Near-infrared monitoring of roller compacted ribbon density: Investigating sources of variation contributing to noisy spectral data

Mary Ellen Crowley, Avril Hegarty, Michael A.P. McAuliffe, Graham E. O'Mahony, Luke Kiernan, Kevin Hayes, Abina M. Crean

PII: S0928-0987(17)30100-8
DOI: doi: 10.1016/j.ejps.2017.02.024
Reference: PHASCI 3922

To appear in: *European Journal of Pharmaceutical Sciences*

Received date: 6 August 2016
Revised date: 18 January 2017
Accepted date: 14 February 2017

Please cite this article as: Mary Ellen Crowley, Avril Hegarty, Michael A.P. McAuliffe, Graham E. O'Mahony, Luke Kiernan, Kevin Hayes, Abina M. Crean , Near-infrared monitoring of roller compacted ribbon density: Investigating sources of variation contributing to noisy spectral data. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Phasci(2017), doi: 10.1016/j.ejps.2017.02.024

Near-Infrared monitoring of roller compacted ribbon density: Investigating sources of variation contributing to noisy spectral data

Authors: Mary Ellen Crowley[1 *], Avril Hegarty[2*], Michael A. P. McAuliffe[3], Graham E. O' Mahony[1], Luke Kiernan[4], Kevin Hayes[2], Abina M. Crean[1,5]

[*] Joint first authors

[1] Synthesis and Solid State Pharmaceutical Centre, School of Pharmacy, University College Cork, Ireland

[2] MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland

[3] Centre for Advanced Photonics and Process Analysis, Applied Physics & Instrumentation Department, Cork Institute of Technology, Cork, Ireland,

[4] Innopharma Labs, 311 Q House, 76 Furze Road, Sandyford Industrial Estate, Dublin 18, Ireland.

[5] Pharmaceutical Manufacturing Technology Centre, School of Pharmacy, University College Cork, Ireland.

Abstract

The aim of this study was to highlight how variability in roller compacted ribbon quality can impact on NIR spectral measurement and to propose a simple method of data selection to remove erroneous spectra. The use of NIR spectroscopy for monitoring ribbon envelope density has been previously demonstrated, however to date there has been limited discussion as to how spectral data sets can contain erroneous outliers due to poor sample presentation to the NIR probes.

In this study compacted ribbon of variable quality was produced from three separate blends of microcrystalline cellulose (MCC)/lactose/magnesium stearate at 8 roll force settings (2 – 16 kN/cm). The three blends differed only in the storage conditions of MCC prior to blending and compaction. MCC sublots were stored at ambient (41% RH/20^{o}C), low humidity (11% RH/20^{o}C) and high humidity (75% RH/40^{o}C) conditions prior to blending. Ribbon envelope density was measured and ribbon NIR spectral data was acquired at line using a multi-probe spectrometer (MultiEye™ NIR).

Initial inspection of the at-line NIR spectral data set showed a large degree of variability which indicated that some form of data cleaning was required. The source of variability in spectral measurements was investigated by subjective visual examination and by statistical analysis. Spectral variability was noted due to the storage conditions of MCC prior to compaction, roll force settings and between individual ribbon samples sampled at a set Roll Force/Blend combination. Variability was also caused by ribbon presentation to probes, such as differences in the presentation of broken, curved and flat intact ribbons.

Based on the subjective visual examination of data, a Visual Discard method was applied and was found to be particularly successful for blends containing MCC stored at ambient and low humidity. However the Visual Discard method of spectra cleaning is subjective and therefore a non-subjective method capable of screening for erroneous probe readings was developed. For this data set a Trimmed Mean method was applied to set a limit on how data is cleaned from the data set allowing

for the removal of a faulty probe reading (25% of data) or a poor sample (33% of data). The 33% Trimmed Mean reduced the impact of spectral variation or misreads between samples or probes and was found to be as successful as the Visual Discard method at cleaning the data set prior to development of the calibration equation.

1. Introduction

Roller compaction is a method of dry granulation used to improve material flow properties for downstream processing such as tablet compression. Dry granulation is particularly advantageous for blends containing active pharmaceutical ingredient (API) sensitive to heat and moisture, for which wet granulation would be unsuitable (Kleinebudde, 2004). Roller compaction involves feeding a powder blend between two counter rotating rollers, densifying the material into ribbon compacts. The roller compaction process is a continuous granule production method; the ribbon intermediate material is immediately milled in-line to give granule particles with improved flow.

The density of the compacted ribbon is a critical quality attribute for granules which are progressed to a subsequent tablet compression step. Ribbon density has a direct impact on granule flow and tablet hardness. Over densification of the blend at the roller compaction stage may result in tablets with lower hardness (Acevedo et al., 2012; Teng et al., 2009) as the compactibility of the roller compacted material may be reduced on tableting. Likewise under-densification at roller compaction fails to adequately improve the materials flow. The greater the ribbon density, the lower the fine fraction and therefore, the better the flowability of the resulting granules (Peter et al., 2010).

A number of reports have investigated the use of NIR (near infrared) spectroscopy for monitoring roller compacted ribbon density (Acevedo et al., 2012; Gupta et al., 2004, 2005; Khorasani et al., 2015a; Khorasani et al., 2015b; Lim et al., 2011). These studies have taken advantage of the fact that the NIR absorption of a material is affected by the density of the material being analysed and therefore offers the potential of measuring/monitoring roller compacted ribbon density during production. Kirsch and Drennen (Kirsch and Drennen, 1999) first used the slope of the best-fit line through the NIR spectrum to quantify an observed upward shift in the baseline in response to the hardness of cimetidine tablets produced with increasing compaction pressure. The advantages in using the entire spectral data over single wavelength regression calibrations was investigated, as calibrations based on single wavelength regression models can be more susceptible to variation in chemical composition. Gupta et al (2004) later applied NIR spectroscopy to monitor roller

compacted ribbons. The slope of the best-fit line through the NIR spectra, measured off-line, was correlated to the mechanical strength of microcrystalline cellulose roller compacted samples. A comparison of NIR spectral slopes acquired off-line and in-line, in real-time for a compacted 10% tolmetin blend showed good agreement. However, more variation and outlier slopes were reported for the in-line spectral slope data compared to off-line data.

Gupta et al (2005) also examined the feasibility of using NIR spectra to detect changes in ribbon tensile strength due to variability in the ambient moisture content of microcrystalline cellulose ribbons using the spectral slope approach (Gupta et al., 2005). Additionally, a multivariate analysis partial least squared (PLS) model of NIR data, acquired off-line, was used to construct a calibration relationship between ribbon relative density, tensile strength and Young's modulus. In-line NIR spectral measurements which focused on the site of ribbon production, exiting the rolls were acquired. It was found that PLS model predictions of ribbon properties from NIR data acquired in-line, showed large variance from actual values measured. The authors attribute differences between the PLS calibrations off-line spectral data and in-line data to differences in the elastic recovery of the material 24 hours after compression.

