Preprints of the 2012 20th Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, July 3-6, 2012

Clustering Data Procedure for the Prediction of the Recovered Volume of the Light Gasoil of a Visbreaking Column

Silvia M. Zanoli, Lorenzo Orlietti, Giacomo Astolfi

Abstract— **In this work the model identification of a visbreaking column for the estimation of the recovered volume at 360°C of Gasoil is considered and a filtering procedure for the selection of the identification data set is presented. A high valuable product for the visbreaking process is the light gasoil; its purity can be measured by the recovered volume at 360°C and, for control purposes, an on-line estimation of this property is very important. In this paper a new procedure for predicting the light gasoil recovered volume is presented; the approach is based on the use of a clustering Fuzzy C-Means algorithm for the selection of the input data used in the identification process. Results are presented which prove the goodness of the proposed procedure and the reliability of the estimated model in the prediction of the gasoil recovered volume.**

I. Introduction

Visbreaker processing units in oil refinery plants are employed to reduce the quantity of residual oil produced in the distillation of crude oil and to increase the yield of more valuable middle distillates. In visbreaking process, the productivity is quantified by specifications of some property of the final products, like purity, or physical or chemical properties. Usually specification on a high valuable product that is the Light GasOil (LGO) is given to assess the quality of the visbreaking process. These specifications regard the LGO recovered volume at 360°C. In general to measure this variable could be a difficult, very expensive and possibly time-consuming task. As a consequence limitations on the design of complex control strategies for supervisory purposes and model-based fault diagnosis techniques arise. In fact, in the considered case, the quality measures, i.e. the recovered volume at 360°C measurements, are available only at spots from laboratory analyses. In this scenario, automatic control and optimization schemes cannot be implemented.

This has motivated the present study that aims at the development of a prediction model of a visbreaking process; in particular the interest is in the prediction of the valuable light Gasoil recovered volume at 360°C.

The approach followed for the development of the predictor is based on the use of an inferential regression model which makes use of independent variables such as pressure, flow and temperature to infer light gasoil recovered volume. To approach the problem, different approaches could be adopted; in particular the adoption of a first principles modelization was discarded given the complexity of the process while neural networks were not taken into account given their "black block" approach and the consequent lack of physical interpretation of the parameters. In fact, in the considered case, as often the case in industrial practice, losing the "contact" with real variables generates suspicious and difficulties on practical adoption of the proposed solution arise.

Thus, suitable solutions for the proper choice of the independent variables have been adopted and a linear regression model has been applied. The validity of the linear model fitting is assessed by the limited range of the independent variables used as predictor. The main novelty of the proposed procedure is the methodology for the selection of the identification data set, based on a Fuzzy C-Means clustering algorithm [1], [2]. Instructing the regressor with the resulting filtered data allows enhancing estimation performances. The best choice for the number of clusters to retain in the data set is discussed in the paper.

The paper is organized as follows: after having defined the process at issue (section II), in section III the linear model of the visbreaking process is discussed. In the same section the clusters generation by a Fuzzy C-Means (FCM) algorithm is depicted. An ad hoc solution for the choice of the number of clusters to be retained that fix it a priori is given. The results relative to the prediction of the amount in terms of volume concentration of the Recovered volume at 360°C of LGO are presented in the Results section (section IV). Finally conclusions are summarized in the last section (section V).

II. Problem Definition

The focus of this study is a visbreaking process located in the refining area of a petrochemical plant. Visbreaking is a non-catalytic thermal process that converts vacuum or atmospheric residues via thermal cracking to gas, distillates and visbroken heavy residue (tar). Vacuum and atmospheric residues are typically charged to a visbreaker to reduce (i.e. to break) fuel oil viscosity increasing distillation process performances: the amount of residual fuel oil produced is reduced while more valuable middle distillates are yield. Visbreaking significantly lowers the viscosity of heavy crude-oil residue without affecting the boiling point range [3], [4], [5]. The thermally cracked heavy residue tar, which accumulates in the bottom of the fractionation tower, is vacuum flashed in a stripper and refluxed in the distillation process. The process is shown in Figure 1.

Previous work concerning the optimization (i.e. minimization of the head pressure) of the visbreaking distillation column head has been presented in [6].

