Quality Monitoring Using Principal Component Analysis and Fuzzy Logic Application in Continuous Casting Process

[1]Salah Bouhouche, [1]Malek Lahreche, [2]Abdelkrim Moussaoui, and [3]Jürgen Bast

[1] Iron and Steel Applied Research Unit – CSC, BP 196 Annaba, 23000, Algeria
[2] Electrical Engineering Laboratory (LGEG), University of Guelma, BP 401, 24000, Algeria
[3]HGUM, Institut für Maschinenbau, TU Bergakademie Freiberg, Cotta Stasse 4, D-9596, Germany

Abstract: This paper deals with non linear system monitoring, based on a combined use of Principal Components Analysis (PCA) and fuzzy logic to process and quality monitoring. PCA coupled to fuzzy logic was used to estimate the fault or defect according to the dynamic changes in the process inputs – outputs characterized by T^2 Hoteling and Squared Prediction Error (SPE). Correlation between the relevant process variables and the importance of defects/faults was obtained by a reliable selection of a reduced set of relevant descriptors. The effectiveness of the computing procedure based on fuzzy rule proved by its application to quality estimation of the solidification process in continuous casting.

Key words: Principal Component Analysis (PCA), Fuzzy Logic, Fault Detection and Diagnosis (FDD), Quality Monitoring, Continuous Casting Process.

INTRODUCTION

Diagnosis and monitoring of a complex system consists of detecting and identifying its working mode from one or more situations characterized by measurements. Systems operating conditions are generally defined by:

- Huge quantity of data which are given by monitoring, control and optimization processes,
- Modern computer and automation systems able to analyze that information, something in the past was not possible.

Therefore, efficient methods to on time fault/defect detection and identification has been one of the main targets in industry.

Data mining drives different classes defining abnormal and normal status. Sometimes, it is difficult to find the consequences of small abnormal conditions on the product quality particularly in noisy situation.

We consider in this work a contribution to improve the classification method using PCA and fuzzy logic. Application in continuous casting will be developed on the quality estimation when a breakout is occurred.

PCA is known as projection to latent structures, it is a dimensionally reduction technique for maximizing the covariance between different process variables summarized in the matrix X [1-3]. A combination of PCA and fuzzy rule is useful in quality classification of industrial system, because product quality control is generally a fuzzy system defining classification as very good quality, good, medium etc.

Process variables are arranged in the matrix X and pre-processed using PCA method.
Residual characteristics such as T^2 Hoteling and SPE values which are used as inputs to a fuzzy system that gives a quality classification in basis of fuzzy rule.

The parameterization of the measurement signal(s) by means of one or more external models in order to obtain a data representation space with a reduced dimension is then considered. In most cases, this phase is followed by an additional feature selection procedure, necessary to keep only relevant descriptors which retain a maximum of information. This method leads to new descriptors determined by linear or non linear combination of initial ones according to an appropriate criterion [1-16]. The second phase consists in building the system diagnosis. Different approaches can be adopted to solve this problem. The first one builds an automatic classifier which assigns the signal to one of the predefined classes according to the estimated posterior probability. Unfortunately this technique doesn't give any information about the position of the observation in the representation space. As shown in Fig.1, Fuzzy sets

Corresponding Author: Dr. Salah Bouhouche,Welding and Control Research Center, URASM-CSC, BP. 196, Annaba, 23000, Algeria

are suggested in this work to keep of the diagnosis procedure according to the importance of T² Hotelling and SPE.
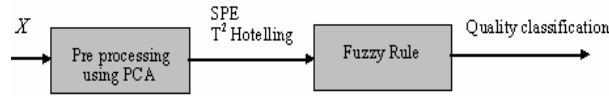

Fig.1: Principle of PCA – Fuzzy rule combination

MODELLING AND MONITORING USING PCA – FUZZY APPROACH

Data Pre-processing Using PCA: We consider a data set X of n observations parameterised by p parameters and m outputs variables denoted by Y, $Y \in R^m$. All variables are normalised, pre-processing on the matrix X is carried out by PCA method. We suppose that:

$X = [X_1, X_2..., X_i,....X_p], \forall i$, $E[X(i)] = 0$, $\sigma[X(i)] = 1$ and

$E[Y] = 0$, $\sigma[Y] = 1$.

By projecting the data into a low-dimensional space that accurately characterises the state of the system, dimensionally reduction techniques can greatly simplify and improve process monitoring procedures. PCA is such a dimensionally reduction technique. It produces a lower-dimensional representation in a way that preserves the correlation structure between the process variables, and is optimal in terms of capturing the data variability. This technique is a linear method of system reduction; it is optimal in term of the capturing of the process variability which is important to detect the process fault.

