MULTICRITERIA OPTIMIZATION OF BEER QUALITY USING THE ROUGH SET METHOD

Shadi Vafaeyan[1], Jules Thibault[1] and Mariana Titica[2]

[1]Department of Chemical Engineering, University of Ottawa, Ottawa Canada K1N 6N5
[2]Centre de Recherche et de Transfert de Technologie CRTT, Laboratoire de Génie des Procédés Boulevard de l'Université, BP 406 44602 Saint-Nazaire Cedex (France)

Abstract: Volatile compounds like higher alcohols and esters contribute to beer's organoleptic profile. Most of these compounds are produced during alcoholic fermentation and operating conditions play an important role on their formation. In this study, a multicriteria optimization method, the Rough Set Method (RSM), is successfully used to determine the optimal region of fermentation conditions (temperature, pressure and initial yeast concentration) with respect to a specified beer aroma profile.

Keywords: multicriteria optimisation, rough set method, beer fermentation process.

1. INTRODUCTION

The complexity of multicriteria optimization problems in different fields of medicine, systems science, biotechnology and many more, has spurred the need to develop powerful methods that can be easily applied to solve different types of complex problems and provide the decision maker with the an acceptable compromised solution.

The optimization of complex processes normally involves minimizing and/or maximizing numerous conflicting objectives; in cases like this, there is no unique solution that can provide the optimal values for all the objective criteria simultaneously. Therefore the decision maker must find a reasonable compromise. In a previous study, the optimization of brewing process with respect to the beer's organoleptic profile has been formulated as a multicriteria optimization problem (Titica et al., 2001; Trelea et al., 2004). The beer quality has been defined in term of flavour-active components content. Six aroma compounds have been selected with respect to their organoleptic threshold: two higher alcohols (isoamyl alcohol and phenyl ethanol), three esters (ethyl acetate, ethyl hexanoate and isoamyl acetate) and one vicinal diketone (diacetyl). These components are mainly produced during alcoholic fermentation and operating conditions play an important role on their formation. Thus, the optimization of the operating conditions like temperature, pressure and initial yeast concentration in the fermentation tank, has been proposed as a way to modify beer organoleptic profile and to guarantee its regularity. The best compromise of operating temperature, pressure and yeast inoculum's size will ensure the best quality beer. The selection of the optimal operating conditions by the decision maker is a very complex task. To cope with these complex problems multicriteria optimization methods can be used, such as the Net Flow Method (NFM) and the Rough Set method (RSM) that can capture the knowledge that the decision maker has on the fermentation process in order to locate the optimal zone of operation.

In this paper, the aroma production model is first presented. Then, the optimization protocol involving the RSM is presented. Finally, the RSM is used to determine the optimal concentrations of the six major aromatic compounds responsible for the organoleptic properties of beer and to specify the optimal fermentation operating conditions. Results are compared with those obtained with the NFM and least squares methods.

2. AROMA PRODUCTION MODEL

In brewing, during the alcoholic fermentation where the yeast consumes sugars and amino acids of the wort, primary products such as ethanol, CO_2, biomass and various secondary metabolites are produced. Some of the secondary products such as higher alcohols, esters, and vicinal diketones are flavour related and are responsible for the taste of the final product. In this study six of these products, as listed in the introduction, that are known to be more significant for beer flavour were selected based on their organoleptic threshold in beer (Titica et al., 2001). A three-factor 2-level full factorial experimental design has been performed (Titica, 2000). The three factors were the main three operating variables: temperature ($T \in 10\text{-}18^oC$), pressure (50-800 mbar gage), and yeast inoculum concentration (0.5-2.0 cell/mL). From this information, kinetic models to account for the evolution of the biomass, sugars, CO_2 and the six selected aroma compound concentrations were derived. These models were used in this investigation to optimise the fermentation process.

In terms of optimisation, the fermentation process can be schematically represented by Figure 1. Given a set of input variables, the concentration of the six selected beer flavour products can be predicted. It is desired to have all of these products as close as possible to their defined targets. The target values for each compound, as determined by a group of experts, are given in Table 1. It is therefore desired to minimize the difference between the predicted values and target values for the first five compounds and to ensure that the concentration of the last compound is less than 0.2 mg/L. These requirements represent the six individual competing objective functions and will be referred as C_1 to C_6.