The principal distillates yield by the visbreaking process are gasoline, kerosene, Light GasOil (LGO) and Heavy GasOil (HGO). To assess the quality of the visbreaking process it is common practice to refer to the LGO recovered volume at 360°C (a product with a high economic value). Due

978-1-4673-2529-5/12/$31.00 ©2012 IEEE

to the high cost of on-line analyzers, in many plants the measurements of the LGO recovered volume are performed off-line by means of laboratory analysis and repeated, typically, every twelve hours. Therefore, between successive measurements, the process operates on real time without any quality indicator. To cope with this problem a predictor of the LGO recovered volume has been built using routinely measured process variables.

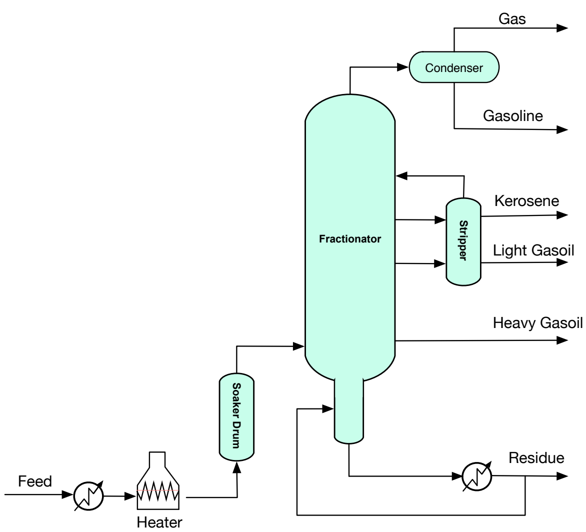

Fig. 1. Scheme of the visbreaking process

III. Visbreaking Process Model

Linear regression models are simple but extremely powerful, and have the power to empirically tease out very complicated relationships between variables [7]. In order to predict the recovered volume of LGO at 360°C (REC360), a linear regression model has been chosen for representing the steady-state behaviour of the system. In the considered visbreaking process, operative conditions are such to justify the adoption of a linear model. The choice of a linear and steady-state model for the prediction of a quality properties of the processes, commonly applied in many industrial cases [8], is here mainly motivated by the difficulties to trust on the time stamp related to the laboratory analyses. The following linear regression model has been considered:

$$REC360 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \qquad (1)$$

where the variable $REC360$ is the recovered volume value of LGO at 360°C, x_i are the process variables, β are the regression coefficients to be estimated and ε is the statistical estimation error. The least squares parameter estimates, as suggested by the ordinary least squares (OLS) algorithm, are then obtained from:

$$\widehat{\beta} = (X^T X)^{-1} X^T REC360 \qquad (2)$$

where X is the measurement data set matrix [9]. Together with the laboratory analyses of the recovered volume value of LGO at 360°C, historical measurements of several process variables such as flows, temperatures and pressures were available. When building the regression model three main issues had to be solved: a proper choice of the number of process variables to be retained in the model, their right selection and the adoption of a meaningful data set.

A. Variables Selection

Many process data could be considered as predictor variables in (1) and the first encountered issue was their proper selection. Twenty variables were at disposal but it was chosen to keep the number of these variables limited to six. In fact, in industrial context, taking into account the incidence of malfunctioning in the instrumentation and/or their (periodic) calibrations, experimentally it was observed that keeping the number of these variables limited to six was the best compromise in order to guarantee a reliable LGO estimation while keeping a good index fitting.

For the choice of the six variables out of the twenty, two different approaches have been followed and their performances have been compared. Firstly, the expertise of the plant operators has been taken into account. As a result three tray temperatures (kerosene extraction, LGO extraction, LGO reflux cold extraction) together with two flow ratios (LGO cold and hot reflux) and a pressure (head column) were considered as predictors. In Table II (upper part) the selected variables are shown. In the paper, this set of predictors will be called *process knowledge variables* and the models estimated using this set of variables will be addressed as *process knowledge models*.

In the second approach, the choice of the six variables to be used as predictor for the LGO recovered volume was performed in a more objective way: considering all possible combinations out of the twenty variables taken six at a time, different linear regression models (1) have been computed and compared considering different index criteria. In the paper, this set of predictors will be called *index knowledge variables* and the models estimated using this set of variables will be addressed as *index knowledge models*.