Given a training set of *n* observations and *m* process variables stacked into a matrix X,

$$X = \begin{bmatrix} x_{11}, x_{12},.....x_{1m} \\ x_{21}, x_{22},.....x_{2m} \\ . \\ x_{n1}, x_{n2},.....x_{nm} \end{bmatrix} \quad (1)$$

We characterise the measured matrix X by:

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \quad (2)$$

Where Λ is a diagonal matrix:

$$\Lambda = \Sigma^T \Sigma \quad (3)$$

$\Sigma \ \Xi \ R^{mxm}$ contains the non-negative real singular values of decreasing magnitude along its main diagonal

($\sigma_1 \geq \sigma_2 \geq,....\sigma_{\min(m,n)} \geq 0$), and zero off diagonal elements. The loading vectors are the orthogonal column vectors in the matrix V, and the variance of the training set projected along the ith column of V is equal to σ_i^2. Solving equation (2) is equivalent to solve an eingenvalue equals the square of the ith singular value (i.e. $\lambda_i = \sigma_i^2$).

In order to optimally capture the variations of the data while minimizing the effect of random noise corrupting the PCA representation, the loading vectors corresponding to the *a* largest singular values are typically retained.

Selecting the columns of the loading matrix $P \in R^{mxa}$ to correspond to the loading vectors associated with the first a singular values, the projection of the observations in X into lower-dimensional space are contained in the score matrix,

$$T = XP \quad (4)$$

and the projection of T back into *m*-dimensional observation space,

$$\hat{X} = TP^T \quad (5)$$

The difference between X and $\hat{X}$ is the residual matrix

$$E = X - \hat{X} \quad (6)$$

The residual matrix *E* captures the variations in the observation space spanned by the loading vectors associated with the *m*-smallest singular values. The subspaces spanned by X and $\hat{X}$ are called the score space and residual space respectively.

The residual matrix captures the variations in the observation space spanned by the loading vectors associated with the m-smallest singular values. The subspaces spanned by $\hat{X}$ and X are called the score space and residual space respectively. T² Hoteling's is an index charactering the variability.

$$T^2 = X^T V D^{-1} V^T X \quad (7)$$

Q − statistic or square prediction error (*SPE*) is defined as:

$$Q = EE^T \quad (8)$$

The process is considered normal if:

$$Q \leq \delta^2 \text{ and } T^2 \leq T^2_{\min} \quad (9)$$

Where δ^2 denotes the upper control limit for *SPE* and $T^2_{\min}$ denotes the upper control limit for T² Hoteling's.

Thresholds of T² and Q -statistic can be defined using normal status. If thresholds are reached, alarm is acted qualifying the abnormal status.

There are many possibilities to optimize PC's [1]. There are different methods such as Cumulative Percent Variance (CPV), Residual Percent Variance (RPV), Minimum Description Length (MDL), Average Eigenvalue (AE), Parallel Analysis (PA), Autocorrelation (AC)

To separate the noisy eigenvectors from the smooth ones, we use in this work cumulative percent variance (CPV) and residual percent variance (RPV) defined by the following relations:

$$CPV(l) = 100 \left[\frac{\sum_{j=1}^{l} \lambda_j}{\sum_{j=1}^{m} \lambda_j} \right] \% \qquad (10)$$

$$RPV(l) = 100 \left[\frac{\sum_{j=l+1}^{m} \lambda_j}{\sum_{j=1}^{m} \lambda_j} \right] \% \qquad (11)$$

l, is the index of the principal components, m is the number of process variables and λ_i is the einegenvalue. Optimal number of PC's is obtained generally when 98% of global change of the considered index (CPV or RPV) is attained.

Quality Classification Using PCA and Fuzzy System: The development of a soft sensor for quality control according to the importance of the process changes is a challenge. This permits a reduction of quality cost management in different branches of industry. Sometimes, it is very difficult or impossible to measure a certain quality parameters in real time. PCA of the input block (X) connected to a real fuzzy rules is a tool that we develop in this part to quality monitoring. Let a process characterised by its process variables X and its quality index Y. The objective is to find a complex relationship between X and Y data. Operating conditions are defined as normal (N), fault 1 (F_1), fault 2 (F_2),….fault n (F_n) according to process changes characterised by the importance of T^2 Hotelling and SPE. Each situation is defined by:

- An equivalent change of X
- An equivalent change of Y

PCA algorithm is applied to pre processing the process variables defined by the X matrix. Residual is computed as the difference between the input X and its estimated values $\hat{X}$ obtained by the application of PCA procedure. Statistical properties of residual such as T^2

