Industrial Monitoring by Evolving Fuzzy Systems

Lara Aires[1], João Araújo[2], António Dourado[1]

[1] Center for Informatics and Systems, Department of Informatics Engineering,
University of Coimbra, Polo II 3030-290 Coimbra Portugal
[2] Galp Sines Refinery, Process Control,
Sines, Portugal
Email: lccaires@gmail.com, dourado@dei.uc.pt, joao.araujo@galpenergia.com

Abstract *— Industrial monitoring of complex processes with hundreds or thousands of variables is a hard task faced in this work through evolving fuzzy systems. The Visbreaker process of the Sines Oil Refinery is the case studied. Firstly dimension reduction is performed by multidimensional scaling, obtaining the process evolution in a three dimensional space. Then an evolving fuzzy system (eFS) is developed to detect eventual malfunction of sensors. This eFS takes the data in the three reduced dimensions as antecedents and classifies the process state into normal and abnormal states. A software platform- the eFSLab (Evolving Fuzzy Systems Laboratory) - , with which this work has been developed, is presented and discussed. Several strategies for rule creation and evolution of rules, for Takagi-Sugeno (and for Mamdani system obtained from these) , are implemented in eFSLab. The obtained eFS shows a promising performance in the case studied, classifying in some simulations the state of the process into abnormal-normal condition in about 95% of the cases, with a number of rules between 5 and 8.*

Keywords — Evolving fuzzy systems; multidimensional scaling; industrial monitoring; fuzzy logic software.

1 Introduction

The growing complexity of industrial processes, the market competition, the more strict quality and safety rules, enhance the needs for sophisticated maintenance functions, including online process monitoring, failure detection and failure diagnosis [1] [11] [12]. This turns to be a complex problem mainly because of its non-linear high dimensional characteristics, since industrial systems are described using a large number of variables, which can be determined by various parameters and measurements.

According to [2] the main problem is to find the clusters of states in order to reflect the measurements and determine the general behavior of the system. Fuzzy systems may play here an important role, if iterative, real time techniques for rule base creation and evolution can be found. The eFSLab, Evolving Fuzzy Systems Laboratory, is a computational framework with that aim, briefly presented in this work.

The Visbreaker process of the Sines Refinery (Galp) is used to obtain real data for testing the performance of eFSLab to classify the state of the process into abnormal-normal condition. This classification is done through fuzzy systems produced by an evolving algorithm which is implemented in the platform.

An analysis is made of the performance of the fuzzy systems as classifier for this problem concluding that this algorithm has the potential to support maintenance activities as a failure detector.

The paper is organized as follows. Section 2 reviews briefly the eFS algorithm and in section 3 the eFSLab environment will be described. In section 4 the results obtained with eFSLab for data from Sines Refinery are presented and in section 5 these results are discussed and some conclusions and directions for future work are proposed.

2 The eFS algorithm

The algorithm to create Takagi-Sugeno fuzzy models is based on the approach proposed by Angelov et al. for online learning [3] and improved by [4] [5]. This algorithm allows creating zero or first orders Takagi-Sugeno fuzzy models with global or local recursive estimation parameters.

Fuzzy models are created with a data-driven approach, composed of an iterative process with 7 main stages [3] .

The procedure begins with the initialization of the rule-base structure (antecedent part of the rules) with a single rule, based on the first data sample. The second stage is to read the next data sample. Then, the potential of each new data sample is recursively calculated and in stage 4 occurs a recursively update of the potentials of the focal points (centers) of the existing rules/clusters.

After this, the potential of the new data sample is compared with the updated potential of existing centers, in order to decide if the rule-base is modified or upgraded. In the next stage consequence's parameters are recursively updated by Recursive Least Squares (RLS) or weighted RLS for globally or locally optimal parameters, respectively. Finally, the last stage is to predict the output of the system. By reading the next data sample at the next time step, the procedure returns to stage 2.