Different testing criteria of the model quality, such as Final Prediction Error (FPE), Akaike Information Criterion (AIC) and Minimum Description Length (MDL) [9] have been considered in order to choose the best sextuple of variables. According to Akaikes theory, the most accurate model has the smallest FPE, AIC, MDL. Indexes equations are listed in Table I

TABLE I
FPE, AIC, MDL Indices Expressions

FPE:	AIC:	MDL:
$\dfrac{n+p'}{n-p'} \cdot J$	$2\dfrac{p'}{n} + \ln J$	$(\ln n)\dfrac{p'}{n} + lnJ$

where n is the size of the data set, $p' = p + 1$ the number of predictor and $J = \dfrac{1}{n}\sum_i (\widehat{e}_i)^2$ is the quadratic index.

From the application of the above testing criteria for the selection of the predictor variables, the six variables shown in Table II were selected; the predictor variables consist of four tray temperatures (11^{th} tray, 22^{th} tray, LGO reflux hot extraction, LGO extraction) and two flow ratios (head column reflux, HGO hot reflux).

Comparing the two sextuples of process variables it can be noted that just one variable, i.e. the tray temperature of the LGO extraction, is in common. In spite of the significant mismatch between the two sets of possible regressor, the plant operators commented the *index knowledge set* as a feasible choice, just not their first choice. The common variable of the *process knowledge set* and the *index knowledge set* is the tray temperature of Light Gasoil extraction, a variable with a straightforward correlation to the recovered volume of LGO. Given its influence, it is therefore advisable that the variable is chosen.

TABLE II

DATA SET FOT THE MODEL DEVELOPMENT

Process Knowledge	
PI1809	Head column pressure
TI18992	Tray temperature of Kero extraction
TI18994	Tray temperature of Light Gasoil extraction
TI18995	Tray temperature of Cold Lateral Reflux of Light Gasoil
REFL_GAS1	Flow Ratio Hot Lateral Reflux of Light Gasoil
REFL_GAS2	Flow Ratio Cold Lateral Reflux of Light Gasoil
$FPE = 5,77$ $AIC = 1,75$ $MDL = 1,91$	
Index Knowledge	
TI18985	11^{th} tray temperature
TI18986	22^{th} tray temperature
TI18993	Tray temperature of Hot Lateral Reflux of LGO
TI18994	Tray temperature of Light Gasoil extraction
REFL_HEAD	Flow Ratio Head Column Reflux
REFL_HGO1	Flow Ratio Hot Lateral Reflux of Heavy Gasoil
$FPE = 4,47$ $AIC = 1,49$ $MDL = 1,65$	

B. *Identification Data Set*

Linear regression performances strongly relay on the adoption of meaningful regressor data, both for the identification phase and the prediction one. A main contribution of this paper is the development of a procedure for the selection of the training data (TDFS - Training Dataset Fuzzy Selection) based on data clustering at the scope to improve the quality of the identified model as resulting from the application of the linear regression equations (1) and (2). Formerly, with the aim to identify abnormal values and filter fast dynamics, a pre-processing data procedure is performed. The pre-processing is made by a consistency check on the typical operating ranges. Subsequently, in order to perform a filtering selection on the raw data, a clustering procedure has been developed which is based on the fuzzy C-means algorithm [2]. Fuzzy C-means (FCM) is a clustering method which allows each element from a dataset to belong to two or more clusters. This method (developed by Dunn in 1973 [10] and improved by Bezdek in 1981 [1]) is frequently used in pattern recognition [11]. FCM is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^m \|x_i - c_j\|^2 \quad , \quad 1 \leq m < \infty \qquad (3)$$

where $m > 1$ is the fuzziness index, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i^{th} of d-dimensional measured data, c_j is the d-dimensional center of the cluster, and $\| * \|$ is the Euclidian norm expressing the similarity between any measured data and the center. In this work the fuzziness index has been set at 2 (for more information about this parameter see [12]). The centroids are then calculated by minimization of the above object function. In the considered application the algorithm typically converged to the minimum of the function within 60 steps.

An important issue in the application of the FCM is the setting of the optimal number of clusters which is a priori, usually unknown. A common approach to this choice is the generation of different sets of data partitions characterized by different numbers of clusters and then selecting that particular partition which provides the best result according to a quality criterion. In this work the well-known Xie-Beni (XB) fuzzy clustering validity index is used [13].

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^m \|x_i - c_j\|^2}{n \cdot \min_{i,j} \|x_i - c_j\|^2} \qquad (4)$$

where c is the centroid's number. The optimal number of clusters should minimize the value of the index.

