


Spring 5-11-2018

Designing A Calibration Set in Spectral Space for Efficient Development of An NIR Method For Tablet Analysis

Md Anik Alam

Follow this and additional works at: <https://dsc.duq.edu/etd>

 Part of the [Analytical Chemistry Commons](#), [Design of Experiments and Sample Surveys Commons](#), [Multivariate Analysis Commons](#), and the [Pharmaceutics and Drug Design Commons](#)

Recommended Citation

Alam, M. (2018). Designing A Calibration Set in Spectral Space for Efficient Development of An NIR Method For Tablet Analysis (Doctoral dissertation, Duquesne University). Retrieved from <https://dsc.duq.edu/etd/1430>

This One-year Embargo is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection. For more information, please contact phillips@duq.edu.

DESIGNING A CALIBRATION SET IN SPECTRAL SPACE FOR EFFICIENT
DEVELOPMENT OF AN NIR METHOD FOR TABLET ANALYSIS

A Dissertation

Submitted to the Graduate School of Pharmaceutical Sciences

Duquesne University

In partial fulfillment of the requirements for
the degree of Doctor of Philosophy

By

Md Anik Alam

May 2018

Copyright by
Md Anik Alam

2018

DESIGNING A CALIBRATION SET IN SPECTRAL SPACE FOR EFFICIENT
DEVELOPMENT OF AN NIR METHOD FOR TABLET ANALYSIS

By

Md Anik Alam

Approved April 12, 2018

Carl A. Anderson, PhD
Associate Professor of Pharmaceutics
Duquesne University
(Committee Chair)

James K. Drennen, III, PhD
Associate Professor of Pharmaceutics
Duquesne University
(Committee Member)

Peter L.D. Wildfong, PhD
Associate Professor of Pharmaceutics
Duquesne University
(Committee Member)

Ira S. Buckner, PhD
Associate Professor of Pharmaceutics
Duquesne University
(Committee Member)

Dongshen Bu, PhD
Senior Research Investigator
Bristol-Myers Squibb
(Committee Member)

David Johnson, PhD
Professor of Pharmacology
Duquesne University

ABSTRACT

DESIGNING A CALIBRATION SET IN SPECTRAL SPACE FOR EFFICIENT DEVELOPMENT OF AN NIR METHOD FOR TABLET ANALYSIS

By

Md Anik Alam

May 2018

Dissertation supervised by Carl A. Anderson

Designing a calibration set is the first step in developing a spectroscopic calibration method for quantitative analysis of pharmaceutical tablets. This step is critical because successful model development depends on the suitability of the calibration data. For spectroscopic-based methods, traditional concentration based techniques for designing calibration sets are prone to have redundant information while simultaneously lacking necessary information for a successful calibration model. The traditional method also follows the same design approach for different spectroscopic techniques and different formulations, thereby lacks the optimizing capability to be technique and formulation specific.

A method for designing a calibration set in the Near Infrared (NIR) spectral space was developed for quantitative analysis of tablets. The pure component NIR spectra of a tablet

formulation were used to define the spectral space of that formulation. This method minimizes sample requirements to provide an efficient means for developing multivariate spectroscopic calibration.

Multiple comparative studies were conducted between commonly employed experimental design approaches to calibration development and the newly developed spectral space based technique. The comparisons were conducted on single API (Active Pharmaceutical Ingredient) and multiple API formulation to quantify model drugs using NIR spectroscopy. Partial least squares (PLS) models were developed from respective calibration designs. Model performance was comprehensively assessed based on the ability to predict API concentrations in independent prediction sets. Similar prediction performance was achieved using the smaller calibration set designed in spectral space, compared to the traditionally designed large calibration sets. An improved prediction performance was observed for the spectrally designed calibration sets compared to the traditionally designed calibration sets of equal sizes. Spectral space was also used to incorporate physico-chemical information into the calibration design to provide an efficient means of developing robust calibration model. Robust calibration model is critical to ensure consistent model performance during model lifecycle. A weight coefficient based technique was developed for selecting loading vector in PLS model to aid in building robust calibration model.

It was also demonstrated that the optimal structures of calibration sets are different between NIR and Raman spectroscopy for the same tablet formulation. The optimum calibration structures are also different between two APIs for the same spectroscopic technique, indicating the criticality of the calibration design to be formulation and

technique specific. This study demonstrates that a calibration set designed in spectral space provides an efficient means of developing spectroscopic multivariate calibration for tablet analysis. This study also provides opportunity to design formulation and technique specific calibration sets to optimize calibration capability.

DEDICATION

This dissertation is dedicated to my family.

ACKNOWLEDGEMENTS

Since the earliest US Ph.D. in 1861, Yale, we have been seeing that Ph.D. life is full of sweat and tears, and lots of coffee. What we cannot see are the love and care from the dearest ones that make the journey worthwhile. It's a beautiful opportunity to thank them at the very first place.

First, I would like to thank Dr. Carl Anderson for his trust on me. He believed in me when I was an unknown from eight thousand miles away. He believed in me when I was known for my weakness, ignorance and incompetency. His trust helped me to face my limitations and look for improvement instead of frustration. His research skills and critical thinking have always helped me to find a way in the toughest problems. I am fortunate to have such advisor in my graduate life.

The second person to be thankful to is Dr. James Drennen. Each person needs someone to admire and follow, and he was right there showing me the way to think, talk and act, not only about scientific challenges, but also about every other challenge we face in our lives. He was always there to help, support, care and guide. I am immensely thankful to him for helping me to evolve in my graduate life.

I am thankful to Dr. Ira Buckner for his academic guidance and intellectual support. He taught me to break down every problem to fundamental science. I am thankful to Dr. Peter Wildfong for his thoughtful remarks and intriguing critics. He taught me critical thinking and reasoning. I am thankful to Dr. Dongsheng Bu for his valuable suggestions and technical recommendations. His analytical insights were found very helpful for this dissertation project.

I would like to thank Nayeem for being a brother. My graduate life had changed since he joined in the lab. I am thankful to my longtime friend Shikhar for walking beside me in this long journey. Truly, he is the most helpful person I have ever seen. I am thankful to my first ever friends in graduate life Drs. Patel and Vasa for their unconditional supports in my toughest times. I am thankful to Tanvir, Junayed and Tasdique for making the graduate school a beautiful happy place. I am thankful to Ryanne for her care from the very first day. I am thankful to my past lab mates, Drs. Talwar, Zacour, Bondi, Igne for their enormous support and guidance in my academic life. I am thankful to my current lab mates, Hanzhou, Yi, Doug, Henry, Natasha, Su Yang for sharing all the stressed moments and making these as beautiful memories. I am thankful to Jackie, Deb and Mary for tolerating all my administrative tortures.

The person whom I owe the most in my life is my mom, Afroza Pervin. She raised me strong, taught me wisdom and inspired me to follow my dreams. My dad, Rafiqul Alam, was always there to support and care for me. They are the greatest blessings in my life. I am extremely thankful to my in-laws, Shafiqur Rahman and Noorjahan Begum for their great sacrifices for me. I am thankful to my brother, Swadipta Alam, for his love and care, for taking care of the family and allowing me to pursue higher studies. I am thankful to my sweet sister, Ahona Abontika, for her love for me and for the family. I want to thank my uncle Idris Bhuyan and Muhibur Rahman for their unconditional support in my life. It would be impossible to survive and finish Ph.D. without the support from all these beautiful people.

I would not be here without my wife, Maria. She was the reason I started this journey. She was the reason behind so many good things in my life. She brought happiness and

joy, and most importantly, a dream of doing great thing in life. Ph.D. is a part of this dream. I want to thank her for being with in every single moment of this journey. I want to thank Fiorella, for completing my life.

And lastly, I want to thank Allah for every single thing.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	viii
LIST OF FIGURES	xiv
LIST OF TABLES	xix
LIST OF ABBREVIATIONS	xx
1 Chapter 1: Introduction	1
1.1 Statement of the Problem	1
1.2 Hypothesis and Objectives	5
1.3 Literature Survey	6
1.3.1 Calibration set requirements.....	8
1.3.2 Strategies for developing calibration set.....	13
1.3.3 Effects of calibration set on model performance.....	39
1.3.4 Conclusion	49
2 Chapter 2: Method development of designing calibration set in spectral space	50
2.1 Theory.....	50
2.2 Spectral calibration set for single API formulation.....	62
2.3 Spectral calibration set for multiple API formulation	67
2.4 Conclusion	72
3 Chapter 3: Comparison between the model performance of spectral calibration and traditional calibration sets	73
3.1 Introduction.....	73
3.2 Single API formulation.....	75
3.2.1 Material and Methods.....	75
3.2.2 Results	82
3.2.3 Conclusion	93
3.3 Multiple APIs formulation	95
3.3.1 Material and Methods.....	95
3.3.2 Results and Discussion.....	103
3.3.3 Conclusion	120
4 Chapter 4: Selecting loading vector for PLS model during quantitative analysis of tablets using NIRS.....	122
4.1 Introduction.....	122

4.2	Theory.....	124
4.3	Material and Methods.....	128
4.3.1	Scale Variation	128
4.3.2	Environmental Variation	129
4.3.3	Physical Variation (Density)	129
4.3.4	Chemical Variation (Degradant)	131
4.3.5	Raw Material Variation	132
4.3.6	Spectral collection	133
4.3.7	Modeling strategy	133
4.3.8	Cross validation.....	133
4.3.9	Model evaluation and design comparison parameters.....	134
4.4	Results	134
4.4.1	Scale Variation	134
4.4.2	Environmental Variation	138
4.4.3	Physical Variation (Density)	138
4.4.4	Chemical Variation (Degradant)	143
4.4.5	Raw Material Variation	145
4.5	Discussion	148
4.6	Conclusion	150
5	Chapter 5: Method development for incorporating physico-chemical variation into the spectral calibration set.....	151
5.1	Introduction.....	151
5.2	Material and method.....	155
5.2.1	Incorporation of physico-chemical information.....	155
5.2.2	Selection of calibration candidates	163
5.2.3	Spectral calibration set for Acetaminophen.....	176
5.2.4	Traditional calibration set for Acetaminophen.....	177
5.2.5	Spectral calibration set for Caffeine	178
5.2.6	Traditional calibration set for Caffeine	180
5.2.7	Test sets to evaluate model performance.....	180
5.2.8	Quantitative model development.....	181
5.3	Results and discussion.....	182
5.3.1	Prediction of Acetaminophen	182

5.3.2	Prediction of Caffeine.....	185
5.4	Conclusion	188
6	Chapter 6: Optimum calibration structure for pharmaceutical formulation and spectroscopic techniques.....	190
6.1	Introduction.....	190
6.2	Material and Method:	191
6.2.1	Calibration and test set	191
6.2.2	Quantitative model development.....	194
6.3	Results and discussion.....	209
6.3.1	Comparison between optimum calibration sets of NIR and Raman spectroscopy 209	
6.4	Conclusion	220
7	Chapter 7: Summary.....	221
	Appendices	229
	Appendix A (Motivation for using scores as design factors for calibration)	229
	Appendix B (Derivation of equation 2.7).....	230
	Appendix C (Derivation of equation 2.8).....	230
	References.....	232

LIST OF FIGURES

Figure 1-1. TNIR spectra of double-layered tablet consisting of various ratios of forms I and III CBZ (A) and Model regression vector (B) [100]	20
Figure 1-2. Tablet chemical composition. Squares were the calibration data points, and circles were the prediction data points [111]	24
Figure 1-3. Accuracy profiles obtained from the work of Feng and Hu [27] for the quantification of Roxithromycin (A) and Erythromycin ethylsuccinate (B) using different NIRS equipment. Dotted horizontal lines (black): $\pm 5\%$ acceptance limits; dotted and dashed horizontal lines (green): $\pm 15\%$ acceptance limits; dashed lines (blue): 95% prediction intervals; dots (black): relative predicted concentrations of the corresponding NIRS methods [138].....	31
Figure 1-4. Improvement in calibration performance as a function of the number of production sample included into the calibration set with optimized latent variables numbers (A) and fixed latent variable numbers (B). Each symbol represents a different sequence of sample addition [62].	34
Figure 1-5. Relative residuals (%) plots obtained from the prediction of four external sample sets, using also four calibration models with different concentration ranges (% w/w, API). Vertical lines indicate the MDL/MQL for each model. Four predictions are shown for calibration ranges: (a) 0–5.00%, (b) 0–2.00%, (c) 0–0.25%, and (d) 0–0.10% (w/w) [156].....	41
Figure 1-6. Prediction errors of five calibration designs (first column) predicting three prediction sets (first row). The asterisks indicate experimental or predicted areas [28]	44
Figure 1-7. Score plot from second derivative reflectance (A) and transmittance (B) spectra by principal component analysis. Open symbols—calibration set: (○) flat-faced tablets with various content; closed symbols—prediction set: (●) flat-faced tablets with various thickness, (Δ) flat beveled edge tablet, (□) convex tablet. The numbers by the data points correspond to the batch no [160]	46
Figure 2-1. Concentration space of a ternary system	50
Figure 2-2. Spectral data (direct spectral space (A)) presented in PCA score space (factor space (B))	51
Figure 2-3. Hypothetical pure component spectra (A) and respective spectral space (B).....	53
Figure 2-4. Hypothetical concentration space (left) and respective spectral space (right).....	54
Figure 2-5. Procedure for designing scores around the model tablet scores in spectral space	58
Figure 2-6. Effect of rotation matrix in spectral simulation	60
Figure 2-7. Pure component spectra and respective basis vectors.....	62
Figure 2-8. A full factorial design around the model tablet score in the orthogonal spectral space	63
Figure 2-9. Orthogonal rotation in spectral space to select calibration candidates	66
Figure 2-10. Pure component spectra and respective basis vectors.....	68
Figure 2-11. PCA scores of the simulated spectra and selected calibration candidates by KNN..	69
Figure 3-1. Raw data for two calibration sets, solid line (-) represents the full factorial design calibration samples and broken line (--represents the spectral design calibration samples.....	82
Figure 3-2. Projection of the full factorial calibration, spectral calibration and three prediction sets samples on the PC1 and PC2 of the combined PC space.	83

Figure 3-3. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 1 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 1 (lower panel).....	85
Figure 3-4. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 2 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 2 (lower panel).....	87
Figure 3-5. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 3 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 3 (lower panel).....	89
Figure 3-6. Confidence interval for bias (A) and standard error of prediction (B) comparison between full factorial design and spectral design for prediction sets analysis.	91
Figure 3-7. Experimental procedures for designing a calibration set in spectral space.....	98
Figure 3-8. Quantitative predictions of NIR images of tablet at target formulation and center compaction force.	103
Figure 3-9. Solid fraction of tablets compressed at different compaction forces.	104
Figure 3-10. Tablet compositions of the traditional full factorial calibration, the spectral calibration and the test set.	105
Figure 3-11. Calibration model performance for traditional full factorial design during Acetaminophen prediction.....	106
Figure 3-12. Calibration model performance for spectral design during Acetaminophen prediction	107
Figure 3-13. Calibration model performance for traditional full factorial design during Caffeine prediction.....	109
Figure 3-14. Calibration model performance for spectral design during Caffeine prediction. ...	110
Figure 3-15. Tablet compositions of the spectral and traditional calibration designs (filled circles), and the test set (open circles).....	113
Figure 3-16. Calibration model performance of the spectral design and traditional designs during Acetaminophen prediction.....	114
Figure 3-17. Calibration model performance of the spectral design and traditional designs during Caffeine prediction.	115
Figure 3-18. Cross design performance of acetaminophen model (left) and caffeine model (right)	116
Figure 4-1. RMSECs and random subset (left), contiguous block (middle) and venetian blind (right) RMSECVs during calibration development for scale variations.	134
Figure 4-2. Cumulative variances explained at each loading vector during calibration development for scale variations.	135
Figure 4-3. Weight co-efficient and percent of variances explained by each loading vector during calibration development for scale variations.	136
Figure 4-4. RMSEPs of models predicting Acetaminophen in weekly runs prepared at different scale.	137
Figure 4-5. RMSEPs of models predicting Acetaminophen in monthly runs prepared at different environment.....	138
Figure 4-6. Weight co-efficient of each loading vector during calibration development for density variations.	140
Figure 4-7. RMSEPs of models predicting Theophylline in tablets with density variations.....	142

Figure 4-8. RMSECs and random subset (right), contiguous block (middle) and venetian blinds (left) RMSECVs during calibration development for chemical variations.	143
Figure 4-9. Cumulative variances explained at each loading vector during calibration development for chemical variations.....	144
Figure 4-10. Weight co-efficient and percent of variances explained by each loading vector during calibration development for chemical variation.....	145
Figure 4-11. RMSEPs of models predicting Niacinamide in presence of chemical variations (Niacin).....	146
Figure 4-12. Weight co-efficient and percent of variances explained by each loading vector during calibration development for raw material variation.....	147
Figure 4-13. Cumulative variances explained at each loading vector during calibration development for raw material variation	148
Figure 4-14. RMSEPs of models predicting Theophylline in presence of raw material variations.	149
Figure 5-1. Effects of tablet density and hardness on reflectance and transmission NIR [193]..	152
Figure 5-2. Spectral calibration strategy for incorporating tablet density variation into the spectral space	154
Figure 5-3. Effect of compression force on the pure component tablet spectra	155
Figure 5-4. Effect of compression pressure on the tablet of target formulation.....	157
Figure 5-5. Effect of residual spectra on the simulation of spectral response	158
Figure 5-6. Simulated spectral response of 225 design points	159
Figure 5-7. Principal component analysis of the simulated spectra of 225 design points. Left: Color coded based on compaction forces (2000, 4000 and 6000 lb), Right: Color coded based on Acetaminophen concentration levels (18.75%, 25%, 31.25%, 37.5% and 43.75% w/w).....	160
Figure 5-8. Principal component analysis of the simulated spectra of 225 design points. Left: Color coded based on Caffeine concentration, Right: Color coded based on MCC to Lactose ratio	161
Figure 5-9. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Left: Color coded based on compaction forces, Right: Color coded based on Acetaminophen concentration	162
Figure 5-10. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Left: Color coded based on compaction forces, Right: Color coded based on Acetaminophen concentration	163
Figure 5-11. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Figure is color coded based on MCC to Lactose ratio.....	164
Figure 5-12. Score plot from the PLS model. The figure is color coded based on Acetaminophen concentration level	165
Figure 5-13. Score plot from the PLS model. Left: Color coded based on compaction force. Right: Color coded based on MCC:Lactose ratio.....	166
Figure 5-14. Selection of calibration candidates (indicated by black circle) in the PLS score space. The figures are color coded based on compaction force (Left), Acetaminophen concentration level (upper right) and MCC:Lactose ratio (bottom right).....	167
Figure 5-15. Distribution of Acetaminophen in global calibration set and spectrally selected calibration set	168
Figure 5-16. Distribution of Caffeine in global calibration set and spectrally selected calibration set	169

Figure 5-17. Distribution of MCC (left) and Lactose (right) in global calibration set and spectrally selected calibration set.....	170
Figure 5-18. Distribution of compaction forces in global calibration set and spectrally selected calibration set	171
Figure 5-19. Distribution of Acetaminophen (left) and Caffeine (right) concentration in global calibration set and spectrally selected calibration set.....	172
Figure 5-20. Distribution of MCC (left) and Lactose (right) concentration in global calibration set and spectrally selected calibration set	172
Figure 5-21. Distribution of compaction forces in global calibration set and spectrally selected calibration set	173
Figure 5-22. Solid fraction of tablets at different compaction forces (A). Spectral variation of the tablet sets prepared at different compaction forces (B).....	174
Figure 5-23. Spectral variation before (A) and after MSC correction of the spectral data.	175
Figure 5-24. PLS scores of first two loading vectors at three different compaction forces	176
Figure 5-25. KNN distance scores of samples at three different compaction forces for PLS models of Acetaminophen (A) and Caffeine (B).....	177
Figure 5-26. Selection and preparation method of spectral calibration set and test set for Acetaminophen and Caffeine	180
Figure 5-27. PLS model performance of the global calibration set predicting Acetaminophen in test set 1	183
Figure 5-28. PLS model performance of the spectral calibration set predicting Acetaminophen in test set 1	184
Figure 5-29. PLS model performance of the traditional (A) and spectral calibration sets (B) predicting Acetaminophen in test set 2	185
Figure 5-30. PLS model performance of the traditional calibration set predicting Caffeine in test set 1	186
Figure 5-31. PLS model performance of the spectral calibration set predicting Caffeine in test set 1	187
Figure 5-32. PLS model performance of the traditional (A) and spectral calibration sets (B) predicting Caffeine in test set 2.....	188
Figure 6-1. NIR (A) and Raman (B) spectra of the calibration sets.....	193
Figure 6-2. NIR PLS model for Acetaminophen from traditional calibration set.....	194
Figure 6-3. Reference vs prediction plot of Acetaminophen from traditional calibration set.....	195
Figure 6-4. Search for optimum calibration set.....	196
Figure 6-5. Prediction performance of randomly selected calibration sets and optimally selected calibration set	197
Figure 6-6. NIR PLS model for Acetaminophen from optimum calibration set.....	198
Figure 6-7. Reference vs prediction plot of Acetaminophen from optimum calibration set.....	200
Figure 6-8. NIR PLS model for Caffeine from traditional calibration set.....	201
Figure 6-9. Reference vs prediction plot of Caffeine from traditional calibration set.....	201
Figure 6-10. NIR PLS model for Caffeine from optimal calibration set.....	202
Figure 6-11. Reference vs prediction plot of Caffeine from traditional calibration set.....	203
Figure 6-12. Raman PLS model for Caffeine from traditional calibration set	203
Figure 6-13. Reference vs prediction plot of Acetaminophen from traditional calibration set...	204
Figure 6-14. Raman PLS model for Acetaminophen from optimal calibration set.....	205

Figure 6-15. Reference vs prediction plot of Acetaminophen from optimum calibration set	206
Figure 6-16. Raman PLS model for Caffeine from traditional calibration set	207
Figure 6-17. Reference vs prediction plot of Caffeine from traditional calibration set.....	208
Figure 6-18. Raman PLS model for Caffeine from optimum calibration set	208
Figure 6-19. Reference vs prediction plot of Caffeine from optimum calibration set.....	209
Figure 6-20. Compositional points between full global, NIR optimized and Raman optimized calibration sets for Acetaminophen.	210
Figure 6-21. Acetaminophen distribution between NIR and Raman optimized calibration sets	211
Figure 6-22. Concentrations of Acetaminophen and MCC in full traditional, NIR optimized and Raman optimized calibration sets.....	212
Figure 6-23. MCC (left) and Lactose (right) distribution between NIR and Raman optimized calibration sets for Acetaminophen	212
Figure 6-24. NIR (left) and Raman (right) spectra of MCC and Lactose.....	213
Figure 6-25. Compositional points between full global, NIR optimized and Raman optimized calibration sets for Caffeine.....	214
Figure 6-26. Caffeine distribution between NIR and Raman optimized calibration sets	215
Figure 6-27. Concentrations of Caffeine and MCC in full traditional, NIR optimized and Raman optimized calibration sets.....	216
Figure 6-28. MCC (left) and Lactose (right) distribution between NIR and Raman optimized calibration sets for Caffeine.....	217

LIST OF TABLES

Table 1-1. Advantages and disadvantages of different strategies for NIR calibration set development	38
Table 1-2. Advantages and disadvantages of different strategies for NIR calibration set development	49
Table 2-1. Composition of the spectral calibration set for single API.....	65
Table 2-2. Composition of the spectral calibration set for multiple API.....	70
Table 2-3. Correlation coefficient between composition and compression force.....	72
Table 3-1. Composition (% w/w) of the full factorial calibration set.....	76
Table 3-2. Model performance comparison between two methods of experimental design; spectral and full factorial.....	90
Table 3-3. Full-factorial calibration design and test set	96
Table 3-4. Model performance comparison between full factorial and spectral calibration design for quantitative analysis of Acetaminophen	108
Table 3-5. Model performance comparison between full factorial and spectral calibration design for quantitative analysis of Caffeine.....	111
Table 3-6. Calibration performance of spectral and traditional calibration designs.....	118
Table 4-1. Calibration for Density variation	141
Table 5-1. Full factorial calibration design	157
Table 5-2. Design of the two test sets.....	179
Table 5-3. Summary of the calibration model performance of the traditional and spectral calibration set	189
Table 6-1. Calibration and test design.....	192
Table 6-2. Model statistics for global, optimum and cross calibration sets NIR and Raman during the prediction of Acetaminophen	217
Table 6-3. Model statistics for global, optimum and cross calibration sets NIR and Raman during the prediction of Caffeine.....	218
Table 6-4. Number of common samples between the optimum calibration sets of different techniques and different APIs	219

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
ANOVA	Analysis of Variance
API	Active Pharmaceutical Ingredient
CCD	Central Composite Design
CLS	Classical Least-Squares
DOE	Design of Experiments
DOPT	D Optimal Design
FDA	The United States Food and Drug Administration
HPC	Hydroxypropyl Cellulose
HPLC	High Pressure Liquid Chromatography
HPMC	Hydroxypropyl Methylcellulose
ICH	International Conference on Harmonisation
IOPT	I Optimal Design
Mg St.	Magnesium Stearate
MCC	Microcrystalline Cellulose
MDL	Multivariate Detection Limit
MQL	Multivariate Quantification Limit
NAS	Net Analyte Signal
NIR	NIR
NIRS	NIR Spectroscopy
PAT	Process Analytical Technology

PCA	Principal Component Analysis
PLS	Partial Least-Squares
PSD	Particle Size Distribution
QbD	Quality by Design
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
RMSEC	Root Mean Squared Error of Calibration
RMSECV	Root Mean Squared Error of Cross Validation
RMSEP	Root Mean Squared Error of Prediction
SEC	Standard Error of Calibration
SEP	Standard Error of Prediction
SNV	Standard Normal Variate
SVD	Singular Valued Decomposition
USP	United States Pharmacopeia

1 Chapter 1: Introduction

1.1 Statement of the Problem

NIR spectroscopy (NIRS) is a well-established analytical tool in the pharmaceutical industry for quantitative analysis of tablets [1-5]. NIRS offers substantial advantages in terms of time and cost compared to the wet chemical methods. However, successful application of this technology depends on an appropriate calibration method, and an appropriate calibration method is built upon appropriate calibration sample design. Typical steps to develop a NIR quantitative calibration method for pharmaceutical tablets are: calibration design, sample set preparation, spectral data collection, reference analysis, and model development [2]. Defining the calibration sample set is the first step for calibration model development. There are a number of requirements for an appropriate calibration set for tablet analysis; examples include spanning the concentration range for an active pharmaceutical ingredient (API) and excipients, multiple lots of raw materials, different processing conditions, mutually independent samples, etc [6]. A proper calibration set is a pre-requisite for generating appropriate spectral data and achieving desirable predictive performance from NIRS. Therefore, the size and direction of the calibration spectral subspace is of critical importance for successful quantitative model development [7].

Other elements of NIR method development have been the subject of many comprehensive reviews and continuous advancements (e.g. pretreatment techniques, regression methods). Yet, analysis of calibration set design and its effect on model performance has received far less attention in the literature and regulatory guidelines. There is no standardized reference method for developing spectroscopic calibration sets

for quantitative analysis of tablets. Practical approaches used for designing calibration sets include experimental design entirely in concentration space (e.g. full-factorial design, central composite design) [8-11], calibration sample selection from large data set based on a selection algorithm [12-16], preparation of sample set that spans target concentrations and anticipated variance (including physical sources of variance) over the product lifecycle [17-21]. The objective of these techniques is to generate an appropriate calibration set that contains ‘sufficient’ information about the system to develop a successful quantitative model. As sufficient information is challenging to define, the general idea is to incorporate as much variance as possible in an effort to ensure that all future variance is captured in the calibration set [22]. This idea often leads to the creation of a large calibration set. The reported numbers of calibration samples for NIR quantitative methods are typically large for pharmaceutical tablets, such as 500 [23], 450 [24], 414 [25], 297 [26] and 276 [27]. Preparation of such large calibration sets is costly, especially in the product development phase when API is expensive. It is also very time consuming to perform the reference analysis (typically HPLC) of such large calibration sets. A small calibration set with desired predictive performance would be helpful to save time and cost associated calibration method development for tablet analysis using NIRS.

Ironically, the traditionally developed large and expensive calibration sets typically contain redundant information. Redundant information is not necessary for developing a successful quantitative model. However, samples containing redundant information are typically identified in posterior analysis from their spectral responses. An analyst often ends up preparing and analyzing samples that are unnecessary for the intended purpose. Currently, there is no published technique to identify the redundant samples at the outset

of calibration development. Such a technique would be helpful to reduce the efforts in preparing redundant calibration samples and save considerable time and cost associated with calibration development.

It is also possible to miss the critical samples in the large calibration sets leading to a lack of robustness of the method despite specific efforts to the contrary. A calibration set may simultaneously contain redundant spectral information and lack the necessary information to provide model robustness. Current approaches for calibration development do not provide opportunity to identify the most critical calibration samples at the outset of calibration development. Such a technique is useful to identify the smallest possible calibration set with desired predictive performance.

Information about critical sample requirements of a calibration set is most useful at the outset of method development to design a small and efficient NIR calibration set for tablet analysis. This requirement is formulation specific and is dependent on the spectral response of the individual constituent. Current strategies for designing NIR calibration sets do not account for variations in pharmaceutical formulations and their spectral responses. Regulatory guidelines and general rules of thumb are usually followed for all types of formulations resulting in similar concentration ranges, levels and number of samples for all types of calibration sets. However, an optimal calibration design for one formulation can be sub-optimal for a different formulation. Different NIR responses produced by different formulations should be considered during the selection of calibration range, size, concentration levels and variance information. Variation of the spectrally similar components might give redundant information and need not to be varied simultaneously in the calibration set. Concentration of a component with relatively

weak spectral response might not need to be varied extensively whereas the component with dominant spectral features might require wide variation in the calibration set to develop a successful calibration model. Non-linearity between the concentration and spectral response can be highly dependent on the formulation and analytical technique of interest. The same calibration designs of different formulations can be mapped differently into the spectral space of an analytical technique, generating redundant information for one formulation while lacking necessary information for the other. A formulation specific calibration strategy should be developed to best utilize NIR capability for quantitative analysis of tablets.

A calibration set should also be specific to analytical technique. Currently, little or no differences are usually found between NIRS and other spectroscopic calibration designs. The same calibration design of the same formulation can perform differently depending on the analytical technique. Practically, there is no best calibration design that works for all purposes. A great deal depends on the method requirements and resource availability [28]. The traditional practices for developing calibration sets are inefficient in utilizing the available resources and minimizing redundant information. The calibration strategies are also too generalized to account for formulation variability and variation in spectral responses. A formulation and technique specific calibration strategy is required to make the calibration process efficient and optimize the calibration performance.

This dissertation demonstrates a strategy for designing a calibration set in spectral space for quantitative analysis of tablets using NIRS. This strategy requires fewer calibration samples to achieve desired prediction performance of the calibration model. This strategy

also offers a formulation and technique specific calibration experiments for calibration method development.

1.2 Hypothesis and Objectives

The central hypothesis of this dissertation is that designing experiments in spectral space provides a more efficient means compared to current practices for developing spectroscopic calibrations for quantitative analysis of pharmaceutical tablets.

Given the central hypothesis, the objectives of this dissertation work were,

1. Develop a strategy to design NIR calibration sets in spectral space for pharmaceutical tablets
 - a. Calibration design for a model drug product containing one API.
 - b. Calibration design for a model drug product containing multiple API.
2. Evaluate calibration performance between spectral calibration sets and traditional calibration sets for quantitative analysis of tablets using NIR
 - a. Comparative studies between calibration strategies for single API.
 - b. Comparative studies between calibration strategies for multiple API.
3. Develop a strategy to incorporate physical information into the spectral design of NIR calibration set
 - a. Develop strategy to incorporate physical information into the spectral space
 - b. Evaluate calibration model robustness between spectral calibration sets and traditional calibration sets

4. Compare optimized calibration sets of NIR and Raman spectroscopy for quantitative analysis of pharmaceutical tablets

1.3 Literature Survey

The first calibration set development and quantitative analysis of pharmaceutical components using NIRS was reported in 1961 [29]. A series of reports followed later on the quantification of amine salts (first application in a solid system) [30], allylisopropylacetureide and phenacetin [31], meprobamate [32], carisoprodol [33] and water content [34] in pharmaceutical compounds using NIRS¹. Since then, application of this technology has been exponentially expanding due to advancements in computational power, multivariate statistics and also information contained in the spectral range [35]. The pharmaceutical industry has adapted NIRS due to its potential to offer fast, easy and, most importantly, non-invasive and non-destructive analysis technique. However, an appropriate calibration set is required to deal with complex and overlapping features of NIR spectra and non-specificity inherent to the technique [36]. An appropriate calibration set is also critical to ensure consistent calibration performance during product life cycle. Any change in the formulations, physical properties of the components or manufacturing unit operations may cause the validated NIR method to become unusable [37].

Current strategies for developing calibration sets suffer from the redundancy of unnecessary information and the inadequacy of critical information in the calibration design. Redundancy of the calibration set and appropriateness of the calibration designs are critical factors that need to be considered during NIR calibration development. Bondi et al. demonstrated improved accuracy of full factorial and I-optimal designs over other

¹ Calibration development for these reported methods used a univariate approach

types of factorial and optimal designs for NIR calibration development [9]. Although the I-optimal design was a subset of the full factorial design, the difference in predictive performance was statistically insignificant. This work suggests the presence of redundant information in the full factorial calibration design. Xiang et al. analyzed the effect of correlated design on the selectivity and robustness of the calibration models [38]. They demonstrated that the model selectivity and robustness of a randomized design was superior to that of a correlated design. Low correlation in the randomized design was critical to avoid fitting of the partial least squares (PLS) model to the non-selective correlation between spectra and excipient concentration. Naes et al. compared five mixture designs for developing calibration models and concluded that similar variation was required between the calibration and the prediction sets for successful model development [9, 39]. The same author also demonstrated the tendency of traditional calibration methods to contain redundant information in a subsequent paper; similar prediction performance was achieved with 20 samples out of 114 samples from the calibration set [40]. All of these examples demonstrate that a calibration set may simultaneously contain redundant spectral information and lack the necessary information to provide calibration model robustness. However, this information is available only in retrospective analyses, after considerable time and effort has been spent on the calibration development.

The current strategies for developing NIR calibration set need significant improvement to make the calibration process efficient and optimize calibration performance. Optimization of the calibration strategy requires an in-depth understanding about the requirements of an appropriate calibration set and the effects of calibration structure on

calibration model performance. This understanding can be helpful to identify the necessary information for a calibration set and challenge the current philosophy of incorporating all possible variances into the calibration set. The current recommendation of a large calibration set can be replaced by an efficient strategy for developing NIR calibration set to quantitatively analyze pharmaceutical tablets.

This chapter is aimed at reviewing the general requirements for an NIR calibration set designed to quantitatively analyze pharmaceutical formulations, current strategies for developing calibration sets and their effects on calibration model performances based on the published literature and regulatory guidelines. This discussion is particularly focused on the calibration sets designed to use inverse least squares regression techniques (partial least squares, principal component regression), considering their wide applicability and most frequent usage for quantitative analysis of pharmaceutical formulation.

1.3.1 Calibration set requirements

There are several requirements for an appropriate NIR calibration set. Blanco et al. indicated wide concentration variation and physico-chemical variability as critical requirements [41]. Mutual independence of samples was added to this list by other groups [6, 38]. Naes et al. discussed these requirements in terms of calibration spectral space and associated eigenvectors due to inherent collinearity of the NIR spectra [39]. Presence of all relevant eigenvectors, wide spanning of the eigenvector's directions and even spread of the spectral space over the eigenvector subspace were considered necessary parameters for NIR calibration set. However, calibration set requirements can be further discussed in terms of range and level of concentration, number of samples and variance information.

1.3.1.1 Range and levels of concentration

An appropriate concentration range and variability in calibration set is necessary to ensure NIR method robustness [42]. The concentration range must be selected appropriately to cover the anticipated concentration range in the routine samples. An extrapolation due to inappropriate calibration range is detrimental to calibration model performance. The effect of extrapolation is more critical to multivariate NIRS method compared to univariate counterpart [43]. Although a wide concentration range is encouraged, there are limitations in terms of time and resources available to develop a large calibration set. The typical production samples do not cover wide concentration range as the Pharmacopoeias often allow 5% deviation of the nominal content for the API [42]. Synthetic samples are often prepared in lab to extend the concentration range. International Conference on Harmonisation (ICH) suggests a range of $\pm 20\%$ of the nominal content of API for the assays method and $\pm 30\%$ for the content uniformity method [44]. The concentration range can be further extended by the inclusion of pure component (100%) or placebo tablets (0%) [45]. However, extreme points have high leverage during calibration development. High leverage samples can significantly change model direction and affect prediction performance in the concentration range of interest. Although a wide concentration range is considered to increase model robustness, it can introduce non-linearity. Good fit is almost guaranteed for a small set with narrow region. However, the predictions outside this small range are of little value considering the samples being outside of the validated range of the method [46]. Multiple ranges of a single calibration set are also used for different prediction samples [47]. A prediction sample at the calibration center is predicted by a narrow calibration range samples (local

calibration) whereas, prediction samples at the edge are predicted by a broad calibration range samples (global calibration). This technique is known as bracketing. Bracketing can also be performed by selecting an optimal subset from global calibration set [48]. Optimal subset can be changed each time a new test sample encountered.

There is no guideline reported for excipient concentration range in the calibration set. Scheibelhofer et al. considered the variation of API more critical than the small deviation of the excipients from the target [49]. However, all possible variations in excipient concentrations were recommended to be included into the calibration set [49]. Components having common spectral features were emphasized in this aspect. A regulatory guideline for excipient variation would be helpful in designing an appropriate calibration set with desired predictive performance.

The concentration levels of a calibration set are also critical to develop appropriate calibration model. Uniform distribution of the concentration levels is favored over Gaussian or normal distribution [50, 51]. High leverage samples in the normal distribution are considered undesirable during model development. The ICH Guideline recommends at least five concentration steps to prove linearity [44]. Several samples should be available per source of variance in NIR calibration set [47, 51]. The link between API concentration range and clinical performance has not been discussed either in literature or in guidelines. The calibration range is selected primarily based on acceptable API range in the final dosage forms. A risk-based approach for defining the range would be advantageous to bridge the gap between acceptable API range, calibration range and clinical performance [52].

1.3.1.2 Number of samples

Appropriate number of calibration samples is critical to calibration model performance [53]. ‘Appropriate number’ has been defined as the number providing sufficient chemical and physical variability resulting in appropriate spectral variability [54]. The number is formulation, method and technique dependent and usually decided based on experience. A set of 20–30 calibration points was considered sufficient for model accuracy by several groups [47, 55, 56]. Agelet et al. recommended 100 samples for a robust NIR calibration [50]. This requirement can be smaller for homogeneous samples. ASTM recommends multiplying the number of principal component (PC) by a factor of 6 to determine appropriate sample number [57]. However, a *priori* data set is a prerequisite to calculate the PCs. Mathematical approaches have been reported to determine the number of calibration samples [49]. For a mixture system having “P” levels for each chemical component, the appropriate number of calibration sample ‘N’ can be calculated as,

$$N = \frac{1}{2} * p * (p + 1) \quad (1.1)$$

Li et al. used full factorial approach to determine the number of calibration samples for a system containing “C” chemical components, each at “P” level, using following equation [58].

$$N = P^{C-1} \quad (1.2)$$

A degree of freedom is lost due to the mixture constraint. Such approaches give general direction for calibration sample number. However, these strategies require each varying component to have same number of levels, which may not always be feasible.

The appropriate numbers of calibration samples are also determined by splitting the available data into calibration and test set. Alvarenga et al. suggested that, the calibration should not be smaller than the 2/3 of the data set [59]. A smaller set is preferred to save time and cost, however, a minimum number must set to meet the statistical requirement. A two-points (extreme) calibration set was also reported during NIRS method development [60], for which statistical validity was questionable.

1.3.1.3 Variance information

The information content of the calibration set and quality of the calibration samples are considered to be more important than the number of calibration samples [22, 61]. A large variation in all directions at the region of interest is a critical requirement for an appropriate calibration set [59]. The information should include variation due to concentrations as well as physical properties, moisture contents [51, 62] of the samples and other factors that are bound to influence NIR spectra [63]. The calibration samples should have a wider range of variation than the production samples to ensure model robustness [19]. The wider range samples can be prepared at the laboratory scale or by spiking (under-dosing/over-dosing) the production samples with appropriate amount of API and/or excipients. However, lab-engineered samples were criticized due to its potential to introduce greater than nominal variances and increase prediction errors [62].

A calibration model has the potential to fail during the prediction of future samples having different interference information such as temperature effects, season to season variation, different tablet for the same active component, and batch effects in industrial

production processes [64-67]. The structure and information content of the calibration set are critical for achieving appropriate representation of the future samples [46].

Several requirements are reported and they provide a general guideline for NIR calibration set development. However, many of the calibration requirements can be method specific and can vary depending on method requirements. Instead of a general calibration design guideline, a standard method specific protocol is essential to meet the calibration requirements. It is also critical to establish appropriate metrics to ensure the fulfillment of calibration set requirements.

1.3.2 Strategies for developing calibration set

The main objective of developing a NIR calibration set is to represent routine production samples and aid in developing a successful predictive model. Routine samples should not be used alone to develop calibration set as these samples contain API and excipients in amounts very close to the target composition, which precludes spanning a wide enough concentration range for calibration set [41, 63, 68]. Several strategies are used to develop calibration sets. Sarraguca et al. reported four strategies [63] for developing calibration set using:

- 1) Production samples with a wide concentration range [24, 69, 70]
- 2) Production samples spiked with API/excipients [71-73]
- 3) Designed laboratory samples [73, 74]
- 4) Combined set of production and laboratory samples [42]

Blanco et al. discussed similar strategies and recommended a combined approach [75]. The author also demonstrated a strategy using laboratory samples with a calculated process spectrum to explain the spectral difference between laboratory and production samples [41]. Spiking production samples and designing laboratory samples are the two most common strategies. These strategies were found to be similar in terms of model performance [63, 72, 73, 75]. Spiked samples usually match with the routine production samples in terms of physical properties offering additional advantages [73]. However, a high correlation between component's concentrations in spiked samples can lead to poor model specificity. Experimental design of the laboratory samples can be used to reduce the correlation between component concentrations [42]. Another widely used strategy involves the selection of calibrations set samples from a large data using different selection techniques [74]. In this report, strategies are classified into three major categories based on the techniques used for developing NIR calibration set:

- A. Use of design of experiment to prepare a calibration set
- B. Use of sample selection techniques to select a calibration set
- C. Under dosing/over dosing (spiking) production samples to prepare a calibration set

1.3.2.1 Use of design of experiment to prepare calibration set

Design of experiments (DoE) is a statistical tool to design a controlled and limited number of experiments to gain in-depth understanding about the relationship between factors and responses of an underlying system. In the field of multivariate calibration (MVC), DoE is used to design calibration samples to define the relationship between

concentration and spectral response. The implementation of DoE is challenging in MVC due to the inconsistency between the theory and practical application. In the calibration design, the concentrations are considered as factors and spectra are considered as response. In contrast, in the model system (regression) the concentration and spectral response act as response and factors, respectively. Although this inconsistency has been previously addressed [76], a practical solution is yet to be introduced. There are several types of experimental designs employed to develop NIR calibration set.

1.3.2.1.1 Full Factorial Design

During spectroscopic calibration development, a full factorial design is used to prepare calibration set using all possible combinations of the varying components. The details of full factorial design can be found elsewhere [77]. Although this design is intended to provide orthogonality, it is not possible to attain complete orthogonality in calibration design of a pharmaceutical system due to the mixture constrain (total is 100%), especially in solid samples. Therefore, the orthogonality between API and major excipients are emphasized over the minor excipients. Complete orthogonality can be achieved using sample solutions. A full factorial orthogonal design was created to build NIR calibration dataset for predicting human serum albumin, gamma-globulin and glucose in control serum and phosphate buffer solution [78, 79]. However, the calibration set was developed by randomly selected 80 samples out of 125. Thus, the orthogonality of the design was not maintained in the calibration set. The effect of randomized selection on calibration model performance was not analyzed.

Selection of design factors and levels are critical during calibration development. In most cases, the API and major excipients are considered as the factors of the design, because the primary objective of the calibration design is to develop mathematical relationship between API concentration and spectral response. The variations in the major excipients are required to train the calibration model against non-specific variation in the spectral response. Other excipients usually act as fillers to fulfill the constraint. For multiple API, only the APIs can act as factors [80]. A component ratio can also act as a factor of the calibration design. Using a ratio as factor helps to reduce the number of design points while maintaining wide range of concentration. Igne et al. considered API concentration and major excipient's ratios as two factors of the design during the development of a calibration set for blend monitoring [81, 82]. The first calibration set contained spectra from different time points of blending runs whereas the second set contained spectra at the end of the blending runs assuming that the blend reached homogeneity.

Aside from the compositions, external factors such as moisture content, batch, day, operator and collection temperature can also be included in the design to introduce external variance into the calibration set [83]. Inclusion of the external factors usually leads to a robust calibration model. However, it can also affect model accuracy due to spectral sensitivity to the external factors. Kamran et al. included the API concentration, particle size distributions (PSD) and compression pressures as the design factors during the quantification of ibuprofen in tablets [84]. The calibration performance was found to be sensitive to PSD variation. The model performance was improved by developing the calibration set on uniform PSD of API and filler. In this study, the sample set was randomly split into calibration and validation set. An independent validation set should

have been used to retain the orthogonality of the design in calibration set. Such validation set was used to assess calibration performance during the quantification of red iron oxide in tablets [85]. The calibration model was found to be sensitive to the thickness variation built into the calibration design. A uniform distribution of the tablet thicknesses in calibration set would be helpful to make the model robust against thickness variation.

Full factorial design can be combined with a fractional factorial design during calibration set development. Cogdill et al. used a combination of full factorial (in production scale) and fractional factorial design (in lab scale) during API quantification in tablets [23]. The common design factors were the source of API and Magnesium stearate, granules' moisture content and compression force. Additionally, in the fractional factorial design, API content was varied at $\pm 30\%$ of the nominal content. The use of a fractional design helped to reduce the number of design points. The calibration model was demonstrated to be robust against source, time, and compaction pressure variation. Another calibration set was developed for predicting Radial Tensile Strength (RTS) of the tablet. Compression force was varied in the calibration set to produce tablets with different RTSs.

1.3.2.1.2 D-Optimal Design

D-optimal design is widely used for developing calibration set with optimum number of samples when the classical symmetrical designs cannot be used because of the large number of design points or the irregular shape of the experimental region [86]. This technique is used to maximize the information content of a design by maximizing the determinant (D) of the information matrix ($X'X$). The details of the design can be found elsewhere [87].

Karande et al. used a D-optimal design to develop a NIR calibration set for blend monitoring [88]. A set of 24 design points was created by D-optimal design strategy out of which, 3 design points were removed for the test set leaving 21 design points for the calibration development. These 21 design points might not be the optimum set of 21 points that had maximum determinant. It would be ideal to create a separate design for the test set to maintain optimality of the calibration set. In this study, the calibration range was 2-6 % and the resulting calibration model was used for predicting a full blend run where API concentration outside the range was expected to be encountered during the blending process. Prediction outside the calibration range should have been considered with limited confidence. D-optimal design was also used to create a calibration set for monitoring a five component blending run [89].

Tomuta et al. used D-optimal design strategy to develop a calibration set for quantitative analysis of API and excipients in meloxicam tablets [90]. The ranges of variation were kept $\pm 20\%$ from the target composition for all components and levels of variation were different for different components. The calibration model was able to predict meloxicam and isomalt concentration accurately, but unable to predict sodium starch glycolate and magnesium stearate due to low concentration. However, meloxicam and sodium starch glycolate had similar concentrations (6.25% and 5%) but different levels of variation in the calibration design. The effect of level variations on prediction performance was not analyzed. In another study, D-optimal design was used to prepare a set of placebo samples during the determination of ibuprofen content in production granules, tablet cores and coated tablets [91]. The calibration samples were prepared by mixing different levels of API with randomly selected placebo matrix and compressing them at different

pressures. The randomization of the placebo mixture helped to mitigate the possibility of compositional correlation in the calibration design.

D-optimal design was also used to create a calibration set during flow property analysis of pharmaceutical powder [92]. A sample set with compositional variations was designed to cause variation in flow properties e.g., angle of repose, compressibility index and hausner ratio. A similar approach was taken to analyze the PSD of a pharmaceutical powder system [93]. Compositional variation was designed and assumed to create variation in the PSD of the final formulation. The use of a D-optimal design ensured maximum concentration variation with given number of design points. However, it did not ensure maximum variation in the PSD with the given design points. D-optimal design was also used to prepare a calibration set during the determination of simvastatin and excipients in lyposomes [94], diacetylmorphine, Caffeine and Acetaminophen [95] and cetirizine dihydrochloride in powder samples [96].

1.3.2.1.3 Binary Design

In the statistical literature, binary design refers to a certain type of block design [97]. In this report, binary design is defined as a “design having high correlation between concentrations of two/multiple components in the calibration set”. The correlation is termed as chance correlation in the literature [38]. Although there is poor model specificity associated with this design due to the presence of correlation, it is the most widely used design in multivariate calibration development.

Patel et al. used binary design to create a NIR calibration set for sulfamethoxazole (SMZ) form I prediction in binary polymorphic mixtures and multi-component mixtures [98]. Because of the concentration correlation between two forms of SMZ, model predictive performance for SMZ form I was depended on signals from both forms of SMZ, leading to poor model specificity. Although sample mass has been reported to affect NIR baseline [47], sample mass variation between the binary samples (300mg) and multicomponent samples (330 to 6000mg) were not considered during the assessment of model performance. In the multicomponent mixtures, the diluent concentration was kept constant. A mixture design would be appropriate to allow different levels of the diluent concentrations with same number of design points. Binary design was used to prepare a calibration set for predicting two polymorphs of carbamazepine in bilayer tablets [99].

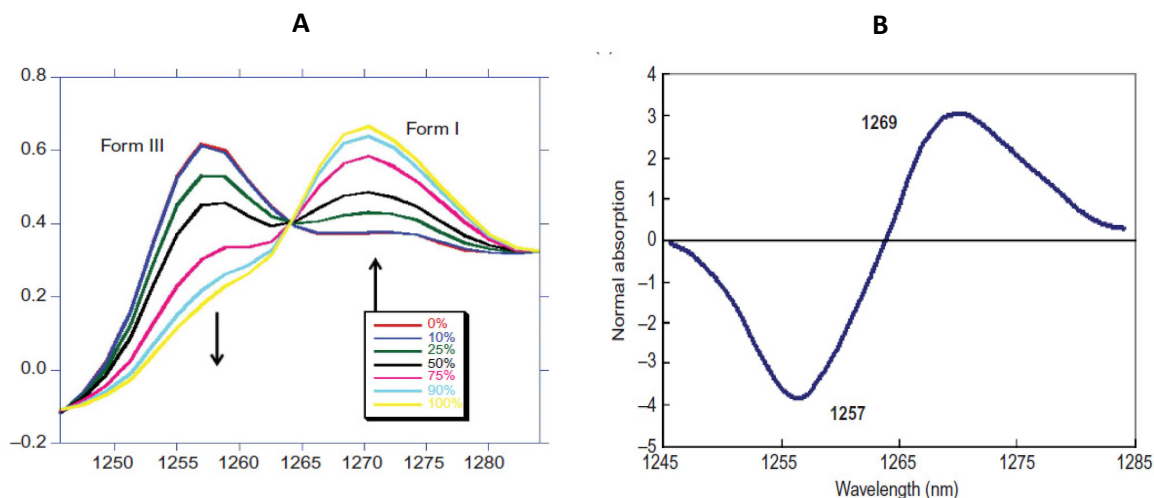


Figure 1-1. TNIR spectra of double-layered tablet consisting of various ratios of forms I and III CBZ (A) and Model regression vector (B) [100]

The calibration set was prepared by varying the ratio of two polymorphs. The spectral changes were highly correlated between two polymorphs due to their perfectly correlated concentrations as shown in Figure 1-1.

A positive regression co-efficient explained the appearance of form I (also disappearance of form III) and negative regression co-efficient explained the opposite scenario. A similar approach was taken to vary crystalline and amorphous cephalexin in a calibration set for pharmaceutical powder [101]. Although these studies reported acceptable model performance, the effect of chance correlation on model specificity was overlooked in the analysis. If the amount of one form were changed while keeping the concentration of the other form constant or vice versa, the model would fail in predicting any of the forms accurately. An appropriate calibration design without correlation could be helpful to overcome this limitation.

Binary design was used during the comparison between reflectance and transmission mode of NIRS based on the ability to predict Theophylline anhydrous concentration in tablets [102]. This design was also used to prepare the calibration sets for determining meloxicam [103], bromazepam and clonazepam [104] and riboflavin [18] content in pharmaceutical preparations. In these studies, several factors such as non-uniform distribution of API, tablet weight variation should have been investigated besides the chance correlation due to their potentials to affect NIR calibration performance.

Introducing a small amount of random variation using replication in the design can reduce the effect of correlation on model specificity. Varying the excipient mixtures instead of an excipient as one factor can also minimize this effect. The correlation

between API and excipient mixtures can be less detrimental to model specificity due to relatively small changes in a particular component concentration. Replication and excipient mixture method was used during a calibration set development for NIR content uniformity method [105]. Model specificity can also be improved by selecting wavelengths that are unique to the analyte of interest [106]. However, this technique can affect model accuracy. In the binary design, full wavelength selection usually provides improved accuracy due to all types of correlated signals. Specificity is an important figures of merit and should always be assessed since it is the underlying chemistry basis for the method [105].

Binary design was also used to prepare a calibration set for quantitative analysis of mixtures containing components of similar or different PSD [60]. Only the two extreme point calibration samples were demonstrated to be sufficient for quantitative analysis of components having similar PSD. However, statistical validity of the two-point calibration is questionable. A larger sample set was required for components with different PSD. Lack of specificity was observed which might be the result of binary calibration design. A similar calibration approach was undertaken to monitor powder mixing between two components of different PSD [107]. Binary design can be less detrimental to a mixing analysis compared to the quantitative analysis of a single component. For a mixing analysis, it is desired for the method to be sensitive to all components of the system whereas for quantitative analysis it should be sensitive to the analyte of interest only.

A calibration set can also contain compositional correlation due to the processing of the calibration samples. For instance, a NIR method was developed to quantify drug content

during a continuous blending process of API, granules and lubricant [108]. Concentration of the API was varied by varying the API flow rates whereas flow rates for the granules and lubricant were kept constant. However, their concentrations must have been changed in the volume of sample scanned by the spectrometer. This change might be directly correlated with the API concentration thus introduced binary structure in the calibration set.

1.3.2.1.4 Mixture Design

Mixture design is also widely used to develop calibration set for NIR method. Mixture design reduces the number of calibration samples while applying constraint on mixture composition using an optimality criterion. The details of the design can be found elsewhere [109].

A mixture design was used to prepare a calibration set during simultaneous determination of chondroitin, glucosamine and ascorbic acid in capsule powder [110] and acetylsalicylic acid, ascorbic acid and Caffeine monohydrate in powder mixtures [96]. Shi et al. used a mixture design to develop a calibration and prediction sets while quantifying Theophylline, MCC and Lactose in tablets [111]. A ternary diagram was used to generate six design points as shown in Figure 1-2. The three outmost points were not included in the calibration because of their lack of leverage for other components. As the same calibration set was used for all components, this led to extrapolated prediction outside the calibration range for MCC and Lactose. Component specific calibration set would be helpful to avoid extrapolation of prediction range.

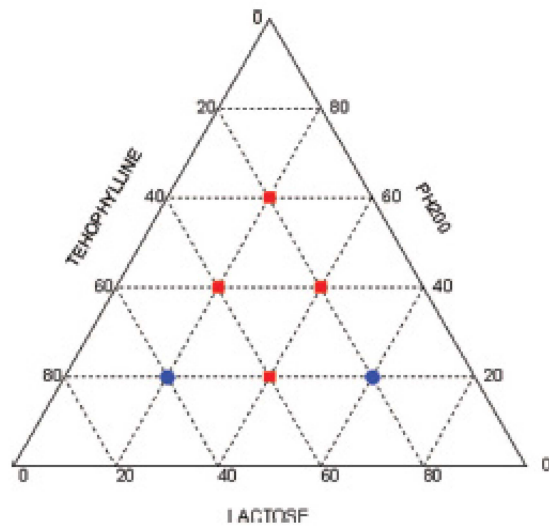


Figure 1-2. Tablet chemical composition. Squares were the calibration data points, and circles were the prediction data points [111]

Short et al. also used mixture design to develop a calibration set for quantifying Theophylline in tablets [11]. A set of 29 calibration sample was designed to cover all concentration ranges of a four-component system. The same design was used during the prediction of crushing strength and density of tablets using NIR reflectance [112] and imaging[113]. Ozdemir et al. used mixture design to develop a calibration set for quantitative analysis of a complex mixture [114]. A set of 21 design points was created and three sets of samples were prepared at the same composition resulted a total of 63 samples. Two calibration sets were developed, one using the first set of 21 samples (set A) and the other using the first and second set of 21 samples (total 42 samples in set B), both predicting the third set of 21 samples (set C). Better model accuracy was observed from set A compared to set B. This was contra-indicatory to the statement that more number of calibration samples helps to minimize model prediction error [48, 115]. Calibration performance does not solely depend on the sample size.

Alexandrino et al. used a combined approach of a mixture and D-optimal designs to develop a calibration set for quantitative analysis of polymer thin films [116]. An experimental domain of polymer mixture was generated by a mixture design with constraints on concentration range. Subsequently a D-optimal design was used to generate a set of 17 design points inside the domain. Different amounts of API were randomly assigned to each design point. Subsequent splitting of the design points into calibration and test was performed. This procedure might not generate the optimum calibration set with appropriate variance. A straightforward approach would be to use a D-optimal design with concentration range constraint. This would allow systematic variance of API along with the excipients.

1.3.2.1.5 Correlation Design

Correlation design is used to assess the correlation between different variables. Although the objective of calibration design is far from finding the correlation between API and excipient concentrations, the correlation between component concentrations is critical to the specificity of the calibration model. Calibration sets have been designed to reduce correlation between component concentrations in NIR calibration set. In this report, such designs are termed as correlation designs. Khan et al. used a correlation design to develop a calibration set for on line monitoring of pharmaceutical blends [117]. The formulation contained API with three major excipients resulted six pairs correlation statistics. A set of 21 design points was selected from 10,000 simulations based on the minimization of the sum of square of six pair correlations. Equal flexibility was given to each component in terms of allowable concentration variation. The resulting ranges of concentration

variations were different for different components, API being the third in descending order. As API concentration variation is more critical than the excipients concentration variation [49], large variation in the API concentration and small variation in the excipients concentration would be advantageous. In another study, the ranges of concentration variations were set higher (70-130%) for the API compared to the excipients (90-110%) [118]. The API and excipients were varied such that all correlation coefficients were below 0.5. The ranges of variation can also be set according to the target concentrations. Correlation design was also used to develop a calibration set for monitoring powder blend [17]. The blend contained 7 chemical components including two APIs. In the design, concentration variations for the most and least concentrated components were $\pm 10\%$ and $\pm 80\%$, respectively. However, the effect of these range variations on NIR model performance was not discussed in any of the studies.

1.3.2.1.6 Other Designs

Several other designs have been used to prepare calibration sets for NIR method. Ito et al. used a central composite design to develop an NIR calibration set while predicting Acetaminophen and Caffeine concentrations in bi-layer tablets [26]. Shi et al. used a Plackett-Burman design to develop a calibration set for quantitative analysis of powder mixture [119] and for blend monitoring [120]. Blend monitoring was also performed using an extreme vertices design [121] and a simplex lattice design [122]. A simplex lattice design of 5 degree and 21 mixture points was used to generate a set of 126 calibration samples varying each component at 6 levels. The same design (excluding mg stearate) was carried out for determining chlorpheniramine concentration in tablets [123].

Alcala et al. used a multivariate design approach for developing a calibration set during API prediction in tablets [124]. The main difference between multivariate design and other conventional designs lies in the data structure and subsequent chemometric analysis (3-way vs 2-way, respectively). The multivariate designs are analyzed through extended modeling techniques (multiway PLS), whereas conventional designs are analyzed through traditional PLS. Osorio et al. used a science based calibration technique where a small calibration set was demonstrated to be sufficient for monitoring a powder blending process [125]. Although a high correlation between component concentrations was present in the calibration design, specificity was assured by selecting unique peak of API and developing an univariate prediction model.

1.3.2.2 Use of sample selection techniques to select calibration set

Calibration sets for a NIR method can also be developed by selecting samples with or without constraints from a given dataset. A random selection process without constraint was performed during the determination of ginsenosides in pharmaceutical tablets [126]. A set of 93 production samples was randomly selected from the manufacturing site for method development. Subsequent randomization was performed to split the data into calibration and test set. Calibration set selection without constraints usually have limited applications. Certain constraints are usually applied during the selection process to ensure regulatory compliance, appropriate concentration range, widespread variance and sufficient representability of the future dataset.