Studies are limited with respect to a number of the practical issues encountered when applying a process analytical technology (PAT) or NIR technique to monitor roller compacted ribbon in-line. Research carried out on a single component compacted material, such as microcrystalline cellulose (MCC), does not address the spectral variability observed with multi-component blends, such as used in pharmaceutical products. During in-line analysis, compacted samples presented to the probe can be broken, split or curved depending on the blend compaction properties and processing parameters. MCC alone is capable of producing intact, uniform ribbon samples for presentation to the NIR probe in-line. Ribbon sample presentation to the NIR probe in-line becomes problematic for many pharmaceutical blends that show little deformation giving brittle, broken ribbon pieces or a high level of deformation producing curved ribbons. Off-line NIR spectral data has been successfully used to construct calibration models to predict compacted ribbon physical properties (Gupta et al.,

2004; Lim et al., 2011; McAuliffe et al., 2014). Although this approach proves the validity of the use of NIR spectral data to detect and quantify differences between ribbon samples, it ignores the variability in spectral data due to sample presentation when monitoring compacted ribbon in-line. Pre-selection of ribbon sections or defining sections for scanning is feasible during off-line analysis but considerably more challenging when monitoring ribbon in-line. Variable ribbon presentation can generate variable NIR spectral data which can be challenging to screen, analyse and interpret.

The aim of in-line PAT technologies is to monitor a process and detect outliers which indicate the process is moving out of specification (OOS). However the outliers flagged by the PAT method must be a true representation of changes in the sample quality. It is futile and misleading to perform data based analysis when the data is contaminated with sampling error outliers. These outliers, which do not represent the sample, may be due to sensor noise, human error or issues with sample presentation. Inclusion of these outliers can lead to model misspecification (poor calibration) and incorrect analysis results (Liu et al., 2004). Therefore the effect of sample presentation to the NIR probe is an important consideration when designing a method of data processing to deal with the inevitable sampling error expected when applying PAT to roller compacted ribbons in-line.

Data processing in real time is challenging as sampling error outliers must be either filtered or cleaned from the data set before further analysis. Filtering implies the use of past and current data to estimate the true result in place of an outlier. Cleaning involves the deletion of outliers (Liu et al., 2004). Visual inspection of the data and removal of sampling error outliers though justifiable is not practical for real time data processing. All spectra should be subject to a statistical spectral quality test to determine whether the characteristics of the sample fall within the range of variation for which the model was calibrated and validated (EMA, 2014). Simple pre-treatment solutions are desirable to allow for rapid analysis of data to enable prompt process adjustment. Simple pre-treatment solutions are also desirable from a quality assurance viewpoint to allow for ease of interpretation of raw data manipulation techniques by quality assurance and regulatory staff.

In this study we investigate the challenges presented during the development of a calibration model for the prediction of compacted ribbon envelop density from NIR spectra data. Spectral data is acquired at-line from ribbon samples produced with variable quality. We highlight how variability in a compacted ribbon can impact on NIR spectral measurements. The study interrogates the source of variability in these spectral measurements and proposes a simple data processing approach to manage this variability based on Trimmed Mean method.

2. Materials and Methods

2.1 Materials

Microcrystalline cellulose (Avicel PH102®) was supplied by FMC BioPolymer, Ireland, anhydrous lactose (Supertab 21 AN®) obtained from DFE Pharma, Germany and magnesium stearate obtained from Merck KGaA, Germany. Placebo blends were prepared composed of 49.875% microcrystalline cellulose, 49.875% lactose and 0.25% magnesium stearate. Blends differed only in the storage conditions of microcrystalline cellulose prior to blending and compaction. MCC sublots were stored at ambient (41% RH/20^{o}C), low humidity (11% RH/20^{o}C) and high humidity (75% RH/40^{o}C) conditions prior to blending **Table 1**. Following storage the moisture content of microcrystalline cellulose was determined by thermogravimetric analysis (TGA) using a TA Q500, TA instruments **Table 1**. All samples were analysed by heating from ambient temperature to 120°C in an inert Nitrogen atmosphere at 5°C/min. The moisture content was determined using TA Universal Analysis Software by determining the loss in mass of the samples across this temperature range. Each blend was prepared using a Erweka AR 403 and double cone mixer DKM (Germany) operated at 30 rpm. Microcrystalline cellulose and lactose were blended for 10 min prior to the addition of magnesium stearate which was blended for a further 1 min.

2.2 Roller compaction set-up

Blends were compacted using a Fitzpatrick CCS220 roller compactor, USA. A combination of a smooth roll and serrated roll both with a diameter of 20 cm and a width of 2 cm were used to ensure good material flow into the nip region of the rolls and to avoid material slippage during compaction. A horizontal feed screw with a pitch of 2.5 cm was used to feed material from the hopper into the vertical feed screw with a pitch of 1.5 cm set at a speed of 300 rpm. The horizontal feed screw speed was controlled automatically by the Fitzpatrick roller compactor control system to obtain the set roll gap of 0.5 mm. The roll speed was kept constant at 2 rpm. The process room relative humidity was monitored throughout the study duration (43 ± 4 % RH). Ribbons of variable quality were produced by varying the compaction roll force and moisture content of the microcrystalline cellulose component of the blend. Examples of ribbons of variable quality produced during this study are presented in **Figure 1**. Ribbons were produced at 24 Roll Force/Blend combinations; from 3 blends which varied in the pre-blending storage conditions of the microcrystalline cellulose component (as detailed in section 2.1) and 8 roll force settings which varied between 2 and 16 kN/cm, increasing in 2 kN/cm intervals.

2.3 Ribbon Envelope Density

Ribbon envelope density was determined using the GeoPyc™ 1360, USA. Envelope density measurements were carried out on three samples of ribbons for each of the 24 Roll Force/Blend combinations. The weight of ribbon was predetermined prior to analysis. The pycnometer determines the volume of the ribbon by measuring the volume of Dryflo™ displaced by the ribbon. The internal diameter of the sample chamber was 25.4 mm, a consolidation force of 51 N and a conversion factor of 0.5153 cm^3/mm respectively were used for analysis.

2.4 Acquisition of NIR spectra at-line

Ribbon NIR spectral data was acquired at-line. A MultiEye™ NIR spectrometer, Ireland scanned ribbon samples at-line as they were produced by the roller compactor, i.e. samples were not pared

to a predetermined sample size. The objective in presenting ribbon samples in this manner was to mimic as closely as possible in-line ribbon presentation. The MultiEye™ instrument is based on Fabry-Perot interferometer technology. A halogen light source and a four point reflectance probe was used to analyse the roller compacted ribbon in reflection mode. The four fibre optic probes were positioned at a fixed distance above the ribbon samples in series, obtaining spectra along the length of the smooth sided ribbon sample (**Figure 2**). Variation in spectral absorbance across the ribbon could not be studied due to the limited width of the ribbon samples relative to the distance between the four IR probes. A wavelength scan of 1550 nm to 2110 nm with a resolution of 20 nm was selected to ensure a fast integration time (<3 seconds). Reference dark and light spectra were obtained with a standard white reference disk. Ribbon samples were scanned as they were produced i.e. good, broken or split to give a sample which would be representative of ribbon quality encountered during real time, in-line analysis. For each of the 24 Roll Force/Blend combinations, 3 ribbons were selected and scanned 5 times by the MultiEye™ instrument using 4 probes, giving a set of 60 NIR spectra for each Roll Force/Blend combination.