The algorithm intends to approach the subtractive clustering technique applied in a batch onset. Since it is to be applied in real-time environments, the data appears successively and at each time the new available data is used to improve the fuzzy system, in such a way that the new informative content of data should be reflected in the fuzzy rules, whether by change some of them, whether by creating new rules if there is sufficiently novelty in the process behavior. There are here some scientific challenges. The main one is how to

ISBN: 978-989-95079-6-8

measure the informative novelty in the new data. The second is how to change/create the rules consequently. The proposed techniques by [3] and [4] are based basically on heuristics and experimentation and as a consequence are case dependent. In the present work the same approach is followed. More conditions for rule update/create have been investigated and implemented in the computational framework eFSLab.

eFSLab allows to develop a Mamdani fuzzy model from a zero order Takagi-Sugeno (TS) fuzzy model. The Mamdani system is made based on the same antecedents of TS system previously created and by centering the fuzzy sets of Mamdani consequents in the Takagi-Sugeno constant consequent (only zero order models can be considered).

3 The eFSLab interface

The eFSLab interface is presented in Fig. 1(at the end of the paper). It was intended to create an interface as complete as possible, where the user could set the great majority of parameters needed to create a Takagi-Sugeno fuzzy system and transform it into a Mamdani system (in case of zero order TS systems), in a simple way.
The first step to use eFSLab interface is to select the file to import data. This file can be an Excel or a text file.
Then the user can choose TS order (zero or first), the recursive parameter estimations method, and set some other parameters needed to the recursive process.
Conditions for creation or modification of rules can also be chosen or defined by the user. There is a set of predefined conditions or the user can build up his (her) own conditions, as shown in Fig. 2

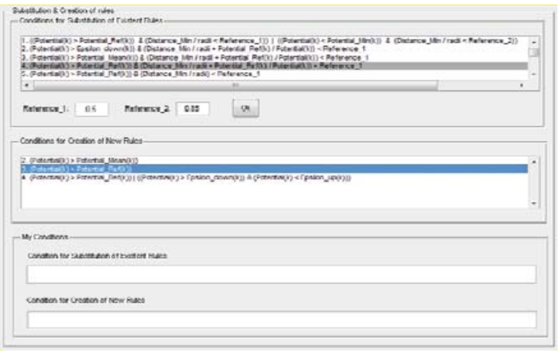
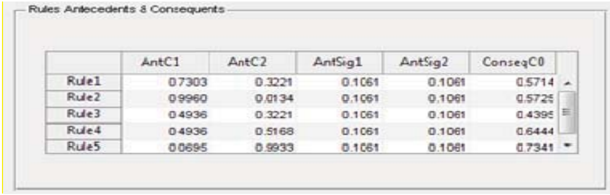


Figure 2. Creating or modifying rules in eFSLab.

The fuzzy system is created and its characteristics can be analysed using the Matlab Fuzzy Logic Toolbox, with all its functionalities (viewing the graphics, consulting antecedents and consequents fuzzy sets, etc). A diagnosis text file is created by eFSLab. The created rules are shown in a table in the interface, shown in Figure 3 for the case of two antecedents, Gaussian membership functions and five rules.



Rules Antecedents & Consequents

	AntC1	AntC2	AntSig1	AntSig2	ConseqC0
Rule1	0.7303	0.3221	0.1061	0.1061	0.5714
Rule2	0.9960	0.0134	0.1061	0.1061	0.5725
Rule3	0.4936	0.3221	0.1061	0.1061	0.4395
Rule4	0.4936	0.5168	0.1061	0.1061	0.6444
Rule5	0.0695	0.9933	0.1061	0.1061	0.7341

Figure 3. Table containing antecedents and consequents for each rule created by eFS procedure.

eFSLab produces a set of graphical information about the evolution of the TS systems, such as the number of rules, the instants of modification or creation of rules, the potential evolution of the data samples, the model output versus real output for validation data, the norm of covariance matrix and rules parameters evolution, the obtained cluster in input space. Some of them are shown in figures below.
The algorithm is based on recursive subtractive clustering [3]. At the end of the learning phase, the collection of points that has been used, together with the obtained clusters, can be observed as in Fig. 4.

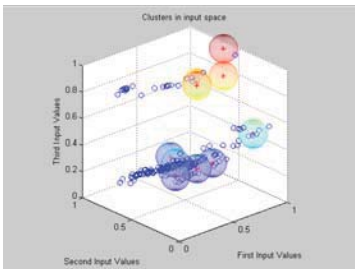


Figure 4. Clusters in input space after training. The dimensions of the spheres depend on the parameters of the clusters in a subtractive clustering environment.