After selecting the training data set (TDFS), the output clusters are processed applying a least square (LS) algorithm on the regression model (1). As it will be discussed in the next section, the proposed TDFS-LS procedure, characterized by the filtering selection feature performed by means of the data clusterization, assures a better validation fitting index than the *classical OLS* approach which uses, for the identification purposes, the whole data set.

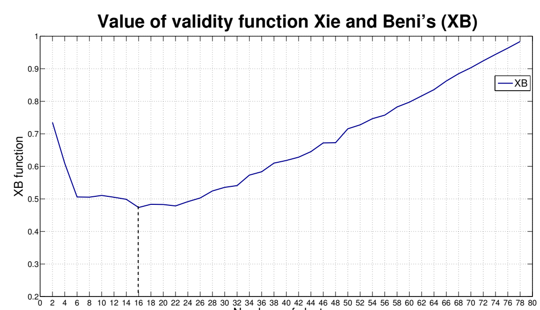

Fig. 2. Value of validity function Xie and Beni (XB)

IV. RESULTS

The available data set (116 samples with time interval of one hour) covers about four months of process activity. In order to test performances of the proposed procedure on different situations, three different records subsets (about 30-50-70% of the total data set, respectively) were extracted from the data set and used for the estimation of the LGO recovered volume. This is motivated by the interest to verify whether good estimation quality is granted even in the case few records are available, e.g. after a stop plant. A first model can still be found using the centroids of this few data as the identification data set. From time to time, as a significant increment of the data is available, a new identification process can be performed. The results obtained with the different data set are shown in the sequel.

For each one of the three considered subsets, two different clustering procedures have been performed: in the first procedure a fixed number of clusters, e.g. half of the size of the identification data set, has been used while the Xie-Beni (XB) criteria has been adopted in the second approach. Finally, in all the above cases both the sets of predictor variables described in section III-A, i.e. the *process knowledge variables* set and the *index knowledge variables* set, have been considered and the resulting estimated models compared.

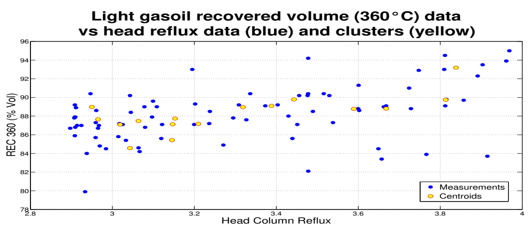

Fig. 3. The centroids relative to the flow ratio reflux of head column

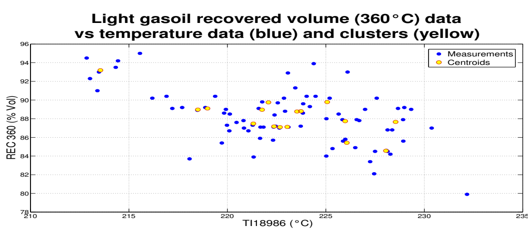

Fig. 4. The centroids relative to the 22^{th} tray temperature

The Root Mean Square Error of Prediction ($RMSEP$) has been adopted to evaluate the goodness of the predictors:

$$RMSEP = \sqrt{\frac{\sum_i e_i^2}{n}} \quad , \qquad e_i = y - \hat{y} \tag{5}$$

where y is the measured output, $\hat{y}$ is the estimation of the output and n is the size of the data set.

A. First identification data set

As explained in the previous, the first test refers to the situation just after a restart of the plant when few sample data are available. In this case a sample set of 40 (about 1 month) has been considered for the model identification while the remaining data have been used for its validation. When performing data clustering and adopting the XB criteria, the minimum XB index value was found to be in correspondence of ten centroids. In the following, this approach will be refered to as *XB-centroids*. Figure 3 and 4 shows the light gasoil recovered volume (360°C) and the 22^{th} tray temperature versus the reflux of the head column respectively (the whole data set and the centroids).

A different clusterization (called *Fixed-centroids* in the sequel) has been performed in addition to the *XB-centroids* one (that suggested ten centroids). In fact, the interest was to test the influence of the number of centroids in the performances of the proposed procedure. The number of centroids was then set equal to half of the data set dimension. To perform model identification, regression analysis has thus been applied both on the *XB-centroids* and *Fixed-centroids* clustered data.

