Desalination 182 (2005) 131–142

DESALINATION

www.elsevier.com/locate/desal

Faults analysis for MSF plants

E.E. Tarifa[ab]*, S. Franco Domínguez[a], D. Humana[a], S.L. Martínez[a], A.F. Núñez[a], N.J. Scenna[bc]

[a]*National University of Jujuy, Gorriti 237, 4600 San Salvador de Jujuy, Argentina*
Tel. +54 388 4221588; Fax 54 388 4221588; email: eetarifa@arnet.com.ar
[b]*CONICET, Avellaneda 3657, 3000 Santa Fe, Argentina*
[c]*INGAR, Avellaneda 3657, 3000 Santa Fe, Argentina*

Received 21 February 2005; accepted 29 March 2005

Abstract

This work analyses the faults consequences in MSF plants. This analysis will help to design an optimal action plans in order to mitigate the faults effects. The nature of the faults consequences are studied at two levels: process level and mathematical level. At the first one, faults are grouped according its effects over the process state; at the second level, they are grouped according the mathematical features of its consequences analyzed as functions of time (i.e. signals).

The study was carried out by using a dynamic simulator for MSF desalination plants. It takes into account the heaters and stages dynamic, hydraulic, standard instrumentation and control systems. This simulator was developed to study the effects of faults that may affect a MSF plant. In order to extend the results scope, the simulator allows the modification of topology and parameters into a wide range. Indeed, it is possible to change the number of stages belonging to the recovery and rejection sections, controllers' parameters (set point, integral time and gain), valves size, pumps characteristics, sea water conditions, stages and heater dimensions, etc. Since faults simulation is the main simulator goal, the model and its resolution were carefully designed to enhance stability and speed. The user can select the fault to simulate among a set of possible faults (fault in controllers, sensors, pumps, etc.), and can specify the activating time (at which the fault starts), the developing time (time elapsed from the fault start up until the fault reaches its maximum magnitude) and the fault magnitude. Thus, it is possible to simulate step and ramp perturbations. The simulator was tested with data from real plants and it has shown a good performance.

Keywords: Fault diagnosis; Process supervision; Artificial neural networks; MSF

---

*Corresponding author.

Presented at the Conference on Desalination and the Environment, Santa Margherita, Italy, 22–26 May 2005. European Desalination Society.

0011-9164/05/$– See front matter © 2005 Elsevier B.V. All rights reserved

1. Introduction

This work studies the faults consequences in MSF (Multi-Stage Flash) plants. This analysis will help to design an optimal action plan, for each abnormal situation, in order to mitigate the faults effects. Those plans may be given to the operator as a manual of emergency procedures, or may form part of the knowledge base of a supervision system [1]. Fig. 1 shows the structure of the supervision system considered in this work.

The detector processes the sensors signals provided by the acquisition system in order to normalize them. Those normalized signals are then analysed by the diagnostic system and the evaluator. The diagnostic system processes the normalized signals provided by the detector to determine whether the process state is normal or not. In the last case, the diagnostic system determines the cause of the abnormal process state; this cause may be: a damaged piece of equipment, an operator's action, an alteration of the process input, or any other perturbation. Here, all those causes are considered as faults [2]. Finally, the evaluator elaborates the best action plan according to the detector and diagnostic system outputs.

When a given fault affects a plant, a process parameter or variable is directly affected by it; then, that original perturbation propagates itself throughout the plant taking the process variables away from their normal values. The evolution of the process variables is a function of the process and the fault characteristics. In this way, there will be different evolutions depending on which process parameter or variable is directly affected by the fault, and on the form of this original perturbation. In this work, the form of that original perturbation is established by three parameters: the activating time *ta*, the magnitude *M* and the developing time *td*. The first one is the time at which the fault happens. The second one specifies the maximum magnitude by which the normal value of the process parameter or variable directly affected by the fault will be changed. The last one designates the time elapsed between the perturbation start-up with null magnitude and the perturbation full development when it reaches the maximum magnitude [3].