In an on-line (iterative) implementation, the new rules can be created or the existent rules can be modified, and the instant when that happens can be shown as in Fig. 5

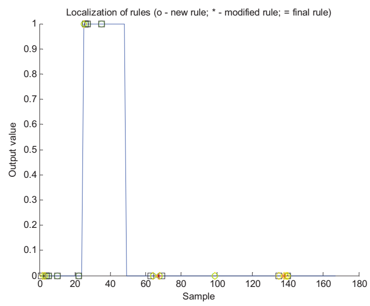


Figure 5. Localization of rules (new rules, rules modified and rules deleted).

The evolution of the number of rules during the process can be visualized as in Fig. 6.
To transform a zero-order TS fuzzy system into a Mamdani one, the *SugenoToMamdani* interface is opened as Fig. 7.
The user can choose the type of membership functions, the correspondent parameters and can preview their shape.

ISBN: 978-989-95079-6-8

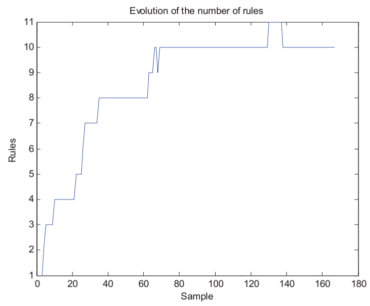


Figure 6. Evolution of the number of rules.

The Mamdani fuzzy system produced can be opened in the Matlab Fuzzy Logic Toolbox, benefiting from its functionalities for a deeper analysis .

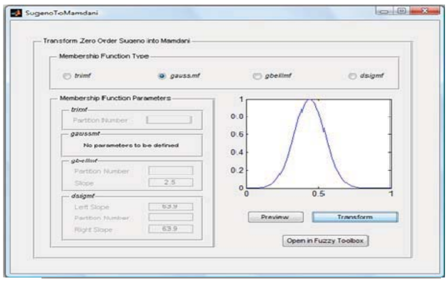


Figure 7.The *SugenoToMamdani* interface.

4 Applicatiom to the Visbreaker process

The Visbreaker Unit is intended to reduce the viscosity of the residual coming from the vacuum column. With this objective a thermal cracking process is used with a relatively low temperature, and a long residence time. As a result of the thermal process a low viscosity Visbreaker residual is obtained, as well as lighter products, such as hydrocarbonets (gas oil diesel, gasoline and gases).The great economical advantage of the Visbreaker process is in fact that it produces a residual with a lower viscosity that the load feed. By this way, it is possible to use a lower quantity of “cutterstocks” (some of them of high benefit) for the production of fuel oil.
Fig. 8 [9] presents its flow sheet. It is composed of several sub-processes, and its main part is a kiln operating at about 310 ºC.

Data averaged for each hour, between the 1st of January and the 5th of June of 2008 was collected. Actually, data from 160 tags is available. From these, after correlation analysis and process expertise, 59 were selected as sufficiently representative of the process. Multidimensional scaling is then applied to those 59 dimensions to obtain a three dimensional representation by using VisRed [6]. The three dimensional data is then fed to eFSLab that finds a zero-order TS fuzzy model.

Multidimensional scaling reduces an n-dimensional data set to p-dimensional data set, with $p<<n$, such that the distance (for example Euclidian) between each pair of points in the reduced space is similar to the distance in the high dimensional space. This means that the topology of the original data set is preserved, as well as the information embedded in the topology. The technique consists in the minimization of the Euclidian distance between two matrices: the dissimilarity matrix in the original space and the dissimilarity matrix in the reduced space. The (i,j) element of a dissimilarity matrix is the distance between points i and j.

Three simulated scenarios have been considered with increasing time horizon and number of faults: one week with one fault in one sensor (Case 1), two weeks with two faults in two sensors (Case 2), and one month with two sensors faults during one day (Case 3). In Case 1 a 20% reduction in data from one temperature sensor in the kiln, during one day, was introduced into the data between the 1st and the 6th of June 2008. In Case 2 a 20% reduction in data from two temperature sensors in the kiln, during one day, was introduced into the data between the 22nd of May and the 4th of June 2008. In Case 3, the process data between the 5th of May and the 5th of June 2008 was changed in the same way in the same two sensors. These scenarios allow to measure the diagnosis capability of the eFS, as a classifier, as the size of data set increase.

