

UNIVERSITI TEKNOLOGI MARA

**CHEMOMETRICS METHODS TO
DETERMINE PALM OILS IODINE
VALUE (IV) USING FT-NIR AND
FTIR SPECTRA**

NOR FAZILA BINTI RASARUDDIN

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science

Faculty of Applied Sciences

August 2018

ABSTRACT

The authentication of the palm oils quality is indirectly crucial to consumer health. There are issues in palm oils quality within consumer macrocosm such as recycled and adulteration of cooking oils. Recently in Malaysia, there are claims of used cooking oils reprocessed and sold as new with lower price to consumers. This creates a threat to consumer health because they might use the reprocessed cooking oils unknowingly. In this study, the palm oils quality was determined through the iodine value (IV) which measured the total unsaturation of the oils. The characteristic of double bond in the oils was examined before and after oxidation process through the combination of spectroscopic method of Fourier Transform Near Infrared (FT-NIR) and Fourier Transform Infrared (FTIR) analysis and chemometrics methods. The IV was chemically determined through MPOB P3.2: 2004 method which later incorporated in the chemometrics models. The chemometrics multivariate approach of Partial Least Squares (PLS) regression model was applied to predict the IV of palm oils. The models were developed with the application of the data preprocessing, variable selection, data splitting and multiblock analysis. The performance of each method was assessed throughout the model development in the IV determination process. From the result, different data preprocessing provided different quality of prediction model. Either the application of the row scaling, column scaling individually or a combination of both might have improved the quality of the model. It is concluded that data preprocessing is dependent on the nature of the dataset and there can be no single method for general use. In this study, for the FT-NIR spectra data of non-fried palm oils, mean centre is the best data preprocessing with R^2_{cv} of 0.9472 and $RMSECV$ of 0.3496. Meanwhile for the FT-NIR spectra data of fried palm oils, standardisation is the best data preprocessing with R^2_{cv} of 0.9105 and $RMSECV$ of 0.4435. For the FTIR spectra data of non-fried palm oils, mean centre is the best data preprocessing with R^2_{cv} of 0.9614 and $RMSECV$ of 0.2966. Meanwhile for the FTIR spectra data of fried palm oils, mean centre and application of row scaling produced the best PLS model with R^2_{cv} of 0.9551 and $RMSECV$ of 0.3181. In comparison of variable selection methods, correlation coefficient (COR) variable selection method is the best method compared to PLSRC and PLSW in all PLS model. Variable selection is proven improved PLS model compared to the full spectrum model. On other hand, the best data splitting method for non-fried model is Kennard-Stone (KS) method while Duplex is the best method for fried palm oils PLS model. In the FT-NIR spectra data of non-fried palm oils, KS method give the best PLS model with R^2_{test} of 0.9355 and $RMSEP$ of 0.6813. Meanwhile, Duplex method resulted the best PLS model in the FT-NIR spectra data of fried palm oils with R^2_{test} of 0.9701 and $RMSEP$ of 0.2153. In the FTIR spectra data of non-fried palm oils, KS method give the best PLS model with R^2_{test} of 0.9693 and $RMSEP$ of 0.2871. Meanwhile, Duplex method resulted the best PLS model in the FTIR spectra data of fried palm oils with R^2_{test} of 0.9734 and $RMSEP$ of 0.3192. In multiblock analysis, similarity analysis through Mantel test, Procrustes Error Analysis, R_v and R_v modified showed both spectra are very similar. Multiblock analysis to predict the IV showed PLS model with single block was found to be advantageous compared to the multiblock PLS (MBPLS). FTIR block showed the best performance, followed by employing the FT-NIR block and employing both blocks in MBPLS.

ACKNOWLEDGEMENT

Firstly, I wish to thank Allah S.W.T. for giving me the opportunity to embark on my Master degree and for completing this long and challenging journey successfully. My gratitude and thanks go to my supervisor, Dr. Mohd Zuli Jaafar and co-supervisor, Prof Mohamed Noor Hasan. Thank you for the support, patience and ideas in assisting me with this project. I also would like to express my gratitude to the staff of the UiTM Cawangan Negeri Sembilan (Beting Campus), especially Mr. Noor Azman Tharim, Mr. Zubir bin Othman, Mr. Mohd Nazim bin Luwi and Mrs Sabariyah Zainal for providing the facilities, knowledge and assistance.

My appreciation goes to the Head of Nanoelectronics and Photonics Laboratory (NAPL) Department, Malaysian Institute of Microelectronic Systems (MIMOS), especially Puan Zalhan Md. Yusof and the crew members who provided the facilities and assistance during FT-NIR instrument analysis. Special thanks to my colleagues and friends for helping me with this project.

Finally, this thesis is dedicated to the loving memory of my very dear father and mother for the vision and determination to educate me. This piece of victory is dedicated to both of you. Alhamdulillah.

TABLE OF CONTENT

	Page
CONFIRMATION BY PANEL OF EXAMINERS	ii
AUTHOR'S DECLARATION	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
TABLE OF CONTENT	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF SYMBOLS	xii
LIST OF ABBREVIATIONS	xiii
LIST OF NOMENCLATURES	xv
CHAPTER ONE: INTRODUCTION	1
1.1 Research Background	1
1.2 Problem Statement	6
1.3 Objectives	7
a) To preprocess FTIR and NIR spectra for PLS model to predict IV of palm oils.	8
b) To select the wavelength region for predicting the quality parameter using correlation coefficient, PLSRC and PLSW as variable selection tools.	8
c) To study the effect of the data splitting in PLS model performance by applying Duplex, SPXY and Kennard-Stone.	8
d) To compare similarity between FTIR and FT-NIR spectra and develop BPLS model to predict the IV of edible oils.	8
1.4 Significance of Study	8
1.4.1 Data Preprocessing	8
1.4.2 Variable Selection	8
1.4.3 Data Splitting Methods	9
1.4.4 Multiblock Partial Least Squares Analysis	9
1.5 Scope and Limitation of Study	9

CHAPTER TWO: LITERATURE REVIEW	10
2.1 Frying Palm Oil	10
2.2 Determination of Frying Oils Physical Property	13
2.3 Instrumentation Techniques in Oils Quality	14
2.4 Chemometrics Methods	17
2.4.1 Partial Least Squares Regression	17
2.4.2 Data Preprocessing	20
2.4.3 Data Splitting Method	23
2.4.4 Variable Selection	25
2.4.5 Model Validation	26
2.4.6 Block Similarity Measure	28
2.4.7 Multiblock PLS Regression (MBPLS) Method in Chemometrics	29
CHAPTER THREE: RESEARCH METHODOLOGY	31
3.1 Introduction	31
3.2 Samples	31
3.3 Programs	31
3.4 FT-NIR and FTIR Analysis	32
3.5 Chemometrics Methods	33
3.5.1 Dataset	33
3.5.2 Regression Method	34
3.5.3 Data Preprocessing	35
3.5.4 Variable Selection Methods	36
3.5.5 Data Splitting Methods	37
3.5.6 Model Validation	39
3.5.7 Block Similarity Measure	41
3.5.8 Multiblock PLS Regression (MBPLS)	41
CHAPTER FOUR: RESULTS AND DISCUSSION	43
4.1 Determination of Iodine Value (IV) of Palm Oils	43
4.2 FT-NIR and FTIR Spectral Analysis	44
4.3 Chemometrics Methods	48
4.3.1 Data Preprocessing	48
4.3.2 Variable Selection	54