

Near infrared transmittance analysis for the assay of pharmaceutical tablet and Multivariate Analysis

Abstract

This paper is a case study illustrating how the active substance in pharmaceutical tablets can be quantified from NIR spectroscopy data using multivariate methods like Principal Component Analysis (PCA) and Partial Least Squares Regression (PLS), as these techniques can remove the high degree of redundancy seen in NIR data. A variety of spectroscopic data pre-treatment methods are applied to separate information from noise, like classical scatter correction techniques, as well as the more recently developed extension, called extended multivariate scatter correction. It will be shown that the choice of pre-treatment is crucial to success, especially since NIR data is often subject to light scatter.

Introduction

The Process Analytical Technology (PAT) initiative promotes the pharmaceutical industry to perform large-scale quality control analysis on a routine basis. Such analysis requires fast and reliable measurement techniques in combination with multivariate data analysis.

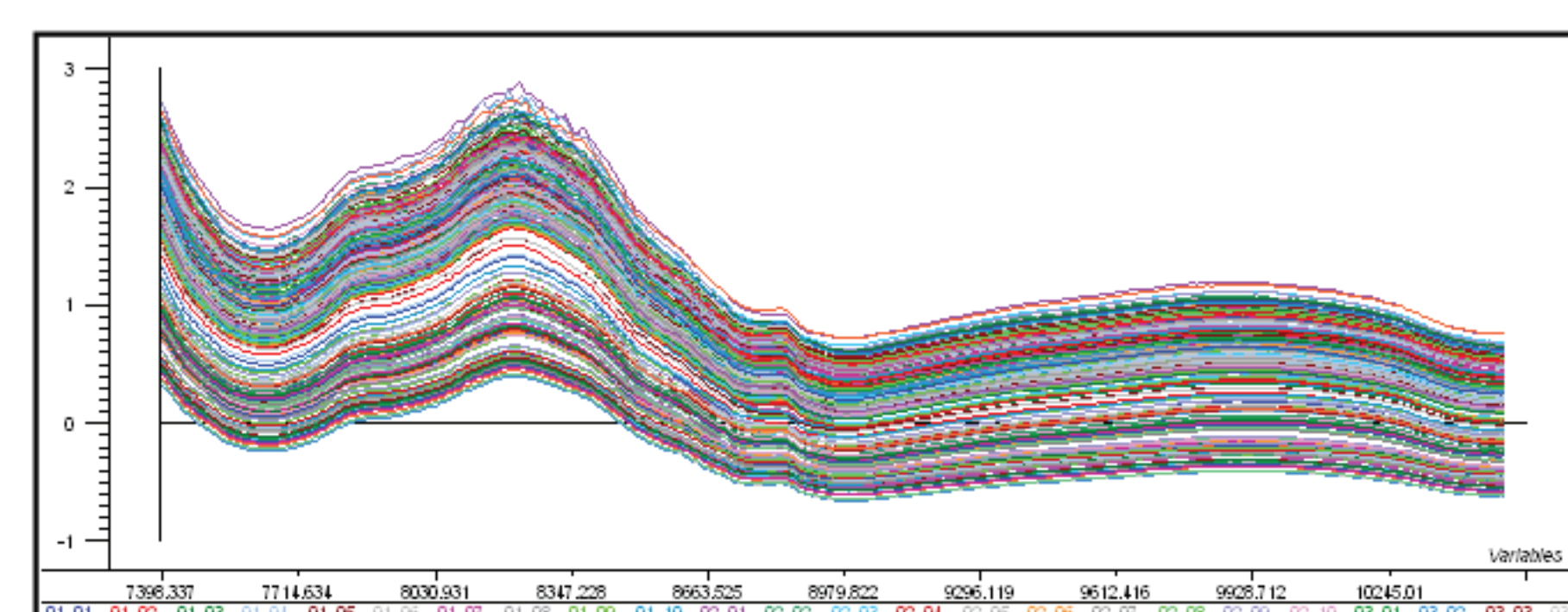
Near Infrared (NIR) spectroscopy has gained considerable attention in recent years due to speed and the fact that no special sample preparation is needed. In addition, NIR spectroscopy is a nondestructive measurement device, and has successfully been applied in many areas of production in the chemical industry. Since typical NIR spectra are not readily interpretable compared to other spectroscopic techniques, multivariate analysis of the spectroscopic data is required to build quantitative models.

