BEST: International Journal of Humanities, Arts, Medicine and Sciences (BEST: IJHAMS)
ISSN (P): 2348-0521, ISSN (E): 2454-4728
Vol. 5, Issue 12, Dec 2017, 65-70

OPTIMIZATION OF RSM CALIBRATION MODELS FOR MULTIPLE PROCESSES, IN THE SUGAR INDUSTRY

JOHN WESLEY. S[1], RANIES CHARLES SELVARAJ. V[2] & S. ELIZABETH AMUDHINI STEPHEN[3]

[1 & 2] M. Tech Food Processing & Engineering, Department of Biosciences & Technology, Karunya University, Coimbatore, India

[3]Associate Professor, Mathematics, Karunya University, Coimbatore, India

ABSTRACT

The measurements of Response surface methodology, combined with data analysis techniques, are widely used for quality control in food production processes. This paper presents a methodology to optimize the calibration models of RSM in four different stages, in a sugar factory. The models were designed for quality monitoring, particularly °Brix and Sucrose, and both common parameters in the sugar industry. The proposed models improve the prediction for the test set (unseen data), compared to the previously published models, resulting in a more accurate quality assessment of the inter-mediate products of the process, in the sugar industry.

KEYWORDS: Quality Control, Calibration Models, Brix and Sucrose & Quality Monitoring

INTRODUCTION

The production flow in the sugar industry encompasses several processes and sub processes that need to be analyzed, in order to maintain a quality standard [1]. The agro-industrial plants require cost-efficient and non-destructive systems, to monitor the quality of their production process, food safety and compliance with the technical specifications [2]. One of these non-destructive systems aimed at ensuring quality is chemo metrics, which has been developing since the 1970s as an interdisciplinary field of study. This field covers a wide and varied range of mathematical and statistical techniques, for analyzing the chemical composition of materials [3]. To analyze the quality of organic raw materials, a commonly-used technique is the RSM associated with chemo metrics; however, the relationship between the absorption in the spectral region of the near infrared and the analyte is frequently of a non-linear type [4]. The origin of these non-linear relationships is diverse and difficult to identify, in some cases due to the differences in viscosity, temperature, pH, particle size and the chemical nature of the analyte. For this reason, calibration is commonly performed using non-linear methods and multivariate analysis [5]. A proper selection of the variables aimed at gathering a small subgroup with lower sensitivity to non-linearities or at discarding the most pronounced wavelengths is usually effective, to improve the performance of the models [6,7]. A recent study conducted by Tange et al. [9] has shown that, the use of calibration models with Support Vector Machines (SVM) for regression is efficient in order to predict °Brix and Sucrose values, the quality parameters of the industrial process of sugar. The use of SVM improves in terms of RMSE compared to the technique of Partial least Square (PLS); however, the proposed model uses the entire RSM, which leads us to believe that the optimization of the model is still possible, by implementing an appropriate preprocessing technique, feature selection and optimization of the parameters of the machine support vectors.

The aim of this study is to optimize the global calibration RSM models, in order to improve quality control of the °Brix and Sucrose parameters. A global calibration RSM model is capable to predict a value in the sugar production process.

MATERIALS AND METHODS

Generally, the sample processing consists of the following steps: acquisition of spectral data, preprocessing of the data to reduce the noise, thereby increasing the signal-to-noise ratio (S/N) [10], selection of relevant features, and development of the calibration model using a set of spectra from which the values of target analytes, obtained by reference techniques are known, and, finally, the model validation using data different to those of calibration [8].

Data Description

The database was published by Tange et al. [9], and the employed data were obtained in a Japanese sugar factory (Daito Togyo Co), where sugar cane is processed. The samples were obtained throughout three months in the harvest season, and during each of the process steps: after grinding (juice), after the process of evaporation (syrup), after crystallization (massecuite) and after centrifugation (molasses), three cycles of crystallization and centrifugation were carried out, resulting in a higher number of samples of massecuite and molasses than those obtained in the other two stages. Immediately after sampling, the NIR signals were extracted, along with the reference technique in relation to process temperature.

The °Brix quality parameter expresses all the dissolved solids (sugar and non-sugar), as a percentage of total weight, its scale reflecting the percentage of sucrose in pure solutions. In any sugar materials (juice, honey, etc.) the °Brix parameters are always higher than those of Sucrose, whereas in high-purity materials, such as spirits from a refinery, the difference between these indicators is minimal. For the current work, the °Brix quality parameter was measured using an Abbemat-WR refractometer, developed by Anton Paar GmbH in Germany.

The Sucrose quality parameter (POL) refers to the amount of sucrose contained in a solution, expressed as percent of the weight; in pure solutions, the POL percentage is equivalent to the percentage of Sucrose, while in other impure solutions, like cane juice and honey, there is a difference between these two values, the more impure is the solution, the higher the difference. For this reason, the POL value is internationally accepted as apparent sucrose. In the current paper, Sucrose was measured using a MCP500 polarimeter, developed by Anton Paar GmbH in Germany.