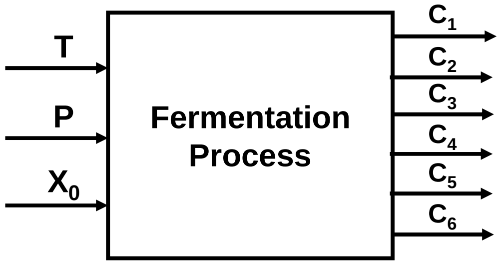

Fig. 1 3-input 6-output optimization model

Table 1: Concentration target for each product

Criteria	Compound	Target [mg/L]
C_1	Isoamyl alcohol	93.8
C_2	Phenyl alcohol	31.6
C_3	Ethyl acetate	23.9
C_4	Isoamyl acetate	2.04
C_5	Ethyl hexanoate	0.25
C_6	Diacetyl	< 0.2

3. OPTIMIZATION METHOD USED FOR BEER AROMA PRODUCTION

In this investigation, a multicriteria optimisation method, known as Rough Set Method (RSM), is used to determine the optimal operating region of the beer fermentation process. A flow chart of the typical procedure for the multicriteria optimisation of a process is presented in Figure 2. After obtaining a proper model of the process, the optimisation method boils down to: (1) circumscribing the Pareto domain approximated by a sufficiently large number of non-dominated solutions, and (2) ranking the entire Pareto domain by order of preferences. The Pareto domain (PD) represents the collection of solutions taken from the total solution set that are not dominated by any other solution within this set. In this respect, a point is said to be dominated by another point if the values of all optimization criteria (six in this investifgation) are worst than those of the second point (Thibault et al., 2002b). A genetic algorithm is often used to obtain the desired number of non-dominated points in order to adequately represent the entire Pareto domain (Halsall-Whitney and Thibault, 2006; Viennet et al., 1996). This first step is common to the majority of multicriteria optimisation techniques and is performed in absence of any biased preference of an expert or decision maker. It is only required to know if a given criterion should be minimized, maximized or as close as possible to a target value.

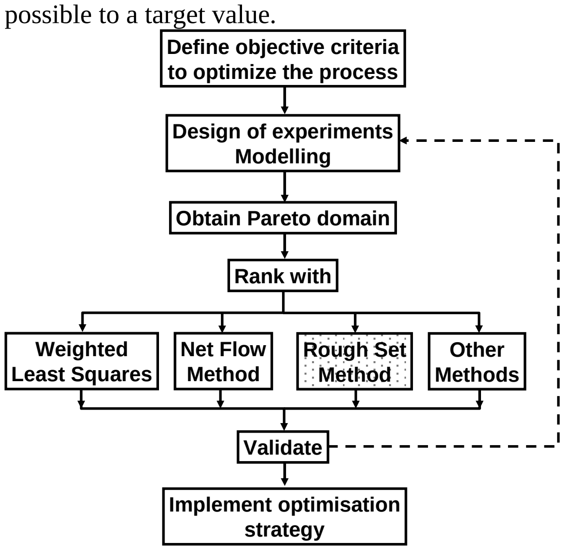

Fig. 2 Flow chart of multicriteria optimisation

The second step consists of ranking the entire Pareto in order of preferences based on the conscious, and sometimes unconscious, knowledge that an expert has on his/her process. There exist many multicriteria optimization methods such as Rough Set method (RSM), Net Flow method (NFM), Least Square method and many others. In this study, the method of interest is RSM.

The Rough Set method is able to transform the preferences of a human expert, who ranks a small set of possible solutions extracted from different regions of the PD, to a simple set of rules for ranking the entire domain (Yanofsky et al., 2005). In the traditional RSM, the set of selected points are presented to the expert all at once (batch). Thibault et al. (2002b) used the traditional RSM to classify the PD for the optimization of a thermomechanical pulping process. They presented the expert, who had a profound knowledge of the process, with seven points all at once, taken randomly from different regions of PD. This set was then ranked by the expert from the most preferred to the least preferred points. From this classified ranked set, rules were established and these rules were used to rank all the members of the PD.