Material and Methods

The samples used were Escitalopram® coated tablets of four different dosage values (5, 10, 15 and 20 mg in dose proportional tablets). Samples were collected at pilot, Full scale and Laboratory studies. 10-15 tablets from each dosage value were used for the study. NIR Transmittance spectra recorded in the range 700-2500 nm from each tablet. Traditional analysis of Assay content analysis has been performed by using HPLC.

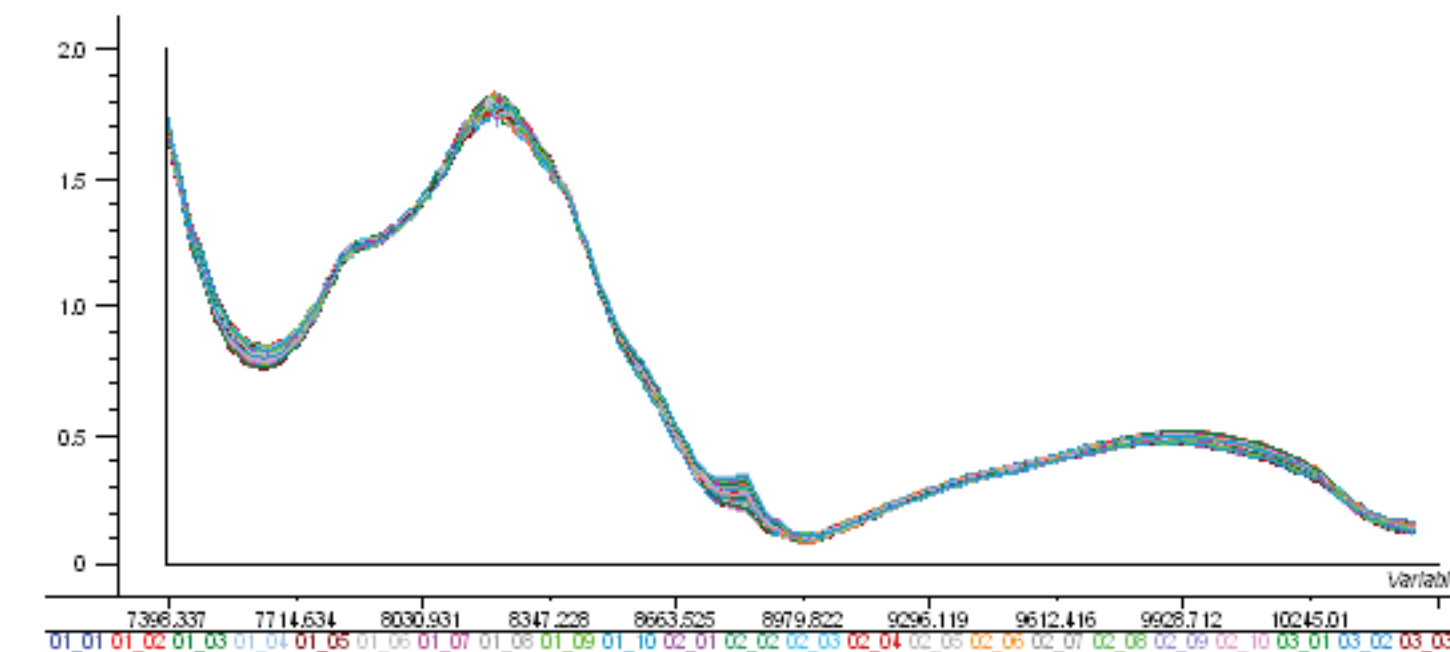
The Unscrambler® was used to develop regression model for the assay content of the Escitalopram tablets. The different preprocessing techniques were applied and comparative analysis done. Below (fig:1, fig:2, fig:3) are the screenshots of different pre-processing techniques applied over data.