Moffat et al. used a random selection method with constraints to prepare a calibration set for Acetaminophen tablets [74]. The calibration set was selected from fifteen production

and five development batches to include scale variations as well as concentration variation into the calibration model. A similar randomized approach was used to prepare calibration set ranging 80-120% of the nominal API content [21]. Meza et al. prepared a calibration set by randomly selecting samples from specific concentration levels of ibuprofen during tablet quantitation [127]. A randomized approach was also used for splitting compositionally varying dataset into calibration and test set during the quantification of propranolol hydrochloride in powder mixture [128].

The random selection method can be combined with an experimental design to develop a calibration set. Ferreira et al. selected a set of calibration samples randomly from a designed sample set of powder mixtures during hydrochlorothiazide (HCTZ) quantification in powder mixtures and tablets [129]. Certain constraints were applied during the selection step to ensure homogenous distribution of HCTZ in the calibration set. The developed model was used to predict three sets of pharmaceutical samples: powder samples before compression, intact tablets and powder samples obtained by grinding tablets. Tablet prediction had the least accuracy among three sample sets due to the density difference between tablet and powder mixture (calibration samples). This study indicated the criticality of the physico-chemical properties of the calibration samples. However, high correlation ($r^2 \sim 0.99$) between HCTZ and MCC concentrations was overlooked during specificity analysis. A similar method was used to prepare two calibration sets for the quantification of multiple API and tablet RTS [130]. Certain constraints were applied to ensure uniform distribution of APIs and tablet RTS.

The calibration samples can also be selected via different selection algorithms. The Kennard-Stone algorithm was used to select a set of representative calibration samples

for determining amoxicillin content in suspension [131]. Kennard stone algorithm is a widely-used technique for selecting a representative subset from a large dataset. The selection algorithm provides uniform coverage over the dataset by selecting the most separated samples at successive iterations. The algorithm provides unique solution, thereby results in the same representative set for all iterations. The details of the Kennard stone algorithm can be found elsewhere [132]. Li et al. used this algorithm to select the calibration set during azithromycin quantification in tablets of different colors and shapes prepared by different manufacturers [133]. Four different calibration sets were selected from four data sets containing spectra of coated tablets, powdered tablets, uncoated tablets and powdered uncoated tablets. Among the four calibration sets, prediction errors in descending order were coated tablets, powdered coated tablets, uncoated tablets and powdered uncoated tablets. Spectral interference from coating component decreased model's prediction performance. The Kennard Stone algorithm was also used to select a set of representative samples from calibration set for developing a calibration transfer method [134].

Principal Component Analysis (PCA) can also be used to select calibration samples based on the variance information of individual sample [135]. Blanco et al. performed PCA on a dataset containing laboratory and production samples spanning concentration as well as physical variability [136]. Samples that showed large variability in the first two PCs, were included into the calibration set to incorporate maximum spectral variance. However, maximum spectral variance may not always indicate concentration variation. In such scenario, PCA based selection should be performed on the PCs that primarily explain concentration information. Some other physico-chemical parameters along with

PCA analysis can also be used to select the calibration set. Alvarenga et al. prepared a calibration set based on PCA and group statistics on API content, water content and tablet hardness [59]. Samples having wider variability were included into the calibration set to develop a robust calibration model. Shi et al. used PLS scores for selecting the process samples to include into the calibration set during content uniformity analysis of tablets [137]. This selection strategy was compared with a random selection process. The mean difference between two sampling approaches was found insignificant. PLS score based selection strategy was found to be cost and time effective. Selecting samples from PCA or PLS scores space should be performed carefully, as this can exclude samples with extreme concentrations and narrow the concentration range leading to undesired extrapolation.

Feng et al. used a classification technique to select a calibration set during the determination of roxithromycin and erythromycin in tablets from different manufacturers [27]. The initial calibration set contained samples from random manufacturers. The resulting calibration model failed to predict the concentration of APIs in tablets prepared by manufacturers that were not included into the calibration set. A cluster analysis was performed on average spectra of each batch based on wards algorithm. Samples from different clusters were selected for the calibration set such that, uniform distribution of API content and manufacturer was maintained. Calibration models were demonstrated to be accurate, robust, specific, linear, and transferrable. However, the validity of this method was challenged by Bleye et al. due to inappropriate prediction interval [138].

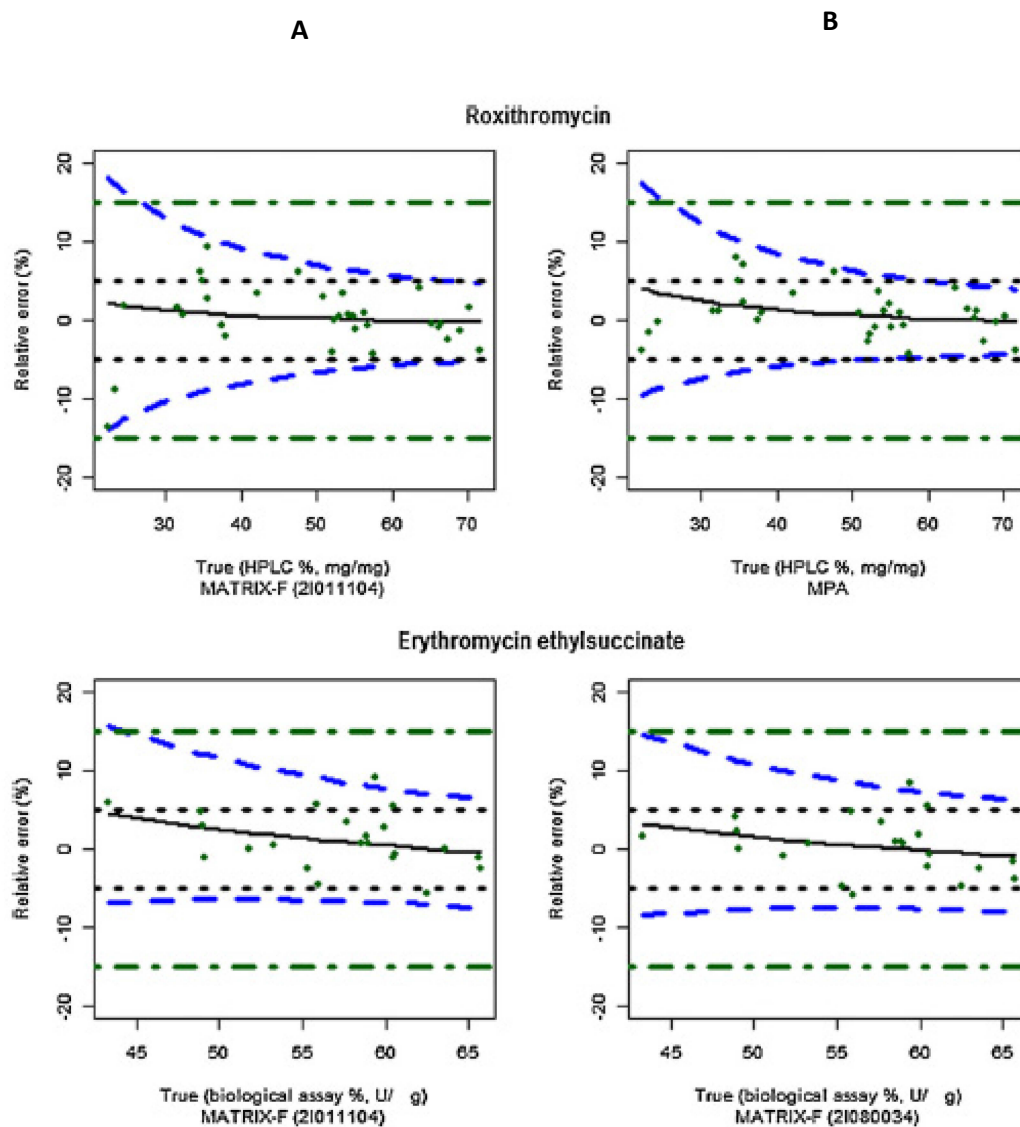


Figure 1-3. Accuracy profiles obtained from the work of Feng and Hu [27] for the quantification of Roxithromycin (A) and Erythromycin ethylsuccinate (B) using different NIRS equipment. Dotted horizontal lines (black): $\pm 5\%$ acceptance limits; dotted and dashed horizontal lines (green): $\pm 15\%$ acceptance limits; dashed lines (blue): 95% prediction intervals; dots (black): relative predicted concentrations of the corresponding NIRS methods [138]

Although having low prediction errors, this method was demonstrated unfit for the intended purpose, because the 95% prediction intervals were outside the acceptable range as shown in Figure 1-3 for two different NIR instruments. The prediction interval should be included in the list of figures of merit for assessing model performance.

1.3.2.3 Under dosing / over dosing (spiking) production samples

Calibration sets can also be prepared by spiking (under dosing / overdosing) production samples. Spiking is performed to extend the concentration range of production samples that are typically prepared at target concentration. Moes et al. used spiking strategy to prepare two calibration sets for analyzing blend uniformity of powder and content uniformity of tablet cores, respectively [139]. A high correlation between the API and excipients concentrations was present in both calibration sets. Along with the high correlation, calibration samples for content uniformity had uneven distribution of API concentration levels and were compressed at different compaction pressures compared to the validation samples. Poor prediction performance was observed during the prediction of tablets compressed at compaction pressures that were not included into the calibration set. Blanco et al. spiked production samples (64.5% amorphous miokamycin) with amorphous/crystalline miokamycin and excipients to prepare a calibration set for determining crystalline and total miokamycin content in tablets [140]. Different levels of PSD were included into the calibration set to make the model robust against particle size variations. Calibration design for determining one polymorph in presence of another polymorph is challenging. High correlation between the concentrations of different polymorphs can diminish calibration model specificity. In this study, the selectivity of the

calibration model was demonstrated qualitatively by using spectral correlation coefficient method.

A similar approach was undertaken to develop a calibration set for quantifying Caffeine in tablets [141]. Due to the presence of high correlation in calibration set, model specificity was assessed by band assignment for Caffeine and comparing spectral features of Caffeine with first loading vector. A set of spiked samples was also used to prepare a calibration set during ranitidine quantification at different steps of tablet manufacturing process; in granules for compression, in tablet cores and in coated tablets [19]. The effect of high correlation between ranitidine and excipients was minimized by selecting appropriate wavelengths. Sarraguca et al. used spiked samples to prepare multiple calibration sets for quantifying Acetaminophen, acetylsalicylic acid, folic acid and neomycin in respective tablet formulation [142]. The correlation coefficient between Net Analyte Signal and pure API spectrum was used as a metric for model specificity. Specificity was found to be low for all formulations due to the high correlation between API and excipients concentrations. A high correlation between API and excipients concentrations is usually found after spiking the production samples. The correlation can be minimized by using multiple placebo mixtures and randomly select one at a time for spiking production samples [135].

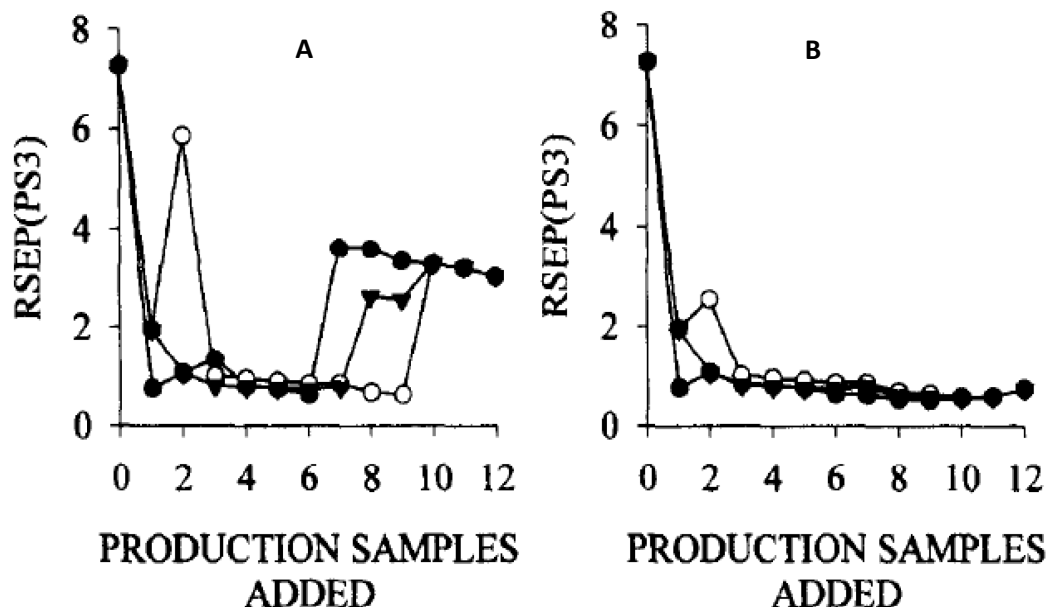


Figure 1-4. Improvement in calibration performance as a function of the number of production sample included into the calibration set with optimized latent variables numbers (A) and fixed latent variable numbers (B). Each symbol represents a different sequence of sample addition [62].

The inclusion of production samples in the calibration set can mitigate the difference between calibration and routine samples in terms of physical properties. The calibration set prepared by spiking the production samples was found to perform better than the calibration set prepared from laboratory sample during quantification of commercial tablets [143]. Spiking was performed with randomly selected placebo samples from a set of placebo mixtures to reduce the correlation between API and excipients concentrations. A set of production samples can also be combined with the laboratory samples into the calibration set and improve calibration model performance [62, 144].

Figure 1-4 shows the improvement in calibration performance as a function of the number of production samples included into the calibration set while predicting otilonium bromide in production tablets [62].

Production samples were also used to develop a calibration set for moisture content determination of tablets [145]. The samples were stored in humidity chamber to vary the moisture content. The samples were also varied in terms of different APIs, excipients, and shapes to develop a global calibration model for moisture content determination. Model developed from tablets of API I failed to predict moisture content of tablets containing different APIs. The most robust model was found by including tablets from all three APIs into the calibration set. However, the effect of uneven distribution of APIs, moisture contents and tablet shapes was not investigated.

A combined strategy of using production and laboratory samples into the calibration set can be helpful, because production samples introduce physical variances whereas laboratory samples extend the calibration concentration range. Calibration sets containing both laboratory and production samples was used for successful quantification of tianeptine in coated tablets [146], Acetaminophen in production tablets [68], hydrocortisone in pharmaceutical powders [147] and isoniazid and rifampicin in capsules [86]. In the last study, production samples from several months were used to introduce possible changes in sample composition, suppliers, process changes or variations in storage conditions. Placebo samples were also used as calibration samples to deduce the interference structure for respective APIs.

A similar preparation method between laboratory and production samples can enhance the performance of the calibration model developed from laboratory samples only. During the determination of dexketoprofen in laboratory and production granules [148], model developed from laboratory powder failed due to differences between granules and powders in terms of physical properties. Although inclusion of production granules into the calibration set improved the prediction performance, the most accurate model was developed from laboratory granules that were prepared using similar process of production granules. This was due to the wider concentration range of laboratory granules compared to production granules. Preparation of laboratory samples encompassing similar physical variability of production samples at wider concentration range was found to be an effective strategy for developing calibration set with desired predictive performance.

Cardenas et al. used process spectra instead of production sample spectra to introduce physical variance into the calibration set during cetirizine quantification in different stages of tablet manufacturing process; blending, compaction and coating [149]. The calibration set was initially developed with powder samples for blend prediction and optimized with process spectra at later stages for core and coated tablet prediction. The process spectrum was calculated as the spectral difference between lab scale powder and industrial samples (core and coated tablets) following the idea developed by Blanco et al [41]. Although this approach offers some advantages for reducing the sample size for calibration, some associated assumptions might have been violated during the calibration development. The addition of a process spectrum assumes that the extent of the effect of physical variation on the spectra remains consistent for all concentration ranges. It also

assumes that the physical effect and chemical effect on NIR spectra are orthogonal. These assumptions are not often true as the physical properties of the samples can vary with compositional variation. The process spectrum approach was compared with conventional approaches for preparing calibration set with laboratory, production and spiked samples [15]. Although similar model performance was achieved, the new strategy was favored for its time and cost effectiveness and simplicity.

Blanco et al. compared different strategies for developing calibration set while predicting nimesulide content in pharmaceutical granulate [75]. Laboratory samples, spiked samples and production samples were used for developing the calibration sets. A calibration model built from spiked samples predicted the production samples with acceptable accuracy; however, model from lab samples failed in the prediction. Prediction performances of both models were improved by including production samples into the calibration sets.

The discussion on different strategies for calibration set development is summarized in Table 1-1 in terms of advantages and disadvantages of each strategy.

Strategies for developing NIR calibration set	Advantages	Disadvantages
Use of design of experiment	<ul style="list-style-type: none"> • Generates systematic variance into the calibration set. • Can be used to minimize the effect of chance correlation. • Can be optimized to meet regulatory requirements in terms of calibration range, levels. 	<ul style="list-style-type: none"> • Usually requires large calibration set. • Can introduce greater than nominal variance leading to reduced accuracy. • Usually employed with laboratory samples which leads to poor robustness against scale variation.
Use of sample selection techniques	<ul style="list-style-type: none"> • Can be employed with any existing dataset. No prior requirements for designing data generation. • Constraints can be applied to meet regulatory requirements. 	<ul style="list-style-type: none"> • Model performance is sensitive to the selection techniques. It can be difficult to select the appropriate combination of selection technique and constraints. • Appropriate samples can be missing in the dataset leading to poor model performance.
Under dosing/over dosing (spiking) production samples	<ul style="list-style-type: none"> • Similar physical properties of calibration samples compared to the routine production samples, which leads to improved model robustness. • Calibration transfer is comparatively easier due to similarity in physical characteristics across scales and manufacturing plants. 	<ul style="list-style-type: none"> • Can introduce chance correlation in the calibration set leading to poor model specificity. • Requires more time and cost to implement in the production floor due to the requirements of sample preparation (production samples and calibration samples must be treated in the same way).

Table 1-1. Advantages and disadvantages of different strategies for NIR calibration set development

1.3.3 Effects of calibration set on model performance

A calibration set has significant effects on NIR model performance. These effects can be discussed from three perspectives as the effects of the calibration design structure, effects of calibration factors and effects of calibration sample properties.

1.3.3.1 Effect of calibration design structure

The structure and parameters of the calibration design can significantly affect NIR model performance. The effect of several parameters including distribution of design points, absence of relevant information and simulation of natural population on NIR model performance were investigated [39]. Although the study did not include pharmaceutical samples, it investigated the general aspects of NIR calibration set, that were applicable for pharmaceutical materials. Even span and all possible combinations of constituents were shown to be necessary for successful model performance. Appropriate experimental design was demonstrated to be a prerequisite for accurate estimation of regression coefficient and adequate predictability. Prediction in the center and corners of the design depended on the sample distribution and overall variance of the design. More samples near the center point gave better prediction performance in the center point compared to evenly spread sample distribution; a finding similar to that of Araujo et al [28]. This trend was opposite while predicting samples outside the center points.

Chance correlation of the calibration design has been pointed out as a limiting factor of multivariate calibration by several researchers [150-154]. Chance correlation can be presented in the calibration set due to inappropriate experimental design, sample properties and the condition underlying the spectroscopic measurement. Xiang et al.

assessed the effect of chance correlation on an NIR model performance [38]. The correlation between API and excipients was set high and low in correlated and randomized calibration set, respectively. The correlated design performed better than the randomized design during calibration development. However, the former design failed during cross-prediction of tablets with excipient and density variations. The randomized design succeeded in cross-prediction indicating that the predictability achieved with the correlated design was dependent on correlated signals not specific to the API. The effect of chance correlation on model specificity is often overlooked due to the traditional practice of univariate model development using HPLC where specificity is achieved by the experimental technique itself.

The interference structure of a calibration set was demonstrated to be critical during NIR calibration development [46]. A poor prediction performance was observed in the validation set containing different interference structure. *A priori* information about the component's pure spectrum and/or interference structure helped to improve the calibration model performance. Li et al. compared the performance between two calibration sets having different levels of API but same levels of excipients [58]. Model performances were similar between the two calibration sets. Quantifications of the minor excipients were more accurate compared to the major excipients and API. It was concluded that the variation of API and major excipients were used during the quantification of the minor excipients leading to improved accuracy. During the quantification of API and major excipients, variability of the minor excipients was not usable due to low concentration leading to biased prediction. However, prediction using other component's concentration profile might have led to poor model specificity. A

modification of the same scheme was used to predict Acetaminophen content in powder blends [155]. Following the previous scheme a set of 625 design points was generated and 50 most distant points in the PCA scores were selected for the calibration set.

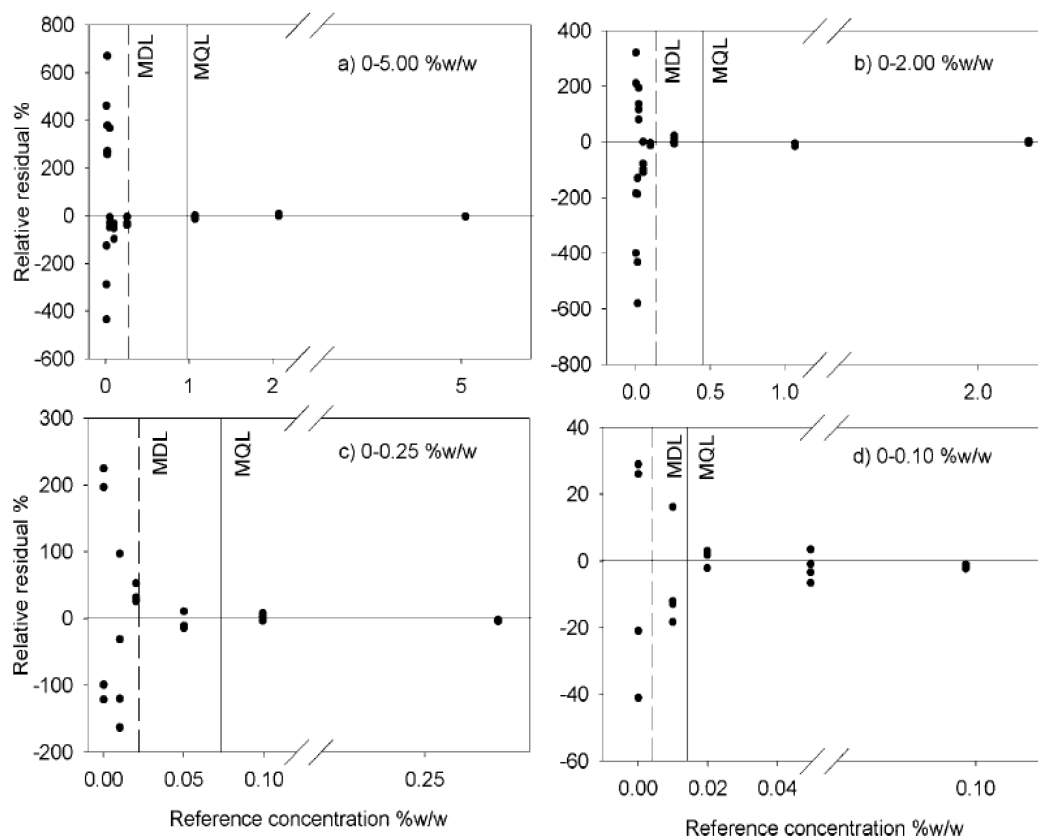


Figure 1-5. Relative residuals (%) plots obtained from the prediction of four external sample sets, using also four calibration models with different concentration ranges (% w/w, API). Vertical lines indicate the MDL/MQL for each model. Four predictions are shown for calibration ranges: (a) 0–5.00%, (b) 0–2.00%, (c) 0–0.25%, and (d) 0–0.10% (w/w) [156]

The concentration range of the calibration set can be critical to model’s predictive performance in terms of multivariate detection limit (MDL) and quantification limit (MQL) and should be selected based on the prediction set property [156]. The relative prediction errors, MDLs and MQLs of four calibration models with different concentration ranges are given in Figure 1-5.

A narrow range of concentrations should be used for predicting low drug concentration and the range should be updated according to the prediction range. Although no explanation was provided, mathematically it can be explained from the equation used for calculating MDL and MQL. As the concentration range decreased, calibration error (SEC) decreased, which in turn decreased the MDL and MQL. Concentration range can also affect the calibration structure. In a calibration set for Theophylline prediction, Theophylline and Magnesium stearate were varied orthogonally and cellactose was considered as filler [157]. However, variation in cellactose concentration had high correlation (0.99) with variation in Theophylline concentration due to its larger range compared to Magnesium stearate concentration. This undesired chance correlation might have led to poor model specificity. Uniform distribution of concentration levels in the calibration set is usually encouraged [50, 51]. However, a similar predictive performance was observed between similar and random distribution of concentration levels [55].

Bondi et al. analyzed the effect of experimental design on the predictive performance of NIR calibration model [9]. The five level full factorial and I-optimal models were statistically similar and had the lowest errors, followed by the CCD, D-optimal and three level full factorial models, respectively. The model generated from the I-optimal design was superior considering the balance between performance and efficiency due to low sample requirement. A more comprehensive assessment was reported by Scheibelhofer et al [49]. Four types of experimental design each with two subtypes were used to develop NIR calibration sets. Model performances in increasing order of prediction error were: five level factorial < extended CCC < simplex centroid < simplex lattice < CCF < I-Optimal with center < three level factorial < I-optimal < D-optimal. A calibration set with

a regular spread between different concentration levels was preferred. It was also concluded that, the similarity between calibration and test structure improves prediction performance. However, as crude pre-processing techniques were used for all designs and the same model was used for all three components, the results might not be the optimum. Araujo et al. also analyzed the effect of experimental design on calibration model performance as shown in Figure 1-6 and found that number of unique experimental run is more important than the number of replicates, so that the calibration can span a wide range of variance [28]. Lorber et al. assessed these effects in terms of a vector “hun”, where “hun” is dependent on the orthogonality in the calibration space. In a poorly designed calibration set, the reduced orthogonality decreases the net concentration signal (orthogonal concentration), which in turns increases “hun” and prediction error [48].

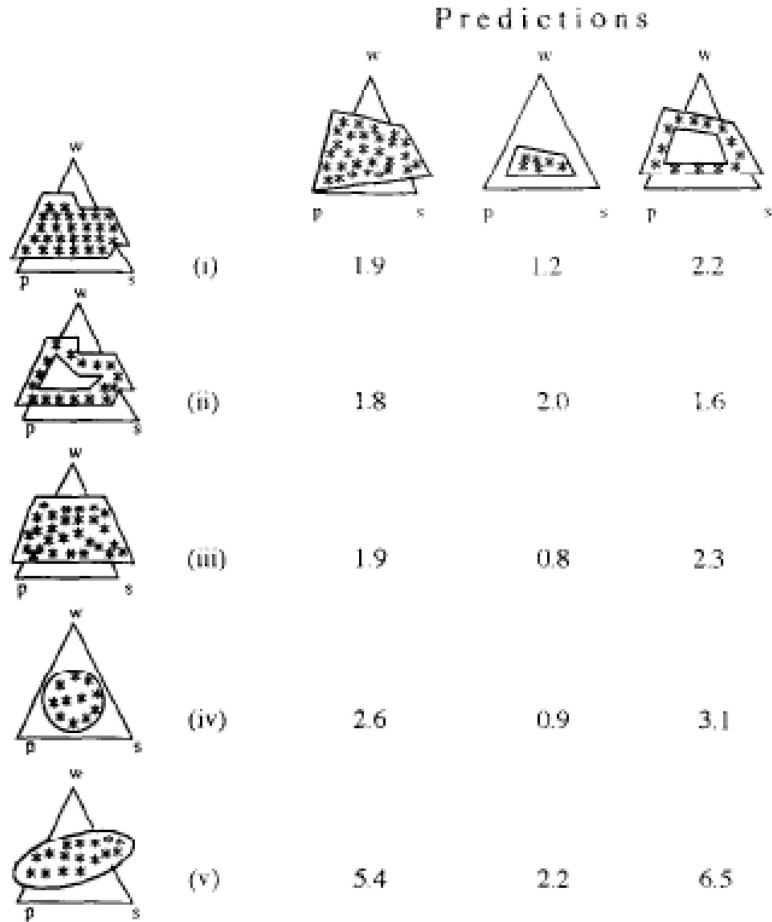


Figure 1-6. Prediction errors of five calibration designs (first column) predicting three prediction sets (first row). The asterisks indicate experimental or predicted areas [28]

1.3.3.2 Effect of calibration samples properties

Different properties of calibration samples can significantly affect NIR calibration model performance. The effect of spectral collection state and scale of calibration set samples on calibration model performance was assessed during blend monitoring [158]. A calibration set was initially developed by collecting spectra at static state. Model performance was improved by incorporating dynamic spectra into the calibration set. Dynamic data from same scale as prediction set was found to improve model performance in a greater extent

compared to the dynamic data from different scale. Sarraguca et al. demonstrated scale and physical characteristics of the calibration set as critical to NIR quantitative method [42]. Calibration set developed from the lab scale powder failed to predict API concentration in the pilot scale powder and tablet due to the scale and physical state difference. Variation in scales and physical attributes were also found to affect calibration model performance during the determination of mirtazapine [135] and ibuprofen [91, 131]. A calibration set developed on one set of tablet thickness and shape, was found to be spectrally different from other set of tablet thickness and shape for reflectance and transmission NIR as shown in score plots in Figure 1-7. This effect was minimized by extending the calibration design to include tablets of different thicknesses, shapes and embossing. Although inclusion of physical variations into a calibration set increases model robustness, it can also affect model accuracy. Model accuracy was found to be decreased upon the inclusion of tablet thickness variations into a calibration set [26]. Accuracy of a model developed on pilot scale samples was affected upon the inclusion of laboratory and production samples of different dosage units into the calibration set [159]. These findings demonstrate that, the current philosophy of including as much variation as possible into the calibration set, can be detrimental to calibration performance. These effects were minimized by bracketing where a local calibration set near the mean tablet thickness [26] and calibration set specific to dosage units [159] were developed. Absence of variation due to size, shape, thickness, as well as narrow concentration range in the local calibration sets helped to improve prediction performances of the respective models. Similar to bracketing, selection of an optimal subset from large calibration set was

proposed by Lorber et al [48]. However, bracketing or subset selection techniques were criticized for causing narrow concentration range and poor robustness of calibration [28].

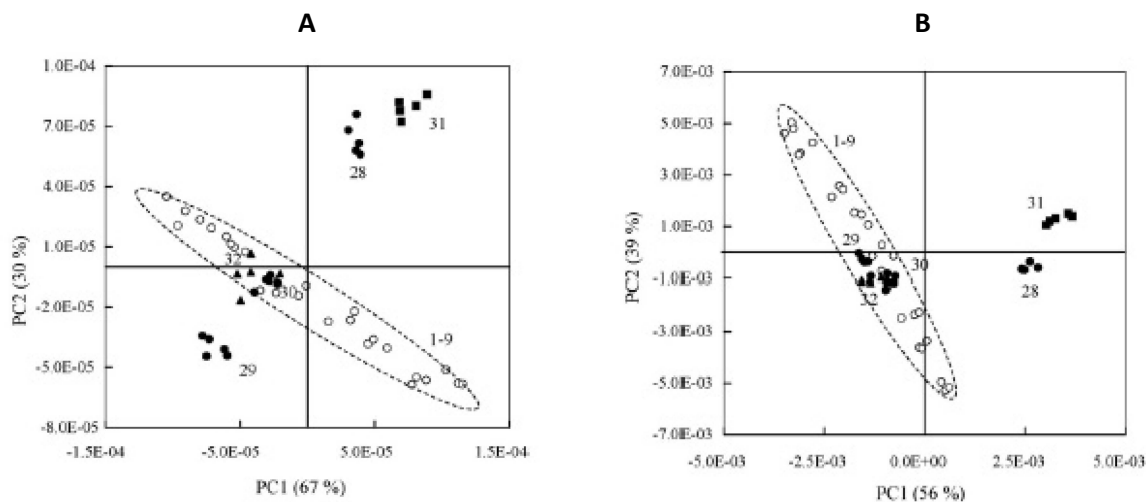


Figure 1-7. Score plot from second derivative reflectance (A) and transmittance (B) spectra by principal component analysis. Open symbols—calibration set: (○) flat-faced tablets with various content; closed symbols—prediction set: (●) flat-faced tablets with various thickness, (△) flat beveled edge tablet, (□) convex tablet. The numbers by the data points correspond to the batch no [160]

Granules PSD and hardness of calibration tablets were found critical and recommended to be included into the NIR calibration set [25]. However, effects of these factors were not analyzed independently. In another study, the effect of PSD was found to be low compared to that of hardness on calibration model performance [38]. Pan et al. independently assessed the effect of variations in tablet hardness, granules PSD and relative humidity on calibration model performance [161]. NIR predictions were found to change up to 10% due to 16% variation in tablet hardness between calibration and predictions sets. For a 20% increase in RH, the change was around 10-20%. No correlation between PSD of granules and NIR predictions were found within the range investigated. However, in another study, PSD of calibration samples was found to be

critical to model performance due to its ability to cause variation in tablet hardness [162]. Reflectance NIR was found to be more sensitive to tablet hardness and packing density compared to transmission NIR [69].

Pieters et al. assessed the effect of intra and inter batch variability of a calibration set on calibration model performance [163]. The presence of these variations in the calibration set helped to improve model performance. The performance was further improved by augmenting the calibration set with noise matrix calculated from net analyte preprocessing. The prediction performance for samples having moisture variation was improved by updating the calibration set with samples having different moisture contents. Mainali et al. assessed the effect of calibration set complexity during the prediction of water content in tablets. The complexity was introduced by using multiple API and tablet shapes into the calibration sets [145]. Slight increase in prediction error was observed with increased complexity. The most complex calibration set (III) was preferred over others due to its increased robustness against API and tablet shape variation. Effect of nonhomogeneous distribution of different APIs and tablet shapes in the calibration set II and III were not discussed.

1.3.3.3 Effect of calibration factors

Although it is recommended to include all factors into the calibration set for developing a robust calibration model [51], this approach requires large sample set, which can be inconvenient on particular instances [164]. Factors that do not affect NIR spectra should be excluded to reduce the size of the calibration set. Factors such as processing parameters, temperature, moisture content, raw material variations etc. can have a

detrimental effect on NIR prediction performance. Tablet moisture content can vary during different steps of the manufacturing process, storage and analysis [165] and excipients variation can occur between batches [163]. All these relevant factors should be included into the calibration set. A small prediction error can be obtained by including relevant information in the calibration set [66]. However, it may not always be possible for the calibration set to contain all relevant factors that can occur in the industrial scales [166], making the models less robust and requiring frequent updates. Inclusion of non-relative information into the calibration set can increase model complexity and affect model performance [163].

Roggo et al. demonstrated process changes as the critical factors for calibration set during dissolution study [167]. Four types of process related changes were included into the calibration. The changing parameters were time and temperature changes during melt granulation, compaction pressure, coating formulation and coating times. Riley et al. assessed the effect of number of varying components in the calibration set on calibration model performance [168]. An increase in the number of varying components from 2 to 6 and 2 to 10, increased the error by approximately 50% and 340%, respectively. However, a calibration set containing higher number of varying components could be useful for predicting complex samples. El-Hagrasy et al. assessed the effect of environmental factor such as blend storage time on calibration model performance [169]. The NIR spectra were collected immediately after blend preparation and after storing for a week to develop two calibration sets, respectively. The predicted blend profiles were significantly different between two sets. As the environmental factor was found to be critical for developing calibration set for blend monitoring, a D-optimal design was used to vary

relative humidity along with concentrations to develop a calibration set in later studies [170, 171]. The effects of NIR calibration set on model performance is summarized in terms of affected validation characteristics and analytical figures of merit in Table 1-2.

Calibration set properties	Affected calibration metrics
Calibration design and structure	Accuracy, specificity, range, linearity, robustness, detection limit and quantitation limit.
Calibration samples properties	Accuracy, specificity, robustness and linearity.
Calibration factors	Accuracy, specificity, robustness and linearity.

Table 1-2. Advantages and disadvantages of different strategies for NIR calibration set development

1.3.4 Conclusion

There are several strategies for developing NIR calibration sets for tablet analysis. The selection of right strategy is critical to model performance, however, less emphasized during calibration development. An in-depth understanding about the calibration set requirements, interaction between calibration structure and model performance, and a formulation and NIRS specific outlines can be helpful for developing an appropriate NIR calibration set. Appropriate calibration set is a prerequisite for successful model performance and must be designed carefully at the outset of calibration development to ensure efficient and maximum utilization of NIRS capability in pharmaceutical analysis.

2 Chapter 2: Method development of designing calibration set in spectral space

In this study, a method was developed to use the pure component spectra of a tablet to identify the critical samples necessary for an efficient calibration set development. The principal difference between current techniques and the proposed technique lies in the space used for designing a calibration set. Calibration sets are typically designed in concentration space where orthogonality between component concentrations is imposed to minimize collinearity and maximize model specificity. In the proposed method, a calibration set is designed in spectral space in order to create orthogonality of the spectral response.

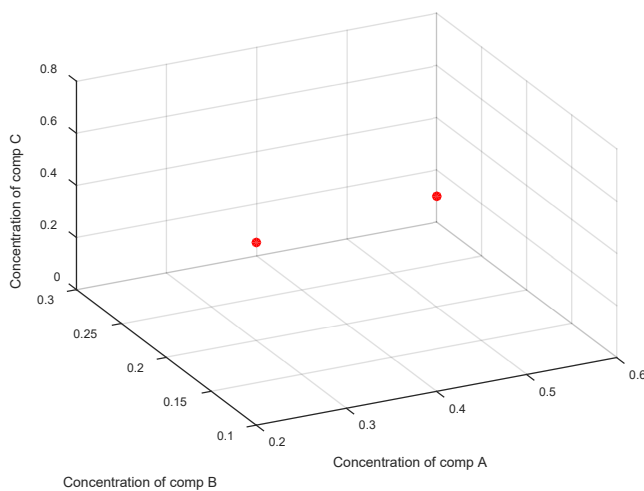


Figure 2-1. Concentration space of a ternary system

2.1 Theory

Concentration space: Concentration space is a multi-dimensional space used to describe the composition of a composite structure. The structure of the composite system is described by relative positions and directions along each of the dimensions in

concentration space coordinates. For example, Figure 2-1 illustrates a concentration space describing two composite structures of a ternary system containing A, B and C components at [0.2, 0.1, 0.7] and [0.6, 0.3, 0.1] concentration levels.

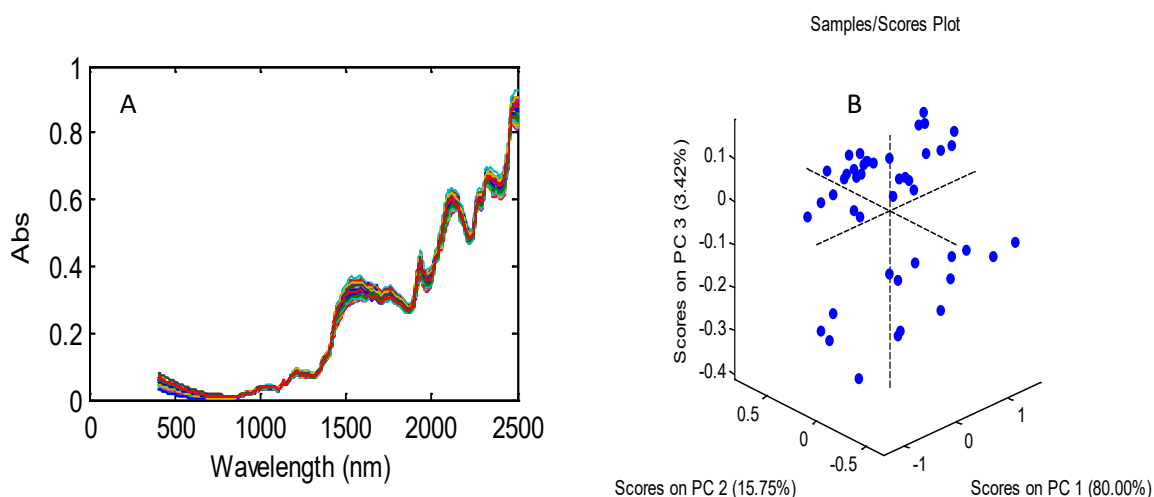


Figure 2-2. Spectral data (direct spectral space (A)) presented in PCA score space (factor space (B))

Spectral space: Spectral space is a multi-dimensional space to describe the spectral response of a system. The spectral response at each wavelength/frequency is described along each dimension of the spectral space coordinates. This space has higher order dimensions, typically on the order of 100 to 10,000 depending on the analytical technique of interest. Therefore, spectral space is often decomposed to a lower dimensional orthogonal subspace using chemometric techniques such as orthonormalization, principal component analysis (PCA), singular valued decomposition, eigen value decomposition etc. Principal component analysis efficiently describes the spectral variance, revealing the maximum variance in the first principal component (PC) and remaining orthogonal variance in each succeeding PC in descending order [172]. The scores (projection of each

spectrum) on the PCs are used to represent the high-dimensional spectral space into a much lower dimensional score space. For example, a typical series of 45 NIR spectra (4200 data points in each spectrum) representing 45 Acetaminophen tablets can be represented in three dimensional score space instead of 4200 dimensional spectral space without losing any significant spectral information as illustrated in

Figure 2-2 (99.18 % of total spectral variance is explained by 3 PCs).

The use of pure component spectra during calibration set development follows the fundamental theory of spectroscopic multivariate calibration and experimental design. A spectroscopic calibration model is developed to predict the chemical/physical properties of a sample from its respective spectrum using regression equation. For a system with “m” samples and “n” spectral variables (wavelengths), the regression equation is as follows,

$$\hat{\mathbf{y}}_{m \times 1} = \mathbf{X}_{m \times n} * \mathbf{b}_{n \times 1} \quad (2.1)$$

Where $\hat{\mathbf{y}}$ is the predicted chemical/physical properties, \mathbf{X} is the spectral data matrix and \mathbf{b} is the regression vector. Concentration will be considered as an example of physico-chemical property in rest of the chapter. The regression equation (2.1) can be expanded as follows,

$$\hat{y}_{i=1..m}^i = b_0 + x_1^i \times b_1 + x_2^i \times b_2 + \dots + x_n^i \times b_n \quad (2.2)$$

To approximate the relationship between concentrations (y) and spectral variables (x) by estimating the regression coefficients b_i , the theory of design of experiments indicates that the spectral responses ($\mathbf{x}_{1..n}$) should vary orthogonally and the output concentrations ($\mathbf{y}_{1..m}$) measured. Orthogonal variation in the independent variable (spectral response)

would allow better estimation of regression coefficients with minimum number of samples/ experiments.

Currently, there is no published technique available for orthogonally varying the spectral responses. The conventional approach to multivariate calibration is to vary concentrations orthogonally in the concentration space and assume that this orthogonality translates to the spectral space. For a system with “k” components, “m” samples and “n” spectral variables using **C** for the concentration of samples, **P** as the pure component spectra, and **X** for the matrix of spectral responses, the current design follows the equation below,

$$\mathbf{C}_{m \times k} * \mathbf{P}_{k \times n} \rightarrow \mathbf{X}_{m \times n} \quad (2.3)$$

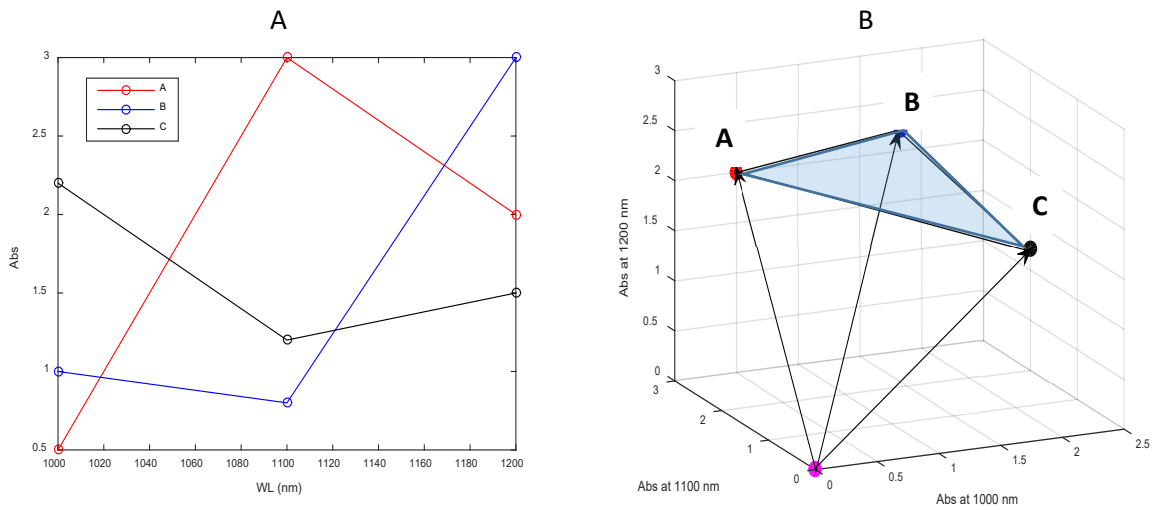


Figure 2-3. Hypothetical pure component spectra (A) and respective spectral space (B)

This does not necessarily create orthogonal variance in the spectral space due to external factors other than concentrations affecting the spectral signals. Moreover, depending on the pure component spectra and their relative contribution to the sample spectrum, experimental design points created in a concentration space can result in highly redundant

and/or inadequate spectral information. A hypothetical scenario is illustrated for a three component system (A, B and C) analyzing with a spectroscopic technique of 3 wavelength channels (1000 nm, 1100 nm and 1200 nm).

Figure 2-3 represents the hypothetical pure component spectra of A, B and C and their respective projections on the spectral space. The spectrum of any sample mixture containing A, B and C should fall in the triangular plane (shaded region) formed by A, B and C in the spectral space due to mixture constraint (sum of A, B and C is always 1).

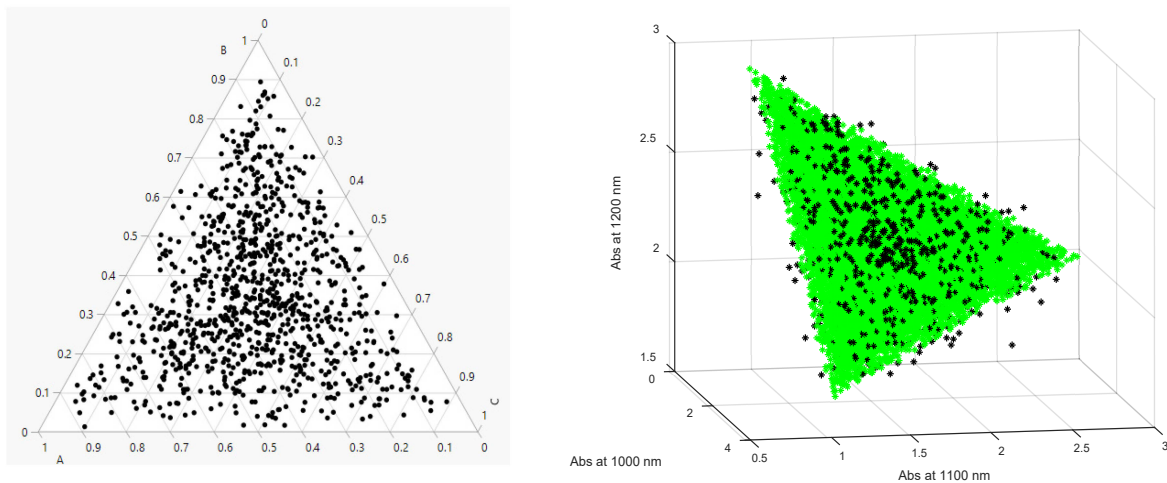


Figure 2-4. Hypothetical concentration space (left) and respective spectral space (right)

Experimental design points created in the concentration space should retain its structure and linearly mapped in this hypothetical plane. However, this does not occur experimentally due to non-linearity between concentration and spectral response,

different spectral contribution of each component into the mixture spectrum, physico-chemical properties of the samples etc.

Figure 2-4 shows a hypothetical scenario where experimental design points in the concentration space (Figure 2-4 (left)) fall outside the hypothetical plane (green) in the spectral space (Figure 2-4 (right)). Thus, an orthogonal design in the concentration space might lose its orthogonality in the spectral space, provide redundant information as well as lack in critical information in the calibration design.

The proposed approach is an attempt to create the orthogonal variations in the spectral space. Orthogonal variation in the spectral space helps to minimize the sample requirement by identifying a set of critical samples necessary to develop a successful calibration model. For instance, compositional variation in the spectrally similar excipients will have unidirectional variation in the orthogonal spectral space. Designing calibration samples in that direction by varying either of them would give sufficient spectral variation to model against the concentration of API and reduce the calibration sample size by a factor of two. Moreover, orthogonal spectral space helps to identify the most spectrally dissimilar samples to be included into the calibration set to cover the maximum possible spectral variation. Training a model against such variation is critical to develop a robust calibration model at the outset of calibration development. The most spectrally dissimilar samples are often different from the most compositionally dissimilar samples. Building calibration set with compositionally different samples and training a model against large compositional variation may not always be successful in developing a calibration model that is robust against future spectral variation.

However, in NIR spectroscopy, spectral responses at each wavelength cannot be varied orthogonally due to highly correlated signals. Alternatively, the correlated spectral responses can be decomposed into orthogonal directions using different techniques such as Orthonormalization, Principal Component Analysis (PCA), Partial Least Square (PLS), Singular Valued Decomposition (SVD) etc. In the proposed method, the pure component spectra of a formulation are used to determine the spectral space for that formulation and Orthonormalization technique is used to decompose the pure component spectral response into orthogonal directions. The spectral projections onto the orthogonal directions are termed as “scores”. The score is used as a metric of spectral variance in the orthogonal spectral space of the formulation. Designing experiments by varying the scores is considered equivalent to the design in orthogonal spectral space. In this study, pure component spectra were acquired by compressing the pure components into tablets and scanning with NIR instrument. When the pure component tablets are difficult to prepare, pure component powders can be used (adjustments for spectral differences may be necessary).

After defining the orthogonal spectral space of a formulation using pure component spectra, a model tablet is compressed at the target formulation and an NIR spectrum is acquired. The spectrum of the model tablet is projected onto the orthonormal basis vectors of the pure component tablet spectra. The projections are termed as “model tablet scores.” Scores of the ‘model tablet’ are calculated based on the following equation

$$\mathbf{t}_{1 \times r}^{\text{Model tablet}} = \mathbf{X}_{1 \times n}^{\text{Model tablet}} * \mathbf{W}_{n \times r}^{\text{Pure component basis}} \quad (2.4)$$

Where, \mathbf{t} is the scores of the “model tablet” for “ r ” factors, \mathbf{x} is the spectrum of the “model tablet” and \mathbf{W} is the orthonormal basis of the pure component spectra.

The idea behind the model tablet score is to design a calibration set such that the scores of the calibration samples are centered around the model tablet scores. This ensures the calibration model to be centered at the target formulation. An in silico full factorial design is created to achieve this objective. In the design, the numbers of factors are kept equal to the number of chemical components. We assume that the signal from each component contains orthogonal features with respect to other components. Thus, full chemical rank is achieved by creating a design where the number of factors is equal to the number of components. The numbers of levels in the design are selected to balance between computational burden and simulation of spectral variance. Simulation of spectral variance is required to select the calibration candidates from a large dataset. Addition of one level can significantly increase the design points and subsequent simulation time. For instances, five levels for each factor in a five-factor design will result in $(5^5) = 3125$ design points. An increase by one in the number of levels will result $(6^5) = 7776$ design points. The effect of number of levels is minimal as long as it is kept equal for all the factors to prevent any confounding effect and generate all the possible spectral variation during simulation.

The center of the in-silico design is then set equivalent to the model tablet scores \mathbf{t} and all other design points are scaled from the center. The transformed design is defined as the ‘target score design’. Centering the design around model tablet scores ensures that the target scores span the space around the model drug. The proposed approach centers the regression model at the target composition (“model tablet”) and then creates balanced

spectral variance around this center. The spectra corresponding these score values vary orthogonally (or very nearly so), because the target score design is built from a full factorial orthogonal structure. Figure 2-5 shows the procedure for designing scores around the model tablet scores. Use of scores as the calibration design factors can also be justified from chemometric point view. A mathematical description is given in Appendix A.

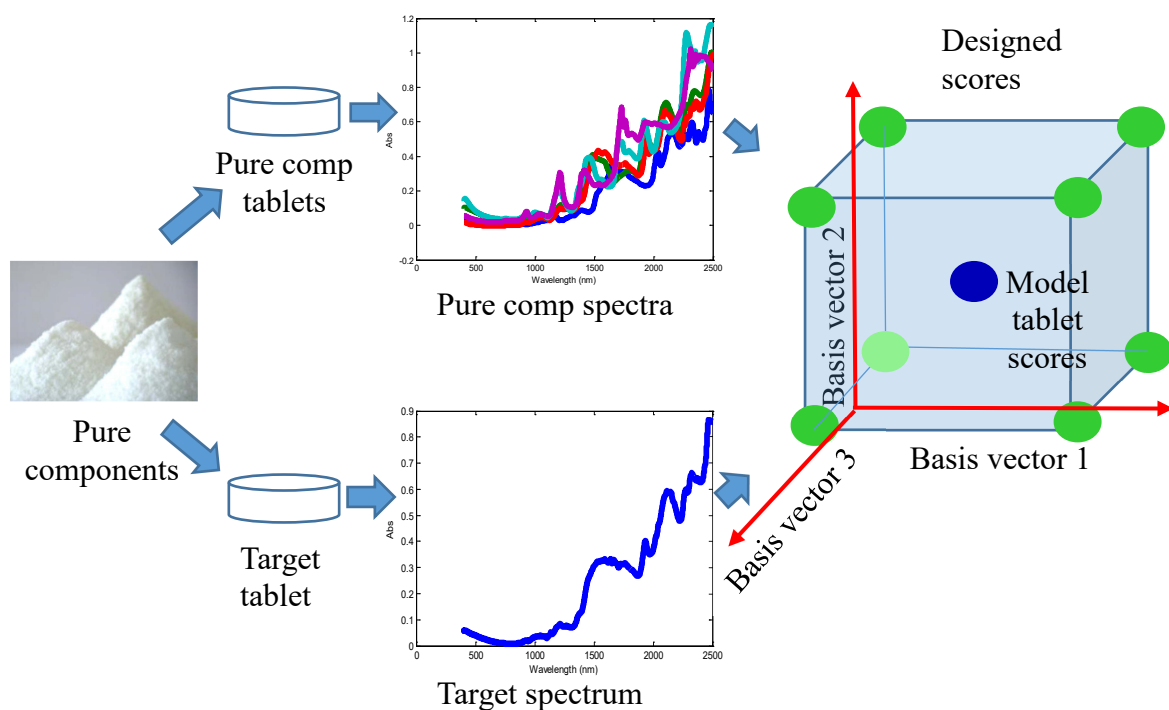


Figure 2-5. Procedure for designing scores around the model tablet scores in spectral space

However, the challenge remains to find the compositional requirements for the sample from which the spectrum is collected and scores are derived. Although the calibration samples are designed in the spectral space, the design must be carried out in the concentration space to prepare those samples.

$$\mathbf{C}_{m \times k}^{\text{Composition}} * \mathbf{P}_{k \times n}^{\text{Pure component spectra}} = \mathbf{X}_{m \times n}^{\text{Sample spectra}} \quad (2.5)$$

The relationship between the compositional points and the spectra can be expressed by equation 2.5. However, this relationship is not always true due to different physical and chemical contributions to the spectral features in the NIR region. A rotation matrix “**R**” is introduced in the above expression to explain all other contributions on spectral features except concentration. The rotation matrix rotates the estimated spectra towards the actual spectra. “**R**” is calculated based on equation 2.7 (the detail derivation is given in appendix B). Where, **U** is the left singular values, **S** is the singular values and **V** is the right singular values for the concentration matrix, **X** is the spectral matrix containing the pure component tablets and model tablet spectra and **P** is the pure component tablets spectra.

$$\mathbf{C}_{m \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} * \mathbf{P}_{k \times n}^{\text{Pure component spectra}} = \mathbf{X}_{m \times n}^{\text{Sample spectra}} \quad (2.6)$$

R_{k×k}^{Rotation matrix}

$$\begin{aligned} &= \mathbf{V}_{k \times k}^{\text{Right singular}'} * \left\{ \left(\mathbf{S}_{k \times m}^{\text{Singular}} \right)' * \left(\mathbf{S}_{m \times k}^{\text{Singular}} \right) * \right\}^{-1} \\ &* \left(\mathbf{S}_{k \times m}^{\text{Singular}} \right)' * \left(\mathbf{U}_{m \times m}^{\text{Left singular}} \right)' * \mathbf{X}_{m \times n}^{\text{Sample spectra}} \\ &* \left(\mathbf{P}_{n \times k}^{\text{Pure component}} \right)' \\ &* \left\{ \left(\mathbf{P}_{k \times n}^{\text{Pure component}} \right) * \left(\mathbf{P}_{n \times k}^{\text{Pure component}} \right)' \right\}^{-1} \end{aligned} \quad (2.7)$$

The effect of rotation matrix in simulating spectral response is demonstrated in Figure 2-6.

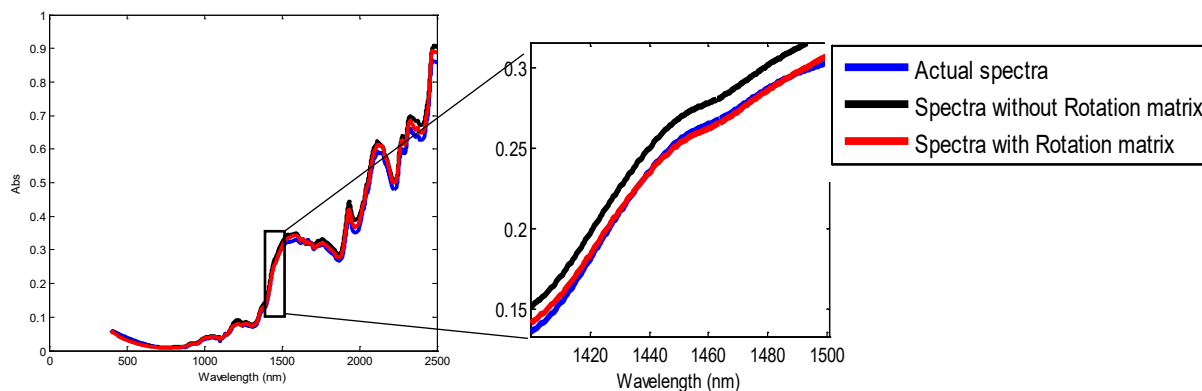


Figure 2-6. Effect of rotation matrix in spectral simulation

Equation 2.8 is solved for “**C**” to calculate the compositional requirements using a global optimization technique, assuming that the variation in pure component spectra dictates the primary spectral variation in the model, (the detail derivation is given in appendix C). Where, **C** is the compositional matrix for the targeted score design having sample size “**u**”, **R** is the rotational matrix, **T** is the target design in the score space, **P** is the pure component spectra, **W** is the basis vector set and $r = k$ (number of latent variables = number of components). Two constraints are applied during the solution by the optimization technique. The first one is, for all “**u**”, $\sum_{i=1}^k C_i = 1$. This is a general constraint for formulation, since the sum of the concentrations of all the components in each design point must always be 1. The second constraint is set to each constituent individually to meet the appropriate concentration range criteria for a feasible tablet preparation and regulatory guidelines. For instances, API concentration ranges should vary at a range around 70-130 % w/w of the label claim and the lubricant concentration

must be limited to a very small range of concentration. Each row of the resultant matrix “C” provides the compositional requirements (concentrations) of a sample that has been designed in the orthogonal spectral space. This approach is referred to as spectral design throughout the rest of this article.

$$\begin{aligned}
 \mathbf{C}_{u \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} \\
 = \mathbf{T}_{u \times r}^{\text{Score space}} \\
 * (\mathbf{P}_{k \times n}^{\text{Pure component spectra}} * \mathbf{W}_{n \times k}^{\text{Pure component basis}})^{-1}
 \end{aligned} \tag{2.8}$$

With the constraint: for all “u”, $\sum_{i=1}^k C_i = 1$

After determining the compositional requirements of the design points, a small representative subset is selected from the design points as the calibration candidates. Different selection strategies can be employed to select the calibration samples such as the Kennard stone algorithm, orthogonal rotation, maximum distance in the score space etc.

2.2 Spectral calibration set for single API formulation

The spectral space strategy was used to design a NIR calibration set for quantitative analysis of a model drug containing Acetaminophen as Active Pharmaceutical Ingredient (API) and Microcrystalline cellulose (MCC), spray dried Lactose, Hypromellose and Magnesium stearate as excipients. The spectral design was created based on the theory discussed in section 2.1.