Hotelling and SPE are used as inputs to fuzzy system. According to Fig.2, quality index Y is obtained by the application of fuzzy rules.
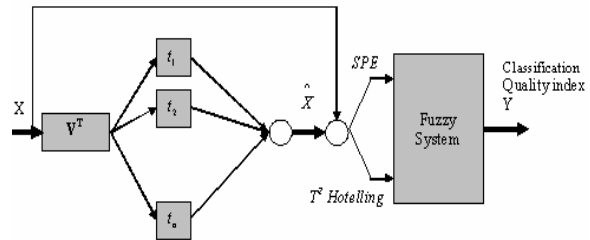

Fig.2: Principle of Quality Classification Using PCA-Fuzzy System

Modelling Using Fuzzy Sets: Many fuzzy modelling methods have been proposed in the literature [8], [13]. Most are based on collections of fuzzy IF-THEN rules of the following form:

IF x_1 is B^1 and.and x_n is B^n THEN y is C (12)

Where $x = [x_1,....x_n]$ and y are the input and output linguistic variables respectively, and B^i and C are the linguistic values characterising the membership functions. It is considered that this fuzzy rule representation provides a convenient framework to incorporate human expert's knowledge. Systems consisting of many rules are more conveniently expressed using relational arrays. However, the use of relational models in engineering application has a number of limitations. Firstly, their use is normally limited to systems with a small number of variables in view of their large size and computing requirements. Another problem posed by relational fuzzy models is that there is no simple approach for deriving numerical optimisation search techniques.

An alternative method of expressing fuzzy rules proposed by Takagi and Sugeno has fuzzy set only in the premise part and a regression model as the conclusion:

IF x_1 is B^1 and….and x_n is B^n THEN

$$y = C_0 + C_1 x_1 + ... + C_n x_n \qquad (13)$$

Where, x, y and B^i are defined in above, and C_i are real – valued parameters. Since this form of rule representation contains more information, the number of rules required will typically be much less than relational fuzzy models (a complex high dimensional non linear model valid within certain operating regimes defined by fuzzy boundaries). Fuzzy inference is then used to interpolate the outputs of the local models in a smooth fashion to get a global model. This modelling

approach provides better modelling accuracy than relational fuzzy models and it is free of the problem arising from model incompleteness which limits the usefulness of relational fuzzy models.

The principle of quality classification given in Fig.2 is summarised by a computing procedure.

As shown in the flowshart of Fig.3, starting by data acquisition as an input this method gives a computed quality index Y as an output.
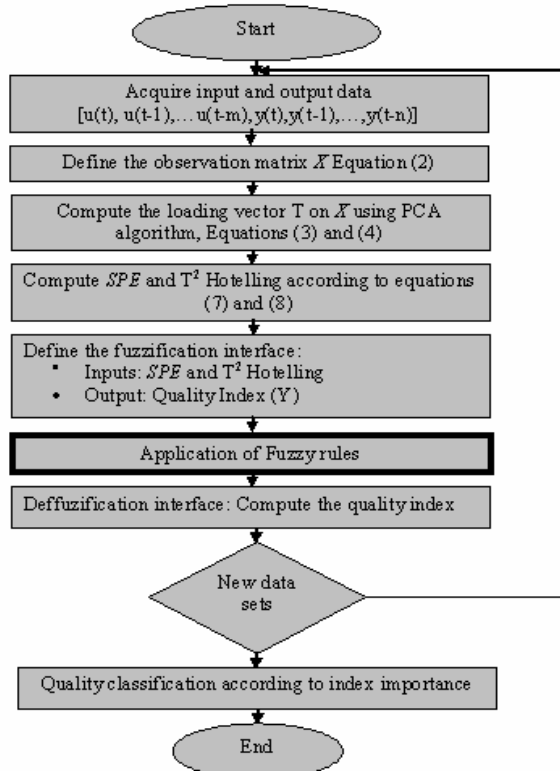

Fig. 3: Principle of quality index computing procedure

Process data are acquired and stacked in an observation matrix X. PCA method computes the *SPE* and T^2 Hotelling which are used as inputs to fuzzy system, this system is defined by the fuzzification interface, the fuzzy rules and the deffuzification interface. Fuzzy rules give relational fuzzy model between the input and the output. Output is the quality index

APPLICATION IN CONTINUOUS CASTING

Control and Monitoring of Solidification in the Mould: In the steel industry, the continuous casting process permits the formation of ingots of solidified metal called slabs that are obtained by the passage of liquid steel through several cooling zones. In a first phase, the liquid steel is poured in the mould, cooled by water, after getting cold enough penetrates the cooling zones at constant casting speed and receiving optimal flow water quantity. The final solidified ingot quality depends on its thermal history during its stay in the different cooling zones. It is, therefore, necessary to lead the cooling according to the casting events, variations of thermal loss, casting speed and different heat and mass dissipation.

