Data-based Design of Inferential Sensors for an Industrial Depropanizer Column with Data Pre-treatment Analysis

Martin Mojto[a,*], Karol Ľubušký[b], Miroslav Fikar[a], Radoslav Paulen[a]

[a]Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Radlinského 9, 812 37 Bratislava, Slovakia, +421 2 59 32 53 49, martin.mojto@stuba.sk
[b]Slovnaft, a.s., Vlčie hrdlo 1, 824 12 Bratislava, Slovakia

Keywords: Data Treatment, Inferential Sensor, Regression

Abstract: The potential of data-based modeling techniques gains importance in the chemical industry whereas the development of first-principles models requires too much effort due to the complex physics and chemistry. The data-based inferential (or soft, surrogate) sensors play a significant role in this field. Such sensors can accurately and frequently estimate the hard-to-measure variables (e.g., product quality) by using the data from other online sensors (e.g., temperature or pressure sensors) within the same plant. In this contribution, we focus on developing the inferential sensors for the industrial depropanizer column, which is a part of the Fluid Catalytic Cracking unit in the oil refinery Slovnaft, a.s. in Bratislava, Slovakia.
The raw industrial data usually involve systematic errors and outliers representing, for example, the process shutdowns due to the maintenance. The presence of outliers can negatively affect the accuracy of the designed inferential sensors. Therefore, it is essential to perform the data treatment before the data is used in further analysis. We use the minimum covariance determinant (MCD) approach which searches for the set with a maximum number of measurements in which the covariance matrix has a minimum determinant. The rest of the measurements are considered outliers and are not included in the inferential sensor design.
The inferential sensors are designed using several methods. Initially, the ordinary least squares regression (OLSR) is applied to provide the most accurate inferential sensor on the training dataset. Subsequently, the OLSR approach is compared with two more advanced design methods searching for the optimal compromise between the accuracy and sensor structure complexity. The first method is principal component regression (PCR), exploring the covariance of the training dataset. The second method is the least absolute shrinkage and selection operator (LASSO), which eliminates the input candidates with the smallest contribution within the structure of the inferential sensor.
We compare the performance of the approaches above with the reference inferential sensor currently used in the refinery. The comparison is carried out on the testing dataset, which is not used during the inferential sensor design. The results indicate that the designed inferential sensors outperform the accuracy of the reference sensor by around 20%.

Acknowledgments: This research is funded by the Slovak Research and Development Agency under the projects APVV-20-0261 and APVV SK-FR-2019-0004 and by the Scientific Grant Agency of the Slovak Republic under the grants VEGA 1/0691/21 and VEGA 1/0297/22.