For each sampling time, every Δt, the plant readings are processed by the detector whose mission is to compare the current vector of readings X (with an element for each measured variable) with the respective vector of values of the normal evolution Xn. The detector output is a vector of normalized deviations δX. This vector is sent to the

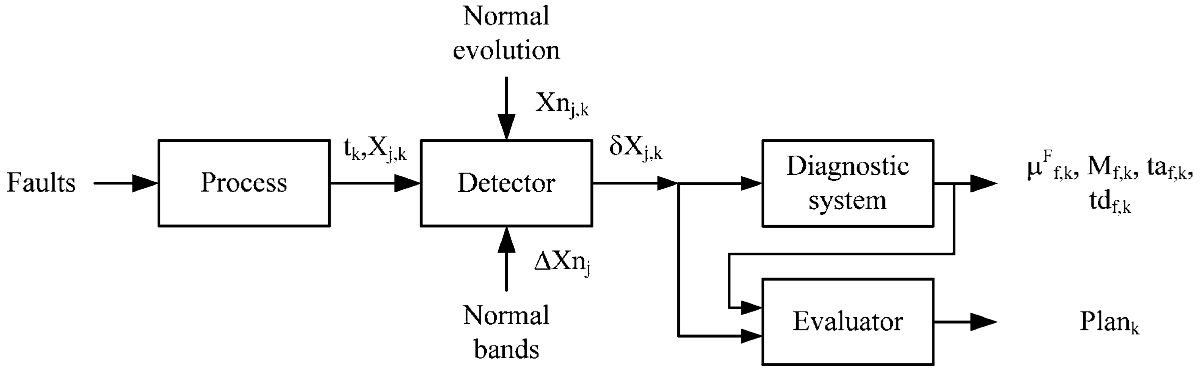

Fig. 1. Supervision system structure.

diagnostic system, which determines for each potential fault the magnitude *M*, the activating time *ta*, the developing time *td* and the certainty supporting the considered fault μ^F (a number between 0 and 1). At a given time, the certainty supporting a particular fault indicates the degree of matching between the already observed process evolution and the expected one for that fault. A null value means that there is not matching at all; conversely, a value equal to one implies a full matching; intermediate values indicate intermediate degrees of matching. Therefore, the fault with maximum μ^F is the most serious candidate to be the real fault affecting the plant [4,5].

On the other hand, the evaluator processes both the detector and diagnostic system outputs in order to evaluate the gravity of the present process state. Then the best action plan, according to the situation, will be selected–or generated–and presented to the operator. This plan can be divided in two parts: the first one aimed to the consequences, and the second one aimed to the causes. The first part considers preventive actions to avoid further consequences, whereas the second part takes into account corrective actions to eliminate the fault affecting the process.

To accomplish its mission, the evaluator analyses the symptoms (i.e. abnormal values) shown by the process, and immediately elaborates a plan of preventive actions. As the process evolves, the diagnostic system may count with enough information to identify the fault. This new datum may be used by the evaluator to reformulate the preventive plan, and to elaborate a corrective one.

The design of suitable action plan demands a deep knowledge of the faults consequences in the process as well as appropriate criterions to evaluate their gravity. This work present criterions and tools to perform that evaluation. Those elements were tested by apply them to a MSF plant. A dynamic simulator was utilized to emulate the acquisition system outputs. As result, a complete ranking of fault gravity is obtained, where the faults are classify from not dangerous (they only require corrective actions or no action at all) to extremely dangerous (they require both corrective and preventive actions).

2. The studied process

Fig. 2 shows the MSF plant structure adopted in this work. That plant is a series of flash unities (stages) where sea water is evaporated to obtain distilled water [6]. Additional vapour streams are produced from partial evaporation of condensate. To reduce corrosion problems, the heater temperature must not be higher than 120°C and the sea water stream must be properly treated. To minimize costs, a fraction of the concentrated brine is recycled. This recycled flow is limited by the salt maximum admissible concentration, which depends on materials and the equipment useful life. The brine and product evolve in counter-current flow in relation to the recycled brine. The temperature difference is the driver force to produce vapour condensation and, simultaneously, feed preheating, which take place in the condenser that is located at the top of each stage. Vapour condensation and a vacuum pump preserve the pressure gradients among successive stages.

The plant has *N* stages; the first *M* ones belong to the recovery section, whereas the remaining ones belong to the rejection section. There are also six $P + I$ controllers, which enable to set the operation conditions for the heater, feed, recycle and last stage level.