Firstly, the *process knowledge variables* have been considered for the model identification. In Figure 5 the results relative to the predictions of the recovered volume of the LGO at 360°C relative to ten centroids (*XB-centroids*) are compared to the results of the "classical" regression analysis obtained using the whole data set of 40 samples (called *classical OLS*). Predictions are evaluated on the remaining 76 samples not considered for the identification. Similarly, in Figure 6 the results on the prediction of the recovered volume of the LGO at 360°C relative to the application of the *Fixed-centroids* approach are shown and compared with the *classical OLS*. From the computed RMSEP index it can be deduced that proposed clusterization approach give better results than the *classical OLS* and that the clusterization obtained by adopting the XB number of clusters (10 in this case) has resulted to have better performances than the choice of using a fixed number of centroids as reported in Figures 5 and 6.

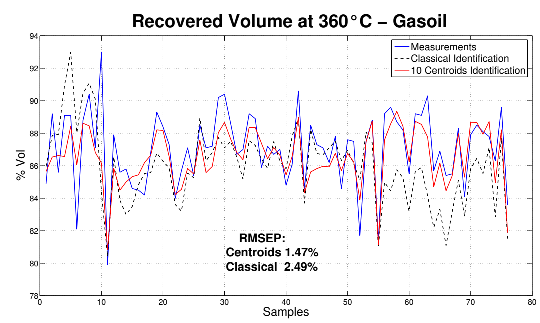

Fig. 5. Prediction of the recovered volume at 360°C of LGO: *XB-centroids* (red line) and *Classical OLS* (black dot line) in first data set

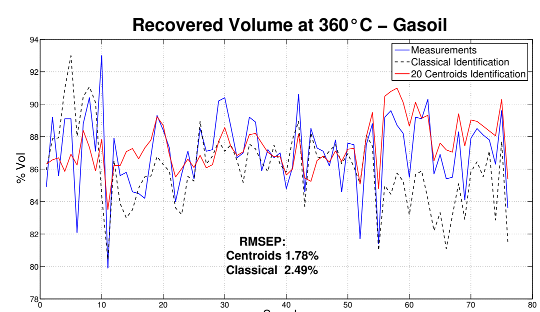

Fig. 6. Prediction of the recovered volume at 360°C of LGO: *Fixed-centroids* (red line) and *Classical OLS* (black dot line) in first data set

B. Second identification data set

The second test consists in using 58 samples (about 2 months) as input data and the *index knowledge variables* have been adopted in (1); in this case the XB index indicates twelve centroids as the best choice. In Figure 7, predictions of the recovered volume of the LGO at 360°C obtained by the *XB-centroids* clustered data are compared with the *classical OLS*. As in the previous case, from the evaluation of the RMSEP values it results that the clustering approach performs better. Finally, the results of the *Fixed-centroids* approach exhibit a better quality than the previous two, as summarized in Table III, confirming what observed in the first case.

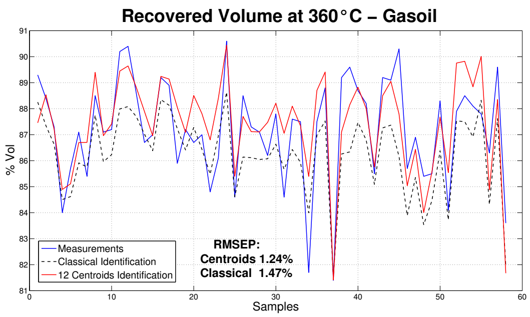

Fig. 7. Prediction of the recovered volume at 360°C of LGO: *XB-centroids* (red line) and *Classical OLS* (black dot line) in second data set

C. Third identification data set

The input data set has been further increased to 80 samples (about 3.5 month). The XB index in this case indicats sixteen centroids as the best clusterization as depicted in Figure 2. In Figure 8, the results relative to the prediction of the Recovered Volume of the LGO at 360°C are shown together with the *classical OLS*. In Figure 9, the *Fixed-centroids* approach is adopted instead. Here 40 clusters are used as input data for identification data set and prediction results are compared with the *classical OLS*.