2.5 Acquisition of NIR spectra off-line

A higher resolution scan of ribbon samples was performed using a Spotlight 400 FT-IR (PerkinElmer, USA) in diffuse reflectance mode. Spectra were recorded from 4000 to 7800 cm^{-1} (1282-2500 nm) with a resolution of 2 cm^{-1} (1.25 nm) and averaged over 8 scans. Ten spectra were measured per ribbon at 1 cm^{-1} intervals along the length of the ribbon. The resulting reflectance data were converted to absorbance (*A*) using equation 1.

$$A = log_{10}\left(\frac{1}{R}\right) \qquad \text{Equation 1}$$

where A is absorbance and R is reflectance.

2.6 Statistical methods

Detailed statistical analyses were used to interrogate sources of variability between the NIR spectra.

The dataset consisted of the 24 ribbon types (Section 2.2). For each of these ribbon types, 3 ribbons were selected and scanned 5 times under MultiEye™ (NIR Spectrometer) using 4 probes at 29 wavelengths (1550 - 2110 nm in steps of 20 nm), giving a set of 60 NIR spectra for each combination. See Appendix 1. Envelope density measurements were carried out off-line on three samples of ribbons for each Roll Force/Blend combination (Section 2.3).

2.6.1 Ribbon sample variability

Variability between the spectral slopes of individual ribbon samples produced at each ribbon Roll Force/Blend combination was investigated. Three ribbon samples were analysed for each of the 24 combinations. Two possible models were selected. The structural difference between these two models is that the alternative model contains a variance term allowing for random differences between samples. The null model does not allow for this source of variation (equations 2 and 3),

$$\text{Null Model}: y_{ijk} = b_j + \varepsilon_{ijk} \text{ , with Var}(b_j) = 0, \qquad \text{Equation 2}$$

$$\text{Alternative Model}: y_{ijk} = b_j + \varepsilon_{ijk} \text{ , with Var}(b_j) \neq 0 \qquad \text{Equation 3}$$

where y_{ijk} is the spectral slope response for each probe $i = 1, \dots 4$, sample $j = 1, \dots 3$, scan $k = 1, \dots 5$, $b_j \sim N(0, \sigma_1^2)$ is a random effect for sample *j*, and $\varepsilon_{ijk} \sim N(0, \sigma^2)$ is a model error term. We compared the models for each roll force/blend/probe combination (96 combinations) using the **exactRLRT** command in the RLRSim package in R (Scheipl et al., 2008). (This uses an (exact) restricted likelihood ratio test based on 10000 simulated values from the finite sample distribution.

2.6.2 Linear mixed effect analysis

Following some exploratory data analysis, analysis of variance was carried out using linear mixed model analysis (Gałecki and Burzykowski, 2013; Pinheiro and Bates, 2006) with the spectral slopes as responses, ribbon samples as a random effect and probes as a fixed effect. Two models for each of

three ribbon samples from the 24 Roll Force/Blend combinations were compared. The model formulation was as described in equations 4 and 5.

$$\text{Model 1: } y_{ijk} = \beta_i + b_j + \varepsilon_{ijk} \qquad \text{Equation 4}$$

$$\text{Model 2: } y_{ijk} = \beta_i + b_j + b_{ij} + \varepsilon_{ijk} \qquad \text{Equation 5}$$

where y_{ijk} is the spectral slope response for probe $i = 1, \dots 4$, sample $j = 1, \dots 3$, scan $k = 1, \dots 5$, and β_i is a fixed effect for probe *i*, $b_j \sim N(0, \sigma_1^2)$ is a random effect for sample *j*, $b_{ij} \sim N(0, \sigma_2^2)$ is a random effect for the interaction of probe and samples, and $\varepsilon_{ijk} \sim N(0, \sigma^2)$ is a model error term. The model was fitted using the **lmer** command in the LME4 (Bates et al., 2014 and 2015) package in R (R Core Team., 2013).

2.6.3 Linear Contrasts

During the study the question arose as to the effect of the observed curvature of the ribbons on variation in the NIR spectral data acquired by each of the four probes. To investigate this effect we used the method of linear contrasts as defined in equation 6 (Montgomery, 2001).

$$C = \frac{1}{2}\mu_{p_1} - \frac{1}{2}\mu_{p_2} - \frac{1}{2}\mu_{p_3} + \frac{1}{2}\mu_{p_4} \qquad \text{Equation 6}$$

where *C* is the contrast, μ_{p_i} denotes the mean for probe *i*, and two possible models were compared for differences between the probes (equations 7 and 8).

$$\text{Model 1: } C = 0\ , (\mu_{p_1} + \mu_{p_4} = \mu_{p_2} + \mu_{p_3}) \qquad \text{Equation 7}$$

$$\text{Model 2 : } C \neq 0\ , (\mu_{p_1} + \mu_{p_4} \neq \mu_{p_2} + \mu_{p_3}) \qquad \text{Equation 8}$$

using an F-test where $F = \frac{\text{MS}_c}{\text{MS}_e}$ where MS_c is the contrast mean square error and MS_e is the mean square error. The 95% confidence limits for *C* were calculated using the standard equation (equation 9)

$$C \pm t_{\frac{\alpha}{2}} \sqrt{\text{Var}\,(C)} \qquad \text{Equation 9}$$

where *C* was estimated by $C = \frac{1}{2}\,(\bar{p}_1 - \bar{p}_2 - \bar{p}_3 + \bar{p}_4)$, and $\bar{p}_i = \sum_{j=1}^{15} p_{ij}$ is the average for probe *i*. Var(*C*) is the variance of *C* which we estimate by equation 10

$$\mathrm{Var}(C) = \frac{\mathrm{MS}_e}{15} \sum_{i=1}^{4} \left(\frac{1}{2}\right)^2 \qquad \text{Equation 10}$$

and the equation simplifies to equation 11

$$\frac{1}{2}\left(\bar{p}_1 - \bar{p}_2 - \bar{p}_3 + \bar{p}_4\right) \pm t_{0.25}\sqrt{\frac{\mathrm{MS}_e}{15}} \qquad \text{Equation 11}$$

2.6.4 Moisture analysis

For moisture content analysis a standard normal variate (SNV) transformation (equation 12) was applied to all at line NIR spectra using The Unscrambler™ software (version 10.3), CAMO Software AS, Oslo, Norway. The SNV transformation was applied to remove slope variations between spectra by mean centring each spectrum which reduces the effect of the physical variation in the ribbon (density effects and surface effects). The impact of moisture content on the NIR spectra collected for each of the three blends could then be explored.

$$Z = \frac{(X-\mu)}{\sigma} \qquad \text{Equation 12}$$

where Z is the standard normal variant, X is a normal variant with mean μ and standard deviation σ.

