ATR – FTIR Chemometrics for Biological Samples

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in

Nanoscience

at Massey University, Manawatū,

New Zealand.

Josiah David Cleland

2018

Abstract:

This project has used the analytical infrared reflectance technique of *Attenuated Total Reflectance Fourier Transform Infrared* (ATR-FTIR) spectroscopy, for the prediction of chemical components in a range of biological samples. Data collection was carried out on 40 hyperaccumulator samples, 56 chicken feed samples, 54 lamb faecal samples and 188 forage feed samples. Predictions were made using several different regression methods including: Ridge, Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, Principal Components (PCR) and Partial Least Squares (PLS). The best methods were identified as LASSO, Elastic Net and PLS. Several spectral data pre-treatments were explored, the best of which combined Standard Normal Variant scaling (SNV) with a first-order Savitzky – Golay (SG) spectral derivative and smoothing filter. Several of the resulting models illustrated high quality predictions ($R^2 > 0.8$, Relative Performance Deviation (RPD) ≥ 2). The SNV and SG pre-treatment almost completely reduces the contribution of strong water-based signals to the regression model, allowing the possibility of *in situ* prediction of forage feed composition with minimal sample preparation. ATR-FTIR spectrometers are available in a hand-held form, and the results of this research suggest that in situ forage quality analysis could be performed using mid – infrared (MIR) reflectance spectroscopy.

Preface:

A portion of this research thesis has been submitted to the Journal of Animal Feed Science and Technology for publishing. The article is titled *"Mid-Infrared Reflectance Spectroscopy as a tool for forage feed composition prediction"* and was authored by Josiah D. Cleland, Ellie Johnson, Patrick C. H. Morel, Paul R. Kenyon, Mark R. Waterland.

Acknowledgements:

First and foremost, I would like to acknowledge my wonderful wife, Jo and daughter Ayla, I wouldn't be where I am today without their love and support. I would like to also thank my parents, Darryl and Trish, for all their help and support throughout my studies.

A special thanks is in order for my supervisor, Associate Professor Mark Waterland, for his guidance and expert advice. Mark's enthusiasm for the theory involved in this project has led to many productive discussions, and contributed greatly to the successes of this research.

The project included a variety of sample types, an additional thanks is extended to the people who kindly provided them, namely Dr Fifi Zaefarian, Dr Felicity Jackson, Miss Antoinette Danso and Associate Professor Chris Anderson. I am also very thankful to Professor Patrick Morel, Professor Paul Kenyon and again Associate Professor Chris Anderson for the specific knowledge they were able to input into the project regarding the forage feed samples and hyperaccumulator samples, respectively.

I would like to also acknowledge Massey University for selecting me as a recipient of the *Massey University Masterate Scholarship*, the funds helped greatly during my studies.

Abbreviations:

PT	Pre – Treatment
SG	Savitsky – Golay filter
SNV	Standard normal variate scaling
MSC	Multiplicative Scatter Correction
DWT	Discrete Wavelet Transform
AsLS	Asymmetric Least Squares
PLS	Partial Least Squares
PCR	Principal Component Regression
PCA	Principal Component Analysis
RR	Ridge Regression
LASSO	Least Absolute Shrinkage and Selection Operator
EN	Elastic Net
NDF	Neutral Detergent Fibre
ADF	Acid Detergent Fibre
GE	Gross – Energy
ME	Metabolisable – Energy
OM	Organic Matter
DOMD	In vitro Digestible Organic Matter
PC	Principal Component
ATR FTIR	Attenuated Total Reflectance Fourier Transform Infrared
MIR	Mid – Infrared
NIR	Near – Infrared
FIR	Far – Infrared
MPAES	Microwave Plasma Atomic Emission Spectrometer
RMSEP	Root Mean Squared Error of Prediction
RMSECV	Root Mean Squared Error of Cross – Validation
SEP	Standard Error of Prediction
SECV	Standard Error of Cross – Validation
RPD	Relative Performance Deviation

Table of Contents: