, 1988.
- [103] R. J. Patton, "Fault-tolerant control: the 1997 situation," *IFAC Proc. Vol.*, vol. 30, no. 18, pp. 1029–1051, 1997.
- [104] A. Evsukoff, J. Montmain, and S. Gentil, "Dynamic model and causal knowledge-based fault detection and isolation," *IFAC Proc. Vol.*, vol. 30, no. 18, pp. 687–692, 1997.
- [105] P. Balle, M. Fischer, D. Fussel, O. Nelles, and R. Isermann, "Integrated control, diagnosis and reconfiguration of a heat exchanger," *IEEE Control Syst.*, vol. 18, no. 3, pp. 52–63, 1998.
- [106] P. M. Frank, "Analytical and qualitative model-based fault diagnosis—a survey and some new results," *Eur. J. Control*, vol. 2, no. 1, pp. 6–28, 1996.
- [107] K. Patan, *Artificial neural networks for the modelling and fault diagnosis of technical processes*. Springer, 2008.
- [108] S. Methnani, "Diagnostic, reconstruction et identification des défauts capteurs et actionneurs: application aux station d'épurations des eaux usées," 2012.
- [109] I. T. Jolliffe, "Principal component analysis: a beginner's guide—I. Introduction and application," *Weather*, vol. 45, no. 10, pp. 375–382, 1990.
- [110] N. Pessel, J. Balmat, F. Lafont, and J. Bannal, "An improved fault detection for the diagnosis," presented at the Proceeding of the 9th International Conference on Automatic Control, Modeling & Simulation, 2007.
- [111] M. Basseville, "Information criteria for residual generation and fault detection and isolation," *Automatica*, vol. 33, no. 5, pp. 783–803, 1997.
- [112] E. Assaf, R. Giordano, and C. Nascimento, "Thermal runaway of ethylene oxidation reactors: prevision through neuronal networks," *Chem. Eng. Sci.*, vol. 51, no. 11, pp. 3107–3112, 1996.
- [113] J. A. Barton and P. F. Nolan, "Runaway reactions in batch reactors," presented at the The Protection of Exothermic Reactors and Pressurised Storage Vessels, IChemE Symp. Series, 1984, vol. 85, pp. 13–21.
- [114] J. Gustin, "Retour d'expérience chimie fine," *Recomm. Ministère L'environnement Union Des*, 1996.
- [115] D. Yu, J. Gomm, and D. Williams, "Sensor fault diagnosis in a chemical process via RBF neural networks," *Control Eng. Pract.*, vol. 7, no. 1, pp. 49–55, 1999.
- [116] F. A. Cubillos and E. L. Lima, "Adaptive hybrid neural models for process control," *Comput. Chem. Eng.*, vol. 22, pp. S989–S992, 1998.
- [117] Y. Power and P. A. Bahri, "A two-step supervisory fault diagnosis framework," *Comput. Chem. Eng.*, vol. 28, no. 11, pp. 2131–2140, 2004.
- [118] V. Venkatasubramanian, R. Vaidyanathan, and Y. Yamamoto, "Process fault detection and diagnosis using neural networks—I. Steady-state processes," *Comput. Chem. Eng.*, vol. 14, no. 7, pp. 699–712, 1990.
- [119] J. C. Hoskins and D. Himmelblau, "Artificial neural network models of knowledge representation in chemical engineering," *Comput. Chem. Eng.*, vol. 12, no. 9–10, pp. 881–890, 1988.



- [120] K. Watanabe, S. Hirota, L. Hou, and D. Himmelblau, "Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks," *AIChE J.*, vol. 40, no. 5, pp. 839–848, 1994.
- [121] P. Hugo, J. Steinbach, and F. Stoessel, "Calculation of the maximum temperature in stirred tank reactors in case of a breakdown of cooling," presented at the Tenth International Symposium on Chemical Reaction Engineering, 1988, pp. 2147–2152.
- [122] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [123] A. Villemeur, "Sûreté de fonctionnement des systèmes industriels: fiabilité-facteurs humains, informatisation," 1988.
