**Machine Learning-Based Soft Sensor for a Sugar Factory’s Batch Crystallizer**

Mohammad Reza Boskabadi, Abhishek Sivaram, Gürkan Sin, Seyed Soheil Mansouri\*

*Department of Chemical and Biochemical Engineering, Søltofts Plads, Building 228A, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

\*seso@kt.dtu.dk

Abstract

In the contemporary industrial landscape, the utilization of real-time data for the surveillance and enhancement of operational processes stands as an imperative, contributing significantly to the refinement of operational efficiency and product quality across a multitude of sectors. This work presents the development of a machine-learning soft sensor utilizing Multivariate Linear Regression (MLR), Generalized Regression Neural Network (GRNN), Decision Tree, and Support Vector Regression (SVR) based on real historical data from a sugar factory. The soft sensor is designed to estimate the Brix index in the vacuum batch crystallizer. Various models have demonstrated robust performance in predicting Brix sensor values. The framework involves three key steps: data pre-processing, model construction employing selected algorithms, and the evaluation of well-performing models. While non-linear techniques, specifically Generalized Regression Neural Network (GRNN) and Decision Trees, exhibited superior performance in line with evaluation criteria, linear methods, such as Multivariate Linear Regression (MLR), closely matched the effectiveness of these advanced approaches.

**Keywords**: Data-driven Modelling, Soft Sensor, Plant-Wide Operation, Crystallization

* 1. Introduction and process description

Sugar production from sugar beets is a complex and intricate process that involves several crucial stages. At the heart of this process is a sugar factory equipped with various specialized units. In Figure 1, we present an overview of the key stages in sugar beet processing. The journey begins with the arrival of trucks at the weight bridge, where the beets undergo initial assessments. Subsequently, the beets proceed through a series of stages including sampling, unloading, beet washing, slicing, and extraction tower. The extracted juice undergoes purification in the juice purification unit before moving through stages such as evaporation, crystallization, centrifuging, and drying. The final product is then stored in silos before being dispatched from the service center. Each stage plays a pivotal role in the overall sugar production process, contributing to the high-quality sugar that results from this intricate industrial sequence. A comprehensive understanding of each stage is essential for optimizing production efficiency and ensuring the consistent quality of the final product.



*Figure 1- Sugar factory schematic from sugar beet to final product including sampling, washing, extraction, juice purification, evaporation, crystallization, centrifugation, silos and service center. (Credit NurdZucker)*

Between the above-mentioned processes, crystallization stands out as one of the most pivotal unit operations in sugar production, exerting a profound influence on the final product specifications, energy consumption, and the overall carbon footprint of the factory. The sugar beet crystallization process is a multifaceted operation, characterized not only by the transformation of syrup into massecuite but also by its inherently nonlinear and non-stationary nature (Meng, et al., 2019). The complex interactions between nucleation, growth, and aggregation mechanisms of crystals unfold during this process. Despite its paramount significance, the sugar beet crystallization process lacks a reliable mechanistic model that comprehensively describes the intricate relationship between crystallization dynamics and operating conditions. Consequently, the control strategies employed in sugar beet crystallization often hinge on online measurements of key process parameters. It is noteworthy that any deviation from the normal operation of the sugar crystallization process can have cascading effects on the plant. Anomalies might prompt the centrifuge to reintroduce small crystals to different batch boilers, necessitating additional energy input and consequently elevating CO2 production. In essence, the balance of sugar crystallization underscores the critical need for a nuanced understanding and effective control measures to optimize efficiency and minimize environmental impact.

In sugar production, the crucial crystallization stage involves interconnected processes. This study employs two strategies to enhance this operation: leveraging insights from mother liquor crystal size distribution through process optimization and first-principle modelling, and integrating data from various sensors to design a robust soft sensor for superior process control. Soft sensors serve as a reliable backup in case of sensor failures, crucial for uninterrupted operation and adaptability to changes. Their significance is particularly evident in complex processes like sugar crystallization, where continuous control is vital for optimal performance and resource efficiency. However, the lack of soft sensors for batch and continuous processes remains a major challenge in implementing various control strategies (Paengjuntuek et al., 2008). This work delves into the intricacies of the developed soft sensor for the Brix index in sugar crystallization batch unit operations, where the Brix index is a key parameter for the control system, representing the mass of total dissolved solids per 100 mass units of solution.

