Fuzzy Knowledge-Based Prediction of Yeast's Morphological Characteristics for Sparkling Wine Manufacturing

Svetla Vassileva

Ass.Prof., Ph.D.

Institute of Control and Systems Research -BAS

Acad.G.Bontchev str., bl.2, P.O.Box 79, 1113 Sofia, Bulgaria

Tel. (+3592) 71- 401 / ext.353

E - mail: CONSYS@BGCICT.ACAD.BG

Abstract

The paper deals with fuzzy and neuro-fuzzy modeling technique implementation in the microbiological research. An original fuzzy knowledge-based model was built and optimized by ANFIS from FUZZY TOOLBOX of MATLAB with SIMULINK for prediction of morphological parameters of a starter sparkling wine yeast culture.

Keywords

Fuzzy knowledge-based system, neuro-fuzzy model parameter identification, biotechnology, sparkling wine manufacturing.

1 INTRODUCTION

Morphological characteristics define living state and activity of microbial population. Basic morphological characteristics of yeast's populations confirm concentration of

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35390-6_58](https://doi.org/10.1007/978-0-387-35390-6_58)

L. M. Camarinha-Matos et al. (eds.), *Intelligent Systems for Manufacturing*
© IFIP International Federation for Information Processing 1998

cells – living, budding and dead; specific growth and budding rate; budding energy, etc. Sparkling wine starter cultures shows a specific sensitivity cultivation conditions.

Morphological research of yeast's cultures is a serious scientific problem and their automation could be impossible without applications of recent soft-computing methods because of the uncertainty, imprecise data and ill-defined knowledge about influence of different cultivation conditions on the morphological characteristics of yeast populations (Vassileva et al., 1994)

The paper deals with fuzzy knowledge-based systems (FKBS) implementation in the morphological research of microbial populations. The applied in our paper modeling approaches use numerical information and expert knowledge about studied phenomena and show a higher efficiency and accuracy of recent soft - computing methods. In the presented work it was appointed through the comparison of fuzzy and neuro-fuzzy knowledge-based modeling techniques.

2 PREDICTION PROBLEM STATEMENT

Main problem in the modeling of multivariate nonlinear complex systems is existence of different ill-defined relationships between process variables. Through the use of fuzzy sets and membership functions, the overall behavior of a FIS is pieced together to form a smooth input-output relationship (Jang, 1993, 1994, Vasileva et al., 1994).

The structure of an original FIS was developed in Fuzzy Toolbox of Matlab 4.2 with Simulink. By using the adaptive neuro-fuzzy inference system (ANFIS) the structure and parameters of the original FIS were optimized.

The morphological characteristic prediction is a multivariate nonlinear regression problem where several input attributes confirm the continuous output variable behavior. The different by nature input attributes with their ranges (universe of discourse) are presented in Table 1.

Table 1 Input attributes for prediction of yeast's morphological characteristics

<i>Input</i>	<i>Range</i>	<i>Type</i>
Cultivation temperature, T [°C]	13-19	continuous
Oxygen aeration volume, aer [vol/vol/h]	0-1	continuous
Agitation speed, n [min ⁻¹]	0-800	continuous
Total biomass concentration, X[mln cells/ml]	18-24	multi-valued discrete
Nitrogen concentration, N ₂ [mg/ml]	0-100	multi-valued discrete

The attribute to be predicted was budding cells X_b concentration. The data were obtained from 3 time repeated 21 experiments with a batch propagation of

sparkling wine yeast strain. It is well known that fermentation processes are non-stationary and non-linear and obtained information is uncertain and often inaccurate. Moreover the duration of studied in our case fermentation processes is approximately 90-180 hours. For this reason the problem of data scarcity is the first problem to be solved. After removing the missing values the size of used data set was 306. In Table 2 are randomly chosen instances from the data set.

Every input variable in our FIS was represented with two different by type membership functions chosen from an expert. The output variables were represented by linear functions. The primary knowledge base includes $2^5=32$ fuzzy rules extracted from expert knowledge. The accuracy reached by using the original FIS was seriously unsatisfied. In order to achieve a higher accuracy the adaptive neural-network approach by using the command *anfis* in Fuzzy Toolbox of Matlab was implemented.

Table 2 Instances of the morphological characteristics training data set

T	aer	n	X	N_2	X_h
19	1	800	240.0	100	8.6
13	1	800	176.0	100	7.0
19	0	800	37.50	100	18.6
13	1	0	124.0	0	13.2
16	0,5	400	135.0	50	17.5
16	0,5	400	143.0	50	22.1

3 ANFIS FOR MORPHOLOGICAL PARAMETERS PREDICTION

Before implementing the command *anfis* the structure of input database was changed. The data was normalized and two of Bellman type membership functions for each variable were used. The entire database was divided into two equals by size data sets: a training data set and a checking data set with 153 points. The training data set is implemented for model building and the checking data set is used for model validation. According (Jang, 1993) it is likely to reach a better generalization capability for unseen data.

The input partition for five input attributes with two linguistic values leads to 32 rules. If we want to stick to a first order predictive fuzzy model it result in $(5+1)*32=192$ linear parameters. Constructing 10 predictive fuzzy model ($C_2^5=10$) trained by *anfis* and compared by their checking errors did the selection of relevant input attributes in our case (see a part of them in Table 3). Comparing checking error we find the most important input attributes for morphological characteristic prediction in sparkling wine manufacturing: X and T . The modeling results are unsatisfied but the studied phenomena are highly nonlinear and complex. It is known that not only X and T are very important for yeast's growth, but also the rest input attributes from original database.

Table 3 Training and checking data errors for input attribute selection

<i>Output</i>	<i>Input1</i>	<i>Input2</i>	<i>Training error</i>	<i>Checking error</i>
X_b	O ₂	n	3,9297	4,5468
	O ₂	T	2,1432	6,4018
	T	X	4,2192	4,9636
	n	N ₂	3,1458	5,0232

On Figure 1 are presented checking and training data errors and on Figure 2 - three-dimensional surface with minimal RMSE=0,006331 for budding cells concentration (X_b) dependence on biomass concentration (X) and cultivation temperature (T), presented with 3 of Belmann type linguistic values after 500 epochs of training.

Figure 1 Checking and training data errors Figure 2 Optimal 3-D surface

FKBS applications in morphological analysis of yeast's helps us to understand new microbiological phenomena by using gathered from many years human experience, numerical and linguistic information about microbial populations. The incorporation of FIS with neural network excellent approximation increases efficiency and accuracy of this modeling technique.

REFERENCES

Jang, J.-S. R. (1993). ANFIS: Adaptive-network-based FIS. IEEE Trans. Systems, Man and Cybernetics, **23**(03), 665-685.

Jang, J.-S. R., Sun C.T. (1995) Neuro-fuzzy modeling and control. IEEE, **83**(3), 379-406.

Vasileva S., V.L.Huong, J.Votruba. (1994) An Exp. System Applied to the Physiol. Analysis of Early Stage of Beer Fermentation. F. Microbiol, **39** (6), 489-492.