- [124] H. Farreny, "L'intelligence artificielle: Une discipline originale avec ses problèmes et ses techniques," presented at the Annales des télécommunications, 1989, vol. 44, pp. 207–217.
- [125] J. Korbicz, D. Uciński, A. Pieczyński, and G. Marczevska, "An integrated approach to fault detection and diagnosis in power plant," *IFAC Proc. Vol.*, vol. 27, no. 5, pp. 133–139, 1994.
- [126] M. Othman, M. Ali, and M. Kamsah, "Process fault detection using hierarchical artificial neural network diagnostic strategy," *J. Teknol.*, vol. 46, pp. 11–26, 2007.
- [127] S. Bailey, "FROM DESK-TOP TO PLANT FLOOR, A CRT IS THE CONTROL OPERATORS WINDOW ON THE PROCESS," *Control Eng.*, vol. 31, no. 6, pp. 86–90, 1984.
- [128] A. O. Molina, "Méthodologie pour le placement des capteurs à base de méthodes de classification en vue du diagnostic," 2005.
- [129] W. Chine, A. Mellit, V. Lughi, A. Malek, G. Sulligoi, and A. M. Pavan, "A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks," *Renew. Energy*, vol. 90, pp. 501–512, 2016.
- [130] Y. Chetouani, "Model selection and fault detection approach based on Bayes decision theory: Application to changes detection problem in a distillation column," *Process Saf. Environ. Prot.*, vol. 92, no. 3, pp. 215–223, 2014.
- [131] S. Rahman, F. Yusof, and M. Bakar, "The method review of neuro-fuzzy applications in fault detection and diagnosis system," *Int. J. Eng. Technol.*, vol. 10, no. 3, pp. 50–52, 2010.
- [132] V. Venkatasubramanian and K. Chan, "A neural network methodology for process fault diagnosis," *AIChE J.*, vol. 35, no. 12, pp. 1993–2002, 1989.
- [133] R. Li, J. H. Olson, and D. L. Chester, "Dynamic fault detection and diagnosis using neural networks," presented at the Intelligent Control, 1990. Proceedings., 5th IEEE International Symposium on, 1990, pp. 1169–1174.
- [134] W. R. Becraft and P. Lee, "An integrated neural network/expert system approach for fault diagnosis," *Comput. Chem. Eng.*, vol. 17, no. 10, pp. 1001–1014, 1993.
- [135] D. Leung and J. Romagnoli, "Dynamic probabilistic model-based expert system for fault diagnosis," *Comput. Chem. Eng.*, vol. 24, no. 11, pp. 2473–2492, 2000.
- [136] D. R. Baughman and Y. A. Liu, *Neural networks in bioprocessing and chemical engineering*. Academic press, 2014.
- [137] D. M. Himmelblau, *Fault detection and diagnosis in chemical and petrochemical processes*, vol. 8. Elsevier Science Ltd, 1978.
- [138] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-based fault diagnosis in dynamic systems using identification techniques*. Springer Science & Business Media, 2013.



- [139] R. MAHADOU, "DIAGNOSTIC INDUSTRIEL PAR NEURO-FLOU-APPLICATION A UN SYSTEME DE PRODUCTION," 2007.
- [140] R. J. Schalkoff, *Artificial neural networks*, vol. 1. McGraw-Hill New York, 1997.
- [141] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," presented at the Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 249–256.
- [142] J. A. Hertz, *Introduction to the theory of neural computation*. CRC Press, 2018.
- [143] G. Dreyfus, J. Martinez, M. Samuelides, M. Gordon, and F. BADRAN, "S. THIRIA et L. HERAULT," *Réseaux Neurones Méthodologie Appl. Paris Ed. Eyrolles*, 2002.
- [144] F. Rosenblatt, *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.
- [145] L. Noriega, "Multilayer perceptron tutorial," *Sch. Comput. Staffs. Univ.*, 2005.
- [146] G. E. Hinton and S. J. Nowlan, "The bootstrap Widrow-Hoff rule as a cluster-formation algorithm," *Neural Comput.*, vol. 2, no. 3, pp. 355–362, 1990.