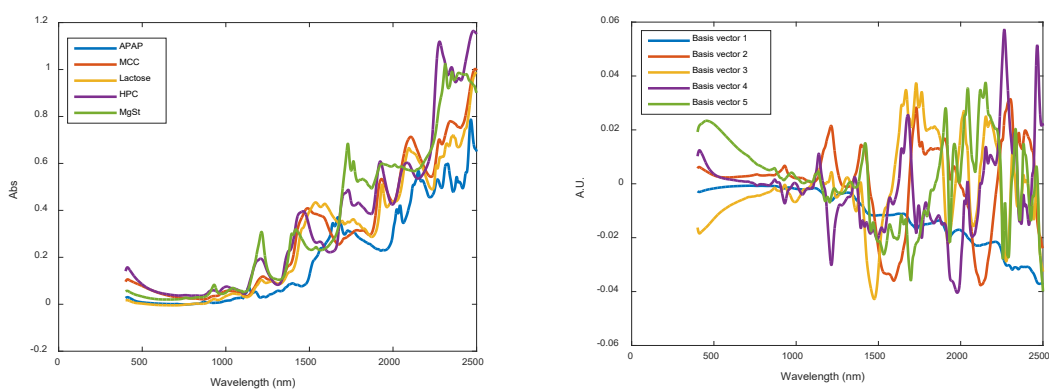


Figure 2-7. Pure component spectra and respective basis vectors

The pure component powders were collected from different sources, Acetaminophen (Mallinckrodt Inc., Raleigh, NC, USA), Hypromellose (HPMC; Pharmacoat 606, Shin-Etsu Chemical Co. LTD, Tokyo, Japan), Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA), Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA) and Magnesium stearate (MgSt; Fisher Scientific, Waltham, MA, USA). The pure component powders were compressed using a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5000 lb force using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg. The target formulation was set as Acetaminophen (27.3% w/w), MCC (34.15% w/w), Lactose (34.15% w/w), HPMC (3.9% w/w) and MgSt (0.5% w/w). A model tablet was

prepared at the target formulation. All these 6 tablets (5 pure component tablets and one model tablet) were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum.

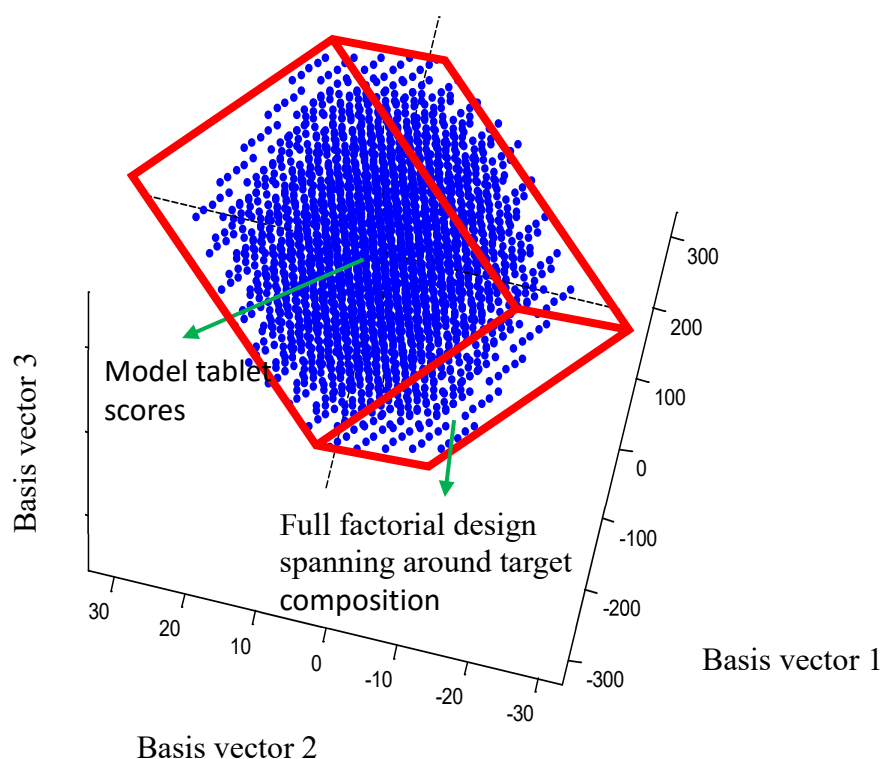


Figure 2-8. A full factorial design around the model tablet score in the orthogonal spectral space

The pure component spectra were orthonormalized to determine the basis vector of the pure components. Figure 2-7 shows the pure component spectra and basis vector. These basis vectors are orthonormal (orthogonal and unit vector) to each other. The basis vectors define the orthogonal spectral space of the pure component. Instead of the

orthonormalization technique, PCA can be done on the unscaled pure component spectra and the principal component of the PCA would give the exact same vector as the basis vectors. The model tablet spectrum was projected on the basis vector set of the pure components. The resulted scores are termed as target scores.

A 5-factor full factorial design with 5 levels per factor was calculated in the orthogonal spectral space where the “model tablet” scores were set as the center of the design as shown in Figure 2-8 for the first 3 basis vectors. This resulted in a 3125 (5^5) design points. A rotation matrix was calculated from the pure component and model tablet spectra. The rotation matrix, pure component spectra and orthonormal basis were used to calculate the required composition for the 3125 design points based on equation (2.2.8). A sample selection technique was employed to select a small representative subset of maximum spectral variance within the full design. NIR spectra were simulated for all the 3125 design points using the concentrations, rotation matrix and pure component spectra following equation (2.6) without adding additional noise. Principal component analysis was performed and a 3-dimensional score space was created using scores on PC1, PC2 and PC3. A set of 3 PCs was able to explain all the variation in the spectra. Delaunay triangulation was used to create a convex hull around the calculated score space. A straight line through the center was rotated orthogonally and the two most distant samples from center along the line were selected to capture maximum possible variance. Figure 2-9 shows the orthogonal rotation of the line in the PCA score space and selected design points.

Design #	Acetaminophen (% w/w)	Hypromellose (% w/w)	Lactose (% w/w)	Microcrystalline cellulose (% w/w)	Magnesium Stearate (% w/w)
1	32.02	4.42	31.43	31.95	0.19
2	20.67	1.00	42.36	35.79	0.19
3	40.13	7.95	21.66	29.53	0.73
4	35.17	0.94	36.06	27.63	0.20
5	30.90	8.05	26.60	33.65	0.80
6	24.30	4.90	31.11	38.97	0.71
7	38.02	2.43	33.15	26.17	0.23
8	21.62	1.01	37.93	39.22	0.21
9	39.85	8.03	26.77	25.13	0.23
10	24.17	1.03	42.04	32.54	0.23
11	27.30	3.90	34.15	34.15	0.50

Table 2-1. Composition of the spectral calibration set for single API

A calibration set composed of 10 design points was selected using this technique. The target composition along with the 10 design points provided 11 compositional points for the spectral design as indicated in Table 2-1. This design was labeled as the ‘spectral design’ and was used to develop a quantitative calibration method.

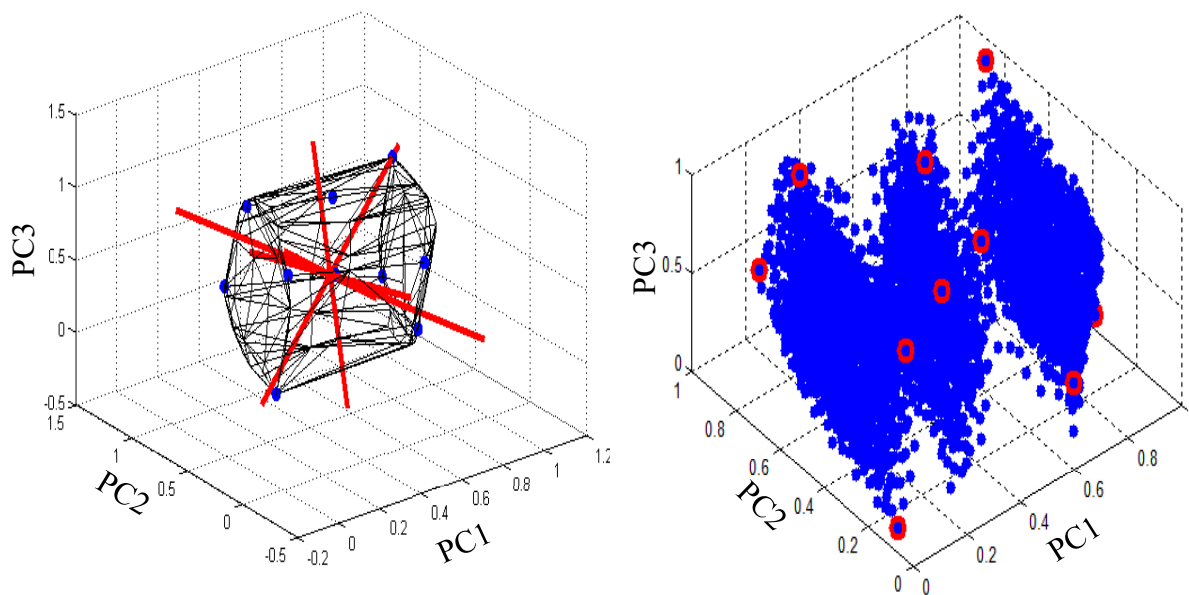


Figure 2-9. Orthogonal rotation in spectral space to select calibration candidates

The spectral design samples were prepared by direct compression. All the components (Acetaminophen, MCC, Lactose, HPMC and MgSt) for each design point were weighed and placed in a 10 ml scintillation vial according to Table 2-1. The ingredients were mixed by rotating the scintillation vial. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5000 lb force using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg. This tablet set was defined as the ‘spectral calibration set’.

All the tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum.

Spectra from both faces of each compact were averaged to produce a single representative spectrum. Spectral calibration set and their respective spectra were used to develop a quantitative NIR calibration model to quantify the amount of Acetaminophen in tablets of interest.

2.3 Spectral calibration set for multiple API formulation

The spectral design strategy was also used to create a calibration set for quantitative analysis of multiple API tablets. The multiple API tablets contained two APIs as Acetaminophen (Acetaminophen; Mallinckrodt Inc., Raleigh, NC, USA) and Caffeine anhydrous (Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA). The excipients were Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA), Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA), Crosscarmellose sodium (Crosscarmellose Na, Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA) and Magnesium stearate (MgSt; Fisher Scientific, Waltham, MA, USA). The target formulation was set as Acetaminophen (31.25% w/w), Caffeine (4.05% w/w), MCC (37.32% w/w), Lactose (24.89% w/w), Crosscarmellose Na (2% w/w) and MgSt (0.5% w/w). All the materials were stored in room temperature and relative humidity. Anhydrous caffeine was reported to be stable under 75% RH for 7 weeks [173]. No hydration and anhydrous caffeine was expected considering lower room RH (~60%) and shorter storage time and analysis.

The spectral design was created based on the theory discussed in section 2.1. Initially, the pure component tablets (six tablets) and one model tablet at the target composition were compressed at 5000 lb using a Carver Automatic Tablet Press. The tablets were stored for two weeks for viscoelastic relaxation based on a previous study on similar formulation [174]. After the viscoelastic relaxation, tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum.

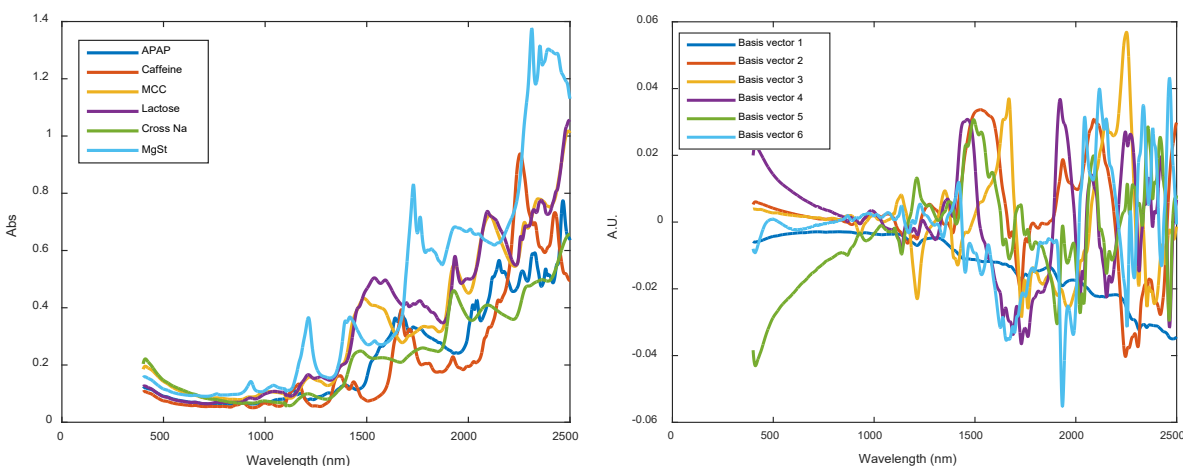


Figure 2-10. Pure component spectra and respective basis vectors

The pure component spectra were orthonormalized to derive the basis vector of the spectral space. Figure 2-10 shows the pure component spectra and basis vector. The model tablet spectrum was projected on the basis vector to calculate model tablet scores on the orthonormal basis of pure component spectra. A 6-factor full factorial design with 5 levels per factor was calculated in the orthogonal spectral space where the “model

tablet” scores were set as the center of the design. This resulted a total of 15625 (5^6) design points.

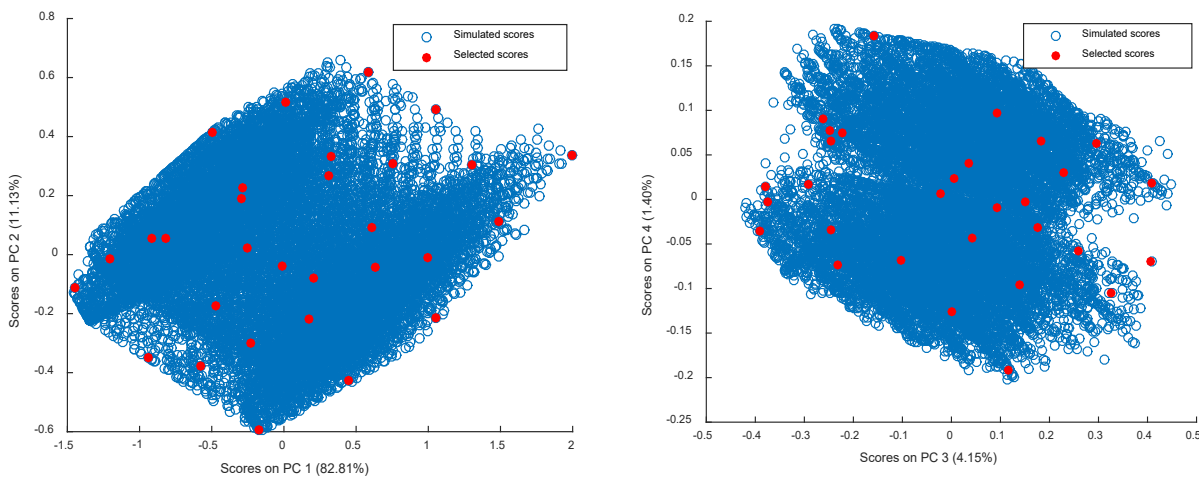


Figure 2-11. PCA scores of the simulated spectra and selected calibration candidates by KNN

A rotation matrix was calculated from the pure component and model tablet spectra to help explaining spectral variance coming from sources other than the spectra of the pure components. The rotation matrix, pure component spectra and orthonormal basis vectors were used to calculate the compositional requirements for the 15625 target design points based on eq. 2.8.

After determining the compositional requirements of the target design, the target design spectra were simulated using the compositional requirements, the pure component spectra and the rotation matrix using equation 2.6. A PCA analysis was performed on the simulated spectra to determine the principal component scores.

Acetaminophen (% w/w)	Caffeine (% w/w)	MCC (% w/w)	Lactose (% w/w)	Crosscarmellose Na (% w/w)	MgSt (% w/w)
34.25	3.79	36.77	21.99	2.40	0.80
19.17	0.81	38.31	37.40	3.50	0.80
43.75	7.29	30.17	15.64	2.95	0.20
30.93	0.81	27.89	38.97	1.20	0.20
27.88	0.81	51.96	15.49	3.08	0.78
39.60	7.29	23.23	28.89	0.80	0.20
24.39	7.29	32.90	31.13	3.50	0.80
36.59	0.81	24.85	35.74	1.81	0.20
34.31	7.29	38.61	15.49	3.50	0.80
24.52	0.81	51.96	21.41	0.50	0.80
40.49	0.81	24.38	30.62	3.50	0.20
28.34	7.29	47.58	15.49	0.50	0.80
35.11	7.29	23.23	33.07	0.50	0.80
32.39	0.81	39.48	24.52	2.60	0.20
43.62	1.95	35.83	15.49	2.32	0.80
24.59	0.81	32.54	37.75	3.50	0.80
41.67	7.29	34.25	15.49	0.50	0.80
28.03	2.57	41.25	27.45	0.50	0.20
41.42	0.81	41.58	15.49	0.50	0.20
29.53	7.29	27.82	33.43	1.73	0.20
29.99	7.29	42.93	15.49	3.50	0.80
33.22	7.29	29.42	26.37	3.50	0.20
31.08	2.32	46.37	15.93	3.50	0.80
26.15	5.46	47.05	19.03	1.51	0.80
43.75	7.29	24.29	23.96	0.51	0.20
36.55	0.81	31.15	27.79	3.50	0.20
34.56	0.81	30.85	32.48	0.50	0.80
41.38	3.23	31.13	23.56	0.50	0.20
35.95	0.81	43.29	16.25	3.50	0.20
21.71	4.18	35.95	33.86	3.50	0.80

Table 2-2. Composition of the spectral calibration set for multiple API

The Kennard Stone algorithm was used to select 30 design points in the PCA score space of the target design spectra (Spectral design). Kennard stone algorithm is a widely-used technique for selecting a representative subset from a large dataset. The selection algorithm provides uniform coverage over the dataset by selecting the most separated samples at successive iteration. The details of the Kennard stone algorithm can be found elsewhere [132]. Figure 2-11 shows the simulated scores of the 15625 points and selected scores of 30 points by Kennard stone algorithm. The composition of the selected 30 points is given in Table 2-2. The target formulation was compressed at 3 different compaction forces resulting in total 33 compositional points for the spectral design.

The 33 tablets of the spectral design were prepared by direct compression on a Carver Automatic Tablet Press using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg. The compression forces range was 4000 lb to 6000 lb. The compression force for each design point was selected such that the correlation between compression force variation and concentration variation of each component is minimal and variance in compaction force is maximal.

Table 2-3 provides the correlation co-efficient between compression force variation and concentration variation. This tablet set was defined as the ‘spectral calibration set’. All the tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum.

	Acetaminophen	Caffeine	MCC	Lactose	Cross carmellose Na	MgSt	Force
Acetaminophen	1.00	0.13	-0.46	-0.35	-0.27	-0.48	0.00
Caffeine	0.13	1.00	-0.23	-0.22	-0.13	0.14	0.02
MCC	-0.46	-0.23	1.00	-0.60	0.10	0.45	0.00
Lactose	-0.35	-0.22	-0.60	1.00	0.01	-0.16	-0.01
Cross carmellose Na	-0.27	-0.13	0.10	0.01	1.00	0.12	0.00
MgSt	-0.48	0.14	0.45	-0.16	0.12	1.00	0.03
Force	0.00	0.02	0.00	-0.01	0.00	0.03	1.00

Table 2-3. Correlation co-efficient between composition and compression force

2.4 Conclusion

A method was developed to design calibration set in spectral space using pure component spectra. Two calibration sets were designed in spectral space, one for a single API formulation and the other for multiple API formulation. The spectral calibrations sets for single API and multiple API were used to developed quantitative NIR methods for respective APIs. The performances of the methods were compared with traditional calibration strategies. The details of the method development and comparative results are discussed in the following chapter.

Copyright ©: Part of this chapter has been reprinted from [175]. Copyright clearance is provided at the end the dissertation.

3 Chapter 3: Comparison between the model performance of spectral calibration and traditional calibration sets

3.1 Introduction

The first step in developing NIR quantitative method for tablet analysis is to design the calibration set. The performance of the NIR method depends on the suitability of the data generated from the calibration set. Calibration sets are traditionally designed in concentration space using experimental design plan. Traditional calibration sets are prone to have redundant samples while simultaneously lacking necessary samples for a successful calibration model. Traditional calibration sets are also developed based on generic design. Similar calibration designs are followed for different spectroscopic techniques and different formulations. Such a calibration strategy lacks the optimizing capability to be technique and formulation specific. A new calibration strategy was developed for designing an NIR calibration set for quantitative analysis of tablets. Following this strategy, NIR calibration sets were designed in spectral space instead of concentration space. The pure component NIR spectra of a tablet formulation were used to define the spectral space of that formulation. The performance of this strategy was compared with commonly employed experimental design approaches to calibration development. The comparisons were conducted on single API (Active Pharmaceutical Ingredient) and multiple API formulation to quantify model drugs using NIR spectroscopy.

In the single API formulation, the comparison was based on a system to quantify a model drug, Acetaminophen, in pharmaceutical compacts using NIRS. A 2-factor full factorial

design (Acetaminophen with 5 levels and MCC:Lactose with 3 levels) was used for calibration development as an example of traditional calibration design. Three replicates at each design point resulted in a total of 45 tablets for the calibration set. Using the newly developed spectral based method, 11 tablets were prepared for the calibration set. Partial least squares (PLS) models were developed from respective calibration sets. Model performance was comprehensively assessed based on the ability to predict Acetaminophen concentrations in multiple prediction sets. One prediction set contained similar information to calibration set while the other prediction sets contained different information from calibration set in order to assess the model accuracy and robustness.

In multiple API formulation, NIRS was used to develop a quantitative method for a model drug containing two APIs as Acetaminophen and Caffeine. The traditional calibration design was developed in the concentration space using concentration-based approach. A full factorial design of experiment was used to vary the APIs concentrations, each at five levels and excipient concentration ratio at 3 levels. Each tablet was compressed at three compaction pressures resulting in a total of 225 tablets in the calibration set. Three other calibration sets were also developed using traditional optimal experimental designs (central composite, D-Optimal and I-Optimal designs). A calibration set of 33 design points was developed from each experimental design. Another calibration set containing 33 samples was developed in the spectral space following the spectral space based method. Partial Least Square (PLS) models were developed from the respective calibration sets to predict APIs concentrations in an independent test set. This chapter aims at comparing the performance of the spectral

calibration sets with traditionally developed calibration sets. The experimental details and results of these comparative studies are provided in the following sections.

3.2 Single API formulation

The single API formulation described in chapter 2 was used for this study. The single API formulation contained a model drug Acetaminophen and four excipients such as Microcrystalline cellulose (MCC), spray dried Lactose, Hypromellose and Magnesium stearate. The target formulation was set as Acetaminophen (27.3% w/w), MCC (34.15% w/w), Lactose (34.15% w/w), HPMC (3.9% w/w) and MgSt (0.5% w/w). A traditional full factorial calibration design and a spectral calibration design were used to develop quantitative NIR method for this formulation to predict Acetaminophen concentration in pharmaceutical tablets. The prediction performances of the respective calibration models were compared to evaluate the utility of the spectral design strategy.

3.2.1 Material and Methods

3.2.1.1 Full factorial Calibration

The traditional calibration design was a 5 by 3 level, 2-factor full factorial design. Samples were prepared in the laboratory. The composition of these samples is provided in Table 3-1. The factors were Acetaminophen concentration (% w/w) and MCC to Lactose ratio. Each design point was replicated three times (45 samples total). The use of a full factorial design of experiment for the calibration set provided orthogonality between the active ingredient and the excipient ratios.

Design #	Acetaminophen	Hypromellose	Intra-granular Lactose	Micro-crystalline cellulose	Extra-granular Lactose	Magnesium Stearate
1	19.11	2.73	5.46	51.77	20.43	0.50
2	23.21	3.32	6.63	48.65	17.70	0.50
3	27.30	3.90	7.80	45.53	14.97	0.50
4	31.40	4.49	8.97	42.41	12.24	0.50
5	35.49	5.07	10.14	39.29	9.51	0.50
6	19.11	2.73	5.46	38.83	33.37	0.50
7	23.21	3.32	6.63	36.49	29.86	0.50
8	27.30	3.90	7.80	34.15	26.35	0.50
9	31.40	4.49	8.97	31.81	22.84	0.50
10	35.49	5.07	10.14	29.47	19.33	0.50
11	19.11	2.73	5.46	25.89	46.31	0.50
12	23.21	3.32	6.63	24.33	42.02	0.50
13	27.30	3.90	7.80	22.77	37.73	0.50
14	31.40	4.49	8.97	21.21	33.44	0.50
15	35.49	5.07	10.14	19.65	29.15	0.50

Table 3-1. Composition (% w/w) of the full factorial calibration set

Granules of Acetaminophen (Mallinckrodt Inc., Raleigh, NC, USA), Hypromellose (HPMC; Pharmacoat 606, Shin-Etsu Chemical Co. LTD, Tokyo, Japan) and intra-granular Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA) were manufactured using a fluid bed processor (model WSG 5, Glatt, Binzen, Germany). Granules, Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA) and Lactose (modified spray-dried; Foremost Farms USA,

Rothschild, WI, USA) were mixed in a bin blender for 15 min. Magnesium stearate (Fisher Scientific, Waltham, MA, USA) was added to this blend and blended for an additional 2 min. Tablets were compressed at lab scale on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5500 lb force using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg. The manufacturing steps were tested to produce reliable tablets and reported in a previous study [174].

3.2.1.2 Spectral Design

The preparation method of the spectral design is described in section 2.2 in chapter 2. The spectral calibration set contained 11 tablets. The composition of these 11 tablets is given in Table 2-1 in chapter 2. The spectral design samples were prepared by direct compression. All the components, Acetaminophen, MCC, Lactose, HPMC and Mg stearate for each design point were weighed and placed in a 10 ml scintillation vial. The ingredients were mixed by rotating the scintillation vial. The intra tablet homogeneity was less of a concern due to the large spot size (~10 mm) of the NIR instrument. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5000 lb force using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

3.2.1.3 Prediction sets

Multiple prediction sets (3 sets) were used to evaluate the performance of the traditional calibration design and spectral calibration design. The first prediction set tablets were prepared at lab scale whereas the other two prediction sets were prepared at the manufacturing scale.

3.2.1.4 Lab scale prediction

The first prediction set (prediction set 1) was developed using the same full factorial design as used for the traditional calibration set; containing 45 samples at five levels of Acetaminophen and 3 levels of MCC:Lactose ratio. These samples were prepared using the same method as calibration set. Granules of Acetaminophen, Lactose and HPMC were prepared and then compressed with MCC and MgSt at lab scale on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5000 lb force using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

3.2.1.5 Manufacturing scale prediction

The second and third prediction sets were prepared at manufacturing scale. Prediction set 2 was prepared by varying the Acetaminophen concentration at 3 levels and keeping the MCC:Lactose ratio constant. A set of 20 tablets was collected from each Acetaminophen level, yielding 60 tablets for this set. Four batches of tablets were prepared using this design, resulted in 240 tablets. A total of 239 samples were used in subsequent analysis, after the accidental destruction of one tablet. Intentional variation resulting from manufacturing scale was introduced in prediction set 2 to challenge the calibration set in presence of unexpected variance. The target composition was used to create two additional batches for the prediction set 3. Tablets were compressed on a weekly basis for 8 weeks at the target composition. A set of 20 tablets were collected each week yielding 160 tablets for prediction set 3. These tablets served as the real time prediction samples.

At the manufacturing scale, after the granulation and blending steps described in the full factorial design section, tablets were compressed on a 38-station rotary tablet press (Elizabeth-Hata International, Inc., North Huntingdon, PA, USA) using round beveled

punches, 3/8 in. (9.5 mm) in diameter. The target tablet weight was 350 mg and the target breaking force was 8 kp.

3.2.1.6 Spectral collection and reference measurements

NIR reflectance measurements for both sides of each compact were collected using a bench top scanning monochromator instrument (XDS Rapid Content Analyzer, FOSS NIR Systems, Inc., Laurel, MD, USA) after tablets reached stable dimensions (viscoelastic relaxation). Spectra corresponding to each side of a compact were averaged to give one spectrum per compact. Acetaminophen reference values for compacts from all data sets except the spectral design were determined using High Pressure Liquid Chromatography (Waters Alliance 2790, Milford, MA, USA), followed by UV detection (Waters 2487). Gravimetric measurement was used as reference for the spectral design compacts.

3.2.1.7 Modeling strategy

Two calibration models were developed, one from the 45 tablets of the full factorial design and another from the 11 tablets of the spectral design. Both calibration models were developed from the lab scale tablets only. To compare calibration model performance between the full factorial set designed in concentration space and the set designed in spectral space, model performance was assessed for lab (prediction set 1 containing similar information as calibration) and manufacturing scale predictions (prediction sets 2 and 3 containing different information from calibration).

Principal Component Analysis (PCA) was used for spectral investigation and Partial Least Squares (PLS) regression was used to develop the calibration models. The NIPALS algorithm was employed. All calculations were performed with MATLAB 2011a (The

Mathworks, Natick, MA, USA) equipped with the PLS_Toolbox v. 7.9.3 (Eigenvector Research Inc., Wenatchee, WA, USA).

3.2.1.8 Model evaluation and design comparison parameters

Comparison of the performance of different calibration designs requires the evaluation of the resulting models. The full factorial and spectral design were compared based on their model predictive performance. Root mean squared error (RMSE) was used to evaluate the model predictive performance in the calibration (RMSEC) and prediction sets (RMSEP). The RMSEPs were separated into bias and Standard Error of Prediction (SEP). The bias and SEPs of the two models (one from full factorial and the other from spectral design) were compared using 95% *t* confidence interval around the bias difference and ratio of SEPs respectively; following the method outlined by Fearn et al. and adapted by Bondi et al [9, 176]. The 95% confidence interval was calculated based on the following equation

$$Bias_1 - Bias_2 \pm t_{n-1,0.025} * S_e \quad (3.1)$$

Where, S_e is the standard error of estimated difference and calculated based on the following equation,

$$S_e = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n * (n - 1)}} \quad (3.2)$$

Here, d_i represents the difference in residuals for sample “i” estimated by models being compared and \bar{d} represents the mean difference in residuals for two models. The biases between two models were considered significantly different only if the confidence

interval for bias difference did not contain zero. The confidence interval for SEP ratio was calculated based on equation (3.3) [9, 176],

$$\frac{SEP_1}{SEP_2} \times \frac{1}{L} \text{ and } \frac{SEP_1}{SEP_2} \times L \quad (3.3)$$

Where,

$$L = \sqrt{K + \sqrt{(K^2 - 1)}} \quad (3.4)$$

and,

$$K = 1 + \frac{2(1 - r^2) * t_{n-2,0.025}^2}{n - 2} \quad (3.5)$$

Here, r represents the correlation co-efficient between the residuals. The SEPs between two models were considered significantly different only if the confidence interval of SEP ratio did not contain one. The prediction performances between the two models were considered significantly different if the bias, or SEPs, or both were significantly different between the models being compared.

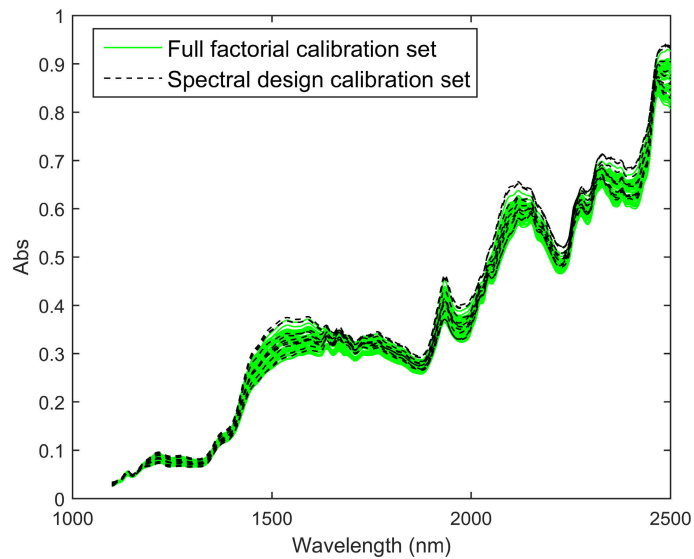


Figure 3-1. Raw data for two calibration sets, solid line (-) represents the full factorial design calibration samples and broken line (--) represents the spectral design calibration samples.

3.2.2 Results

3.2.2.1 Spectral investigations

Figure 3-1 illustrates the raw NIR spectra of the two calibration sets, traditional full factorial and spectral design calibration set. Prediction sets prepared at the manufacturing scale (prediction set 2 and 3) exhibited a baseline shift compared to the lab scale samples (data not shown). The observed differences in baseline were attributed to shape, density and hardness differences between the lab and manufacturing samples. Density differences in the samples are known to cause baseline shift in NIR spectra [177, 178]. Figure 3-2 illustrates the score plots of PCA analysis of the preprocessed spectra from two calibration sets and three prediction sets (total 500 spectra). Multiplicative scatter correction (MSC) followed by Savitzky Golay 1st derivative in a second order polynomial fit over a window size 19 was used as the spectral processing technique. Other

preprocessing methods showed similar results in the PCA analysis. The 95% confidence interval was dominated by the prediction set samples (444 out of 500 tablets). Full factorial, spectral design and prediction set 1 had similar variance explained by PC1 and PC2 and were grouped together by PC scores except for the fact that the full factorial design and prediction set 1 tablets had wider range of variance compared to the spectral design tablets. The full factorial design and prediction set 1 tablets were prepared at the exact same design points and showed similar projections in the PC space. Their wider spectral variance was attributed to the wider range of the excipient concentrations and physical variance compared to the spectral design tablets. The prediction sets prepared at the manufacturing scale (prediction set 2 and 3) grouped separately in the scores plot due to differences in the physical properties such as density, hardness, surface texture and mass.

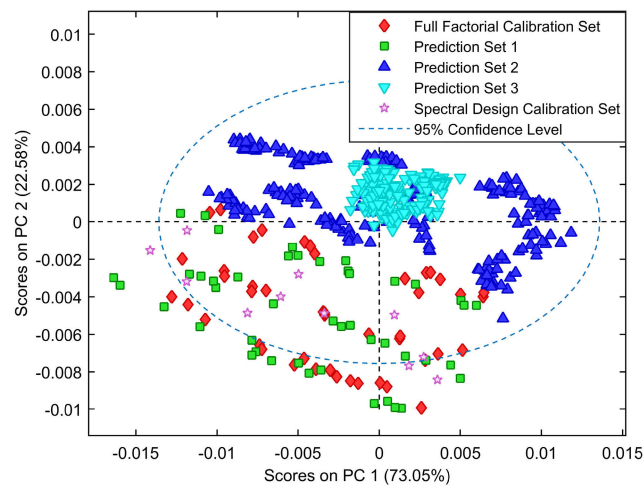


Figure 3-2. Projection of the full factorial calibration, spectral calibration and three prediction sets samples on the PC1 and PC2 of the combined PC space.

The different ranges of score values between the two calibration designs and the separate group of scores for manufacturing scale prediction sets were largely spanned in PC2 whereas all the samples had similar score values in PC1. It is critical for the calibration and prediction samples to have similar score values in the principal components that explain analyte concentration variation to achieve acceptable predictive performance. In this study, PC1 primarily explained the concentration variation of Acetaminophen (the absolute correlation co-efficient between Acetaminophen concentration and scores on PC1 was 0.9125). The different projections on PC2 did not affect the prediction at a large extent due to poor correlation with the Acetaminophen concentration (the absolute correlation co-efficient between Acetaminophen concentration and scores on PC2 was 0.2525).

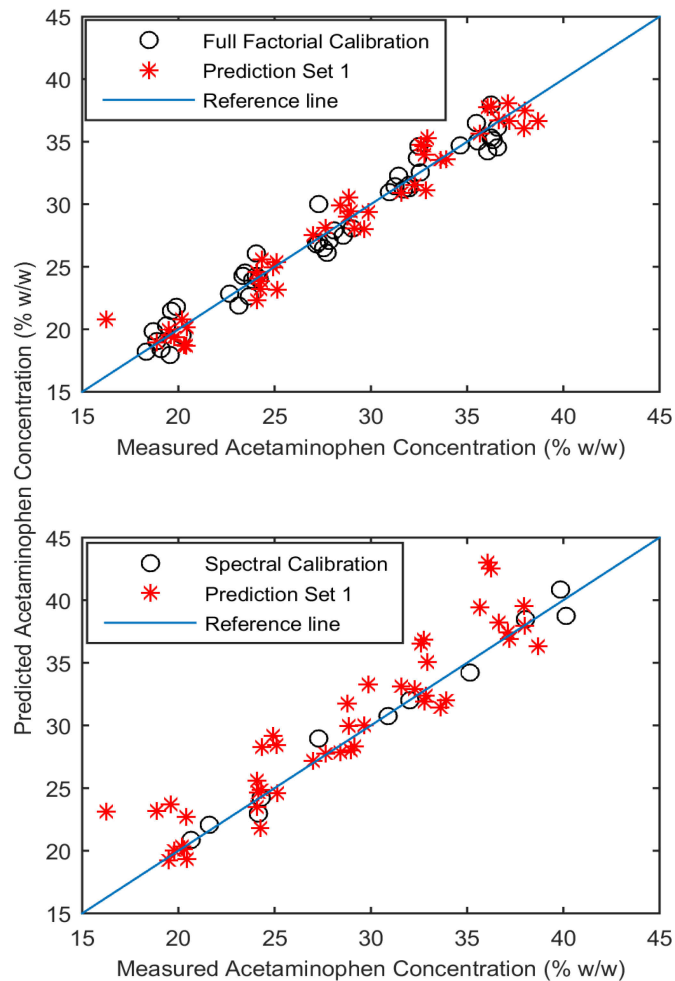


Figure 3-3. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 1 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 1 (lower panel).

3.2.2.2 Model optimization

Models developed from each calibration design were independently optimized. The range of the wavelength used was 1100-2500 nm for both calibration designs. The pre-processing and model optimization focused on maximizing the potential to predict

samples manufactured at different scales. Multiplicative Scatter Correction (MSC) followed by Savitzky-Golay first derivative (window size 19 and second polynomial order) and mean centering was used as the optimized preprocessing technique. Other preprocessing methods were also tried, but scatter correction followed by derivatives provided better predictive performance for the calibration models. This preprocessing technique minimized the baseline shift caused by scale variation in the tablet preparation methods between calibration and prediction sets. However, the spectral difference could not be completely removed as evident by the different projections of calibration and predictions sets (after applying same preprocessing technique) on the PC score space in Figure 3-2. Selection of the latent variables is critical for PLS model performance. A minimum number of latent variables with acceptable prediction performance was selected. The optimum numbers of latent variables for the spectral and full factorial calibration design sets were two (spectral) and three (full factorial), respectively.

3.2.2.3 Prediction performance

The calibration models from the full factorial and spectral design were used to predict the Acetaminophen concentrations in prediction sets. Prediction set 1 samples were prepared at the lab scale; and samples for prediction sets 2 and 3 were prepared at the manufacturing scale.

3.2.2.4 Lab scale prediction

Figure 3-3 shows the reference vs prediction plot for full factorial and spectral designs. The model developed from the full factorial design predicted Acetaminophen in prediction set 1 with RMSEC and RMSEP value of 1.13% and 1.34%, respectively.

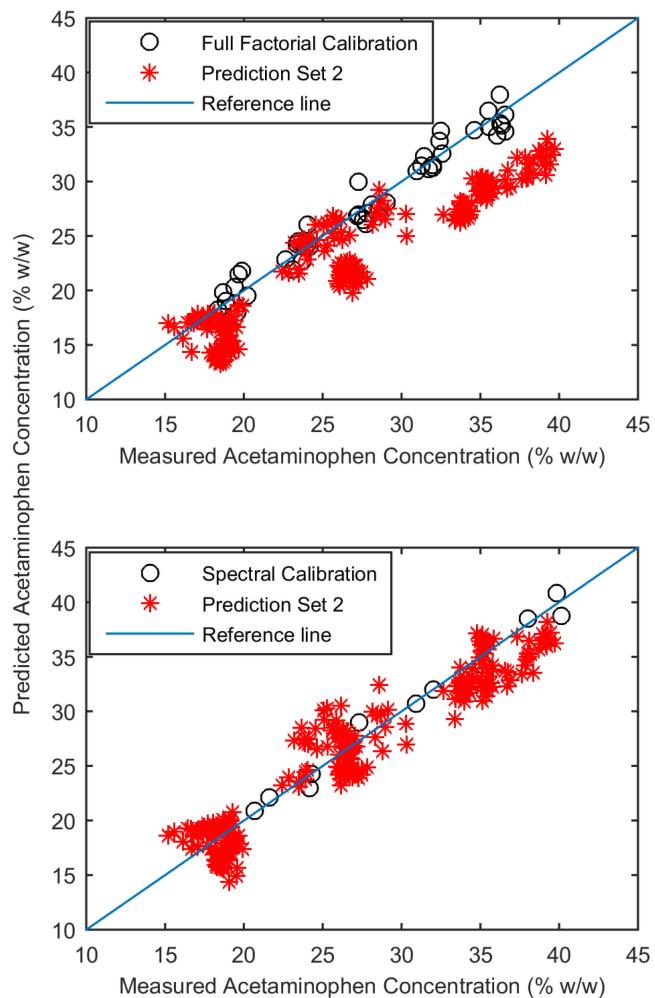


Figure 3-4. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 2 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 2 (lower panel).

The model developed from the spectral design had a lower RMSEC of 0.89% but higher RMSEP of 2.70 %. The improved prediction performance for the full factorial design was due to the chemical and physical similarity between prediction set 1 samples and the full factorial design. The spectral design samples had compositional and physical dissimilarities from prediction set 1 that include composition and preparation method.

The full factorial design and prediction set 1 samples were prepared using wet granulation method whereas the spectral design samples were prepared using dry mixing method.

3.2.2.5 Manufacturing scale prediction

Robustness was tested for models developed from the full factorial and spectral designs. Samples with new sources of variation (manufacturing scale) were predicted using the models developed at the calibration step (lab scale) to test the robustness. Figure 3-4 shows the reference vs prediction plot of Acetaminophen concentrations in prediction set 2 for both models. The model developed from the full factorial design had a RMSEP of 4.68 % compared to the RMSEP of 2.21 % for the model developed from spectral design. A bias is observed in the prediction of samples at the top figure of Figure 3-4 for the full factorial design model. The bias was due to an interference introduced by the scale of the sample preparation method. No significant bias was observed while predicting these samples by the model developed from spectral design as shown at the bottom of Figure 3-4.

Figure 3-5 shows the reference vs prediction plots of prediction set 3, containing 160 tablets. This prediction set was prepared at the target concentration to mimic routine production. Full factorial design model had a higher RMSEP of 5.05 % compared to the RMSEP of 1.68 % of the spectral design model. Table 3-2 shows the prediction results for both full factorial and spectral design.

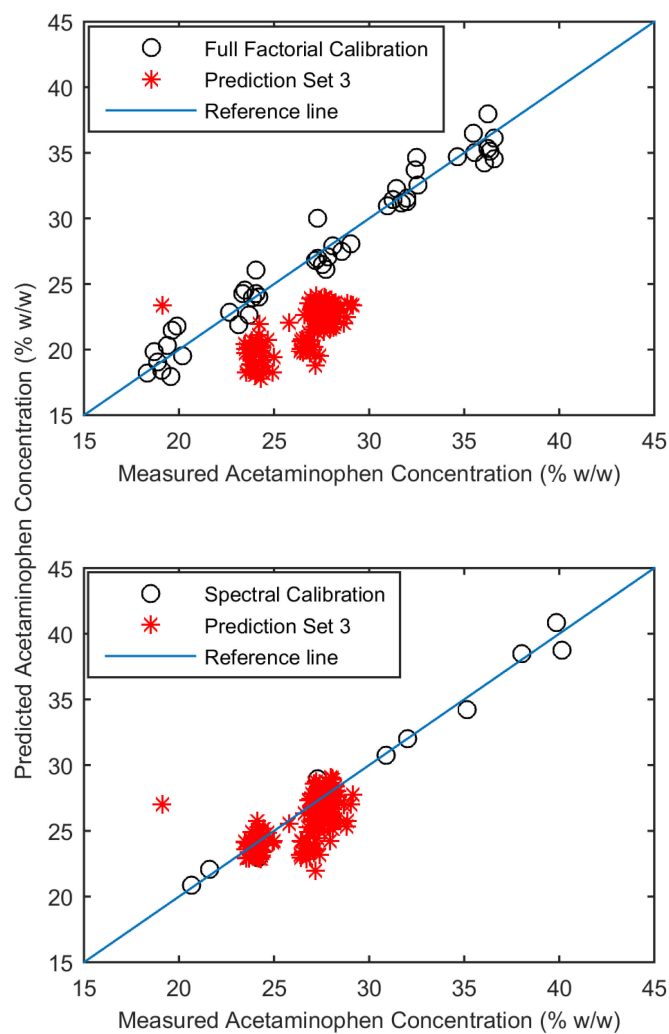


Figure 3-5. Measured vs predicted plot of Acetaminophen concentration for full factorial calibration and prediction set 3 (upper panel). Measured vs predicted plot of Acetaminophen concentration for spectral calibration and prediction set 3 (lower panel).

3.2.2.6 Design comparison

The predictive performances of the optimized models from the traditional concentration based full factorial design and the spectral design were compared. Figure 3-6 shows the confidence intervals for bias and SEP comparison between the optimized models. The

two models were significantly different in terms of both bias and SEP. Full factorial design predictions had a lower bias for the lab scale samples (prediction set 1) and

higher bias for the manufacturing scale samples (prediction set 2 and 3), compared to the spectral design model. While comparing SEPs, full factorial design had lower SEP for prediction set 1 and 3 compared to the spectral design. For prediction set 2, the spectral design had a lower SEP. Although full factorial design had a lower SEP, it had a higher RMSEP due to the higher bias in the

Calibration Design	Spectral design	Full factorial design
No. of samples	11 samples	45 samples
No. of latent variable	2 LV	3 LV
RMSEC % (w/w):	0.889	1.133
RMSECV % (w/w):	1.625	1.288
Prediction Set 1		
RMSEP % (w/w):	2.701	1.342
Prediction Bias	1.270	0.071
SEP	2.384	1.340
Prediction Set 2		
RMSEP % (w/w):	2.212	4.682
Prediction Bias	-0.578	-3.938
SEP	2.135	2.531
Prediction Set 3		
RMSEP % (w/w):	1.676	5.046
Prediction Bias	-0.786	-4.882
SEP	1.480	1.274

Table 3-2. Model performance comparison between two methods of experimental design; spectral and full factorial

prediction sets. For all the manufacturing scale predictions, full factorial design always had a higher RMSEP compared to the spectral design. The spectral design approach was found to be more robust than the traditional full factorial design in spite of having fewer calibration samples.

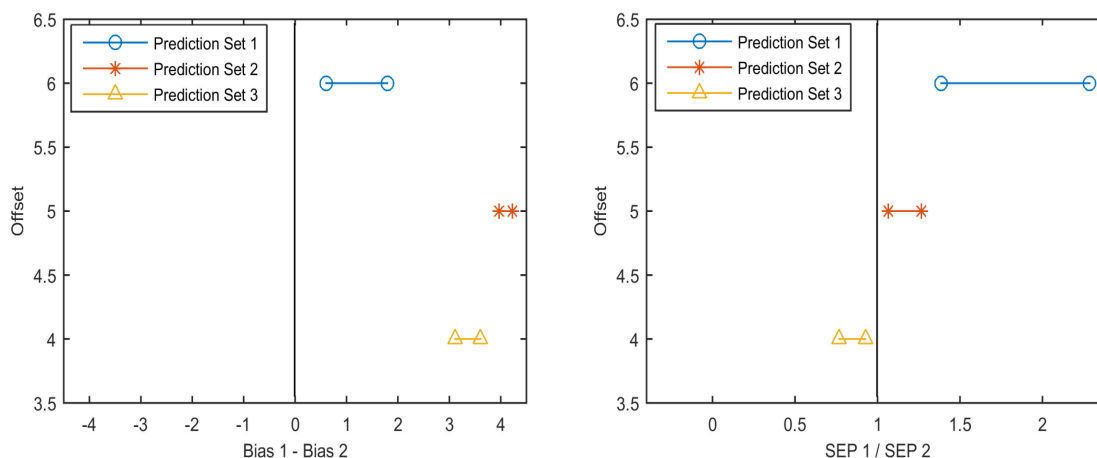


Figure 3-6. Confidence interval for bias (A) and standard error of prediction (B) comparison between full factorial design and spectral design for prediction sets analysis.

3.2.2.7 Discussion

The main objective of this project was to develop and test a spectral design method and compare it with a traditional full factorial design method for NIR spectroscopic multivariate calibration. The use of a traditional design, such as full factorial, in concentration space is a common practice for calibration sample preparation. However, this practice is inefficient and includes redundant information in the calibration model. The spectral design approach offers an efficient enhancement for NIR calibration development by using pure component spectra. Information about the pure component spectra was used to define an orthogonal spectral space for the formulation and identify a set of critical samples necessary for successful calibration model development.

Designing a calibration sample set in this orthogonal spectral space was found to be an efficient approach for developing an NIR calibration method for pharmaceutical tablets. Fewer samples were required when this approach was used. A reduction of 75% in raw material requirement was demonstrated (11 vs 45) while achieving similar (if not improved) predictive performance compared to the traditional full factorial design.

Considering the RMSEPs in manufacturing scale samples, the spectral design was found to be more robust than the full factorial design. The full factorial design was very sensitive to latent variable selection, scale variation and prone to biased prediction. The sensitivity of this concentration based design was due to the information that is unique to the calibration set and not always exists in the prediction sets. Fitting this information into the model introduced biased predictions in the presence of new information (lack of robustness). Biased corrected prediction performance (SEP) of the full factorial design were found to be similar to the spectral design for manufacturing samples.

The spectral calibration set showed similar SEP to the full-factorial calibration set due to similar span of Acetaminophen concentration between two calibration sets. The full-factorial calibration set contained wider range of excipient variation compared to the spectral calibration set, however, such information was redundant for calibration performance. The limited excipient variation in the spectral calibration set was representative of the additional excipient variation in the full factorial calibration set.

Prediction results at lab scale emphasized the similarity between the calibration and prediction samples. For the full factorial design, calibration and lab samples came from the same exact design points and were prepared using the same methodology. As expected, the calibration model was very accurate in predicting this set of samples. In spectral design method, errors from the lab scale predictions based on the spectral design model were within errors observed in other prediction sets due to the inherent robustness of the design.

Selection of the latent variables is critical to model performance. However, in practice, this selection process occurs only during model development. Once the model is

developed and optimized, further investigations of predictive performance are conducted based on that optimized model only. The number of latent variable does not change during the assessment of prediction performance. Following this practice, the comparative analyses were performed between the optimized models with fixed latent variables from respective designs.

3.2.3 Conclusion

A technique for designing a calibration set using pure component spectra for NIR spectroscopic calibration method was described and tested for quantitative analysis of pharmaceutical tablets. This technique was found to be efficient in terms of sample requirements compared to the traditional (full factorial) concentration based technique. A reduced number of calibration samples yielded a model with improved predictive capabilities for this individual system. In particular, a calibration model developed in spectral space demonstrated improved robustness in presence of variation introduced by a change of manufacturing scale, with respect to a model developed using full factorial design.

Calibration model performance can be highly system specific. The improved prediction performance of the spectral calibration set is not expected for every formulation. However, it is expected that the spectral calibration set will provide equivalent prediction performance with fewer number of samples while comparing with traditional calibration sets. It is also expected that the spectral calibration set will provide improved prediction performance with an equal number of calibration samples while comparing with traditional calibration sets. In general, spectral calibration is expected to be efficient in

sample number compared to the traditional calibration sets. Such efficiency is more critical for multiple API formulation, since the traditional calibration sets for multiple API requires larger calibration sets compared to the single API. The following section describes a comparative study conducted on multiple API system to assess the performance of spectral calibration set and compare with traditional calibration sets.

3.3 Multiple APIs formulation

The utility and performance of the spectral calibration technique was assessed for a multiple API formulation. The performance of the spectral calibration set (33 calibration samples) was compared with traditional full factorial calibration set of larger size (225 calibration samples) as well as traditional optimal calibration sets of similar sizes (33 calibration samples) to further investigate the applicability of the newly developed spectral calibration strategy. The multiple API formulation described in section 2.3 in chapter 2 was used for this study. The multiple API formulation contained two model drug Acetaminophen (Acetaminophen) and Caffeine and four excipients such as Microcrystalline cellulose (MCC), spray dried Lactose, Crosscarmellose Na and Magnesium stearate. A full factorial calibration design and three optimal calibration designs were used as examples of traditional calibration set. A spectral calibration design was used to develop another calibration set. All these calibration sets were used to develop quantitative NIR methods for this formulation to predict Acetaminophen and caffeine concentration in pharmaceutical tablets. The prediction performances of the respective calibration models were compared to evaluate the utility of the spectral design strategy.

3.3.1 Material and Methods

3.3.1.1 Full factorial Calibration

The model drug product contained two APIs as Acetaminophen (Mallinckrodt Inc., Raleigh, NC, USA) and anhydrous Caffeine (Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA), and four excipients as Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA), Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA), Crosscarmellose sodium (Crosscarmellose Na, Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA) and Magnesium stearate (Fisher Scientific, Waltham, MA, USA). The

target formulation was set as Acetaminophen (31.25% w/w), Caffeine (4.05% w/w), MCC (37.32% w/w), Lactose (24.89% w/w), Crosscarmellose Na (2% w/w) and MgSt (0.5% w/w). The traditional calibration design was developed using a 4-factor full factorial design. The factors were Acetaminophen concentration, Caffeine concentration, MCC: Lactose ratio and compaction force. Each of the APIs (Acetaminophen and Caffeine) concentrations was varied at five levels and other two factors (MCC: Lactose ratio and compaction force) were varied at three levels resulting in 225 design points in the traditional full factorial calibration set. A test set was developed by varying all the factors except MCC: Lactose ratio, each at three levels resulting in total 27 design points. Table 3-3 provides the details of the full factorial calibration and test design.

	Design Factors	Design Levels				
Calibration Design (5x5x3x3) 225 samples	Acetaminophen (%)	L.C.-40%	L.C.-20%	L.C.	L.C.+20%	L.C.+40%
	Caffeine (%)	L.C.-80%	L.C.-40%	L.C.	L.C.+40%	L.C.+80%
	MCC/Lac	1		1.5		2
	Force (lb)	4000		5000		6000
Test Design (3x3x3) 27 samples	Acetaminophen (%)	L.C.-35%		L.C.		L.C.+35%
	Caffeine (%)	L.C.-70%		L.C.		L.C.+70%
	Force (lb)	4000		5000		6000
	MCC/Lac	1.5				

Table 3-3. Full-factorial calibration design and test set

Tablets were individually prepared by direct compression. All the components for a single design point were weighed using a digital weighing machine (Data Range, Model No. AX504DR, Mettler Toledo) and placed in a 10 ml scintillation vial. The ingredients were mixed in a bin

blender (L.B. Bohle LLC, Warminster, PA, USA) for 10 mins followed by a high shear mixing using a vortex machine (Vortex-2 Genie, Model G-560, Scientific Industries, IN, USA). The final mixing was done in the bin blender before compression. Additional blending steps were included in the process to ensure well mixing of the low dose caffeine. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at respective compression forces (4000/ 5000/ 6000 lb) using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

3.3.1.2 Optimal calibration designs

Three optimal calibration designs were developed in this study as examples of traditional calibration designs. The factors of the designs were Acetaminophen concentration, Caffeine concentration, MCC: Lactose ratio and compaction force. The CCD design was developed by creating a CCD structure containing 11 design points in the concentration space and varying the compaction force at three levels for each design point resulting a total of 33 design points in the calibration set. The D-Optimal and I-Optimal designs were developed using JMP (v. 12, SAS, Cary, NC, USA). A set of 33 design points was selected from a full factorial space such that the determinant of the information matrix is maximized (D-Optimal) and average prediction variance is minimized (I-Optimal). All the calibration design tablets were subsets of the full-factorial design described earlier.

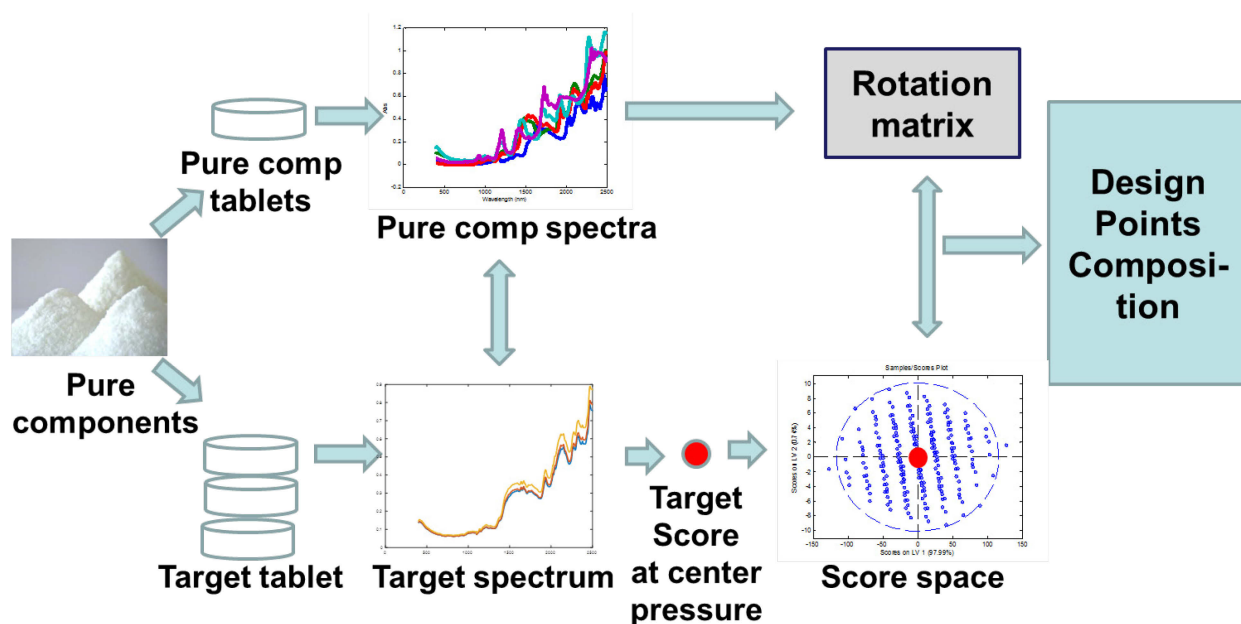


Figure 3-7. Experimental procedures for designing a calibration set in spectral space

3.3.1.3 Spectral Design

The preparation method of the spectral design of multiple API is described in section 2.3 in chapter 2. Figure 3-7 illustrates the procedure for designing a calibration set in the spectral space. In summary, the pure component tablets (six tablets) and one tablet at the target composition were compressed at the middle compression force (5000 lb) using a Carver Automatic Tablet Press. A spectrum at the target composition was used for calculating model tablet scores (target scores) on the orthonormal basis of pure component spectra. A 6-factor full factorial design with 5 levels per factor was calculated in the orthogonal spectral space where the “model tablet” scores were set as the center of the design. The full factorial structure resulted a total of 15625 (5^6) design points. The compositional requirements for the target design were calculated using eq. 3.6 (same as equation 2.8).

$$\mathbf{C}_{u \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} = \mathbf{T}_{u \times r}^{\text{Score space}} * (\mathbf{P}_{k \times n}^{\text{Pure component spectra}} * \mathbf{W}_{n \times k}^{\text{Pure component basis}})^{-1} \quad (3.6)$$

Here, **C** is the compositional matrix for the targeted score design having sample size “u”, **R** is the rotational matrix, **T** is the target design in the score space, **P** is the pure component spectra, **W** is the basis vector set and $r = k$ (number of latent variables = number of components). After determining the compositional requirements ‘C’ of the target design, the target design spectra were simulated using the compositional requirements, the pure component spectra and the rotation matrix. The Kennard Stone algorithm was used to select 30 design points in the PCA score space of the target design spectra (Spectral Design) [132]. The composition of these 30 design points is provided in Table 2-2 in section 2.3 of chapter 2. The target formulation was replicated 3 times resulting in 33 design points in the spectral design. The 33 tablets of the spectral design were prepared by direct compression on Carver Automatic Tablet Press using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg. The compression forces were selected such that the correlation between compression force variation and concentration variation of each component is minimal and variance in compaction force is maximal.

3.3.1.4 Experimental control

Intra homogeneity of the tablets can be critical in case the respective spectrum represents only a portion of the tablet. The blending was performed in three steps in a bin blender followed by high shear mixing and subsequent blending to avoid inhomogeneity of the tablet. The intra homogeneity of the tablet was analyzed to test the appropriateness of the blending process. A set of 15 tablets were selected from the 225 calibration samples. NIR images of these tablets were

collect by NIR chemical imaging system (MatrixNIR, Malvern Inc., MD) with 0.5× objective lens (field of view 1.72 cm × 2.15 cm). An integration time of 256 ms and a count of 16 co-adds were utilized throughout the image collection of all 15 tablets. The wavelength range was from 1050 to 1620 nm with a 5 nm interval. A typical Classical Least Square approach was used to develop a quantitative method for image analysis.

All the materials were stored in room temperature and relative humidity. Anhydrous caffeine was reported to be stable at 75% RH for 7 weeks [173]. No hydration and anhydrous caffeine was expected considering lower room RH (~60%) and shorter storage time and analysis. The tablets were stored for 2 weeks to allow viscoelastic relaxation. Two weeks duration was set based on a previous study [174].

Solid fraction of each of the 225 tablets was calculated based on the following equation to test the appropriateness of the compaction force range. A wide range of solid fraction was expected as a result of compaction force variation.

$$\text{Solid fraction of tablet} = \frac{\text{Tablet density}}{\text{True density}} \quad (3.7)$$

3.3.1.5 Spectral collection and reference measurements

NIR reflectance measurements for both sides of each compact were collected using the bench top scanning monochromator instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc., Laurel, MD, USA) after tablets reached stable dimensions (viscoelastic relaxation). Spectra corresponding to each side of a compact were averaged to give one spectrum per compact. Gravimetric measurement was used as reference for all tablets.