For each test, there are some parameters of eFSLab that had to be set in order to produce the fuzzy system. The first parameter to be set is the order of the TS fuzzy model (zero) and the parameter estimation (Global estimation) for all the three cases. Other parameters controlling the rule base evolutions were fixed at their default values in eFSLab.
Radii value (controlling the region of influence of each cluster in input space) changed from test to test in order to evaluate its influence on the process. The conditions for creation and modification of rules were chosen to have the capability to produce rules in regions with less data points, improving the classification performance.
Some results of the three cases are presented in the following. In Case 1 and Case 3was introduced perturbation both in test and train data sets. In Case 2 the perturbation was only introduced in test data set

Case 1. Data is from the 1st to the 6th of June 2008 (one week). In train data set it was added a reduction of 20% to data from the 2nd of June and in test data set the same perturbation was added to data from the 4th of June, in a temperature sensor.
Dimension reduction by multidimensional scaling was performed with cosine dissimilarity metric without normalization. Fig. 9 represents data after dimension reduction. Two different groups corresponding to normal and abnormal data can be observed.
The reduced data is then presented to eFSLab, obtaining a zero order TS system with 8 rules to classify the data into two classes 1 and 2. A Radii value of 0.3 was used. These rules can be visualized in Fig. 10 (using Matlab Fuzzy Logic Toolbox).
In Fig. 11 shows the real output and the fuzzy system output It can be seen that the fault was detected. This classifier shows here 95% of successful classification, although there

ISBN: 978-989-95079-6-8

are also some false faults detected, reducing the performance of the system.

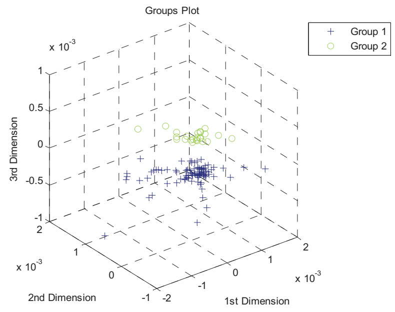


Figure 9. The three-dimensional state space for Case 1. Blue points (Group 1) are normal operating point. Group 2 are points of a simulated sensor fault.

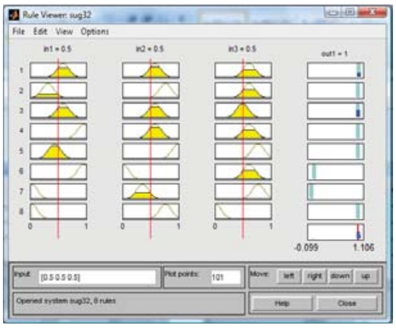


Figure 10. The 8 rules Takagi-Sugeno eFS obtained to classify the points in Case 1.

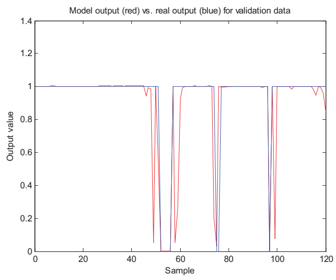


Figure 11. Real output (blue) and simulation output (red) obtained with zero-order TS system created for Case 1.

Case 2. Data from 22nd of May to 4th of June 2008 (two weeks) is considered, with a perturbation of 20% in data from two sensors in 23 of May and 4 of June.

Dimension reduction was made with Euclidean dissimilarity metric without normalization. Fig. 12 is a three dimensional representation of the reduced data.

Giving to eFSLab the reduced data, with Radii of 0.1, a 10 rules zero order TS system was obtained, shown in Fig. 13.

Case 3. Now all the data from 5th of May to the 5th of June (one month) is considered. Perturbation is the same as in the Case 2, but in 20th of May and 4th of June. Also the Radii value and the dissimilarity metric for dimension reduction are the same as in Case 2. In Fig. 15 is represented data set after dimension reduction.

Fig.14 compares the real output with the fuzzy system output. The classification is successful in 98.8% of the points.

The obtained zero-order TS system has now 11 rules, presented in Fig. 16.

Fig. 17 compares the output of the fuzzy system with the prepared data. The performance of the created fuzzy system is now substantially decreased: although the fault was detected, so have been many false positives. Overall 80.80% of the points have been correctly classified.