- [147] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, p. 533, 1986.
- [148] S. E. Vt and Y. C. Shin, "Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems," *IEEE Trans. Neural Netw.*, vol. 5, no. 4, pp. 594–603, 1994.
- [149] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man-Mach. Stud.*, vol. 7, no. 1, pp. 1–13, 1975.
- [150] J. B. Kiszka, M. M. Gupta, and P. N. Nikiforuk, "Energetic stability of fuzzy dynamic systems," *IEEE Trans. Syst. Man Cybern.*, no. 6, pp. 783–792, 1985.
- [151] J. M. Kohcielnny, "Application of fuzzy logic for fault isolation in a three-tank system," *IFAC Proc. Vol.*, vol. 32, no. 2, pp. 7754–7759, 1999.
- [152] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, vol. 4. Prentice hall New Jersey, 1995.
- [153] W. Pedrycz and F. Gomide, *An introduction to fuzzy sets: analysis and design*. Mit Press, 1998.
- [154] R. BABUŠKA, "FUZZY AND NEURAL CONTROL DISC Course Lecture Notes (October 2001)," 2001.
- [155] L. A. Zadeh, "Fuzzy logic," *Computer*, vol. 21, no. 4, pp. 83–93, 1988.
- [156] O. Castillo, P. Melin, J. Kacprzyk, and W. Pedrycz, "Type-2 fuzzy logic: theory and applications," presented at the Granular Computing, 2007. GRC 2007. IEEE International Conference on, 2007, pp. 145–145.
- [157] L. A. Zadeh, "Soft computing and fuzzy logic," *IEEE Softw.*, vol. 11, no. 6, pp. 48–56, 1994.
- [158] W. Van Leekwijck and E. E. Kerre, "Defuzzification: criteria and classification," *Fuzzy Sets Syst.*, vol. 108, no. 2, pp. 159–178, 1999.
- [159] K. BEDOUD, "Modélisation et diagnostic de défaillance d'une électrovanne pneumatique," 2010.
- [160] Y. Chetouani, "A neural network approach for the real-time detection of faults," *Stoch. Environ. Res. Risk Assess.*, vol. 22, no. 3, pp. 339–349, 2008.
- [161] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2–3, pp. 191–203, 1984.
- [162] Y. Chetouani, "Using Artificial Neural networks for the modelling of a distillation column," *IJCSA*, vol. 4, no. 3, pp. 119–133, 2007.



- [163] A. Muller, M.-C. Suhner, and B. lung, "Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system," *Reliab. Eng. Syst. Saf.*, vol. 93, no. 2, pp. 234–253, 2008.
- [164] F. Ciarapica and G. Giacchetta, "Managing the condition-based maintenance of a combined-cycle power plant: an approach using soft computing techniques," *J. Loss Prev. Process Ind.*, vol. 19, no. 4, pp. 316–325, 2006.
- [165] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [166] M. Terminology, "European Standard," *CEN Eur. Comm. Stand. EN*, vol. 13306, 2001.
- [167] K. Medjaher, R. Gouriveau, and N. Zerhouni, "A procedure for failure prognostic in dynamic systems," *IFAC Proc. Vol.*, vol. 42, no. 4, pp. 1597–1602, 2009.
- [168] K. W. Hipel and A. I. McLeod, *Time series modelling of water resources and environmental systems*. Elsevier, 1994.
- [169] D. P. Lettenmaier and E. F. Wood, "Hydrologic forecasting," *Handb. Hydrol.*, pp. 26–1, 1993.
- [170] J. Luo, M. Namburu, K. Pattipati, L. Qiao, M. Kawamoto, and S. Chigusa, "Model-based prognostic techniques [maintenance applications]," presented at the AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference. Proceedings, 2003, pp. 330–340.
- [171] J. Yan, J. Lee, and M. Koç, "Predictive algorithm for machine degradation detection using logistic regression," presented at the Fifth International Conference on Managing Innovations in Manufacturing (MIM), 2002, pp. 172–178.