In the gravimetric method, all the weights of the individual chemical components at each design point were recorded from digital weighing machine (Data Range, Model No. AX504DR, Mettler Toledo). The tablet weights were recorded after compression using the same digital weighing machine. The concentrations % (w/w) were calculated for APIs from the respective individual weights and tablet weights, and used as reference values for model development.

3.3.1.6 Modeling strategy and optimization

Quantitative models were developed using Partial Least Squares (PLS) modeling technique in MATLAB 2015a environment (The Mathworks, Natick, MA, USA) using PLS_Toolbox v. 7.9.3 (Eigenvector Research Inc., Wenatchee, WA, USA). Data independent spectral preprocessing techniques were used to optimize model performance. Two calibration models were developed for each API, one from the 225 tablets of the full factorial design and another from the 33 tablets of the spectral design. Models developed from each calibration design were independently optimized for each API. The processing techniques and loading vectors were selected independently. Selection of the loading vector is critical for PLS model performance. Latent variables were chosen based on a parsimonious approach. A minimum number of latent variables with an acceptable performance level was selected. Model performance was assessed based on the prediction of Acetaminophen and Caffeine concentration in calibration and test set tablets.

3.3.1.7 Model evaluation and design comparison parameters

Comparison of the performance of different calibration designs requires the evaluation of the resulting models. Root mean squared error (RMSE) was used to evaluate the model predictive performance in an independent test sets (RMSEP). A two-way analysis of variance (ANOVA) test was performed to compare the prediction errors of the full factorial and spectral calibration designs for each API [22]. The underlying model for the ANOVA analysis was following, where

index 'i' refers to the calibration model and index 'j' refers to the sample number. The symbol α_i and β_j refer to the effect of model number i and sample number j on prediction error, respectively.

$$(\hat{y}_{ij} - y_{ij})^2 = \mu + \alpha_i + \beta_j + e_{ij} \quad (3.8)$$

The calibration models are considered significantly different in terms of prediction performance, if the calibration model parameter α is found to be significant in the ANOVA analysis.

Robustness of the calibration models from respective calibration designs were tested based on cross design prediction. During cross design prediction, calibration model was developed from spectral calibration set and used to predict the API concentration in traditional calibration sets. Subsequently, calibration models were developed traditional calibration sets and used to predict the API concentration in spectral calibration set. The objective was to assess the prediction performance of a calibration model in presence of new test structure. Robustness of the calibration model was tested for both Acetaminophen and Caffeine PLS model.

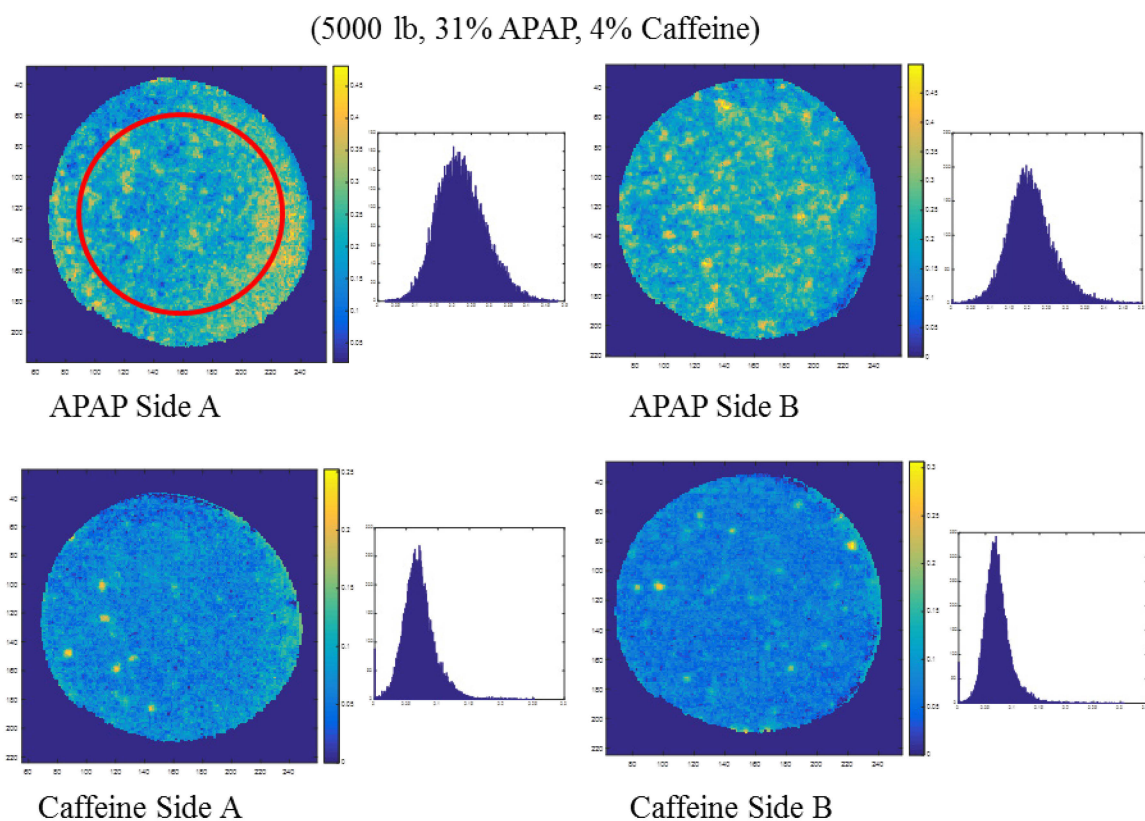


Figure 3-8. Quantitative predictions of NIR images of tablet at target formulation and center compaction force.

3.3.2 Results and Discussion

3.3.2.1 Experimental control

Figure 3-8 shows the quantitative predictions of APAP and caffeine at both sides of the tablet prepared at the target formulation and center compaction force. The quantitative results were obtained from the NIR images of the tablet and CLS model of the pure component images. It was shown that the APAP and caffeine were homogeneously distributed in the tablet. The red circle shows the spot size (10 mm) of the NIR diffuse reflectance measurement. It was shown that the NIR diffuse reflectance measurement reflects majority of the tablets due to large spot size. The criticality of the intra tablet homogeneity was minimized by such large spot size.

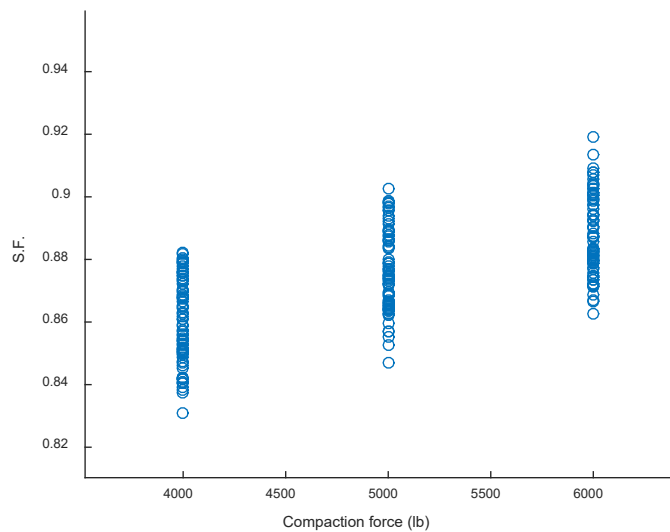


Figure 3-9. Solid fraction of tablets compressed at different compaction forces.

Figure 3-9 shows the solid fraction measurements of 75 tablets at each compaction force. The compaction force range (4000 lb-6000 lb) results in a wide range of solid fraction (0.83-0.92). This range of solid fraction was able to produce wide range of spectral variation related to tablet density.

3.3.2.2 Calibration design comparison between full factorial and spectral calibration sets

Figure 3-10 shows the tablet compositions for the traditional full factorial calibration design, spectral calibration design and test design. The resultant compositions from the spectral design approach contained similar ranges of Acetaminophen and Caffeine concentrations. Spectral design had wider range of excipient variation compared to the traditional full factorial design. The wider excipient ranges were the result of distant design points in the target score space, that could only be obtained by large variation in the excipient concentrations. Although wide concentration range is considered to increase calibration model robustness, it can introduce non-linearity and affect prediction performance. Good fit is almost guaranteed for a small set with

narrow region. Moreover, a typical production batch is manufactured at the target composition and large excipient variation is not expected. So, it is recommended to use two calibration sets, one for coarse prediction of concentration (global calibration set) and the other for precise and accurate prediction in the critical range (local calibration set). This technique is known as bracketing [47]. However, maintaining two calibrations and selecting which one to use is challenging. A single robust calibration capable of predicting at the critical range would be advantageous over the bracketing technique. In this study, spectral calibration set containing large excipient variation was tested for the critical range prediction around target composition. Successful model performance from the spectral calibration set would offer a single calibration set with sufficient robustness and accuracy.

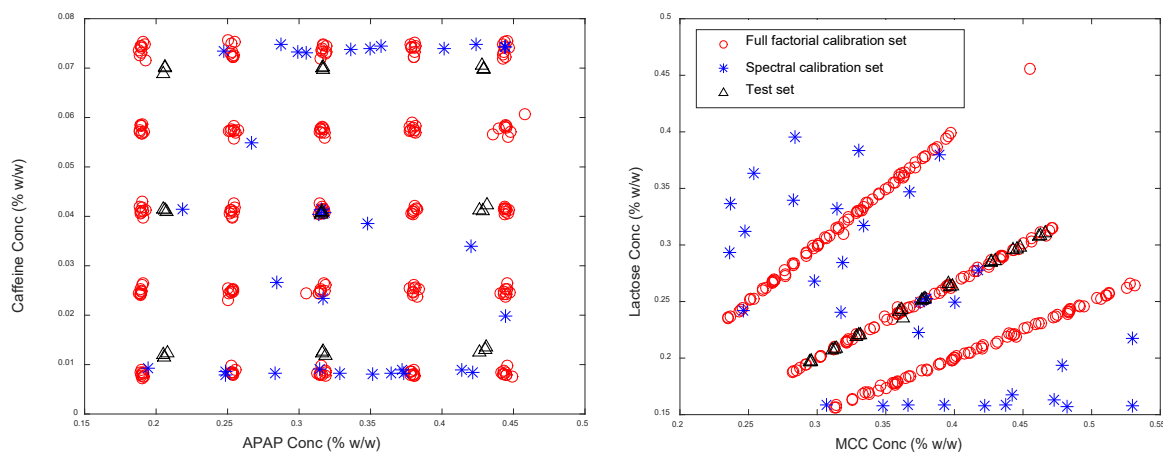


Figure 3-10. Tablet compositions of the traditional full factorial calibration, the spectral calibration and the test set.

3.3.2.2.1 Prediction performance

The calibration models from the full factorial and spectral design were used to predict the Acetaminophen and Caffeine concentrations in the prediction set.

3.3.2.2.1.1 Acetaminophen prediction

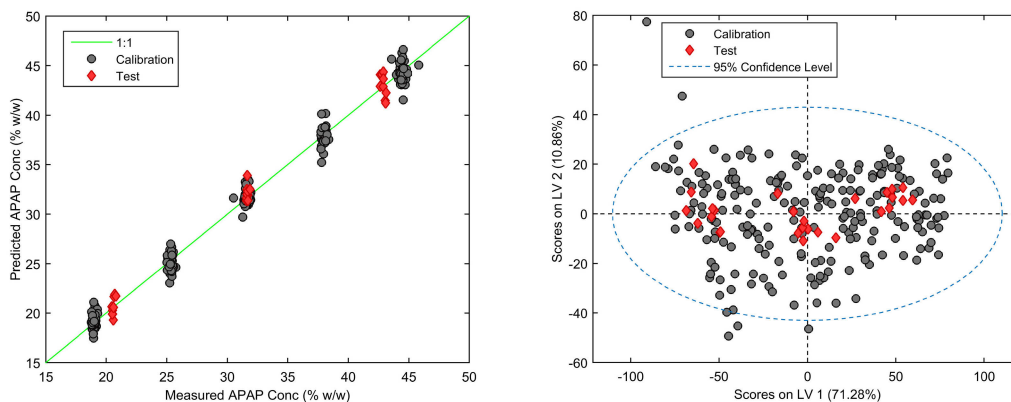


Figure 3-11. Calibration model performance for traditional full factorial design during Acetaminophen prediction

Figure 3-11 shows the calibration performance of the traditional full factorial design for Acetaminophen prediction in the calibration and test set. Figure 3-11 (right) shows the projection of the calibration and test samples on the first two loading vectors of the PLS model. The calibration set and test set had similar spectral variance as indicated by their similar projections on the first two loading vectors. The calibration set had wider range of scores on LV2 compared to the test set. This effect was attributed to the wide excipient variation present in the calibration set.

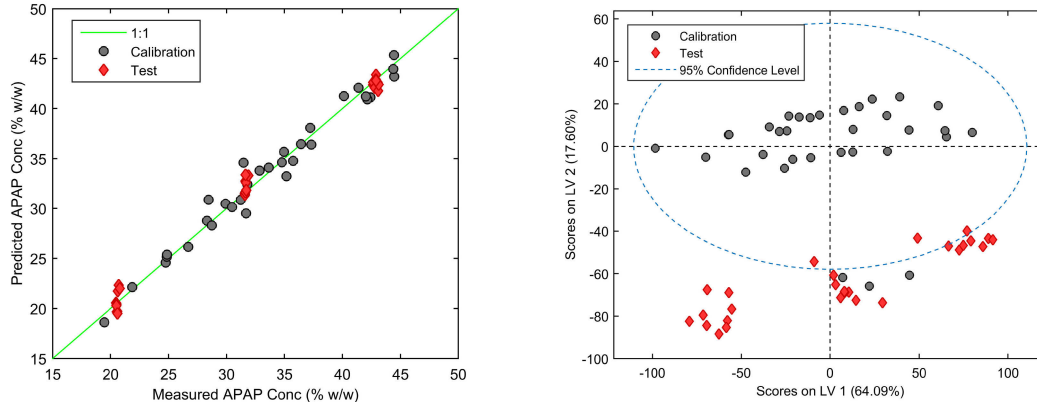


Figure 3-12. Calibration model performance for spectral design during Acetaminophen prediction

The RMSEP for the traditional Acetaminophen calibration was around 1% (w/w) and no significant bias was observed during prediction. Figure 3-12 shows the calibration performance of the spectral design for Acetaminophen prediction in the calibration and test set. The score plot (Figure 3-12 (right)) shows similar projections of the calibration and test set on LV1 and different projections on the LV2. However, this difference in the scores on LV2 did not affect the prediction performance due to compensating effect of LV4 on the opposite direction (data not shown). The information in LV2 and LV4 was majorly attributed to the wide range of excipient variation in spectral calibration set. The RMSEP was 0.83% (w/w) and no significant bias was observed. The prediction errors of the full factorial and spectral calibration design for Acetaminophen prediction was found to be similar in the ANOVA analysis (p value 0.1968). Table 3-4 provides the details of the calibration models developed from traditional full factorial and spectral design.

	Traditional (225 samples)	Spectral (33 samples)
Preprocessing techniques X block Y block	MSC*, Auto scale Auto scale	MSC*, SG (15,2,1)**, Auto scale Auto scale
LV	6	4
RMSEC (% w/w)	0.829	1.092
RMSECV (% w/w)	0.867	1.336
RMSEP (% w/w)	0.982	0.832
Bias (% w/w)	0.221	0.115
R ² Calibration	0.991	0.973
R ² Prediction	0.989	0.991
p values (ANOVA)	0.1968	

Table 3-4. Model performance comparison between full factorial and spectral calibration design for quantitative analysis of Acetaminophen

3.3.2.2.1.2 Caffeine prediction

Figure 3-13 shows the calibration performance of the traditional full factorial design for Caffeine prediction in the calibration and test set. The calibration set and test set had similar spectral variance; indicated by their similar projections on the first two loading vectors as shown in Figure 3-13 (right). The score ranges were similar in both latent variable directions as opposed to the scenario in the Acetaminophen prediction model.

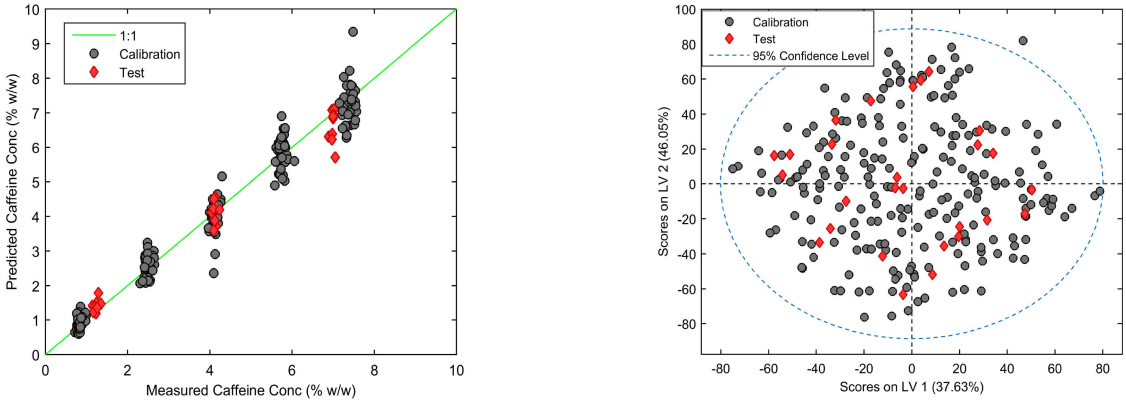


Figure 3-13. Calibration model performance for traditional full factorial design during Caffeine prediction

This effect was attributed to the similar ranges of APIs (Acetaminophen and Caffeine) concentrations between the calibration and test set. The first two loading vectors primarily explained the direction of variation in Caffeine concentrations even at low dose, due to maximization of covariance between spectra and Caffeine concentrations. These loading vectors also explained the direction of variation in Acetaminophen concentrations due to its unique spectral feature and relatively large dose. The RMSEP for the traditional Caffeine calibration was 0.41% (w/w) and no significant bias was observed during prediction. Figure 3-14 shows the calibration performance of the spectral design for Caffeine prediction in the calibration and test set. Similar trend in the score plot Figure 3-14 (right)) was observed as compared to the traditional calibration design. The first two loading vectors primarily explained the direction of the variation in APIs concentrations. The RMSEP was 0.41% (w/w) and no significant bias was observed. The prediction errors of the full factorial and spectral calibration design for Caffeine prediction were found to be similar in the ANOVA analysis (*p* value 0.9527). Table 3-5 provides the details of the calibration models developed from traditional full factorial and spectral design.

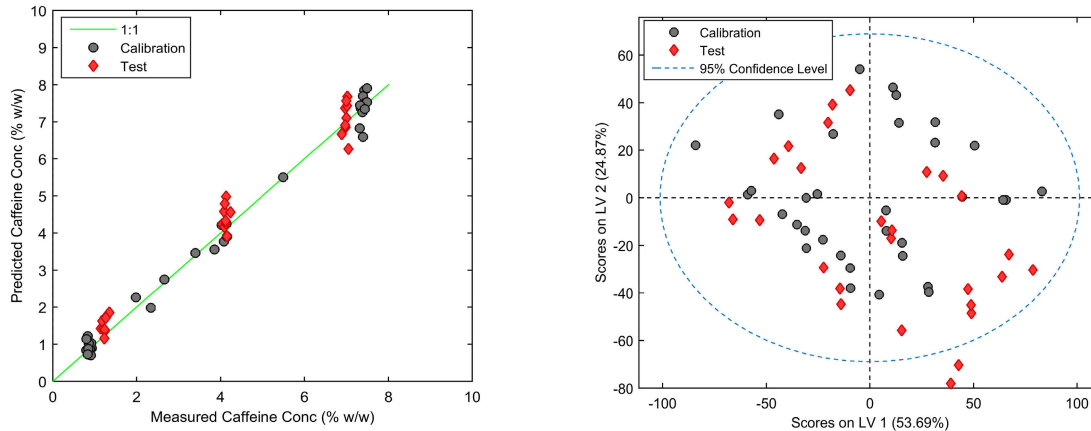


Figure 3-14. Calibration model performance for spectral design during Caffeine prediction.

3.3.2.2.2 Discussion

The main objective of this project was to develop and test a spectral design method and compare it with a traditional full factorial design method for NIR spectroscopic multivariate calibration for multiple API system. The use of a traditional design such as full factorial in concentration space is a common practice for calibration sample preparation. However, this practice is inefficient and provides redundant information to the calibration model. This effect is exacerbated for the tablets containing multiple API. The spectral design approach offers an efficient enhancement for NIR calibration development by using pure component spectra. Information about the pure component spectra defined an orthogonal spectral space for the formulation and was used to identify a set of critical samples with necessary compositional variation for successful calibration model development.

	Traditional (225 samples)	Spectral (33 samples)
Preprocessing techniques X block Y block	MSC*, Auto scale Auto scale	MSC*, SG (15,2,1)**, Auto scale Auto scale
LV	6	6
RMSEC (% w/w)	0.394	0.261
RMSECV (% w/w)	0.410	0.473
RMSEP (% w/w)	0.406	0.411
Bias (% w/w)	0.087	0.232
R ² Calibration	0.971	0.991
R ² Prediction	0.976	0.979
<i>p</i> values (ANOVA)	0.9527	

Table 3-5. Model performance comparison between full factorial and spectral calibration design for quantitative analysis of Caffeine

Variation in the physico-chemical properties of the tablet (tablet density) was also added into this study as a potential source of variation for NIR analysis of tablets. The objective was to test the performance of the spectral calibration design technique in presence of physical variation. Physical variation was added into the spectral design by minimizing correlation between composition and compaction pressure to maximize design orthogonality. Physical variation can also be added into the spectral space by compressing pure components using different compaction forces. While this effort helps to build robustness against density variation, effect of

tablet density variation on NIR spectra were also minimized by preprocessing techniques (scatter correction and derivatives). Both spectral and traditional calibration models were found to be robust against density variation. However, density variation was added more efficiently in the spectral design technique (minimizing correlation) compared to the traditional full factorial technique (replicating each design at different compaction forces).

Designing a calibration sample set in spectral space was found to be an efficient approach for developing a NIR calibration method for tablets containing multiple API. Fewer samples were required when this approach was used. A reduction of 85% in raw material requirement was demonstrated (33 vs 225 samples) while achieving similar predictive performance compared to the traditional full factorial design.

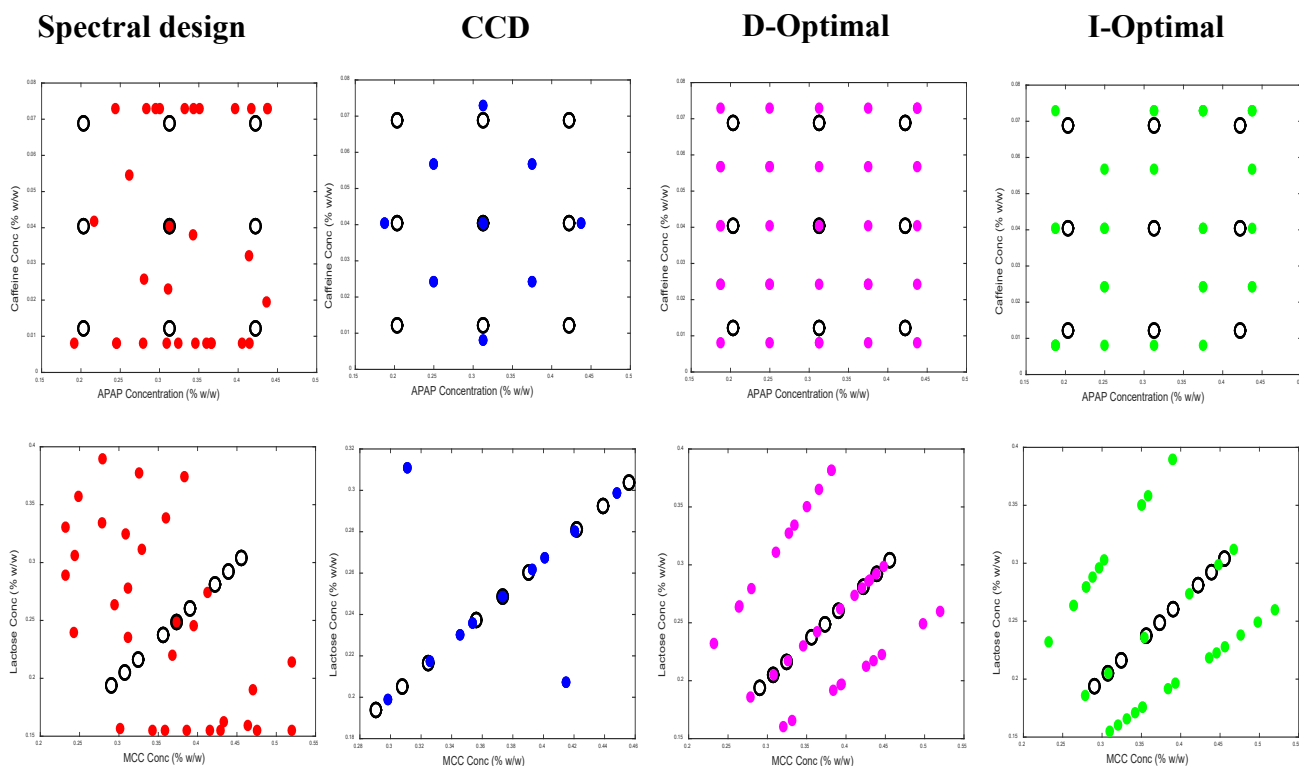


Figure 3-15. Tablet compositions of the spectral and traditional calibration designs (filled circles), and the test set (open circles)

3.3.2.3 Calibration design comparison between optimal and spectral calibration sets

Figure 3-15 shows the tablet compositions for the spectral and traditional designs along with the test set compositions. The CCD design had maximum resemblance with the test set design. The other designs have wider range of excipient variation compared to the CCD and test design. The wide excipient range in the spectral design and traditional optimal designs (D-Optimal and I-Optimal) were incorporated to maximize the spectral variance and concentration variance, respectively.

3.3.2.3.1 Prediction performance

The calibration models from the optimal designs and spectral design were used to predict the concentration of Acetaminophen and Caffeine in the prediction set.

3.3.2.3.1.1 Acetaminophen prediction

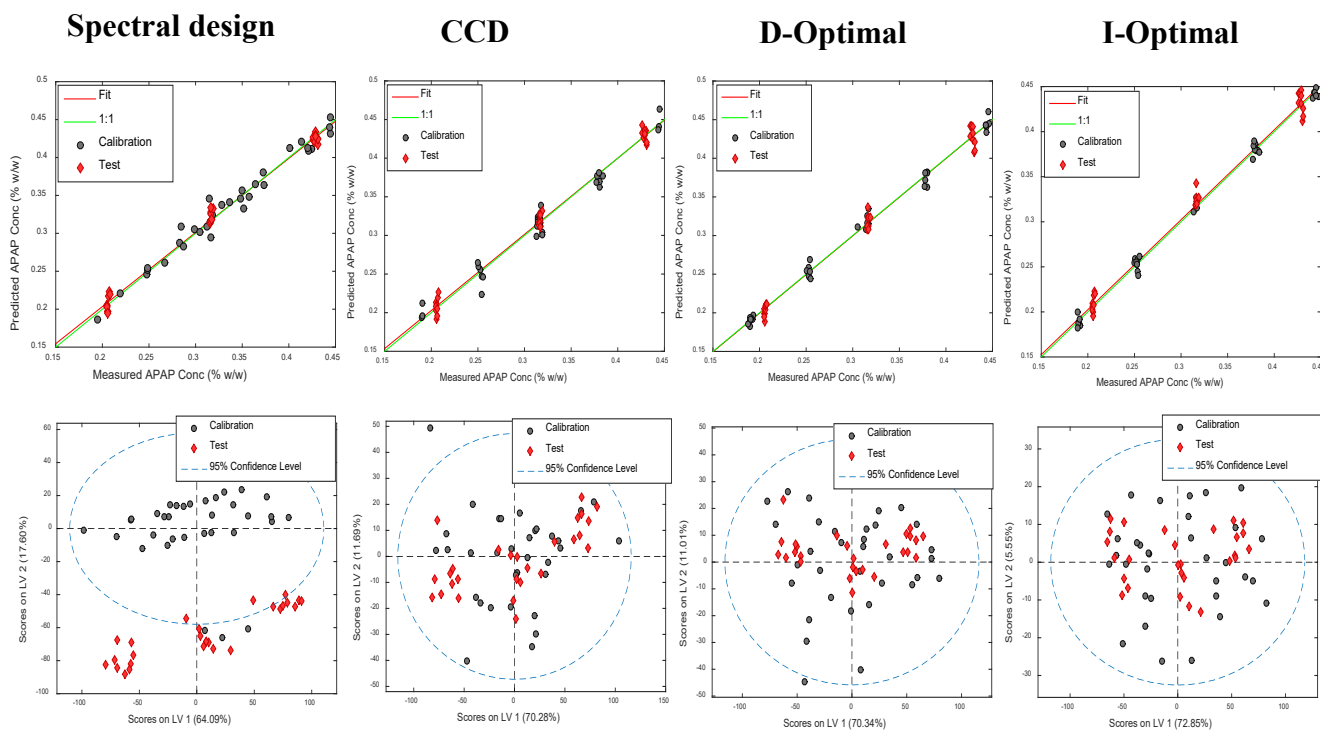


Figure 3-16. Calibration model performance of the spectral design and traditional designs during Acetaminophen prediction.

Figure 3-16 shows the calibration performance of the traditional and spectral calibration designs for Acetaminophen prediction in the calibration and test set. The traditional calibration sets (CCD, D-Optimal and I-Optimal) and test set had similar spectral variance as indicated by their similar projections on the first two loading vectors. For the spectral design, the score plot shows similar projections of the calibration and test set on LV1 and different projections on the LV2. Large excipient variation in the spectral design could contribute to the score differences. However, this difference in the scores on LV2 did not affect the prediction performance due to its small weight coefficients and minimal contribution to the regression vector and prediction.

The ascending order of the prediction error of Acetaminophen (RMSEP) was Spectral design < CCD < D-Optimal < I-Optimal.

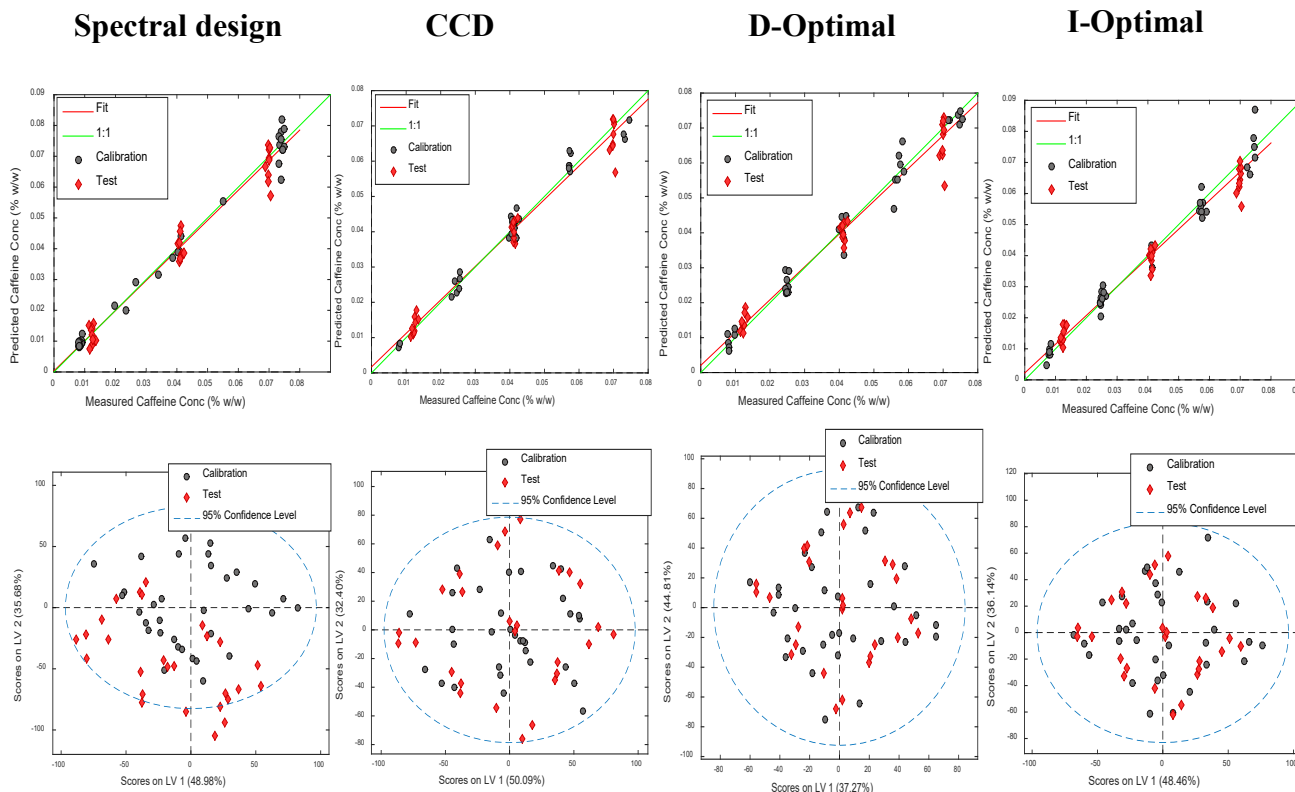


Figure 3-17. Calibration model performance of the spectral design and traditional designs during Caffeine prediction.

3.3.2.3.1.2 Caffeine prediction

Figure 3-17 shows the calibration performance of the traditional and spectral calibration designs for Caffeine prediction in the calibration and test set. All the calibration sets and the test set had similar spectral variance as indicated by their similar projections on the first two loading vectors. The score ranges were similar in both latent variable directions as opposed to the scenario in the Acetaminophen prediction model. This effect was attributed to the similar range of APIs (Acetaminophen and Caffeine) concentrations between the calibration and test sets. The first two

loading vectors primarily explained the direction of variation in Caffeine concentrations even at low dose, due to maximizing the covariance between spectra and Caffeine concentrations. These loading vectors also explained the direction of variation in Acetaminophen concentrations due to its unique spectral feature and relatively large dose. The ascending order of the prediction error of Caffeine (RMSEP) was CCD < Spectral design < D-Optimal < I-Optimal.

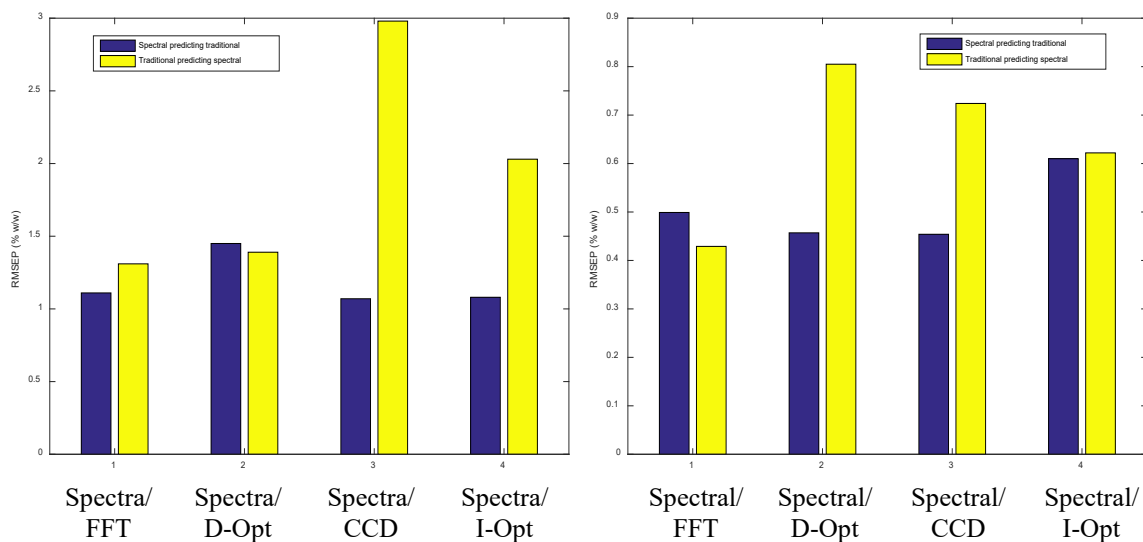


Figure 3-18. Cross design performance of acetaminophen model (left) and caffeine model (right)

3.3.2.3.1.3 *Assessment of model robustness*

Model robustness was assessed by cross design prediction. The calibration model developed from spectral calibration set was found to be more robust than the calibration models developed from optimal calibration sets. The robustness between spectral and full factorial calibration set was found to be similar. Figure 3-18 shows the results of the cross design prediction for both Acetaminophen (left) and Caffeine (right). The blue bars indicate the prediction errors of the spectral calibration model while predicting API in traditional calibration sets. The yellow bars indicate the prediction errors of the traditional calibration model while predicting API in spectral calibration. The spectral calibration model showed prediction errors similar to the full factorial calibration model in spite of having fewer sample number (33 vs 225). The full factorial calibration model was trained with larger sample size resulting in a robust calibration model. The prediction error was relatively low when a different test structure (spectral calibration set) was presented in front of the calibration model. The spectral calibration set was robust due to its wide range of spectral variation built into the calibration set. The CCD and I-optimal calibration models showed poor prediction performance for both Acetaminophen and Caffeine when a different test structure (spectral calibration set) was presented. D-optimal calibration model showed poor prediction performance while predicting Caffeine concentration in spectral calibration set. The spectral calibration model showed reasonable prediction performance while predicting Acetaminophen and Caffeine in all of the traditional calibration sets (Full factorial (FFT), CCD, I-optimal and D-optimal) indicating model robustness against different types of test set structure as shown in Figure 3-18.

3.3.2.3.2 Discussion

The main objective of this project was to develop and test a spectral design method and compare it with traditional calibration design methods for NIR spectroscopic multivariate calibration for multiple API system. Table 3-6 provides the calibration performances of the spectral and traditional designs for Acetaminophen and Caffeine. Spectral design and CCD were found to outperform the optimal designs in terms of prediction performance. The use of a traditional design in concentration space is a common practice for calibration sample preparation. However, this practice is inefficient and provides redundant and deleterious information to the calibration model as seen in the case of D-Optimal and I-Optimal design.

	Acetaminophen prediction				Caffeine prediction			
	Spectral	CCD	D-Opt	I-Opt	Spectral	CCD	D-Opt	I-Opt
RMSEC	1.092	1.187	0.758	0.611	0.260	0.286	0.335	0.368
RMSECV	1.336	1.381	1.004	0.882	0.473	0.345	0.456	0.513
RMSEP	0.832	0.893	1.050	1.110	0.410	0.379	0.463	0.481
Bias	0.115	0.259	0.085	0.484	0.232	0.090	0.084	0.174
R ² Cal	0.973	0.967	0.993	0.995	0.991	0.971	0.977	0.974
R ² Pred	0.991	0.991	0.987	0.988	0.979	0.977	0.967	0.975

Table 3-6. Calibration performance of spectral and traditional calibration designs

The optimality in the concentration space does not guarantee optimality in the spectral space. This effect is exacerbated for the tablets containing multiple API. In multiple API formulation, two individual APIs might have respective optimal design spaces that are very different from each other. However, in the optimal concentration based design, the optimality is defined in a single compositional space and same weightages are used for all chemical components including APIs and excipients to generate the optimality criterion (e.g. levels 1,2,3... to calculate D-optimal or I-optimal criterion). The resultant design points of concentration based optimal design are only indicative of optimal compositional structure but lack critical information specific to APIs.

The improved performance of the CCD calibration set compared to the other optimal calibration sets emphasizes the importance of similarity between calibration and prediction set structures. The calibration model provides improved prediction performance when it encounters known information in the prediction set. The CCD calibration model was trained on the information that was available in the prediction set, resulting in improved performance. However, such similarity is not always guaranteed. The calibration model from CCD calibration set showed poor prediction performance when a different test structure (spectral calibration set) was presented during cross design prediction for both Acetaminophen and Caffeine. Although the recommendation is to include all anticipated variations into the calibration set, there is always a possibility for the calibration set to encounter new information in the prediction set during model life cycle. It is advantageous to practice robust modelling technique instead of anticipating similarity between calibration and prediction set during calibration development.

3.3.3 Conclusion

A technique for designing a calibration set for NIR spectroscopic calibration method for tablets containing multiple API was described and tested. The pure component spectral information of the tablet was used to develop a spectral design method to design the critical calibration samples necessary for desired prediction performance. The spectral design strategy was found to be efficient in terms of sample requirements compared to the traditional full factorial technique. A reduced number of calibration samples yielded a model of similar predictive capabilities for both APIs. The spectral design strategy was found to provide better performance compared to the traditional optimal designs such as D-optimal and I-optimal designs. The spectral design strategy was found to provide similar prediction performance compared to the CCD design. Overall, the spectral design approach offers an efficient enhancement for the NIR calibration development for pharmaceutical tablets.

This approach has the potential to be useful for other tablet formulations and spectroscopic techniques. Infrared, Raman and THz spectroscopy have been used frequently to develop quantitative method for pharmaceutical tablets. Spectral design strategy can be useful to minimize the calibration sample requirements for such techniques.

PLS modeling technique was used in all comparative studies. Selecting the appropriate number of loading vector is a critical step during PLS model development. Appropriate number of loading vector is critical to define the model variance space. The current strategies for selecting the appropriate number of loading vector include cross-validation and total variance techniques. However, these techniques were unable to select the appropriate number of loading vector for traditional PLS model in section 3.2.1 for single API formulation. The poor model performance was due to inappropriate number of loading vector in the PLS model. A new strategy was

developed to select the appropriate number of loading vector based on weight co-efficient technique. This strategy was tested on multiple NIR dataset. The details of the new method and its performance are provided in the next chapter.

An improved prediction performance with lower RMSEPs was achieved from the traditional PLS model for single API formulation following the new method of loading vector selection (RMSEP2: 2.21 vs 4.68 and RMSEP3: 1.43 vs 5.05). The improved performance was found to be similar to the spectral design calibration model (RMSEP2: 2.21 vs 2.21 and RMSEP3: 1.43 vs 1.68). Spectral design was considered to be efficient considering its fewer sample requirement (11 vs 45) compared to the traditional calibration design.

Selecting the appropriate number of loading vector was critical to find the optimum calibration models and compare the respective calibration strategies. The weight based co-efficient technique was found to be an efficient alternative to the current loading vector selection techniques. The following chapter discusses this technique in further details.

Copyright ©: Part of this chapter has been reprinted from [175]. Copyright clearance is provided at the end the dissertation.

4 Chapter 4: Selecting loading vector for PLS model during quantitative analysis of tablets using NIRS

4.1 Introduction

NIR spectroscopy is a well-established analytical tool in the pharmaceutical industry for quantitative analysis of tablets [1-3, 5]. It is fast, non-destructive and requires little or no sample preparation. However, successful implementation of this technique requires analytical insight and chemometric tools to analyze the complex dataset. The broad overtones and combination vibrations in the NIR region are very difficult to assign to a specific chemical feature. Besides, NIR spectra are very sensitive to both physical and chemical properties of the samples [47, 135, 178, 179]. A quantitative model developed to extract the chemical information from the tablet spectrum is often affected by different sources of variations such as physical, chemical, scale, instrumental and environmental variations during tablet analysis. Batch-to-batch variation in the raw material properties such as particle size distribution of the granules and excipients, different vendors of the same excipient leading lot to lot variation, polymorphic conversion of the API and excipients during the manufacturing steps, all serve as potential sources of physico-chemical variation in the samples affecting NIR model performance. Environmental variations such as differences in relative humidity between different batches also affect NIR model performance. Difference in the scales of tablet preparation is a very common source of variation in the NIR spectra. It is quite possible for the analyst to develop a calibration model at one scale during the product development phase and intent to predict samples throughout the product life cycles from a different scale. A robust NIR model is critical to encounter such scale variations. In addition to table manufacturing scale, relative tablet density of the tablet have significant impact on NIR spectra [112]. Variation in the tablet density serves as a potential source of spectral variance. It is

also possible for the API in tablets to undergo degradation and produce degradants during its shelf life. The presence of a new chemical entity such as a degradant in the sample, can affect NIR model performance. Similar spectral features of the API and degradant can have detrimental effect on model selectivity. All these types of variation have the potentials to affect spectral shapes and thereby NIR model performance. So, model robustness has long been discussed as a critical merit during quantitative analysis using NIR spectroscopy [38, 54].

Special pre-processing techniques and multivariate calibration methods are often employed to develop robust modeling technique and extract desired information from the NIR spectra [180-182]. Besides, the analyst seeks to include as much variance as possible in the calibration step to train the model against all possible variances within defined scopes of drug product, NIR instrument and measurement environment that are forthcoming [183]. While, it is often impossible to forecast all the possible variances at the outset of calibration development, updating a calibration model with the newly encountered variance is a well-established practice [184]. Calibration update is a slow process and requires re-validation. Moreover, it is often difficult to get modification of an already approved method from the regulatory agencies [185]. So, it is desired to undertake the most robust modeling approach at the outset of calibration.

Partial Least Squares (PLS) is a widely used chemometric modeling technique used for quantitative analysis. PLS is used to develop a quantitative model in the calibration step that can be used to predict any future samples. A critical step during PLS model development is the selection of optimum number of loading vectors (model components) [186]. A parsimonious approach is frequently undertaken for this purpose. The minimum number of loading vector with acceptable model performance is selected to mitigate the model complexity and prevent over

fitting of the calibration model. Cross-validation technique is mainly employed to indicate the model over fitting and select the optimum number of loading vectors for the model development [187]. In addition to the cross-validation error, the amount of spectral and concentration variance explained by each loading vector is also considered during the selection process.

There are various cross-validation techniques including random subset, venetian blinds, contiguous blocks etc. However, all these techniques describe the error rate only in the calibration subspace. These techniques cannot anticipate model predictive ability in any future data set bringing new information to the calibration. The current techniques of cross-validation may fail to indicate the optimum number of loading vector to develop the most robust model at the outset of calibration development. In this study, a new method using the weight coefficient of each loading vector is proposed during calibration development of a pharmaceutical tablet. This method was tested to develop a calibration model robust against scale variation, environmental variation, physico-chemical variation, chemical variation and raw material variation. The robustness of the calibration model was compared with a traditionally built calibration model using current cross-validation techniques. The new weight coefficient based method was demonstrated to be more efficient than the current cross-validation techniques to develop the most robust model at the outset of calibration development.

4.2 Theory

A quantitative spectroscopic calibration model is developed to quantify physico-chemical properties of a sample from its respective spectrum using regression model. For a system with “m” samples and “n” spectral variables (wavelengths), the regression model is as follows,

$$\hat{\mathbf{y}}_{m \times 1} = \mathbf{X}_{m \times n} * \mathbf{b}_{n \times 1} \quad (4.1)$$

Here, “ \hat{y} ” is the predicted physico-chemical property, “ \mathbf{X} ” is the spectral data matrix and “ \mathbf{b} ” is the regression vector. Concentration will be considered as an example of physico-chemical property in rest of the chapter.

In PLS, spectral data (\mathbf{X}) and concentration data (\mathbf{y}) are decomposed into scores and loading vectors. The loading vectors describe the covariance structure between the spectra and concentration. The scores serve as a metric for spectral variance of the individual sample in the direction of the respective loading vectors. The details of the PLS theory is discussed elsewhere [188].

$$\mathbf{X}_{m \times n} = \mathbf{T}_{m \times k} * \mathbf{P}'_{n \times k} \quad (4.2)$$

$$\mathbf{Y}_{m \times u} = \mathbf{U}_{m \times k} * \mathbf{Q}'_{u \times k} \quad (4.3)$$

Here, “ \mathbf{T} ” and “ \mathbf{U} ” are the scores of the spectral matrix “ \mathbf{X} ” and concentration matrix “ \mathbf{Y} ” respectively and “ k ” is the number of loading vectors and weight vectors to be included into the model. “ \mathbf{P} ” and “ \mathbf{Q} ” are the loading vectors for “ \mathbf{X} ” and “ \mathbf{Y} ” as well. For a single predictor, the concentration matrix “ \mathbf{Y} ” is a vector and does not need to be decomposed.

NIPALS is a widely used algorithm in PLS model development [189]. In NIPALS, the loadings and scores are iteratively solved. The first loading “ \mathbf{P} ” indicates the primary direction of covariance between spectra and concentrations, and the first set of scores “ \mathbf{T} ” indicates the projections of samples into that “ \mathbf{P} ” direction. The second loading vector indicates the direction of maximum covariance between the residual spectra and residual concentration after spectral reconstruction with the first loading vector. The third loading vector indicates the same after spectral reconstruction with the first and second loading vector. Once the number of loading

vector is fixed, the regression model of Eq (4.1) is developed by calculating the regression vector as follows,[190]

$$\mathbf{b}_{n*1} = \mathbf{W}_{n*k} * (\mathbf{P}'_{k*n} * \mathbf{W}_{n*k})^{-1} * (\mathbf{T}'_{k*m} * \mathbf{T}_{m*k})^{-1} * \mathbf{T}'_{k*m} * \mathbf{y}_{m*1} \quad (4.4)$$

Here, “**W**” is the weight vectors expressing both the “positive correlation” between spectral matrix “**X**” and concentration matrix “**Y**” and the “compensation correlation” needed to predict “**Y**” from “**X**” clear from the secondary variation in “**X**” [191]. In the PLS toolbox from Eigenvector Research Inc., the weight vector “**W**” is expressed as “**W***”,

$$\mathbf{W}_{n*k}^* = \mathbf{W}_{n*k} * (\mathbf{P}'_{k*n} * \mathbf{W}_{n*k})^{-1} \quad (4.5)$$

which provides the same information as “**W**”. So, Eq 4.5 can be written as,

$$\mathbf{b}_{n*1} = \mathbf{W}_{n*k}^* * (\mathbf{T}'_{k*m} * \mathbf{T}_{m*k})^{-1} * \mathbf{T}'_{k*m} * \mathbf{y}_{m*1} \quad (4.6)$$

$$\mathbf{b}_{n*1} = \mathbf{W}_{n*k}^* * \mathbf{s}_{k*1} \quad (4.7)$$

Here, “**s**” is considered as the weight coefficients and “**k**” is the number of loading vector to be included in the PLS model. The same weight coefficient is also used in the prediction equation as follows (replacing “**b**” in Eq 4.1,

$$\hat{\mathbf{y}}_{m*1} = \mathbf{T}_{m*k} * \mathbf{s}_{k*1} \quad (4.8)$$

In a PLS model, each weight vector, its associated loading vector and scores are associated with a weight coefficient which can serve as a metric for its contribution in regression vector calculation (Eq (4.7)), as well as in prediction equation (Eq (4.8)).

Optimum selection of the loading vector number “k” is critical for model performance.

Traditional methods for selecting “k” involve either one of the two approaches or both:

1. Root Mean Squared Error of Calibration (RMSEC) and Root Mean Squared Error of Cross Validation (RMSECV) at each loading vector and weight vector inclusion into the model.
2. Variance explained by each loading vector during the model development.

However, these two techniques have the potential to fail in selecting the appropriate number of loading vectors to develop a robust PLS model at the outset of calibration development. The new method uses the weight coefficients “s” as the basis for selecting appropriate number of loading vector “k” to be included into the model. It uses the weight coefficients ‘s’ for isolating loading vectors that explain a small percent of the spectral and concentration variance while having high influence (high weight coefficient) in the regression vector calculation and prediction equation. Inclusion of such loading vectors can be detrimental to the model performance, because such inclusion tends to over fit the model towards a direction that explains minimal spectral and concentration variations. Overfitting model towards irrelevant information can affect model specificity and prediction performance.

According to the new weight coefficient based method, any weight coefficient larger than the coefficient of the maximum variance explaining loading vector (often the first one) was considered inappropriately high and PLS model was developed with the preceding loading vectors. If ‘nth’ weight co-efficient is higher than the co-efficient of the maximum variance explaining loading vector, the optimum number of loading vectors was considered to be “n-1” according to the weight coefficient based method.

In the traditional method, PLS models were developed following the current two techniques for selecting the optimum number of loading vector. In the new weight coefficient based method, PLS model was developed using the weight coefficient along with the associated percent of variance explained by each loading vector. The performances of the new and old methods were compared in selecting the optimum number of loading vectors and developing the most robust PLS model at the outset of calibration development.

4.3 Material and Methods

Five different datasets were used to test the NIR model robustness against different types of critical variations including scale variation, environmental variation, physical variation, chemical variation and raw material variation. The details of the dataset are given below

4.3.1 Scale Variation

The scale variation test was designed to analyze and compare model performance between the traditional method and new weight coefficient based method while predicting Acetaminophen concentrations in tablets prepared at different scales. The calibration set was prepared at laboratory scale while the prediction set was prepared at manufacturing scale. The calibration design was a 5 by 3 level, 2-factor full factorial design to make 45 tablets. The prediction set was prepared over consecutive eight weeks at the label claim. 20 tablets collected from every week's run resulted in a total of 160 tablets in prediction set. These samples are the same samples as traditional calibration set and prediction set 3 for single API. The details of the experimental method are described in section 3.2.1.1 (calibration set) and 3.2.1.5 (prediction set).

4.3.2 Environmental Variation

The environmental variation test was designed to analyze and compare model performance between the traditional method and new weight coefficient based method while predicting Acetaminophen concentrations in tablets prepared in different environmental conditions. The calibration set was prepared in the winter (November, 2012) whereas the prediction set was prepared in the summer in 3 consecutive months (May, June and July 2013).

The same calibration set mentioned in the scale variation test was used for calibration model development. A 3 by 1 level full factorial design was used to generate the monthly prediction sets. The factors of the design were kept same as in the calibration and the levels were inside the calibration range. A set of 20 tablets at each design point resulted in a total of 60 tablets for each month's prediction set. The tablets were prepared at the manufacturing scale in the same fashion described in the scale variation section in section 4.3.1.

4.3.3 Physical Variation (Density)

The density variation test was designed to analyze and compare model performance between the traditional method and new weight coefficient based method while predicting Theophylline concentrations in tablets of different densities. The capability of these two methods in selecting the optimum number of loading vector were compared during calibration development and calibration update. A previously published dataset was used for this purpose [112].

A fully balanced, quaternary mixture design comprising of Theophylline anhydrous, Lactose 316 Fast Flo Monohydrate, Microcrystalline cellulose and soluble starch was generated. A set of 29 design points was chosen to cover a wide range of all constituents. At each design point, five

tablets were compressed at five compaction pressures (67.0, 117.3, 167.6, 217.8, and 268.1 MPa) on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) using a 13-mm die and flat-faced punches. One additional tablet was also compressed at each design point, with a compaction pressure chosen pseudo-randomly from the five levels resulted in a total of 174 tablets. This calibration set was divided into five groups (Cal 01 to Cal 05) based on the compaction pressure levels. The first group contained samples compacted at 67.0 MPa, while the second group contained samples compacted at both 67.0 and 117.3 MPa, and subsequently, the fifth group contained samples compacted at all five compaction pressures. The primary calibration was developed only using samples compacted at 67.0 MPa and then the calibration set was updated by samples compacted at different compaction levels.

The same quaternary design was used to create a test set. At each design point, two tablets were compressed at two randomly chosen compaction pressure out of five (67.0, 117.3, 167.6, 217.8, and 268.1 MPa) using a Carver Automatic Tablet Press. A total of 58 samples were prepared in the test set and used for assessing the prediction performance of all the models prepared in the calibration step. Tablet density was measured for each compact and solid fraction was calculated. A wide range of solid fraction (0.8 – 0.98) was reported along the range of the compaction forces [112].

The performance of the traditional method and new weight coefficient based method were compared in selecting the appropriate number of loading vectors for the models developed from these five calibration sets (primary and updated sets).

4.3.4 Chemical Variation (Degradant)

The chemical variation test was designed to analyze and compare the model performance between the traditional method and new weight coefficient based method while predicting Niacinamide concentration in tablets prepared with and without its degradant Niacin. The calibration set was prepared using samples containing Niacinamide and excipients including starch, Microcrystalline cellulose (MCC), Di-tab and Magnesium stearate. The prediction set was prepared using samples containing Niacinamide, Niacin and excipients including starch, Microcrystalline cellulose (MCC), Di-tab and Magnesium stearate.

The Niacinamide concentration was varied at 7 levels in the calibration set while keeping the excipients ratios (starch:MCC:Di-tab) constant. The concentration of Magnesium stearate was kept constant at 1%. The nominal target tablet weight was 500 mg.

At each design point, 5 gm samples mixture was prepared and placed in a 15 ml scintillation vial. The samples were mixed by placing the vial in a bin blender and rotating for 15 minutes. At each design point, three tablets were compressed from the same mixture on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at 5000 lb. using a 13-mm die and flat-faced punches.

A 3 by 3 levels, two-factor full factorial design was used to prepare the test set. The factors were Niacinamide concentration and Niacin concentration. At each design point, 3 tablets were compressed in the same fashion as mentioned earlier. Out of the total of 27 tablets, tablets from one design point were damaged leaving 24 tablets in the test set. The test set was subdivided into three sets based on the degradant (Niacin) concentration.

4.3.5 Raw Material Variation

The raw material variation test was performed on a previously published dataset to analyze and compare model performance between the traditional method and new weight coefficient based method while predicting Theophylline concentration in tablets prepared with different raw material properties [164]. The calibration set was prepared with a specific polymorph of API as Theophylline anhydrous, a specific particle size distribution of Lactose as 100 microns and a specific source of starch as EMD Chemicals. The prediction set was prepared using a different polymorph of API as Theophylline monohydrate, different particle size distribution of Lactose as 50 micron and different vendor of starch as Acros Organics.

A fully balanced, quaternary mixture design comprising of Theophylline anhydrous, Lactose 316 Fast Flo Monohydrate, Microcrystalline cellulose and soluble starch was generated. A set of 29 design points were chosen to cover a wide range of all constituents. At each design point, 2 tablets were compressed at 2 randomly selected pressures out of five (67.0, 117.3, 167.6, 217.8 and 268.1 MPa) on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) using a 13-mm die and flat-faced punches. In total, 58 tablets were compressed for calibration set with a target weight of 800 mg.

A 2 by 2 by 3 levels, three-factor full factorial design was used to prepare the test set at a single composition inside the calibration range. The factors were selected as API polymorphs, sources of starch and Lactose particle size distributions to ensure different types of raw material variability in the test set. All the tablets were compressed at 167.6 MPa using the same setup as calibration. At each design point, 3 tablets were compressed, resulted in a total of 36 tablets.

The test set was subdivided into three sets based on the difference with calibration set in raw material properties. The first set was comprised of all the samples containing Theophylline monohydrate, second set comprised of all the samples containing Lactose of 50 microns particle size distribution and third set comprised of all the samples containing starch from vendor 2 (Acron Organics).

4.3.6 Spectral collection

NIR reflectance measurements for both sides of each compact were collected using a bench top scanning monochromator instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc., Laurel, MD, USA). Spectra were collected over the wavelength range of 400 – 2,500 nm at 0.5-nm increments, averaging 32 co-adds per spectrum. Spectra corresponding to each side of a compact were averaged to give one representative spectrum per compact.

4.3.7 Modeling strategy

All calculations were performed with MATLAB 2015a (The Mathworks, Natick, MA, USA) equipped with a PLS Toolbox v. 7.9.4 (Eigenvector Research Inc., Wenatchee, WA, USA).

4.3.8 Cross validation

PLS Toolbox was used to implement 3 different cross-validation techniques as random subset, contiguous block and venetian blind. For each technique, the data set was divided into 3 different ways to remove the aliasing effect between data subset and cross-validation error. The details of the cross-validation technique using PLS toolbox can be found elsewhere [192].

4.3.9 Model evaluation and design comparison parameters

The root mean squared error (RMSE) was used to evaluate model predictive performances in the calibration (RMSEC), cross-validation (RMSECV) and prediction (RMSEP) sets.

$$\text{RMSE} = \sqrt{\frac{\sum_1^n (\hat{y} - y)^2}{n}} ; \quad (4.9)$$

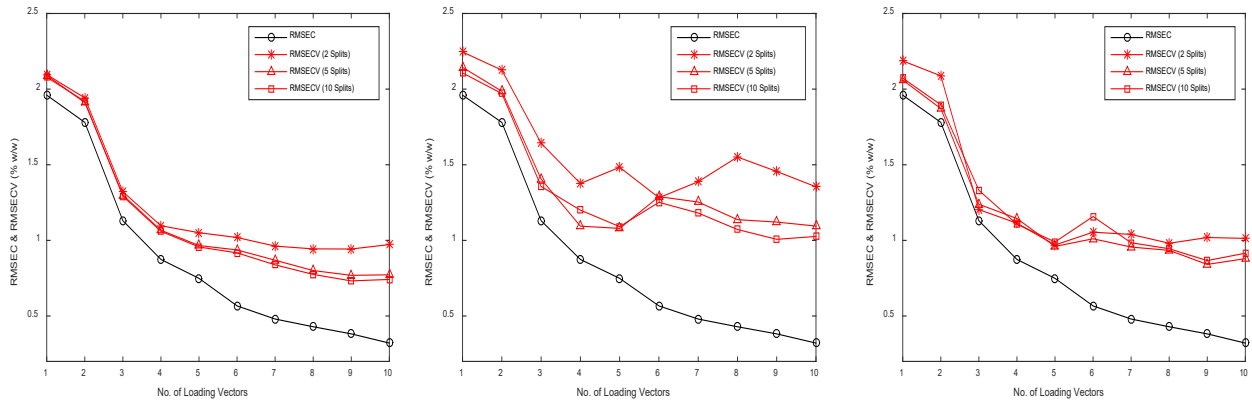


Figure 4-1. RMSECs and random subset (left), contiguous block (middle) and venetian blind (right) RMSECVs during calibration development for scale variations.