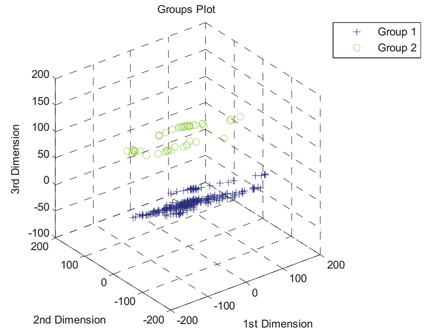


Figure 12: The three-dimensional state space representation of the Visbreaker process in the last week of May and the first week of June 2008. Blue points are normal operating point. Group 2 are points of a simulated sensor malfunction situation.

Figure 13: The 10 rules Takagi-Sugeno eFS obtained to classify the points in Case 2.

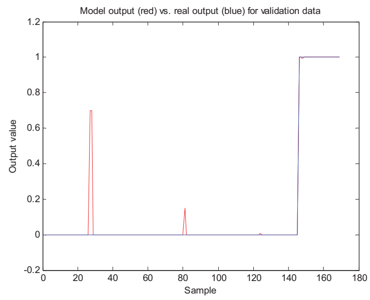


Figure 14. Real output (blue) and simulation output (red) obtained with fuzzy system created for Case 2.

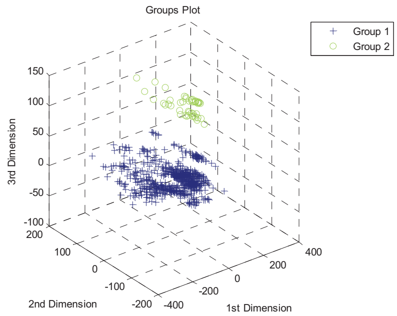


Figure 15: The three-dimensional state space representation of the Visbreaker process during the 5th of May and the 5th of June 2008. Blue points are normal operating point. Group 2 are points of a simulated sensor malfunction situation.

ISBN: 978-989-95079-6-8

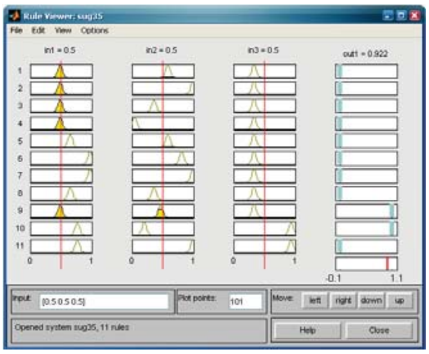


Figure 16: The 11 rules Takagi-Sugeno eFS for Case 3.

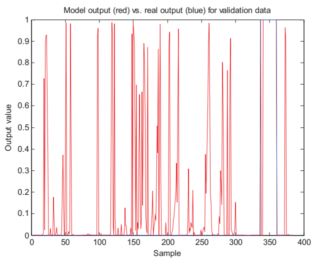


Figure 17. Case 3 :real output (blue) and fuzzy system output in Case 3.

Table 1 presents a summary of the results.

Table 1: Summary of results for the three cases.

Test	Data Set Size	MD Scaling Dissimilarity Metric	Radii	Output Matching (%)
1	1 week	Cosine	0.3	95%
2	2 weeks	Euclidean	0.1	98.8%
3	1 month	Euclidean	0.1	80.80%

When the time horizon increases (the data sets size increases proportionally), the performance of the fuzzy system as classifier seems to be worse. The way how the rules are created probably needs improvements. This is also supported by the distribution of the final clusters in the input space (see for example Fig. 4), that does not cover all the relevant space. However the results show a promising performance of the fuzzy system in terms of output (classification) matching. Furthermore, the low number of rules produced in each test contributes to a lower fuzzy sets superposition, increasing eventually interpretability and transparency, which is a main concern in fuzzy modelling [7] [8] [9]. Further work is needed to improve the iterative clustering technique, looking for better approximations to the subtractive clustering method.

5 Conclusions

Development of fuzzy systems from data is still an unattained aim of the fuzzy logic communities. Accuracy, interpretability and transparency are characteristics with an important practical relevance, allowing to give sense and credibility to the fuzzy systems. In the case of industrial monitoring, in big complexes of basic process industry, they may play an important role supporting the decision making in daily operation of the plants. *eFSLab* can be a useful computational framework to this goal. It is built with friendly user interfaces and allows an easy configuration and application to any data set built in proper form.