- [172] D. Lin and V. Makis, "Recursive filters for a partially observable system subject to random failure," *Adv. Appl. Probab.*, vol. 35, no. 1, pp. 207–227, 2003.
- [173] H. Sohn, C. R. Farrar, F. M. Hemez, G. Park, A. N. Robertson, and T. O. Williams, "A coupled approach to developing damage prognosis solutions," presented at the Key Engineering Materials, 2003, vol. 245, pp. 289–306.
- [174] G. Vachtsevanos and P. Wang, "Fault prognosis using dynamic wavelet neural networks," presented at the AUTOTESTCON Proceedings, 2001. IEEE Systems Readiness Technology Conference, 2001, pp. 857–870.
- [175] C. S. Byington and M. J. Roemer, "Prognostic enhancements to diagnostic systems for improved condition-based maintenance [military aircraft]," presented at the Aerospace Conference Proceedings, 2002. IEEE, 2002, vol. 6, pp. 6–2815.
- [176] A. Mathur, K. F. Cavanaugh, K. R. Pattipati, P. K. Willett, and T. R. Galie, "Reasoning and modeling systems in diagnosis and prognosis," presented at the Aerospace/Defense Sensing, Simulation, and Controls, 2001, pp. 194–203.
- [177] O. E. Dragomir, R. Gouriveau, N. Zerhouni, and F. Dragomir, "Framework for a distributed and hybrid prognostic system.," presented at the 4th IFAC Conference on Management and Control of Production and Logistics, MCPL'2007., 2007, vol. 3, pp. 431–436.
- [178] O. Dragomir, R. Gouriveau, and N. Zerhouni, "Adaptive Neuro-Fuzzy inference system for mid term prognostic error stabilization.," presented at the International Conference on Computers, Communications & Control, ICCCC'08., 2008, vol. 3, pp. 271–276.
- [179] J. G. De Gooijer and R. J. Hyndman, "25 years of time series forecasting," *Int. J. Forecast.*, vol. 22, no. 3, pp. 443–473, 2006.
- [180] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel, "Metrics for offline evaluation of prognostic performance," *Int. J. Progn. Health Manag.*, vol. 1, no. 1, pp. 4–23, 2010.



- [181] R. Adeline, R. Gouriveau, and N. Zerhouni, "Pronostic de défaillances: Maîtrise de l'erreur de prédiction.," presented at the 7ème Conférence Internationale de MObilisation et SIMulation, MOSIM'08., 2008, vol. 1, p. 10–pages.
- [182] K. Goebel and P. Bonissone, "Prognostic information fusion for constant load systems," presented at the Information Fusion, 2005 8th International Conference on, 2005, vol. 2, p. 9–pp.
- [183] G. J. Vachtsevanos, F. Lewis, A. Hess, and B. Wu, *Intelligent fault diagnosis and prognosis for engineering systems*. Wiley Online Library, 2006.
- [184] S. Létourneau, C. Yang, C. Drummond, E. Scarlett, J. Valdés, and M. Zaluski, "A domain independent data mining methodology for prognostics," presented at the Essential technologies for successful prognostics: proceedings of the 59th Meeting of the Society for Machinery Failure Prevention Technology, Virginia Beach, Virginia, April 18-21, 2005, 2005.
- [185] W. Caesarendra, A. Widodo, H. T. Pham, and B.-S. Yang, "Machine degradation prognostic based on RVM and ARMA/GARCH model for bearing fault simulated data," presented at the Prognostics and Health Management Conference, 2010. PHM'10., 2010, pp. 1–6.
- [186] D. Tobon-Mejia, K. Medjaher, and N. Zerhouni, "CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks," *Mech. Syst. Signal Process.*, vol. 28, pp. 167–182, 2012.
- [187] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Health assessment and life prediction of cutting tools based on support vector regression," *J. Intell. Manuf.*, vol. 26, no. 2, pp. 213–223, 2015.
- [188] E. Ramasso and R. Gouriveau, "Prognostics in switching systems: Evidential Markovian classification of real-time neuro-fuzzy predictions," presented at the Prognostics and Health Management Conference, 2010. PHM'10., 2010, pp. 1–10.
