INFERENTIAL ESTIMATION AND CONTROL OF CHEMICAL PROCESSES USING PARTIAL LEAST SQUARES BASED MODEL

LIM WAN PIANG

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Engineering (Chemical)

Faculty of Chemical and Natural Resources Engineering
Universiti Teknologi Malaysia

JUNE 2005

To my beloved MOTHER and FATHER

ACKNOWLEDGEMENT

First of all, I would like to express my sincere appreciation to my research project supervisor, Professor Dr. Arshad Ahmad for his invaluable guidance, advice, support as well as encouragement throughout this research.

I am also indebted to the Ministry of Science, Technology and Environment, Malaysia for funding my research through National Science Foundation scholarship.

My heartiest appreciations are also for my fellow friends who have directly and indirectly contribute to the success of this project.

ABSTRACT

The use of inferential estimation model as a strategy to overcome the lack of efficient on-line measurement for product qualities is proposed. This strategy makes use of easy to measure secondary variables, such as temperature and pressure to infer the value of non-measurable primary variables such as chemical composition. As a case study, a fatty acid fractionation column from a local company was considered. The plant that was simulated using HYSYS™ simulator provided all the required process data throughout the study. To provide the necessary process insights, analyses of dynamic behaviour were carried out. Appropriate secondary measurements with significant relationships with the product composition were then identified for the construction of the inferential estimator within MATLAB® environment. A number of models were considered but nested neural network partial least squares (NNPLS) model was found most proficient. The model was tested on-line and reasonable performances were obtained. Further refinements were proposed to improve the accuracy and robustness of the estimator. In particular, the issue of data scaling was elaborately addressed. Following the success implementation of the estimator, inferential control of the product quality was examined. In both regulatory and servo controls, better performances were obtained compared to the indirect strategy of controlling product composition using selected tray temperature. This was further improved by employing cascade control. The results obtained throughout this work have illustrated the potential of inferential control strategy and the capability of the hybrid neural network-PLS model as the process estimator. This should therefore serve as an alternative solution to the lack of measurement in chemical process industry. The model developed from the simulation stage is specified to a particular case and it should be verified against the actual process before practical implementation.

ABSTRAK

Penggunaan model anggaran taabir sebagai satu strategi untuk menyelesaikan masalah pengukuran kualiti produk secara dalam talian yang berkesan telah dikemukakan. Strategi ini menggunakan pembolehubah sekunder yang mudah diukur, seperti suhu dan tekanan untuk meramal nilai pembolehubah utama yang tidak dapat diukur seperti komposisi produk. Sebagai kajian kes, sebuah turus pemecahan asid lelemak dari industri tempatan telah digunakan. Loji yang telah diselaku dengan menggunakan perisian HYSYS™ digunakan untuk membekal semua data proses yang diperlukan dalam kajian ini. Untuk memahami proses tersebut dengan lebih mendalam, analisis sambutan dinamik telah dilaksanakan. Pembolehubah sekunder yang berhubung rapat dengan komposisi produk telah dikenalpasti bagi tujuan pembangunan model anggaran taabir yang dilakukan dengan perisian MATLAB®. Penggunaan beberapa model telah dinilai tetapi model rangkaian saraf kuasa dua terkecil separa bergelung didapati paling berkesan. Model tersebut diuji secara dalam talian dan prestasi yang munasabah telah diperolehi. Beberapa pembaikkan telah dikemukakan untuk meningkatkan kejituan dan ketangkasan model anggaran. Secara khusus, isu penskalaan data telah dikaji dengan mendalam. Ekoran dari kejayaan perlaksanan model anggaran itu, kawalan taabir kandungan produk telah diuji. Dalam kedua-dua masalah gangguan dan servo, prestasi yang lebih memuaskan telah dicapai berbanding dengan strategi kawalan kandungan produk secara tidak langsung yang menggunakan suhu dulang. Prestasi tersebut seterusnya dipertingkatkan dengan menggunakan kawalan lata. Keputusan yang diperolehi dalam penyelidikan ini telah menunjukkan potensi strategi kawalan taabir dan keupayaan model hibrid rangkaian saraf kuasa dua terkecil separa sebagai penganggar proses. Kaedah ini seharusnya mampu menjadi salah satu daripada penyelesaian kepada kekurangan alat pengukuran dalam industri proses kimia. Model yang dibina daripada tahap perselakuan adalah terhad kepada kes yang tertentu dan ia harus dinilai dengan process sebenar sebelum diamalkan secara praktikal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