During the cooling phase, slabs maintained at high temperature are in direct contact with the cooling water provoking the formation of oxides, called calamine, which involves some important variations in heat exchange, and, affects the surface temperature stability**.** According to results of metallurgical studies, surface defects such as cracks and segregation are the result of the deviation from the target temperature in different cooling zones beginning by the thermal history in the mould. The continuous casting process is shown in Fig.4.
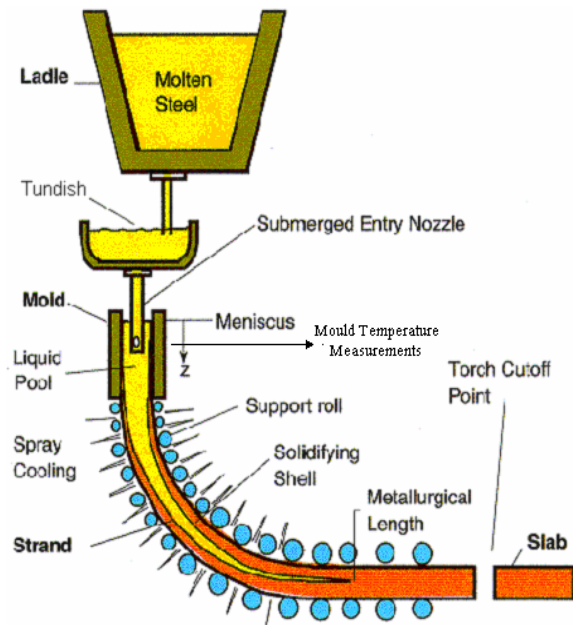

Fig. 4: Continuous casting process scheme

Defect Propagation: The mechanism for the original sticking can be explained by the existing conditions at the meniscus such as variations of casting speed, mould bath level of liquid steel, steel temperature and lubrification. Changes of casting speed have an important influence. Procedures for start-up and speed changes have been altered to slowly ramp up the speed. A breakout appears generally during metal sticking on the copper plate of the mould followed by perforation

of the solid shell due to a solidification disturbance. Sticking breakout is propagated with various speeds in various directions and particularly in casting direction. Example of breakout propagation affecting the slab quality is shown in Fig. 5.
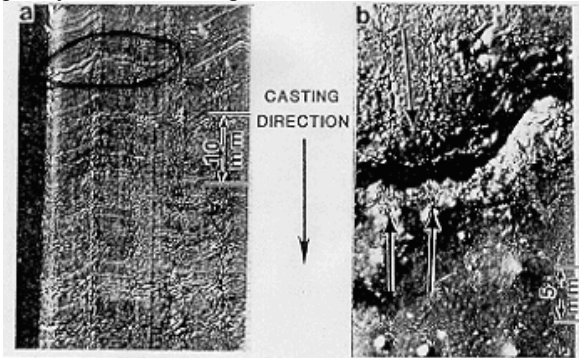

Fig. 5: Example of defect generated by breakout propagation.

In this complex situation, it is practically impossible to describe the development of a breakout in the geometrical space of the mould using an analytical model based on heat transfer, solidification and mechanical laws. The measurement and acquisition of process variables, i.e., temperature in different points at the mould surface constitute a tool for analysis and comprehension of the phenomenon. This experimental approach is also used for the development of a reliable quality monitoring system. The technique is the basis of the MTM technology that considers the mould as a thermal reactor and the appearance of defect (breakout) is a result of an imbalance of the distributed thermal reactions. The dynamics of process data that have generated such defect are affected by these random terms. Generally when a breakout is generated, the upper thermocouple records a higher temperature due to the local breakout, followed by a reduction in temperature that is also due to a partial solidification. Under the effect of the casting speed, the crack propagates and the same phenomenon is observed at lower thermocouples. Alarms and reductions of casting speed are activated. In the case of conventional techniques, when the difference between the measured temperatures and those calculated by a model reaches a fixed threshold, a series of alarms is activated. When the error reaches dangerous levels, the casting speed is automatically reduced to zero. Quality of solidified shell depends on the importance of the defect, i.e., the importance of the process variables changes. Temperature field of the mould is measured by a whole of thermocouples that give information about changes. Defect importance is characterised by:

- The number of thermocouples that have recorded changes
- The temperature dynamic changes, variations can be noised, PCA takes the significant changes components of process variables.