Evolving fuzzy systems, where rules are based on clusters, need more elaborated techniques for iterative clustering methods. This is not an easy task, since here neighbour points do not appear at the same time. Some kind of manageable memory must be developed to approach the iterative results to the non-iterative ones. This is one research direction of the authors.

New ways to measure novelty in data is needed. In a real time set the size of datasets tends to infinity. It is possible to maintain in memory only a finite time window of data Deleted points are forgotten history and new points may simply repeat pas behaviour but leading to new rules that are not only unnecessary but also prejudicial. This is another challenge to future work. The fuzzy community can have here a rich field of research to make fuzzy systems a credible and useful tool for data mining and knowledge discovery, whether in batch, whether in real-time.

Acknowledgments. This work was supported by Project CLASSE POSC/EIA/58162/2004 including EU FEDER Support.

References

[1] N. Palluat *et al*. A neuro-fuzzy monitoring system application to flexible production systems. *Computers in Industry*, 57:528-538, 2006.

[2] O. Simula *et al*, Monitoring and Modeling of Complex Processes using Hierarchical Self-Organization Maps. Helsinki University of Technology, Laboratory of Computer and Information Science. Espoo, Finland. 1996.

[3] P. Angelov and D. Filev. An Approach to Online Identification of Takagi-Sugeno Fuzzy Models. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, 34(1), 2004.

[4] J. Victor and A. Dourado, "Evolving Takagi-Sugeno fuzzy models," Technical Report, CISUC, (2003, September). Available: http://cisuc.dei.uc.pt/acg/view_pub.php?id_p=760.

[5] J. V. Ramos and A. Dourado, "On line interpretability by rule base simplification and reduction," Eunite Symposium, Aachen, 2004.

[6] Dourado, A. and Ferreira, E. and Barbeiro, P. , "VISRED - Numerical Data Mining with Linear and Nonlinear Techniques", in Advances in Data Mining. Theoretical Aspects and Applications,, pp. 93-106, LNAI, Vol. 4597, Springer, Industrial Conference on Data Mining 2007, Leipzig, Germany. Available for download from

[7] Mikut, R., J. Jäkel, L. Gröll "Interpretability issues in data-based learning of fuzzy systems", Fuzzy Sets and Systems , vol. 150, 179–197, 2005.

[8] C. Mencar and A.M. Fanelli. Interpretability constraints for fuzzy information granulation. *Information Sciences*, 178:4585–4618, 2008.

[9] González, J., I.Rojas, H. Pomares, L. J. Herrera, A. Guillén, J. M. Palomares, F. Rojas, "Improving the accuracy while preserving the interpretability of fuzzy function approximators by means of multi-objective evolutionary algorithms", International Journal of Approximate Reasoning ,Volume 44 , Is.1, , pp32-44, 2007.

[10] Galp (2006), Sines Refinery Flow Sheets.

[11] Filev D. P. and F. Tseng, Novelty Detection Based Machine Health Prognostics, Proc. of the 2006 International Symposium on Evolving Fuzzy Systems, Lake District, UK, pp. 193--199, 2006

[12] Lughofer, E. and C. Guardiola, On-Line Fault Detection with Data-Driven Evolving Fuzzy Models. Control and Intelligent Systems, Vol. 36 (4), pp. 307-317, 2008.