SYMBOLS

a	-	Last dimension in PLS
b	-	Regression coefficient in PLS
b_0	-	Bias weight of neuron model
$\hat{b}_{ak}$	-	Regression coefficient in inner PLS
c_i	-	Coefficients of polynomial function.
$\boldsymbol{d}$	-	Weight vector in QPLS
$\mathbf{e}$	-	Matrix of mismatch between the $\mathbf{u}$ and $\hat{\mathbf{u}}$
$\mathbf{E}$	-	Residue matrix of $\mathbf{X}$ blocks in PLS
$\mathbf{F}$	-	Residue matrix of $\mathbf{Y}$ blocks in PLS
$\mathbf{f}$	-	Residual matrix of inner PLS
$\mathbf{g}_{ak}$	-	Loading scores of e-block in inner PLS
$\mathbf{h}_{ak}$	-	Loading scores of Z-block in inner PLS
k	-	Last dimension in the inner PLS
K_c	-	Controller gain
K_p	-	Steady state gain
N	-	Non-linear function in NNPLS
$\mathbf{p}$	-	Input loading factors matrix in PLS
$\mathbf{q}$	-	Output loading factors matrix in PLS
$\mathbf{r}_{ak}$	-	Latent scores of e-block in inner PLS
$\mathbf{s}$	-	Column matrix in QPLS
$\mathbf{s}_{ak}$	-	Latent scores of Z-block in inner PLS
s_x	-	Standard deviation
$\mathbf{t}$	-	Input latent scores matrix in PLS

t_s	-	Settling time
$\mathbf{u}$	-	Output latent scores matrix in PLS
$\hat{\mathbf{u}}$	-	Predicted output latent scores matrix in PLS
$\mathbf{v}$	-	Column vector in QPLS
$\mathbf{v}_{ak}$	-	Weight in the inner PLS
$\mathbf{w}$	-	Input weights matrix in PLS
w_i	-	Neuron weight
$\mathbf{X}$	-	Independent variables matrix in PLS
$\boldsymbol{x}$	-	Input variables
$\bar{x}$	-	Average value
x_i	-	Neuron input
x_{ms}	-	Mean-scaled data
$x_D^{C_{12}}$	-	Composition of C_{12} fatty acid in the distillate product
$\mathbf{Y}$	-	Dependent variables matrix in PLS
y	-	Neuron output
$\hat{y}_t^c$	-	Filtered output
$\hat{y}_t$	-	Predicted value at current time
$\hat{y}_{t_f}^c$	-	Previous corrected predicted value at time t_f
$\mathbf{Z}$	-	Weight matrix in QPLS

GREEK SYMBOLS

a	-	Prediction coefficient in PLS
e	-	Residual matrix of inner relation in quadratic PLS
ω	-	Weight of neural network model
ß	-	Bias of neural network model
? **w**	-	Increment weight in QPLS
s	-	Activation function in neural network
t_D	-	Derivative time constant
t_I	-	Integral time constant

ABBREVIATIONS

ANN	-	Artificial neural networks
ARMA	-	Autoregressive moving average
CC	-	Composition controller
CSTR	-	Continuous stir tank reactor
DDE	-	Dynamic data exchange
DNNPLS	-	Dynamic neural network partial least squares
EBWU	-	Error-based weight updating procedure
EKF	-	Extended Kalman filter
EOS	-	Equation of state
EPV	-	Explained prediction variance
FC	-	Flow controller
FFN	-	Feed forward networks
FPM	-	First principle method
LC	-	Level controller
LCC	-	Light-cut column
LM	-	Levenberg-Marquardt method
MIMO	-	Multiple input multiple output
MISO	-	Multiple input single output
MLR	-	Multiple linear regressions
MSE	-	Mean squared error of prediction
NIPALS	-	Non-linear iterative partial least squares
NNPLS	-	Neural network partial least squares
PCA	-	Principal component analysis
PCR	-	Principal component regression
PIC	-	Pressure indicator
PI	-	Proportional-Integral
PID	-	Proportional-Integral-Derivative
PLS	-	Partial least squares regression
QPLS	-	Quadratic partial least squares
RBFPLS	-	Radial basis function partial least squares
SISO	-	Single input single output
SSE	-	Sum Squared Error

