**WP 2 - Supply Chain Planning**

**D 2.1 - OPTIMISATION-BASED TECHNIQUES FOR DEMAND FORECASTING**

Chemical and Process Industries are an important sector in the economy of developed and emerging countries. Because of this reason, there has been considerable effort to improve the design and operation of production plants in order to operate safely and efficiently, to ensure the best quality of intermediate and final products, maintaining the economic viability and competitiveness, and reducing the environmental impact of their activities.

In recent years, the amount of operating data collected and available in operational level has increased enormously, as a result of the adoption and continued development in sensor technology and the improvements in data collection equipment and storage systems. Several authors have recognized the potential of knowledge embedded in these data to support the operation and improvements applicable to the process, but have also highlighted the lack of current capacity for successful exploitation of information from these data. This has caused the need to develop, adopt and extend data analysis techniques in order to manage these data to extract useful information and, using this information, to support the implementation of several operational tasks.

In this framework, the thesis titled *“Estrategias de Análisis y Exploración de Datos Como Soporte a la Operación y Supervisión de Procesos Químicos”*, presented by Rodolfo Tena, presents a set of new methodologies that tries to exploit the information embedded in process historical data and effectively support process analysis and supervision tasks.

The data rectification problem was considered in first place. The adequacy of some type of wavelet for univariate filtering of different signal patterns was studied. Then, a strategy to determine the best decomposition level was proposed and consequently, an initial step to improve current wavelet filtering approaches was found. The obtained results expand the applicability and reliability of existing filtering schemes with wavelets for on-line applications without losing of accuracy on signal estimation. Additionally, an alternative strategy was proposed to solve the problem of which wavelet to choose. This last strategy consist on a weighted combination of different wavelets functions with only one output.

The data rectification problem was also studied through a Data Reconciliation (DR) approach. The focus was set on DR developments for Dynamics and Linear Systems. The proposed strategy consists on first applying a trend extraction step, to identify measured process variables trends and then reconciling these trends to make them consistent with the dynamic process model studied. For the trend extraction step, filtering using wavelets was adopted. To reconcile the estimated variables trends, an extended polynomial approach was used. The comparison with existing RD approaches shows promising results in terms of accuracy and computing efficiency. Further extensions that contemplate nonlinear cases were also introduced, showing also satisfactory results.

Process Supervision problems were considered in second place. Primarily, Principal Components Analysis (PCA) based monitoring strategies for treatment of processes frequently affected by slowly appearing disturbances or small relative shifts were compared. This comparison included some new proposals combining wavelets filtering approaches and PCA. One of the proposed approaches was capable of handling the detection of small disturbances as good as other existing approaches, but dramatically reducing the problem of false alarms generation.

Multioperational process supervision strategies were also considered and studied. First, a comparison of different strategies from literature was considered. The aim was to determine the strategy that produced better results in front of issues like identification of clusters with different forms or its performance facing outliers. The considered strategies are based on the combination of PCA with clustering techniques (PCA-clustering). Not only existing approaches were studied but also some extensions of them were also considered. Finally it was shown how new modified strategies lead to improve handling of all the considered issues. In addition, cluster number estimation problem was studied and some successfully strategies were proposed to perform it.

Finally, the integration of the above PCA-clustering strategies with multigroup PCA for supervising of multioperational process was proposed and evaluated. The aim was to allow good process supervision capabilities for handling operating changes situations and to facilitate fault diagnosis tasks together with additional capabilities like data transitions treatment. Additionally, an extension of the above-integrated strategy for analysis and supervision of process with decaying performance was evaluated. The resulting strategy was shown as potentially useful to extract useful knowledge from data and to support supervising, planning and maintenance process tasks.

The motivation of the work developed has been the challenge of extracting and exploiting the available information from historical data for using as a support tool in decision making of multiple tasks associated with Chemical and Process Industries. A careful review of the literature on Chemical and Process Engineering reveals the great academic and industrial interest in the development of data analysis strategies that respond to this challenge. The number of techniques and strategies proposed in current literature is enormous, such as in Knowledge Discovery in Databases (KDD). Thus, throughout this work, comparative studies of existing strategies have been emphasized. However, it could be observed that:

* Most of the proposals have been diverted to the theoretical refinements of the techniques used rather than effective responses to solve problems.
* For some problems, there are a high number of proposals. In this context, it is necessary to compare existing proposals to improve the resolution of the problem.