2.7 Calibration model to predict ribbon envelop density

A linear regression of the spectral slopes to the envelope density was used to model the relationship between the NIR at-line spectra from the ribbon envelope density. Note that since this is a calibration equation, it was assumed that the envelope density was the independent variable, measured without error, and used inverse prediction (Mullins, 2003). Three different options were considered when selecting spectral data to include in the calibration data set (removal of outliers); (1) the average of all 60 spectral slopes acquired for each Roll force/Blend combination without removing any outliers (Full Mean), (2) the average slope was calculated by inspecting the 60 individual spectra for each Roll Force/Blend combination visually and discarding those that appeared

to have been misread (Visual Discard) and (3) the 33% Trimmed Mean which was calculated by ranking the 60 spectral slopes for each Roll force/Blend combination and discarding the highest and the lowest 20 samples (33%), and averaging the remaining 20 samples. Common examples of the use of the Trimmed Mean include the scoring method in sports such as boxing and gymnastics where the highest and lowest judges' scores are discarded and the LIBOR interest rate (https://www.theice.com/iba/libor) which, based on the rates for 18 banks, discards the rates for the highest and lowest four banks.

3. Results

3.1 Variability between ribbon samples

The initial visual inspection of the NIR spectral plots in **Appendix 1** indicated that at a number of the Roll Force/Blend combinations individual samples showed wide variation between spectra. Prior to the development of a calibration equation to relate NIR spectral data acquired at-line to ribbon envelope density measurements, the variability in ribbon density and spectral slope data was examined. Initial visual inspection of the data showed that ribbon envelope density increased with increase in roll force. This relationship is in agreement with previous studies, as the ribbon density increases there is less diffuse scattering of the NIR beam due to the reduction in air particle boundaries in the ribbon sample, less light reaches the detector and thus give an apparently higher absorbance (Donoso et al., 2003; Short et al., 2009). This phenomenon has been used in a number of studies applying NIR spectroscopy to measure roller compacted ribbon density (Acevedo et al., 2012; Gupta et al., 2004, 2005; Khorasani et al., 2015a; Khorasani et al., 2015b; Kirsch and Drennen, 1999; Lim et al., 2011; McAuliffe et al., 2014). The increase in ribbon envelope density was observed for all blends at lower roll forces, the increase in density was less evident for roll forces greater than 8 kN/cm (**Figure 3**).

Two way ANOVA analysis was performed to investigate the effect of roll force setting and blend differences (i.e. microcrystalline cellulose storage conditions prior to blending) on resultant ribbon envelope density. Results of this analysis showed that the ribbon envelope density samples produced at different roll forces are significantly different ($p < 0.001$). Also, a significant difference was observed due to the storage conditions of the microcrystalline cellulose prior to blending when compacted at roll forces 4, 12 and 14 kN/cm ($p < 0.001$). The overall interaction between the factors of roll force and microcrystalline cellulose storage conditions was not found to be significant.

Plots of the envelope densities and the NIR spectral slopes by roll force for individual ribbon samples show that while there does appear to be a relationship between the variables, there is also considerable variation between samples (**Figure 3**). Plots of the envelope densities against the spectral slopes also show much variation in individual results (**Figure 4**)**.** Note the envelope densities plotted in **Figure 4** are averages of individual readings for 3 dependent ribbon samples. For the Blends 1 and 2, containing MCC stored under ambient and low relative humidity conditions, there does appear to be a linear relationship between spectral slopes and envelope densities. However, this linear trend was not evident for Blend 3 containing MCC stored at a high % relative humidity and there is a large variability in spectral slopes for spectra acquired from samples produced at the same Roll Force/Blend combination.

When investigating the variability between spectral slopes and envelope densities for Blend 3 the influence of higher levels of moisture in these samples warrants consideration. The moisture content of microcrystalline cellulose contained in blend 3 was 6.96% w/w and higher than moisture content of microcrystalline cellulose included in blends 1 and 2, which was 4.86% and 4.12% respectively. The impact of moisture content on spectral absorbance of ribbons was explored using standard normal variate (SNV) pre-treatment to correct for scattering due to density differences between samples. SNV pretreatment highlighted increased spectral absorbance in the OH region at 1950 nm for compacted blend 3 compared to blends 1 and 2 (**Figure 5**). The presence of higher moisture content would increase spectral slope values does not explain the variability in the spectral slopes

observed for blend 3 ribbons. Variability is also attributed to the impact of increased MCC moisture content on the microcrystalline cellulose compaction (Sun, 2008) and hence quality of the ribbons produced.

The variability between the spectral slopes of individual ribbon samples produced at each ribbon Roll Force/Blend combination was further analysed by selecting between two statistical models with and without a variance term which allowed for random differences between probes (**equations 2 and 3**) (section 2.6.1). The results of these restricted likelihood tests show that for ribbons produced from blend 3, or acquired by probes 2 or 3 (with the exception of spectral data acquired by probe 3/blend 2/roll force 6 kN/cm) there are differences in the spectral slope data between the three samples and therefore a random intercept should be included in the model (**Table 2**). This is also true for the majority of the remaining Roll Force/Blend combinations (probes 1 and 4, Blends 1 and 2). The conclusion from this is that the spectral slopes of the samples are significantly different within most Roll Force/Blend/Probe combinations. Applying even the conservative Bonferroni correction for multiple testing (Pocock et al., 1987) would not significantly alter this conclusion.

3.2 Source of NIR spectral slope variability- exploratory analysis

Visual inspection of NIR spectral plots in **Appendix 1** from which the spectral slopes are calculated in some cases, show flat spectra and may have been misread. Some exploratory analysis was performed to investigate the source of these differences in spectra. Variation within the sets of 60 spectra for each Roll Force/Blend combination can arise from variation between the three ribbon samples (as described above), between the 4 probes and between the 5 scans acquired by each probe. Further visual examination of the NIR spectra for some Roll Force/Blend combinations also showed large variations both between the three samples of ribbons and the four probes. For example (**Figure 6**) shows boxplots of the spectral slopes for ribbons produced from blend 3 at a roll force of 12 kN/cm and spectra for ribbons produced from blend 1 compacted at a roll force of 4 kN/cm. There is a very large variation between spectral slopes within ribbon samples of blend 3, and

much lower variation was seen for ribbons for blend 1. This variation in spectral slopes for ribbon samples produced for the same Roll Force/Blend combination was attributed to the quality of the ribbon sample and increased moisture content in the case of blend 3.

the impact of high moisture content on the spectral slope was also explored by SNV analysis (section 3.1) Sample quality when presented to the NIR probes at-line was seen to influence the quality of the NIR data set acquired. To relate NIR spectral data to ribbon density it is crucial that the NIR spectra captured truly reflect the physical properties of the ribbon being produced. Ribbon presentation to the NIR probe should allow all four probes to acquire a representative spectrum. Ideally to achieve this ribbons should be flat and long as seen in **Figure 1(a-b)**. Issues arose when ribbon produced was curved or broken and split as seen in **Figure 1(c-d)**. Broken ribbons were particularly relevant for blend 3 when compressed at higher roll forces. However, even ribbons that visually appeared to be flat in some cases showed higher absorbance for Probes 1 and 4, than for Probes 2 and 3. Curves in the ribbon surface resulted in lower absorbance values than expected which was attributed to the angle the NIR reflected beam hits the detector. **Figure 7a** shows an example of the type of NIR spectra sampling error obtained for one piece of ribbon scanned under the 4 probes, similar to the ribbons in **Figure 1(c-d)**. In this example the spectral data collected by the beam spot of probe 2 is attributed to an area broken or split ribbon where only a fragment of sample was scanned by probe 2.