Raw Spectrum



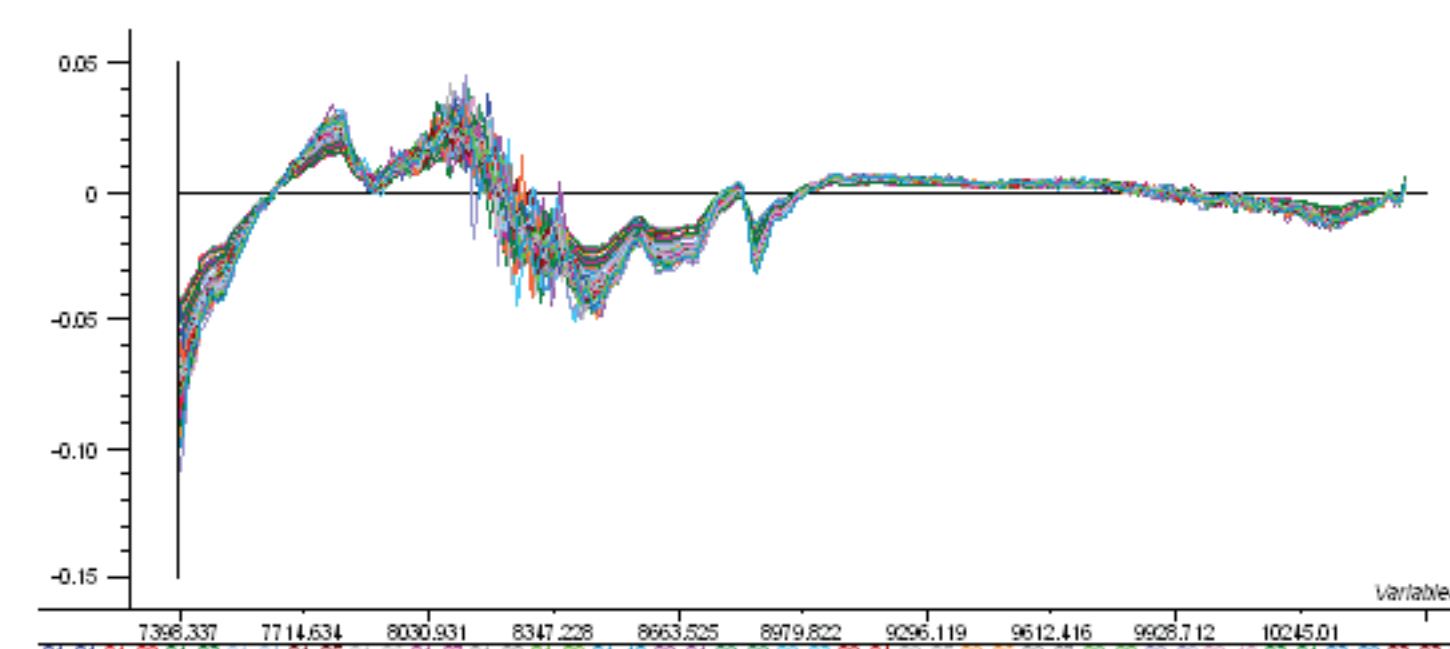
(fig:1)

Multivariate Scatter Correction (MSC)



(fig:2)

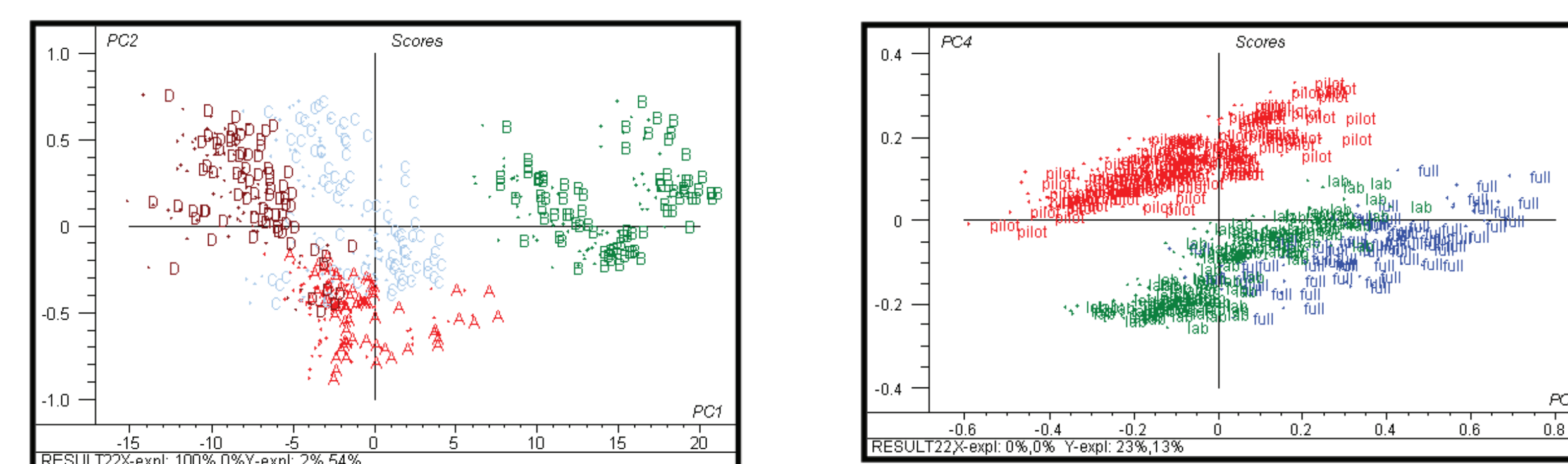
1st or 2nd derivative



(fig:3)