4.4 Results

4.4.1 Scale Variation

Figure 4-1 shows the RMSECs and RMSCEVs for loading vectors 1 to 10 calculated by different cross-validation techniques such as random subset (left), contiguous block (middle) and venetian blinds (right). The no. of splits was also changed as the optimization parameter as shown in the figures. All the techniques (except contiguous block with 2 splits) showed a plateau between RMSEC and RMSECV at fifth loading vector and there was no significant reduction in cross-

validation errors after that. A PLS model was developed using five loading vectors as per current technique for selecting optimum number of loading vectors for model development.

Figure 4-2 shows the cumulative percent of variances explained after the addition of each loading vector into the model. Besides RMSECV, this metric was also used for selecting appropriate number of loading vectors. It was shown that, most of the variances in spectral matrix (X) and concentration matrix (Y) were explained by five loading vectors, indicating five as the optimum number of loading vectors for model development.

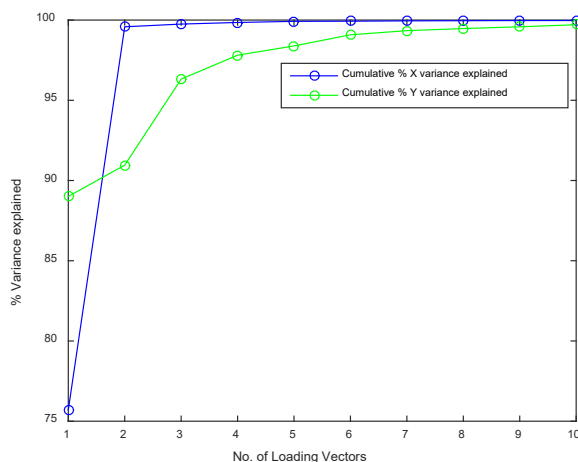


Figure 4-2. Cumulative variances explained at each loading vector during calibration development for scale variations.

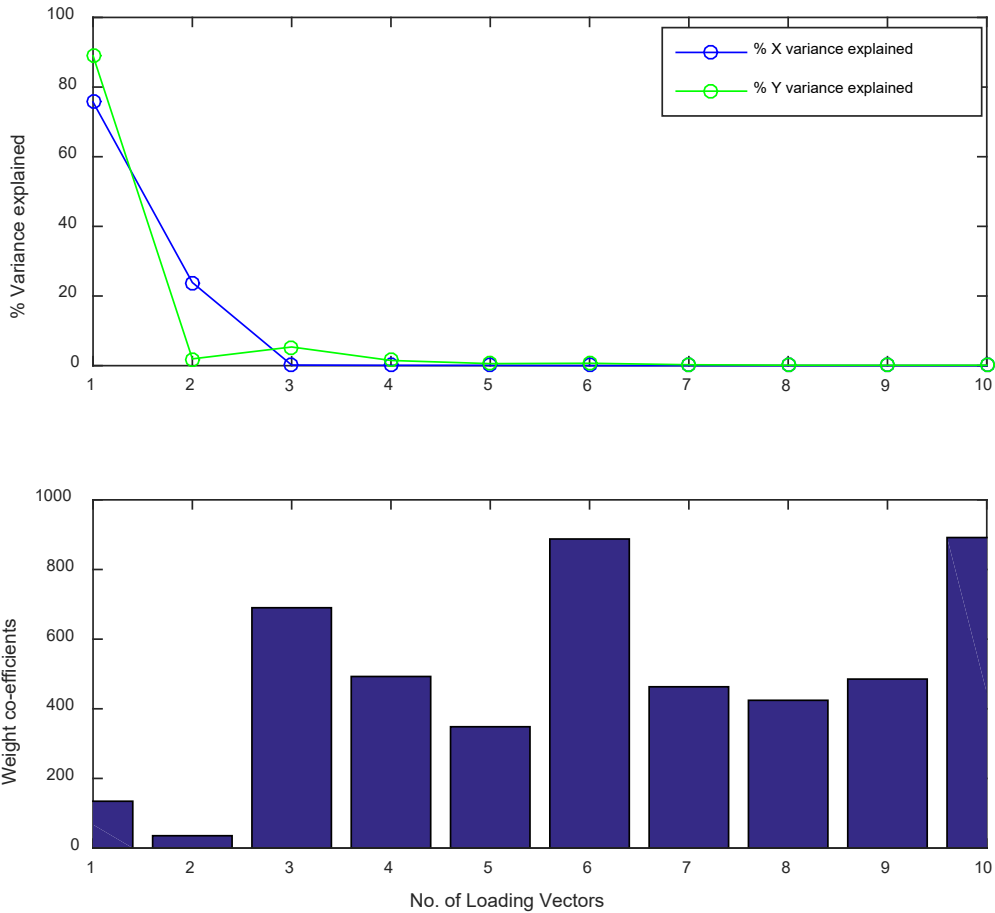


Figure 4-3. Weight co-efficient and percent of variances explained by each loading vector during calibration development for scale variations.

Figure 4-3 shows the percent of variances explained by each loading vector and its respective weight coefficient. It was found that, the third loading vector had a much higher weight coefficient than first two loading vectors. High weight coefficient of the third loading vector indicated its large influence on the regression vector calculation and prediction equation compared to the first and second loading vector. However, only a small portion of the total variances was explained by the third loading vector indicating that, inclusion of this loading vector could over fit model using irrelevant information and affect model performance in the

future samples. Two loading vectors were considered optimum according to the new weight coefficient based method and a PLS model was developed using first and second loading vector. Another PLS model was also developed with three loading vectors to analyze the effect of including third loading vector into the model. These three PLS models were developed on lab scale tablets and used to predict tablet set prepared at manufacturing scale for eight consecutive weeks. The RMSEPs served as a metric for the model performance and robustness.

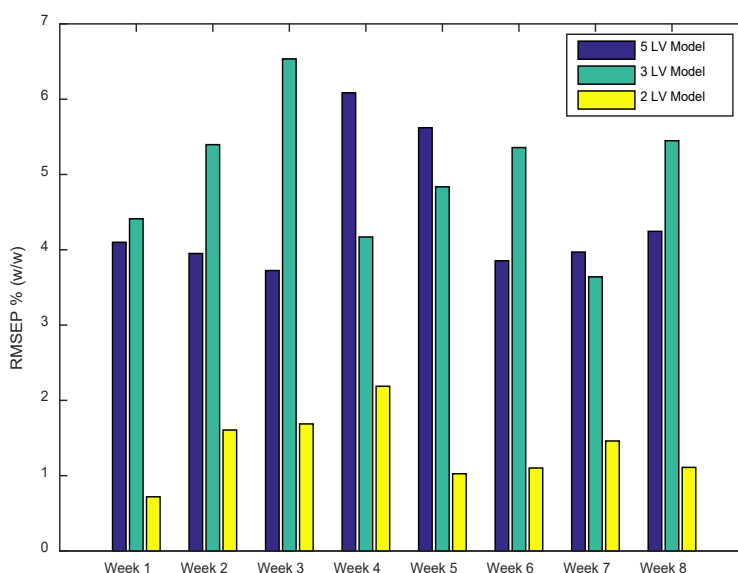


Figure 4-4. RMSEPs of models predicting Acetaminophen in weekly runs prepared at different scale.

Figure 4-4 shows the RMSEPs for models developed according to traditional method (5 LV model), new weight coefficient based method (2 LV model) and model showing the effect of third loading vector inclusion (3 LV model). The model developed using the new weight coefficient based method was found to be more robust against scale variation and more accurate through eight weeks compared to the traditional method. Inclusion of the third loading vector affected the mode performance as anticipated during model development.

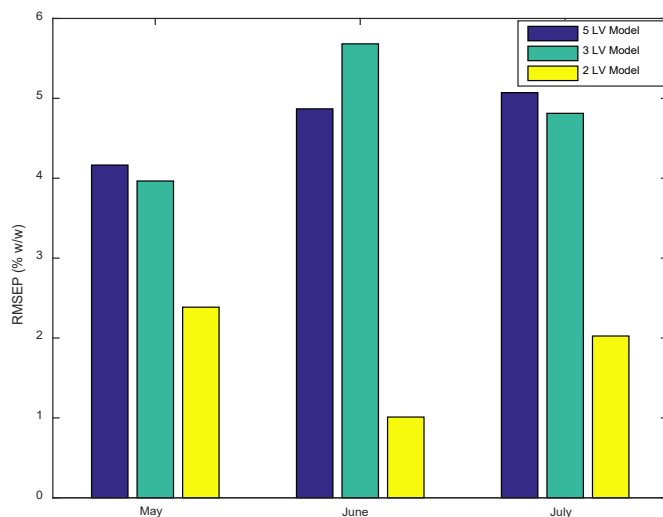


Figure 4-5. RMSEPs of models predicting Acetaminophen in monthly runs prepared at different environment.

4.4.2 Environmental Variation

The robustness of these calibration models were also tested against environmental variation. All of the three models were developed in November, 2012 and used to predict the tablets prepared in May, June and July 2013 at the manufacturing scale. Figure 4-5 shows the RMSEPs for all three models predicting the tablets from different months. The model developed according to the new weight coefficient based method (2 LV model) was found to be more robust and accurate compared to other two models.

4.4.3 Physical Variation (Density)

Model robustness was tested against density variation as the physical source of variation in NIR analysis. The primary calibration set contained samples made at a single compaction force and then the calibration set was sequentially updated using samples made at four different compaction forces.

A total of 5 calibration sets were prepared. For each calibration set, two PLS models were developed following the traditional and new weight coefficient based method, respectively. The calibration models were used to predict the test set containing samples made at five different compaction forces. This analysis compared the performance of the traditional method and new weight coefficient based method in selecting optimum number of loading vectors during calibration development and calibration update.

In the traditional method, the optimum number of loading vectors for each of the five models was selected based on RMSEC and RMSECV plots and the percent of variances explained by each loading vector. In the new weight coefficient based method, weight coefficient of each loading vector was used to select the optimum number of loading vectors to be used during model development.

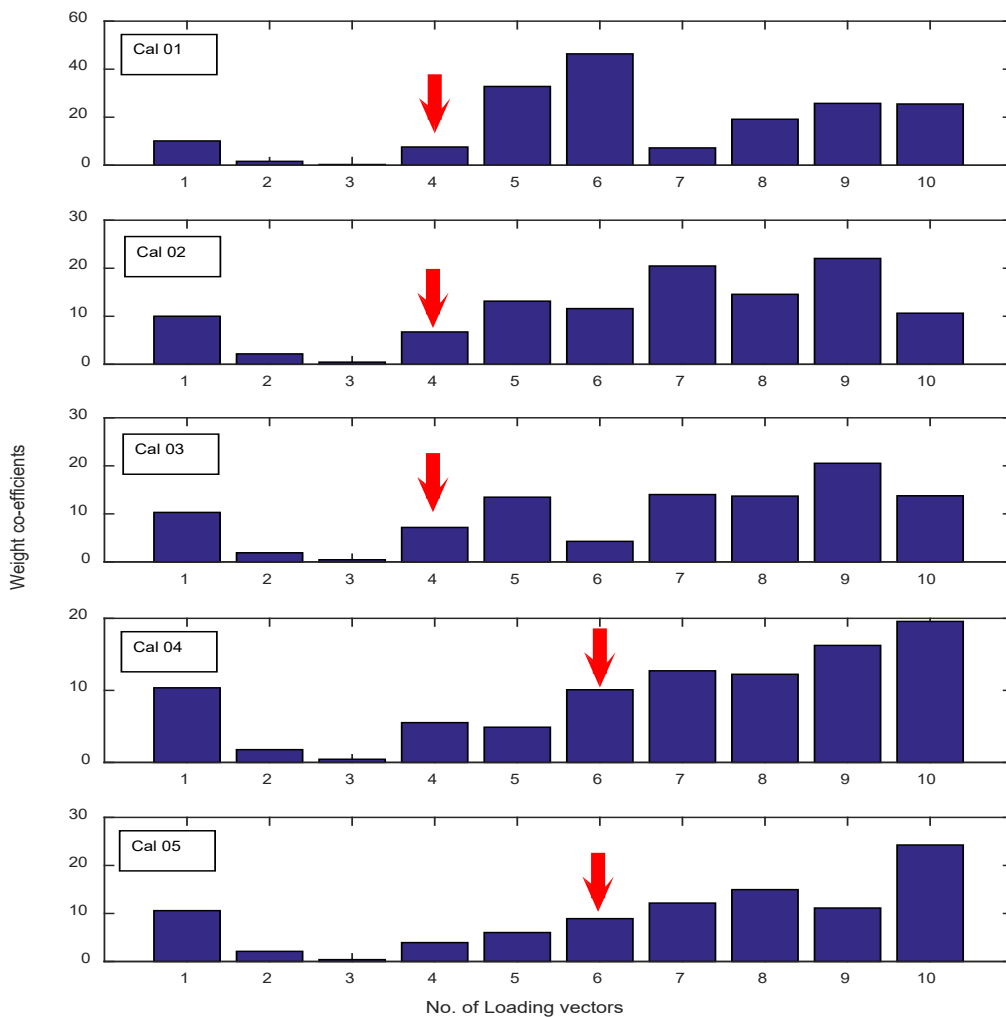


Figure 4-6. Weight co-efficient of each loading vector during calibration development for density variations.

Figure 4-6 shows the weight coefficient of each loading vector for five calibration models. The red arrows indicate the number of loading vectors used during model development following the new weight coefficient based method. For “Cal 01”, four loading vectors were used, because the fifth loading vector had higher weight coefficient than the previous four, indicating its large influence on the regression vector calculation and prediction equation with very little information

on the spectral and concentration variance. This approach was followed for the rest of the calibration sets (Cal 02 to Cal 05).

Calibration set	Sample no.	Sample Force Included (MPa) in Calibration Set	Loading Vector		RMSEP % (w/w)	
			Traditional Method	New Weight Coefficient Based Method	Traditional Method	New Weight Coefficient Based Method
Cal 01	35	67.0	7	4	3.48	2.82
Cal 02	70	67.0, 117.3	8	4	3.31	2.37
Cal 03	105	67.0, 117.3, 167.6	9	4	2.94	2.26
Cal 04	139	67.0, 117.3, 167.6, 217.8	5	6	2.39	2.26
Cal 05	174	67.0, 117.3, 167.6, 217.8, 268.1	6	6	2.23	2.23

Table 4-1. Calibration for Density variation

Table 4-1 provides a summary of the calibration sets, the respective number of loading vectors used by the traditional method and new weight coefficient based method and the RMSEPs of the respective models. Figure 4-7 shows the RMSEPs of the calibration models developed by the traditional method and new weight coefficient based method. The models developed by the new weight coefficient based method were found to be more robust and accurate compared to the models developed by the traditional method while predicting the same test set having density variation. For the global set “Cal 05”, that contained samples prepared at all compaction forces, both methods indicated the same optimum number of loading vector thereby providing equal RMSEP. It was found that, with a fewer samples, the model developed using the new weight coefficient based method was able to achieve similar prediction performance compared to the global model (new method for Cal 03 (105 samples) vs traditional method for Cal 05 (174

samples)). This indicated that, the new weight coefficient based method of model development has the potential to reduce the burden of calibration update.

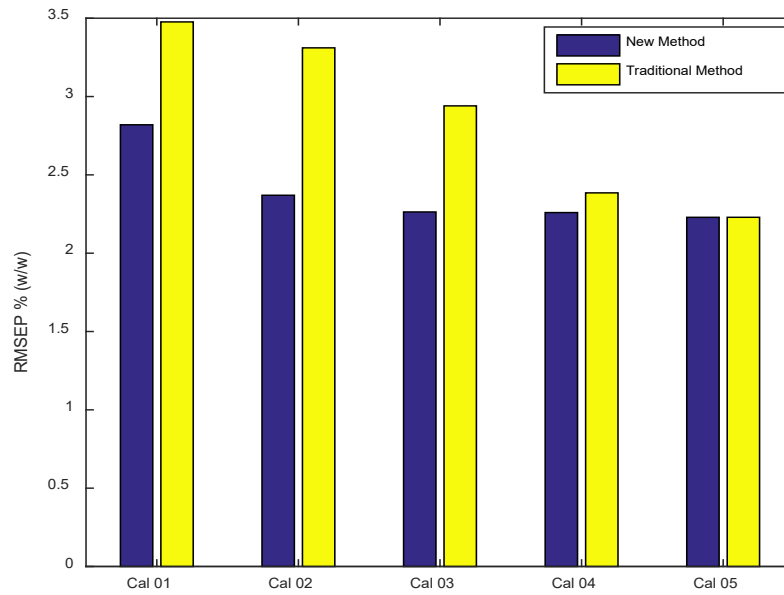


Figure 4-7. RMSEPs of models predicting Theophylline in tablets with density variations.

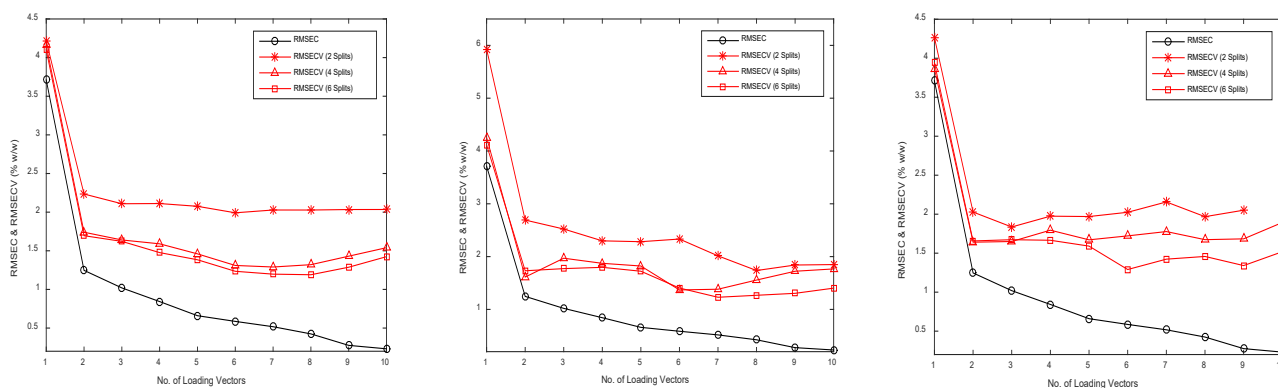


Figure 4-8. RMSECs and random subset (right), contiguous block (middle) and venetian blinds (left) RMSECVs during calibration development for chemical variations.

4.4.4 Chemical Variation (Degradant)

Model robustness against chemical variation was tested by introducing a new chemical entity in the prediction set. A degradant of the API was included in the prediction samples whereas the calibration model was developed using samples without the degradant. shows the RMSEC and RMSECVs of each loading vector calculated by different cross-validation techniques during calibration development. For all these techniques, RMSEC and RMSECV reached a plateau at second loading vector. A PLS model was developed using two loading vectors as per the traditional method of model development.

Figure 4-9 shows the cumulative percent of variances explained after the addition of each loading vector into the model. Five loading vectors explained most of the variances and no significant variance was added after that. Another PLS model was developed using five loading vectors as per the traditional method of model development.

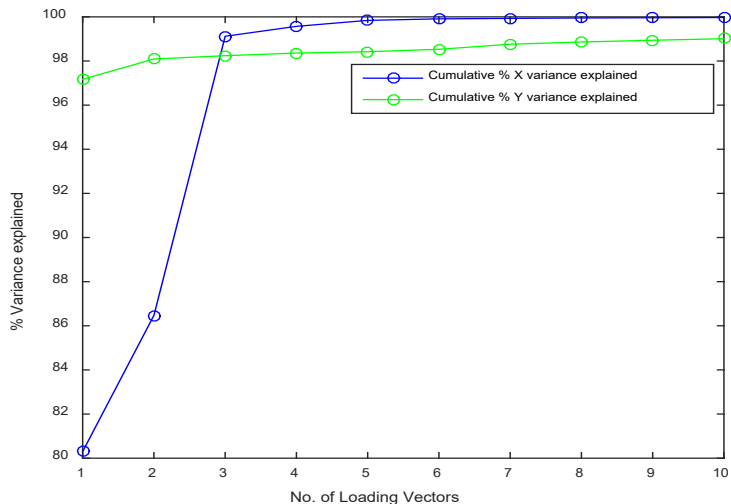


Figure 4-9. Cumulative variances explained at each loading vector during calibration development for chemical variations

Figure 4-10 shows percent of variances explained by each loading vector and its respective weight coefficient. It was found that, the second loading vector had higher weight coefficient than the first loading vector indicating its higher influence on the regression vector calculation and prediction equation. However, it explained a small portion of the X and Y variance compared to the first loading vector. Inclusion of the second loading vector was expected to over fit the model towards a direction that explained a small portion of the spectral and concentration variances. A third PLS model was developed using only one loading vector as per the new weight coefficient based method.

Two models developed by the traditional method and one model developed by the new weight coefficient based method were used to predict Niacinamide concentration in test samples containing different amount of Niacin. The RMSEPs served as a metric of the model performance and robustness in presence of chemical variation coming from the Niacin as a degradant of the API. Figure 4-11 shows the RMSEPs for all three models predicting

Niacinamide concentration in samples containing 2%, 5% and 10% w/w of Niacin. It was found that, for all concentrations of Niacin samples, the model developed by the new weight coefficient based method was more robust and accurate in comparison to the other two models developed by the traditional methods. As expected, all model performances deteriorated as the Niacin content increased in the test samples.

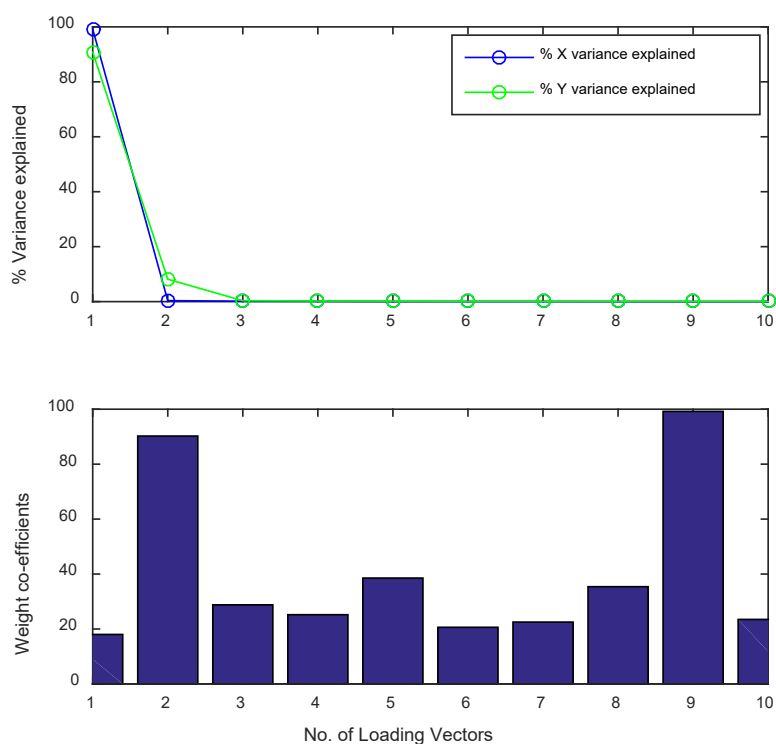


Figure 4-10. Weight co-efficient and percent of variances explained by each loading vector during calibration development for chemical variation

4.4.5 Raw Material Variation

Model robustness was tested against raw material variation. The calibration models were developed using the samples containing a specific polymorph of API as Theophylline anhydrous, a specific particle size distribution of Lactose as 100 microns and a specific vendor of starch as

EMD Chemicals. The models were used to predict the samples containing different polymorph of API, different particle size distribution of one excipient and different vendor of the other excipient.

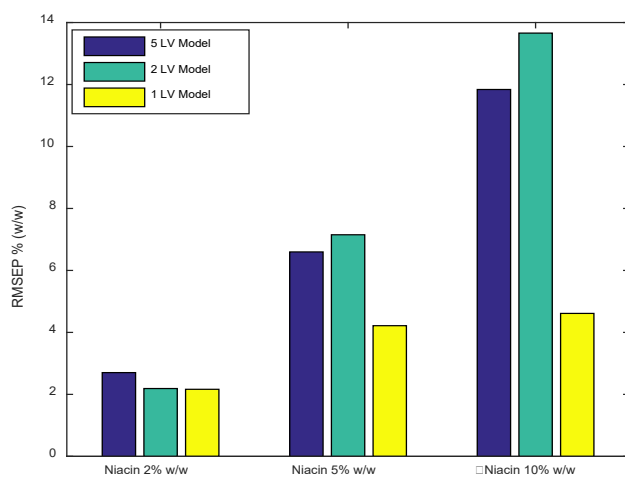


Figure 4-11. RMSEPs of models predicting Niacinamide in presence of chemical variations (Niacin)

In this analysis, both the traditional method using RMSEC and RMSECV plot and the new weight coefficient based method using weight coefficient indicated five loading vectors as the optimal number for model development. Figure 4-12 shows the percent of variances explained by each loading vector and its respective weight coefficient. Five loading vectors were selected for model development because the sixth loading vector had higher weight coefficient than the preceding loading vectors indicating its high influence on the regression vector calculation and prediction equation. Inclusion of such loading vector would over fit the model towards a direction that explained a very small portion of the spectral and concentration variance. Although fourth and fifth loading vectors did not explain a significant portion of spectral or concentration

variances individually, they were kept into the model due to their smaller weight coefficients compared to the first loading vector.

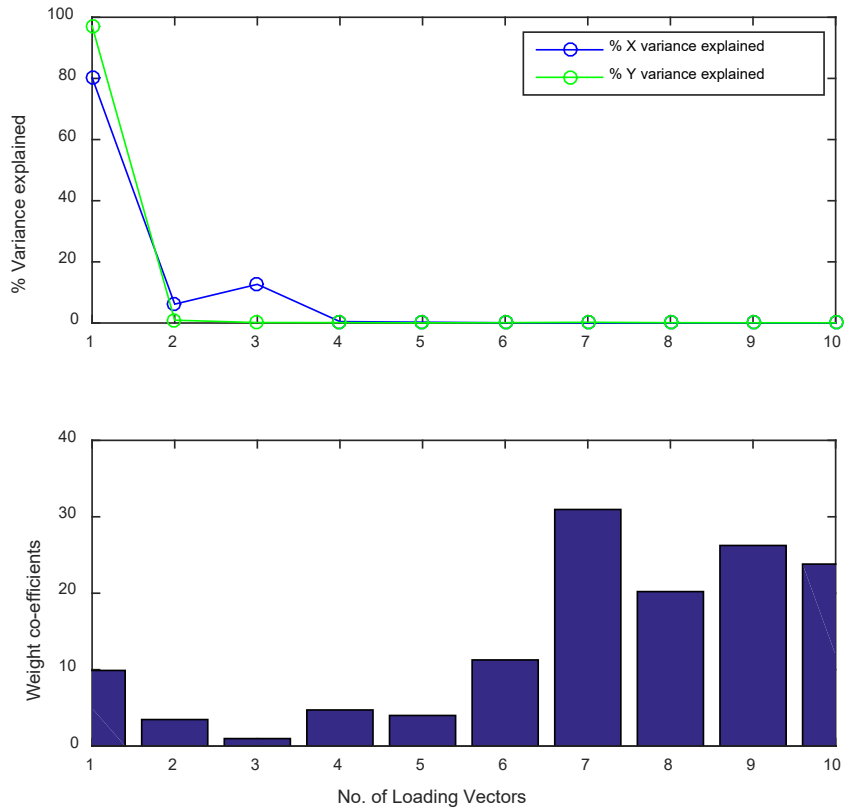


Figure 4-12. Weight co-efficient and percent of variances explained by each loading vector during calibration development for raw material variation

Figure 4-13 shows the cumulative percent of variances explained at each successive loading vector addition into the model. It was found that, after the third loading vector inclusion, no significant variance was explained with the addition of a new loading vector. So, three loading vectors were used for model development as per one of the traditional techniques.

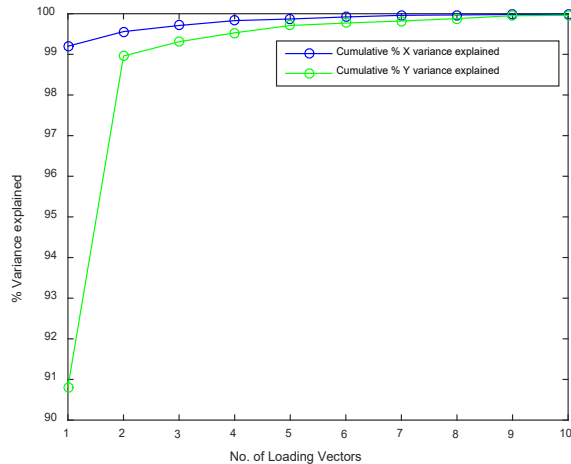


Figure 4-13. Cumulative variances explained at each loading vector during calibration development for raw material variation

Figure 4-14 shows the RMSEPs of the two models developed by one of the traditional methods (3 LV model) and the new weight coefficient based method (5 LV) for predicting Theophylline concentration in test set containing new variances coming from different polymorph of API (Theophylline monohydrate), different particle size distribution (Lactose 50 micron) and different source (Starch vendor 2).

4.5 Discussion

In the above case studies, the new weight coefficient based method was found to be more effective in selecting the optimum number of loading vectors and indicating the most robust PLS model compared to the current techniques of cross-validation and variance calculation. Cross validation provides the error rate at successive loading vectors, however, that error rate is specific to the calibration space. Although cross-validation indicates model performance as a function of loading vector and prevents over fitting of the model in the calibration space, it provides little or no information about the model performance on any future data set containing new information to the calibration. Thus, cross-validation often fails to aid in the development of

the most robust PLS model at the outset of calibration. Upon such failure, it is usually assumed that the working model is the best case in place and model update is necessary. This is often not the case as shown in this paper. Moreover, it is also shown that, calibration update might not even be necessary at some points (density variation) following the new weight coefficient based method. This can be helpful in saving time and cost associated with calibration model update.

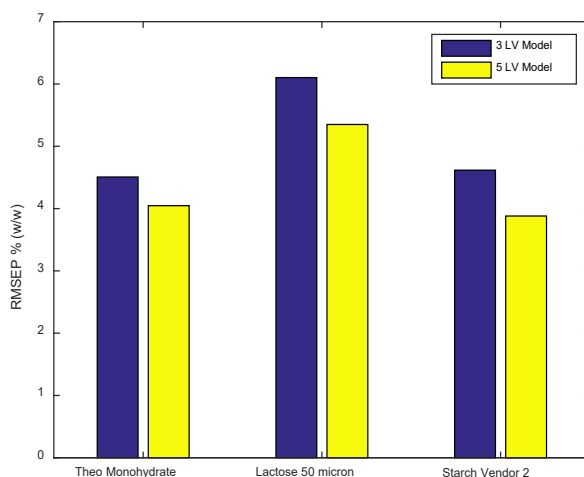


Figure 4-14. RMSEPs of models predicting Theophylline in presence of raw material variations.

During PLS model development, the goal is to maximize co-variance between the spectra and the concentration. Being an iterative technique, the first iteration of PLS seeks to estimate the direction of maximum co-variance between the spectra and the concentration, and the succeeding steps seek to estimate the co-variance between the residual information of the spectra and the concentration left from the preceding steps. Most often, the initial loading vectors explain majority of the variance presented in the spectra and concentration assuming that appropriate pre-processing has already been applied to mitigate other effects. While later loading vectors explain a blend of information that is a small part of the total calibration variance, they still help

to improve the model performance in the calibration space. However, outside the calibration space, it is often possible that the new samples show different projections on the direction of the later loading vectors for being different in terms of any of the blended information explained. This might be detrimental, especially when the particular loading vector has high influence (high weight co-efficient) in the regression vector calculation and the prediction equation. In that case, the regression vector will be sensitive to the information that might not be related with the concentration. The prediction can also be biased by the difference in sample scores on that loading vector. Exclusion of these loading vectors is expected to prevent the model from poor predictability for new samples and improve model robustness as demonstrated in this study. The weight coefficient based method is mostly applicable when certain loading vector has high weight coefficient but explains little spectral and concentration information. This method can identify and eliminate such loading vector from model space. If no such latent variable exists in the model space, the weight coefficient based method and current cross validation technique would provide similar solution to the loading vector selection problem.

4.6 Conclusion

Model robustness is very critical for successful implementation of NIR method. Different sources of variation can be present during product lifecycle and affect the NIR model performance. It is best practice to develop the most robust model at the outset of calibration development. Optimum number of loading vectors selection is necessary for indicating the most robust model. A new method using the weight coefficient was found to be more effective in selecting the optimum number of loading vectors and developing the most robust model compared to the current techniques of cross-validation and variance calculation.

5 Chapter 5: Method development for incorporating physico-chemical variation into the spectral calibration set

5.1 Introduction

The NIR spectral response of a pharmaceutical tablet is affected by its physico-chemical properties such as PSD of the API and excipients, tablet density, tablet hardness, moisture content etc [26, 42, 159]. These physico-chemical properties are critical factors for calibration design due to their effects on NIR spectra [25, 161, 162]. A calibration model typically shows poor accuracy in quantifying tablets with new physico-chemical properties that were not built into the calibration set. A calibration set should contain variation in the physico-chemical properties of the tablets in order to develop a robust calibration model. In the previous chapters, calibration sets were not controlled for certain critical factors such PSD of the API and raw materials. Calibration sets were also used to predict samples with different physico-chemical properties (lab scale vs manufacturing scale). The objective was to test and compare spectral calibration strategy with traditional calibration strategy in presence of new information in the test set. It was found that the presence of new information showed similar impacts on the spectral calibration set and traditional calibration set. Thus, the spectral calibration strategy was considered efficient due to its fewer sample requirements. However, the best practice for calibration set development is to incorporate the critical physico-chemical information into the calibration set. The current approach for incorporating physico-chemical variation into the calibration set is to allow a systematic variation of the respective factors in the calibration design [164]. This approach often leads to a large calibration set with redundant information; a limitation observed during the incorporation of compositional variation as well. Designing calibration set in the spectral space minimizes redundancy and offers an efficient strategy to

incorporate compositional variation into the calibration set. However, it is also critical for the spectral space calibration strategy to efficiently incorporate physico-chemical variation into the calibration set to build a robust calibration model. Incorporation of physico-chemical variation into the spectral calibration set requires the spectral space to contain physico-chemical information. The NIR spectral signature of the physico-chemical properties of the samples can be utilized to incorporate such information into the spectral space. A spectral space containing physico-chemical information provides the basis for selecting a small set of calibration samples with critical physico-chemical information. Such small calibration set would be efficient in sample requirement to develop a robust calibration model for quantitative analysis of pharmaceutical tablets.

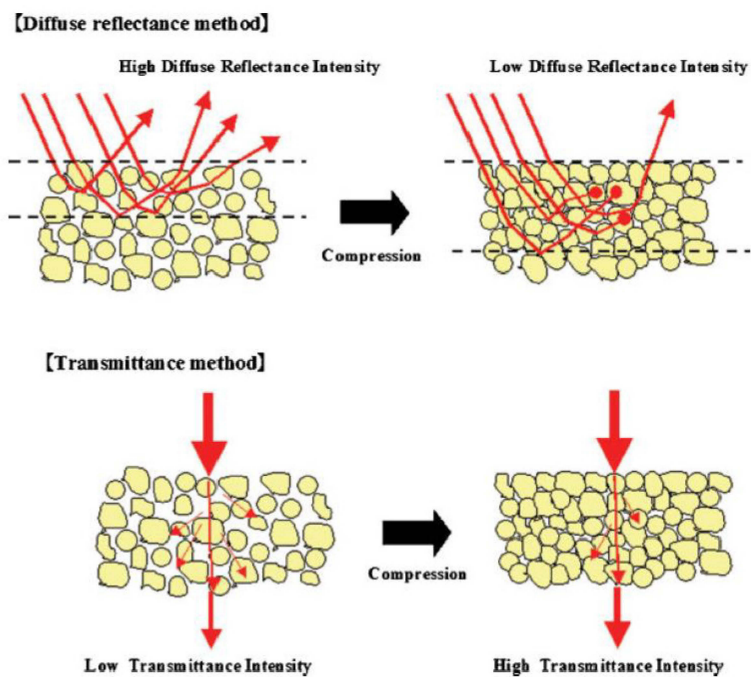


Figure 5-1. Effects of tablet density and hardness on reflectance and transmission NIR [193].

In this study, a method was developed to incorporate physico-chemical variation into the spectral calibration design strategy. Tablet density was selected as an example of critical physico-chemical property of the tablet. Tablet density variation affects the NIR spectral signal due to the changes in air-particle interfaces associated with density variation [11]. As the tablet density increases, the number of air-particle interfaces inside the tablet decreases. Air-particle interface acts as a scattering center for NIR light. As the number of air-particle interface decreases, the extent of scattering also decreases, causing the NIR light to penetrate deeper into the sample and resulting in an increased absorbance. The opposite effect is also observed when the tablet density decreases. Reflectance NIR method was found to be affected more by the tablet density compared to the transmission method [69]. Figure 5-1 shows the effect of tablet density and inter particle interface on the reflectance and transmission NIR spectra.

It is critical for the calibration set to contain information regarding table density variation in order to build a robust calibration model. Tablet density variation is usually incorporated into the calibration set by compressing calibration samples at different compaction forces to prepare tablets of different densities. In the spectral calibration design strategy, tablet density variation was incorporated into the calibration set by selecting calibration samples from a spectral space that contained information regarding tablet density variation. Tablet density variation was incorporated into the spectral space by compressing pure component tablets at different compaction forces. The pure components tablet spectra were used to simulate spectra of compositionally varying tablets at different densities. The simulated tablet spectra of different densities were decomposed to derive the spectral space for selecting calibration samples. The covariance between spectra and API concentration was maximized prior to spectral

decomposition in order to incorporate compositional variation besides the tablet density variation into the spectral space.

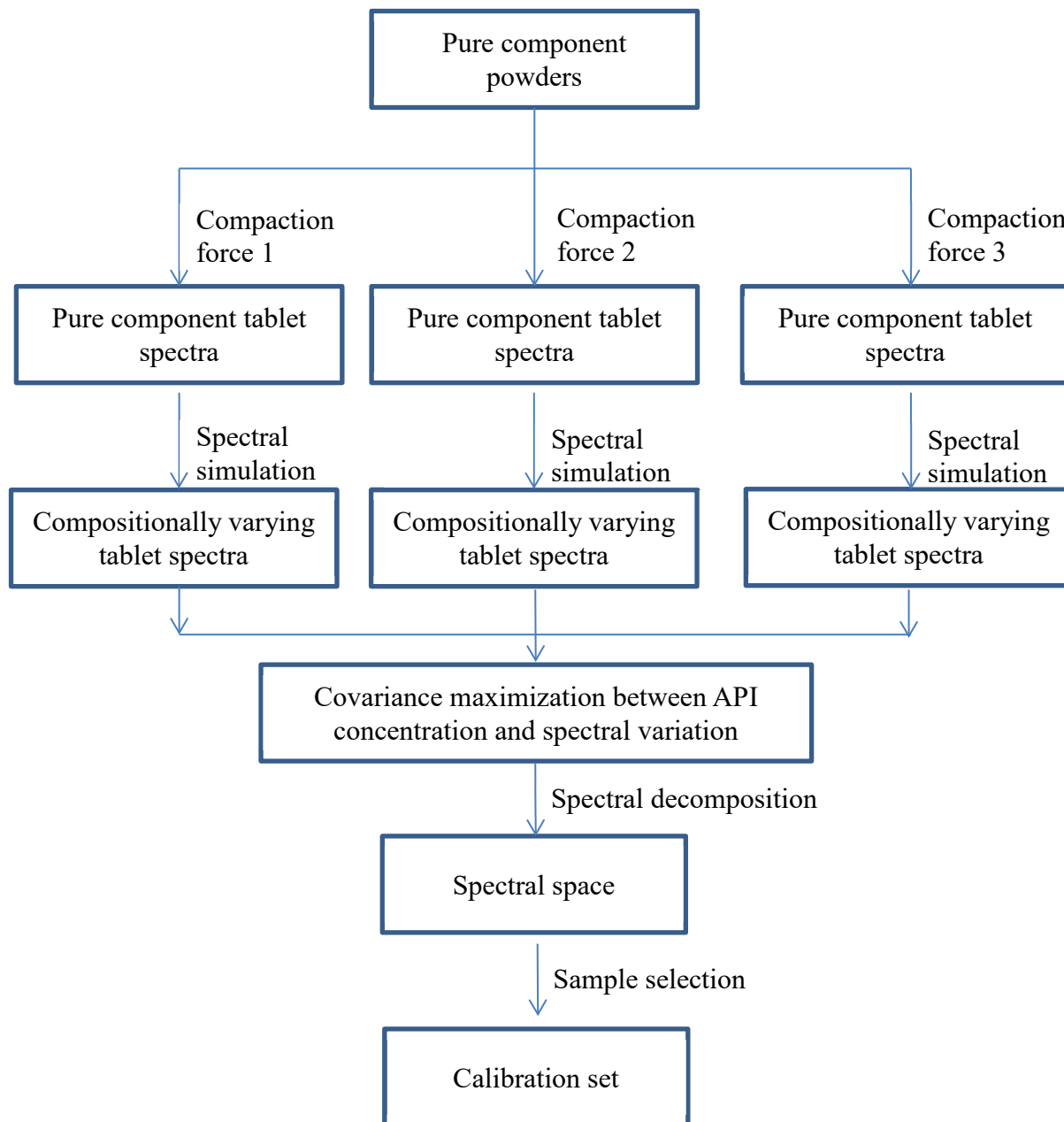


Figure 5-2. Spectral calibration strategy for incorporating tablet density variation into the spectral space

The Kennard Stone algorithm was used to select a representative subset of samples as calibration candidates from the spectral space. This strategy was tested to quantify API in tablets with

different densities. This strategy was also compared with the current strategy for incorporating density variation into the calibration set. The following diagram depicts the spectral calibration strategy for incorporating tablet density variation into the spectral space.

5.2 Material and method

5.2.1 Incorporation of physico-chemical information

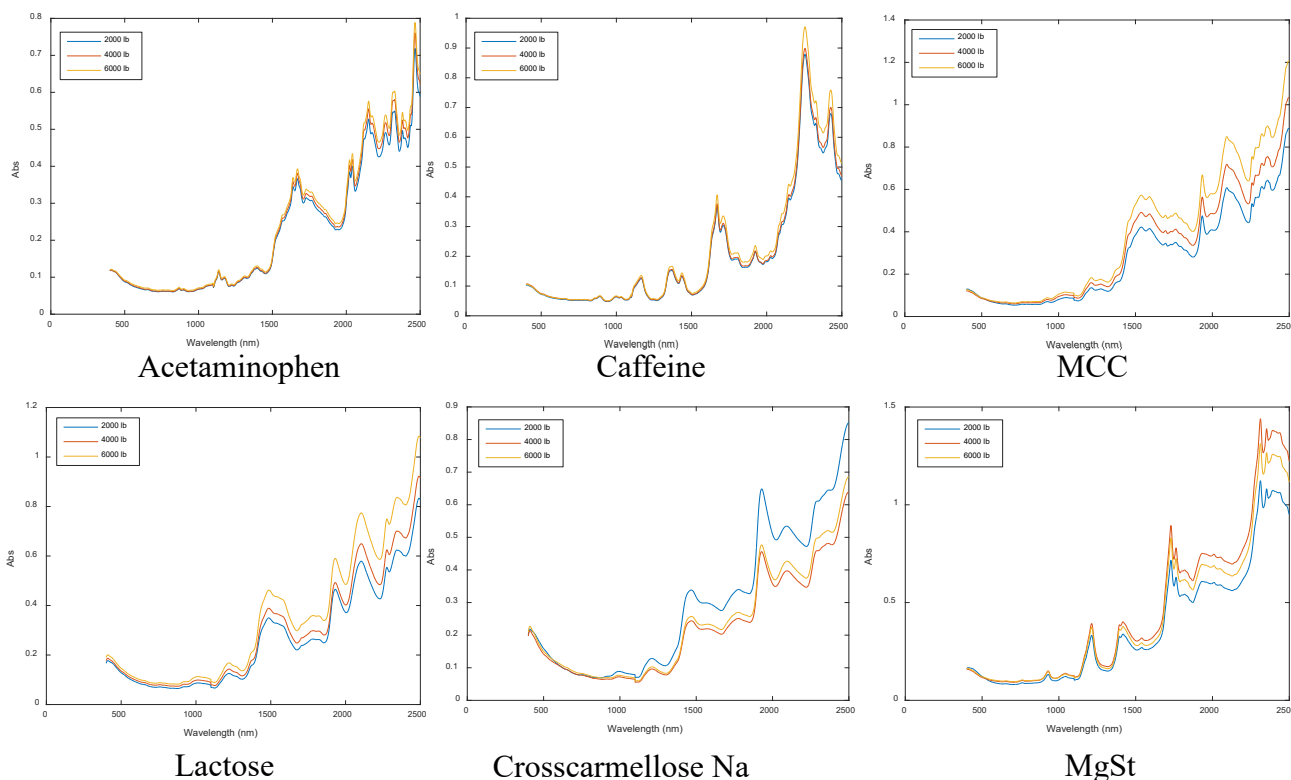


Figure 5-3. Effect of compression force on the pure component tablet spectra

The multiple API formulation was selected for this study. The multiple API tablets contained two APIs as Acetaminophen (Mallinckrodt Inc., Raleigh, NC, USA) and Caffeine anhydrous (Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA). The excipients were Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA),

Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA), Crosscarmellose sodium (Crosscarmellose Na, Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA) and Magnesium stearate (MgSt; Fisher Scientific, Waltham, MA, USA). The target formulation was set as Acetaminophen (31.25% w/w), Caffeine (4.05% w/w), MCC (37.32% w/w), Lactose (24.89% w/w), Crosscarmellose Na (2% w/w) and MgSt (0.5% w/w). The pure component tablets and a tablet at target formulation were compressed at three different compaction forces (2000, 4000 and 6000 lb) on a Carver Automatic Tablet Press using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

The experiments were controlled for tablet homogeneity and viscoelastic relaxation as described in section 3.3.1.4 in chapter 3. All the materials were stored in room temperature and relative humidity. Anhydrous caffeine was reported to be stable at 75% RH for 7 weeks [173]. No hydration and anhydrous caffeine was expected considering lower room RH (~60%) and shorter storage time and analysis. After a viscoelastic relaxation period of two weeks, tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum. Figure 5-3 shows the pure component tablet spectra at different compaction forces. The pure component spectra showed consistent increases in absorbance with increases in compaction pressure except the two cases of Crosscarmellose Na and MgSt. Crosscarmellose Na and MgSt were very difficult to compress, leading to some inconsistent behavior between compaction pressures and spectral responses. However, this spectral effect was expected to be insignificant in the matrix tablet due to very small concentration of Crosscarmellose Na and MgSt in the target formulation tablet.

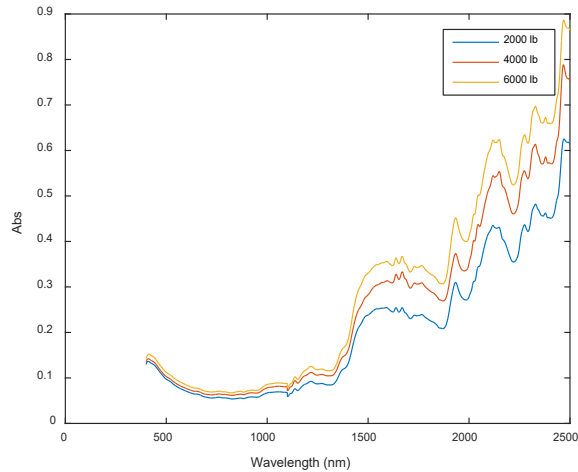


Figure 5-4. Effect of compression pressure on the tablet of target formulation

Figure 5-4 shows the target tablet spectra at three different compaction pressures. A consistent increase in absorbance with the increase in compaction force was observed in the target tablet formulation. A full-factorial experimental design was created to vary the concentration of Acetaminophen, Caffeine and MCC:Lactose ratio. The Acetaminophen and Caffeine concentrations were varied at 5 levels and MCC:Lactose ratio was varied at 3 levels resulting in a total of 75 design points. The details of the design is given in Table 5-1

$$C_{m \times k}^{\text{Composition}} * P_{k \times n}^{\text{Pure component spectra}} = X_{m \times n}^{\text{Sample spectra}} \quad (5.1)$$

	Design factors	Factor levels				
Calibration (5x5x3) 75 samples	Acetaminophen (%)	L.C.-40%	L.C.-20%	L.C.	L.C.+20%	L.C.+40%
	Caffeine (%)	L.C.-80%	L.C.-40%	L.C.	L.C. +40%	L.C.+80%
	MCC/Lac	1		1.5		2

Table 5-1. Full factorial calibration design

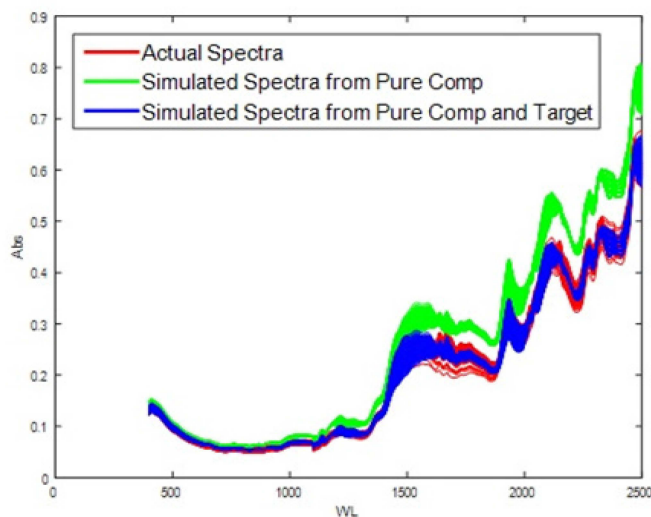


Figure 5-5. Effect of residual spectra on the simulation of spectral response

The spectra of these 75 design points were simulated using equation 5.1. The pure component spectra of each compression force were used during the simulation, resulting in 3 sets of simulated spectra for 3 compression forces (2000, 4000 and 6000 lb).

Significant differences were observed between the simulated spectral response and actual spectral response of the target tablet formulation at all three compaction forces. The spectral difference was calculated between the actual and simulated spectral response of the target formulation for each compaction force. These residual spectra (3 residual spectra from 3 compaction forces) were added to the respective simulated spectral responses. Figure 5-5 shows the actual and simulated spectral responses of the compaction pressure 2000 lb with and without the contribution from the residual spectrum. The actual spectral response were obtained by preparing 75 actual tablets at the compositional points and collecting NIR spectra following the method described in section 2.3 in chapter 2. In total, there were 225 simulated spectra (5 levels

of Acetaminophen, 5 levels of Caffeine, 3 Levels of MCC : Lactose, and 3 levels of compaction forces). Figure 5-6 shows the simulated spectra of 225 design points.

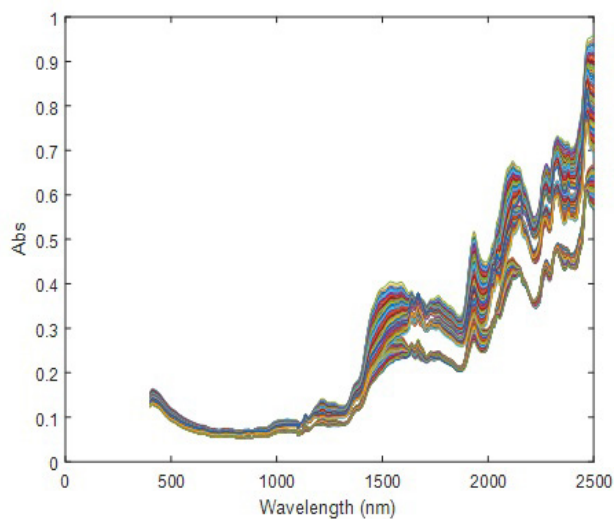


Figure 5-6. Simulated spectral response of 225 design points

The spectral data was truncated from 1100-2499.5 nm. A Principal component analysis was performed on the simulated data using MATLAB 2015a (The Mathworks, Natick, MA, USA) equipped with the PLS_Toolbox v. 7.9.3 (Eigenvector Research Inc., Wenatchee, WA, USA).

Figure 5-7 shows the score plot of first two principal components. In the PCA score space (left), the in-silico samples are separated based on the compaction forces indicating that the majority of the simulated spectral variation was generating from different compaction forces.

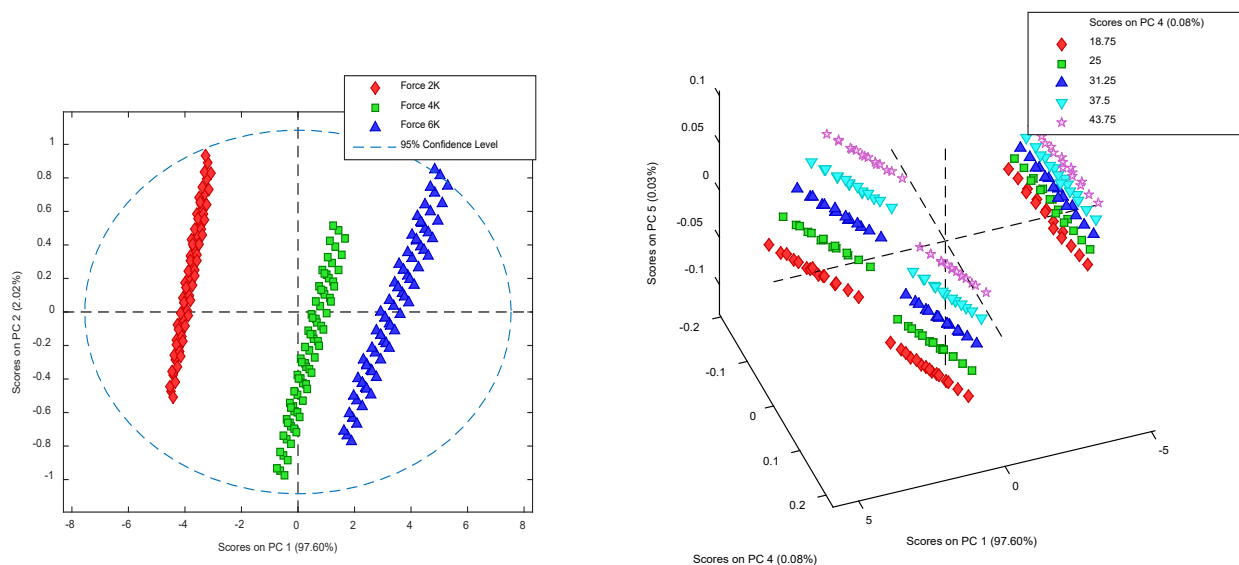


Figure 5-7. Principal component analysis of the simulated spectra of 225 design points. Left: Color coded based on compaction forces (2000, 4000 and 6000 lb), Right: Color coded based on Acetaminophen concentration levels (18.75%, 25%, 31.25%, 37.5% and 43.75% w/w)

The later principal components also explained variation related to composition as shown in

Figure 5-7 (right) (Acetaminophen), Figure 5-8 (left) (Caffeine) and Figure 5-8 (right) (MCC:Lactose ratio). However, the spectral variation was dominated by compaction force variation. The PC1 that explained 97.60% of the total spectral variation, clearly separated the in-silico samples based on the compaction forces. The objective of this study was to incorporate physico-chemical information into the calibration design by selecting the calibration candidates from a score space (aka spectral space) that contained information related to physico-chemical properties. In this score space, such selection would be dominated by compaction force information only. The selected calibration design points would represent insufficient information related to compositional variation. However, the compositional variation is also critical to calibration set, since the primary objective of the calibration set development is to build a

quantitative calibration model for compositional analysis. A balance between compositional and physico-chemical variation is desired to develop a robust calibration model.

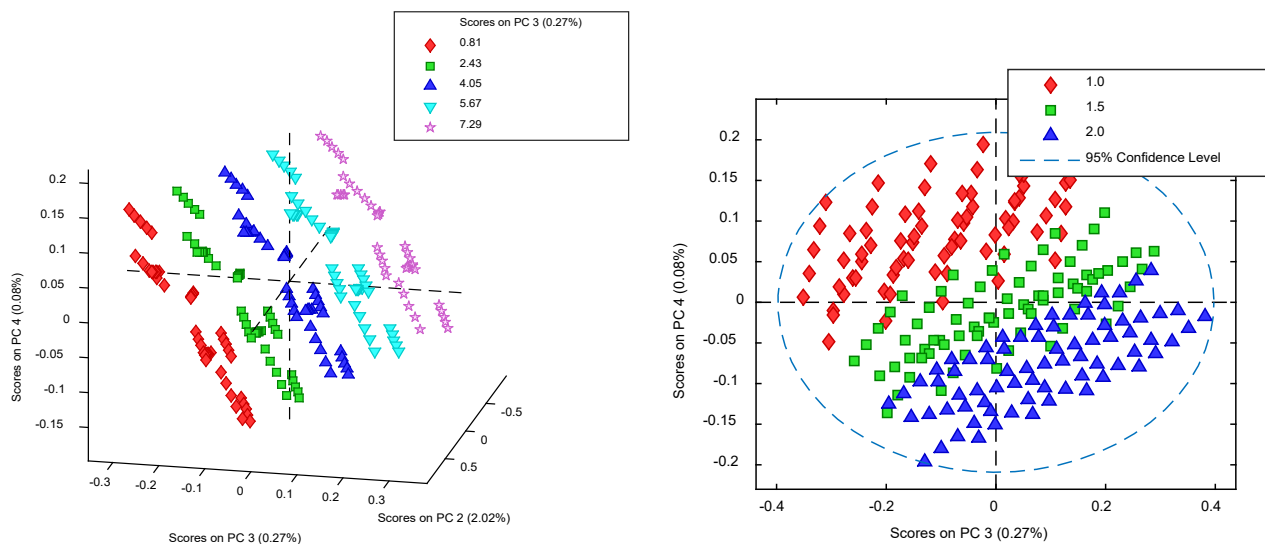


Figure 5-8. Principal component analysis of the simulated spectra of 225 design points. Left: Color coded based on Caffeine concentration, Right: Color coded based on MCC to Lactose ratio

Signal preprocessing technique can be utilized to minimize the spectral effect due to compaction force and offer a balance between physico-chemical and compositional variation. Multiple scattering corrections were applied to minimize the baseline variation caused by the compaction force variation. A PCA analysis was performed on the preprocessed spectra. The PCA analysis showed that the information related to compaction force variation was minimized after preprocessing.

Figure 5-9 shows the PCA score plots after applying MSC correction on the spectral data. No dominance of a single source of spectral variation was observed (no separation in the first two PCs). However, it was also critical for the relevant information to be present in the spectral space

for guiding the calibration candidate selection. A comprehensive PCA analysis was performed.

Figure 5-10 shows the PCA score plots generated from subsequent PCs.

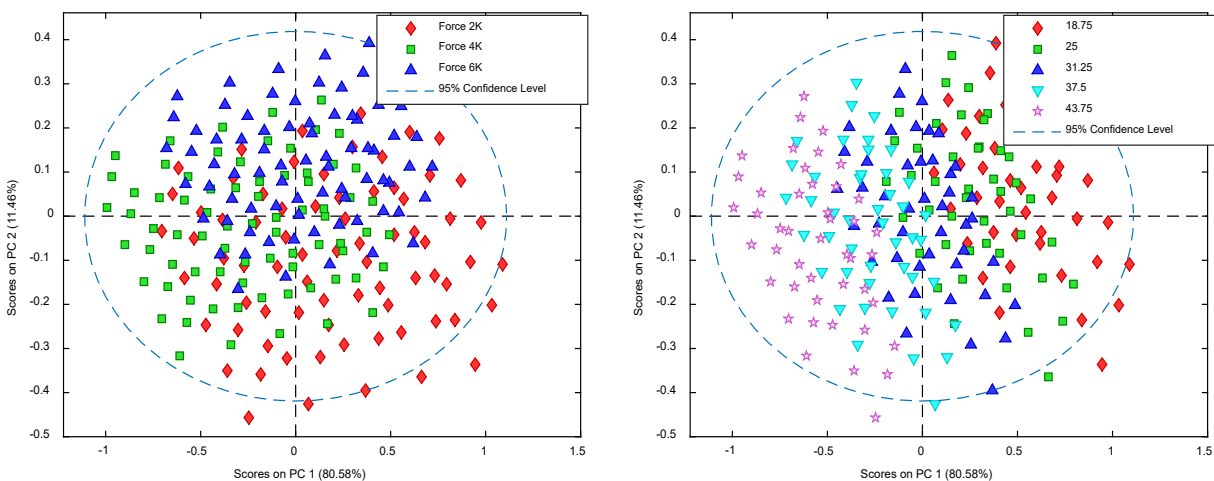


Figure 5-9. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Left: Color coded based on compaction forces, Right: Color coded based on Acetaminophen concentration

Figure 5-10 (left) and Figure 5-10 (right) show that the spectral data set contained information related to compaction force and Acetaminophen concentration variation. Figure 5-11 shows that the spectral response also contained information related to the MCC to Lactose ratio of the in-silico samples. The simulated spectra of in-silico samples were found to contain both physical and chemical variation and it was possible to incorporate a balance between these two sources of spectral variation by implementing an appropriate preprocessing technique.

