*All pages strictly confidential*

# Prediction of Variance in the Product Quantity and Quality of HDPE through PLS Modeling - An Introductory Study

Final Report
(24th of July 2003)



by
Henky Muljana
(s1311816)
Department of Chemical Engineering
Rijksuniversiteit Groningen



RuG
July - 2003

H.Muljana

# Prediction of Variance in the Product Quantity and Quality of HDPE through PLS Modeling - An Introductory Study

Rijksuniversiteit Groningen
Department of Chemical Engineering

Groningen, July 24, 2003

**Supervisors :**
Prof.dr.ir. B. Roffel - Rijksuniversiteit Groningen
Dr.ir. Ton Kroezen - Sabic Euro Petrochemicals

# Summary

In this thesis, a study was undertaken to model the product quantity and quality in a HDPE production process, Partial Least Squares and Recursive Least Squares were investigated. The modeling steps are data collection, data conditioning, performing a Principal Component Analysis for variable selection, performing the PLS/RLS modeling and the model validation.

Data was collected for five different production types, although modelling was primarily done for one production type with model validation for the other production types.

There are 75 variables including two output variables (production level and melt index) that were collected from the plant. There are 19 variables that have to be removed from the input matrix because they show a zero and a constant values. Some of data the variables also need to be changed because these data can be marked as potentially bad data. Some correlated variables were found after performing a Principal Component Analysis. As the result of this analysis, the new input matrix with 44 variables was ready to be used for modelling purposes.

For the production quantity, a good model to explain the relationship between the conversion (1-R) and the input variables was obtained. This model can be used to predict the conversion for another condition. The model can predict actual values very well by capturing 85% of the variance and producing not more than 5% of the error. It was found that only seven variables of the data set of 44 variables were needed for a good prediction of the conversion.

For prediction of the melt index, a relatively good model with 15 variables was obtained. This model can capture 76% of the variance of the actual value. However, the model can not be used to predict the melt index of another production type. Application of the PLS model show very poor prediction, the predicted value deviates far from the real value.

An adaptive model based on the recursive least squares method was developed and resulted in a very good prediction. However, a recursive least squares method can not be used to predict far away in the future since the model always adapting all the time. Further investigation is therefore required to identify how the melt index can best be modeled, such as the use of hierarchical models.

# Table of Contents

# Acknowledgements

Finally, after spending almost two years, my study period at Rijksuniversiteit Groningen will come to the end. This thesis is the final work before I can finish and get the Master of Engineering degree.

I wish to express my personal gratitude to Professor Roffel for his kindly supervision, ideas, valuable advise and suggestion during the past eight months of my research period. I learn a lot from him. I also want to thank Dr. Ton Kroezen for his kindly assistances during my staying in Geleen, thanks for all the help and for valuable discussion we had during this period.

Thanks due to many people for their support and courage that has been given to me for all this time. I would like to thank all of my friends here in Netherlands : Ari and Selvy, Asaf, Andreas, Buana, Fesia, Farhad, Henk and Poppy, Ita, Iman, Iwan, and Ronny, thank you for everything. Especially for Ari and Selvy, I owe you two very much.

For the poeder hall boys : Edwin Zondervan, George van Dam, Roelof Bergsma and for my study companions : Holly, Paola, Lara. I am very happy that I have the opportunity to meet you and study together with all of you. I appreciate this time very much. Thank you.

To the one I love, Hedy, thank you for the understanding and also for the support you give to me. I am sorry that I cannot be there with you in the hard period of your life.

Last but not least, to my parents and families goes my eternal gratitude for their love, support and dedication.

H.Muljana
Rijksuniversiteit Groningen,
24th of July 2003

# Chapter 1
# Introduction

## 1.4 General

In polyetylene production, a reaction in a long tubular reactor at high pressure or in slurry reactor, which work at a medium pressure can be conducted.

The product quantity is controlled by controlling the ethylene mass flow going into the reactor. However, from field experiences, the ratio between these two measurements (the product quantity and the ethylene mass flow) seems to be influenced by others process measurements.

The product quality is determined by several measurements such as the melt index, the flow index and the density. The product quality is not constant, it can fluctuate considerably between different campaign. A model that can be used to predict the product quality, in this case the melt index is desireable.

Through PLS (Partial Least Squares) modeling a model can be developed to predict the output variable, by relating the output variable with all available measurements. In this project, two models are being developed, first is the product quantity model, the other is the product quality model. All important measurements have been collected for modeling purposes. Prior to modeling, the process data needs to be conditioned and it has to be determined which process variables do affect the product quantity and quality.

After a suitable model has been developed, it should be concluded whether a sufficient part of the variance in the product quantity and the quality can be explained from the process measurements. If so, the process can be controlled, if not, additional measurements may be required.

All data required for the research project are taken from the HDPE (High Density Polyethylene) production plant at SABIC Euro Petrochemicals.

## 1.5 Research Goals

The research goals are development of model, which can predict the output variable and to determine which part of the variance in the output variable can be explained from measured process data and which part of variance remains unknown.

## 1.6 Outline of the Report

The report is divided into ten chapters as describe in the following paragraph :

Chapter 2 represents a general explanation about the polyethylene processes including the High Density Polyethylene (HDPE) process technology.

Chapter 3 shows a general information about process identification. The identification steps used in the research project is described in this chapter.

Chapter 4 until chapter 7 are the main part of this report. Chapter 4 shows the information about the data collection from the plant.

Chapter 5 represents the general information about data conditioning. In this chapter, several important thing in data conditioning such detecting and removing bad data are included. The data conditioning results are described in chapter 5.2.

Chapter 6 introduces the Principal Component Analysis as the method for reducing the dimensionality and eliminating the correlated variables. The result of the PCA is described in chapter 6.3.

Chapter 7 deals with description and mathematical background of the Partial Least Squares modeling. The evaluation of the model is explained in chapter 8 and chapter 9.

Chapter 10 shows the conclusions and suggestions of this research project.