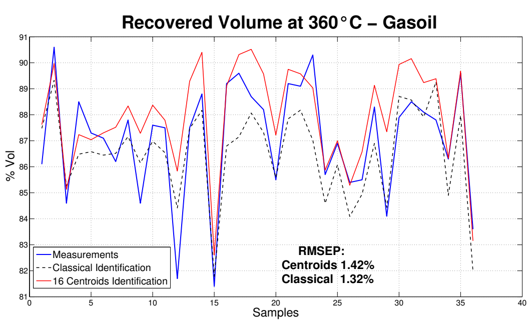

Fig. 8. Prediction of the recovered volume at 360°C of LGO: *XB-centroids* (red line) and *Classical OLS* (black dot line) in third data set

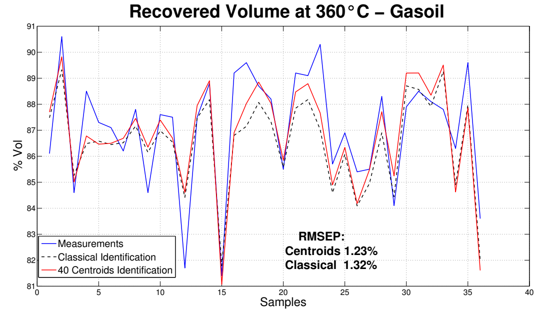

Fig. 9. Prediction of the recovered volume at 360°C of LGO: *Fixed-centroids* (red line) and *Classical OLS* (black dot line) in third data set

In Figure 10 the *Fixed-centroids* prediction model developed with the *process knowledge variables* and the *index knowledge variables* are compared. As it can be noted, using the variables computed by the quality indexes analysis rather than considering the ones suggested by operators from their process knowledge better performances are obtained.

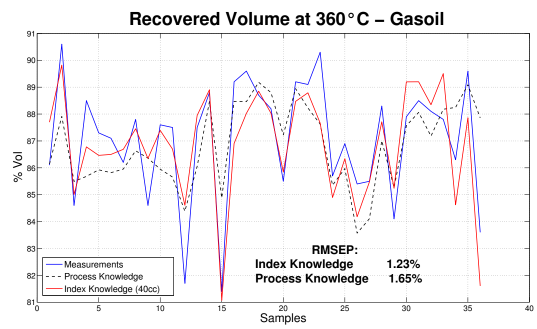

Fig. 10. Prediction of the recovered volume at 360°C of LGO with *Fixed-centroids*: *index knowledge model* (red line) and *process knowledge model* (black dot line) in third data set

D. Results discussion

The performance indexes of the proposed TDFS-LS approach are presented. The reported results refer to both the case the Xie-Beni index criteria is adopted for the selection of the number of the clusters (*XB-centroids*) and the case a certain number of clusters fixed a priori is adopted (*Fixed-centroids*) instead. Furthermore, identification results as from the application of the *classical OLS* approach, are reported. To correctly validate data results, in addition to the original data set considered in Section IV, two supplementary data sets have been analyzed. Each data set has been chosen of 116 samples, covering a period of about four months consistently with the first data set.

From data inspection it results that the *Fixed-centroids* methods allows to achieve better performances with respect to both the *XB-centroids* and the *classical OLS* approaches. The average improvements that can be obtained by adopting the *Fixed-centroids* TDFS-LS can be computed to be about 14% and 20% when considering the *XB-centroids* and the *classical OLS* approach, respectively.

TABLE III
FIT INDICES

First subset (40 samples)

RMSEP	CLUSTER APPROACH 10cc	CLUSTER APPROACH 20cc	CLASSICAL APPROACH
1°set	1,47%	1,78%	2,49%
2°set	1,70%	1,29%	1,39%
3°set	1,95%	1,47%	1,59%

Second subset (58 samples)

RMSEP	CLUSTER APPROACH 12cc	CLUSTER APPROACH 30cc	CLASSICAL APPROACH
1°set	1,24%	1,19%	1,47%
2°set	2,01%	1,26%	1,62%
3°set	1,91%	1,60%	1,61%

Third subset (80 samples)

RMSEP	CLUSTER APPROACH 16cc	CLUSTER APPROACH 40cc	CLASSICAL APPROACH
1°set	1,42%	1,23%	1,32%
2°set	1,33%	1,28%	1,47%
3°set	1,61%	1,25%	1,90%

The *XB-centroids* solution shows that good performances are not always achieved, with the exception of the first data set. In this case, the computed *XB-centroids* index shows the better results while the worst results are obtained by the *classical OLS*. Nevertheless, the *Fixed-centroids* maintain good performances. The motivation of this behavior can be found in the presence of relevant uncertainties in the selected data set successfully handled by the filtering action performed by the clusterization.