A dynamic simulator was implemented by using Delphi 5.0 (Inprise Borland), a

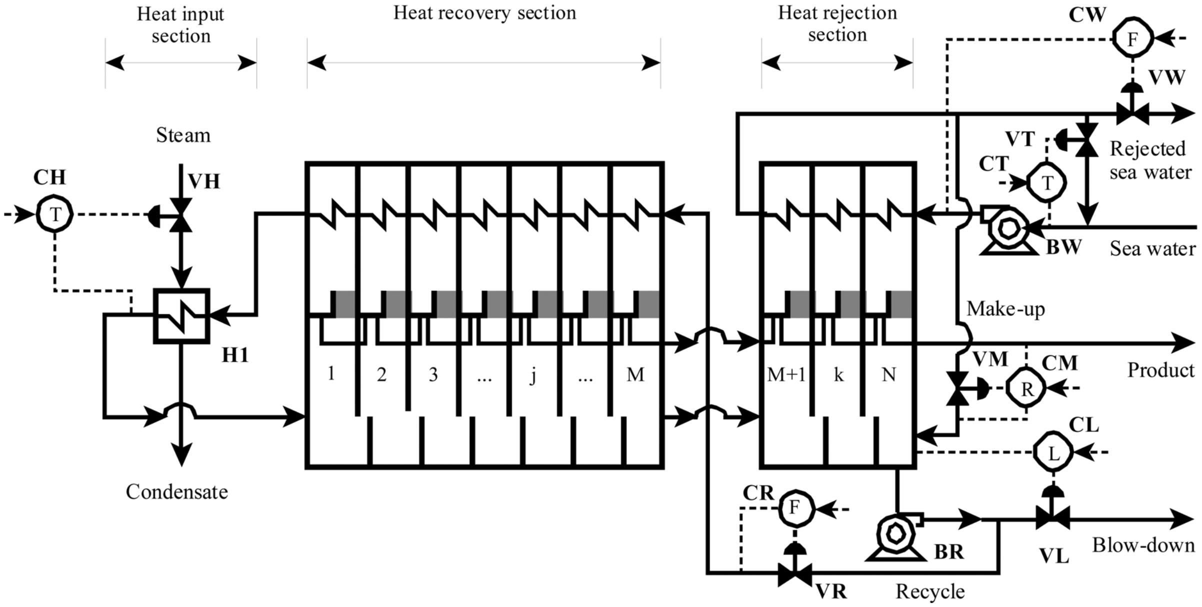

Fig. 2. MSF plant. Controllers: *T*, temperature; *L*, level; *F*, flow; *R*, rate.

computer visual language [3]. The simulator interface allows the full specification for both the process conditions (i.e. topology, equipment parameters, operation conditions, etc.) and the simulation conditions (i.e. initial state, simulation final time, faults, etc.). It also plots outputs and can export them in traditional spreadsheet format.

To evaluate the simulator performance the system presented by Thomas et al. [6] was simulated. This system has 15 stages in the recovery section and 3 in the rejection section. The simulator showed a good performance during the test; so, it was used to simulate the MSF plant in this work. In this way, the vector X can be obtained for each studied fault.

3. The detector

When a fault happens, the process state evolves away from its normal state (continuous processes) or normal evolution (batch processes). In this work, each possible abnormal evolution is called dynamic state. For a given fault *f*, infinite dynamic states exist since infinite possible magnitudes, activating times and developing times exist. In other words, a dynamic state is associated to each possible (*f*, *M*, *ta*, *td*) combination, and the possible combinations are infinite.

The mission of the detector is to detect the presence of a dynamic state, and to calculate the normalized deviations vector δX. In order to do that, a data acquisition system (or the simulator) provides every Δt the values of the measured variables of the process X. The corresponding sampling time for the reading k is $t_k = k{\cdot}\Delta t$. In addition, the user must provide the normal values expected for the measured variables Xn. This vector will be constant for continuous processes, and it will be a function of the time for batch processes. With these values, the quantitative deviations vector is calculated as $\Delta X = X\text{-}Xn$. By working with deviations, the proposed method does not depend on the type of process (continuous or batch).

The normalized deviation for the variable j at the time t_k is defined by:

$$\delta X_{j,k} = \frac{\Delta X_{j,k}}{\Delta Xn_j} \tag{1}$$

where $\Delta Xn_j > 0$ defines the interval ($-\Delta Xn_j$, ΔXn_j) as the band of normality for the quantitative deviation of the variable j. That means, the value of a given variable is classified as normal if ΔX belongs to that band or, which is equivalent, δX belongs to $(-1, 1)$. The selection of these values must consider the trade-off between the sensibility and reliability of the system. By working with normalized deviations, which are dimensionless numbers, the proposed method does not depend on the type of variable (pressure, temperature, etc.).

