Information Sciences 355–356 (2016) 229–236

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A multivariate statistical combination forecasting method for product quality evaluation

Shen Yin [a,c], Lei Liu [b], Jian Hou [a,*]

[a] College of Engineering, Bohai University, Jinzhou 121013, China
[b] College of Mathematics and Physics, Bohai University, Jinzhou 121013, China
[c] School of Astronautics, Harbin Institute of Technology, Harbin 150001, China

ARTICLE INFO

Article history:
Received 2 July 2015
Revised 12 January 2016
Accepted 16 March 2016
Available online 4 April 2016

Keywords:
Prediction
Combination forecasting
Quality evaluation
Multivariate statistical methods

ABSTRACT

In this paper, a multivariate statistical combination forecasting method is proposed for key performance evaluation in process industry. This method is developed based on some of the most popular multivariate statistic approaches. It merges the advantages of the principal component regression method (PCR), the partial least squares regression method (PLSR) and the modified partial least squares regression method (MPLSR). We test the proposed method with a numerical example and also an actual wine production process. The results indicate that the prediction accuracy of the optimal combination forecasting method is superior to those of individual methods.

1. Introduction

Over the past years, forecasting has played a crucial role in performance evaluation in process industry. Many techniques have been presented to deal with the related issues in forecasting, classification and detection, including multivariate statistic methods, general exponential smoothing, stochastic time series and support vector machines [8], etc. Furthermore, the filter theory has been successfully applied in the forecasting branch. In this direction, many filtering methods have been proposed in the literature. While in [14] a continuous-time Kalman filtering method is proposed for a class of linear, uncertain time-lag systems with randomly jumping parameters, other works, e.g., [11,13,17,24,26], concentrate mainly on the H_∞ filter. In addition, [19] focuses on the problem of robust H_∞ filtering for a class of systems with parametric uncertainties and unknown time delays under sampled measurements and proposes an approach to design H_∞ filters. On the other hand, [9] is concerned with the problem of H_∞ fuzzy filtering of nonlinear systems with intermittent measurements and focuses on the design of an H_∞ filter to make the filter error system stochastically stable and preserve a guaranteed H_∞ performance. Moreover, some fault detection filters are designed in [12,25,30,32]. Obviously, forecasting model is helpful to solve some industrial problems, e.g., performance evaluation [10,20,21].

Due to the extreme importance of product quality evaluation, a lot of related methods have been proposed. The combination forecasting has been a hot research topic in this field and achieved impressive results. To some extent, combination forecasting methods with prediction accuracy as the objective function are able to improve the prediction accuracy. Some methods have been proposed to determine the weighting coefficients of individual methods in combination forecasting. These works help choose suitable and operable forecasting methods and improve the accuracy of prediction and decision

* Corresponding author. Tel.: +86 4163719182.
E-mail addresses: shen.yin2011@gmail.com (S. Yin), bhushxll@gmail.com (L. Liu), dr.houjian@gmail.com (J. Hou).

http://dx.doi.org/10.1016/j.ins.2016.03.035

making. In our point of view, however, these methods can be further improved by making use of some intelligent methods, e.g., multivariate statistic methods.

In this paper, the multivariate statistical combination forecasting methods are used to evaluate wine quality. Some typical multivariate statistical methods include principal component regression method (PCR) [27], partial least squares regression method (PLSR), and modified partial least squares regression method (MPLSR) [29]. The combination forecasting methods are able to merge the advantages of individual multivariate methods and use weighted combination to generate better prediction results. Therefore different weighting methods, including the optimal combination forecasting methods and the sub-optimal combination forecasting methods, are studied in this paper. The optimal combination forecasting models concentrate on the criteria of minimizing error sum of squares (MESS), minimizing the sum of absolute error (MSAE) and minimizing the maximum absolute error (MMAE). Whereas the sub-optimal combination forecasting ones include arithmetic average value method (AAVM), the reciprocal prediction error sum of squares method (RPESSM), the reciprocal mean square error method (RMSEM), simple weighted average method (SWAM) and binomial coefficient method (BCM).

The remainder of this paper is organized as follows. Section 2 introduces the combination forecasting methods, the optimal combination forecasting methods and the sub-optimal combination forecasting methods briefly. In Section 3, a numerical example is presented to illustrate the proposed method, and then the individual forecasting models and combination forecasting methods are presented in details and their prediction abilities are analyzed. Finally, Section 4 concludes this paper.

2. Combination forecasting

The starting point of combination forecasting is that different forecasting models are combined to obtain better forecasting results than any individual one. Different methods usually characterize systems from different and somewhat complementary perspectives. This makes it possible to combine different methods to obtain better results. Existing forecasting methods can be categorized in the following ways.

(1) Based on the function relationship between combination forecasting methods and individual forecasting methods, combination forecasting can be categorized into linear combined forecasting and nonlinear combined forecasting.
(2) According to the computing methods of weighting coefficients, the combination forecasting methods can be divided into optimal combination forecasting ones and sub-optimal combination forecasting ones.
(3) Based on whether the weighting coefficients vary over time or not, the combination forecasting can be classified into permanent weight combined forecasting and changeable weight combined forecasting.
(4) According to the forecasting quality evaluated by some criteria, combination forecasting can be classified into non-inferior combination forecasting and superior combination forecasting.

As a widely accepted way to improve forecasting accuracy, combination methods are proposed to combine the individual forecasts. Apparently, combination forecasting methods can be built based on the given information. It can collect the information of each individual forecasting method to combine. The optimal combination forecasting methods are direct and it is easy to obtain the weighting coefficients. The sub-optimal combination forecasting methods are easier than the optimal combination forecasting ones, but perform inferior to the latter. So in this section, the optimal combination forecasting methods and the sub-optimal combination forecasting ones are introduced, and we focus on the optimal combination forecasting methods.

2.1. The optimal combination forecasting methods

The combination forecasting methods use the weighted individual forecasting methods to make the prediction more accurate. Therefore, the key is to obtain the appropriate weighting coefficients. In this article, the optimal combination forecasting methods are mainly concerned. The optimal combination forecasting methods are built based on three criteria, i.e., MESS[18], MSAE [6] and MMAE [3].

2.1.1. Minimizing error sum of squares (MESS)

Prediction error sum of squares (PESS) is one of the important indicators to reflect the prediction accuracy. It is widely used in practical prediction problems.

In general, $[x_1, x_2, \cdots x_n]$ is the index sequence of the forecasting objects. Let x_{it} denote the forecasting output of the ith individual forecasting model at time point t for time series x_t. $\omega = [\omega_1, \omega_2, \ldots, \omega_n]$ is the weighting coefficient vector in the combination forecasting model. In order to ensure that the combined model is reasonable, the constrains that $\sum_{i=1}^{n} \omega_i = 1, 0 \le \omega_i \le 1$ must be satisfied [4]. The residual of the combination model is denoted as $e_i = x_t - \hat{x_t} = \sum_{i=1}^{n} \omega_i e_{it}$, and x_t actually is the vector of the wine qualities evaluated by the wine taster. The PESS Q_1 is denoted by $Q_1 = \sum_{t=1}^{N} e_t^2 = \sum_{t=1}^{N} \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i e_{it} \omega_i e_{jt}$, where n is the number of individual forecasting methods. N is the number of the time points.