- [189] E. Ramasso, M. Rombaut, and N. Zerhouni, "Joint prediction of continuous and discrete states in time-series based on belief functions," *IEEE Trans. Cybern.*, vol. 43, no. 1, pp. 37–50, 2013.
- [190] M. El Koujok, "Contribution au pronostic industriel: intégration de la confiance à un modèle prédictif neuro-flou.," 2010.
- [191] O. E. Dragomir, "Contribution au pronostic de défaillances par réseau neuro-flou: maîtrise de l'erreur de prédiction," 2008.
- [192] R. Yam, P. Tse, L. Li, and P. Tu, "Intelligent predictive decision support system for condition-based maintenance," *Int. J. Adv. Manuf. Technol.*, vol. 17, no. 5, pp. 383–391, 2001.
- [193] C. S. Byington, M. J. Roemer, M. J. Watson, T. R. Galie, and C. Savage, "Prognostic enhancements to diagnostic systems (PEDS) applied to shipboard power generation systems," presented at the ASME Turbo Expo 2004: Power for Land, Sea, and Air, 2004, pp. 825–833.
- [194] A. Muller, A. C. Marquez, and B. lung, "On the concept of e-maintenance: Review and current research," *Reliab. Eng. Syst. Saf.*, vol. 93, no. 8, pp. 1165–1187, 2008.
- [195] A. Muller, "Contribution à la maintenance prévisionnelle des systèmes de production par la formalisation d'un processus de pronostic," 2005.
- [196] D. Chelidze and J. P. Cusumano, "A dynamical systems approach to failure prognosis," *J. Vib. Acoust.*, vol. 126, no. 1, pp. 2–8, 2004.



- [197] M. Luo *et al.*, "Model-based fault diagnosis/prognosis for wheeled mobile robots: a review," presented at the Industrial Electronics Society, 2005. IECON 2005. 31st Annual Conference of IEEE, 2005, p. 6–pp.
- [198] M. I. Rafiq, M. K. Chryssanthopoulos, and T. Onoufriou, "Performance updating of concrete bridges using proactive health monitoring methods," *Reliab. Eng. Syst. Saf.*, vol. 86, no. 3, pp. 247–256, 2004.
- [199] N.-H. Kim, D. An, and J.-H. Choi, "Data-Driven Prognostics," in *Prognostics and Health Management of Engineering Systems*, Springer, 2017, pp. 179–241.
- [200] M. Schwabacher and K. Goebel, "A survey of artificial intelligence for prognostics," presented at the Aai fall symposium, 2007, pp. 107–114.
- [201] X. Zhang, R. Xu, C. Kwan, S. Y. Liang, Q. Xie, and L. Haynes, "An integrated approach to bearing fault diagnostics and prognostics," presented at the American Control Conference, 2005. Proceedings of the 2005, 2005, pp. 2750–2755.
- [202] S. Charbonnier, C. Garcia-Beltan, C. Cadet, and S. Gentil, "Trends extraction and analysis for complex system monitoring and decision support," *Eng. Appl. Artif. Intell.*, vol. 18, no. 1, pp. 21–36, 2005.
- [203] M. Dong and D. He, "Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis," *Eur. J. Oper. Res.*, vol. 178, no. 3, pp. 858–878, 2007.
- [204] C. Elbers and G. Ridder, "True and spurious duration dependence: The identifiability of the proportional hazard model," *Rev. Econ. Stud.*, vol. 49, no. 3, pp. 403–409, 1982.
- [205] D. Frechtling, *Forecasting tourism demand*. Routledge, 2012.
- [206] O. Dragomir, F. Dragomir, and E. Mincă, "An application oriented guideline for choosing a prognostic tool," presented at the AIP Conference Proceedings, 2009, vol. 1107, pp. 257–262.
- [207] R. Silva *et al.*, "Proton exchange membrane fuel cell degradation prediction based on adaptive neuro-fuzzy inference systems," *Int. J. Hydrog. Energy*, vol. 39, no. 21, pp. 11128–11144, 2014.