The detector calculates the vector δX for every sampling time. While the process state is normal, the absolute value of every element of that vector is less than 1, because every process variable is inside its corresponding band of normality. When a variable goes beyond its band, the process state is no longer normal, and a dynamic state is detected. Therefore, the detecting or observation time *to* is defined as the first time at which the absolute value of some element of the vector δX is equal or greater than 1. At that time, the diagnostic system is turned on, and so the search for the fault that affects the process begins [4,5].

4. The diagnostic system and the evaluation system development

Both the diagnostic system as the evaluation system development comprises several stages. They are:

1. *Determination of the potential faults*: The systems are prepared to deal only the faults determined by this analysis.
2. *Selection of sensors to be supervised by the systems*: Only the sensors that provide relevant information are selected.
3. *Systems building*: They can be an expert system, or a system based on qualitative or quantitative models.
4. *Systems evaluation*: Generally, this stage is carried out by using a process simulator that replaces the plant.

The following sections are concerned about each of those stages.

5. Potential faults

The first step in the development of the supervision system is the determination of the scope of application. This scope is given by the list of faults that must be recognised by the system and the desired format for the report.

Due to their characteristics and importance, some faults must be studied by *ad hoc* systems; this is the case of heat exchanger fouling [7]. In this work, except the ‘Recovery low’ and ‘Rejection low’, all the faults modelled in the simulator MSF 2000 [3] are considered.

The desired format for the report is the certainty μ^F, a number belonging to [0, 1], for each potential fault. This number represents the certainty about the corresponding fault is affecting the plant. The higher that value, the higher is the certainty of the affirmation.

6. Selection of sensors

For a given scope of application, not all variables provide useful information. Irrelevant data may complicate the structure and affect the performance of the supervision system. To identify the relevant variables, the outputs of the simulator MSF 2000 [3] are analysed. Once identified those variables, only the corresponding sensors are selected to be supervised by the supervision system. Tarifa and Scenna [5] perform the selection of sensor for the considered MSF plant. From the analysis of more than 270 dynamic states and the normal bands showed in Table 1, they selected the sensors showed in Table 2.

7. Criterions of faults gravity

In order to be able to assign a gravity grade to each potential faults, it is required to establish how the faults gravity will be measured. Besides the nature of the fault and process, the gravity of a given fault depends on the task for which that information is generated. For instance, a given fault can be quite serious for the operator that must control de process, whereas the same fault can be a trivial one for the personal that must repair the piece of equipment that has failed. For this reason, there are almost as criterions as operators in the plant. In this work, three particular criterions are outlined: the integral, the derivative, and the classification. The two firsts consider the process point of view, whereas the last one considers the mathematic point of view.

Table 1
Normal bands

2°C for temperatures.
500 tn/h for flow rates.
5 cm for levels.
2% for outputs of controllers.
0.1 for Rmus[frac] (set point of make-up controller).
Others variables are not measured.

7.1. The integral criterion

For a given fault, this criterion attempts to measure the effects over the process states and equipment. The larger and more lasting the effects, the more difficult will be the elimination of them to return to the normal conditions. Thus, this information is fairly interesting for the operator that must bring back the process state and equipment to the normal conditions. Fig. 3 shows how to calculate an index according this criterion. Each variable contributes to it with the area under the curve δX that is outside of the band of normality. Finally, all contributions are added to obtain the index.

For a given dynamic state i, the index can be evaluate by processing the result of the corresponding simulation with activating time null ($ta = 0$); that is:

$$Iit_i = \sum_j Ii_{i,j} \tag{2}$$

$$Ii_{i,j} = \int_0^{tf} fe\left(\delta X(t)_{i,j}\right) dt \simeq \Delta t \sum_k fe\left(\delta X_{i,j,k}\right) \tag{3}$$