In RSM, rules are established based on the expert's ranked set and the range of indifference for each criterion which is defined as the difference between two values of that criterion that is not considered significant enough to rank one value as preferred over the other; these are also established by the expert. To obtain a rule, each point is compared to every other point within that set in order to define "rules of preference" (P) and "rules of non preference" (NP). A rule is represented in the form of a binary set of values (i.e. 0 or 1), one for each criterion. A value of 1 indicates that the first point is better than the second point, while a value of 0 indicates that the first point is worse than, or not significantly different than the second point with respect to a particular criterion (Thibault et al., 2002b).

Once all possible rules from the expert's ranked set have been established, some rules need to be eliminated. This elimination is necessary for two reasons. First, if two preference or non preference rules are identical, only one copy of the rule is retained. Second, if a preference rule is identical to a non-preference rule, they are both eliminated since the expert cannot rank one point better than another point for the same reason that he considers a point worse than another point.

One perceived problem with the RSM is the selection of the small random set of points that are presented to the expert. These points must be discriminative enough to allow the generation of a representative set of rules (Renaud et al., 2005). This cannot be guaranteed since the points are selected randomly and it is very much possible that there is not enough distance between some of the criteria of these points to allow the generation of different and efficient number of rules, i.e. a set that one could use to reliably order the entire Pareto domain.

The method that is suggested and implemented in this paper for selection of points that were presented to the expert, to optimise the beer quality, seems to have effectively resolved the problem involved with the RSM. In this method, instead of presenting a subset of points from the PD to the expert all at once, the points were presented two at a time. Since two points can only be associated to two rules, one preference and one non preference, the rules are automatically generated after the points are ranked by the expert. Then, another pair of points is selected in a way to generate a different pair of rules, thus overcoming the duplication and elimination of rules. This process is continued until the desired number of rules or all possible rules are generated. If only a fraction of all possible rules is desired, the selection process would start with the most frequently encountered rule and progress towards the least frequent rules contained in the PD. It is important that all criteria of the selected pairs of points be sufficiently spaced to allow proper discrimination between the two points. In this investigation, the pair of points was selected in such a way that the value of one criterion for one point was as close as possible to one quarter of the total range of the criterion and the value of the other point was as close as possible to three quarter of the total range. The pair of points that best met this requirement for each set of two rules was selected.

In this investigation, we did not have readily access to an expert. It was therefore decided to use the Net Flow Method as the expert such that the ranking obtained with NFM was used as the expert's ranking. The NFM is a fairly robust multicriteria optimisation technique that uses the relative weighting of each criterion along with three thresholds for each criterion (indifference, preference and veto) to rank the entire PD. The reader is referred to Thibault et al. (2002a) for more details on the NFM.

4. RESULTS

The Pareto domain for the six-criterion beer fermentation process was approximated with 5000 Pareto-optimal solutions. The existence of multiple Pareto-optimal solutions only occurs when the objectives are conflicting to each other. Otherwise a unique solution is obtained. The PD only contains non-dominated solutions, i.e. in a pairwise comparison there is at least one criterion for each point in the PD that is better than one of the criteria for all the other points. When several criteria are considered simultaneously there is no unique optimal solution but a set of mathematically equivalent Pareto-optimal solutions. According to Pareto, a solution is optimal if no criterion can be improved without impairing some other criterion. This reduced search space was then ranked using the RSM.

The PD for the six criteria is graphically represented by projecting the six-dimensional criterion space onto three two-dimensional spaces as shown in Figures 3 to 5. On each plot, the black points correspond to the best 10%, the dark grey area represents the subsequent 40% and the light grey area shows the last 50% of the ranked PD. The intended target for each criterion is shown as a horizontal or vertical line on each plot.

As clearly shown in Figures 3 to 5, the solutions circumscribed by the best 10% satisfy very well five out of the six criteria. The only criterion that has a ranking pattern that is contrary to the desired one is C_1 where the optimal zone is the farthest from the desired target value. It is important to remember that all these criteria are competing and it is not possible to satisfy all criteria simultaneously. The optimization procedure has undoubtedly compromised the quality of the C_1 solution to obtain relatively good solutions for all the other criteria. The trade-off that must inevitably exist between some criteria is very well illustrated in Figure 4 where the criterion C_4 is plotted as a function of C_3. An increase in C_3 must also be accompanied by an increase in C_4 to remain in the optimal region and vice-versa. This highlights the advantage of using a multicriteria optimisation technique that uses the entire PD because it provides valuable information on the interrelationship that exists between the various criteria, in contrast with traditional optimisation where a unique solution is obtained.