A comparison of at-line and off-line ribbon NIR spectral data for ribbon produced at the same Roll Force/Blend combination showed that this sampling issues encountered during at-line NIR spectral acquisition was not encountered with off-line spectral analysis. **Figure 7b** shows the off-line data of ribbon sample produced the same Roll Force/Blend combination as shown in **Figure 7a**. The off-line NIR spectral data showed no similar sampling issues. It was concluded that the instrumental set up was responsible for sampling error in the at-line spectral data. The offline, Spotlight 400 FT-IR (PerkinElmer) system, uses an incident angle between 10 and 28 degrees, the distance between the sample and the detector is fixed hence a more consistent signal is detected for each ribbon. The spot

size of the offline instrument is 0.01 μm^2 so therefore curvature of the ribbon has little to no impact on the reflectance measurement. For the at line MultiEye™ set-up the path length is much longer (approximately 15 cm) and the spot size is larger, approx. $1mm^2$. Both these factors influence the amount of signal reflected from the ribbon hence causing a different off set between the at-line and off-line measurements and less consistent signal detected for each ribbon.

It was concluded that misread spectra did not reflect true measurements of ribbon samples and it was therefore assumed that the deletion of erroneous spectral data from the spectral data for ribbon physical analysis set was justified before attempting a calibration equation to relate NIR spectral data acquired at-line to ribbon envelope density measurements. However, this leads to the questions of how to decide which spectra to remove and how many should be removed. Also, since the exploratory analysis only considered a small number of (representative) samples, we next undertook some more rigorous statistical analysis on the full NIR spectral database prior to deciding on a strategy to remove non-representative spectra from the data set.

3.3 Source of NIR spectral slope variability- statistical analysis

Linear mixed model analysis. To investigate the variability between ribbon samples and probes, analysis of variance was carried out using linear mixed model analysis (Gałecki and Burzykowski, 2013; Pinheiro and Bates, 2006) with the spectral slopes as responses, ribbon samples as a random effect and probes as a fixed effect (see **section 2.6.2**). Comparison of Model 1 (equation 4) and Model 2 (equation 5) using the likelihood ratio test in the anova function showed that in all cases Model 2 explained the spectral slope variation better than Model 1 (p<0.001 for Chi-square test for all ribbons except for ribbons samples for those produced by the blend 2/roll force 6 kN/cm, where p<0.01), indicating that a random effect for the interaction of probe and samples should be included. Appendix 2 shows detailed results including AIC, BIC, Log likelihood. These show that in all cases, Model 2 is to be preferred and therefore it was decided to use Model 2. (Application of the

Bonferroni correction for multiple testing (Pocock et al, 1987) would not significantly alter this conclusion).

The variance components of Model 2 for the random effects (samples and interaction) are shown in **Table 3** for representative samples of the full set of ribbons. Including the random term for the samples in the model, reduces the spectral slope variation. For example variation is reduced from 196 x 10^{-10} to 1 x 10^{-10} for blend 3 ribbons compacted at 12 kN/cm and from 6.35 x 10^{-11} to 2.77 x 10^{-11} for blend 1 ribbons compacted at 4 kN/cm. The percentage reduction is much greater for the blend 3 ribbons, reflecting the greater variation between samples evident in **Figure 6.**

Analysis of the fixed effects (probes) shows that Probe 1 (Intercept) had the highest estimated value, followed by Probe 4, with Probes 2 and 3 lower, **Table 4**. In summary, this linear mixed effect analysis shows that the interaction term between samples and probes should be included (different samples behave differently for different probes), and also that probes scanning the ends of the ribbon (Probes 1 and 4) tend to give higher absorbance intensities than do those in the middle. This statistical analysis supports the finding from informal investigations in (**section 3.2**).

Linear Contrasts. The linear mixed model analysis above indicated that the variance in spectral slopes for probes 1 and 4 (at the ends of the ribbons) tended to be higher than for probes 2 and 3 (in the middle of the ribbons). A linear contrasts method was used to investigate the magnitude of any difference in variance between the end probes (probes 1 and 4) and the other two probes (**section 2.6.3**).

Confidence limits for the linear contrast for each of the 24 Roll Force/Blend combinations ribbon samples are given in **Table 5**. The results show that for most ribbons the confidence limits are both positive indicating that $\mu_{p_1} + \mu_{p_4} > \mu_{p_2} + \mu_{p_3}$ which means that the variance for Probes 1 and 4 is significantly greater than for Probes 2 and 3, while only for ribbons produced from blends 1 and 3 compacted at 2 kN/cm and from blend 2 at 4 and 6 kN/cm the confidence limits were negative indicating that $\mu_{p_1} + \mu_{p_4} < \mu_{p_2} + \mu_{p_3}$. None of these confidence intervals included zero (which would have indicated no difference between the pairs of probes).

3.4 Developing a calibration equation between spectral slopes and envelope density

Exploratory and statistical analysis of spectral data set concluded that the deletion of erroneous misread spectral data from the at-line spectral analysis set was justifiable (**Figure 7**). Inspection and analysis of the data set revealed significant differences between samples, between probes and the overall level of variation. Therefore it was decided to develop the calibration equation between spectral slopes and envelope density using a range of approaches to select a spectral data set to be incorporated in this model and data to be discarded.

Two calibration approaches were undertaken. Firstly a Visual Discard method was employed. The average slope for each ribbon type was calculated by inspecting the 60 individual spectra for each Roll Force/Blend combination visually and discarding those that appeared to have been misread; this was termed the Visual Discard method. This approach is not a practical solution for a real-time monitoring of ribbon envelope density due to its subjective nature and to the need for timely visual inspection, it was included in the analysis for comparison with the other methods.

The second method used a 33% Trimmed Mean. The 33% Trimmed Mean was calculated by ranking the spectral slopes for the 60 samples for each ribbon type, discarding the highest and the lowest 20 samples (33%), and averaging the remaining 20 samples. The desired objective was to discard results from an entire sample or to ignore a probe if there has been a misread.

For comparison, the results obtained were compared with those obtained using all 60 results for each ribbon type and with the off-line spectral data. A simple linear regression was performed of the spectral slopes on the envelope density, using the average of the three envelope densities as the independent variable and (i) the average of all 60 spectral slopes acquired off-line (Full Mean) (ii) the average after visual discard (Visual Discard), (iii) the 33% Trimmed Mean (Trim Mean), and (iv) off-line spectral data, as the dependent variables (**Table 6 and Figure 8**).

Use of the 33% Trimmed Mean compared to the Full Mean increased the correlation coefficient (r) from 93% to 96% (and the level of explained variation (R^2) from 87% to 92%) for Blend 1, and r

increased from 89% to 97% (with R^2 increasing from 79% to 93%) for Blend 2. However the overall high level of variation in the NIR spectral slopes was greatest from the high-moisture Blend 3 (**Figure 8c**) which renders it unsuitable for use in a calibration equation and even using the 33% Trimmed Mean does not improve the situation (r = 0.74, R^2 = 54%). The Trimmed Mean results also show an improvement on the Visual Discard results for Blends 1 and 2, which leads to the conclusion that the simpler Trimmed Mean method is preferable to the labour-intensive Visual Discard method. However, for Blend 3 the Visual method gave higher correlation than the Trimmed Mean method (**Table 6**). The poor correlation for the blend 3 ribbons is the result high variability in spectral slopes which is attributed to the combined, related effects of increased moisture and poor quality split broken ribbon, previously discussed in **section 3.1**.

