Opportunities and Challenges of Model Predictive Control in Food Technologies

Želimir Kurtanjek*

University of Zagreb, Faculty of Food Technology and Biotechnology,
Pierotijeva 6, Zagreb, CROATIA

Abstract

Modern food industry has gone transformation from classical production concepts based on intensive manual work and off-line monitoring to a highly automated computer on-line controlled processes. The main focus in process automation is on application of modern process analytical technologies (PAT) and computer models for analysis and synthesis of information from on-line sensor signals with basic engineering principles of heat, mass and momentum, chemical and biochemical reactions, and industrial microbiology. New trends in process information synthesis and analysis of complex multidimensional data are based on: chemometric methods, such as principal component analysis (PCA) and partial least squares (PLS); and artificial intelligence (AI) algorithms, such as artificial neural networks (ANN), and fuzzy logic inference. For incorporation of computer algorithms in model predictive control (MPC) needed are developments of mathematical and statistical models for prediction of future outputs of multivariate nonlinear systems over a finite time horizon based on a set on multivariate inputs. Development of accurate and robust multivariate models for food technologies represents the main challenge and is crucial for MPC applicability. Methodology of MPC requires determination of manipulative inputs by optimization of a control objective function with constraints on manipulative and state output variables. From a practical point of view, main advantage of MPC (and the reason for its industrial success) is its true multivariate structure and ability to handle systems with constraints in a systematic and transparent manner. From process management point of view, MPC control can support process operation in a flexible and dynamic way to meet changing market requirements. The MPC technology is used to steer processes closer to their physical limits to obtain a better economical result. Main opportunities of MPC are in process control for production of food with improved nutritional and organoleptic properties, product quality assurance, environment protection, and increased product market value.

Key words: model predictive control MPC, principal component analysis PCA, partial least squares PLS, neural networks ANN, ARMAX models

Introduction

Application of modern control systems in food industry has reached the highest levels of technical development. Processes are supported by modern process analytical technologies (PAT), numerous on-line measurement (bio) sensors, soft-sensors, process computer vision technologies, robotics, and process computers. This development is driven by global economy demands for more consistent attainment of high food quality and assurance, high productivity, more efficient use of energy, increasing awareness and legal restrictions on impact of food industry on environment.

For food industry to meet these challenges efficient process control is needed. It can be achieved only by integration of on-line multivariate plant information on the level of accurate, reliable and robust mathematical models. The models are on-line incorporated into the MPC control scheme that utilizes the process model for two main objectives (1-2):

- Explicit prediction of future plant behavior, and
- Computation of appropriate corrective control action, based on past information over a finite time interval, required to drive predicted process output in optimal sense to input information (output target sets) in a future finite time interval.

* E-mail: zkurt@pbf.hr Tel: +385 1 4605 294 Fax: +385 1 4836 083

For control of food technologies the most challenging aspects for modeling are:

- Essential process nonlinearities,
- Pronounced transport time delays,
- Large process time constants,
- Strong multivariate interactions,
- Dominant batch type process operations,
- Non-homogeneous properties,
- Lack of explicit on-line data on product quality,
- Lack of accurate parameters for physical and chemical properties.

These modeling challenges are met by application of high throughput on-line process analytical technologies (PAT), most of them are based on process spectrometry technologies (VIS, NIR, IR, fluorescence), and specific biosensors. Instrument information is integrated and reduced into process models by numerically and statistically reliable multivariate algorithms based on PCA decomposition, such as linear partial least squares PLS models, nonlinear principal component artificial neural networks (ANN), and models based on fuzzy logic principles. These methodologies are especially applicable for modeling of explicitly unmeasured control objective variables such as product organoleptic profile, nutritional value and product quality. Complex process dynamics is usually accounted by discrete time auto-regressive moving averages with exogenous inputs (ARMAX models) extending over time windows corresponding to characteristic process times [3-6]. For these models to be incorporated into process control scheme the most important features are their accuracy and robustness of the model predictions in view of possible and predicted process variations [3-6].

Methods

In general terms, the nonlinear optimal control is defined by the following model based relations:

Nonlinear multivariate process model

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x},\mathbf{u},\mathbf{d},\boldsymbol{\theta}) \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{1}$$

where $\boldsymbol{x}$ is a state vector (measured and un-measured states included), $\boldsymbol{x}_0$ is initial state vector, $\boldsymbol{u}$ is a vector of manipulative control input variables, $\boldsymbol{d}$ is a vector of unmeasured process disturbances (due to variations of raw materials composition, process generated disturbances, and process plant environment influences), $\boldsymbol{f}$ are model functions based in process first principles (energy, mass and momentum balances), and $\boldsymbol{\Theta}$ is a vector of known model parameters,

Output function model is

$$\mathbf{y} = \mathbf{g}(\mathbf{x},\mathbf{u},\mathbf{d}) \tag{2}$$

where $\boldsymbol{y}$ is the vector of output variables which are subject to control