Table 2
Selected sensors

Controller CL	Controller CH	Controller CT	Controller CW	Controller CM	Controller CR	Others
Lm [cm]	T0m [C]	Tcwm [C]	Wcwm [tn/h]	Wpm [tn/h]	Wbm [tn/h]	L[17] [cm]
Ls [cm]	T0s [C]	Tcws [C]	Wcws [tn/h]	Wmum [tn/h]	Wbs [tn/h]	Pvh [atm]
EL [cm]	ET0 [C]	ETcw [C]	EWcw [tn/h]	Rmus [frac]	EWb [tn/h]	T [1] [C]
AL [pc]	AT0 [pc]	ATcw [pc]	AWcw [pc]	EWmu [tn/h]	AWb [pc]	Wbd [tn/h]
				AWmu [pc]		Whw [tn/h]

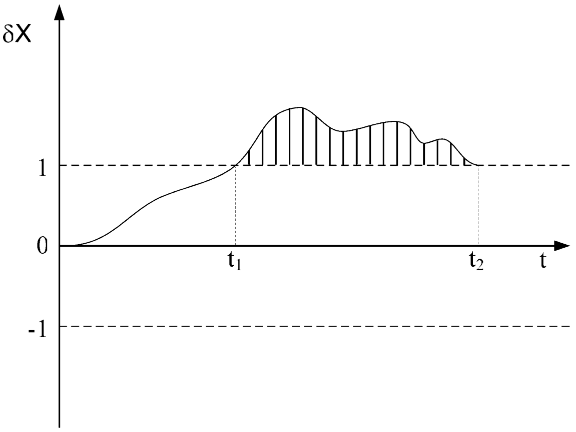

Fig. 3. Integral index.

where Iit_i is the integral index for the dynamic state i, $Ii_{i,j}$ is the contribution of the variable j, and $fe(x)$ is the function that calculates the absolute error of the signal, and it is defined as:

$$fe(x)=\begin{cases}0 & |x|<1\\ |x|-1 & otherwise\end{cases} \qquad (4)$$

On the other hand, the integration is performed along the interval[0, *tf*], where *tf* is the simulation final time. When the simulation is halted before the *tf* due to the process evolves beyond the model scope, the data is completed to *tf* by copying the last value reported by the simulator.

As the integral index is aimed at evaluating the actual fault effects over the process and equipment, the variables belonging to control loops must be excluded from this analysis. Indeed, those variables do not have any evidence of effects over the process, but over the utilities that evolve to avoid changes in the process state. Table 2 shows the variables belonging to the following control loops: CL, CH, CT, CW, CM and CR.

For the studied plant, 874 dynamic states were simulated. Table 3 shows the frequency and the percentagecumulative for the normalized integral index, which is obtained by dividing the integral index by the maximum value. Fig. 4 depicts the corresponding histogram taking out the first class, which is conformed by faults whose effects are well handled by the controllers–the 92.33% -; the remainder has effects that are exponentially distributed. Table 4 lists the 10 more serious faults according the integral index.

Table 3
Normalized integral index histogram

Class	Frequency	% cumulative
0.20	807	92.33
0.40	52	98.28
0.60	12	99.66
0.80	1	99.77
1.00	2	100.00

7.2. The derivative criterion

Another point of view considers the supervisor position. That person must study the sensors signal looking for any abnormal values and must make the appropriate decisions (e.g. keep running process doing nothing at all, shut down it, change set points, etc.) every time that detects a fault. The larger the amount and speed symptoms, the more difficult will be to deal with the situation. Fig. 5 shows how to calculate an index according this criterion. Each variable contributes with the slope of the line defined by two point of the curve δX: the first point that is outside of the band of normality, and the point of maximum deviation in the interval [0, *tf*]. Finally, all contributions are added to obtain the index.

For a given dynamic state i, this index can be evaluate by processing the result of the corresponding simulation; that is:

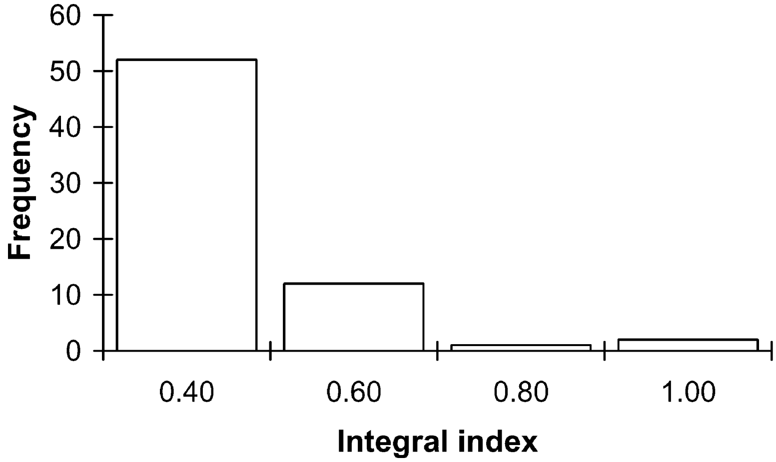

Fig. 4. Normalized integral index histogram.