RESPONSE SURFACE METHOD

Response surface methodology uses statistical models, and therefore practitioners need to be aware that even the best statistical model is an approximation to reality. In practice, both the models and the parameter values are unknown, and subject to uncertainty on top of ignorance. Of course, an estimated optimum point need not be optimum in reality, because of the errors of the estimates and of the inadequacies of the model.

Nonetheless, response surface methodology has an effective track-record of helping researchers improve products and services: For example, Box's original response-surface modeling enabled chemical engineers to improve a process that had been stuck at a saddle-point for years. The engineers had not been able to afford to fit a cubic three-level design to estimate a quadratic model, and their biased linear-models estimated the gradient to be zero. Box's design reduced the costs of experimentation so that a quadratic model could be fit, which led to a (long-sought) ascent direction.

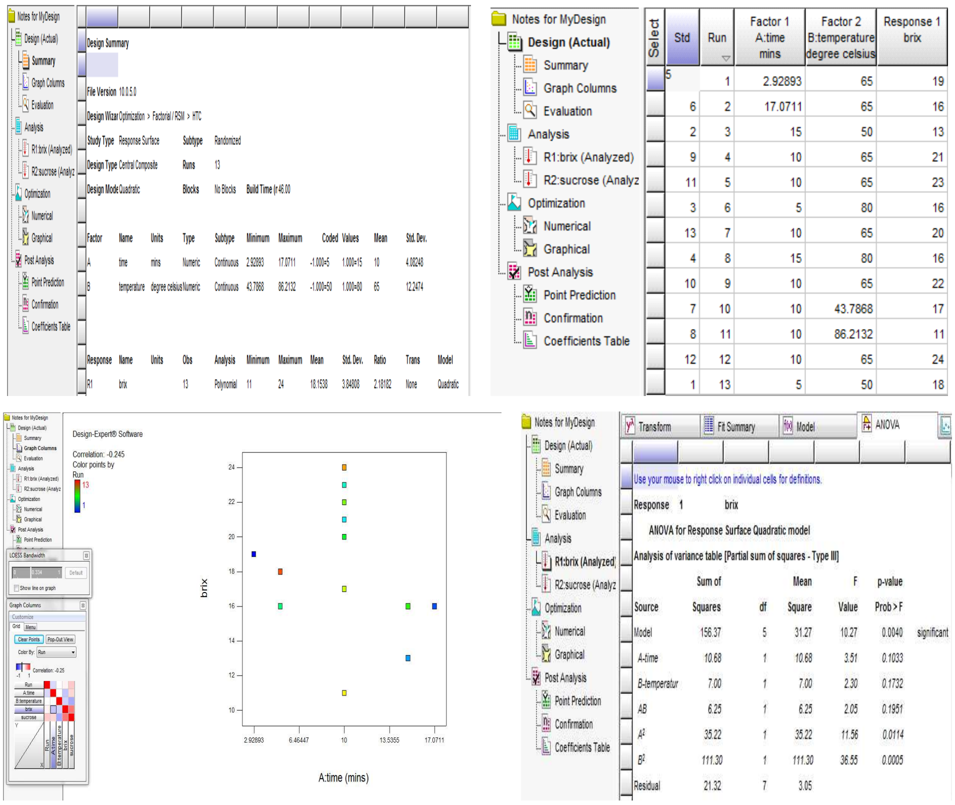
RESULT AND DISCUSSIONS

This study presents a methodology which optimizes RSM calibration models for multiple processes in the sugar industry. This consists of a sequence of steps described below:

At this point, the model is evaluated and its optimal values are set, in order that the C and γ parameters of SVR can be optimized using a grid search technique. Using these optimized values, a tuning of the parameter ε of SVR, Response surface methodology in which central composite is used with 13 runs.

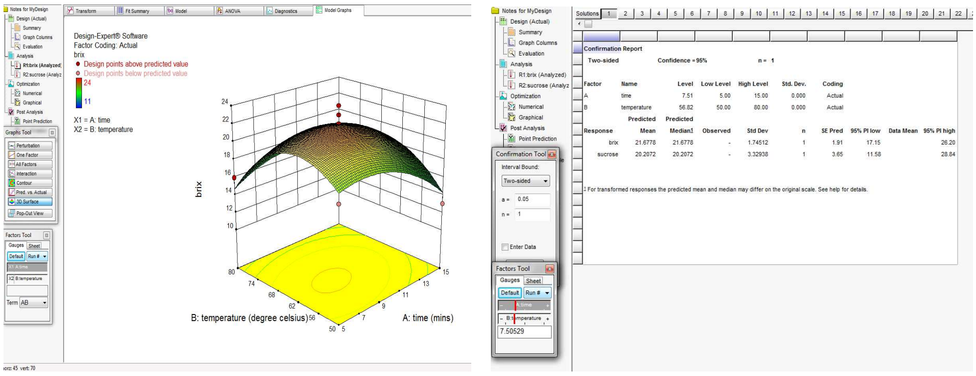
Independent variable: Time, temperature (◦c).