The model of control objective functional is given by

$$\min_{\mathbf{U}} J(\mathbf{u}(t),\mathbf{x}(t),\mathbf{y}(t)) \tag{3}$$

Where J is a scalar objective function accounting for product quality, process economy, and other important control objectives,

Control is constrained by the equality relations

$$\mathbf{h}(\mathbf{x},\mathbf{u}) = 0 \tag{4}$$

and inequality relations

$$\mathbf{q}(\mathbf{x},\mathbf{u}) \le 0 \tag{5}$$

In view of development of new robust numerical algorithms for nonlinear optimization, the main obstacle in application is difficulty in reliable choice of model functions $\boldsymbol{f}$ and parameters $\boldsymbol{\Theta}$. Derivation of models $\boldsymbol{f}$ and parameters $\boldsymbol{\Theta}$ on first principles and their validation is in general very time consuming and prone to numerous modeling errors. A great challenge is in possibility of replacement of the first principle modeling by "standardized" model structures, such as ARMAX, principle component (latent variable) LV-ARMAX, piece wise affine (PWA-ARMAX) models, nonlinear neural network (NN-ARMAX), and fuzzy logic (FL-ARMAX) models. Such models are derived by well established numerical and statistical procedures, and their development is based on extensive process data records which are readily available in modern computer controlled process plants.

The representative discrete model structure for linear ARMAX is given by:

$$y(k)=\sum_{i=0}^{N_y} a(i)\cdot y(k-i)+\sum_{i=0}^{N_u} b(i)\cdot u(k-i-m) \qquad [6]$$

Determination of sizes of time windows N_y and N_u are related to model order and discretization accuracy and is a part of modeling effort performed prior to the on-line optimization. The model parameters a_i and b_i are determined by the regressions method prior to optimization, and usually under open loop designed experiments. Also, the process delay m is assumed known or estimated parameter.

For optimal control by on-line MPC method considered are only predictions of the process model of a finite future interval H, called prediction horizon, for a given sequence of future manipulative input decisions over a finite C control horizon. The discrete form of the objective function is defined by the vector of predicted process outputs:

$$\mathbf{Y}(k)=[y(k+1),y(k+2)....y(k+H)]^T \qquad [7]$$

input manipulative variable decisions:

$$\mathbf{U}(k)=[u(k+1),u(k+2)....u(k+C)]^T \qquad [8]$$

and predetermined input variable decisions:

$$\mathbf{Y}^*(k)=[y^*(k+1),y^*(k+2)....y^*(k+H)]^T \qquad [9]$$

A discrete representation of the objective functional [3] is formulated as a scalar quadratic function:

$$J(k)=\min_U\left\{\begin{aligned}&[\mathbf{Y}(k)-\mathbf{Y}^*(k)]^T\cdot\mathbf{W}_Y\cdot[\mathbf{Y}(k)-\mathbf{Y}^*(k)]+\\&+[\mathbf{U}(k)-\mathbf{U}(k-1)]^T\cdot\mathbf{W}_U\cdot[\mathbf{U}(k)-\mathbf{U}(k-1)]\end{aligned}\right\} \qquad [10]$$

The control objective [10] sums up the effects of process deviations from the optimal desired process profile weighted by a known matrix $\mathbf{W}_Y$ of weighting coefficients, and the effects of deviations of consecutive manipulative variable moves weighted by $\mathbf{W}_U$.
The predicted states and control actions are constrained by lower and upper bounds given inequalities:

$$y_{\min}\le y(k)\le y_{\max} \quad \text{and} \quad u_{\min}\le u(k)\le u_{\max} \qquad [11]$$

Numerical optimization is performed by standard method (software) such as sequential quadratic programming (SQP).

The optimization is performed on-line in the moving horizon framework, in which only the first element (step) of the optimization sequence is applied, first component u(k) of the vector $\boldsymbol{U}$(k), is implemented, and the optimization procedure is repeated for the next time interval.

Process nonlinearity can be modeled by a set of piece wise linear model leading to PWA-MPC algorithm (2). A real challenge in modeling nonlinear processes based on extensive on-line measurement by process analyzers (spectrometers) is in extraction of essential information by principal component analysis PCA leading to nonlinear modeling by latent variables. Due to multidimensional spectral data sets reduction by projection to a space of latent variables leads to significant simplification of mathematical models with reduced sets of model parameters and increased robustness of model predictions. The space of control variables **U** and output variables **Y** are projected to corresponding lower dimensional target spaces **T** and **Z** with uncounted data variability contained in the residual matrixes $\mathbf{E}_U$ and $\mathbf{E}_Y$, which are mostly of random nature do to measurement errors (usually 1-5 % relative error).

$$\mathbf{U}(k) = \mathbf{T}(k) \cdot \mathbf{P}^T + \mathbf{E}_U(k) \qquad \mathbf{Y}(k) = \mathbf{Z}(k) \cdot \mathbf{Q}^T + \mathbf{E}_Y(k) \tag{12}$$

A linear model is established as an inner relation between target variables given by:

$$\mathbf{Z}(k) = \mathbf{T}(k) \cdot \mathbf{B} + \mathbf{E}(k) \tag{13}$$

where **B** are model parameters and **E** is a model error matrix. MPC is applied to target control variables **T** for optimization of the target output variables **Z**. Real values of control variables needed for action in the time span from k to *k+1* are obtained by the projection matrixes given in [12].