Table 4
Normalized integral index partial ranking

Fault	% M	% td	Normalized integral index
CR Output High	75	0	1.00
BR Low	75	25	0.80
CR Output High	50	0	0.65
BR Low	75	0	0.58
CR Output High	50	25	0.58
CR Output Low	50	0	0.55
CL Output Low	100	0	0.52
CR Output High	50	50	0.51
CL Output Low	100	25	0.51
BR Low	100	25	0.50

$$Idt_i = \sum_j Id_{i,j} \tag{5}$$

$$Id_{i,j} = \frac{fe\left(\delta X(t_2)_{i,j}\right) - fe\left(\delta X(t_1)_{i,j}\right)}{t_2 - t_1} \simeq \frac{fe\left(\delta X_{i,j,k_2}\right) - fe\left(\delta X_{i,j,k_1}\right)}{\Delta t(k_2 - k_1)} \tag{6}$$

where Idt_i is the derivative index for the dynamic state i, $Id_{i,j}$ is the contribution of the variable j, and $fe(x)$ is the function that calculates the absolute error of the signal.

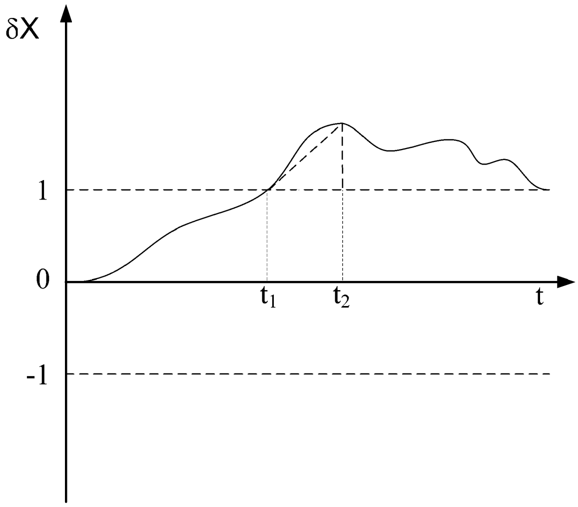

Fig. 5. Derivative index.

As the derivative index is aimed at evaluating the observable effects over the process, assuming pour instrumentation (i.e. the worst case), only the variables belonging to control loops must be included in this analysis (Table 2). Indeed, those variables by their nature will be the first affected ones by the faults.

For the studied plant, the same 874 dynamic states presented for the integral index were simulated. Table 5 shows the

Table 5
Normalized derivative index histogram

Class	Frequency	% cumulative
0.20	865	98.97
0.40	3	99.31
0.60	2	99.54
0.80	2	99.77
1.00	2	100.00

frequency and the percentagecumulative for the normalized derivative index, which is obtained by dividing the derivative index by the maximum value. Fig. 6 depicts the corresponding histogram taking out the first class, which is conformed by faults whose effects are well handled by the controllers–the 98.97% -; the remainder has effects that are uniformly distributed. Table 6 lists the 10 more serious faults according the derivative index.

7.3. The classification criterion

Finally, a point of view that combines the two former ones is outlined. In this case, only the variables belonging to control loops are taken into account (as was the case for the derivative index), as well as the absolute error for all the variables–$fe(\delta X_{i,j,k})$–along the interval[0, tf] (similar to the integral index). The larger and more lasting the effects, the more difficult will be the elimination of them to return to the normal conditions. The larger the amount and speed symptoms, the more difficult will be to deal with the situation. Both aspects are contemplated by this criterion.