- [208] A. Lapedes and R. Farber, "Nonlinear signal processing using neural networks: Prediction and system modelling," 1987.
- [209] H. White, "Economic prediction using neural networks: The case of IBM daily stock returns," 1988.
- [210] R. Sharda and R. Patil, "Neural networks as forecasting experts: an empirical test," presented at the Proceedings of the International Joint Conference on Neural Networks, 1990, vol. 2, pp. 491–494.
- [211] A. S. Lapedes and R. M. Farber, "How neural nets work," presented at the Neural information processing systems, 1988, pp. 442–456.
- [212] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35–62, 1998.
- [213] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *Power Syst. IEEE Trans. On*, vol. 16, no. 1, pp. 44–55, 2001.
- [214] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *Syst. Man Cybern. IEEE Trans. On*, no. 1, pp. 28–44, 1973.



- [215] E. Gorrostieta and C. Pedraza, "Neuro fuzzy modeling of control systems," presented at the Electronics, Communications and Computers, 2006. CONIELECOMP 2006. 16th International Conference on, 2006, pp. 23–23.
- [216] D. Nauck and R. Kruse, "What are neuro-fuzzy classifiers," presented at the Proc. Seventh International Fuzzy Systems Association World Congress IFSA, 1997, vol. 97, pp. 228–233.
- [217] N. Palluat, D. Racoceanu, and N. Zerhouni, "A neuro-fuzzy monitoring system: Application to flexible production systems," *Comput. Ind.*, vol. 57, no. 6, pp. 528–538, 2006.
- [218] W. James *et al.*, "Artificial Intelligence in predicting bladder cancer outcome: A comparison of Neuro-Fuzzy and Artificial Neural Networks," *Clin. Cancer Res.*, vol. 9, pp. 4172–4177, 2003.
- [219] W. Q. Wang, M. F. Golnaraghi, and F. Ismail, "Prognosis of machine health condition using neuro-fuzzy systems," *Mech. Syst. Signal Process.*, vol. 18, no. 4, pp. 813–831, 2004.
- [220] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.
- [221] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 115–118, 2004.
- [222] K. Javed, "A robust & reliable Data-driven prognostics approach based on extreme learning machine and fuzzy clustering.," 2014.
- [223] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.
- [224] Z. Tian, "An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring," *J. Intell. Manuf.*, vol. 23, no. 2, pp. 227–237, 2012.
- [225] H. Sahraie, A. Ghaffari, and M. Amidpour, "Adaptive Neuro-Fuzzy Inference System (ANFIS) Approach to Evaluate the Debutanizer Top Product," *Int. J. Eng.*, vol. 6, no. 02, p. 8269, 2014.
- [226] R. Sivakumar and K. Balu, "ANFIS based distillation column control," *Int. J. Comput. Appl. Spec. Issue Evol. Comput.*, vol. 2, pp. 67–73, 2010.
- [227] Q. Zhu, Y. Wang, D. Zhao, S. Li, and S. A. Billings, "Review of rational (total) nonlinear dynamic system modelling, identification, and control," *Int. J. Syst. Sci.*, vol. 46, no. 12, pp. 2122–2133, 2015.
- [228] Y. Chetouani, "Modeling and prediction of the dynamic behavior in a reactor-exchanger using NARMAX neural structure," *Chem. Eng. Commun.*, vol. 194, no. 5, pp. 691–705, 2007.
- [229] Y. Chetouani, "Use of cumulative sum (CUSUM) test for detecting abrupt changes in the process dynamics," *Int. J. Reliab. Qual. Saf. Eng.*, vol. 14, no. 01, pp. 65–80, 2007.
- [230] Y. Rui and A. El-Keib, "A review of ANN-based short-term load forecasting models," presented at the System Theory, 1995., Proceedings of the Twenty-Seventh Southeastern Symposium on, 1995, pp. 78–82.
- [231] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.
- [232] L. Rutkowski and L. Rutkowski, *New soft computing techniques for system modeling, pattern classification and image processing*. Springer, 2004.