It has been also included the dynamic of the mould bath level and casting speed.

Modelling of Quality Index using PCA and Fuzzy Rules: The principle of On-line quality index evaluation is shown in Fig.6. On-line quality evaluation is an important domain particularly in continuous casting where the quality-control service operates in off-line, i.e., after achieving the production cycle. Control is generally made using statistical sampling method. Analyzed samples are submitted to different tests to characterize its defect. This approach is inefficient in dynamic mode because it is not taking into account all of historical events defined by the manufacturing parameters.

We consider in this part the on-line quality evaluation using PCA coupled to fuzzy system. Pre-processing of process variables gives a loading vectors based on PCA analysis, this permit us to eliminate inadequate noise that it not correlated with product quality in basis of normal and abnormal status (Fig. 7a). Alarms level from breakout detection system are also introduced in the PCA algorithm to evaluate the defect importance.
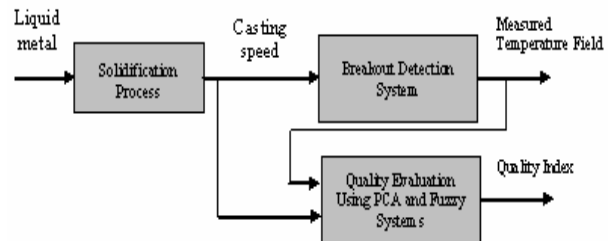

Fig. 6: Quality evaluation using PCA-Fuzzy Systems

PCA Analysis of Normal Status: As defined in equations (1), the process variables are scaled and the matrix X is formed. Fig.7b, Fig.7c and Fig.7d shown the distribution of the twenty three input variables using different indicators such as RPV, CPV and eigenvalues distribution (λ_i). All indicators confirm that only the three first components (from twenty three) of loading matrix T gives 98% of the variations. The projection of input data from first three PCA loading vectors is shown in Fig.7e. The score matrix T is calculated according to equation (4). The advantage to retaining only three principal components is that the process variability can be visualized by plotting t_2 versus t_1 and t_3 versus t_1 (see Fig.7e). t_1, t_2 and t_3 are respectively the first, the second and the third loading vectors of the loading matrix T.
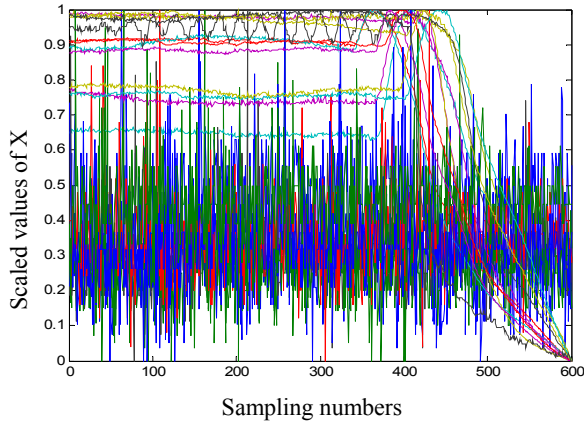
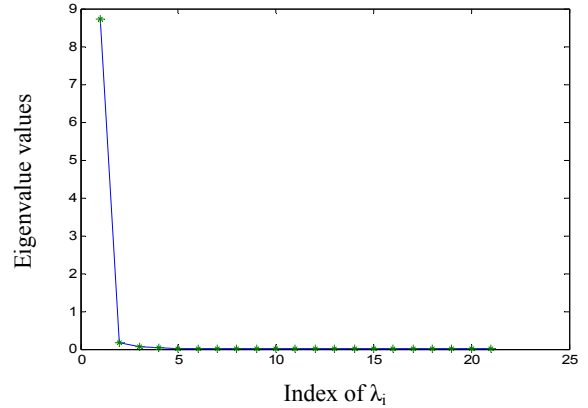Fig. 7a: Measured process variables