* 1. Data-driven modeling

*Table 1- Input and output variables for the data-driven model in the sugar-crystallization unit*

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Description | Range | Unit |
| *X1* | Temperature | [50,90] | ˚C |
| *X2* | Vacuum Pressure | [0.15,0.35] | bar |
| *X3* | Steam Pressure | [0.4,1.5] | bar |
| *X4* | Level | [0,400] | Cm |
| *Y1* | Brix Index | [70,100] | Bx |

|  |  |
| --- | --- |
| Index | Formula |
| MAPE | $$MAPE=\frac{\sum\_{i=1}^{m}\frac{\left|y\_{i}-f\_{i}\right|}{y\_{i}}}{m}×100\%$$ |
| MSE | $$MSE= \frac{1}{n}\sum\_{i=1}^{m}(y\_{i}-f\_{i})^{2}$$ |
| R2 | $$R^{2}= \frac{\sum\_{i=1}^{m}(y\_{i}-f\_{i})^{2}}{\sum\_{i=}^{m}(y\_{i}-\overbar{y\_{i}})^{2}}$$ |
| MAE | $$MAE=\frac{\sum\_{i=1}^{m}\left|y\_{i}-f\_{i}\right|}{m}$$ |

In the development of soft sensors tailored to operational parameters, the dataset underwent a meticulous reduction and categorization process, with distinct batches assigned to elucidate the operational nuances inherent in the data. Table 1 delineates the input and output variables within the final dataset. Employing a versatile approach, the soft sensor design employed various regression techniques, including Linear Regression, Neural Network, Support Vector Machine, and Decision Tree models. To assess the efficacy of these models, the Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Determination Coefficient (R2), and Mean Absolute Error (MAE) served as key performance metrics. The formulae for each index are detailed in Table 2. The ensuing section delves into an exploration of the machine-learning algorithms implemented for soft sensor design.

*Table 2- Model performance evaluation index and their identification*

* + 1. Multivariate Linear Regression (MLR)

Multivariate Linear Regression (MLR) (Höskuldsson, 1996) stands out as an extensively utilized regression technique in the realm of soft sensor design(Wang et al., 2009). This method intricately establishes a linear polynomial relationship between the predictor variable X and the response variable y. The vector X encompasses diverse types of data, including monitoring data, electrochemical data, and univariate operational process data such as temperature, vacuum pressure, steam pressure, viscosity, density, flow rate, and more, denoted as bk (k = 1, …, m). Here, b = [b1, b2, …, bm] represents the regression coefficients, while ε denotes the residual.

|  |  |
| --- | --- |
| $$y=b\_{1}x\_{1}+b\_{2}x\_{2}+…+b\_{m}x\_{m}+ε$$ | (1) |

* + 1. Generalized Regression Neural Network (GRNN)

The utilization of neural networks is a prevalent approach in the development of soft sensors (Kadlec et al., 2009). Specht (Specht, 1991)originally proposed this method for non-linear function approximation. In the general regression algorithm, the relationship between input and output is articulated through a probability density function derived from observed data, as represented by Eq. (2). The Euclidean distance between two input vectors, denoted as D2, is defined by Eq. (3). A typical GRNN comprises four layers: an input layer, a pattern layer, a summation layer, and the output layer. The input layer contains neurons equal to the number of input variables, while the pattern layer features neurons corresponding to the number of training cases.

|  |  |
| --- | --- |
| $$\hat{y}= \frac{\sum\_{i=1}^{m}y\_{i} ×exp⁡(\frac{-D\_{i}^{2}}{2σ^{2}})}{\sum\_{i=1}^{n}exp(\frac{-D\_{i}^{2}}{2σ^{2}})}$$ | (2) |
| $$D\_{i}^{2}=(x-x\_{i})^{T}(x-x\_{i})$$ | (3) |

* + 1. Support Vector Regression(SVR)

Owing to its strong foundation in statistical learning theory, Support Vector Machines (SVM) have garnered growing interest among soft sensor developers (Meng, et al., 2019). The overarching SVR algorithm was initially introduced by (Drucker et al., 1996) The initial phase of the SVR algorithm involves creating a linear function in a high-dimensional space using a kernel function, as depicted by Eq. (4). Here, φ() denotes a kernel function that maps the m-dimensional data to a higher-dimensional space. The objective is to optimize the parameters ω and b, as outlined in Eq. (5).

|  |  |
| --- | --- |
| $\hat{y}=ω^{T}$ φ(x)+b | (4) |
| $$\left|y-ω^{T}φ\left(x\right)-b\right|< ϵ$$ | (5) |

* + 1. Decision Tree Regression(DTR)

A decision tree is a hierarchical structure that categorizes instances by recursively splitting them based on input variables. During training, an algorithm optimizes the tree's fitness by minimizing errors between predicted and actual values(Loh, 2011). This study applies a regression model to the target variable using each independent variable. The dataset undergoes multiple splits, and at each split, the algorithm selects the variable resulting in the lowest error for further division. This iterative process is repeated to construct the decision tree

* 1. Result and discussion

To create a strong soft sensor for sugar crystallization, operational data from a real sugar factory was thoroughly analyzed. The factory used batch-based crystallization designs and a sophisticated process control system monitoring key parameters at one-minute intervals. Initial data analysis involved separating batches based on operational parameters. Time-series data were labeled, and a correlation matrix was used to identify relationships among various sensors, particularly for predicting Brix sensors (depicted in Figure 2). Recognizing the importance of Brix sensors in determining crystallization completion time, regression models were then employed to predict Brix sensor behavior based on data from other sensors—a crucial step toward developing a data-driven soft sensor for real-time process optimization.