Results and Discussions

Initially Partial Least Square Regression (PLS) were applied to the complete range of spectrum i.e. 700- 2500 nm and also different data range were created from 1100 – 1250 nm and developed PLS methods. Below (fig:4) is the scores plot which clearly shows the variability in the dosage forms as well as different studies.

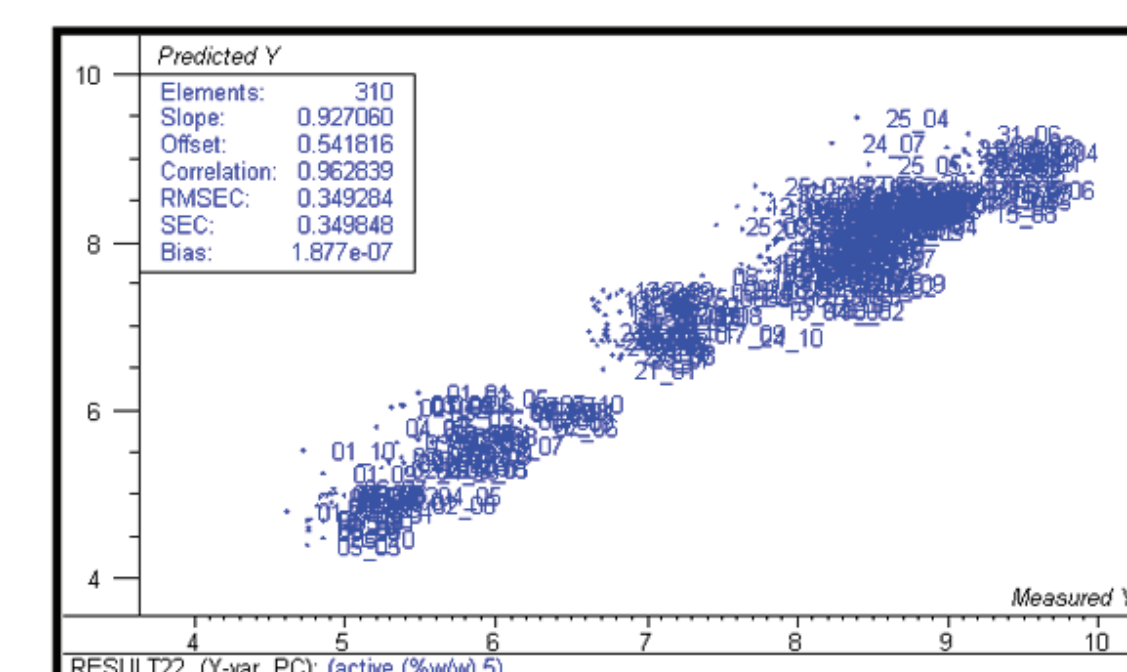


Scores for PC1 and PC2 from the PLS model describe variability among dosage forms.