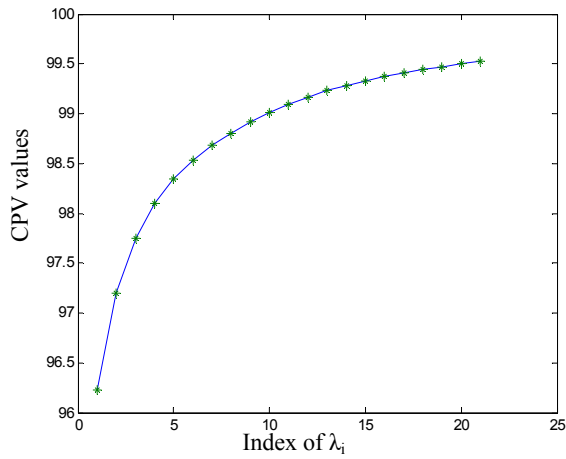Fig. 7d: Distribution of eingenvalues evolution
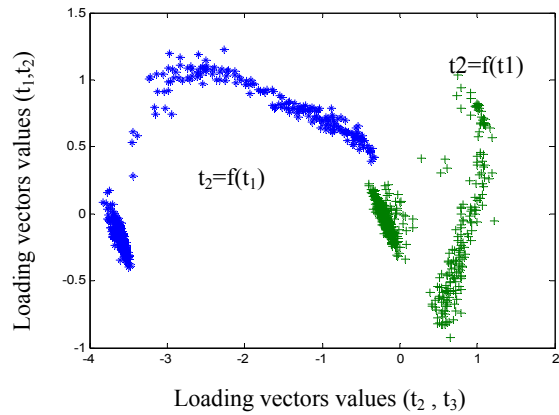

Fig. 7b: Distribution of CPV indicator

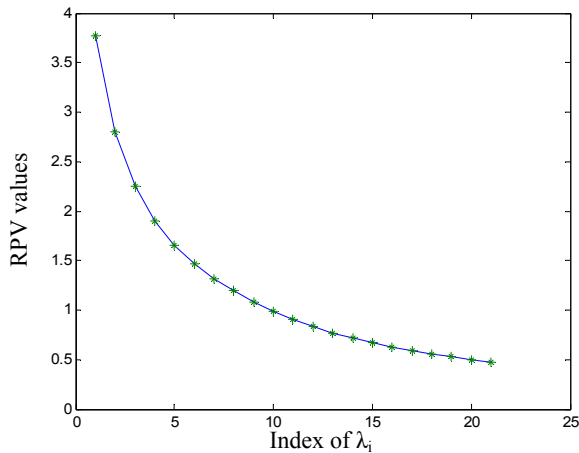
Fig. 7e: Projection of input data

Quality Index Prediction Using Fuzzy Rules: Using input - output data and NN structure, quality index has been obtained by a combined use of PCA based SPE and T^2 Hotelling and fuzzy logic. Fig.8 shows the principle of quality index prediction. This principle uses 02 inputs defined by SPE and T^2 Hotelling, 01 output defining the quality index and a model defined by fuzzy logic reasoning. Fig.9a, Fig.9b and Fig.9c show the membership functions of inputs and output respectively. All data are defined in a normalized range of -1 to 1.
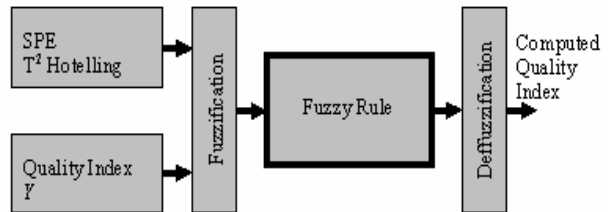

Fig. 7c: Distribution of RPV indicator

Fig. 8: Principle of prediction of quality index

The fuzzy rules are defined as follows:

