2021-2022
Second semester

University of Bahrain
College of Engineering
Department of Chemical Engineering

INDUSTRIAL STATISTICAL DATA ANALYSIS FOR DEBUTANIZER COLUMN

Abstract

The main aims of this project were to study the effect of numerous input variables in a data set on the output that was required to model. Our results regarding the data analysis were very preside and all analyses concluded that all the variables influence the output butane especially the temperature and the flow rate, which is expected. Therefore, it was extremely beneficial to check whether based on the given data sets, can a soft sensor be modelled to continuously determine the quality of the product or not. Several methodologies from different sources were referred to and studied such as the principal component analysis and Akaike Information Criterion to study the time series effect and select the important input variables to model our soft sensor and eliminate the variables that had no effect on butane. It was essential to consider a time series model to model the soft sensor and data analysis was a crucial step to achieving our goal

Objective and Motivation

This research aims to develop an online composition technique that eliminates the significant time delay caused by laboratory samples while also improving online monitoring estimation, by employing an online dynamic neural network to predict the n-butane composition at the top of an industrial debutanizer column in real time. Because the plant generates a variety of data, extensive screening was performed to choose the appropriate input variables for the dynamic model. The Akaike Information Criterion, Principal Component Analysis, and Pearson Correlation were used as model adequacy tests to predict the composition of n-butane in the column, Low Root Mean Square Error (RMSE) analysis was also done to suggest superior prediction and all data sets were standardized to assure consistency and minimize any unexpected and erroneous deviations.

Design and Implementation

It was extremely crucial to do some sort of industrial data analysis for the debutanizer unit for numerous essential reasons. One reason is to study the most effective input variables on the output variable which is the volumetric flowrate of butane in BPH. Another reason is that in the debutanizer unit, there were several variables in the input as well as the output. Therefore, it was extremely crucial for us to study thoroughly each variable and how each one had an effect on the flow rate of butane. With lots of research, several analyses were studied, and implemented in our project. However, before starting any data analysis, we were able to standardize the values we received from the industry in order just to keep the units consistent in case of the inconsistent unit. This inconsistency in the units of the variables could show us that one variable is highly affected by another variable whereas, it might just be because the units are not consistent it is showing a deviation.

the different types of data analysis to do our research such as the Akaike Information Criterion, Principal Component Analysis, and Pearson Correlation. This was done to ensure that take the best decision into choosing the most important input data sets and eliminating any data set that has no effect on the output.

we also used this R-studio software to try to do a test modelling for a dynamic soft sensor with the help of the theory behind machine learning and artificial neural network using the final data sets that we achieved from the data analysis. Designing the ANN model will be used to forecast the top butane composition. Therefore, this concludes that, before training the ANN, the preprocessing procedures are necessary and extremely essential .

Results

When the operating pressure is less than about 3 bar, increasing pressure can increase throughput; when the operating pressure is greater than about 7 bar, decreasing pressure can increase throughput; and when the operating pressure is in the middle, pressure changes have little effect on throughput. The Diversity in the methods finally concluded in one result which stated that the debutanizer's feed flowrate was the most effective input variable on the quality of butane, which was very obvious due to theory, since as the flow rate increases, the effective mass transfer rate increases while decreasing the contact time, which allows more products to be produced. It was also concluded that the overhead pressure is the least effective variable due to the fact that the operating pressure of the column appeared to be at a range that has a minimal effect on the quality of butane.

Numerous statistical analyses were studied and implemented in the debutanizer column in order to determine which of the data sets in the debutanizer column has the most effective on the quality of butane, and which has the least effect. The correlation method for data analysis was studied and implemented, in different software just because each software consists of different types, and we needed every piece of information we need to fit the puzzle and conclude with the most satisfactory results.

Conclusion

In the conclusion , machine learning and dynamic soft sensors require a huge amount of data analysis to precisely understand which of the input variables have an effect on the required output data, to model the neural network. If data analysis was not done, our model would not be designed with full accuracy. Our project required us do research different types of data analysis such as the Akaike Information Criterion, Principal Component Analysis, and Pearson Correlation. This was done to ensure that take the best decision into choosing the most important input data sets and eliminating any data set that has no effect on the output.

Student Name:
1-Ayat Salah Ahmed
2-Eman Salah Albastawy
3-Noor Khaled Isa AlSaqer

Student ID:
1-20170873
2-2017577
3-20175112

Adviser Name:
1-Dr.Bassam Al-Hamad
2-Dr.Raed Al-Jowder