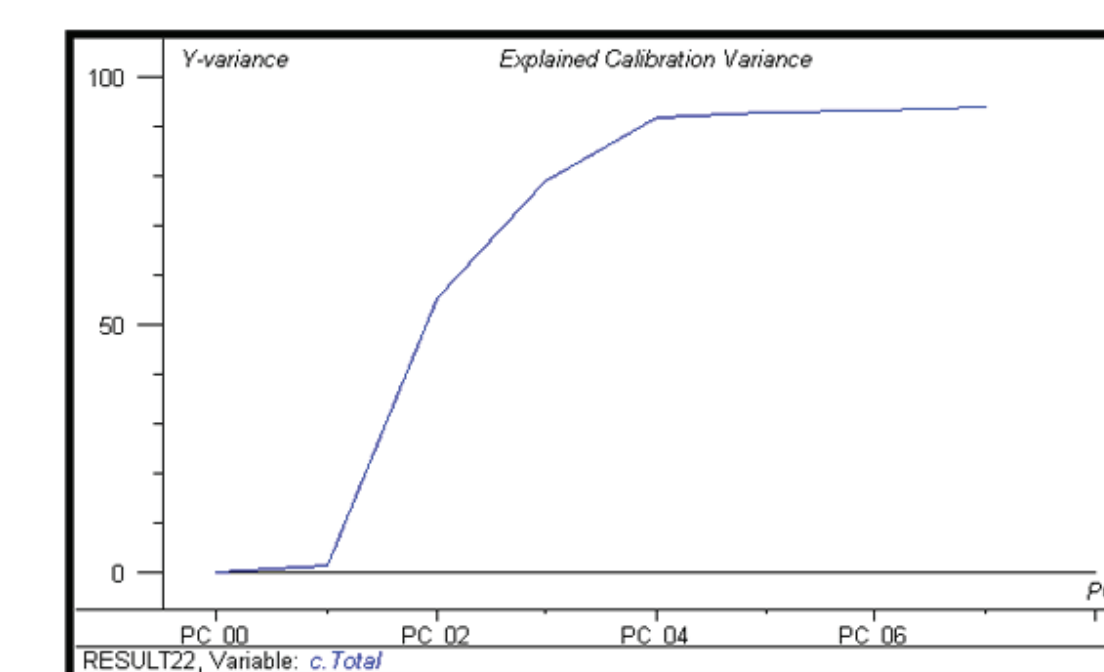
PC3 and PC4 from the model describes variation between lab, pilot and full scale studies.

(fig:4)

Below (fig:5, fig:6) are the experimental results of PLS model applied on whole range of spectrum without any pre processing. The Explained variance curve indicate the amount of variation described by the components in the model. This Model require 5 PCs describe most of the variation with RMSEC of 0.35.

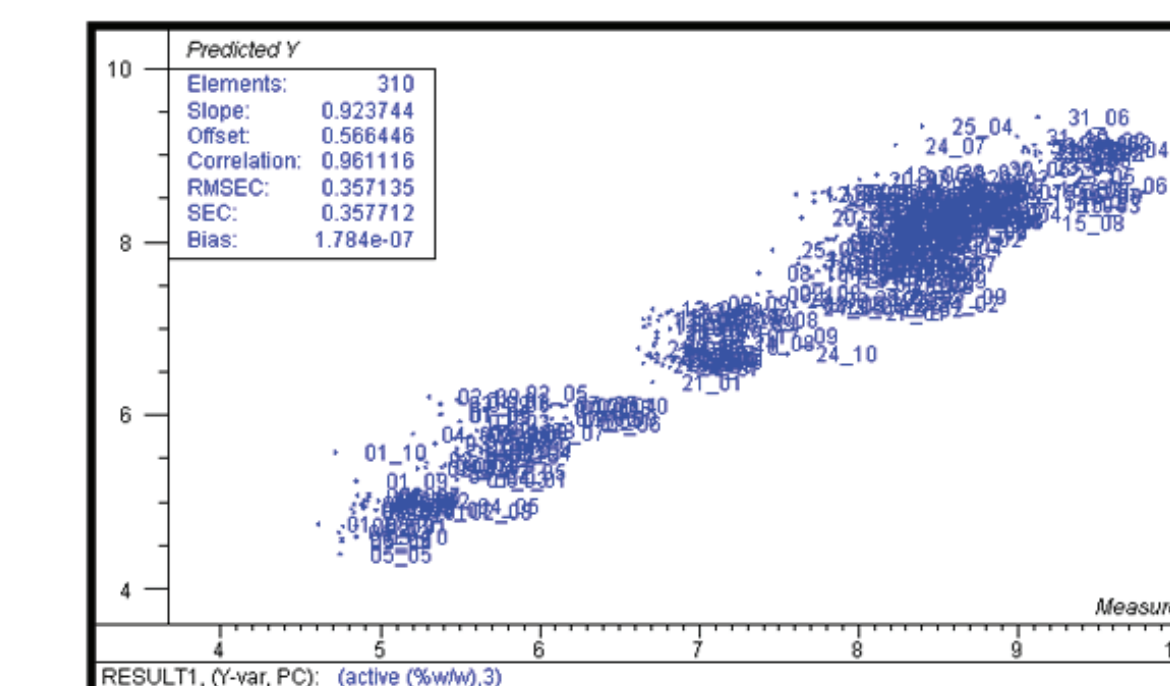


Explained variance curve indicate the amount of variation described by the components in the model. 5 PCs describe most of the variation.