For a given dynamic state i, the result of the simulation is feed to an ANN (Artificial

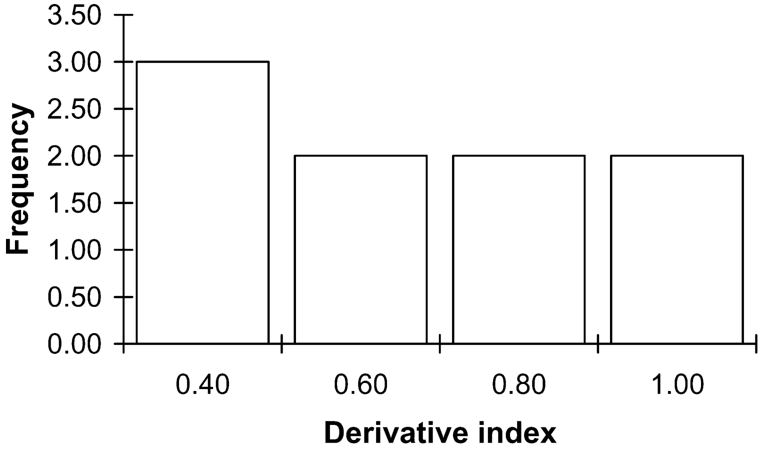

Fig. 6. Normalized derivative index histogram.

Table 6
Normalized derivative index partial ranking

Fault	% M	% td	Normalized integral index
CM Set Point Low	100	0	1.00
CM Set Point High	100	0	1.00
CM Set Point Low	75	0	0.75
CM Set Point High	75	0	0.75
CM Set Point High	50	0	0.50
CM Set Point Low	50	0	0.50
BR Low	75	0	0.28
CM Set Point High	25	0	0.25
CM Set Point Low	25	0	0.25
CR Output Low	75	0	0.17

Table 7
Classification histogram

Class	Frequency	% cumulative
8.00	524	59.95
16.00	179	80.43
24.00	99	91.76
32.00	50	97.48
40.00	22	100.00

Neural Network) [8]. From the several types of ANNs, the SOM 6×8 was selected because it is able to classify data clustering them according their characteristics.

For the studied plant, the same 874 dynamic states presented for the integral index were simulated. Table 7 shows the frequency and the percentagecumulative for the normalized class, which is obtained by dividing the derivative class by the maximum value. Fig. 7 depicts the corresponding Pareto histogram (i.e. histogram ordered by classes frequency), which is exponentially distributed (the same as the integral index). Table 8 lists the 10 more serious faults according this criterion (quite similar to the derivative index). Those results were yielded by the SOM after 500 training steps. The faults belonging to the smaller classes are considerer dangerous since the probability of occurrence for a given fault is, generally, inversely proportional to its gravity.

Table 8
Classification partial ranking

Fault	% M	% td	Normalized class
VL Low	75	100	1.00
VT Low	50	100	0.98
CM Set Point High	75	50	0.95
CM Set Point Low	75	50	0.95
CM Set Point High	75	75	0.93
CM Set Point Low	75	75	0.93
CM Set Point High	25	100	0.90
CM Set Point Low	25	100	0.90
VW Low	100	100	0.88
VW Low	100	25	0.88

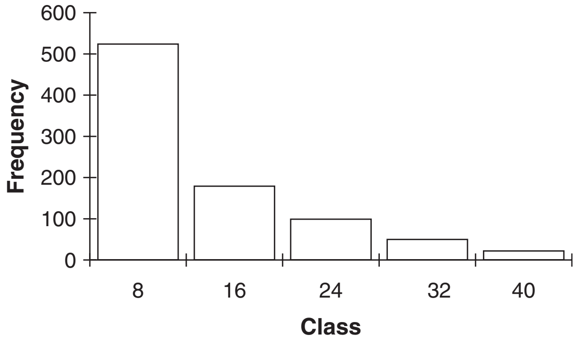

Fig. 7. Classification histogram.

8. System implementation

Known the more serious faults, experts can focus their efforts on them to elaborate appropriate action plans; so, when one of them occurs, the operator will be prepared to overcome it. The plans can also be elaborated automatically by the evaluator system. To enable its task, it can estimate the situation gravity (Fig. 1); that can be done by an ANN [8]. In fact, for each potential fault, a feed-forward/back-propagation $5+1$ with sigmoid activation functions (Fig. 8) was successfully used to estimate the corresponding derivative index from *M* and *td*, which are provided by the diagnostic system. In the learning phase, each ANN is trained with all the dynamic states belonging to the corresponding fault (supervised training). The values *M* and *td* are given as input, while the normalized derivative index is given as the desired output. Those values are taken from the Table 5. Afterward, the ANN is able to estimate the normalized derivative index for the considered fault for any value of *M* and *td*.