As an alternative to the 33% Trimmed Mean, we therefore considered a calibration based on the top 33% of the spectral slopes after ranking. However, while this improves the correlation for Blend 3 from r=0.74 to r=0.93 due to the high level of variation and misreads, the corresponding results for Blend 2 decrease from r= 0.97 to r=0.65 (due mainly to the results for ribbons produced at the lowest roll force of 2kN/cm). After consideration of the overall levels of variation within all the ribbon sets, it was concluded that the 33% Trimmed Mean was preferable to the top 33% and the visual inspection method.

For the off-line NIR spectral slopes, where issues with misread spectra due to sample presentation are not encountered (see **section 2.5**) as shown in **Figure 7b**, data showed a good correlation between spectral slope and increasing ribbon envelope density for each blend as shown in **Table 6**. However, comparison with the correlation coefficients for the Trimmed Mean calibration equations showed that the method using the Trimmed Mean is at least as good as using spectral data acquired using the off-line method.

4. Discussion

The results presented highlight the challenge of interpreting a noisy spectral data set obtained when analysing ribbon samples of variable quality using a multi-probe instrument. This study design aimed to capture the variability presented by roller compacted ribbons when applying a NIR spectroscopy PAT technique to monitor ribbon envelope density at-line. It was anticipated that the results obtained in this study at-line give insight into the variability in ribbon quality and spectral reading that will be obtained in-line.

Due to the nature of ribbon production during roller compaction sampling error is inevitable as few materials compact to form perfect flat ribbons continuously intact for NIR analysis. Many drug substances have poor physical properties and can be particularly brittle or plastic (Teng et al., 2009). A high drug loaded brittle API can form brittle spilt ribbons or alternatively a very plastic API can be too sensitive to the roller forces and over deform. The aim in roller compaction is not to produce perfect ribbons but to produce granules of suitable quality for subsequent processing, such as compaction. Granules must be compacted enough to improve blend flow downstream while avoiding over compaction which can reduce tablet quality (Teng et al., 2009). Ribbon envelope density was influenced by the roller compactor settings, most notably the roll force and the compressibility behaviour of the blend itself. In this study the storage of microcrystalline cellulose under a range of temperature and humidity setting prior to blending and compaction introduced variability in material compaction behaviour while keeping the main constituents of the placebo blend composition constant to facilitate comparisons between blends.

It should be noted that during the study NIR spectra was acquired at-line, with an integration time of < 3 secs. To attain the high data acquisition speed required for monitoring a commercial roller compaction process in real-time, a reduction in integration time would be required. The set-up of the NIR instrument in-line and the identification of an optimal signal to noise ratio would be critical to sufficiently reduce the integration time.

Visual inspection of the full spectral data set obtained (**Appendix 1**) showed much variability in the spectral data and on closer analysis multiple sources of variability existed in the data set. Blend

properties (compression force and microcrystalline cellulose storage conditions) translated into ribbons of differing quality. Increasing roll force produced denser ribbons a relationship which was reflected in increased spectral slopes. The high moisture content blend 3 (MCC stored at 75% RH) in particular compressed strongly producing cracked curved ribbons which were problematic for NIR analysis. Ribbon envelope density analysis showed an inherent variability between samples of the same blend produced at the same roller compaction setting. However this does not account for the extent of the spectral variability. Closer visual inspection of the data obtained from each probe showed that some probes failed to capture sample when scanning due to broken, split or curved ribbon samples. The theory that the curvature in ribbons resulted in lower spectral values then expected was also supported by Linear Contrasts analysis (**section 2.6.3**) as the spectra produced by probes positioned at the end of the ribbon sample (probes 1 and 4) were significantly different to those from the middle probes 2 and 3.

The variability in the blend 3 at line data set was also impacted by the microcrystalline cellulose moisture content. This effect of moisture was not evident for blends 1 and 2 which had lower and similar moisture content. All blend 3 ribbons consistently showed higher absorbance in the moisture region (1950 nm) thus increasing the spectral slope. A possible solution to reduce the impact of moisture on spectral slopes is to reduce the spectral range of analysis and eliminate the wavelength region influenced by moisture i.e. to analyse the slope from 1550-1900nm. However restricting the spectral range would not correct for the variability in spectral slopes due to poor ribbon quality as a result of the higher moisture content of microcrystalline cellulose in blend 3.

Sample preparation and sample presentation during off-line NIR analysis produced more consistent spectra. Sample presentation to the probes was not an issue as the off-line set up sampled much smaller ribbon areas and focused on the sample before scanning thereby giving a good quality scan each time. However while such consist spectral data provides good calibration between NIR slopes and off-line ribbon density data its use is limited for rapid analysis due to equipment costs and duration of analysis

Having determined the main sources of spectral variability the next stage was to explore different pre-treatment and calibration options in order to identify a simple effective solution. All initial analysis of the full data set indicated that some form of data cleaning was required in order to obtain an at-line data set from which to produce a calibration model. The Visual Discard method was applied to the full data set and was found to be particularly successful for blends 1 and 2 using the spectral slope data .The issue of sampling error due to ribbon quality at presentation to the probe remained a challenge for blend 3. However the Visual Discard method was successful but laborious, time consuming and an impractical method of PAT analysis where the ideal is rapid measurement to allow for process adjustments during processing.

The Trimmed Mean was investigated as a non-subjective method capable of screening for these erroneous probe readings. The Trimmed Mean method set a limit to how data is cleaned from the data set allowing for the removal of a faulty probe reading (25% of data) or a poor sample (33% of data). As a spectral slope technique was used to calculate the calibration equation, the conclusion that curved ribbons may lead to lower spectra than straight ribbons is not in itself a reason to discard such spectra (as the slopes may not be affected). However, the flat spectra due to broken ribbons do need to be discarded. This is another reason to prefer the Trimmed Mean method to the Visual Discard method since the distinction between these two types of spectra may not be clear, and the decision on whether to delete is subjective. The correlation coefficients for the Trimmed Mean equations were comparable with the calibration equations obtained using the off-line spectra and Visual Discard pre-treatment.

Major variation is apparent in the spectral slopes for Blend 3 (75%) which renders it unsuitable for prediction and calibration. Using the 33% Trimmed Mean in place of the Full Mean reduces the impact of variation or misreads between samples or probes. Further development of this method when used for calibration could investigate the optimal number of ribbon samples to % data trimmed to achieve a statistically robust balance of data cleaning. Increasing the number of samples used during calibration may allow for a greater retention of data in the Trimmed Mean e.g. 5

samples and keeping 3 giving a 20% Trimmed Mean. The Trimmed Mean method is no less successful than other simple calibration methods but offers the advantage of objective, non-biased means of removing NIR data collected during sampling error at-line/in-line. An alternative to the Trimmed Mean which could also be considered is the use of the median. We preferred the Trimmed Mean to the median as it is less wasteful of the data while still reducing the bias due to erroneous data. A further alternative for future work is an automatic method of outlier detection and deletion. However, an automatic method is unlikely to improve on a manual discard method, and we found the manual method gave results which were no better than the Trimmed Mean. In addition to its effectiveness, a major advantage of the Trimmed Mean method is its simplicity which facilitates understanding of the pre-treatment methodology. Implementation of such an approach in quality environment enables quality, analytical and production personal to understand the rationale of the approach and raw data manipulation to obtain a calibration data set.