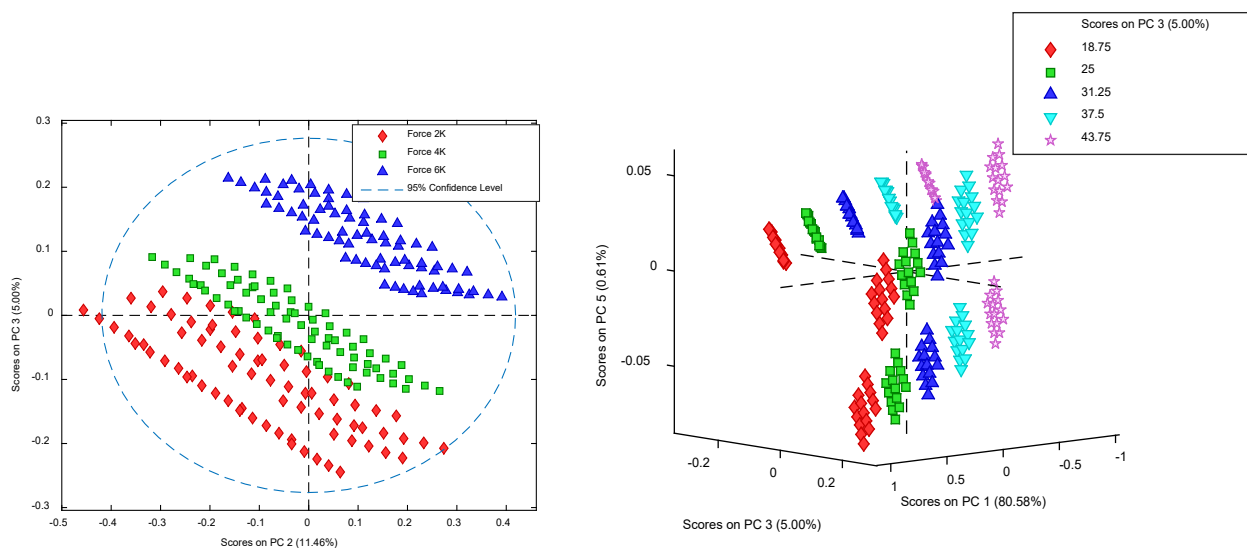


Figure 5-10. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Left: Color coded based on compaction forces, Right: Color coded based on Acetaminophen concentration

5.2.2 Selection of calibration candidates

The next step was to find the calibration candidates for developing quantitative methods for Acetaminophen and Caffeine. It was assumed that the optimum calibration sets are different for different APIs. In other words, the optimum calibration set for Acetaminophen is different from the optimum calibration set for Caffeine (this assumption was tested and the result is shown in the next chapter). However, selecting the calibration candidates in the PCA score space would result the same calibration design points for Acetaminophen and Caffeine. It was also observed that the PCA score space contained little/no information regarding the variation of Caffeine concentration. Selecting the calibration candidates in the PCA score space would result a calibration set without appropriate variation in Caffeine concentration. It is critical to select the calibration candidates from a score space that contains information regarding the variation in the respective API concentration. In PLS, the covariance between respective API (Acetaminophen /

Caffeine) and spectral response is maximized to generate the score space. Therefore, PLS technique was used to generate separate score spaces that are specific to Acetaminophen and Caffeine.

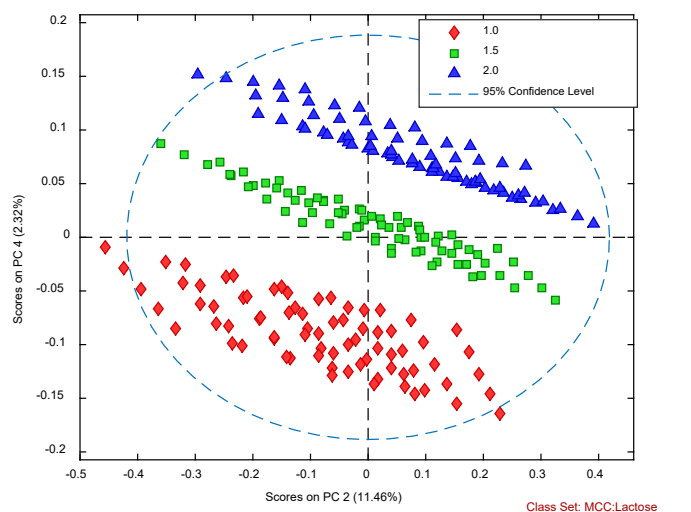


Figure 5-11. Principal component analysis of the MSC corrected simulated spectra of 225 design points. Figure is color coded based on MCC to Lactose ratio

5.2.2.1 Calibration candidates for Acetaminophen

PLS model was developed from the in-silico samples to maximize the covariance between Acetaminophen concentration and simulated spectral variation. MSC and mean centering were used as the preprocessing techniques for spectral data as indicated by the PCA analysis. Auto scaling was used to preprocess the concentration data. The resultant score space from PLS model is shown in Figure 5-12. It was observed that, the first set of LVs explained spectral variation related to Acetaminophen concentration due to the maximization of covariance between

Acetaminophen concentration and spectral response. The spectral variation related to compaction force and excipient variation was also explained by other LVs (LV 3 and LV5) as shown in

Figure 5-13.

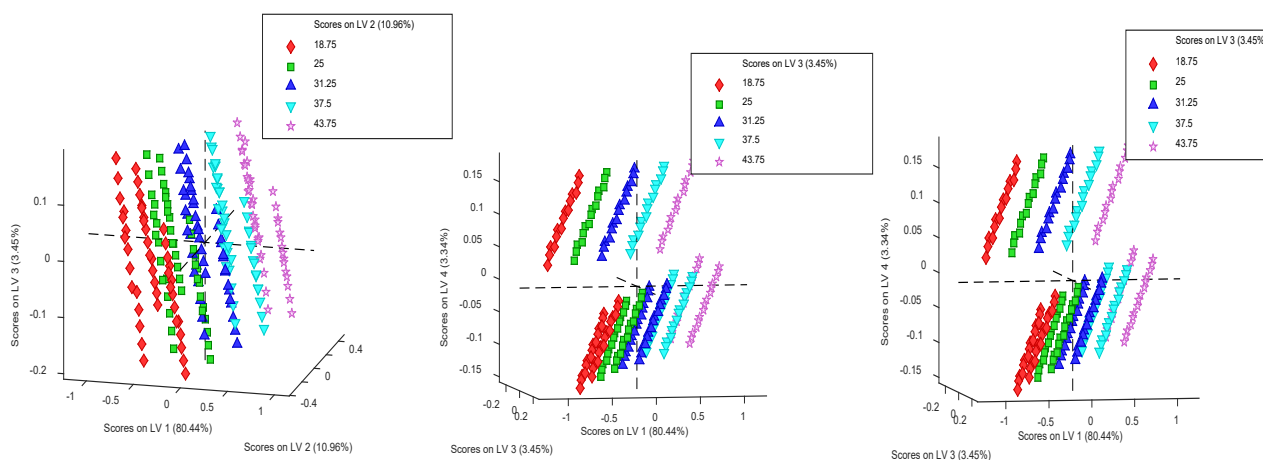


Figure 5-12. Score plot from the PLS model. The figure is color coded based on Acetaminophen concentration level

From this score space, a set of 60 design points was selected as the calibration candidates for Acetaminophen calibration model using the Kennard Stone algorithm. Kennard Stone algorithm selected a representative sub set of in-silico samples from the large dataset based on Euclidian distances. The details of the Kennard stone algorithm can be found elsewhere [132]. Figure 5-14 shows the selected points using Kennard stone algorithm in the score space. It was shown in the score plot that, calibration candidates were selected from all potential sources of variation during the selection process. All levels of Acetaminophen concentrations, MCC:Lactose ratios and compaction forces were included into the selected calibration candidates. It is critical to include all sources of variation into the calibration set to make a robust calibration model against all potential sources of variation. The current strategy to incorporate such variation is to create a wide range of systematic variation in all sources of variation during calibration design.

Following such strategy, the traditional method of incorporating compaction force variation would be to prepare tablets at 3 compaction forces for each of the 75 compositional design points, resulting in a total of 225 calibration points in the calibration set. This traditional calibration set will be referred as ‘global calibration set’ throughout the rest of the chapter. The global calibration set may contain redundant information since tablets of all three compaction forces at each compositional design point might not be necessary to incorporate sufficient amount of spectral variation into the calibration set. The selection of calibration candidates in the simulated score space would allow the incorporation of compaction force variation into the calibration set using limited number of samples.

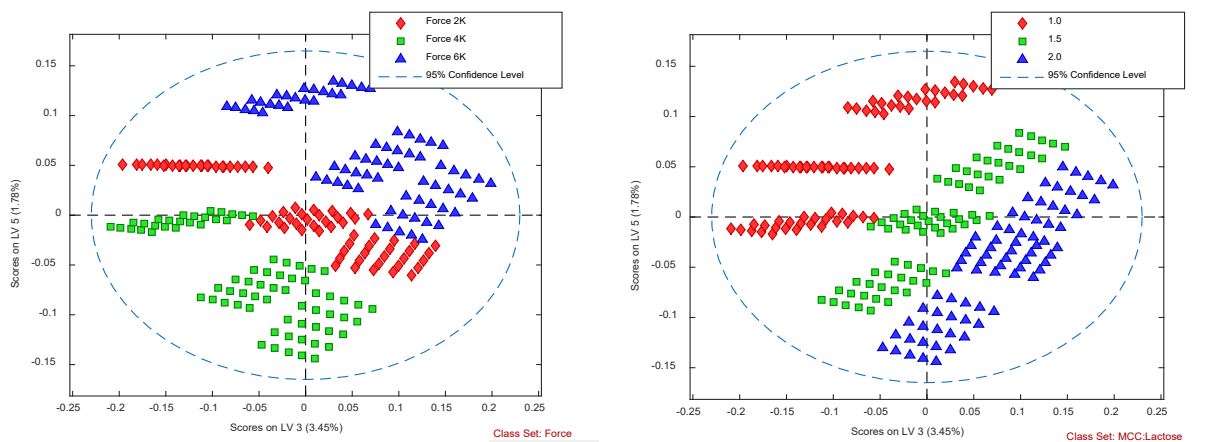


Figure 5-13. Score plot from the PLS model. Left: Color coded based on compaction force. Right: Color coded based on MCC:Lactose ratio

The set of 60 calibration candidates had critical sources of spectral variation including Acetaminophen concentration variation, Caffeine concentration variation, MCC:Lactose ratio and compaction force. In the traditional calibration design approach, the variation of such critical factors is not guided by a priori information and largely depends on user choice (e.g. no of levels

for each factor). However, in the spectral design strategy, such variations in the calibration set were guided by the spectral information from each individual factor and their interaction.

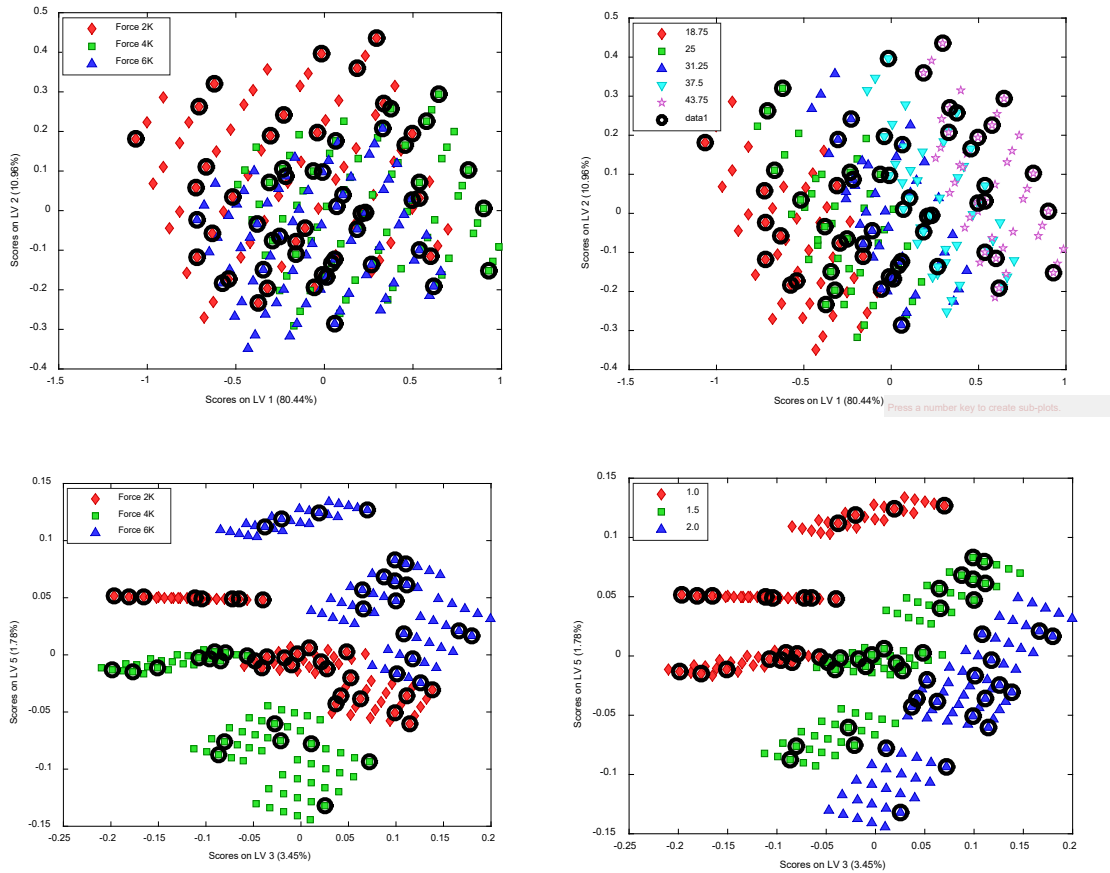


Figure 5-14. Selection of calibration candidates (indicated by black circle) in the PLS score space. The figures are color coded based on compaction force (Left), Acetaminophen concentration level (upper right) and MCC:Lactose ratio (bottom right)

The selection of the calibration candidates was also guided by maximizing the covariance between Acetaminophen concentration and spectral variation. The effect of this guided selection strategy was investigated by analyzing the distribution of different critical factors such as Acetaminophen concentration, Caffeine concentration, MCC:Lactose ratio and compaction force variation in the selected calibration candidates.

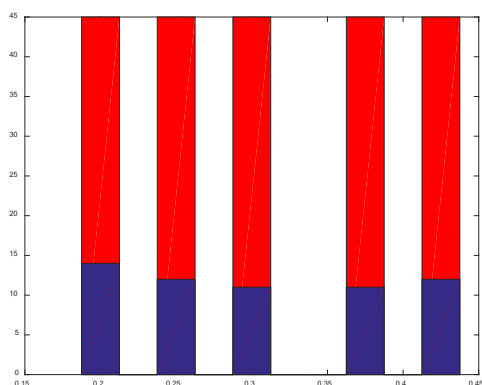


Figure 5-15. Distribution of Acetaminophen in global calibration set and spectrally selected calibration set

Figure 5-15-17 shows the distribution of the critical factors in the global calibration set (225 samples) and in the selected calibration set (60 samples). In the global calibration set (225 design points), all five concentration levels were presented at 45 occurrences ($5 \times 45 = 225$) as shown in Figure 5-15 in red bars. In the spectral calibration set, a set of 60 design points was selected in which all the concentration levels were homogeneously distributed and occurred around 12 times for each concentration level as shown in Figure 5-15 in blue bars.

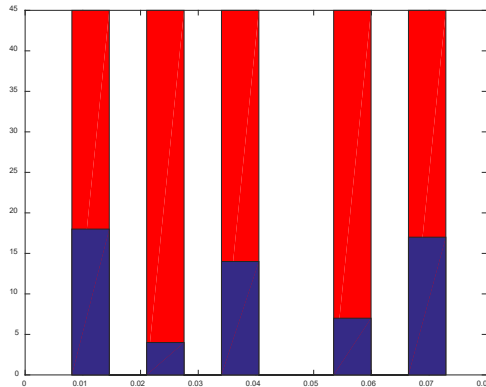


Figure 5-16. Distribution of Caffeine in global calibration set and spectrally selected calibration set

The Caffeine concentrations in the selected design points were mostly distributed at the extreme levels to make the calibration model robust against Caffeine concentration variation as shown in

Figure 5-16. Mid-level concentration points were mostly excluded as the sensitivity to Caffeine concentration was not critical for the calibration performance during Acetaminophen prediction.

Also, the low dose of Caffeine was not able to contribute a significant portion of the spectral variation after the preprocessing and maximization of covariance between spectra and Acetaminophen concentration. The extreme design points were sufficient to cover such small spectral contribution from the Caffeine concentration.

The excipient concentrations (MCC and Lactose) were widely distributed in the selected design points as shown in Figure 5-17. Excipient variation contributed to a significant portion of spectral variation as shown in the score plots.

Excipient variation also causes variation in the physical properties such as density, porosity, hardness, moisture content of the tablets. An even distribution of excipient variation in the design points is critical to cover such variation in the calibration set.

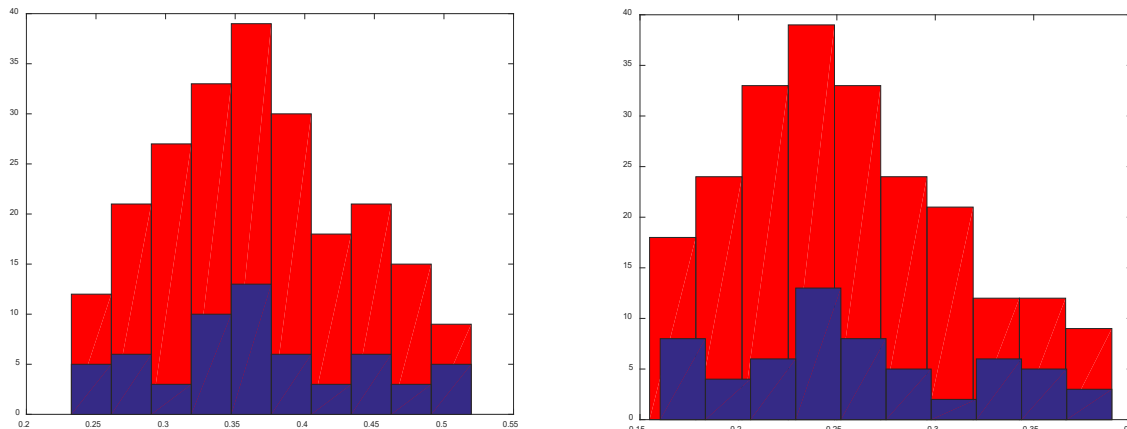


Figure 5-17. Distribution of MCC (left) and Lactose (right) in global calibration set and spectrally selected calibration set

The objective of this study was to incorporate physico-chemical information such as variation in tablet density into the calibration set. The compaction force is usually varied to introduce variation in the tablet density into the calibration set. Traditionally, compaction force is varied comprehensively without any prior direction or optimization in the calibration set. In this study, the spectral effect of the compression force was analyzed and a small representative subset was selected based on the spectral effect of the compression force. The distribution of the compaction forces in the traditional global calibration set and in the spectrally selected calibration set is shown in Figure 5-18. A large number of design points were selected in the low compaction force region compared to the high compaction force region. A lower compaction force causes wider distribution of solid fractions in the tablet set compared to the high compaction force. As the solid fraction affects the NIR spectral response, it was assumed that, more design points were required / selected at the low compaction force region to cover wider distribution of solid fractions and associated spectral responses.

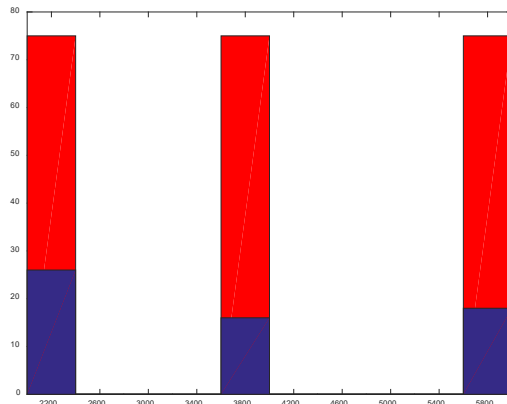


Figure 5-18. Distribution of compaction forces in global calibration set and spectrally selected calibration set

5.2.2.2 Calibration candidates for Caffeine

A similar strategy was used to select calibration candidates for developing a quantitative method for Caffeine. A PLS model was developed from the simulated spectral response of the in-silico tablets and respective Caffeine concentration. A set of 60 design points was selected from the PLS score space using the Kennard stone algorithm. The distribution of critical factors such as Caffeine concentration and Acetaminophen concentration in the global calibration set (225 samples) and selected calibration set (60 samples) are shown in

Figure 5-19. In this calibration set, a homogeneous distribution of Caffeine concentration was observed; contrary to the calibration set of Acetaminophen. A homogeneous distribution of Caffeine concentration was selected from the latent variable score space due to maximization of covariance between Caffeine concentration and spectral variance. Such distribution of Caffeine concentration was critical to make the model sensitive to Caffeine concentration.

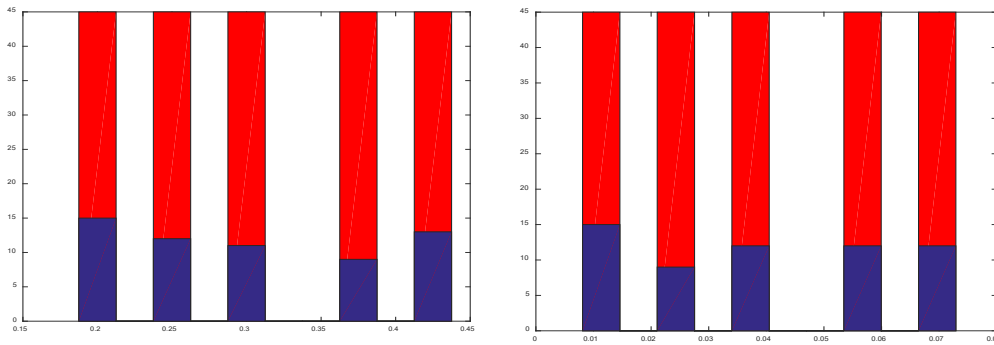


Figure 5-19. Distribution of Acetaminophen (left) and Caffeine (right) concentration in global calibration set and spectrally selected calibration set

Similar distributions of MCC and Lactose concentration were observed in the calibration set of Caffeine compared to the calibration set of Acetaminophen. The distribution of MCC and Lactose concentrations in the global (225 design points) and selected calibration set (60 design points) are shown in Figure 5-20. A wide distribution of excipients was critical to ensure calibration model robustness against excipient variation. Figure 5-21 shows the distribution of compaction forces in the selected calibration set for Caffeine. This calibration set also contained more design points at low compaction force compared to the high compaction force.

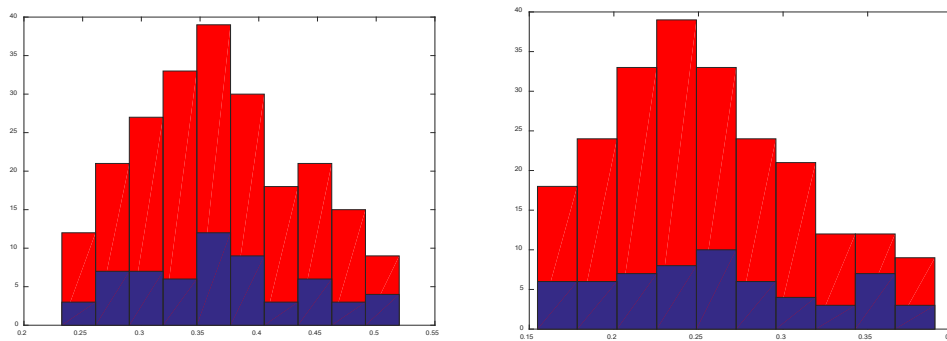


Figure 5-20. Distribution of MCC (left) and Lactose (right) concentration in global calibration set and spectrally selected calibration set

Justification of low compaction force

It was assumed that more design points were selected at the low compaction force due to wide range of spectral variation and solid fraction. These tablet design points were in-silico tablets and its spectra were simulated from actual pure component and target tablets at different compression forces. It was assumed that, tablets prepared at low compaction forces had wider spectral variation compared to the tablets prepared at high compaction forces. And the wide spectral variation was caused by wide range of solid fraction usually seen at low compaction forces.

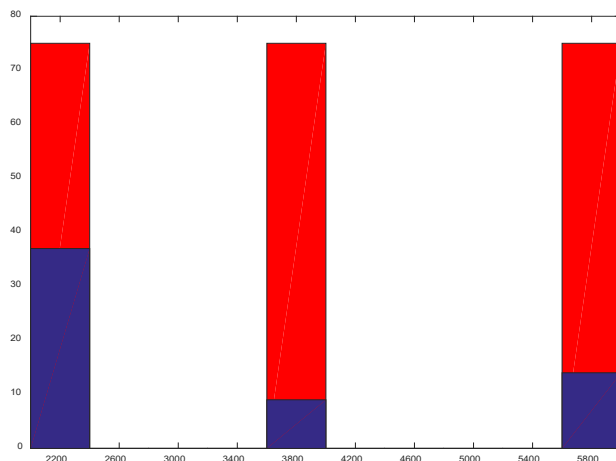


Figure 5-21. Distribution of compaction forces in global calibration set and spectrally selected calibration set

Actual tablets were prepared at different compression forces to test this assumption. The solid fraction of each tablet was measured by the following equation 5.2. Figure 5 22 (left) shows the tablet solid fractions at each compaction force. At low compaction force, a wider range of solid fraction was present as expected. The spectral variances were calculated at each wavelength in the simulated spectral set of each compaction force. Figure 5 22 (right) shows the spectral

variance at each wavelength for all three compaction forces of in-silico tablets. In contrast to the initial assumption, it was found that the wide distribution of solid fractions did not cause wide range of spectral variation.

$$\text{Solid fraction of tablet} = \frac{\text{Tablet density}}{\text{True density}} \quad (5.2)$$

The spectral variation was minimum at the low compaction force in-silico tablets. The spectral variation was maximum at the high compaction force in-silico tablets. It was further investigated to understand the cause behind the selection of more design points from low compaction force in-silico tablets associated with smaller amount of spectral variation.

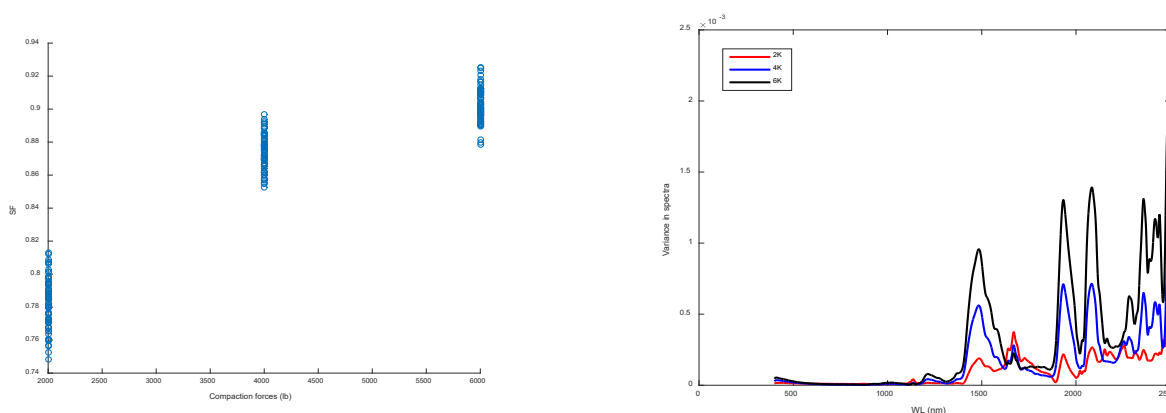


Figure 5-22. Solid fraction of tablets at different compaction forces (A). Spectral variation of the tablet sets prepared at different compaction forces (B).

It was expected to have more design points at high compaction force in the selected calibration set due to the large spectral variance observed at high compaction force. However, after the MSC correction of the spectral data during PLS model development, more spectral variance was observed at low compaction force as shown in Figure 5-23. It might be due to the greater extent of non-linearity caused by the wider distribution of solid fractions at the low compaction force.

MSC preprocessing technique corrects for linear scattering effect. The non-linear scattering effect could not be corrected by MSC preprocessing, resulting in wide range of spectral variation in the preprocess data at low compaction force. Such wide range of spectral variation resulted in wide distribution of scores in the PLS score plots as shown in Figure 5-24.

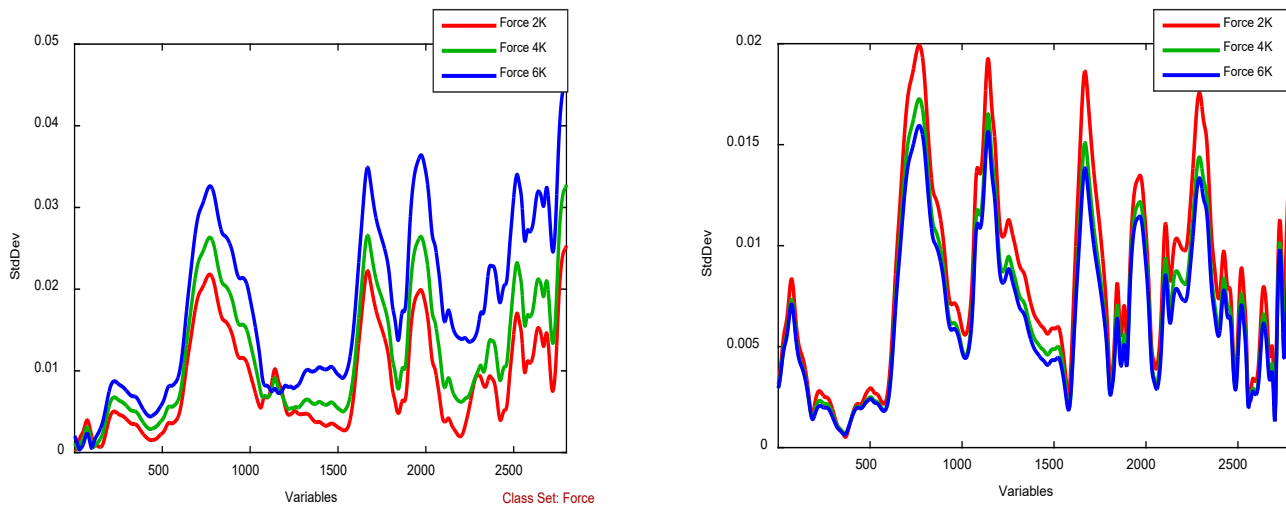


Figure 5-23. Spectral variation before (A) and after MSC correction of the spectral data.

The area inside the box represents the range of score distribution for each respective compression force. The distribution of scores was wider for low compaction force tablets compared to the high compaction force tablets. This wider distribution of scores of low compaction force tablets resulted in higher k-nearest neighbor (KNN) distance scores. Since, the Kennard stone algorithm selects samples based on inter sample distances and the highly distant samples are included in the selected subset, more samples were selected from the low compaction force during the selection of calibration candidates.

Figure 5-25 shows the KNN distance scores of tablets at three different compaction forces for both Acetaminophen and Caffeine. For both APIs, the higher KNN distances of low compaction

force tablets resulted in large number of low compaction design points in the selected calibration candidates.

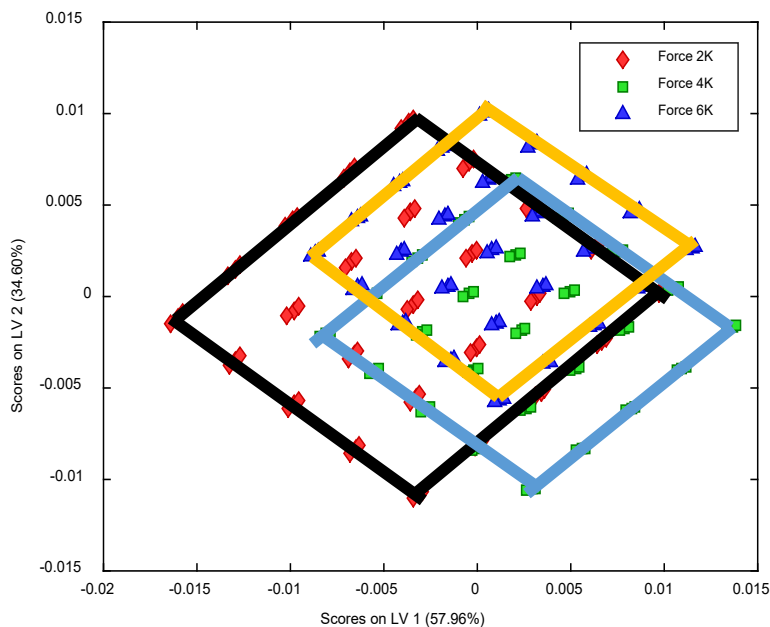


Figure 5-24. PLS scores of first two loading vectors at three different compaction forces

5.2.3 Spectral calibration set for Acetaminophen

A set of 60 design points was selected to develop the calibration set for Acetaminophen prediction. Actual tablets at the selected design points were prepared by direct compression. The required amounts of Acetaminophen, Caffeine, MCC, Lactose, Crosscarmellose Na and MgSt at each design point were weighed and placed in a 10 ml scintillation vial. The ingredients were mixed by rotating the scintillation vial. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at respective compaction force using a 13 mm

die and flat-faced punches. The target tablet weight was 700 mg. This tablet set was defined as the ‘spectral calibration set’ for Acetaminophen.

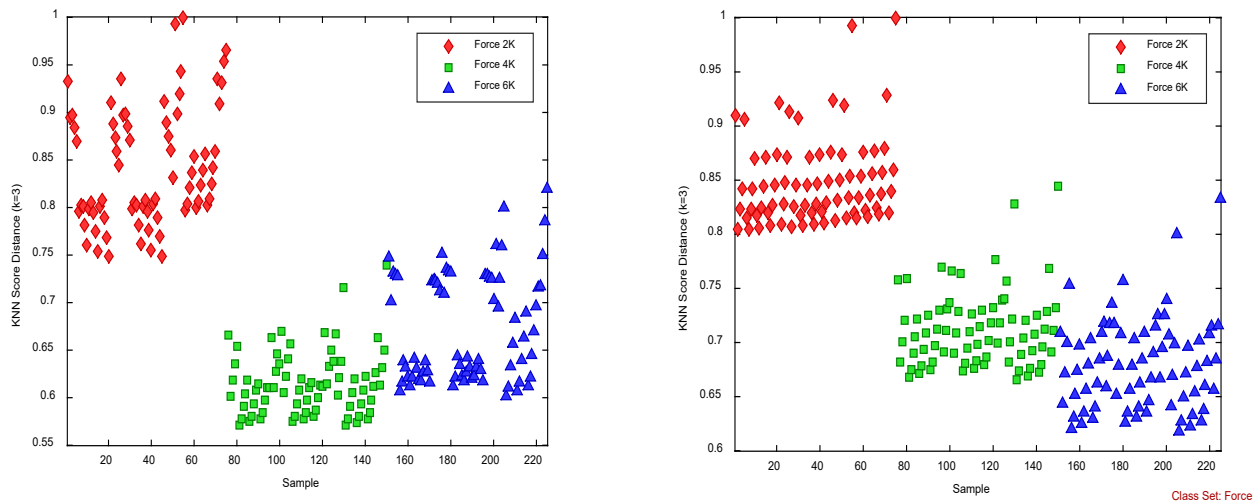


Figure 5-25. KNN distance scores of samples at three different compaction forces for PLS models of Acetaminophen (A) and Caffeine (B)

All the tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum. Spectral calibration set and their respective spectra were used to develop a quantitative NIR calibration model to quantify the amount of Acetaminophen in tablets of interest.

5.2.4 Traditional calibration set for Acetaminophen

The traditional calibration set was developed using the 4-factor full factorial design previously used to simulate the spectral response. The factors were Acetaminophen concentration, Caffeine concentration, MCC: Lactose ratio and compaction force. Each of the APIs (Acetaminophen and Caffeine) concentrations was varied at five levels and other two factors (MCC: Lactose ratio and

compaction force) were varied at three levels resulting in 225 (5x5x3x3) design points in the traditional full factorial calibration set. The calibration design is different from the design in chapter 3 in that, the levels of compaction forces are 2000, 4000 and 6000 lb in the current design contrary to 4000, 5000 and 6000 lb in the earlier design.

Each of the 225 tablets was individually prepared by direct compression. All the components for a single design point were weighed using a digital weighing machine (Data Range, Model No. AX504DR, Mettler Toledo) and placed in a 10 ml scintillation vial. The ingredients were mixed in a bin blender (L.B. Bohle LLC, Warminster, PA, USA) for 10 mins followed by a high shear mixing using a vortex machine (Vortex-2 Genie, Model G-560, Scientific Industries, IN, USA). The final mixing was performed in the bin blender before compression. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at respective compression forces using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

All the tablets were scanned using a bench top NIR instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc) in reflectance mode. Spectral data were collected at 0.5 nm increment over a range of 400 nm – 2499.5 nm with 32 co-adds per spectrum. Spectra from both faces of each compact were averaged to produce a single representative spectrum. Traditional calibration set and their respective spectra were used to develop a quantitative NIR calibration model to quantify the amount of APPA in tablets of interest.

5.2.5 Spectral calibration set for Caffeine

The spectral calibration set for Caffeine was developed using the similar strategy as used for spectral calibration set for Acetaminophen. A set of 60 design points was selected as the

calibration candidates from the PLS scores of Caffeine calibration model of in-silico samples. Only 29 design points were found to be common between the selected design points of Acetaminophen and Caffeine. Actual tablets at these design points were prepared by the direct compression technique described earlier. The NIR spectra of these tablets were collected using the same instrument and method described earlier. This calibration set was described as ‘spectral calibration set’ for Caffeine. Spectral calibration set and their respective spectra were used to develop a quantitative NIR calibration model to quantify the amount of Caffeine in tablets of interest.

	Design Factors	Design levels				
Test 1 (5x5x3x1) 75 samples	Acetaminophen (%)	L.C.-40%	L.C.-20%	L.C.	L.C. +20%	L.C. +40%
	Caffeine (%)	L.C.-80%	L.C.-40%	L.C.	L.C. +40%	L.C. +80%
	MCC/Lac	1		1.5		2
	Force (lb)	5000				
Test 2 (3x3x1x4) 36 samples	Acetaminophen (%)	L.C.-35%		L.C.	L.C.+35%	
	Caffeine (%)	L.C.-70%		L.C.	L.C.+70%	
	MCC/Lac			1.5		
	Force (lb)	2000	4000		5000	6000

Table 5-2. Design of the two test sets

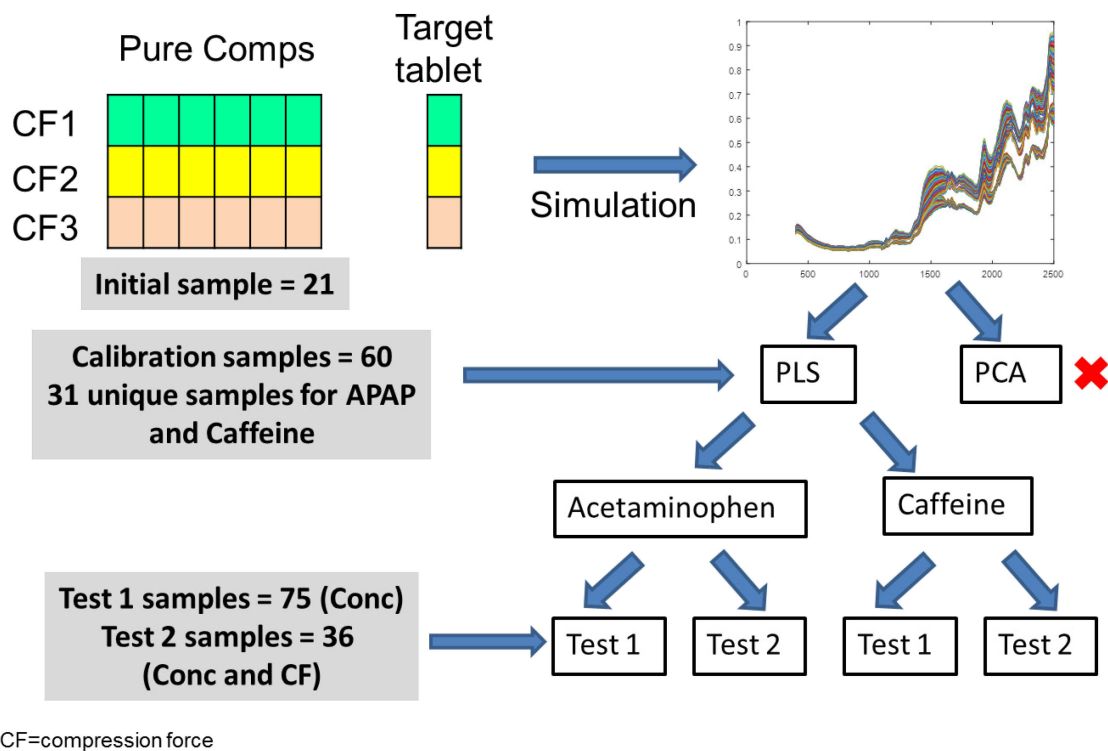


Figure 5-26. Selection and preparation method of spectral calibration set and test set for Acetaminophen and Caffeine

5.2.6 Traditional calibration set for Caffeine

The same full factorial calibration set for Acetaminophen (225 samples) was used to develop a traditional calibration method for Caffeine prediction.

5.2.7 Test sets to evaluate model performance

Two test sets were used to evaluate the performance of the calibration designs. The first test was developed by varying Acetaminophen and Caffeine concentration at five levels and MCC to Lactose ratio at 3 levels resulting in 75 design points. In the first test set, the compaction force was kept constant. The second test set was developed by varying Acetaminophen and Caffeine concentration at three levels and compaction force at four levels resulting in total 36 design

points. In the second test set, the excipient ratio was kept constant. Table 5-2 describes the structure of the two test sets. The method for preparing calibration and test set is depicted in Figure 5-26.

5.2.8 Quantitative model development

Quantitative models were developed using Partial Least Squares (PLS) modeling technique in MATLAB 2015a environment (The Mathworks, Natick, MA, USA) using PLS_Toolbox v. 7.9.3 (Eigenvector Research Inc., Wenatchee, WA, USA). Data independent spectral preprocessing techniques were used to optimize model performance. Two calibration models were developed for each API (Acetaminophen and Caffeine), one calibration model from the 225 actual tablets of the full factorial design and another calibration model from the 60 actual tablets of the spectral design. Models developed from each calibration design were independently optimized for each API. The preprocessing techniques and loading vectors were selected independently. Selection of the loading vector is critical for PLS model performance. Latent variables were chosen based on a parsimonious approach. A minimum number of latent variables with acceptable performance were selected. Model performance was assessed based on the prediction of Acetaminophen and Caffeine concentration in calibration and test set tablets.

The effect of the incorporation of physico-chemical information into the spectral calibration sets was analyzed by evaluating calibration model performance and comparing it with the global calibration model performance. Root mean squared error (RMSE) was used to evaluate the model predictive performance in the two independent test sets (RMSEP). A two-way analysis of variance (ANOVA) test was performed to compare the prediction errors of the full factorial and spectral calibration sets for each API [22]. The underlying model for the ANOVA analysis is

described in equation 5.3, where index ‘i’ refers to the calibration model and index ‘j’ refers to the sample number. The symbol α_i and β_j refer to the effect of calibration model ‘i’ and sample number ‘j’ on prediction error, respectively.

$$(\widehat{y}_{ij} - y_{ij})^2 = \mu + \alpha_i + \beta_j + e_{ij} \quad (5.3)$$

The calibration models are considered significantly different in terms of prediction performance, if the calibration model parameter α is found to be significant in the ANOVA analysis.

5.3 Results and discussion

5.3.1 Prediction of Acetaminophen

A PLS model was developed from the traditional calibration set (225 samples). The preprocessing techniques were selected as MSC and mean centering for spectral data and auto scaling for concentration data. A set of 6 loading vectors was selected for model development. Figure 5-27 shows the performance of the traditional calibration set for predicting Acetaminophen in the first test set (75 samples). The score plot shows that the spectral variation in the test set was covered by the calibration set. The calibration, cross validation and prediction errors were found to be similar (~1% w/w). A reasonable prediction performance was achieved due to similar spectral variation between calibration and test set samples.

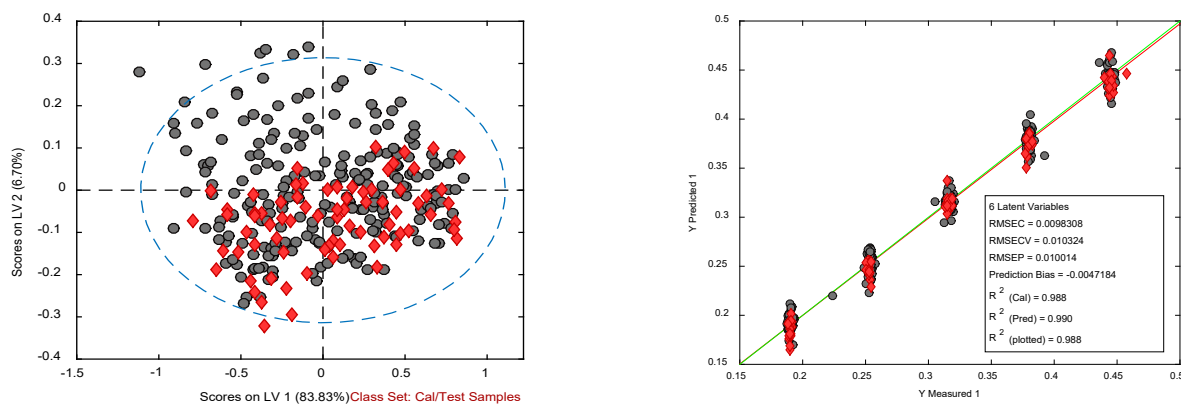


Figure 5-27. PLS model performance of the global calibration set predicting Acetaminophen in test set 1

Another PLS model was developed from the spectral calibration set for Acetaminophen prediction. Same preprocessing techniques (MSC and mean centering for spectral data and auto scaling for concentration data) were found to be optimum for this model. A set of 6 loading vector was selected for model development. Figure 5-28 shows the performance of the spectral calibration set for predicting Acetaminophen in the first test set. The calibration samples had similar and wider spectral variation compared to the test set samples. Similar prediction performance (RMSEP ~ 1% w/w) was achieved by the spectral calibration set compared to the traditional calibration set in spite of having fewer samples (60 vs 225).

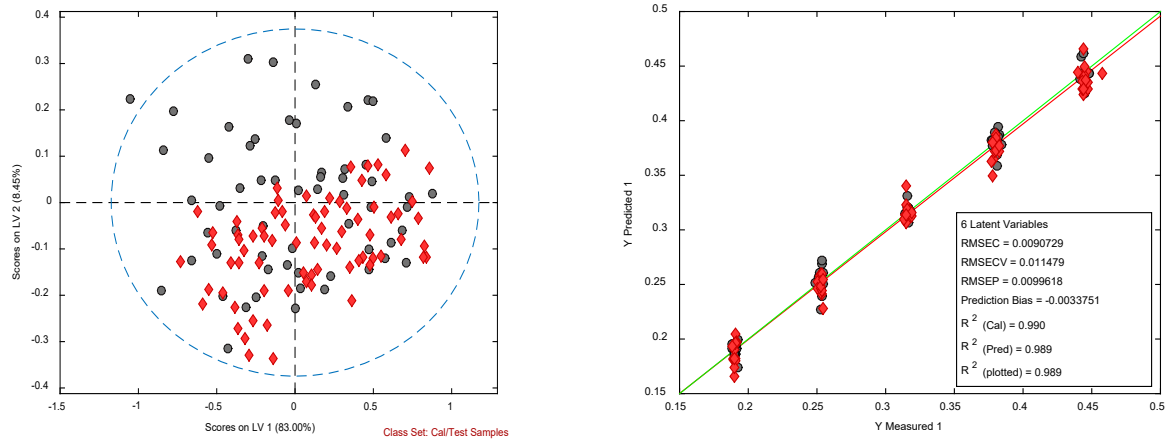


Figure 5-28. PLS model performance of the spectral calibration set predicting Acetaminophen in test set 1

The same calibration models of traditional calibration set and spectral calibration set were used to predict the concentration of Acetaminophen in the second test set. In the second test set (36 samples), the prediction errors of the two models were slightly higher compared to the first test set (RMSEPs ~ 1.2 % w/w). The higher error was caused by the compaction force variation in the test set samples. The prediction performances between the two calibration models were found to be equivalent. Figure 5-29 shows the score plot and reference vs prediction plot for two calibration models predicting Acetaminophen in the second test set.

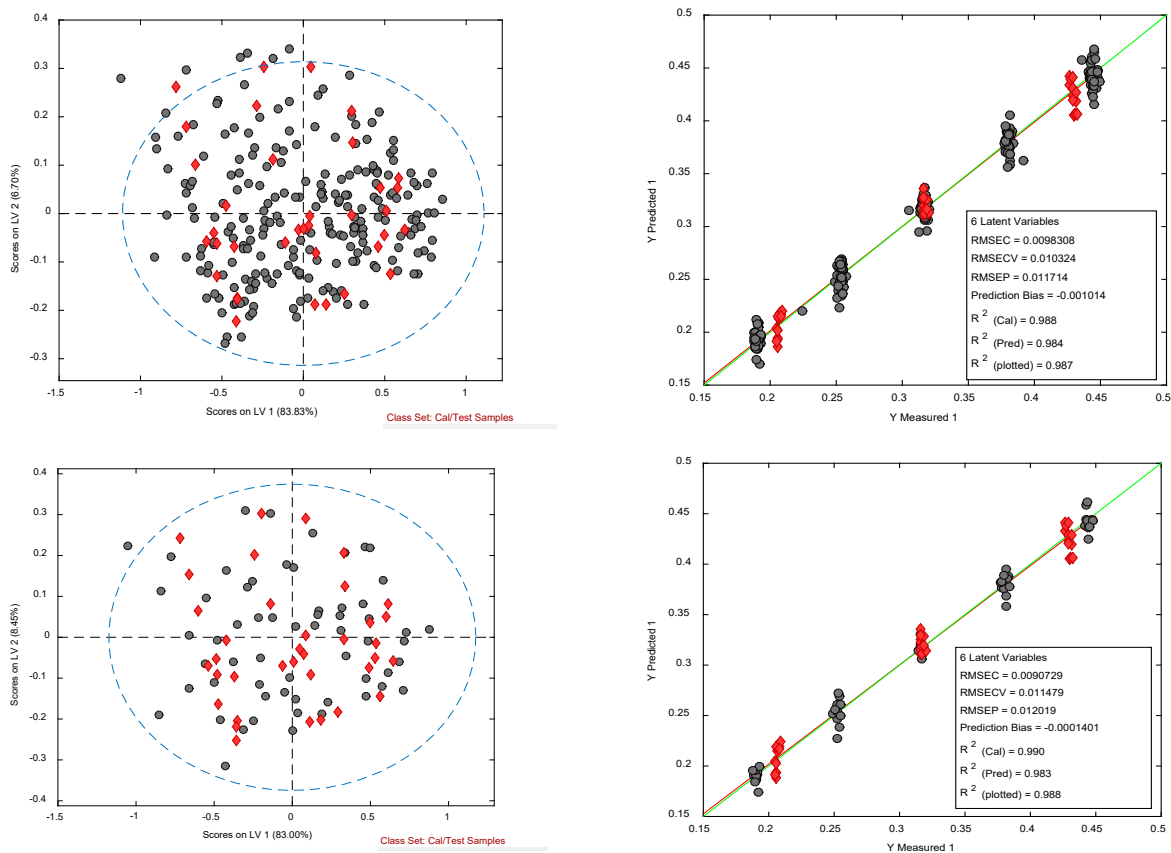


Figure 5-29. PLS model performance of the traditional (A) and spectral calibration sets (B) predicting Acetaminophen in test set 2

5.3.2 Prediction of Caffeine

A PLS model was developed from the traditional calibration set (225 samples) to predict Caffeine concentration in the first test set (75 samples). The preprocessing techniques were MSC, Savitzky-Golay derivative (window size 15, second order polynomial, first derivative) and mean centering for spectral data and auto scaling for the concentration data. A set of 5 loading vectors was selected for model development. Figure 5-30 shows the score plot and reference vs prediction plot for the calibration model. Similar ranges of spectral variation were observed

between traditional calibration and first prediction set. The RMSEC and RMSEP were 0.35% w/w and 0.46% w/w, respectively.

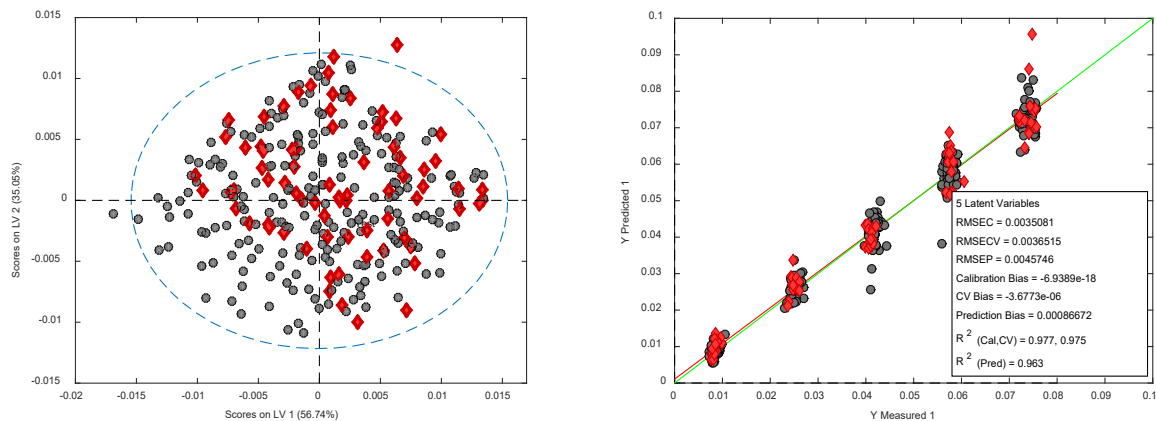


Figure 5-30. PLS model performance of the traditional calibration set predicting Caffeine in test set 1

Another PLS model was developed from the spectral calibration set (60 samples) of Caffeine. The preprocessing techniques were MSC, Savitzky-Golay derivative (window size 15, second order polynomial, first derivative) and mean centering for spectral data and auto scaling for the concentration data. A set of 5 loading vectors was selected for model development. Figure 5-31 shows the score plot and reference vs prediction plot for the calibration model. Similar ranges of spectral variation were observed between the calibration and prediction sets. The prediction performance of the spectral calibration set (RMSEP 0.46% w/w) was found to be similar to the traditional calibration set in spite of having fewer calibration samples (60 vs 225).

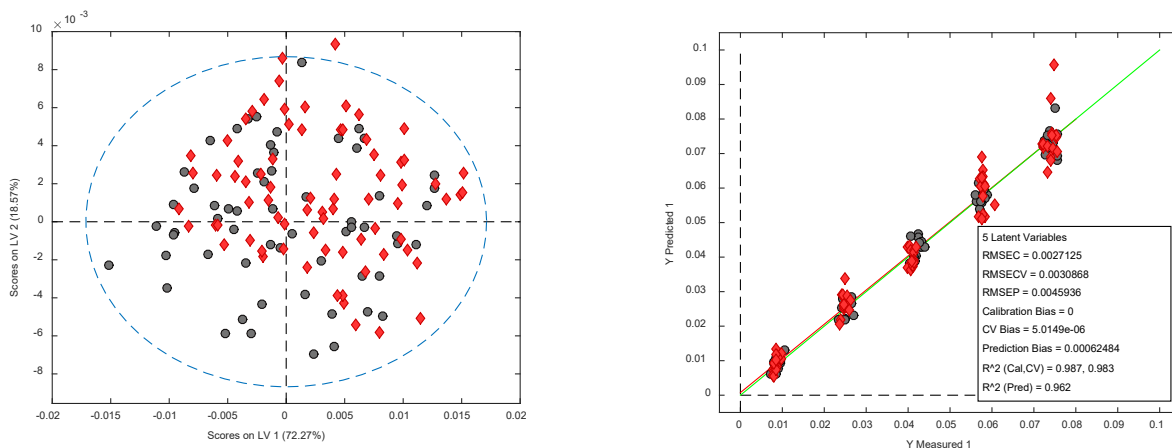


Figure 5-31. PLS model performance of the spectral calibration set predicting Caffeine in test set 1

The same calibration models of the traditional and spectral calibration sets were used to predict the concentration of Caffeine in the second test set (36 samples). The prediction performances for both calibration models were improved in the second calibration set. The extents of improvement were similar between the traditional and spectral calibration sets. The RMSEPs for both calibration models were $\sim 0.35\%$ w/w. Figure 5-32 shows the score plot and reference vs prediction plots for the two calibration models of traditional and spectral calibration sets.

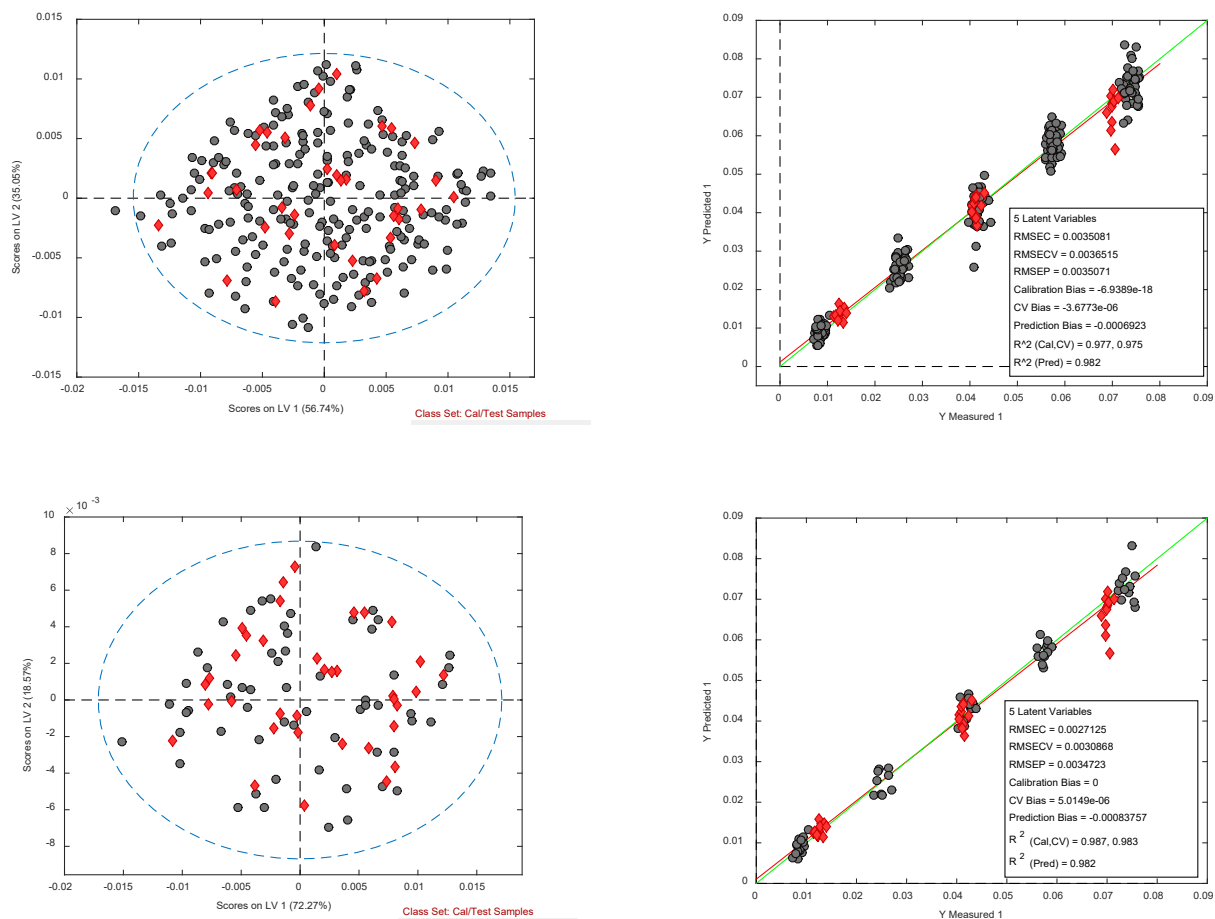


Figure 5-32. PLS model performance of the traditional (A) and spectral calibration sets (B) predicting Caffeine in test set 2

5.4 Conclusion

The spectral calibration sets (60 samples) were found to provide equivalent prediction performances compared to the traditional calibration set (225 samples) for both Acetaminophen and Caffeine. Equivalent prediction performances were observed in both test sets containing compositional and physico-chemical variation. The summary of the prediction performances is given in Table 5-3. The spectral calibration sets were designed by using prior information about the spectral effect of physico-chemical information. This information was utilized to efficiently design a small calibration set to incorporate wide range of information. Such incorporation was

successful in building a robust calibration model against compositional and physico-chemical variation.