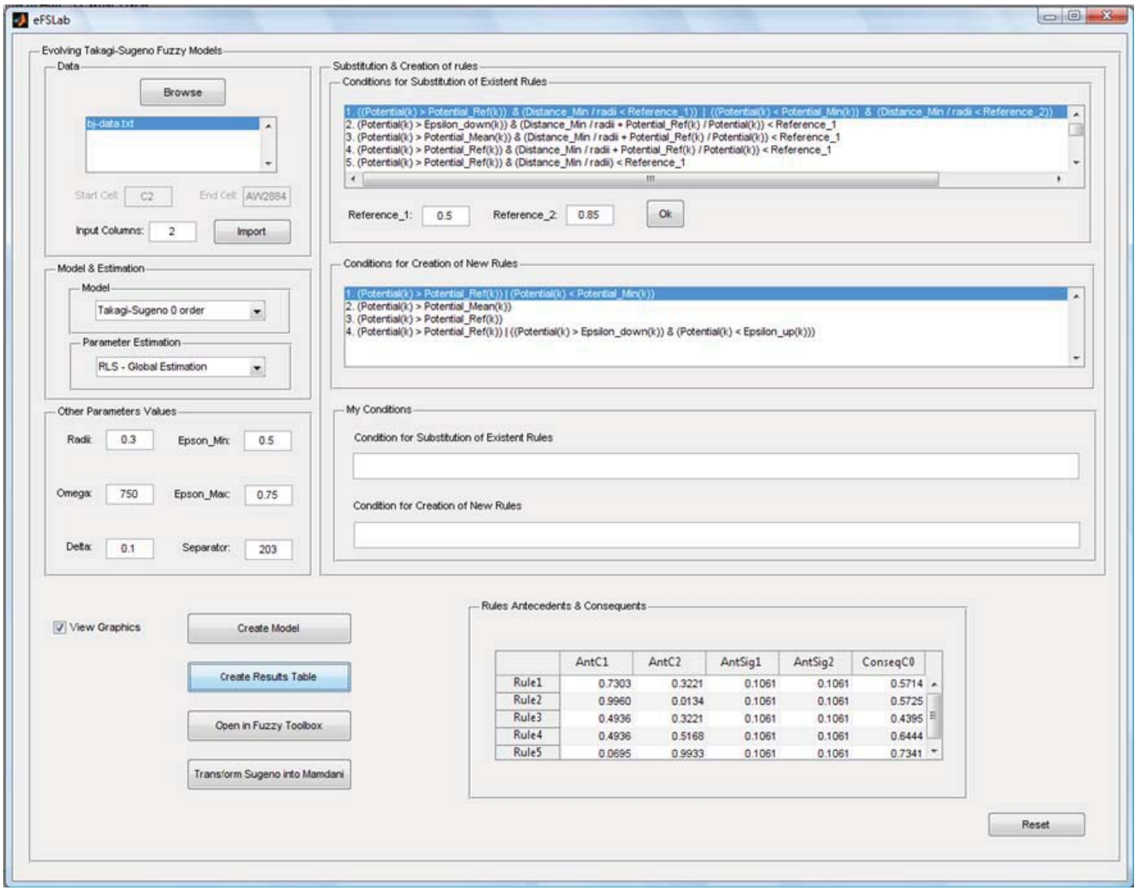


Figure 1: *eFSLab* interface.

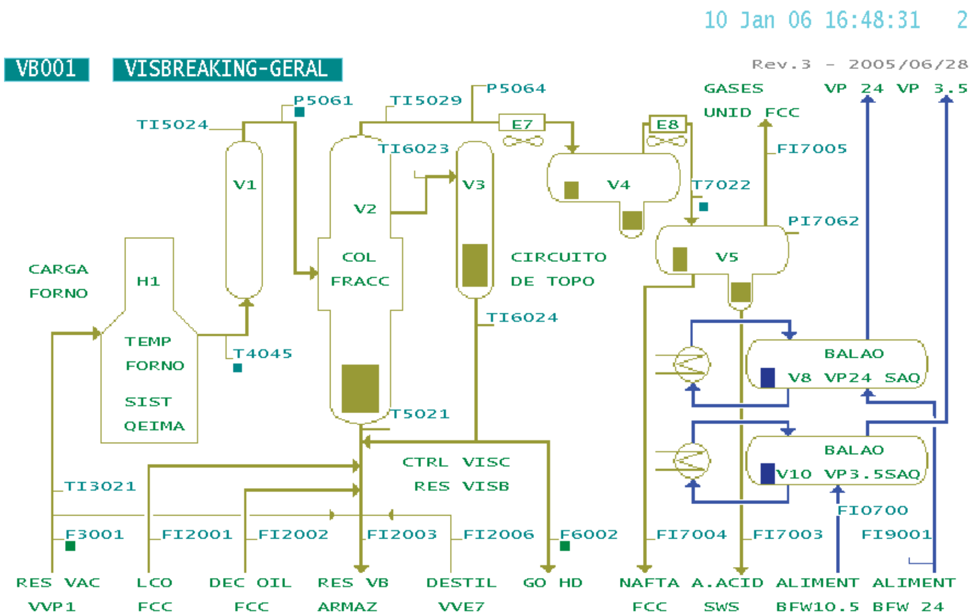


Figure 8: The Visbreaker Process [10].

ISBN: 978-989-95079-6-8