TC	-	Temperature controller
UNIQUAC	-	Universal Quasi Chemical

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Motivation of Study

Stringent product specifications, stiff competition among manufacturers and increasingly strict regulation from local authority in the face of full capacity operation with zero accidents and emissions have forced many existing plants to revamp their existing control system. More advanced control schemes have been introduced, and although small in numbers, real-time optimisations have also been implemented.

Despite these successful implementations, many issues remained as hindrances to efficient process control. For example, the success in the implementation of any optimisation scheme requires adequate performance of all control loops. This is however, sometimes hampered by two issues. The first is related to inadequacy of conventional controllers used since chemical process dynamics are typically non-linear whilst the controllers are based on linear theory. The second issue is associated with process measurements, the accuracy of which is a prerequisite to successful process control.

Since measurement devices are one of the main factors in achieving effective process control, selection of appropriate sensors and their location should be properly considered. However, not all variables in a process plant are readily to be measured on-line. Product quality variables such as chemical composition and molecular weight distribution of polymer are rarely available on-line, and are usually obtained by laboratory sample analyses. This is usually performed at long intervals and is therefore not practical to be used for process control.

Over the years, various on-line measurement devices have been developed. However, many of these on-line devices are still suffering from problems due to the availability, reliability, complexity and large delays. For example, on-line gas chromatograph is a common instrumentation for the on-line measurements of product compositions. However, in many applications this measurement device is not reliable enough to be used for on-line control due to low sampling rate and sometimes inconsistency of results. High operating and maintenance costs add to the disadvantages of such implementations. For some quality variables, existing analytical tools used are simply unavailable for on-line applications. Hence, the development of inferential estimation and control has been advocated as one of the alternative solution to deal with measurement difficulties.

1.2 Problem Statement

A fatty acids fractionation column from a local industry is faced with the product compositions control problems. Currently, indirect control of product compositions is achieved by controlling temperature at selected location in the column. At times, this control scheme cannot cope with disturbances and process uncertainties in the plant. This scenario has created some difficulties in the composition control and occasionally, off-specification products have been produced.

This work is proposed to untangle some of these difficulties. An inferential model, which is built based on partial least squares (PLS) regression is employed for estimating the product composition in the light-cut column to facilitate process control.

1.3 Objective and Scope of Work

The aim of this work is to develop an inferential model using PLS modelling approach and to investigate its application in composition control. The scopes of the study are as follows.

i. Dynamic simulation and analysis of a fatty acid fractionation column using HYSYS™ process simulator.
ii. Development of a base-case inferential model using PLS modelling approach in MATLAB® environment.
iii. Improving the inferential model using a modified PLS model, namely nested PLS model.
iv. Development of inferential control strategy to regulate the product composition using the established estimation model by linking both software packages using DDE interface.

1.4 Contribution of the Work

Successful development of the inferential estimator using PLS modelling approach is the main contribution of this work. Dealing with the issue of on-line implementation of the inferential model, on-line post-processing or rescaling of the predicted values was solved using polynomial regression method. After several refinements, the inferential estimator was therefore able to produce predictions with reasonable accuracy under various operating conditions. The work has also demonstrated the use of neural networks in a hybrid PLS modelling structure as a means of extending the PLS model capability to nonlinear process estimation.

Another contribution of this work is on-line monitoring and controlling of the product composition of a fatty acid fractionation in the simulation platform using the inferential model. Two inferential configurations were investigated in this work and the inferential cascade control showed better performance in both regulator and servo control.

1.5 Organisation of the Thesis

The thesis is organised as follows. Chapter 2 includes literature review and some theoretical background about inferential estimation and control. Previous development and application of PLS-based models are reviewed. The theory and implementation requirements for the proposed method are also discussed. Chapter 3 commences with the description of the selected case study. This is followed by relevant analysis for inferential model development. Chapter 4 then elaborates the development of process estimation model and some proposed improvements to the model. Chapter 5 demonstrates the application of inferential model in process control. The thesis is then concluded with the overall findings and some recommendations for future work.