	Acetaminophen model		Caffeine model	
	Full factorial design	Spectral design	Full factorial design	Spectral design
No. of Samples	225	60	225	60
RMSEC (% w/w)	0.983	0.907	0.351	0.271
RMSECV (% w/w)	1.032	1.148	0.365	0.309
R ² Calibration	0.988	0.990	0.977	0.987
RMSEP1 (% w/w)	1.001	0.996	0.457	0.459
Prediction Bias 1 (% w/w)	-0.472	-0.338	0.087	0.062
R ² Prediction 1	0.990	0.989	0.963	0.962
RMSEP2 (% w/w)	1.171	1.202	0.351	0.347
Prediction Bias 2 (% w/w)	-0.101	-0.014	-0.069	-0.084
R ² Prediction 2	0.984	0.983	0.982	0.982

Table 5-3. Summary of the calibration model performance of the traditional and spectral calibration set

6 Chapter 6: Optimum calibration structure for pharmaceutical formulation and spectroscopic techniques

6.1 Introduction

Calibration structure is a critical factor for ensuring desired performance of the spectroscopic method for quantitative analysis of pharmaceutical formulation. The optimum calibration structure depends on the sample formulation and spectroscopic technique. An optimum calibration structure for an NIR method can be sub-optimum for a different spectroscopic method such as Raman or THz spectroscopy due to inherent differences in spectral responses. An optimum calibration structure for one formulation can also be sub-optimum for another formulation. This is the fundamental concept leading to the development of spectral design strategy for calibration development. In the traditional strategy, same set of samples designed in the concentration space is usually used for both NIR and Raman calibration method development. These samples are designed to allow wide concentration and physical variation to follow the traditional concept of incorporating as much variance as possible into the calibration set. However, all the variance information in the calibration set may not be relevant to the particular technique of interest. For instance, a NIR method can be sensitive to a particular type of physical variation that needs to be incorporated into the calibration set to develop a robust quantitative model, whereas the same physical variation may not affect Raman spectra and be unnecessary to incorporate into the Raman calibration set.

Spectral design strategy allows optimization of the calibration structure depending on the formulation and technique. Since the strategy utilizes pure component spectral information, the resultant calibration structure depends on the interaction between formulation and spectroscopic

technique. A unique calibration set is developed for each formulation and spectroscopic technique. However, the criticality of having unique calibration structure to ensure optimum calibration performance has not been investigated. In this study the optimum calibration structure between NIR and Raman were compared. The optimum calibration structures between two APIs (Acetaminophen and Caffeine) were also compared. This study also provided the basis for utilizing spectral design strategy to design a formulation and technique specific calibration set.

6.2 Material and Method:

6.2.1 Calibration and test set

The multiple API formulation was selected for this study. The multiple API tablets contained two APIs as Acetaminophen (Acetaminophen; Mallinckrodt Inc., Raleigh, NC, USA) and Caffeine anhydrous (Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA). The excipients were Microcrystalline cellulose (MCC; Avicel PH 200, FMC Biopolymer, Mechanicsburg, PA, USA), Lactose (modified spray-dried; Foremost Farms USA, Rothschild, WI, USA), Crosscarmellose sodium (Crosscarmellose Na, Spectrum Chemical Mfg. Corp., New Brunswick, NJ, USA) and Magnesium stearate (MgSt; Fisher Scientific, Waltham, MA, USA). The target formulation was set as Acetaminophen (31.25% w/w), Caffeine (4.05% w/w), MCC (37.32% w/w), Lactose (24.89% w/w), Crosscarmellose Na (2% w/w) and MgSt (0.5% w/w).

A full-factorial experimental design was created to vary the concentration of Acetaminophen, Caffeine, MCC:Lactose ratio and compaction force of the tablets. The Acetaminophen and Caffeine concentration was varied at 5 levels and MCC : Lactose ratio and compaction force were varied at 3 levels resulting in a total of 225 design points. A test set was developed by varying all the calibration factors except MCC: Lactose ratio, each at three levels resulting in a

total of 27 design points. Table 6-1 provides the details of the full factorial calibration and test design.

	Design Factors	Design Levels				
Calibration Design (5x5x3x3) 225 samples	Acetaminophen (%)	L.C.-40%	L.C.-20%	L.C.	L.C.+20%	L.C.+40%
	Caffeine (%)	L.C.-80%	L.C.-40%	L.C.	L.C.+40%	L.C.+80%
	MCC/Lac	1		1.5	2	
	Force (lb)	4000		5000	6000	
Test Design (3x3x3) 27 samples	Acetaminophen (%)	L.C.-35%		L.C.	L.C.+35%	
	Caffeine (%)	L.C.-70%		L.C.	L.C.+70%	
	Force (lb)	4000		5000	6000	
	MCC/Lac	1.5				

Table 6-1. Calibration and test design

Each tablet of the calibration and test design points was individually prepared by direct compression. All the components for a single design point were weighed using a digital weighing machine (Data Range, Model No. AX504DR, Mettler Toledo) and placed in a 10 ml scintillation vial. The ingredients were mixed in a bin blender (L.B. Bohle LLC, Warminster, PA, USA) for 10 mins followed by a high shear mixing using a vortex machine (Vortex-2 Genie, Model G-560, Scientific Industries, IN, USA). The final mixing was performed in the bin blender before compression. Tablets were compressed on a Carver Automatic Tablet Press (Model 3887.1SD0A00, Wabash, IN, USA) at respective compression forces using a 13 mm die and flat-faced punches. The target tablet weight was 700 mg.

NIR reflectance measurements for both sides of each compact were collected using a bench top scanning monochromator instrument (XDS Rapid Content Analyzer, FOSS NIRSystems, Inc., Laurel, MD, USA) after tablets reached stable dimensions (viscoelastic relaxation). Spectra corresponding to each side of a compact were averaged to give one spectrum per compact. Gravimetric measurement was used as reference for all tablets.

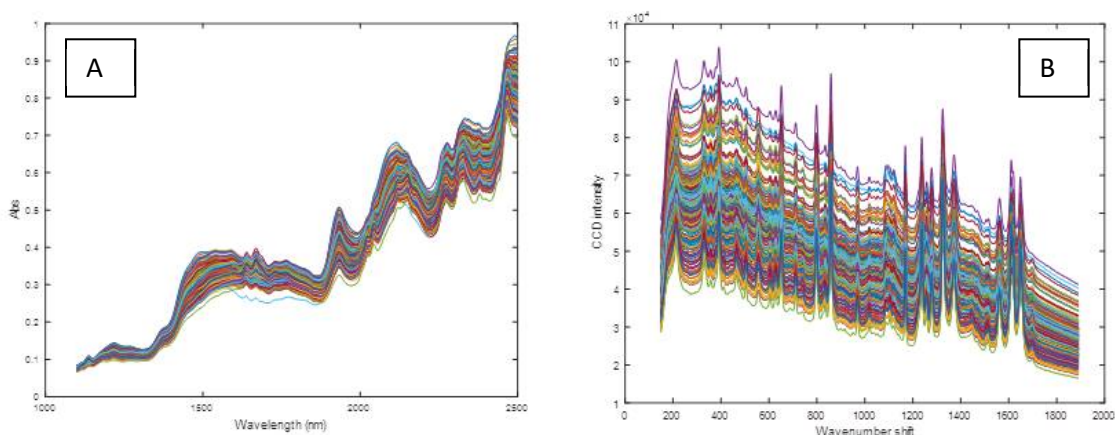


Figure 6-1. NIR (A) and Raman (B) spectra of the calibration sets

Raman measurements for both sides of each compact were collected using a PhAT System spectrometer coupled with a probe head (HoloGRAMS version 4.0, Kaiser Optical Systems, Inc, Ann Arbor, MI). The laser excitation wavelength was 785 nm. The PhAT System samples a spot size of ~6 mm. The integration time and co-adds were 15 seconds and 3, respectively. Spectra corresponding to each side of a compact were averaged to give one spectrum per compact over the range of 150 to 1890 cm^{-1} at a 0.2 cm^{-1} increment. The NIR and Raman spectra of the tablet are given in Figure 6-1.

All the analyses were performed in MATLAB 2015a environment (The Mathworks, Natick, MA, USA). Quantitative models were developed using Partial Least Squares (PLS) modeling

technique using PLS_Toolbox v. 7.9.3 (Eigenvector Research Inc., Wenatchee, WA, USA). Data independent spectral preprocessing techniques were used to optimize the model performance.

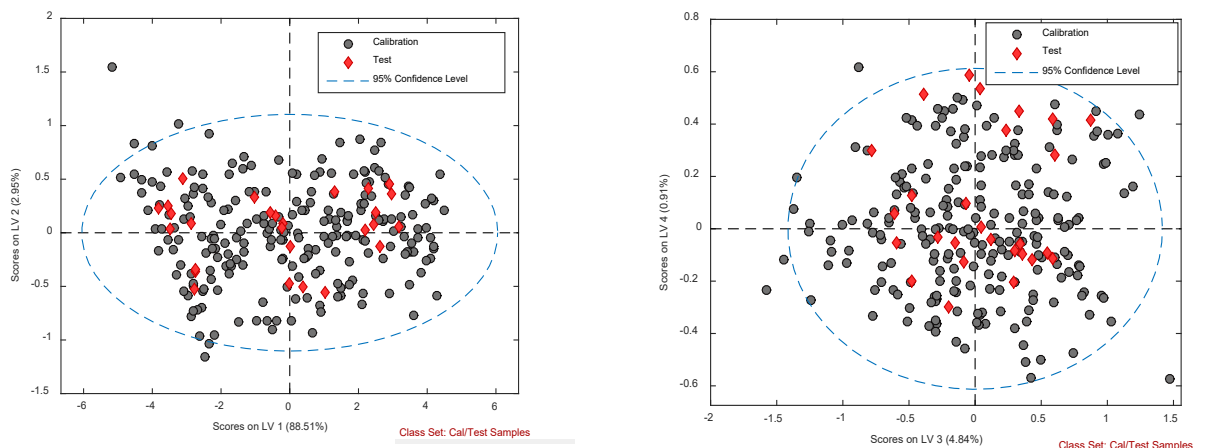


Figure 6-2. NIR PLS model for Acetaminophen from traditional calibration set

6.2.2 Quantitative model development

6.2.2.1 NIR infrared spectroscopy

6.2.2.1.1 Acetaminophen prediction

A PLS model was developed for predicting Acetaminophen concentration in calibration and test set tablets. During PLS model development, the optimum preprocessing techniques were found to be SNV and mean centering for the spectral data and auto scaling for the concentration data. The optimum number of loading vector (6) was selected from a routine cross validation step. The projected scores of the calibration and test samples on the first four loading vectors are given in Figure 6-2. Similar projections were found between calibration and test set tablets indicating similar types of spectral variations on the loading vectors direction. A lower prediction error in

both calibration (0.97% w/w) and test (1.04% w/w) set ensured reasonable model performance. The reference vs prediction plot is given in Figure 6-3.

This calibration structure was designed from a full factorial template without any prior information. The calibration structure was neither optimized for Acetaminophen, nor for NIR spectroscopy. The optimality of the calibration structure and calibration performance was unknown. Although the NIR method provided reasonable prediction performance, the best possible prediction performance with available resources was unknown. A mathematical search was performed to find the optimum calibration set that provided the best prediction performance (lowest prediction error) for the test set. It must be noted that the optimum calibration structure depends on the test set structure. A different calibration set could provide the best performance for a different test set. The search for the optimum calibration set followed the steps provided below:

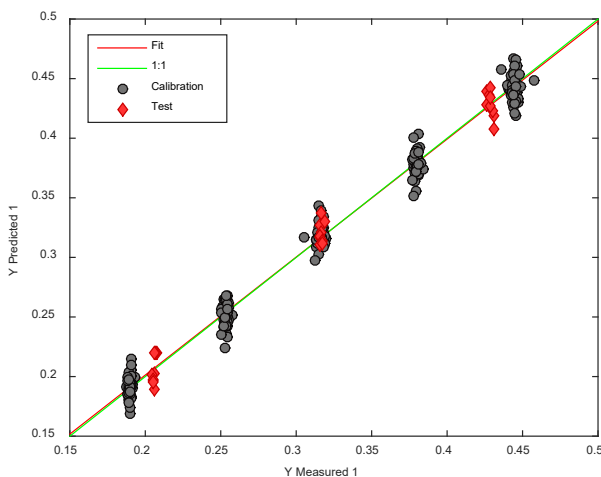


Figure 6-3. Reference vs prediction plot of Acetaminophen from traditional calibration set

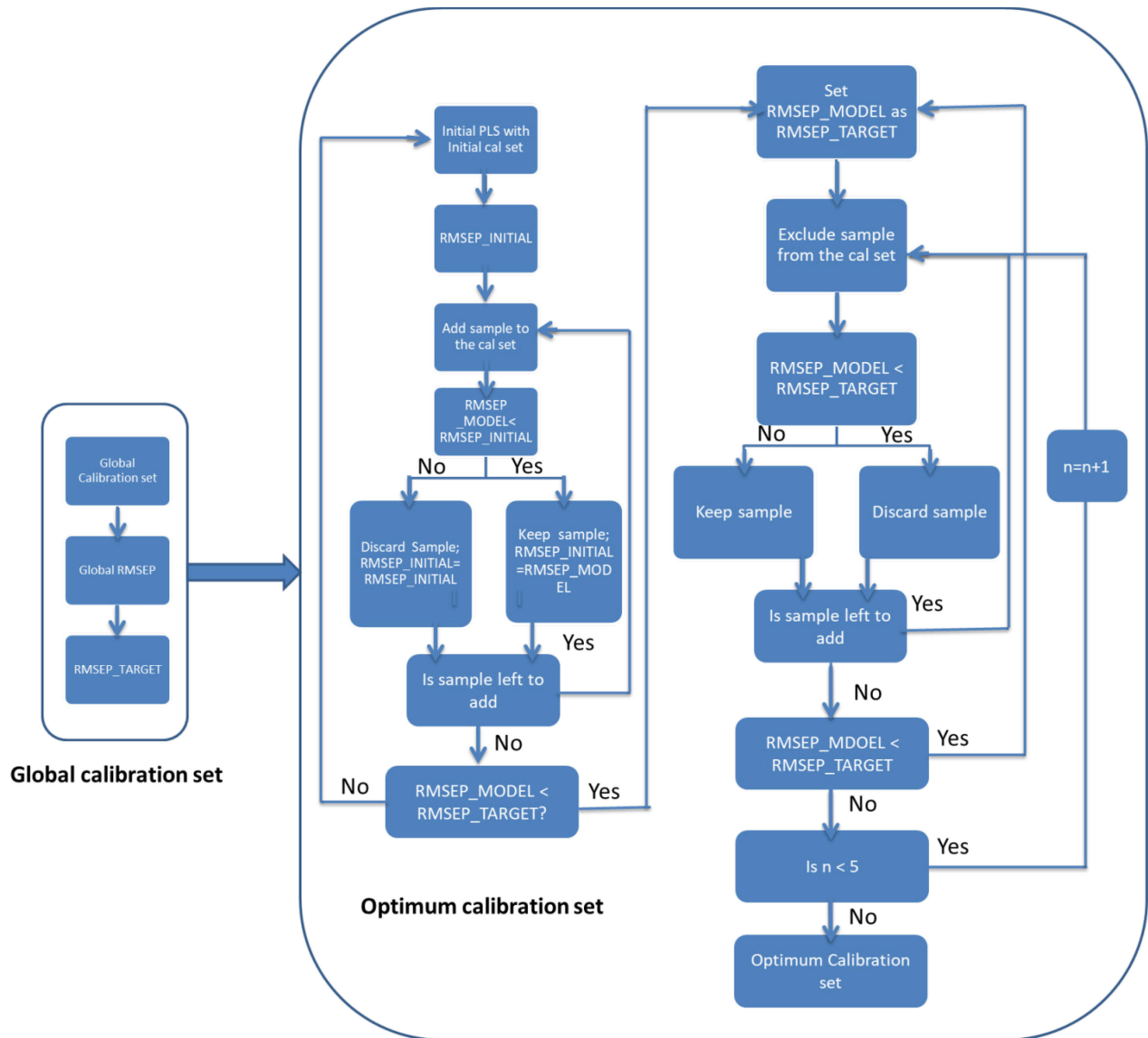


Figure 6-4. Search for optimum calibration set

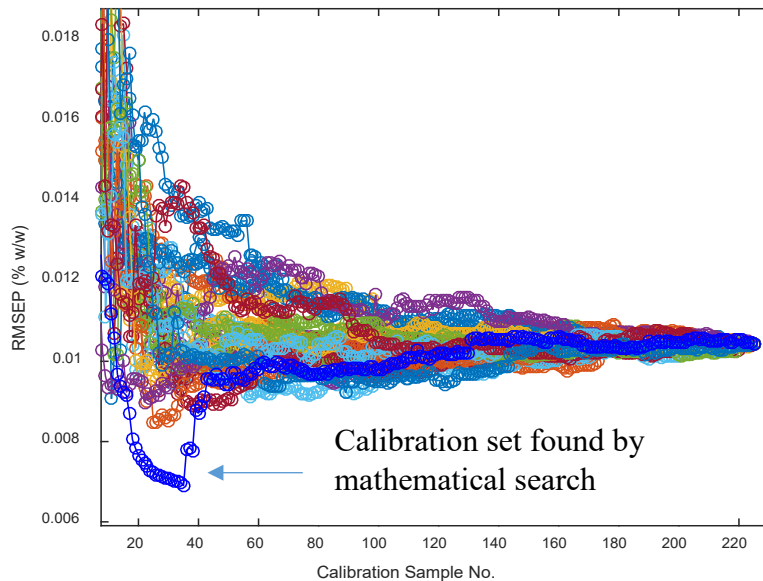


Figure 6-5. Prediction performance of randomly selected calibration sets and optimally selected calibration set

Initially a PLS calibration model was developed using all the calibration samples (global calibration set) and used for predicting Acetaminophen concentration in the test set. The calibration set was set as ‘Optimum calibration set’ and resultant RMSEP was set as ‘RMSEP_TARGET’. Then a second PLS model was developed using randomly selected 3 calibration samples (initial calibration set) and used for predicting Acetaminophen concentration in the same test set. The resultant RMSEP was set as ‘RMSEP_INITIAL’. Then another randomly selected calibration sample was added to the initial calibration set, a PLS model was developed using the updated calibration set and used to predict Acetaminophen concentration in the test set. The resultant RMSEP was set as ‘RMSEP_MODEL’. If RMSEP_MODEL was found to be lower than ‘RMSEP_INITIAL’, the newly added sample was kept in the updated calibration set, otherwise discarded from the updated calibration set. The RMSEP of the updated calibration set was set as ‘RMSEP_INITIAL’ (for sample inclusion it would change (lowered),

for sample exclusion it would remain same). Then another randomly selected calibration sample was added to the updated calibration set and effect of its inclusion on the prediction performance of the test set was calculated. It was kept in case it helped to improve the prediction performance, otherwise discarded. This process was iterated until all the calibration samples were tested for their effects on the prediction performance. A subset of calibration samples was found after the iteration. The RMSEP of the calibration subset (RMSEP_MODEL) was compared with RMSEP of the global calibration set (RMSEP_TARGET). If the RMSEP_MODEL was found to be lower than RMSEP_TARGET, a set of better calibration candidates was found by sample inclusion strategy. The algorithm then went to the next phase to analyze the effect of sample exclusion on the prediction performance. If RMSEP_MODEL was found to be higher than RMSEP_TARGET, the algorithm started from the beginning by randomly selecting a different initial set and calibration candidates.

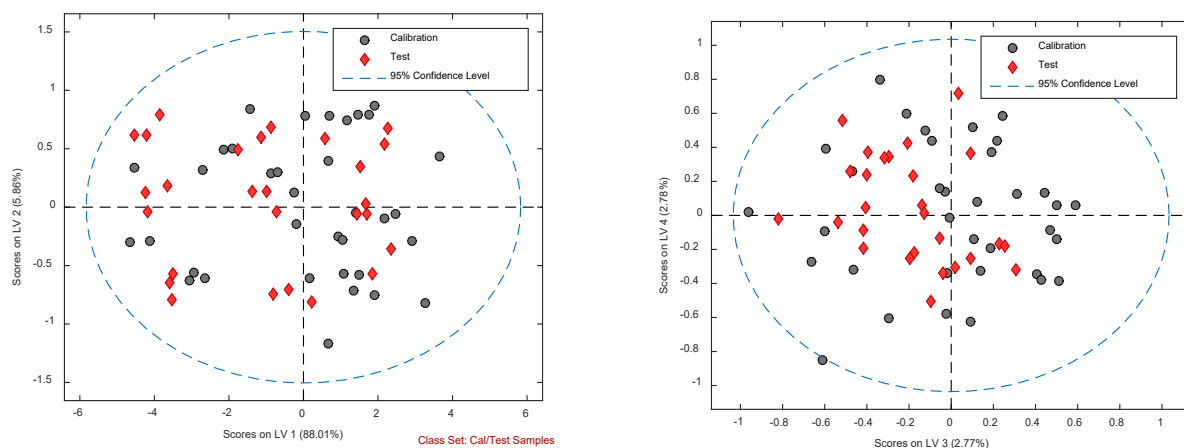


Figure 6-6. NIR PLS model for Acetaminophen from optimum calibration set

In the next phase, a randomly selected sample was excluded from the updated calibration set, a PLS model was developed and used to predict Acetaminophen concentration in the test set. The

sample was excluded in case the exclusion lowered the RMSEP, otherwise it was kept in the updated calibration set. In the next step, another randomly selected sample was tested for exclusion criteria. This process was iterated until all the samples were individually tested to analyze the effect of their exclusion on the prediction performance. A calibration subset containing better calibration candidates was found at the end. All these calibration samples were tested for both inclusion and exclusion criteria and found to improve the prediction performance. The RMSEP of the updated calibration set (RMSEP_MODEL) was calculated and compared with the RMSEP_TARGET. If RMSEP_MODEL was found to be lower than the RMSEP_TARGET, the RMSEP_MODEL was set as the new RMSEP_TARGET. The updated calibration set was stored and set as the 'Optimum calibration set'. If RMSEP_MODEL was found to be higher than the RMSEP_TARGET, the RMSEP_TARGET and 'Optimum calibration set' remained same. This whole process was repeated iteratively, until RMSEP_TARGET (lowest RMSEP at hand) was found to be lower than the RMSEP_MODEL for five consecutive iterations. It indicated that the search algorithm found the 'Optimum calibration set' that provided the best calibration performance (lowest RMSEP). The performance of this guided search algorithm was compared with multiple randomized search algorithms. The resultant RMSEPs of the search algorithms are shown in Figure 6-5 with the calibration sample number. The guided search algorithm was able to find the calibration set that provided the best model performance. Following the guided search, the optimum calibration set for Acetaminophen was found to contain 35 calibration samples.

A PLS model was developed from this optimum calibration set. The preprocessing techniques were SNV and mean centering for the spectral data and auto scaling for the concentration data, respectively. A set of 6 loading vectors was selected for model development. The scores of the

calibration and test samples on the first four loading vectors are given in Figure 6-6. It was found that, the small optimum calibration set was able to explain similar spectral variation as the test set. An improved prediction performance (RMSEC 0.49 % w/w and RMSEP 0.69% w/w) was observed from the optimum calibration set as compared to the global calibration set. The reference vs predicted concentration for calibration and test sample is given in Figure 6-7.

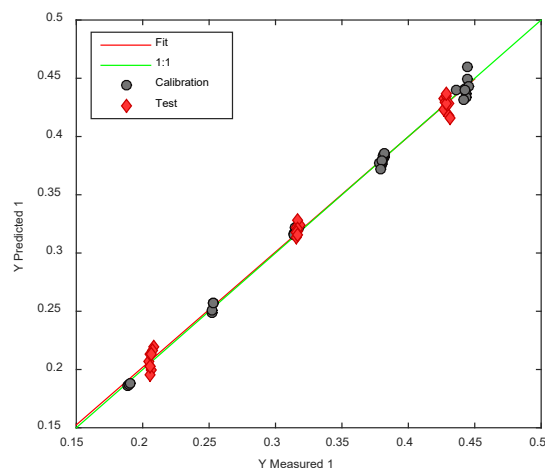


Figure 6-7. Reference vs prediction plot of Acetaminophen from optimum calibration set

6.2.2.1.2 Caffeine prediction

A PLS model was developed from all the 225 samples to predict the concentration of Caffeine in the test set. The optimal preprocessing techniques were SNV and mean centering for spectral data and auto scaling for concentration data. A set of 6 latent variables was selected for model development. The scores of the calibration and test samples on the first four loading vectors are given in Figure 6-8. The calibration and test samples were found to span similar spectral variation on the loading vector directions.

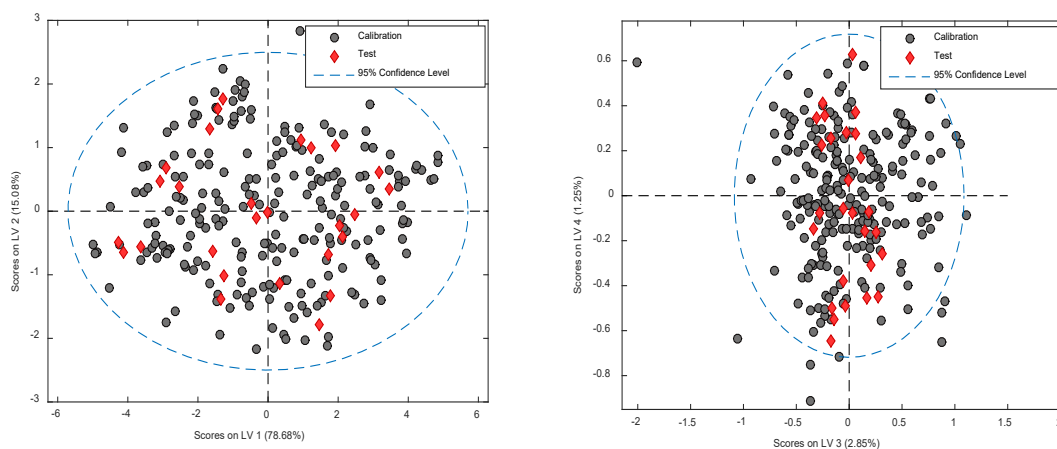


Figure 6-8. NIR PLS model for Caffeine from traditional calibration set

Similar projections of calibration and test samples ensured reasonable prediction performance of the calibration model. The resultant RMSEC and RMSEP were 0.40% w/w and 0.43% w/w, respectively. The reference vs predicted concentration for calibration and test sample is given in Figure 6-9.

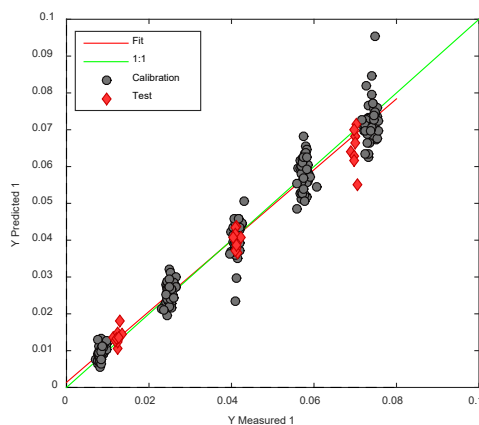


Figure 6-9. Reference vs prediction plot of Caffeine from traditional calibration set

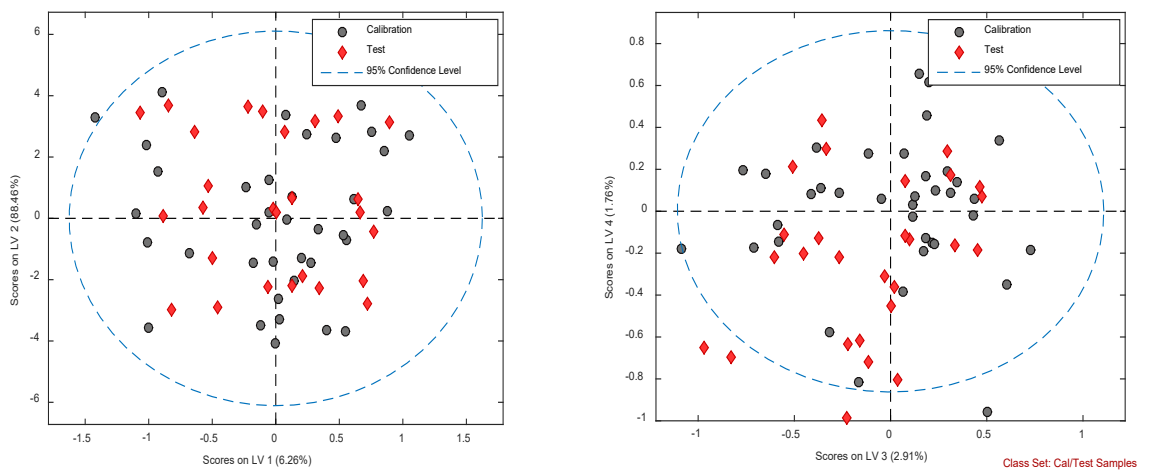


Figure 6-10. NIR PLS model for Caffeine from optimal calibration set

A guided search was performed following the previously described algorithm to find the optimum calibration set for Caffeine prediction. The optimum calibration set of Caffeine was selected based on the prediction performance on the test set. The resultant optimized calibration set contained 37 samples. A PLS model was developed from the optimum calibration set. The preprocessing techniques were SNV and mean centering for the spectral data and auto scaling for the concentration data. A set of 6 loading vectors was selected.

Figure 6-10 shows the calibration and test sample scores on the first four loading vectors. It was found that, the small optimum calibration set spanned the entire range of spectral variation of the test set samples. Such spectral coverage ensured improved model performance compared to the global calibration set in spite of having fewer sample number. The resultant RMSEC and RMSEP were 0.19% w/w and 0.28% w/w, respectively. The reference vs prediction plot is given in Figure 6-11. A small optimum calibration set was able to provide better model performance compared to the global calibration for both Acetaminophen and Caffeine. These optimum calibrations sets were found by the search algorithm. The optimum calibration sets between

Acetaminophen and Caffeine were different. Only 8 samples were found to be common between the optimum calibration set for Acetaminophen (35 samples) and Caffeine (37).

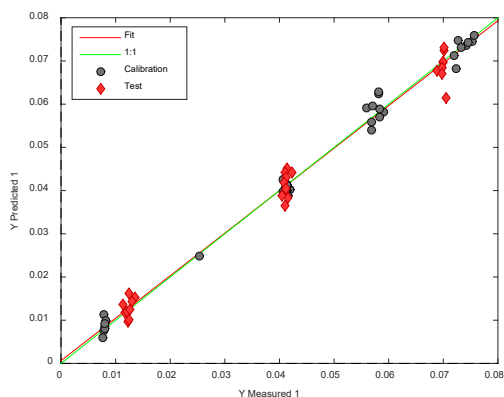


Figure 6-11. Reference vs prediction plot of Caffeine from traditional calibration set

6.2.2.2 Raman spectroscopy

The same dataset was scanned with Raman spectroscopy. Quantitative PLS models were developed from Raman spectra to predict the concentrations of Acetaminophen and Caffeine in the test set.

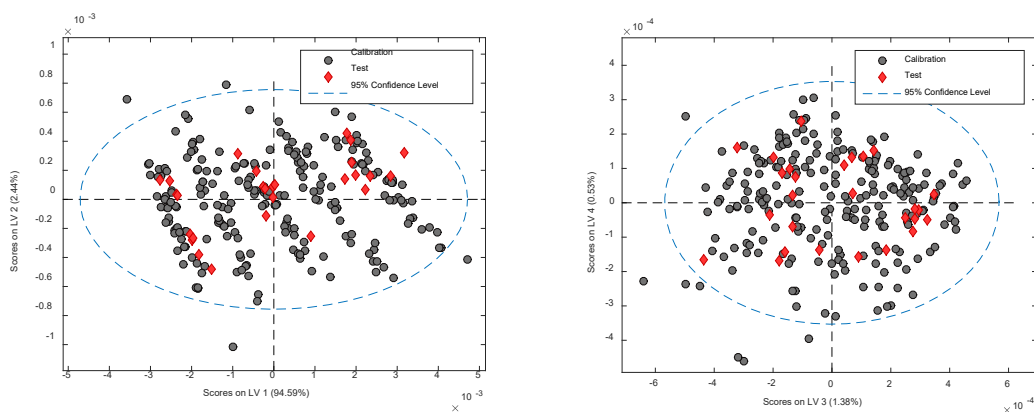


Figure 6-12. Raman PLS model for Caffeine from traditional calibration set

6.2.2.2.1 Acetaminophen prediction

A PLS model was developed to predict the concentration of Acetaminophen in the test set using Raman spectra. The optimum preprocessing techniques were SNV, normalization and mean centering for the spectral data and auto scaling for the concentration data. A set of 4 latent variables were selected for the model development. Figure 6-12 shows the projection of the calibration and test samples on the first four loading vectors. The first two loading vectors grouped the calibration (5 groups) and test samples (3 groups) based on the Acetaminophen concentration. It was also observed that the calibration and test samples spanned similar ranges of spectral variation. A reasonable prediction performance was achieved. The RMSEC and RMSEP were 1.02 % w/w and 1.09 % w/w, respectively.

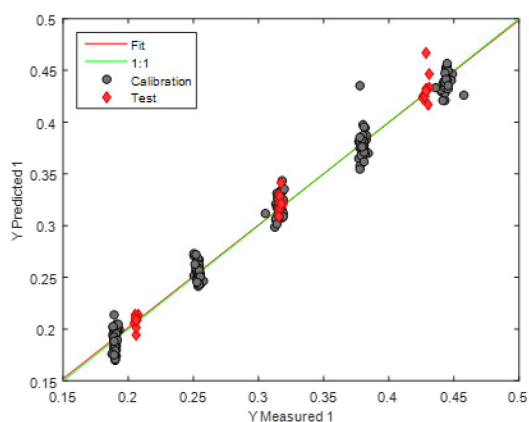


Figure 6-13. Reference vs prediction plot of Acetaminophen from traditional calibration set

The reference vs prediction plot is shown in Figure 6-13. An optimum subset was selected from 225 samples using the search algorithm described earlier. The resultant optimum calibration set contained 71 samples. A PLS model was developed from the selected 71 samples. The preprocessing techniques were SNV, normalization and mean centering for the spectral data and

auto scaling for the concentration data. A set of 4 latent variables was selected for model development.

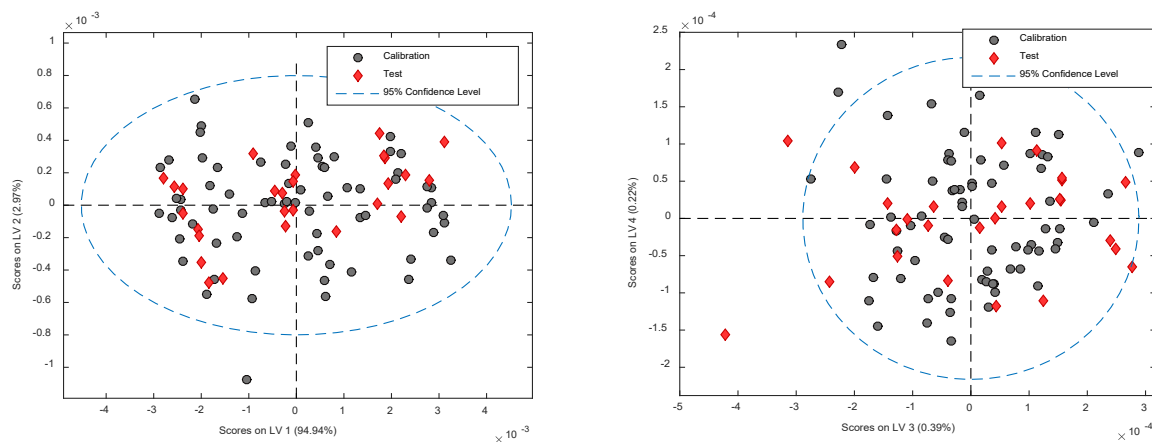


Figure 6-14. Raman PLS model for Acetaminophen from optimal calibration set

Figure 6-14 shows the projections of the calibration and test samples on the first four loading vectors. Similar to the global calibration set (225 samples), calibration and test samples were grouped based on the Acetaminophen concentration on the first two loading vectors. The extents of variation explained by the first two loading vectors were also similar to that of global calibration set. The calibration and test samples had similar projection on the loading vector. An improvement in the prediction performance was observed from the optimum calibration set compared to the global calibration set. The RMSEC and RMSEP were 1.05 % w/w and 0.79 % w/w, respectively. The reference vs prediction plot is shown in Figure 6-15.

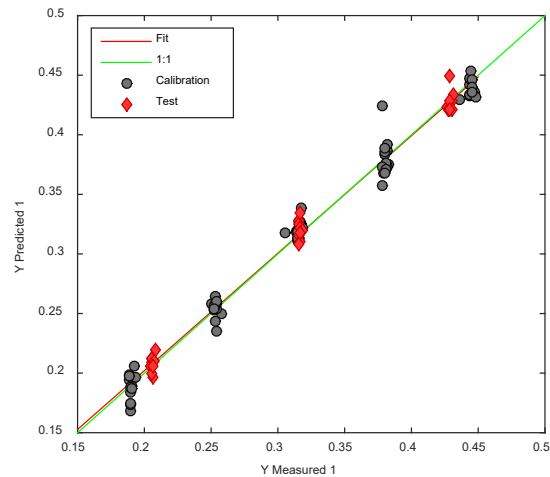


Figure 6-15. Reference vs prediction plot of Acetaminophen from optimum calibration set

6.2.2.2.2 Caffeine prediction

A PLS model was also developed to predict the concentration of Caffeine in the test set using Raman spectra. The optimum preprocessing techniques were SNV and mean centering for the spectral data and auto scaling for the concentration data. A set of 4 latent variables was selected for model development. Figure 6-16 shows the projections of the calibration and test samples on the first four loading vectors. No clear grouping based on the Caffeine concentrations was observed in the score plot. The first four loading vectors explained convoluted information from Caffeine concentrations and other sources of spectral variation. The Caffeine concentration was not a dominant factor due to low concentration in the tablets.

The calibration and test samples had similar ranges of spectral variation. A reasonable prediction performance was achieved from the global calibration set. The RMSEC and RMSEP were 0.59 % w/w and 0.55 % w/w, respectively. The reference vs prediction plot is shown in Figure 6-17.

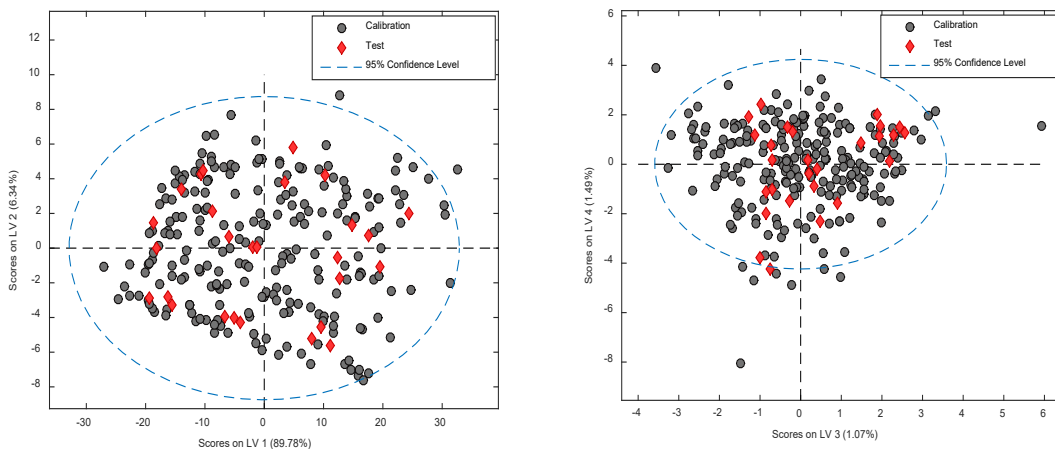


Figure 6-16. Raman PLS model for Caffeine from traditional calibration set

An optimum subset was selected from 225 samples using the search algorithm described earlier. The resultant optimum calibration set contained 58 samples. A PLS model was developed from this optimum calibration set. The preprocessing techniques were SNV and mean centering for the spectral data and auto scaling for the concentration data. A set of 4 latent variables was selected for model development.

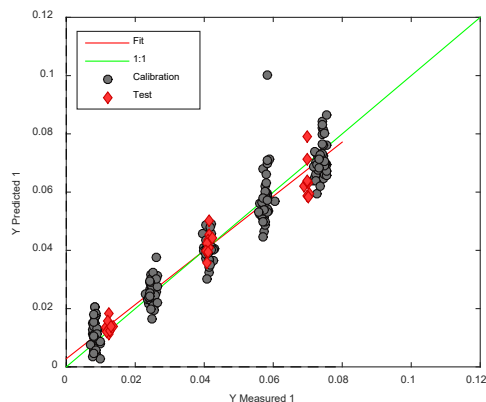


Figure 6-17. Reference vs prediction plot of Caffeine from traditional calibration set.

Figure 6-18 shows the calibration and test scores on the first four loading vectors. No grouping based on Caffeine concentration was observed in the score plot. The calibration and test set had similar ranges of spectral variation as shown in the score plot. An improved prediction performance was observed from the optimum calibration set as compared to the global calibration set. The RMSEC and RMSEP were 0.38 % w/w and 0.46 % w/w, respectively. The reference vs prediction plot is shown in Figure 6-19.

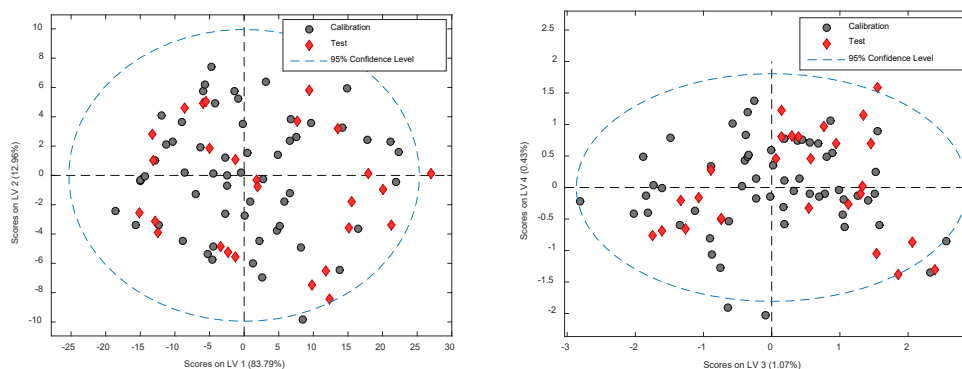


Figure 6-18. Raman PLS model for Caffeine from optimum calibration set

The optimum calibration sets with fewer sample numbers were able to provide better model performances compared to the global calibration sets with higher sample numbers for both

Acetaminophen and Caffeine. The optimum calibration sets between Acetaminophen and Caffeine were found to be different. Only 24 samples were found common between the optimum calibration sets for Acetaminophen (71 samples) and Caffeine (58 samples). This scenario was similar to the scenario in NIR spectroscopy.

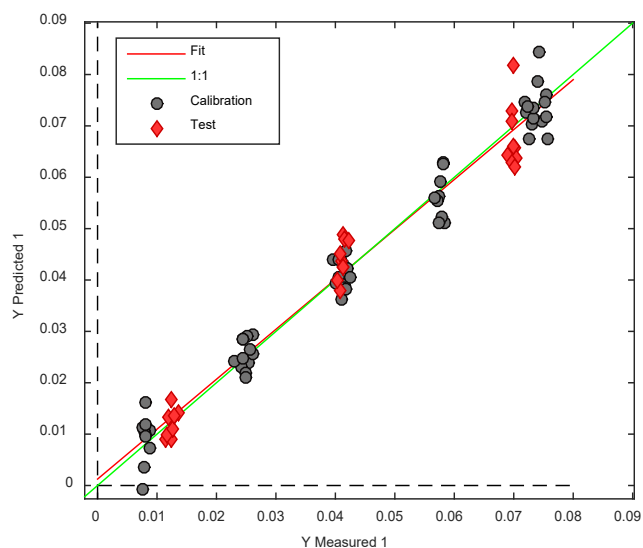


Figure 6-19. Reference vs prediction plot of Caffeine from optimum calibration set.

6.3 Results and discussion

6.3.1 Comparison between optimum calibration sets of NIR and Raman spectroscopy

The optimum calibration sets between NIR and Raman spectroscopy were different for both Acetaminophen and Caffeine. The comparison of the optimum calibration sets are discussed in the following sections.

6.3.1.1 Optimum calibration sets of Acetaminophen

The optimum calibration set of NIR spectroscopy required fewer samples compared to the optimum calibration set of Raman spectroscopy (35 vs 71) during the prediction of Acetaminophen in the test set. The optimum structures of the calibration sets were different from each other. Only 14 samples were common between the two calibration sets.

Figure 6-20 shows the compositional points between full global calibration set, optimum NIR and Raman calibration sets. The optimum NIR calibration set had lower prediction error compared to that of optimum Raman calibration set (RMSEP 0.69% w/w vs 0.79% w/w) in spite of having fewer sample number. However, a significant test was performed following the method described in section 5.2.8 in chapter 5 and no significant difference were found between their prediction performances.

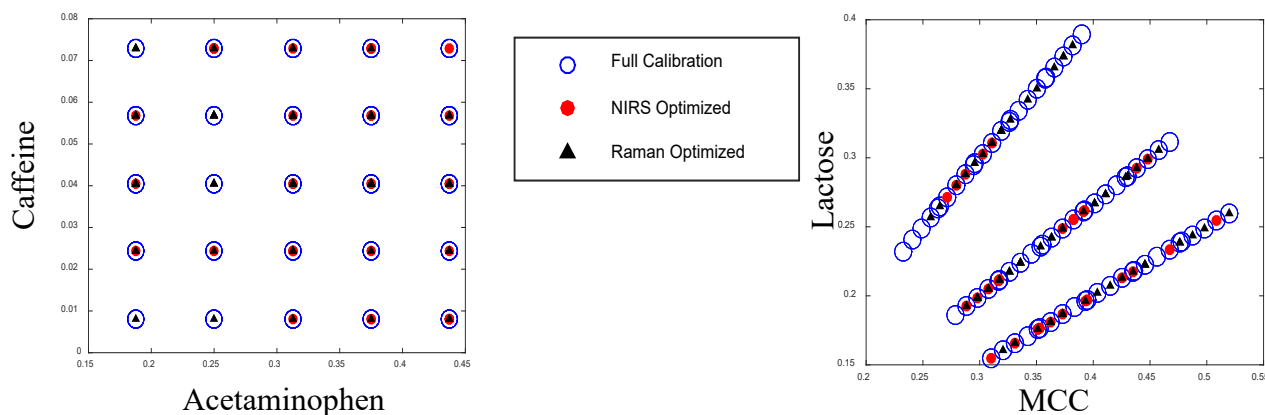


Figure 6-20. Compositional points between full global, NIR optimized and Raman optimized calibration sets for Acetaminophen.

The distribution of Acetaminophen concentration in optimum calibration sets of NIR and Raman is provided in Figure 6-21. The NIR calibration set required fewer samples compared to the

Raman set, especially at low concentration of Acetaminophen. These samples were selected from traditional full factorial calibration set. The traditional full factorial calibration structure was not fully orthogonal. In the calibration structure, there was a correlation between Acetaminophen concentration and MCC concentration as shown in Figure 6-22. At low Acetaminophen concentration, the MCC concentration was usually high. MCC has a strong background features on Raman spectra due to fluorescence. Due to such strong spectral features of MCC, Raman calibration set required more samples at high MCC concentration to make the calibration set robust against fluorescence.

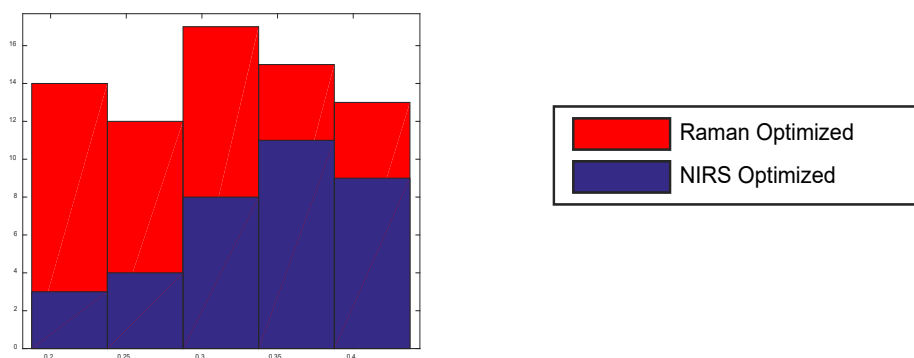


Figure 6-21. Acetaminophen distribution between NIR and Raman optimized calibration sets

MCC does not have any strong fluorescence like unique features on NIR spectra. So, high concentration of MCC at the low level of Acetaminophen did not direct the optimum calibration set to select higher number of samples.

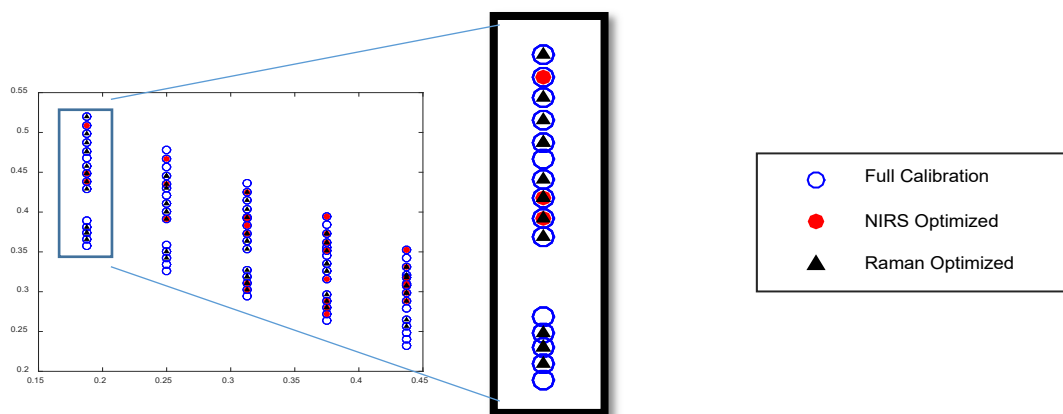


Figure 6-22. Concentrations of Acetaminophen and MCC in full traditional, NIR optimized and Raman optimized calibration sets

These two optimum calibration sets were also different in terms of excipient concentration. The distribution of MCC and Lactose between Optimum NIR and Raman calibration sets are shown in Figure 6-23. The optimum calibration for NIR required narrower distribution of excipients, especially for Lactose.

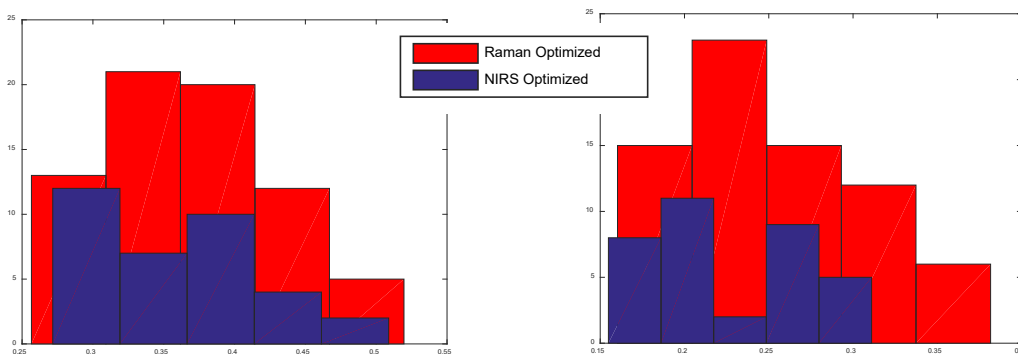


Figure 6-23. MCC (left) and Lactose (right) distribution between NIR and Raman optimized calibration sets for Acetaminophen

The overall required spectral variation from excipient was also small for NIR calibration set compared to that of Raman calibration set. This was due to the similarity of the spectral

responses between MCC and Lactose in NIR spectroscopy. MCC and Lactose has very similar NIRS response as shown in Figure 6-24. Due to such similarity, the excipients were not required to be varied simultaneously. Variance in MCC would provide necessary spectral variation in the calibration structure. As a result, the optimum calibration set provided good prediction performance without significant variation in the Lactose concentration. However, in the Raman spectroscopy MCC and Lactose have very different spectral responses as shown in Figure 6-24. The optimum Raman calibration set required variation in both MCC and Lactose concentration due to their unique information. This resulted in higher number of calibration samples in the optimum Raman calibration set compared to that of NIR spectroscopy.

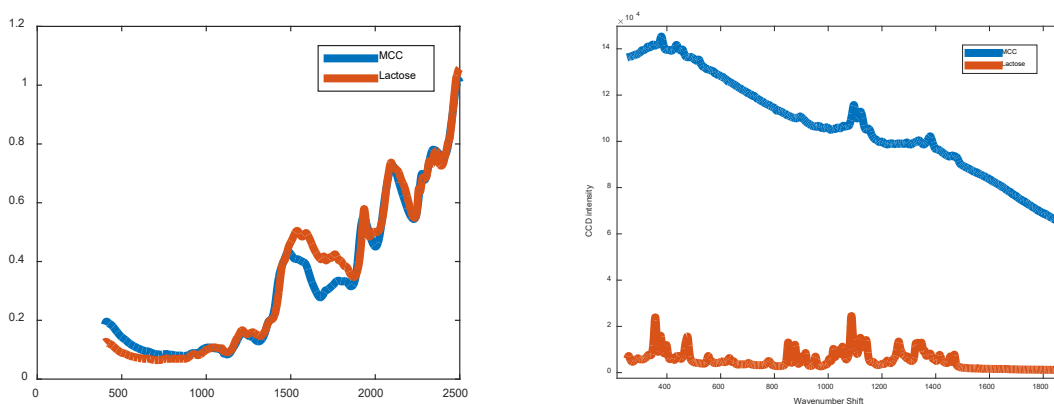


Figure 6-24. NIR (left) and Raman (right) spectra of MCC and Lactose

6.3.1.2 Optimum calibration sets of Caffeine

The optimum calibration structures of Caffeine were found to be different between NIR and Raman spectroscopy. The optimum calibration set of NIR spectroscopy required fewer samples compared to the optimum calibration set of Raman spectroscopy (37 vs 58) during the prediction

of Caffeine; a scenario similar to the Acetaminophen calibration set. Only 10 samples were common between the two calibration sets.

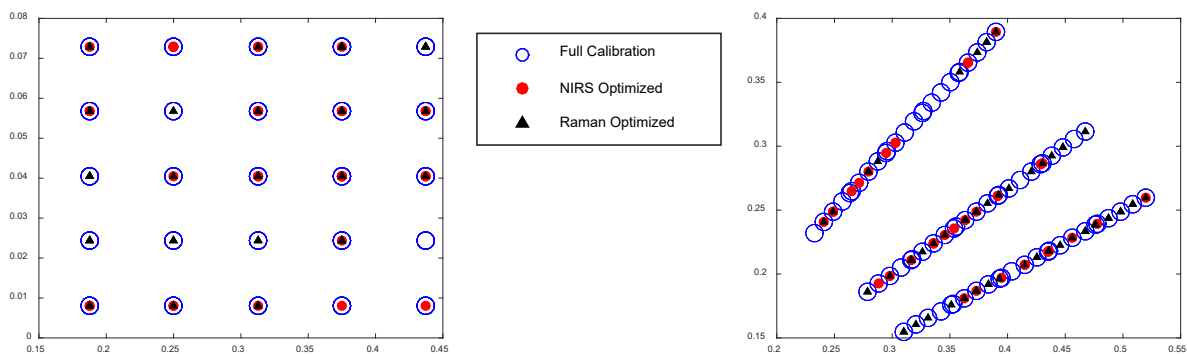


Figure 6-25. Compositional points between full global, NIR optimized and Raman optimized calibration sets for Caffeine.

Figure 6-25 shows the compositional points between global calibration set, optimum NIR and Raman calibration sets for Caffeine. The prediction performance of the optimum NIR calibration set was significantly better than the prediction performance of the optimum Raman calibration set (RMSEP 0.28% w/w vs 0.46% w/w). The significance test was performed following the method described in section 5.2.8 in chapter 5.

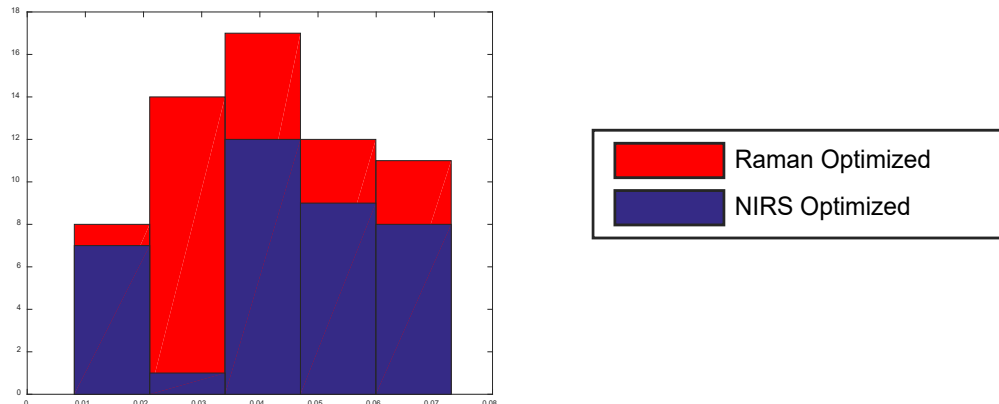


Figure 6-26. Caffeine distribution between NIR and Raman optimized calibration sets

The distribution of Caffeine concentrations between the NIR and Raman calibration sets is shown in Figure 6-26. Similar distributions were observed between NIR and Raman spectra. The NIR calibration set did not require fewer calibration samples at low Caffeine concentration as it did for the optimum calibration set of Acetaminophen. No correlation between Caffeine and MCC concentration was found as shown in Figure 6-27. MCC concentrations were relatively similar at all levels of Caffeine concentration resulting in similar distribution of Caffeine between optimum calibration sets of NIR and Raman spectroscopy.

The excipient distributions (MCC and Lactose) between two calibration sets are given in Figure 6-28. The optimum NIR calibration set required fewer calibration samples due to similarity between the spectral response of MCC and Lactose. However, the optimum Raman calibration set required higher number of calibration samples due to unique information from both excipients.

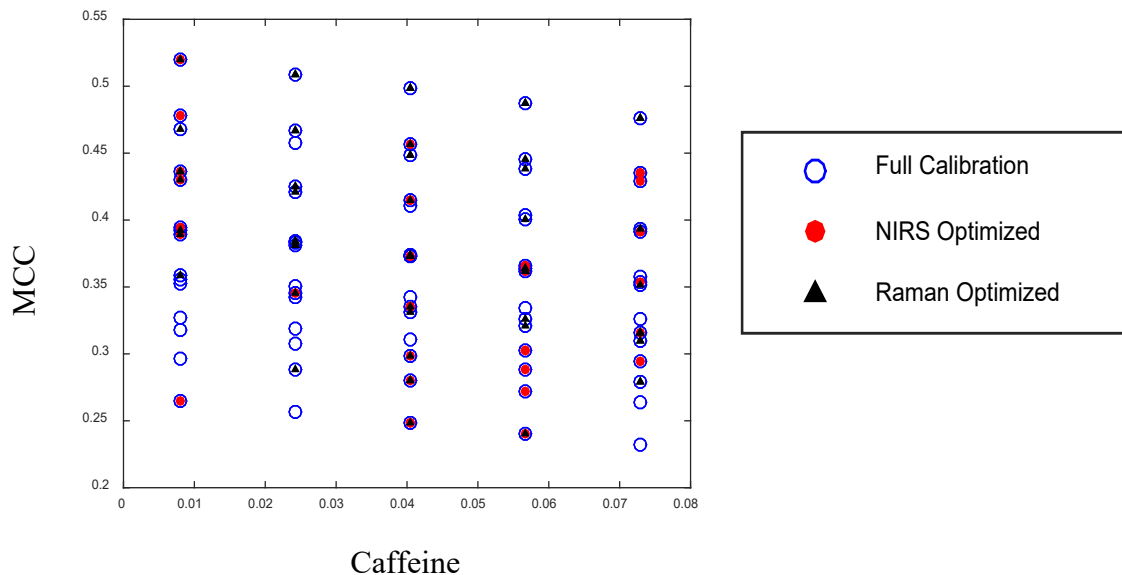


Figure 6-27. Concentrations of Caffeine and MCC in full traditional, NIR optimized and Raman optimized calibration sets

The optimum calibration sets of NIR and Raman spectroscopy were found to be significantly different from each other for both APIs due to inherent differences in the analytical techniques. However, the current practice of calibration design does not account for such variations in the analytical techniques. A similar strategy is followed for both techniques resulting in sub-optimal calibration performance. A study was performed to analyze the criticality of the calibration sets to be technique specific.