- [233] S. Hugueny, D. A. Clifton, and L. Tarassenko, "Probabilistic patient monitoring with multivariate, multimodal extreme value theory," presented at the International Joint Conference on Biomedical Engineering Systems and Technologies, 2010, pp. 199–211.



- [234] H. Sahraie, A. Ghaffari, and M. Amidpour, "Adaptive Neuro-Fuzzy Inference System (ANFIS) Approach to Evaluate the Debutanizer Top Product," *Int. J. Eng.*, vol. 6, no. 02, p. 8269, 2014.
- [235] H. C. Pusey, "Turbomachinery condition monitoring and failure prognosis," *Sound Vib.*, vol. 41, no. 3, pp. 10–15, 2007.
- [236] Q. Gaudel, "Approche intégrée de diagnostic et de pronostic pour la gestion de santé des systèmes hybrides sous incertitude," 2016.
- [237] A. Zaidi, B. O. Bouamama, and M. Tagina, "Bond graph and bayesian networks for reliable diagnosis," *Int. J. Simul. Syst. Sci. Technol.*, vol. 12, no. 5, pp. 1240–1253, 2012.
- [238] H. Fan, C. Hu, M. Chen, and D. Zhou, "Cooperative predictive maintenance of repairable systems with dependent failure modes and resource constraint," *IEEE Trans. Reliab.*, vol. 60, no. 1, pp. 144–157, 2011.
- [239] Y. Chen, S. Alspaugh, and R. Katz, "Interactive analytical processing in big data systems: A cross-industry study of mapreduce workloads," *Proc. VLDB Endow.*, vol. 5, no. 12, pp. 1802–1813, 2012.
- [240] R. G. Batruni, "Crest factor reduction," Feb. 2013.
- [241] J. D. Brown, "Skewness and kurtosis," *Shiken JALT Test. Eval. SIG*, 1997.
- [242] P. Bonato, S. H. Roy, M. Knaflitz, and C. J. De Luca, "Time-frequency parameters of the surface myoelectric signal for assessing muscle fatigue during cyclic dynamic contractions," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 7, pp. 745–753, 2001.
- [243] B. Boashash and M. Mesbah, "Signal enhancement by time-frequency peak filtering," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 929–937, 2004.
- [244] J. Keyantash and J. A. Dracup, "The quantification of drought: an evaluation of drought indices," *Bull. Am. Meteorol. Soc.*, vol. 83, no. 8, pp. 1167–1180, 2002.
- [245] A. Zhang, B. Yang, and L. Huang, "Feature extraction of EEG signals using power spectral entropy," presented at the Biomedical Engineering and Informatics, 2008. BMEI 2008. International Conference on, 2008, vol. 2, pp. 435–439.
- [246] M. El Badaoui, F. Guillet, and J. Daniere, "New applications of the real cepstrum to gear signals, including the definition of a robust fault indicator," *Mech. Syst. Signal Process.*, vol. 18, no. 5, pp. 1031–1046, 2004.
- [247] Z. Peng and F. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mech. Syst. Signal Process.*, vol. 18, no. 2, pp. 199–221, 2004.
- [248] D. Alamedine, M. Khalil, and C. Marque, "Parameters extraction and monitoring in uterine EMG signals. Detection of preterm deliveries," *IRBM*, vol. 34, no. 4, pp. 322–325, 2013.
- [249] S. X. Yang and C. Luo, "A neural network approach to complete coverage path planning," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 34, no. 1, pp. 718–724, 2004.
- [250] K. Watanabe, I. Matsuura, M. Abe, M. Kubota, and D. Himmelblau, "Incipient fault diagnosis of chemical processes via artificial neural networks," *AIChE J.*, vol. 35, no. 11, pp. 1803–1812, 1989.
- [251] J. Hoskins, K. Kaliyur, and D. M. Himmelblau, "Fault diagnosis in complex chemical plants using artificial neural networks," *AIChE J.*, vol. 37, no. 1, pp. 137–141, 1991.
- [252] S. Wang *et al.*, "Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection," *Int. J. Imaging Syst. Technol.*, vol. 25, no. 2, pp. 153–164, 2015.



Normandie Université