When considering the proposed *Fixed-centroids* TDFS-LS approach, from Table III inspection it can be observed that model predictions mostly show an RMSEP coefficient around 1.2 - 1.3%. These values can be considered as a very satisfactory result when considering the reproducibility of the LGO recovered volume at 360°C. In fact, its reproducibility, calculated using the ASTM-D86 standard (see [14]) results to be within 2%.

TABLE IV
INDEX AND PROCESS KNOWLEDGE MODEL

Third subset (80 samples)

RMSEP	INDEX KNOWLEDGE MODEL 16cc	INDEX KNOWLEDGE MODEL 40cc	PROCESS KNOWLEDGE MODEL 16cc	PROCESS KNOWLEDGE MODEL 40cc
1°set	1,42%	1,23%	1,88%	1,65%
2°set	1,33%	1,28%	3,83%	1,41%
3°set	1,61%	1,25%	2,90%	1,98%

Finally, the validity of the choice operated on the variable selection (see Section III-A) has been tested. As depicted in Table IV, for all the data sets the adoption of the *index knowledge models* assured better performances with respect to the *process knowledge models* (from a minimum of 10% for the second data set to a maximum of 36% for the third one) confirming the validity of the adopted approach.

V. CONCLUSIONS

The development of a linear regression model of a visbreaking process for the prediction of the valuable Light Gasoil recovered volume at 360°C is described. The availability of a reliable prediction for the Gasoil purity is a crucial factor, especially for control purposes, given the lack of online measures. A new selection procedure of the identification data set based on Fuzzy C-means (FCM) clustering algorithm that suitably filter the input data has been developed. This procedure selects a restricted number of measurements which represents several operative conditions of the process. The Xie-Beni (XB) validity function has been used in selecting the best identification data set from a raw data of the process. The new identification procedure and the classical identification technique have been compared: the results show an improvement in the measurements fitting of the new approach.

Finally, it can be stated that the new proposed approach that makes use of a Fuzzy C-means clusterization for preliminary filtering the input data, accomplishes a better fitting with respect to the ordinary least square approach. Moreover, the proposed approach for the choice of the number of clusters to retain, that fixes it at a given percentage of the data at disposal has proven to assure better performances with respect the classical Xie-Beni method.

REFERENCES

[1] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, M. Nadler, Ed. Plenum Press, 1981.

[2] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. Wiley, 1973, vol. 7, no. 4.

[3] R. Maples, *Petroleum refinery process economics*. PennWell Corp., 2000.

[4] J. H. Gary and G. E. Handwerk, *Petroleum Refining Technology and Economics*. CRC Press, 2001, vol. 49, no. 16.

[5] J. G. Speight, "The chemistry and technology of petroleum," *Bulk Solids Handling*, vol. 38, no. 8, pp. 1304–1305, 2006.

[6] S. M. Zanoli, G. Barchiesi, D Astolfi, and L. Barboni, "Head pressure minimization of a visbreaking column through an advanced pid controllers architecture," *IFAC Conference on Advances in PID Control (PID'12)*, March 2012.

[7] B. D. R. Helsel and R. M. Hirsch, "Statistical Methods in Water Resources," *Technometrics*, vol. 36, no. 3, p. 323, 1994.

[8] A. Rogina, I. Šiško, I. Mohler, v. Ujević, and N. Bolf, "Soft Sensor For Continuous Product Quality Estimation (In Crude Distillation Unit)," *Chemical Engineering Research and Design*, no. January, pp. 1–8, 2011.

[9] L. Ljung, *System Identification Theory for the User*, T. Kailath, Ed. Prentice-Hall, 1987, vol. 25, no. 3.

[10] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," *Cybernetics and Systems*, vol. 3, no. 3, pp. 32–57, 1973.

[11] K. T. Hayato Nakada and T. Katayama, "Identification of piecewise affine system based on statistical clustering techniques," *Automatica*, vol. 41, pp. 905–913, 2005.

[12] M. J. Fadili, S. Ruan, D. Bloyet, and B. Mazoyer, "On the number of clusters and the fuzziness index for unsupervised fca application to bold fmri time series," *Medical Image Analysis*, vol. 5, no. 1, pp. 55–67, 2001.

[13] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 841–847, 1991.

[14] A. S. D86, *Standard Test Method for Distillation of Petroleum Products at Atmospheric Pressure*, ASTM International Std., Rev. 2011b, 2001. [Online]. Available: www.astm.org