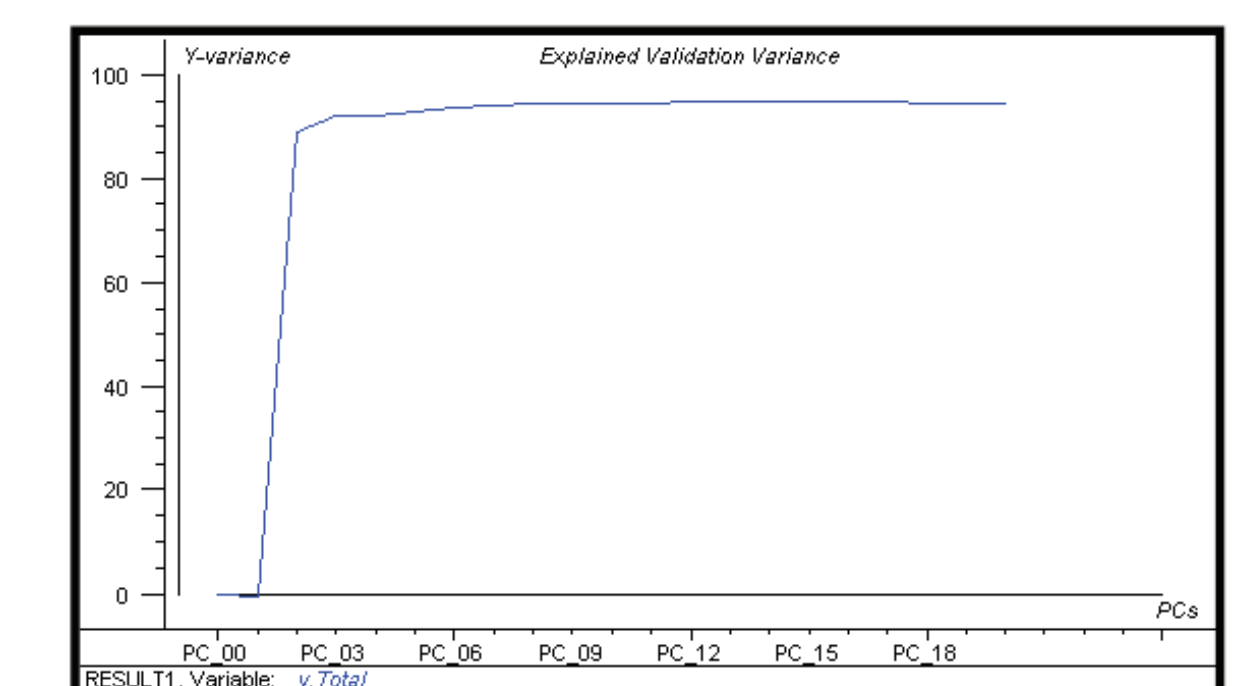


The predicted versus measured plot shows how well the model is described in terms of performance, RMSEC

(fig:5)



Explained variance curve indicates that a smaller number of components are needed for the PLS model using only parts of the data.



The same RMSEC is achieved with the reduced data set as for the full data set, with only 3 instead of 5 components

(fig:6)

Second approach of the data analysis is PLS Model is applied over the spectral region of only 1100-1250 nm with out preprocessing the data. This Model required 3 PC's to explain the most variation in the data with RMSEC of 0.35

Conclusion

Different approaches were considered in the process of method development.

- From the different calibration models the PLS Regression model with Multi Scatter Correction (MSC) for the spectral range 1100-1250 nm were prove to be good one with number of principle components 2 with RMSEC of 0.33. Below is the comparative study chart of different preprocessing techniques.

- From the scores plot, it was clear that the lab study was more similar to the full scale than the pilot study.

Preprocessing	Data set	# PCs	RMSEC
None	All	5	0.35
None	Region	3	0.35
MSC	All	3	0.34
MSC	Region	2	0.33
1st derivative	All	5	0.38
1st derivative	Region	1	0.38
2nd derivative	All	3	0.36
2nd derivative	Region	3	0.35

ACKNOWLEDGEMENTS:

Martin Kermit, CAMO Software AS – Nedre Vollgate, 8 – NO-0158 Oslo- Norway martin.kermit@camo.no
Suresh Kumar B.V., CAMO Software India Pvt. Ltd., Domlur Layout, Bangalore, India skumar@camo.co.in



CAMO Software India Pvt. Ltd.,
14 & 15, Krishna Reddy Colony,
Domlur Layout, Bangalore - 560 071
INDIA

Tel: +91 (80) 4125 4242
Fax: +91 (80) 4125 4181