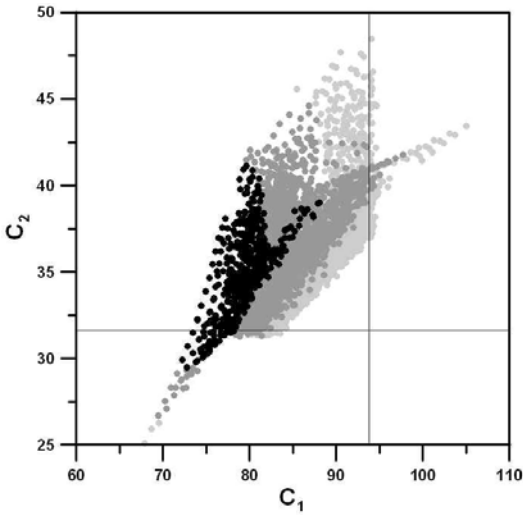

Fig. 3 Rough Set Method C_2 vs. C_1

Figure 5 shows clearly that the best 10% for criterion C_5 is located the closest to the target value. However, no solutions within the domain of exploration were able to satisfy the target. The target was established by experts through tasting trials but, assuming that the model predictions are accurate, the target value for C_5 cannot be obtained with the current range of experimental operating conditions. For criterion C_6, it is necessary for the concentration of diacetyl to be lower than 0.2 which was achieved for all points in the Pareto domain. In the current investigation, this criterion was however minimized, i.e. a value of C_5 closest to 0.

It is believed that the RSM has successfully identified the zone corresponding to the best 10% of all solutions in the PD. It is now important to examine the operating condition space that has given rise to the ranked PD. Figures 6 and 7 present the operating conditions corresponding to the PD using the same colour coding used in Figures 3 to 5. The operating conditions corresponding to the optimal region (best 10%) are in the following ranges: $T \in [12\text{-}16^oC]$, $P \in [60\text{-}310$ mbarg$]$ and $X_0 \in [0.5\text{-}0.75$ cell/mL$]$. To achieve the best 10% of all Pareto-optimal solutions, low pressure and low initial inoculum, compared to the initial range of operation, should be used whereas the temperature spans over a larger range of operation.