5. Conclusions

NIR spectral analyses can be used to monitor the ribbon density of roller compacted ribbons. However studies to date have not reported the challenges of noisy spectral data sets due to variability in ribbon sample quality and presentation of ribbons to probes during at-line and in-line analysis. In this study the NIR data set generated showed a large degree of variance due to variability in ribbon quality. The use of a pre-treatment method, the 33% Trimmed Mean method, was shown to be an effective method to discard spectral results due to broken, split or curved ribbon. The pre-treatment technique was as effective as the removal of non-representative spectra using the Visual Discard method and the off-line spectral data obtained using a high resolution NIR array spectrometer. The 33% Trimmed Mean method offers a simple and practical solution to dealing with the NIR spectral challenges when applying this PAT technique to roller compacted ribbon in-line/at-line.

Acknowledgments

Research is funded by the Synthesis and Solid State Pharmaceutical Centre (SSPC) under grant number 12/RC/2275, the Mathematics Applications Consortium for Science and Industry funded by Science Foundation Ireland Investigator Award 12/IA/1683 and the Enterprise Ireland/Irish Development Authority (IDA) funded Pharmaceutical Manufacturing Technology Centre, Ireland. Microcrystalline cellulose PH102 donated by FMC Corporation. Innopharma Laboratories for use of the MultiEye™ NIR Spectrometer.

.

References

Acevedo, D., Muliadi, A., Giridhar, A., Litster, J.D., Romanach, R.J., 2012. Evaluation of Three Approaches for Real-Time Monitoring of Roller Compaction with Near-Infrared Spectroscopy. AAPS PharmSciTech 13, 1005-1012.

Bates D, Maechler M, Bolker B and Walker S., 2014. _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-7, <URL: http://CRAN.R-project.org/package=lme4>.

Bates D, Maechler M, Bolker B and Walker S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.

Donoso, M., Kildsig, D., Ghaly, E.S., 2003. Prediction of Tablet Hardness and Porosity Using Near-Infrared Diffuse Reflectance Spectroscopy as a Nondestructive Method. Pharmaceutical development and technology 8, 357-366.

EMA, 2014. Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations. European Medicines Agency Website EMEA/CHMP/CVMP/QWP/17760/2009 Rev2 accessed on 05/08/2016.

Gałecki, A., Burzykowski, T., 2013. Linear mixed-effects models using R: A step-by-step approach. Springer Science & Business Media.

Gupta, A., Peck, G.E., Miller, R.W., Morris, K.R., 2004. Nondestructive measurements of the compact strength and the particle-size distribution after milling of roller compacted powders by near-infrared spectroscopy. J. Pharm. Sci. 93, 1047-1053.

Gupta, A., Peck, G.E., Miller, R.W., Morris, K.R., 2005. Influence of ambient moisture on the compaction behaviour of microcrystalline cellulose powder undergoing uni-axial compression and roller-compaction: A comparative study using near-infrared spectroscopy. J. Pharm. Sci. 94, 2301-2313.

Khorasani, M., Amigo, J.M., Sonnergaard, J., Olsen, P., Bertelsen, P., Rantanen, J., 2015a. Visualization and prediction of porosity in roller compacted ribbons with near-infrared chemical imaging (NIR-CI). Journal of Pharmaceutical and Biomedical Analysis 109, 11-17.

Khorasani, M., Amigo, J.M., Sun, C.C., Bertelsen, P., Rantanen, J., 2015b. Near-infrared chemical imaging (NIR-CI) as a process monitoring solution for a production line of roll compaction and tableting. European Journal of Pharmaceutics and Biopharmaceutics 93, 293-302.

Kirsch, J.D., Drennen, J.K., 1999. Nondestructive tablet hardness testing by near-infrared spectroscopy: a new and robust spectral best-fit algorithm. J. Pharm. Biomed. Anal. 19, 351-362.

Kleinebudde, P., 2004. Roll compaction/dry granulation: pharmaceutical applications. European Journal of Pharmaceutics and Biopharmaceutics 58, 317-326.

Lim, H., Dave, V.S., Kidder, L., Neil Lewis, E., Fahmy, R., Hoag, S.W., 2011. Assessment of the critical factors affecting the porosity of roller compacted ribbons and the feasibility of using NIR chemical imaging to evaluate the porosity distribution. Int. J. Pharm. 410, 1-8.

Liu, H., Shah, S., Jiang, W., 2004. Online outlier detection and data cleaning. Comput. Chem. Eng. 28, 1635-1647.

McAuliffe, M., O'Mahony, G., Blackshields, C., Collins, J., Egan, D., Kiernan, L., O'Neill, E., Lenihan, S., Walker, G., Crean, A., 2014. The Use of PAT and Off-line Methods for Monitoring of Roller Compacted Ribbon and Granule Properties with a View to Continuous Processing. Organic Process Research & Development 19, 158-166.

Montgomery, D.C., 2001. Design and Analysis of Experiments, Fifth edition ed. John Wiley & Sons

Mullins, E., 2003. Statistics for the quality control chemistry laboratory. Royal Society of Chemistry.

Peter, S., Lammens, R.F., Steffens, K.-J., 2010. Roller compaction/Dry granulation: Use of the thin layer model for predicting densities and forces during roller compaction. Powder Technology 199, 165-175.

Pinheiro, J., Bates, D., 2006. Mixed-effects models in S and S-PLUS. Springer Science & Business Media.

Pocock, S.J., Geller, N.L., Tsiatis, A.A., 1987. The analysis of multiple endpoints in clinical trials. Biometrics, 487-498.

R Core Team., 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Scheipl, F., Greven, S., Küchenhoff, H., 2008. Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. Computational statistics & data analysis 52, 3283-3299.

Short, S.M., Cogdill, R.P., Wildfong, P.L.D., Drennen Iii, J.K., Anderson, C.A., 2009. A Near-Infrared Spectroscopic Investigation of Relative Density and Crushing Strength in Four-Component Compacts. Journal of Pharmaceutical Sciences 98, 1095-1109.

Sun, C.C., 2008. Mechanism of moisture induced variations in true density and compaction properties of microcrystalline cellulose. Int. J. Pharm. 346, 93-101.

Teng, Y., Qiu, Z., Wen, H., 2009. Systematical approach of formulation and process development using roller compaction. European Journal of Pharmaceutics and Biopharmaceutics 73, 219-229.