REFERENCES

Adebiyi, O.A. and Corripio, A.B. (2003). Dynamic Neural Networks Partial Least Squares (DNNPLS) Identification of Multivariable Processes. *Computers & Chemical Engineering.* 27(2): 143-155.

Ahmad, A., Wong, T.S. and Ling, L.Y. (2001). Dynamic Simulation for a Palm Oil Fractionation Process. *Proceeding of the Regional Symposium on Chemical Engineering*. October 29-31. Bandung, Indonesia. M012-M018.

Amirthalingam, R. and Lee, J.H. (1999). Subspace Identification Based Inferential Control Applied to a Continuous Pulp Digester. *Journal of Process Control*. 9(5): 397-406.

Amirthalingam, R., Sung, S.W. and Lee, J.H. (2000). Two-Step Procedure for Data-Based Modelling for Inferential Control Applications. *AIChE*. 46(10): 1974-1988.

Baffi, G., Martin, E.B. and Morris, A.J. (1999a). Non-linear Projection to Latent Structures Revisited (the Quadratic PLS Algorithm). *Computers & Chemical Engineering.* 23(3): 395-411.

Baffi, G., Martin, E.B. and Morris, A.J. (1999b). Non-linear Projection to Latent Structures Revisited (the Network PLS Algorithm). *Computers & Chemical Engineering.* 23(9): 1293-1307.

Barry, L., Montague, G.A., Frith, A.M., Gent, C. and Bevan, V. (2001). Industrial Application of Neural Networks – An Investigation. *Journal of Process Control.* 11(5): 497-507.

Bhat, N. and McAvoy, T.J. (1990). Use of Neural Nets for Dynamic Modelling and Control of Chemical Process Systems. *Computers & Chemical Engineering.* 14(4/5): 573-583.

Budman, H.M., Webb, C., Holcomb, T.R. and Morari, M. (1992). Robust Inferential Control for a Packed-Bed Reactor. *Industrial Engineering Chemical Research.* 31: 1665-1679.

De Jong, S. (1993). SIMPLS: An alternative pproach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems.* 18: 251-263.

Demuth, H. and Beale, M. (2001). *Neural Network Toolbox User's Guide*. 4th Version. The MathWorks, Inc.

Doyle III, F.J. (1998). Nonlinear Inferential Control for Process Applications. *Journal of Process Control.* 8(5-6): 339-353.

Geladi, P. and Kowalski, B.R. (1986). Partial Least-Squares Regression: A Tutorial. *Analytica Chimica Acta.* 185: 1-17.

Hoo, K.A., Sinzinger, E.D. and Piovoso, M.J. (2002). Improvements in the Predictive Capability of Neural Networks. *Journal of Process Control.* 12(1): 193-202.

James, S., Legge, R. and Budman, H. (2002). Comparative Study of Black-Box and Hybrid Estimation Methods in Fed-Batch Fermentation. *Journal of Process Control.* 12(1): 113-121.

Joseph, B. (1999). A Tutorial on Inferential Control and its Applications. *Proceedings of the American Control Conference.* June 1999. San Diego, California, 3106-3118.

Joseph, B. and Brosilow, C.B. (1978). Inferential Control of Processes. *AIChE Journal.* 24(3): 485-508.

Kano, M., Miyazaki, K., Hasebe, S. and Hashimoto, I. (2000). Inferential Control System of Distillation Compositions Using Dynamic Partial Least Squares Regression. *Journal of Process Control.* 10(2): 157-166.

Kano, M., Showchaiya, N., K., Hasebe, S. and Hashimoto, I. (2003). Inferential Control of Distillation Compositions: Selection of Model and Control Configuration. *Control Engineering Practice.* 11(8): 927-933.

Kresta, J.V., Marlin, T.E. and MacGregor, J.F. (1994). Development of Inferential Process Models Using PLS. *Computers & Chemical Engineering.* 18(7): 597-611.

Lakshminarayanan, S., Shah, S.L. and Nandakumar, K. (1997). Modelling and Control of Multivariable Processes: Dynamic PLS Approach. *AIChE Journal*. 43(9): 2307-2322.