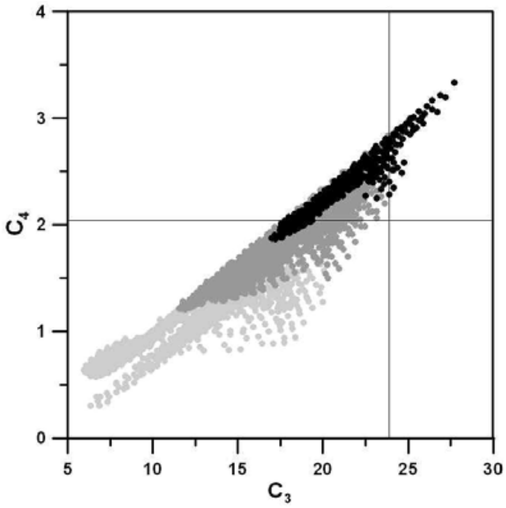

Fig. 4 Rough Set Method C_4 vs. C_3

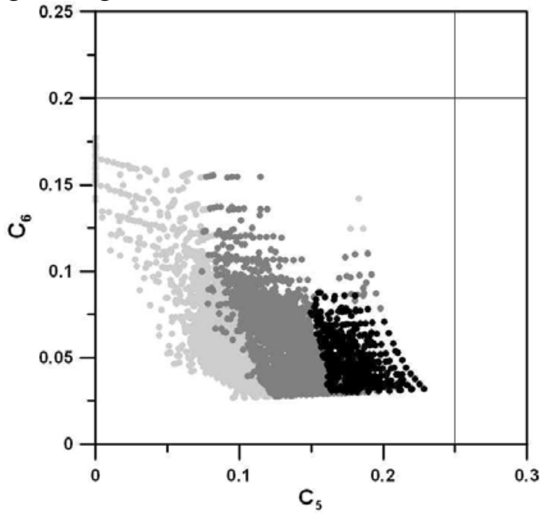

Fig. 5 Rough Set Method C_6 vs. C_5

Table 2 shows the operating conditions and the solution that ranked first amongst all Pareto-optimal solutions using the RSM. The results, obtained with the NFM and the weighted Least Squares method (LSQ), are also presented for the purpose of comparison. Each criterion closer to the target is identified in bold characters. For the LSQ, the solutions in the PD were ranked according the following global objective function:

$$J_{LSQ} = \frac{1}{6}\sum_{n=1}^{6}\left(\frac{\delta_n}{\delta_{n,max} - \delta_{n,min}}\right)^2 \qquad (1)$$

$$\text{with } \delta_n = \left|C_n - C_{n,\,target}\right|$$

In the LSQ, equal weights were used for all criteria. Comparing the criteria of the optimal points (presented in Table 2) obtained by the three methods, for the first five criteria the RSM optimum was closest to the desired target for three criteria, whereas the optimum obtained by the other two methods were closest to the target values only for one criteria. All methods obviously satisfied criterion C_6 and identified a relatively low value.

In this investigation, the Net Flow method played the role of the expert to rank each selected pair of points. The preference and non preference sets of rules were determined from the small ranked set. The two methods rank the entire Pareto domain using the information from an expert or decision maker. In the NFM, the information is in the form of a sensitivity of the differences of each criterion to the target values whereas for RSM, the information is indirectly obtained by forcing the expert to make a choice between solutions taken from the PD. It is therefore interesting to see how well the RSM was able to capture the information portrayed by the NFM. Figure 8 presents the parity plot of the Rough Set ranking as a function of the NFM ranking. It is clear that the best 10% is more or less the same as the best solutions are gathered in the lower left-hand corner of the plot. The correlation coefficient between the rankings of the two methods is 0.925 even though it is much lower for the best 10%. For comparison, the correlation coefficient between the RSM and LSQ is 0.895. The LSQ becomes identical to NFM when the three thresholds are equal to zero.

The RSM strongly relies on the ability of the method to select representative solutions that are presented to the expert to decide which of the two points is better. In previous investigations, a small set of points were presented in batch to the expert and then the rules were established (Thibault et al., 2002b) with the risk of obtaining duplicate rules, not accounting for some rules and finding rules in both the preference and non preference sets. The method proposed in this investigation, for selecting points presented to the decision maker, prevented the presence of duplicate rules as a pair of points was chosen in order that each time two new rules were generated. For a system with six criteria, there exists a maximum of 62 distinct rules, i.e. 2^6-2. The minus two accounts for rules (000000) and (111111) that cannot be part of the PD because of the domination constraint. Each time a pair of points is chosen, two complementary binary rules are generated (Ex.: 110011 and 001100).

Table 2: Optimal point for each method

	LSQ	NFM	RSM
T (^{o}C)	15.3	14.87	13.37
P (mbarg)	284	425	50
X (10^7 cell/mL)	0.5	0.5	0.5
C_1	84.0	**88.18**	73.42
C_2	37.4	37.81	**31.49**
C_3	20.0	15.93	**22.55**
C_4	**2.33**	1.795	2.624
C_5	0.181	0.1523	**0.2069**
C_6	**0.0357**	**0.0377**	**0.05816**

In the current PD, all 62 rules were present such that it was necessary to select 31 pairs of points in the PD to generate all possible rules. Naturally, some rules were more frequent than others. The most frequent set of rules was {(011110);(100001)} with 21% occurrence, whereas the least frequent set was {(111101);(000010)} with 0.0032%. Some of the rules along with their frequency of occurrences (FO) are presented in Table 3. The rules in Table 3 are ordered from most to least frequent for both the preference and non preference sets. The 31 pairs of rules were used to rank all points of the PD as shown in previous Figures 3 to 5.