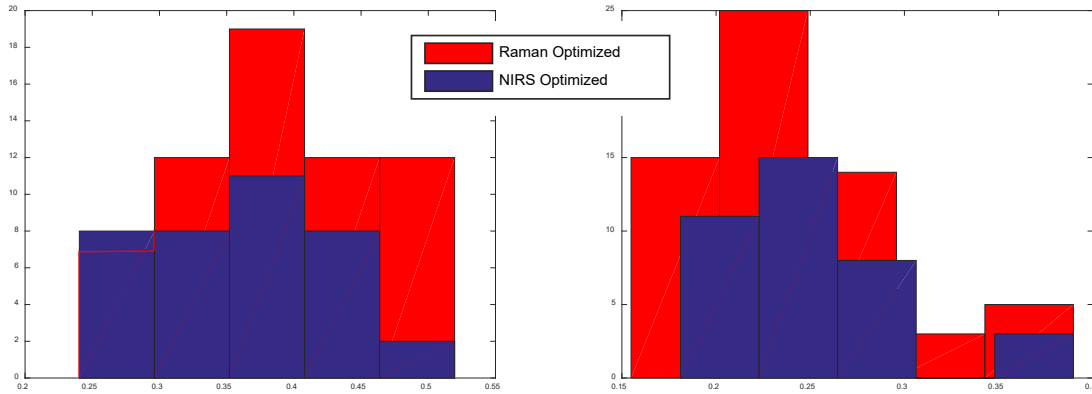


Figure 6-28. MCC (left) and Lactose (right) distribution between NIR and Raman optimized calibration sets for Caffeine

In this study, a Raman calibration method was developed from the optimal calibration structure of NIR spectroscopy for both Acetaminophen and Caffeine. Similarly, an NIR calibration

Calibration set	NIRS			Raman		
	Global	NIR Optimum	Raman Optimum	Global	Raman Optimum	NIR Optimum
Sample no.	225	35	71	225	71	35
LV	6	6	6	4	4	4
RMSEC	0.97	0.49	1.01	1.02	1.05	0.78
RMSECV	1.03	0.88	1.17	1.11	1.27	1.01
RMSEP	1.04	0.69	1.13	1.09	0.79	1.13
Bias	0.08	0.68	0.15	0.39	0.14	0.4
SEP	1.04	0.13	1.12	1.02	0.77	1.05

Table 6-2. Model statistics for global, optimum and cross calibration sets NIR and Raman during the prediction of Acetaminophen

method was developed from the optimal calibration structure of Raman spectroscopy. For instance, during the prediction of Acetaminophen, the optimal calibration sets for NIR and

Raman spectroscopy contained 35 and 71 calibration samples, respectively. A Raman method was developed from the optimal NIR calibration set that contained 35 samples and an NIR method was developed from the optimal Raman calibration set that contained 71 samples. This study was conducted for both Acetaminophen and Caffeine. PLS models were developed from the respective calibration sets. Model performances were compared with the optimal and global calibration performance. The results for Acetaminophen prediction are provided in Table 6-2.

It was shown that, the prediction performance of the calibration set deteriorated when the optimal structure of a different spectroscopic technique was used. The RMSEP of the optimal NIR calibration set (35 samples) was 0.69 % w/w. The prediction performance was affected

Calibration set	NIRS			Raman		
	Global	NIR Optimum	Raman Optimum	Global	Raman Optimum	NIR Optimum
Sample no.	225	37	58	225	58	37
LV	6	6	6	4	4	4
RMSEC	0.4	0.19	0.4	0.59	0.38	0.48
RMSECV	0.42	0.27	0.52	0.63	0.49	0.62
RMSEP	0.43	0.28	0.48	0.55	0.46	0.6
Bias	0.15	0.02	0.14	0.09	0.02	0.05
SEP	0.41	0.28	0.45	0.54	0.46	0.6

Table 6-3. Model statistics for global, optimum and cross calibration sets NIR and Raman during the prediction of Caffeine

significantly (RMSEP 1.13% w/w) when the optimal structure of Raman calibration set (71 samples) was used to develop a NIR calibration model to quantify the same API. The performance of this same calibration set was significantly better when it was used for Raman

spectroscopy (RMSEP 0.79% w/w). And the prediction performance of the optimal NIR calibration set deteriorated significantly when it was used to develop a Raman calibration method (RMSEP 0.69% w/w vs 1.13% w/w).

Similar results were obtained during the calibration model development for Caffeine prediction as shown in Table 6-3. When the optimum calibration set of Raman (58 samples) was used to develop an NIR method, the prediction error was significantly high (RMSEP 0.48% w/w) compared to the prediction error (RMSEP 0.28% w/w) of optimum calibration set of NIR (37 samples). This optimum set of 37 samples was suitable specifically for NIR spectroscopy. When this optimum NIR set of 37 samples was used to develop a Raman calibration model, the prediction error was significantly high (RMSEP 0.6% w/w).

	14		
8	NIR Acetaminophen (35)	Raman Acetaminophen (71)	24
	NIR Caffeine (37)	Raman Caffeine (58)	
	10		

Table 6-4. Number of common samples between the optimum calibration sets of different techniques and different APIs

6.4 Conclusion

The optimum calibration sets for NIR and Raman spectroscopy were found to be different from each other. Only a few common samples were found between two APIs and two spectroscopic techniques. Table 6-4 summarizes the results. These results indicated that the optimum calibration structure depends on the interaction between spectral responses and physico-chemical properties of the tablets. An optimum calibration set for one formulation and/or technique can be sub-optimum for another formulation and/or technique. A formulation and technique specific calibration set should be designed to maximize calibration performance of a spectroscopic system for quantitative analysis of pharmaceutical tablets.

7 Chapter 7: Summary

The theory of experimental design is well established and NIR calibration have been widely used for wide range of applications, however there is a gap between the theory and practical implementation of experimental design during NIR calibration set development. Direct application of classical theory of experimental design is still missing in the problem of multivariate spectroscopic calibration [76, 194]. Bridging this gap can lead towards better understanding of the interaction between calibration structure and NIRS model performance. This understanding can be helpful to identify the necessary information for a calibration set and challenge the current philosophy of incorporating all possible variances into the calibration set. All possible variances may not always be necessary for desired model performance. Moreover, it requires a large calibration set to span all possible variances, which may not be an efficient approach in terms of time and available resources. Besides, it is almost impossible to anticipate and span all possible variances, especially at the early stage of product and process development. The current all-inclusive calibration strategy can be replaced by an efficient strategy for developing NIR calibration set using spectral space.

It was demonstrated that, the pure component spectral response of a formulation can be utilized to define the spectral space of that formulation. The use of pure component spectral information results in a formulation and technique specific spectral space. Such specificity leads towards the development of a formulation and technique specific calibration set. The current strategies for designing NIR calibration sets do not account for such specificity towards formulations and analytical technique. Regulatory guidelines and general rules of thumb are usually followed for all types of formulations resulting in similar concentration ranges, levels and number of samples

for all types of calibration sets. However, an optimal calibration design for one formulation can be sub-optimal for a different formulation. Different NIR responses produced by different formulations should be considered during the selection of calibration structure, range, size, concentration levels and variance information. It was also demonstrated that the spectral space can be used to identify the critical calibration samples. Information regarding the critical sample requirement helps to design and prepare a small and efficient calibration set. This strategy was defined as the spectral space design strategy for calibration set development during NIR quantitative analysis of tablets.

The spectral space design strategy was compared with the current calibration strategies of designing calibration sets. The current strategy is prone to have redundant information. It was demonstrated that, a traditionally designed calibration set contained redundant information whereas a spectrally designed calibration set identified and eliminated redundant information, thus requiring fewer calibration samples to provide similar calibration performance. Multiple comparative studies were conducted between commonly employed experimental design approaches to calibration development and the newly developed spectral space based technique. The comparisons were conducted on single API (Active Pharmaceutical Ingredient) and multiple API formulations to quantify drug tablet API using NIR spectroscopy. Partial least squares (PLS) models were developed from respective calibration designs. Model performance was comprehensively assessed based on the ability to predict API concentrations in independent prediction/validation sets. Similar prediction performance was achieved using the smaller calibration set designed in spectral space (11 for single API and 33 for multiple API tablets), compared to the traditionally designed large calibration sets (45 for single API and 225 for multiple API tablets). An improved prediction performance was observed for the spectrally

designed calibration sets (33 tablets for spectral design) compared to the traditionally designed calibration sets of equal sizes (33 tablets for D-Optimal, CCD and I-Optimal designs). It was demonstrated that a calibration set designed in spectral space provided an efficient means of developing spectroscopic multivariate calibration for tablet analysis. Such strategy provides opportunity to design formulation and technique specific calibration sets to optimize calibration capability.

The spectral design strategy was built on the spectral space of a formulation. The spectral space was defined by pure component spectral information. Initially the spectral space was used to identify only compositional design points that are critical to calibration performance. However, it is also critical for a calibration set to contain physico-chemical variation besides the compositional information to build a robust calibration model. The current strategy for incorporating physico-chemical variation into the calibration requires large calibration set. It was demonstrated that, spectral space can be used to develop an efficient strategy to incorporate physico-chemical information into the calibration set. The density of the tablet was identified as a critical physico-chemical information for the NIR calibration set. The pure component tablet spectra at different densities were utilized to define a latent variable spectral space containing both concentration and density information at different directions. A small set of critical samples was selected from the latent variable spectral space. These samples were designed to prepare a small calibration set containing both concentration and density information. The performance of this small calibration set was compared with the traditionally designed large calibration set during quantitative analysis of Acetaminophen and caffeine using NIR. The small calibration set provided similar prediction performance compared to the large calibration sets indicating the

efficiency of the spectral design strategy to incorporate physico-chemical information into the calibration set.

The spectral design strategy can be generalized to develop efficient calibration sets for other formulations. However, it is critical to have pure component spectral information to define the spectral space of that formulation. In this study, pure component spectral information was collected from pure component tablets. Pure component powder can also be used to collect pure component spectra for uncompressible/less compressible powders. The spectral design strategy can also be generalized to incorporate other types of physico-chemical information such as moisture content, particle size distribution etc. However, it is critical to consider the effects of physico-chemical information in the pure component spectra and tablet spectra during such analysis.

The spectral design strategy identifies the critical samples from spectral information whereas the current calibration design strategy assumes certain samples to be critical for the calibration set development from their compositional information. For instance, the full factorial design assumes the orthogonal compositional samples are most critical to calibration set, the D-optimal design assumes that the compositionally distant samples are most critical to calibration set. The advantage of having large sample set is that, it provides useful and critical information even when these assumptions are violated. Such information from a large calibration set can be used to develop a successful quantitative NIR method. The advantage of spectral calibration set is that, it provides similar performance with a smaller calibration set. It is very critical to consider that spectral design is expected to be 'efficient' not 'better / more accurate' compared to the current calibration approaches.

It is also critical to consider that the spectral design strategy utilizes a selection technique to identify critical calibration samples in the spectral space. Different selection techniques can result in slightly different calibration sets in the spectral space. In this study, multiple selection techniques (Kennard stone, orthogonal rotation, Euclidian distance) were used to identify critical samples and all of these sets gave similar prediction results due to similar spanning of spectral variation. Different spectral variation from different selection algorithm would have resulted differences in prediction performances. It is critical that the selection technique maximizes the variation in spectral space to find the calibration set with optimum performance. Such optimum performance can be obtained with the current calibration strategy in case the same samples maximize variation in the concentration space. Such case would require linear mapping between spectral and compositional spaces.

In this study, it was also demonstrated that the optimal calibration structure depends on the formulation and spectroscopic technique. An optimal calibration set for one formulation can be sub-optimal for another formulation. An optimal calibration set for one spectroscopic technique can be suboptimal for another spectroscopic technique. The current calibration strategies do not allow optimization of calibration set based on formulation variation and specific spectral response. The same calibration design is used for different formulations and different techniques. This strategy does not provide optimum performance of the calibration efforts. The optimum calibration structures between NIR and Raman spectroscopy were compared during the quantitative analysis of Acetaminophen and Caffeine. The optimum NIR calibration set was found to be different from the optimum Raman calibration set for both APIs. The optimum calibration set for Acetaminophen was found to be different from the optimum calibration set for

Caffeine for both spectroscopic techniques. The common calibration designs between different formulations and different techniques do not provide optimal calibration performance.

In the spectral design strategy, the use of pure component spectral information allows the calibration set to be formulation and technique specific. Such specificity would provide additional information to establish method suitability during method validation. However, it would not ease the validation steps or make it more difficult. The current validation steps require the calibration model to demonstrate accuracy, linearity, specificity, precision, range and robustness. The same requirements should be applied to the spectrally design calibration set. It is expected that, the spectrally design calibration set should pass the validation steps with fewer samples.

In this study, Partial Least Square (PLS) was used as the modeling technique. It was demonstrated that robustness of the PLS calibration model is critical for successful implementation of an NIR quantitative method in the pharmaceutical industry. Different sources of variations such as scale variations, physical variations and chemical variations can affect the PLS model performance during product life cycle. A robust PLS model is required to be developed at the outset of calibration to encounter such variations and provide desired predictive performance thus avoiding the needs for recurrent model updates. Selecting the optimum number of model components (loading vectors) is critical to build a robust model with the available dataset. Cross-validation error along with the amount of model variance captured by each model component is currently used to select the optimum number of model components. A new method was developed to select the optimum number of model components based on weight co-efficient of each model component. The new weight coefficient based method was found to be more

effective in selecting the optimum number of model components and improving model robustness compared to the current selection technique. Several datasets possessing different types of critical variations to NIR PLS model were used in this study to demonstrate the efficiency of the new weight coefficient based method in selecting the optimum number of model components.

PLS is the most frequently used modeling technique for calibration development. However, there are other modeling techniques such as Principal Component Regression, Multiple Linear Regression, Support Vector Machine, Artificial Neural Network etc. that can be used for calibration model development. The spectral design strategy was tested for PLS modeling technique considering its wide applicability. However, for other modeling techniques, the efficiency of the spectral design strategy is yet to be tested. This can be a direction of the future studies exploring spectral design strategy as an efficient alternative.

In future, the strategy demonstrated in this dissertation can also be used for other spectroscopic techniques including Raman and terahertz spectroscopy to develop an efficient calibration strategy. Pure spectrum of each component should be used during the design/selection of the calibration set. Comprehensive studies in this direction can lead towards formulation specific calibration development and ensure optimal performance of each calibration set. Comparative studies are suggested to distinguish the interactions between calibration structures and different spectroscopic techniques.

One of the challenges of the spectral design strategy is to calculate the compositional requirements of the spectrally designed samples. In this study, pure component and target tablet information was used to define the rotation matrix and calculate the compositional requirements

for spectrally designed samples. The rotation matrix was defined using target tablet since the calibration set was intended to center around target formulation. In other instances, when calibration sets are not intended to center around target formulation, the target tablet and pure component tablets might not be sufficient to calculate the rotation matrix. Additional sets of actual tablets might be necessary to calculate the rotation matrix and compositional requirements of the spectrally designed samples. Other challenge of the spectral design strategy is that, it utilizes optimization algorithm during the solution of compositional requirement calculations. All the optimization algorithms have the potentials to provide local minima solution. Multiple iterations are required to ensure global minima and expected solution.

There is no best calibration design that works for all purpose. A great deal depends on the method requirements [28]. The traditional practices for developing calibration sets are too generalized. Spectral design approach offers an efficient and formulation specific alternative strategy to the current practices of NIR calibration design for quantitative analysis of pharmaceutical tablets.

Appendices

Appendix A (Motivation for using scores as design factors for calibration)

The motivation for using scores as calibration design factors comes from the multivariate regression equation itself. The regression equation 2.1 can be transformed into a linear model equation with the concentrations as a response and PLS scores (projection on the orthogonal spectral space of the PLS model) as input variables shown in equation 2.9 for the same system with “m” samples, “n” spectral variables and “r” eigen vectors/(selected latent variables in PLS) using \mathbf{y} for the concentrations of samples, \mathbf{T} for the score matrix of “r” latent variables and \mathbf{u} for the regression vector between scores and concentrations

$$\hat{\mathbf{y}}_{m \times 1} = \mathbf{X}_{m \times n} * \mathbf{b}_{n \times 1}$$

$\hat{\mathbf{y}}_{m \times 1} = \mathbf{X}_{m \times n} * \mathbf{W}_{n \times r} * (\mathbf{T}'_{r \times m} * \mathbf{T}_{m \times r})^{-1} * (\mathbf{T}'_{r \times m} * \mathbf{y}_{m \times 1})$ [Here \hat{y} is the predicted concentration for new samples, y is reference concentration of the calibration samples. Please note that, the sample numbers do not have to be equal]

$$\hat{\mathbf{y}}_{m \times 1} = \mathbf{T}_{m \times r} * \mathbf{u}_{r \times 1} \tag{2.9}$$

$$\hat{y}_{i=1..m}^i = t_1^i * u_1 + t_2^i * u_2 + \dots + t_r^i * u_r \tag{2.10}$$

$$\hat{\mathbf{y}}_{m \times 1} = \mathbf{T}_{m \times r} * (\mathbf{T}'_{r \times m} * \mathbf{T}_{m \times r})^{-1} * (\mathbf{T}'_{r \times m} * \mathbf{y}_{m \times 1})$$

Thus, the classical idea of design of experiments is best embodied by orthogonally varying the PLS scores and measuring the concentrations as response according to equation 2.10. These experiments are then used to approximate a linear model between concentrations (y) and scores (t) with a limited number of factors. The challenge for designing orthogonal PLS scores is not

trivial. The designed scores need to be derived from a set of real pharmaceutical samples. The pure component spectra of the samples can be used to define the orthogonal spectral space of the PLS model assuming that, the primary model variance is dictated by the formulation variation after appropriate preprocessing technique has been applied.

Appendix B (Derivation of equation 2.7)

$$\mathbf{C}_{m \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} * \mathbf{P}_{k \times n}^{\text{Pure component spectra}} = \mathbf{X}_{m \times n}^{\text{Sample spectra}}$$

$$\mathbf{U}_{m \times m}^{\text{Left singular}} * \mathbf{S}_{m \times k}^{\text{Singular}} * \mathbf{V}_{k \times k}^{\text{Right singular}} * \mathbf{R}_{k \times k}^{\text{Rotation}} * \mathbf{P}_{k \times n}^{\text{Pure component spectra}} = \mathbf{X}_{m \times n}^{\text{Sample spectra}}$$

$$\mathbf{R}_{k \times k}^{\text{Rotation matrix}}$$

$$\begin{aligned} &= \mathbf{V}_{k \times k}^{\text{Right singular}'} * \left\{ \left(\mathbf{S}_{k \times m}^{\text{Singular}} \right)' * \left(\mathbf{S}_{m \times k}^{\text{Singular}} \right) * \right\}^{-1} * \left(\mathbf{S}_{k \times m}^{\text{Singular}} \right)' \\ &* \left(\mathbf{U}_{m \times m}^{\text{Left singular}} \right)' * \mathbf{X}_{m \times n}^{\text{Sample spectra}} * \left(\mathbf{P}_{n \times k}^{\text{Pure component}} \right)' \\ &* \left\{ \left(\mathbf{P}_{k \times n}^{\text{Pure component}} \right) * \left(\mathbf{P}_{n \times k}^{\text{Pure component}} \right)' \right\}^{-1} \end{aligned}$$

Since, the “S” is a diagonal matrix, a simpler solution can be performed by taking the diagonal values and making an inverse of the diagonal elements. However, it would give the exact same solution. The readers are encouraged to use either of the solutions.

Appendix C (Derivation of equation 2.8)

Using the assumption that the variation in pure component spectra dictates the primary spectral variation in the model and resembles the weight factors in the PLS,

$$\mathbf{X}_{u \times n}^{\text{Sample spectra}} * \mathbf{W}_{n \times k}^{\text{Pure component basis}} = \mathbf{T}_{u \times r}^{\text{Score space}}$$

$$\mathbf{C}_{u \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} * \mathbf{P}_{k \times n}^{\text{Pure component spectra}} * \mathbf{W}_{n \times k}^{\text{Pure component basis}} = \mathbf{T}_{u \times k}^{\text{Score space}}$$

$$\mathbf{C}_{u \times k}^{\text{Composition}} * \mathbf{R}_{k \times k}^{\text{Rotation}} = \mathbf{T}_{u \times k}^{\text{Score space}} * (\mathbf{P}_{k \times n}^{\text{Pure component spectra}} * \mathbf{W}_{n \times k}^{\text{Pure component basis}})^{-1}$$

References

1. Kirsch, J.D. and J.K. Drennen, *Near-infrared spectroscopy: applications in the analysis of tablets and solid pharmaceutical dosage forms*. Applied Spectroscopy Reviews, 1995. **30**(3): p. 139-174.
2. Roggo, Y., et al., *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*. J Pharm Biomed Anal, 2007. **44**(3): p. 683-700.
3. Luypaert, J., D.L. Massart, and Y. Vander Heyden, *Near-infrared spectroscopy applications in pharmaceutical analysis*. Talanta, 2007. **72**(3): p. 865-83.
4. Rasanen, E. and N. Sandler, *Near infrared spectroscopy in the development of solid dosage forms*. J Pharm Pharmacol, 2007. **59**(2): p. 147-59.
5. Jamrogiewicz, M., *Application of the near-infrared spectroscopy in the pharmaceutical technology*. J Pharm Biomed Anal, 2012. **66**: p. 1-10.
6. Kramer, R., *Chemometric techniques for quantitative analysis*. 1998: CRC Press.
7. Kalivas, J.H., *Basis sets for multivariate regression*. Analytica chimica acta, 2001. **428**(1): p. 31-40.
8. Blanco, M. and M. Alcalá, *Content uniformity and tablet hardness testing of intact pharmaceutical tablets by near infrared spectroscopy*. Analytica Chimica Acta, 2006. **557**(1-2): p. 353-359.
9. Bondi, R.W., et al., *Effect of Experimental Design on the Prediction Performance of Calibration Models Based on Near-Infrared Spectroscopy for Pharmaceutical Applications*. Applied spectroscopy, 2012. **66**(12): p. 1442-1453.
10. Ito, M., et al., *Development of a method for nondestructive NIR transmittance spectroscopic analysis of acetaminophen and caffeine anhydrate in intact bilayer tablets*. J Pharm Biomed Anal, 2010. **53**(3): p. 396-402.
11. Short, S.M., R.P. Cogdill, and C.A. Anderson, *Determination of figures of merit for near-infrared and Raman spectrometry by net analyte signal analysis for a 4-component solid dosage system*. Aaps Pharmscitech, 2007. **8**(4): p. 109-119.
12. Honigs, D., et al., *Unique-sample selection via near-infrared spectral subtraction*. Analytical Chemistry, 1985. **57**(12): p. 2299-2303.
13. Puchwein, G., *Selection of calibration samples for near-infrared spectrometry by factor analysis of spectra*. Analytical Chemistry, 1988. **60**(6): p. 569-573.
14. Chen, D., W. Cai, and X. Shao, *Representative subset selection in modified iterative predictor weighting (mIPW) — PLS models for parsimonious multivariate calibration*. Chemometrics and Intelligent Laboratory Systems, 2007. **87**(2): p. 312-318.
15. Càrdenas, V., M. Blanco, and M. Alcalá, *Strategies for Selecting the Calibration Set in Pharmaceutical Near Infrared Spectroscopy Analysis. A Comparative Study*. Journal of Pharmaceutical Innovation, 2014. **9**(4): p. 272-281.
16. Galvao, R.K., et al., *A method for calibration and validation subset partitioning*. Talanta, 2005. **67**(4): p. 736-40.
17. Storme-Paris, I., et al., *Near InfraRed Spectroscopy homogeneity evaluation of complex powder blends in a small-scale pharmaceutical preformulation process, a real-life application*. European Journal of Pharmaceutics and Biopharmaceutics, 2009. **72**(1): p. 189-198.
18. Bodson, C., et al., *Comparison of FT-NIR transmission and UV-vis spectrophotometry to follow the mixing kinetics and to assay low-dose tablets containing riboflavin*. Journal of pharmaceutical and biomedical analysis, 2006. **41**(3): p. 783-790.

19. Rosa, S.S., et al., *Development and validation of a method for active drug identification and content determination of ranitidine in pharmaceutical products using near-infrared reflectance spectroscopy: A parametric release approach*. *Talanta*, 2008. **75**(3): p. 725-733.
20. Trafford, A.D., et al., *A rapid quantitative assay of intact paracetamol tablets by reflectance near-infrared spectroscopy*. *Analyst*, 1999. **124**(2): p. 163-7.
21. Ritchie, G.E., et al., *Validation of a near-infrared transmission spectroscopic procedure: Part B: Application to alternate content uniformity and release assay methods for pharmaceutical solid dosage forms*. *Journal of pharmaceutical and biomedical analysis*, 2002. **29**(1): p. 159-171.
22. Naes, T., et al., *A user friendly guide to multivariate calibration and classification*. 2002: NIR publications.
23. Cogdill, R.P., et al., *Process analytical technology case study: Part II. Development and validation of quantitative near-infrared calibrations in support of a process analytical technology application for real-time release*. *AAPS PharmSciTech*, 2005. **6**(2): p. E273-E283.
24. Corti, P., et al., *Near infrared transmittance analysis for the assay of solid pharmaceutical dosage forms*. *Analyst*, 1999. **124**(5): p. 755-758.
25. Sparén, A., et al., *Matrix Effects in Quantitative Assessment of Pharmaceutical Tablets Using Transmission Raman and Near-Infrared (NIR) Spectroscopy*. *Applied spectroscopy*, 2015. **69**(5): p. 580-589.
26. Ito, M., et al., *Development of a method for nondestructive NIR transmittance spectroscopic analysis of acetaminophen and caffeine anhydrate in intact bilayer tablets*. *Journal of pharmaceutical and biomedical analysis*, 2010. **53**(3): p. 396-402.
27. Feng, Y.-C. and C.-Q. Hu, *Construction of universal quantitative models for determination of roxithromycin and erythromycin ethylsuccinate in tablets from different manufacturers using near infrared reflectance spectroscopy*. *Journal of pharmaceutical and biomedical analysis*, 2006. **41**(2): p. 373-384.
28. Araujo, P.W. and R.G. Brereton, *Experimental design III. Quantification*. *TrAC Trends in Analytical Chemistry*, 1996. **15**(3): p. 156-163.
29. Fernandez, J., et al., *Quantitative Analysis of Acetic Acid-Acetic Anhydride Mixtures in Near-Infrared Region. Statistical Study*. *Analytical Chemistry*, 1960. **32**(2): p. 158-162.
30. Sinsheimer, J.E. and A.M. Keuhnelian, *Near-infrared spectroscopy of amine salts*. *Journal of pharmaceutical sciences*, 1966. **55**(11): p. 1240-1244.
31. Oi, N. and E. Inaba, *Analyses of drugs and chemicals by infrared absorption spectroscopy. 8. Determination of allylisopropylacetate and phenacetin in pharmaceutical preparations by near infrared absorption spectroscopy*. *Yakugaku zasshi: Journal of the Pharmaceutical Society of Japan*, 1967. **87**(3): p. 213.
32. Sherken, S., *Rapid near-infrared spectrophotometric method for determination of meprobamate in meprobamate tablets*. *JOURNAL OF THE ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS*, 1968. **51**(3): p. 616-&.
33. Allen, L., *Quantitative determination of carisoprodol, phenacetin, and caffeine in tablets by near IR spectrometry and their identification by TLC*. *Journal of Pharmaceutical Sciences*, 1974. **63**(6): p. 912-916.
34. Sinsheimer, J.E. and N.M. Poswalk, *Pharmaceutical applications of the near infrared determination of water*. *Journal of Pharmaceutical Sciences*, 1968. **57**(11): p. 2007-2010.
35. Roggo, Y., C. Roeseler, and M. Ulmschneider, *Near infrared spectroscopy for qualitative comparison of pharmaceutical batches*. *Journal of Pharmaceutical and Biomedical Analysis*, 2004. **36**(4): p. 777-786.

36. Blanco, M., et al., *Analytical control of pharmaceutical production steps by near infrared reflectance spectroscopy*. *Analytica Chimica Acta*, 1999. **392**(2): p. 237-246.
37. Li, W., et al., *Applications of NIR in early stage formulation development: Part I. Semi-quantitative blend uniformity and content uniformity analyses by reflectance NIR without calibration models*. *International journal of pharmaceutics*, 2007. **340**(1): p. 97-103.
38. Xiang, D., et al., *Robust calibration design in the pharmaceutical quantitative measurements with near-infrared (NIR) spectroscopy: Avoiding the chemometric pitfalls*. *Journal of pharmaceutical sciences*, 2009. **98**(3): p. 1155-1166.
39. Næs, T. and T. Isaksson, *Selection of samples for calibration in near-infrared spectroscopy. Part I: General principles illustrated by example*. *Applied Spectroscopy*, 1989. **43**(2): p. 328-335.
40. Isaksson, T. and T. Næs, *Selection of samples for calibration in near-infrared spectroscopy. Part II: Selection based on spectral measurements*. *Applied Spectroscopy*, 1990. **44**(7): p. 1152-1158.
41. Blanco, M. and A. Peguero, *Analysis of pharmaceuticals by NIR spectroscopy without a reference method*. *TrAC Trends in Analytical Chemistry*, 2010. **29**(10): p. 1127-1136.
42. Sarraguça, M.C. and J.A. Lopes, *Quality control of pharmaceuticals with NIR: From lab to process line*. *Vibrational Spectroscopy*, 2009. **49**(2): p. 204-210.
43. Faber, N. and R. Rajko, *How to avoid over-fitting in multivariate calibration—The conventional validation approach and an alternative*. *Analytica Chimica Acta*, 2007. **595**(1): p. 98-106.
44. *International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, November 2005*, in *Validation of Analytical Procedures: Text and Methodology Q2(R1)*.
45. Martens, H., J.P. Nielsen, and S.B. Engelsen, *Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures*. *Analytical Chemistry*, 2003. **75**(3): p. 394-404.
46. Sharma, S., et al., *Efficient use of pure component and interferent spectra in multivariate calibration*. *Analytica chimica acta*, 2013. **778**: p. 15-23.
47. Burns, D.A. and E.W. Ciurczak, *Handbook of near-infrared analysis*. 2007: CRC press.
48. Lorber, A. and B.R. Kowalski, *The effect of interferences and calibration design on accuracy: Implications for sensor and sample selection*. *Journal of chemometrics*, 1988. **2**(1): p. 67-79.
49. Scheibelhofer, O., et al., *Designed Blending for Near Infrared Calibration*. *Journal of pharmaceutical sciences*, 2015.
50. Agelet, L.E. and C.R. Hurburgh Jr, *A tutorial on near infrared spectroscopy and its calibration*. *Critical Reviews in Analytical Chemistry*, 2010. **40**(4): p. 246-260.
51. Agency, E.M., *European Medicines Agency. 2014. Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations*. 2014.
52. Cogdill, R.P. and J.K. Drennen, *Risk-based quality by design (QbD): a Taguchi perspective on the assessment of product quality, and the quantitative linkage of drug product parameters and clinical performance*. *Journal of Pharmaceutical Innovation*, 2008. **3**(1): p. 23-29.
53. Roggo, Y., et al., *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*. *Journal of pharmaceutical and biomedical analysis*, 2007. **44**(3): p. 683-700.
54. *Development and Submission of Near Infrared Analytical Procedures Guidance for Industry*, March 2015, FDA, <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM440247.pdf>.

55. Sasakura, D., et al., *Strategic development of a multivariate calibration model for the uniformity testing of tablets by transmission NIR analysis*. Die Pharmazie-An International Journal of Pharmaceutical Sciences, 2015. **70**(5): p. 289-295.
56. Williams, P. and K. Norris, *Near-infrared technology in the agricultural and food industries*. 1987: American Association of Cereal Chemists, Inc.
57. *ASTM E1655-05(2012), Standard Practices for Infrared Multivariate Quantitative Analysis*. ASTM International, West Conshohocken, PA, 2012,.
58. Li, W. and G.D. Worosila, *Quantitation of active pharmaceutical ingredients and excipients in powder blends using designed multivariate calibration models by near-infrared spectroscopy*. International journal of pharmaceutics, 2005. **295**(1): p. 213-219.
59. Alvarenga, L., et al., *Tablet identification using near-infrared spectroscopy (NIRS) for pharmaceutical quality control*. Journal of pharmaceutical and biomedical analysis, 2008. **48**(1): p. 62-69.
60. Berntsson, O., et al., *Quantitative determination of content in binary powder mixtures using diffuse reflectance near infrared spectrometry and multivariate analysis*. Analytica chimica acta, 2000. **419**(1): p. 45-54.
61. Fearn, T., *Chemometrics: an enabling tool for NIR*. NIR news, 2005. **16**(7): p. 17-19.
62. Blanco, M., et al., *Quantitation of the active compound and major excipients in a pharmaceutical formulation by near infrared diffuse reflectance spectroscopy with fibre optical probe*, in *Analytica chimica acta*. 1996. p. 147-156.
63. Blanco, M., et al., *Near-infrared spectroscopy in the pharmaceutical industry*. ANALYST-LONDON-SOCIETY OF PUBLIC ANALYSTS THEN ROYAL SOCIETY OF CHEMISTRY-, 1998. **123**: p. 135R-150R.
64. Roger, J.-M., F. Chauchard, and V. Bellon-Maurel, *EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits*. Chemometrics and Intelligent Laboratory Systems, 2003. **66**(2): p. 191-204.
65. Segtnan, V.H., et al., *Low-cost approaches to robust temperature compensation in near-infrared calibration and prediction situations*. Applied spectroscopy, 2005. **59**(6): p. 816-825.
66. Peirs, A., et al., *Effect of biological variability on the robustness of NIR models for soluble solids content of apples*. Postharvest Biology and Technology, 2003. **28**(2): p. 269-280.
67. Kemps, B.J., et al., *The importance of choosing the right validation strategy in inverse modelling*. Journal of Near infrared spectroscopy, 2010. **18**(4): p. 231.
68. Trafford, A., R. Jee, and A. Moffat, *A rapid quantitative assay of intact paracetamol tablets by reflectance near-infrared spectroscopy*. Analyst, 1999. **124**(2): p. 163-167.
69. Gottfries, J., et al., *Vibrational spectrometry for the assessment of active substance in metoprolol tablets: a comparison between transmission and diffuse reflectance near-infrared spectrometry*. Journal of pharmaceutical and biomedical analysis, 1996. **14**(11): p. 1495-1503.
70. Merckle, P. and K.-A. Kovar, *Assay of effervescent tablets by near-infrared spectroscopy in transmittance and reflectance mode: acetylsalicylic acid in mono and combination formulations*. Journal of pharmaceutical and biomedical analysis, 1998. **17**(3): p. 365-374.
71. Dreassi, E., et al., *Application of near-infrared reflectance spectrometry to the analytical control of pharmaceuticals: ranitidine hydrochloride tablet production*. Analyst, 1996. **121**(2): p. 219-222.
72. Blanco, M., Coello, J, Iturriaga, H, Maspoch, S, C. de la Pezuela,, *Strategies for constructing the calibration set in the determination of active principles in pharmaceuticals by near infrared diffuse reflectance spectrometry*. Analyst, 1997. **122**(8): p. 761-765.

73. Blanco, M., et al., *Influence of the procedure used to prepare the calibration sample set on the performance of near infrared spectroscopy in quantitative pharmaceutical analyses*. Analyst, 2001. **126**(7): p. 1129-1134.
74. Moffat, A.C., et al., *Meeting the International Conference on Harmonisation's Guidelines on Validation of Analytical Procedures: Quantification as exemplified by a near-infrared reflectance assay of paracetamol in intact tablets*The opinions expressed in the following article are entirely those of the authors and do not necessarily represent the views of either The Royal Society of Chemistry or the Editor of The Analyst. Analyst, 2000. **125**(7): p. 1341-1351.
75. Blanco, M., M. Romero, and M. Alcalá, *Strategies for constructing the calibration set for a near infrared spectroscopic quantitation method*. Talanta, 2004. **64**(3): p. 597-602.
76. Vandeginste, B., et al., *Handbook of chemometrics and qualimetrics B*. 1998, Elsevier, Amsterdam.
77. Ryan, T.P. and J. Morgan, *Modern experimental design*. Journal of Statistical Theory and Practice, 2007. **1**(3-4): p. 501-506.
78. Kasemsumran, S., et al., *Near-infrared spectroscopic determination of human serum albumin, γ -globulin, and glucose in a control serum solution with searching combination moving window partial least squares*. Analytica Chimica Acta, 2004. **512**(2): p. 223-230.
79. Kasemsumran, S., et al., *Simultaneous determination of human serum albumin, [gamma]-globulin, and glucose in a phosphate buffer solution by near-infrared spectroscopy with moving window partial least-squares regression*. Analyst, 2003. **128**(12): p. 1471-1477.
80. Sondermann, N. and K.-A. Kovar, *Screening experiments of ecstasy street samples using near infrared spectroscopy*. Forensic science international, 1999. **106**(3): p. 147-156.
81. Igne, B., et al., *Modeling strategies for pharmaceutical blend monitoring and end-point determination by near-infrared spectroscopy*. International journal of pharmaceutics, 2014. **473**(1): p. 219-231.
82. Igne, B., et al., *Online monitoring of pharmaceutical materials using multiple NIR sensors—part ii: blend end-point determination*. Journal of Pharmaceutical Innovation, 2013. **8**(1): p. 45-55.
83. Mantanus, J., et al., *Moisture content determination of pharmaceutical pellets by near infrared spectroscopy: method development and validation*. Analytica chimica acta, 2009. **642**(1): p. 186-192.
84. Kamran, F., et al., *Transmission Near-Infrared (NIR) and Photon Time-of-Flight (PTOF) Spectroscopy in a Comparative Analysis of Pharmaceuticals*. Applied spectroscopy, 2015. **69**(3): p. 389-397.
85. Abrahamsson, C., et al., *Time-resolved NIR spectroscopy for quantitative analysis of intact pharmaceutical tablets*. Analytical chemistry, 2005. **77**(4): p. 1055-1059.
86. Costa, F.S., et al., *Multivariate Control Charts for Simultaneous Quality Monitoring of Isoniazid and Rifampicin in a Pharmaceutical Formulation Using a Portable Near Infrared Spectrometer*. Journal of the Brazilian Chemical Society, 2015. **26**(1): p. 64-73.
87. Fedorov, V.V., *Theory of optimal experiments*. 1972: Elsevier.
88. Karande, A., C. Liew, and P. Heng, *Calibration sampling paradox in near infrared spectroscopy: A case study of multi-component powder blend*. International journal of pharmaceutics, 2010. **395**(1): p. 91-97.
89. Sekulic, S.S., et al., *Automated system for the on-line monitoring of powder blending processes using near-infrared spectroscopy: Part II. Qualitative approaches to blend evaluation*. Journal of pharmaceutical and biomedical analysis, 1998. **17**(8): p. 1285-1309.

90. Tomuță, I., et al., *Quantification of meloxicam and excipients on intact tablets by near infrared spectrometry and chemometry*. 2010, Farmacia.
91. Blanco, M. and A. Peguero, *Influence of physical factors on the accuracy of calibration models for NIR spectroscopy*. Journal of pharmaceutical and biomedical analysis, 2010. **52**(1): p. 59-65.
92. Sarraguça, M.C., et al., *Determination of flow properties of pharmaceutical powders by near infrared spectroscopy*. Journal of pharmaceutical and biomedical analysis, 2010. **52**(4): p. 484-492.
93. Sarraguça, M.C., et al., *Comparison of different chemometric and analytical methods for the prediction of particle size distribution in pharmaceutical powders*. Analytical and bioanalytical chemistry, 2011. **399**(6): p. 2137-2147.
94. Porfire, A., et al., *Simultaneous quantification of simvastatin and excipients in liposomes using near infrared spectroscopy and chemometry*. Journal of pharmaceutical and biomedical analysis, 2015. **107**: p. 40-49.
95. Melucci, D., et al., *Rapid In Situ Repeatable Analysis of Drugs in Powder Form Using Reflectance Near-Infrared Spectroscopy and Multivariate Calibration*. Journal of forensic sciences, 2012. **57**(1): p. 86-92.
96. Alcalà, M., et al., *Qualitative and quantitative pharmaceutical analysis with a novel handheld miniature near-infrared spectrometer*. Journal of Near Infrared Spectroscopy, 2013. **21**(6): p. 445-457.
97. Morgan, J., *Optimal incomplete block designs*. Journal of the American statistical association, 2007. **102**(478): p. 655-663.
98. Patel, A.D., P.E. Luner, and M.S. Kemper, *Quantitative analysis of polymorphs in binary and multi-component powder mixtures by near-infrared reflectance spectroscopy*. International journal of pharmaceuticals, 2000. **206**(1): p. 63-74.
99. Otsuka, M. and Y. Fukui, *Determination of carbamazepine polymorphic contents in double-layered tablets using transmittance-and reflectance-near-infrared spectroscopy involving chemometrics*. Drug development and industrial pharmacy, 2010. **36**(12): p. 1404-1412.
100. Otsuka, M. and Y. Fukui, *Determination of carbamazepine polymorphic contents in double-layered tablets using transmittance- and reflectance-near-infrared spectroscopy involving chemometrics*. Drug Dev Ind Pharm, 2010. **36**(12): p. 1404-12.
101. Fukui, Y. and M. Otsuka, *Determination of the crystallinity of cephalexin in pharmaceutical formulations by chemometrical near-infrared spectroscopy*. Drug development and industrial pharmacy, 2010. **36**(1): p. 72-80.
102. Thosar, S.S., et al., *A comparison of reflectance and transmittance near-infrared spectroscopic techniques in determining drug content in intact tablets*. Pharmaceutical development and technology, 2001. **6**(1): p. 19-29.
103. Tomuta, I., et al., *High-throughput NIR-chemometric method for meloxicam assay from powder blends for tableting*. Scientia pharmaceutica, 2011. **79**(4): p. 885.
104. Chalus, P., et al., *Near-infrared determination of active substance content in intact low-dosage tablets*. Talanta, 2005. **66**(5): p. 1294-1302.
105. Doddridge, G.D. and Z. Shi, *Multivariate figures of merit (FOM) investigation on the effect of instrument parameters on a Fourier transform-near infrared spectroscopy (FT-NIRS) based content uniformity method on core tablets*. Journal of pharmaceutical and biomedical analysis, 2015. **102**: p. 535-543.
106. Otsuka, E., et al., *Nondestructive prediction of the drug content of an aspirin suppository by near-infrared spectroscopy*. Drug development and industrial pharmacy, 2010. **36**(7): p. 839-844.

107. Berntsson, O., et al., *Quantitative in-line monitoring of powder blending by near infrared reflection spectroscopy*. Powder Technology, 2002. **123**(2): p. 185-193.
108. Martínez, L., et al., *Use of near-infrared spectroscopy to quantify drug content on a continuous blending process: Influence of mass flow and rotation speed variations*. European Journal of Pharmaceutics and Biopharmaceutics, 2013. **84**(3): p. 606-615.
109. Eriksson, L., E. Johansson, and C. Wikström, *Mixture design—design generation, PLS analysis, and model usage*. Chemometrics and Intelligent Laboratory Systems, 1998. **43**(1): p. 1-24.
110. El-Gindy, A., et al., *Reflectance Near-Infrared Spectroscopic Method with a Chemometric Technique Using Partial Least Squares Multivariate Calibration for Simultaneous Determination of Chondroitin, Glucosamine, and Ascorbic Acid*. Journal of AOAC International, 2012. **95**(3): p. 724-732.
111. Shi, Z. and C.A. Anderson, *Scattering orthogonalization of near-infrared spectra for analysis of pharmaceutical tablets*. Analytical chemistry, 2009. **81**(4): p. 1389-1396.
112. Short, S.M., et al., *A near-infrared spectroscopic investigation of relative density and crushing strength in four-component compacts*. Journal of pharmaceutical sciences, 2009. **98**(3): p. 1095-1109.
113. Igne, B., C.A. Anderson, and J.K. Drennen lii, *Radial tensile strength prediction of relaxing and relaxed compacts by near-infrared chemical imaging*. International Journal of Pharmaceutics, 2011. **418**(2): p. 297-303.
114. Ozdemir, D. and B. Ozturk, *Genetic multivariate calibration methods for near infrared (NIR) spectroscopic determination of complex mixtures*. Turkish Journal of Chemistry, 2004. **28**(4): p. 497-514.
115. De Maesschalck, R., et al., *The development of calibration models for spectroscopic data using principal component regression*. Internet Journal of Chemistry, 1999. **2**(19): p. 1.
116. Alexandrino, G.L. and R.J. Poppi, *NIR imaging spectroscopy for quantification of constituents in polymers thin films loaded with paracetamol*. Analytica chimica acta, 2013. **765**: p. 37-44.
117. Sulub, Y., et al., *Real-time on-line blend uniformity monitoring using near-infrared reflectance spectrometry: a noninvasive off-line calibration approach*. Journal of pharmaceutical and biomedical analysis, 2009. **49**(1): p. 48-54.
118. Lin, Y., et al., *Development of a NIR-based blend uniformity method for a drug product containing multiple structurally similar actives by using the quality by design principles*. International journal of pharmaceutics, 2015. **488**(1): p. 120-126.
119. Shi, Z., et al., *Optical coefficient-based multivariate calibration on near-infrared spectroscopy*. Journal of Chemometrics, 2010. **24**(5): p. 288-299.
120. Shi, Z., et al., *Process characterization of powder blending by near-infrared spectroscopy: blend end-points and beyond*. Journal of pharmaceutical and biomedical analysis, 2008. **47**(4): p. 738-745.
121. Wu, H., et al., *Quality-by-Design (QbD): An integrated multivariate approach for the component quantification in powder blends*. International Journal of Pharmaceutics, 2009. **372**(1): p. 39-48.
122. Liew, C., A. Karande, and P. Heng, *In-line quantification of drug and excipients in cohesive powder blends by near infrared spectroscopy*. International journal of pharmaceutics, 2010. **386**(1): p. 138-148.
123. Karande, A.D., P.W.S. Heng, and C.V. Liew, *In-line quantification of micronized drug and excipients in tablets by near infrared (NIR) spectroscopy: Real time monitoring of tableting process*. International journal of pharmaceutics, 2010. **396**(1): p. 63-74.

124. Alcalà, M., et al., *Deconvolution of chemical and physical information from intact tablets NIR spectra: Two-and three-way multivariate calibration strategies for drug quantitation*. Journal of pharmaceutical sciences, 2009. **98**(8): p. 2747-2758.
125. Osorio, J.G., et al., *Characterization of pharmaceutical powder blends using in situ near-infrared chemical imaging*. Chemical Engineering Science, 2014. **108**: p. 244-257.
126. Zhang, W., et al., *Near-infrared reflectance spectroscopy (NIRS) for rapid determination of ginsenoside Rg1 and Re in Chinese patent medicine Naosaitong pill*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2015. **139**: p. 184-188.
127. Meza, C.P., M.A. Santos, and R.J. Románach, *Quantitation of drug content in a low dosage formulation by transmission near infrared spectroscopy*. Aaps Pharmscitech, 2006. **7**(1): p. E206-E214.
128. Marques Junior, J.M., et al., *Determination of Propranolol Hydrochloride in Pharmaceutical Preparations Using Near Infrared Spectrometry with Fiber Optic Probe and Multivariate Calibration Methods*. Journal of analytical methods in chemistry, 2015. **2015**.
129. Ferreira, M.H., J.W. Braga, and M.M. Sena, *Development and validation of a chemometric method for direct determination of hydrochlorothiazide in pharmaceutical samples by diffuse reflectance near infrared spectroscopy*. Microchemical Journal, 2013. **109**: p. 158-164.
130. Li, J., et al., *Simultaneous determination of the impurity and radial tensile strength of reduced glutathione tablets by a high selective NIR-PLS method*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2014. **125**: p. 278-284.
131. Silva, M.A., et al., *Development and analytical validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy*. Talanta, 2012. **89**: p. 342-351.
132. Kennard, R.W. and L.A. Stone, *Computer aided design of experiments*. Technometrics, 1969. **11**(1): p. 137-148.
133. Li, P., et al., *Rapid and nondestructive analysis of pharmaceutical products using near-infrared diffuse reflectance spectroscopy*. Journal of pharmaceutical and biomedical analysis, 2012. **70**: p. 288-294.
134. Bouveresse, E., et al., *Standardization of near-infrared spectrometric instruments*. Analytical Chemistry, 1996. **68**(6): p. 982-990.
135. Blanco, M. and M. Alcalá, *Content uniformity and tablet hardness testing of intact pharmaceutical tablets by near infrared spectroscopy: a contribution to process analytical technologies*. Analytica chimica acta, 2006. **557**(1): p. 353-359.
136. Blanco, M., et al., *Identification and quantitation assays for intact tablets of two related pharmaceutical preparations by reflectance near-infrared spectroscopy: validation of the procedure*. Journal of pharmaceutical and biomedical analysis, 2000. **22**(1): p. 139-148.
137. Shi, Z., et al., *A novel sample selection strategy by near-infrared spectroscopy-based high throughput tablet tester for content uniformity in early-phase pharmaceutical product development*. Journal of pharmaceutical sciences, 2012. **101**(7): p. 2502-2511.
138. De Bleye, C., et al., *Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications*. Journal of pharmaceutical and biomedical analysis, 2012. **69**: p. 125-132.
139. Moes, J.J., et al., *Application of process analytical technology in tablet process development using NIR spectroscopy: Blend uniformity, content uniformity and coating thickness measurements*. International journal of pharmaceutics, 2008. **357**(1): p. 108-118.

140. Blanco, M. and A. Villar, *Development and validation of a method for the polymorphic analysis of pharmaceutical preparations using near infrared spectroscopy*. Journal of pharmaceutical sciences, 2003. **92**(4): p. 823-830.
141. Laasonen, M., et al., *Development and validation of a near-infrared method for the quantitation of caffeine in intact single tablets*. Analytical chemistry, 2003. **75**(4): p. 754-760.
142. Sarraguça, M.C. and J.A. Lopes, *The use of net analyte signal (NAS) in near infrared spectroscopy pharmaceutical applications: Interpretability and figures of merit*. Analytica chimica acta, 2009. **642**(1): p. 179-185.
143. Sarraguça, M.C., S.O. Soares, and J.A. Lopes, *A near-infrared spectroscopy method to determine aminoglycosides in pharmaceutical formulations*. Vibrational Spectroscopy, 2011. **56**(2): p. 184-192.
144. Eustaquio, A., et al., *Determination of paracetamol in intact tablets by use of near infrared transmittance spectroscopy*. Analytica chimica acta, 1999. **383**(3): p. 283-290.
145. Mainali, D., et al., *Development of a comprehensive near infrared spectroscopy calibration model for rapid measurements of moisture content in multiple pharmaceutical products*. Journal of pharmaceutical and biomedical analysis, 2014. **95**: p. 169-175.
146. Boiret, M., L. Meunier, and Y.-M. Ginot, *Tablet potency of Tianeptine in coated tablets by near infrared spectroscopy: Model optimisation, calibration transfer and confidence intervals*. Journal of pharmaceutical and biomedical analysis, 2011. **54**(3): p. 510-516.
147. Nikolich, K., C. Sergides, and A. Pittas, *The application of Near Infrared Reflectance Spectroscopy (NIRS) for the quantitative analysis of hydrocortisone in primary materials*. JOURNAL-SERBIAN CHEMICAL SOCIETY, 2001. **66**(3): p. 189-198.
148. Blanco, M., M. Bautista, and M. Alcalá, *Preparing calibration sets for use in pharmaceutical analysis by NIR spectroscopy*. Journal of pharmaceutical sciences, 2008. **97**(3): p. 1236-1245.
149. Cárdenas, V., et al., *Strategy for design NIR calibration sets based on process spectrum and model space: An innovative approach for process analytical technology*. Journal of pharmaceutical and biomedical analysis, 2015. **114**: p. 28-33.
150. Swierenga, H., et al., *Strategy for constructing robust multivariate calibration models*. Chemometrics and Intelligent Laboratory Systems, 1999. **49**(1): p. 1-17.
151. Arnold, M.A., J.J. Burmeister, and G.W. Small, *Phantom glucose calibration models from simulated noninvasive human near-infrared spectra*. Analytical chemistry, 1998. **70**(9): p. 1773-1781.
152. Arnold, M.A., et al., *Pure component selectivity analysis of multivariate calibration models from near-infrared spectra*. Analytical chemistry, 2004. **76**(9): p. 2583-2590.
153. Wold, S., et al., *Some recent developments in PLS modeling*. Chemometrics and intelligent laboratory systems, 2001. **58**(2): p. 131-150.
154. Small, G.W., *Chemometrics and near-infrared spectroscopy: avoiding the pitfalls*. TrAC Trends in Analytical Chemistry, 2006. **25**(11): p. 1057-1066.
155. Li, W., et al., *Mass-balanced blend uniformity analysis of pharmaceutical powders by at-line near-infrared spectroscopy with a fiber-optic probe*. International journal of pharmaceutics, 2006. **326**(1): p. 182-185.
156. Alcalá, M., et al., *Analysis of low content drug tablets by transmission near infrared spectroscopy: selection of calibration ranges according to multivariate detection and quantitation limits of PLS models*. Journal of pharmaceutical sciences, 2008. **97**(12): p. 5318-5327.

157. Kessler, W., D. Oelkrug, and R. Kessler, *Using scattering and absorption spectra as MCR-hard model constraints for diffuse reflectance measurements of tablets*. *Analytica chimica acta*, 2009. **642**(1): p. 127-134.
158. Nakagawa, H., et al., *Verification of model development technique for NIR-based real-time monitoring of ingredient concentration during blending*. *Int J Pharm*, 2014. **471**(1-2): p. 264-75.
159. Dyrby, M., et al., *Chemometric quantitation of the active substance (containing C≡N) in a pharmaceutical tablet using near-infrared (NIR) transmittance and NIR FT-Raman spectra*. *Applied spectroscopy*, 2002. **56**(5): p. 579-585.
160. Ito, M., et al., *Development of a method for the determination of caffeine anhydrate in various designed intact tables by near-infrared spectroscopy: A comparison between reflectance and transmittance technique*. *Journal of pharmaceutical and biomedical analysis*, 2008. **47**(4): p. 819-827.
161. Pan, D., et al., *Low level drug product API form analysis—Avalide tablet NIR quantitative method development and robustness challenges*. *Journal of pharmaceutical and biomedical analysis*, 2014. **89**: p. 268-275.
162. Ebube, N., et al., *Application of Near-Infrared Spectroscopy for Nondestructive Analysis of Avicel® Powders and Tablets*. *Pharmaceutical Development & Technology*, 1999. **4**(1): p. 19.
163. Pieters, S., et al., *Robust calibrations on reduced sample sets for API content prediction in tablets: Definition of a cost-effective NIR model development strategy*. *Analytica chimica acta*, 2013. **761**: p. 62-70.
164. Igne, B., et al., *Effects and Detection of Raw Material Variability on the Performance of Near-Infrared Calibration Models for Pharmaceutical Products*. *Journal of pharmaceutical sciences*, 2014. **103**(2): p. 545-556.
165. Van den Kerkhof, T., et al., *Augmentation of near infrared diffuse reflectance and transmittance spectral data for the development of robust PLSBC models for classifying double blind clinical trial tablets*. *Journal of pharmaceutical and biomedical analysis*, 2006. **42**(4): p. 517-522.
166. Massart, D.L., et al., *Handbook of chemometrics and qualimetrics: Part A*. 1997: Elsevier Science Inc.
167. Roggo, Y., et al., *Characterizing process effects on pharmaceutical solid forms using near-infrared spectroscopy and infrared imaging*. *European Journal of Pharmaceutics and Biopharmaceutics*, 2005. **61**(1): p. 100-110.
168. Riley, M.R., M.A. Arnold, and D.W. Murhammer, *Effect of sample complexity on quantification of analytes in aqueous samples by near-infrared spectroscopy*. *Applied Spectroscopy*, 2000. **54**(2): p. 255-261.
169. El-Hagrasy, A.S., et al., *Near-infrared spectroscopy and imaging for the monitoring of powder blend homogeneity*. *Journal of Pharmaceutical Sciences*, 2001. **90**(9): p. 1298-1307.
170. El-Hagrasy, A.S., F. D'Amico, and J.K. Drennen, *A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending. Part I: D-optimal design for characterization of powder mixing and preliminary spectral data evaluation*. *Journal of pharmaceutical sciences*, 2006. **95**(2): p. 392-406.
171. El-Hagrasy, A.S., M. Delgado-Lopez, and J.K. Drennen, *A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending: Part II: Qualitative near-infrared models for prediction of blend homogeneity*. *Journal of pharmaceutical sciences*, 2006. **95**(2): p. 407-421.
172. Wold, S., K. Esbensen, and P. Geladi, *Principal component analysis*. *Chemometrics and intelligent laboratory systems*, 1987. **2**(1-3): p. 37-52.

173. Trask, A.V., W.S. Motherwell, and W. Jones, *Pharmaceutical cocrystallization: engineering a remedy for caffeine hydration*. *Crystal Growth & Design*, 2005. **5**(3): p. 1013-1021.
174. Igne, B., et al., *Robustness considerations and effects of moisture variations on near infrared method performance for solid dosage form assay*. *Journal of Near Infrared Spectroscopy*, 2014. **22**(3): p. 179-188.
175. Alam, M.A., J. Drennen III, and C. Anderson, *Designing a calibration set in spectral space for efficient development of an NIR method for tablet analysis*. *Journal of pharmaceutical and biomedical analysis*, 2017. **145**: p. 230-239.
176. Fearn, T., *Comparing standard deviations*. *NIR news*, 1996. **7**(5): p. 5-6.
177. Norris, K. and P. Williams, *Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. I. Influence of particle size*. *Cereal Chemistry*, 1984.
178. Kirsch, J.D. and J.K. Drennen, *Nondestructive tablet hardness testing by near-infrared spectroscopy: a new and robust spectral best-fit algorithm*. *Journal of pharmaceutical and biomedical analysis*, 1999. **19**(3): p. 351-362.
179. Morisseau, K.M., *THE EFFECT OF COMPRESSION FORCE ON THE NEAR-INFRARED SPECTRA OF TABLET DOSAGE FORMS*. 1996.
180. Igne, B., et al., *Robustness considerations and effects of moisture variations on near infrared method performance for solid dosage form assay*. *J. Near Infrared Spectrosc*, 2014. **22**: p. 179.
181. Preys, S., J.M. Roger, and J.C. Boulet, *Robust calibration using orthogonal projection and experimental design. Application to the correction of the light scattering effect on turbid NIR spectra*. *Chemometrics and Intelligent Laboratory Systems*, 2008. **91**(1): p. 28-33.
182. Rinnan, Å., F. van den Berg, and S.B. Engelsen, *Review of the most common pre-processing techniques for near-infrared spectra*. *TrAC Trends in Analytical Chemistry*, 2009. **28**(10): p. 1201-1222.
183. Næs, T., et al., *A user-friendly guide to multivariate calibration and classification*. Vol. 6. 2002: NIR publications Chichester.
184. Cogdill, R.P., C.A. Anderson, and J.K. Drennen, *Process analytical technology case study, part III: calibration monitoring and transfer*. *AAPS PharmSciTech*, 2005. **6**(2): p. E284-E297.
185. *Draft guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations*. European Medicines Agency. 2009, (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/02/WC500122769.pdf).
186. Osten, D.W., *Selection of optimal regression models via cross-validation*. *Journal of Chemometrics*, 1988. **2**(1): p. 39-48.
187. Mevik, B.-H. and R. Wehrens, *The pls package: principal component and partial least squares regression in R*. *Journal of Statistical Software*, 2007. **18**(2): p. 1-24.
188. Lorber, A., L.E. Wangen, and B.R. Kowalski, *A theoretical foundation for the PLS algorithm*. *Journal of Chemometrics*, 1987. **1**(1): p. 19-31.
189. Geladi, P. and B.R. Kowalski, *Partial least-squares regression: a tutorial*. *Analytica chimica acta*, 1986. **185**: p. 1-17.
190. Wise, B.M., *Properties of Partial Least Squares (PLS) Regression, and Differences between Algorithms*. 2004.
191. Wold, S., M. Sjöström, and L. Eriksson, *PLS-regression: a basic tool of chemometrics*. *Chemometrics and intelligent laboratory systems*, 2001. **58**(2): p. 109-130.
192. http://wiki.eigenvector.com/index.php?title=Using_Cross-Validation.

193. Otsuka, M., et al., *Chemoinformetrical evaluation of dissolution property of indomethacin tablets by near-infrared spectroscopy*. Journal of pharmaceutical sciences, 2007. **96**(4): p. 788-801.
194. Martens, H. and T. Naes, *Multivariate calibration*. 1992: John Wiley & Sons.

Copyright clearance for [175]

2/22/2018

Rightslink® by Copyright Clearance Center



RightsLink®

Home

Create Account

Help



Title: Designing a calibration set in spectral space for efficient development of an NIR method for tablet analysis

Author: Md Anik Alam, James Drennen, Carl Anderson

Publication: Journal of Pharmaceutical and Biomedical Analysis

Publisher: Elsevier

Date: 25 October 2017

© 2017 Elsevier B.V. All rights reserved.

LOGIN

If you're a [copyright.com](#) user, you can login to RightsLink using your [copyright.com](#) credentials. Already a [RightsLink](#) user or want to [learn more?](#)

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at customer@copyright.com