Li, B., Hassel, P.A., Morris, A.J. and Martin, E.B. (2005). A Non-linear Nested Partial Least-Squares Algorithm. *Computational Statistical & Data Analysis*. 48: 87-101.

Ljung, L. (1999). *System Identification: Theory for the User*. 2nd Edition. New Jersey: Prentice Hall.

Mejdell, T. and Skogested, S. (1991a). Estimation of Distillation Compositions from Multiple Temperature Measurements Using Partial-Least-Squares Regression. *Industrial Engineering Chemical Research*. 30: 2543-2555.

Mejdell, T. and Skogested, S. (1991b). Composition Estimator in a Pilot-Plant Distillation Column Using Multiple Temperatures. *Industrial Engineering Chemical Research*.30: 2555-2564.

Moore, C.F. (1992). Selection of Controlled and Manipulated Variables. In: Luyben, W.L. *Practical Distillation Column*. New York: Van Nostrand Reinhold. 140-177.

Page, G.F., Gomm, J.B. and Williams, D. (1993). *Application of Neural Networks to Modelling and Control*. London: Chapman & Hall.

Pan, Y. and Lee, J.H. (2003). Recursive Data-Based Prediction and Control of Product Quality for A PMMA Batch Processes. *Chemical Engineering Science*. 58: 3215-3221.

Parrish, J.R. and Brosilow, C.B. (1988). Nonlinear Inferential Control. *AIChE Journal*. 34(4): 633-644.

Pollard, J.F., Broussard, M.R., Garrison, D.B. and San, K.Y. (1992). Process Identification Using Neural Networks. *Computers and Chemical Engineering*. 16(4): 253-270.

Psichogios, D.C. and Ungar, L.H. (1992). A Hybrid Neural Network-First Principles Approach to Process Modelling. *AIChE Journal*. 38(10): 1499-1511.

Qi, H., Zhou, X.G., Liu, L.H. and Yuan, W.K. (1999). A Hybrid Neural Network-First Principles Model for Fixed-Bed Reactor. *Chemical Engineering Science*. 54: 2521-2526.

Qin, S.J. and McAvoy, T.J. (1992). Nonlinear PLS modelling Using Neural Networks. *Computers & Chemical Engineering.* 16(4): 379-391.

Stephanopoulos, G. (1984). *Chemical Process Control: An Introduction to Theory and Practice.* Englewood Cliffs, N.J.: Prentice-Hall.

Tham, M.T., Montague, G.A., Morris, A.J. and Lant, P.A. (1991). Soft-sensors for process estimation and inferential control. *Journal of Process Control*. 1(1): 3-14.

Thompson, M.L. and Kramer, M.A. (1994). Modelling Chemical Processes Using Prior Knowledge and Neural Networks. *AIChE Journal.* 40(8): 1328-1340.

Willis, M.J., Montague, G.A., Massimo, C.D., Tham, M.T. and Morris, A.J. (1992). Artificial Neural Networks in Process Estimation and Control. *Automatica*. 28(6): 1181-1187.

Wold, H. (1985). Partial Least Squares. In: Kotz, S. and Johnson, N.L. (ed). *Encyclopedia of Statistical Sciences.* Vol 6. New York: Wiley. 581-591.

Wold, S., Kettaneh-Wold, N. and Skagerberg, B. (1989). Nonlinear PLS Modeling. *Chemometrics and Intelligent Laboratory Systems.* 7(1-2): 53-65.

Yeh, T.M., Huang, M.C. and Huang C.T. (2002). Estimate of Process Compositions and PLantwide Control from Multiple Secondary Measurements using Artificial Neural Networks. *Compouter & Chemical Engineering.* 27(1): 55-72.

Yu, C.C. and Luyben, W.L. (1984). Use of Mutiple Temperatures for the Control of Multicomponent Distillation Column. *Ind. Eng. Chem.Process Des Dev.* 23: 590-597.

Zhang, H. and Lennox, B. (2004). Integrated Condition Monitoring and Control of Fed-batch Fermentation Processes. *Journal of Process Control.* 14(1): 41-50.

Zorzetto, L.M.F., Macial Filho, R. and Wolf-Maciel, M.R. (2000). Process Modelling Development through Artificial Neural Networks and Hybrid Models. *Computers & Chemical Engineering.* 24: 1355-1360.