Dependant variable: Brix



Final Equation in Terms of Coded Factors:

brix =
+22.00
-1.16 * A
-0.94 * B
+1.25 * AB
-2.25 * A^2
-4.00 * B^2



The results of this study were compared to those published by Tange et al. [9], who used SVR to carry out the model calibration for °Brix and Sucrose in the manufacturing process of sugar cane. The results presented herein obtained lower RMSE values than Tange et al. [9], who presented more accurate estimates of the quality parameters. This is explained by the fact that in the present work techniques of preprocessing and selection of features were used, and therefore noise spectra were removed, along with the wavelengths which did not contribute significantly to the model.

The optimization of parameters of SVR proved its importance to obtain minimum RMSE for the model. In our case, it was performed by using a search grid technique, providing the optimal combination of parameters C, γ and ε, which is consistent with the results obtained by Jeng [13] and by Devos et al. [12] who claimed that, the combined values of the parameters of SVM, determined the complexity of the limits and therefore, the performance of the model.

CONCLUSIONS

In this regard, one can deduce that the proposed model is accurate and stable, due to the parameter optimization, which is consistent with the results obtained by Cristianini and Shawe [11] and Devos et al. [12] who stated that, the adjustment of the SVM kernel parameters controlled the complexity of the resulting hypothesis and avoided the over fitting of the model. The evaluation of the models was performed using the repeated cross-validation technique, which, according to Garcia and Filzmoser [14] leads to a suitable method, aimed at choosing the best model, to analyze the mean and standard deviation of the results of repetitions; these results correspond to the test data set, that is, they are data which were not used for calibration, which allows estimating on how the model would behave in the future with new data.

REFERENCE

1. L. S. Polanco, D. F. Day, S. Savoie, S. Bergeron, T. Charlet, B. L. Legendre, Improvements of Raw Sugar Quality Using Double Purge of C-Massecuites Performance Comparison, in: lsu Ag Center Audubon Sugar Institute Factory Operations Seminar, lsuagcenter.com, 2014, p. 46.
2. C. Kumaravelu, A. Gopal, A review on the applications of Near-Infrared spectrometer and Chemo metrics for the agro-food processing industries, 2015, pp. 8–12.
3. B. R. Kowalski, Chemom., Anal. Chem. 52 (1980) 112R–122R.
4. E. Bertran, M. Blanco, S. Maspoch, M. C. Ortiz, M. S. Sánchez, L. A. Sarabia, Handling intrinsic non-linearity in near-infrared reflectance spectroscopy, Chemom. Intell. Lab. Syst. 49 (1999) 215–224.
5. H. Martens, T. Naes, Multivariate Calibration, Wiley, 1992.

6. R. Leardi, R. Boggia, M. Terrile, Genetic algorithms as a strategy for feature selection, J. Chemom. 6 (1992) 267–281.
7. Y. Saeys, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, Bioinformatics 23 (2007) 2507–2517.
8. H. Cen, Y. He, Theory and application of near infrared reflectance spectroscopy in determination of food quality, Trends Food Sci. Technol. 18 (2) (2007) 72–83.
9. R. I. Tange, M. A. Rasmussen, E. Taira, R. Bro, Application of support vector regression for simultaneous modelling of near infrared spectra from multiple process steps, J. Infrared Spectrosc. (2015) 〈http://www.forskningsdatabasen.dk/ en/catalog/2266200601〉.
10. C. D. Brown, P. D. Wentzell, Hazards of digital smoothing filters as a preprocessing tool in multivariate calibration, J. Chemom. 13 (1999) 133–152.
11. N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge, 2000.
12. O. Devos, C. Ruckebusch, A. Durand, L. Duponchel, J.-P. Huvenne, Support vector machines (SVM) in near infrared (NIR) spectroscopy: focus on parameters optimization and model interpretation, Chemom. Intell. Lab. Syst. 96 (2009)27–33.
13. J.-T. Jeng, Hybrid approach of selecting hyper parameters of support vector machine for regression, IEEE Trans. Syst. Man Cybern. B Cybern. 36 (2006)699–709.
14. H. Garcia, P. Filzmoser, Multivariate Statistical Analysis using the R package chemometrics, Vienna, Austria,2015〈 ftp://155.232.191.229/cran/web/packages/chemometrics/vignettes/chemometrics-vignette.pdf〉.