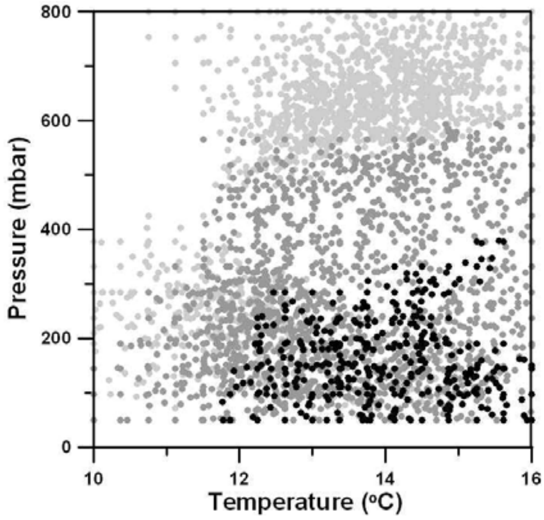

Fig. 6 Rough set Method P vs. T

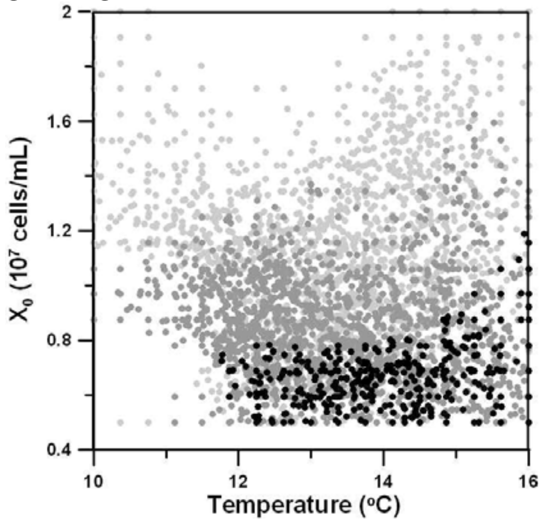

Fig. 7 Rough Set Method X_0 vs. T

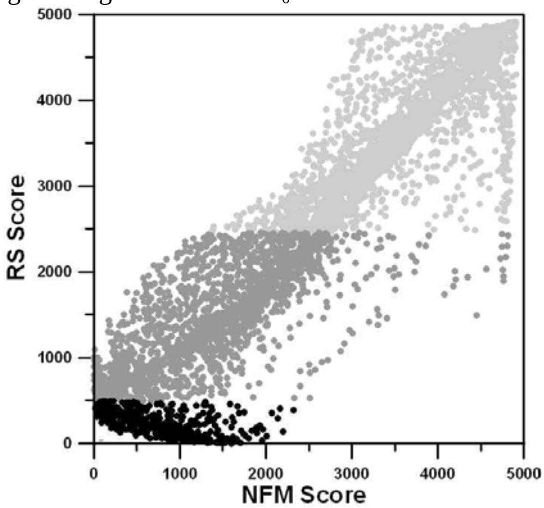

Fig. 8 Ranking comparison between NMF and RSM.

To present 31 pairs of points to the decision maker to account for all possible rules is undoubtedly excessive as it represents a large human effort. In batch mode if 7 points are presented to the expert, a total of 21 point-to-point comparisons must be made simultaneously, which may be overwhelming for the expert. However. it is believed that selecting a reduced number of pairs of points representing the most frequent rules could be sufficient to provide an equivalent ranking of the PD. In the present optimisation problem, if five pairs of points are chosen to represent the ten most frequent rules, 72% of all pairwise comparisons within the PD would be accounted for. This percentage increases to 89% if 10 pairs of points are presented to the expert. The last 30 rules (last 15 complementary pairs) account for only 3% of the total number of occurrences. To test this hypothesis, the ranking of the PD was performed from a reduced number of rules and the results are presented in Figure 9. The ranking with a reduced number of rules is compared in terms of the regression coefficient (R^2) with the NFM ranking and the ranking using all 31 sets of rules. It is obvious that in the present case, using approximately 10 sets of rules leads to an equivalent results than for the 31 sets of rules, and a result, the experts has to examine a much lower number of points for an identical final result.