*Figure 2- Correlation matrix between operational parameterspositive value means direct correlation and negative value shows reverse correlation.*

*Table 3- Model performance evaluation index and their identification*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Index/Method | MLR | GRNN | SVR | DTR |
| MAPE(%) | 1.00 | 0.47 | 0.88 | 0.44 |
| MSE | 1.98 | 0.67 | 2.17 | 0.77 |
| R2 | 0.91 | 0.97 | 0.90 | 0.97 |
| MAE | 0.79 | 0.38 | 0.68 | 0.35 |

In the initial phase, a correlation matrix was used to explore relationships among various parameters in a vacuum batch crystallizer, revealing a significant correlation of 0.95 between the level and Brix sensors. Recognizing the crucial role of Brix sensors in indicating batch cessation, a strategic decision was made to develop a Brix soft sensor based on the level parameter. Various regression models (MLR, GRNN, SVR, DTR) were employed and rigorously assessed, using 40% of the dataset for training and 60% for testing. The dataset comprised 6772 data points and 60 batches.

|  |  |
| --- | --- |
| A graph with a red line  Description automatically generated | A graph with blue dots  Description automatically generated |
|  a |  b |
| A graph showing a line graph  Description automatically generated with medium confidence | A graph showing a line of a graph  Description automatically generated with medium confidence |
|  c d*Figure 3- Actual and predicted Brix index from. a) MLR b) GRNN c) SVR d) DTR*Figure 3- Actual and predicted Brix index from. a) MLR b) GRNN c) SVR d) DTR |  |

 As depicted in Table 3, various models demonstrated their capability to predict Brix sensors based on the input from the level sensor in the crystallizer. In scenarios where the Brix sensor data is unavailable for a particular batch in the crystallizer, the utilization of a soft sensor becomes essential for Brix prediction. Figure 3 shows the result of different regression models for brix index prediction.

* 1. Conclusions

In this research, a data-driven method for creating a customized soft sensor that predicts the Brix index in real-time during the crystallization stage in a sugar factory has been introduced. The model was trained using operational data from an industrial sugar factory, and various machine-learning methods were applied. Non-linear methods like GRNN and DTR showed superior performance, although linear methods like MLR performed similarly well, except for predicting low massecuite levels in the crystallization tank, which has minimal impact on the overall evaluation due to the crystallizer level typically being above 20 Decimeters during most operations. The models generated by these machine learning approaches were thoroughly tested with different datasets, consistently demonstrating robust performance. This comprehensive assessment highlights the effectiveness of both non-linear and linear methods in constructing accurate soft sensor models for real-time Brix index prediction in the sugar crystallization process.

* 1. Acknowledgement

This project receives funding the European Regional Development Fund (REACT-EU RF-21-0025 “Biosolutions Zealand”) for providing financial support to S.S.M.; and The Technical University of Denmark.

References

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in Neural Information Processing Systems*, *9*.

Höskuldsson, A. (1996). *Prediction Methods in Science and Technology.: Vol 1. Basic theory*.

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven soft sensors in the process industry. *Computers & Chemical Engineering*, *33*(4), 795–814.

Loh, W. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *1*(1), 14–23.

Meng, Y., Lan, Q., Qin, J., Yu, S., Pang, H., & Zheng, K. (2019). Data-driven soft sensor modeling based on twin support vector regression for cane sugar crystallization. *Journal of Food Engineering*, *241*, 159–165.

Meng, Y., Yu, S., Zhang, J., Qin, J., Dong, Z., Lu, G., & Pang, H. (2019). Hybrid modeling based on mechanistic and data-driven approaches for cane sugar crystallization. *Journal of Food Engineering*, *257*, 44–55.

Paengjuntuek, W., Arpornwichanop, A., & Kittisupakorn, P. (2008). Product quality improvement of batch crystallizers by a batch-to-batch optimization and nonlinear control approach. *Chemical Engineering Journal*, *139*(2), 344–350.

Specht, D. F. (1991). A general regression neural network. *IEEE Transactions on Neural Networks*, *2*(6), 568–576.

Wang, D., Liu, J., & Srinivasan, R. (2009). Data-driven soft sensor approach for quality prediction in a refining process. *IEEE Transactions on Industrial Informatics*, *6*(1), 11–17