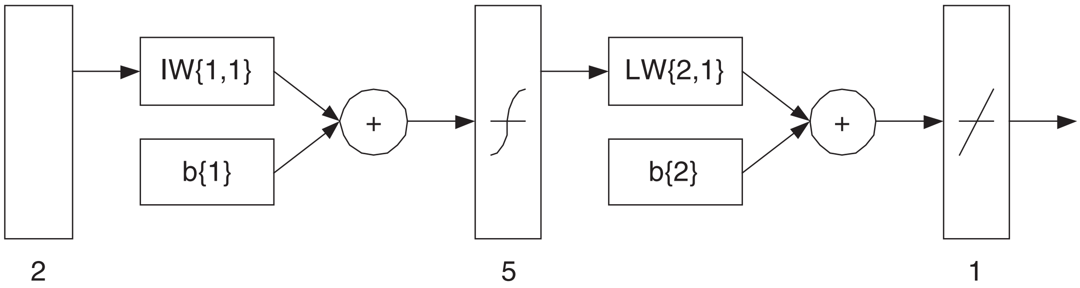

Fig. 8. ANN used to estimate the derivative index (Matlab notation).

9. Conclusions

A possible structure for a supervisor system was proposed. Criterions to measure faults gravity were also introduced. They were tested by using a dynamic simulator. The three outlined criterions yielded an appropriate classification, and helped to identify the most dangerous faults.

Additional plants must be studied in order to verify whether the exponential distribution of the integral index is a general case; in the same way, the uniform distribution of the derivative index and the exponential distribution of the classification index must be confirmed. The exponential distribution for the integral index was expected because the probability of occurrence for a given fault is, generally, inversely proportional to its gravity. However, the uniform distribution of the derivative index was not expected.

The elaboration of suitable action plans is a quite interesting line for further research. That goal can be formulated as an optimization problem, where the manipulable variables of the plant are the decision variables, and the present situation gravity is the objective function. The gravity can be estimated by using one of the presented criterions.

10. Symbols

δX — vector of normalized deviations of process variables.
μ^F — vector of certainties of faults.
Δt — sampling interval.
ΔX — vector of quantitative deviations of process variables.
ΔXn — vector of normal bands of process variables.
f — fault.
t — sampling time.
ta — activating time.
td — developing time.
tf — final time of simulation.
to — detecting time.
M — fault maximum magnitude, or stages belonging to the recovery section of the MSF plant.
N — MSF plant stages.
X — vector of quantitative values of process variables.
Xn — vector or array of normal values of process variables.

Acknowledgments

This research is supported by: Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET), Universidad Nacional del Jujuy (UNJu), International Center for Water and Energy Systems ICWES and Institute Foundation for Water Science and Technology, Abu Dhabi (IFFWSAT).

References

[1] R. Isermann, Supervision, Fault-Detection and Fault-Diagnosis Methods - An Introduction. Control Eng. Prac., 5(5) (1997) 639–652.

[2] E.E. Tarifa, Fault Diagnosis in Complex Chemistries Plants. Plants of Large Dimensions and Batch Processes, PhD Thesis, Universidad Nacional del Litoral (UNL), Santa Fe, Argentina. 1995.

[3] E.E. Tarifa and N.J. Scenna, A Dynamic Simulator for MSF Plants. Desalination, 138 (2001) 349–364.

[4] E. Tarifa, D. Humana, S. Franco, S. Martínez, A. Nuñez and N. Scenna, Fault diagnosis for a MSF using neural networks. Desalination, ISSN 0011–9164, Elsevier, 152 (2002) 215–222.

[5] E. Tarifa and N. Scenna, Fault diagnosis for a MSF using a SDG and fuzzy logic. Desalination, ISSN 0011–9164, Elsevier, 152, 2002 207–214.

[6] P.J. Thomas, S. Bhattacharyya, A. Patra and G.P. Rao, Steady state and dynamic simulation of multi-stage flash desalination plants: a case study. Computers Chem. Engng., 22(10) (1998) 1515–1529.

[7] R. Isermann, Process Fault Detection Based on Modeling and Estimation Methods - A Survey. Automatica, 20(4) (1984) 387–404.

[8] S.J. Russell and P. Norvig, Artificial Intelligence - a Modern Approach, PRENTICE HALL, New Jersey, 1995.