A general nonlinear structure based on latent variables is provided by PCA neural networks NN. Neural network models in the target spaces are:

$$\mathbf{Z}(k) = \mathbf{NN}(\mathbf{T}(k)) \tag{14}$$

where **NN** (neural network matrix) contains nonlinear neuron activity functions, for example given as "sigmoid" functions:

$$y_{out}(k) = \frac{1}{1 + e^{-w_0 - \mathbf{w}_1 \cdot \mathbf{x}_{in}(k)}} \tag{15}$$

where the model parameters are: w_0 a threshold of neuron activity; and $\boldsymbol{w}_1$ a vector of weighting coefficients for a corresponding neuron. Due to highly adaptive nature of multilayer structure of neurons, through network training PCA based NN networks can mimic very complex system behavior.

A very important alternative to model reduction by principal component analysis and nonlinearity of neural networks is offered by fuzzy logic modeling. Control *U* and state *Y* spaces are covered by fuzzy sets U_i and Y_i, usually associated (labeled) with corresponding linguistic variables. Each fuzzy set is defined as an order pair of a variable value *u*(k) and *y*(k) at time instant *k* and membership function μ which is a measure of degree of membership, in the range from 0 to 1, of association of a variable and a fuzzy set.

$$U_i = \{u(k), \mu(u(k)) | u(k) \in U\} \qquad Y_i = \{y(k), \mu(y(k)) | y(k) \in Y\} \tag{16}$$

Fuzzy logic control algorithm (system) integrates the following subsystems: fuzzification subsystems which transform measurement signals (crisp data) into fuzzy sets (linguistic variables). Linguistic variables are input to decision making subsystem which is supported by rule based subsystem and a data base subsystem, both are part of a fuzzy logic knowledge based control system. In the decision making subsystem are applied rules of fuzzy reasoning and results are inputs to defuzzification subsystem where fuzzy controls are transformed into crisp values of control moves for the next sampling interval from *k* to *k*+1. Fuzzy control systems have a number of important advantages over principal component and neural networks. The main advantages are transparency of control reasoning and decision making, incorporation of past controls and expert rules included into the control knowledge based system, and minimal computational load which makes them suitable for control with high sampling rate, i.e. for short control horizons leading to better control stability and robustness.

Results and Discussion

In about last 20 years, in literature is reported that MPC has been used in over 5000 industrial applications, most of them in chemical and electrical industry, and less in fermentation and food process industry (2).

For industrial process control practitioners, MPC has the following advantageous properties:

- it is a true multivariate control algorithm enabling integration of various plant information into process monitoring and control software,
- has ability to naturally and explicitly handle nonlinearities, time delays, and multivariable input and output constraints by direct incorporation into on-line optimization,
- enables early detection of process disturbances and prediction of malfunctions in process equipment improve the safety of the process, minimize the time and resources needed for maintenance, and increase the uniform quality of the products,
- the objective of online-monitoring is to trace the state of the process and the condition of process equipment in real-time, and to detect faults as early as possible,
- available are standard commercial numerical and statistical algorithms needed for MPC

Due to expanding global food market food industry needs to achieve higher levels of productivity, enhanced safety assurance and better quality products, which will expansion of process automation and intelligent on-line control which will inevitably advance application of MPC based control.

Literature

1. A. B. Ogunnaike, W.H. Ray: *Process Dynamics, Modeling and Control*, Oxford University Press, New Yor (2004).
2. M. Morari, M. Barić, *Comp. Chem. Eng*., 30 (2006) 1619-1631.
3. M.J. McGrath, J.F.O'Connor, S. Cummins, *J. Food Eng*., 35 (1998) 313-321.
4. G. Chen, T.J. McAvoy, M.J. Piovoso, *J. Proc. Cont*., 8 (1998) 139-149.
5. R. Simpson, A. Teixeira, S. Almonacid, *Food Control*, 18 (2007) 821-833.
6. T. McAvoy, Annual Reviews in Control, 26 (2002) 75-86.
7. J. Flores-Cerrillo, J.F. MacGregor, *J. Process Control*, 15 (2005) 651-663.
8. C. Bordons, J.R. Cueli, *J. Process Control*, 14 (2004) 305-315.
9. H. Zhang, B. Lennox, *J. Process Control*, 14 (2004) 41-50.
10. X. Liu, X. Chen, W. Wu, Y. Zhang, *Food Control*, 17 (2006) 894-899.