1. If T^2 Hotelling IS Minimum AND SPE is Minimum THEN the quality is Very Good (VG)
2. If T^2 Hotelling IS Minimum AND SPE is Medium THEN the quality is Good (G)
3. If T^2 Hotelling IS Minimum AND SPE is Maximum THEN the quality is Medium (M)
4. If T^2 Hotelling IS Medium AND SPE is Minimum THEN the quality is Good (G)
5. If T^2 Hotelling IS Medium AND SPE is Medium THEN the quality is Medium (M)
6. If T^2 Hotelling IS Medium AND SPE is Maximum THEN the quality is Poor (P)
7. If T^2 Hotelling IS Maximum AND SPE is Minimum THEN the quality is Poor (P)
8. If T^2 Hotelling IS Maximum AND SPE is Medium THEN the quality is Poor (P)
9. If T^2 Hotelling IS Maximum AND SPE is Maximum THEN the quality is Very Poor (VP)
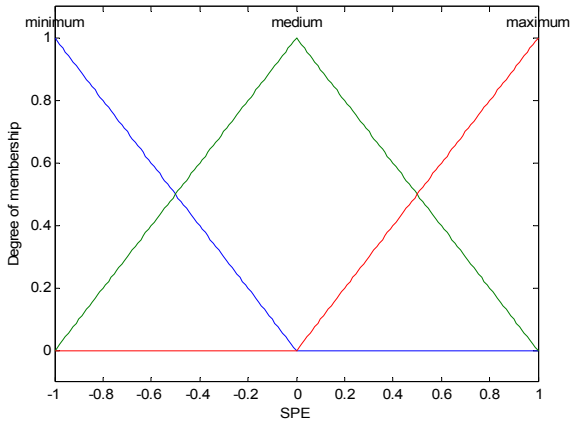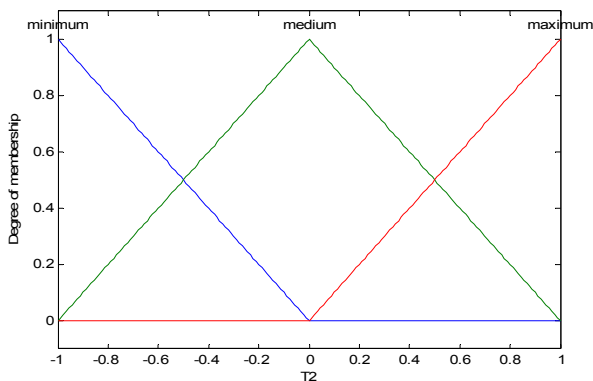

Fig. 9a: Membership function of SPE

Fig. 9b: Membership function of T^2 Hotelling

Simulation based real world measurements is carried out by 05 cycles of breakouts. Fig.10a and Fig.10b give the corresponding computed SPE and T^2 Hotelling

respectively. Using the fuzzy rules of Fig.9a and Fig.9b, a corresponding quality index is computed (Fig.10c).
The obtained results confirm the logical dependence between inputs (SPE and T^2 Hotelling) and output (quality index). It is clear that according to changes importance of process variables, quality index changes are equivalent.
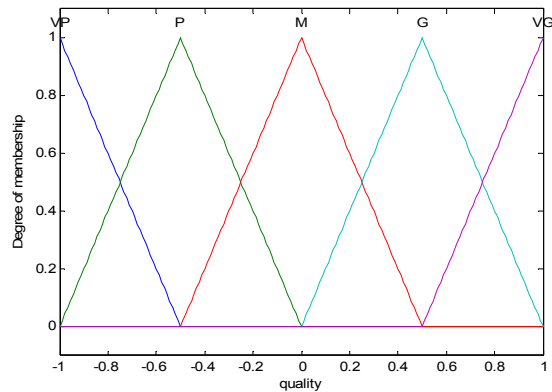

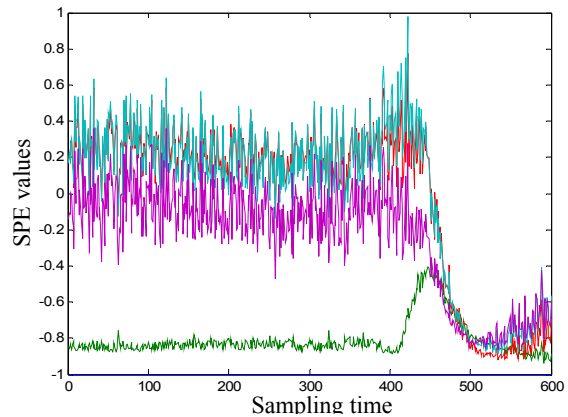Fig. 9c: Membership function of output (quality index)
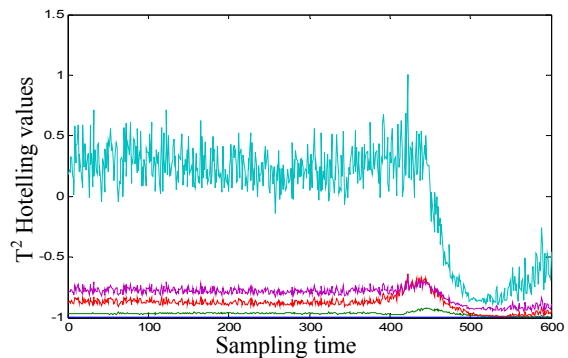

Fig. 10a: Evolutions of typical SPE

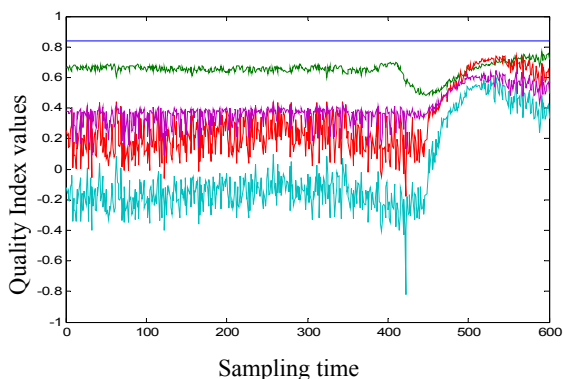Fig. 10b: Evolutions of typical T^2 Hotelling

Fig.10c:Computed Quality Index

CONCLUSION

We developed in this work a combined use of PCA and fuzzy logic to quality evaluation in continuous casting process. Non linear correlation between the significant residual properties such as SPE and T^2 Hotelling and quality index is carried out by a fuzzy system. The obtained results show that the product quality is easily evaluated using the dynamic of process data connected to a fuzzy reasoning. Tests have been made using five sets of input data. The obtained results confirm that the correlation between the defect importance and variations of its historical data is possible. Application in continuous casting seems to be useful to improve the existed package used in breakout detection system by an extension to product quality evaluation. The simulation results obtained from real sets of data acquired from continuous casting process computer confirm the effectiveness of the presented approach.

REFERENCES

1. L.H.Chiang, E.L. Russel land R.D. Braatz, 2001. Fault Detection and Diagnosis in Industrial Systems. Springer-Verlag Second Edition London Berlin Heidelberg, pp: 30-99

2. R. Isermann and P. Ballé, 1997. Trends in the Application of Model- Based Fault Detection and Diagnosis of Technical Processes. Control Engineering Practice., 5: pp 709-719

3. C. Comstock Mathew and E. B James, 1999. Literature Review for Application of Fault Detection and Diagnosis Methods to Vapour Compression Cooling Equipment. Report 4036-2

4. S.Bouhouche, 2002. Contribution to Quality and Process Optimisation using Mathematical Modelling, Ph.D thesis, Institut fûr Maschinenbau, TU Bergakademie Freiberg Germany

5. M.Norgaard, O. Ravn, N.K. Poulsen and L.K. Hansen, 2001. Modelling and Control of Dynamic Systems. Springer-Verlag London Berlin Heidelberg

6. G.Sorgrl, T.Poppe and M.schlang, 1998. Real-time Control with Neural Networks in Steel Processing: European Commission for Technical Steel Research, ECSC Workshop, Proceedings of Application of Artificial Neural Network Systems in the Steel Industry., pp: 209-225

7. D.Pham and X.Liu, 1990. State Space Identification of Dynamic Systems using Neural Networks: Engineering Application in Artificial Intelligence., (3): pp 198-203

8. C.Harris, M.Brown, K.M.Bossley, D.J.Mills and F.Ming, 1996. Advances in Neuro-Fuzzy Algorithms for Real-time Modelling and Control: Engineering Application of Artificial Intelligence., 9, (1):pp 1-16

9. A.P.Loh, K.O.Looi and K.F.Fong, 1995. Neural Network Modelling and Control Strategies for a pH Process: Journal of Process Control., 15 (6): pp 355-362

10. W.Zhenni, D.Christine, T.Ming and J.A.Morris, 1994. A Procedure for Determining the Topology of Multilayer Feed-Forward Neural Networks. Neural Networks, 7 (2), pp 291-300

11. D.Pham and X.Liu, 1990. State Space Identification of Dynamic Systems using Neural Networks. Engineering Application in Artificial Intelligence, 3, pp 198-203

12. S.Bouhouche, M.S.Boucherit and M.Lahreche, 2001. Improvement of Breakout Detection System in Continuous Casting Process using Neural Networks. IEEE Proceedings on Advanced Process Control - Applications for Industry Workshop, Vancouver, Canada: pp 53-63

13. D.Lee, J.S.Lee and T.Kang, 1996. Adaptive Fuzzy Control of the Molten Steel Level in a Strip Casting Process. Control Engineering Practices, 4 (11): pp 1511-1520

14. T.Kim and S.R.T.Kumara, 1997. Boundary Defect Recognition using Neural Networks. International Journal of Production Research, 35 (9): pp 2397-2412

15. S Bouhouche, MS Boucherit, M Lahreche and J Bast, 2003. Controlled Solidification in Continuous Casting using Neural Networks. Book EPMESC IX Conference, Hong Kong

16. Bouhouche.S, Lahreche M, Ziani S and Bast J, 2005. Fault Detection and Monitoring of Length Loop Control System in Pickling Process. Proceedings of the International Conference on Computational Intelligence in Modelling, Control and Automation, 0-7695-2504-0/05©2005 IEEE Computer Society.