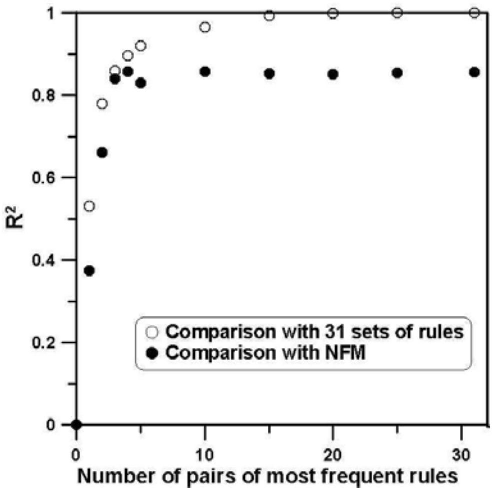

Fig. 9 Performance of RSM using a reduced number of rules from most to least frequent rules.

Table 3: P and NP rules with frequency of occurrences (FO).

	C_1	C_2	C_3	C_4	C_5	C_6	FO (%)
P Rules	0	1	1	1	1	0	10.5
	0	0	1	1	1	1	9.21
	…	…	…	…	…	…	…
	1	1	1	0	1	0	0.0021
	1	1	1	1	0	1	0.0016
NP Rules	1	0	0	0	0	1	10.4
	1	1	0	0	0	0	9.21
	…	…	…	…	…	…	…
	0	0	0	1	0	1	0.0021
	0	0	0	0	1	0	0.0016

Another important aspect of the present optimisation problem is the uncertainty whether or not the experts would be able to assess the quality of beer based on the values predicted by the model since they usually assess the quality of beer through actual testing. The expert needs to associate a given aroma concentration with an actual tasting of beer, which may represent a significant challenge. It is believed that if the model is good, using a robust multicriteria optimization method will lead to acceptable results. It is required to finally perform a validation experiment at the optimal operating conditions where the experts would actually test the final product.

CONCLUSION

This investigation has considered the use of the Rough Set method for selecting the operating conditions that would optimize the quality of beer. The RSM was able to clearly identify an operating zone for which the concentrations of the six most important metabolites, responsible for the organoleptic quality of beer, would be as close as possible to their estimated target values.

The RSM can be very useful for multicriteria optimisation to capture in a very natural way the conscious and, sometimes unconscious, knowledge that the expert has on his/her process. This knowledge is not always efficiently captured by traditional optimisation methods.

REFERENCES

Halsall-Whitney H. and Thibault J. (2006). Multi-Objective Optimization for Chemical Processes and Controller Design: Approximating and Classifying the Pareto Domain. Computers Chem. Eng. 30, 1155-1168.

Renaud J., Thibault J., Lanouette R., Kiss L.N., Zaras K. and Fonteix C., (2007). Comparison of Two Multicriteria Decision Aid Methods: Net Flow and Rough Set Methods in a High Yield Pulping Process. *EJOR*, 177, 1418-1432.

Thibault J., Lanouette R., Fonteix C. and Kiss L.N. (2002a). Multicriteria Optimization of High Yield Pulping Process. Can. J. Chem. Eng., 80, 5, 897-902.

Thibault, J., Taylor D., Yanofsky, C., Lanouette, R., Fonteix, C. and Zaras, K., (2002b). Multicriteria Optimization of a High Yield Pulping Process with Rough Sets. *Chem. Eng. Sci.* 58, 203-213.

Titica M.L. (2000). Contrôle du profil aromatique de la bière lors de la fermentation alcoolique. Thèse de doctorat, INPG, Grenoble, France.

Titica, M., Trelea, L.C. and Cheruy A., (2001). Optimization of Fermentation Operating Conditions with a View to Control Beer Aroma Profile. *8th Int. Con. in Biotechnology.* (Dochain, D., and Perrier, M. (Ed.). 225-230.

Trelea C., Titica M. and Corrieu G. (2004) Dynamic optimal control of the aroma production in brewing fermentation Journal of Process Control, 14(1), pp. 1-16.

Viennet R., Fonteix C. and Marc I. (1996). Multicriteria Optimization Using a Genetic Algorithm for Determining a Pareto Set. Int. J. Sys. Sci., 27, 2, 255-260.

Yanofsky, C., Taylor, D.G. and Thibault, J. (2006). Novel Methodology for Assessing a Ranked Pareto Domain: Drift Group Analysis. *Chem. Eng. Sci.*, 61, 1312-1320.