On the other hand, the range of problems in decision making is very broad in Chemical and Process Industries, keeping a relationship with the treatment of information from historical data process. Within this broad spectrum of problems, this work is related to decision making process in operational level. Within this context, efforts have been focused mainly on ensuring the quality of measurement data through the process of rectification methods and monitoring processes.

**CORRECTION DATA**

Correction methods are being studied for more than 3 decades to explore new developments. This is, on the one hand, because the fact that incorrect data (or contaminated data by noise) can lead to erroneous conclusions from operational tasks, while the use of verify or rectify data minimize the risk of errors in its application. On the other hand, the number of operational tasks that use these data is very large (regulatory control, monitoring, fault diagnosis, quality control, real-time optimization), and for these reasons, verifying data is essential.

In a first approach the univariate filter techniques were analyzed. In this area, studies in recent years have shown the superiority of wavelets to filtering techniques over other classical techniques. This work contributes to fill existing gaps, such as:

* Providing experimentally which are the most appropriate wavelet functions for the filtration of different patterns in order to propose concrete guidelines for selection. In this sense, benefits of daubechies wavelets db4, db8 and especially db1 for filtering signals in the presence of different seasonal patterns have been tested.
* Proposing a way to estimate the level associated with the decomposition of a signal using wavelets, which offers a simple alternative in the decision about the level set when it is required to analyze one or more available process signals.
* Opening a new way to achieve a more effective response to the problem of wavelet function selection through combined correction methods. This method is not valid enough in terms of accuracy, but it opens an alternative way to solve the problem.

In a second approach, the rectification data was addressed to Data Reconciliation for linear and dynamic cases. Developments in Data Reconciliation for dynamic cases are limited and strategies are difficult to apply in real situations. The contribution of this work is significant. In this context, a new strategy that integrates the individual rectification of variables by using wavelets including conciliation of these data using a polynomial representation was proposed. The most significant contributions of this integration are:

* This proposed technique effectively exploits temporal redundancy and functional data through the use of filtered trends of the process.
* Through the use of trends, the complexity in the estimation of the variances of the process was reduced, which has been a subject of continuing debate in the literature.
* As a result, the proposed method improves the estimation of dynamic reconciliation methods based on polynomial representations.

Furthermore, the applicability of this strategy is assured in situations requiring correction online through a moving horizon approach. Finally, a first extension of the method for nonlinear cases that generalize the results of the proposal was proposed

**DATA SUPERVISION**

In the area of ​​process supervision, previous literature has been focused in the review of existing strategies for monitoring specific situations in the chemical industry.

One of these situations is the monitoring of processes that can be commonly affected by abnormalities of slow onset. Although the number of strategies proposed in the literature is not very high, a comparative analysis of these strategies was proposed. All strategies are based on a filtering technique combined with Principal Component Analysis (PCA). The comparison carried out includes several variants in adopting the wavelet filtering. The proposed strategies get superior results comparing with current methods for the detection of such disturbances. In addition, the appearance of problems such as generation of false alarms has been reduced dramatically.

In addition, multithreaded process monitoring has been studied. The amount of proposed monitoring strategies for these processes in the Chemical and Process Industries is considerable, most of which have evolved into the combined use of multivariate statistical techniques, particularly the PCA, with techniques based on Fuzzy Logic.

* A first contribution in this area has been the comparative study of existing strategies in terms of various clustering techniques. This is the first time that a rigorous comparative analysis such strategies and applications for the Chemical and Process Industries has been established.
* As a result of this comparison, it has been established which of them gives a better identification of clusters in different ways and methods of interpretation and efficient management of outliers when they are present in the available data.
* The study adds extensions to existing techniques that involves a better management of all aspects mentioned above.

Finally, strategies to support the analysis and monitoring multioperational process based on a combination of previous results with PCA have been proposed. In addition,