Figures

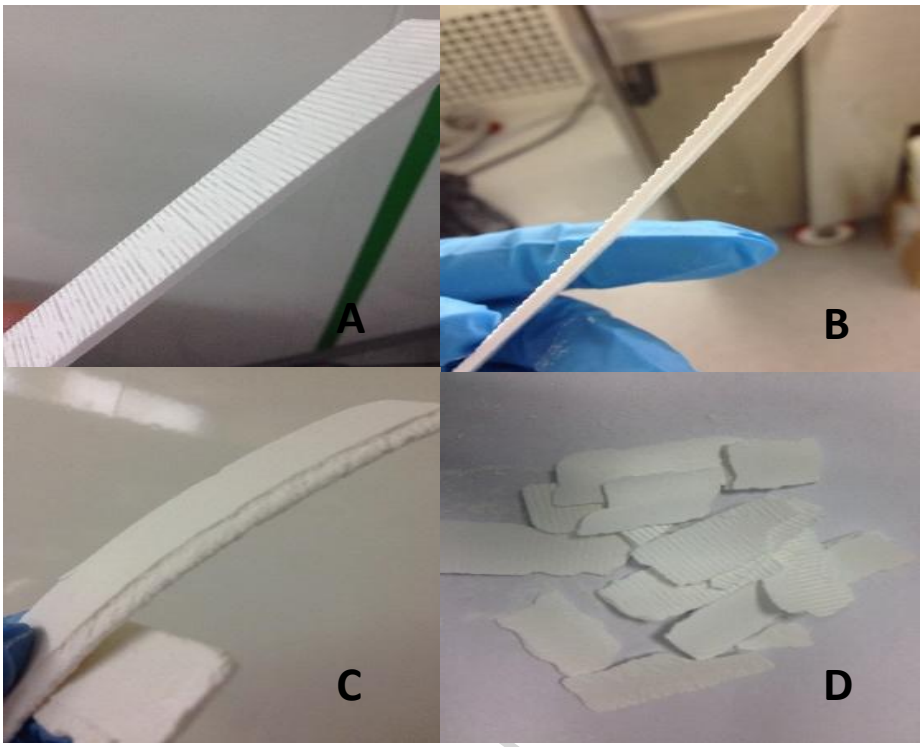


Figure 1. Photographs of examples of good quality flat ribbon samples (A and B), curved ribbons (C) and very poor split broken ribbons (D)

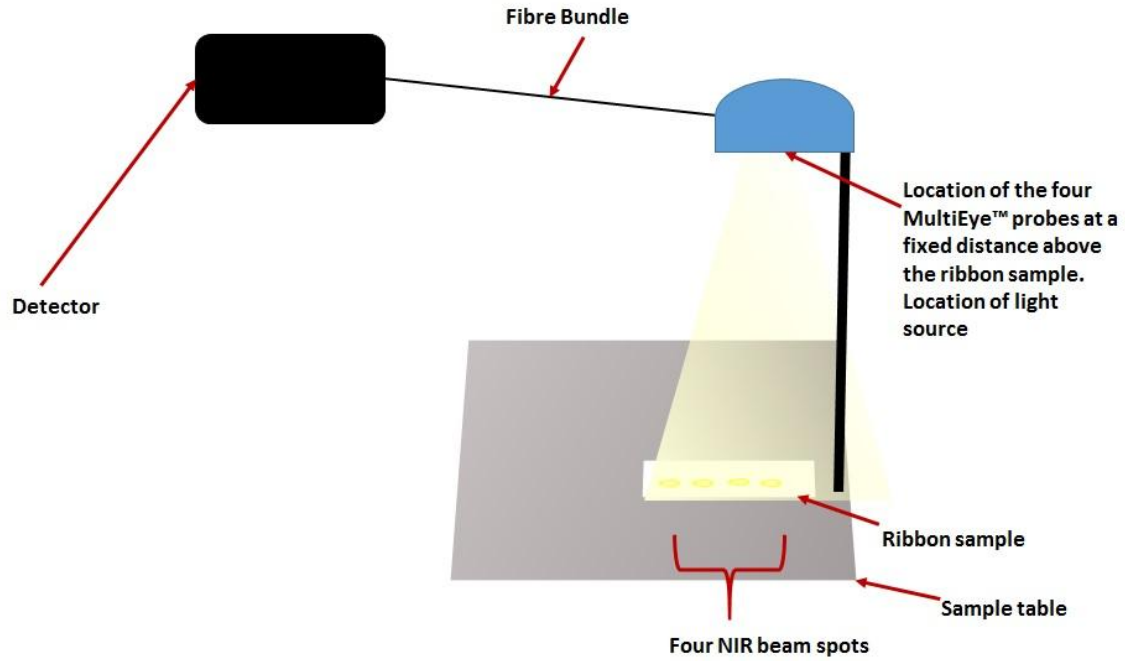


Figure 2. Schematic illustration of MultiEye™ NIR probe set up for at-line ribbon analysis. All four probes were positioned in series along the length of the ribbon at a fixed distance above the sample which was aligned in a fixed position on the sample table.

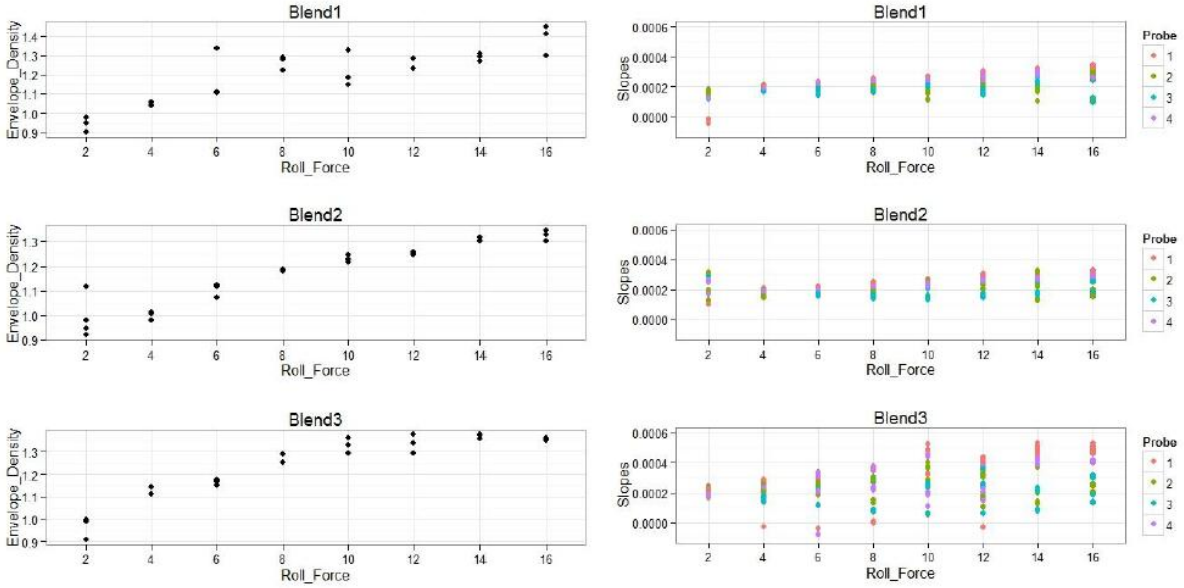


Figure 3. Plots showing the envelope density measurements of individual ribbon sample (left side pane) and spectral slopes for individual probes for each of five scans on three ribbon samples (right side panel) versus roll force. Blends 1-3 differed in the storage conditions of microcrystalline cellulose (MCC) prior to blending with lactose and magnesium stearate; Blend 1 (ambient (41% RH/20^{o}C)), Blend 2 (low humidity (11% RH/20^{o}C)), and Blend 3 (high humidity (75% RH/40^{o}C)).

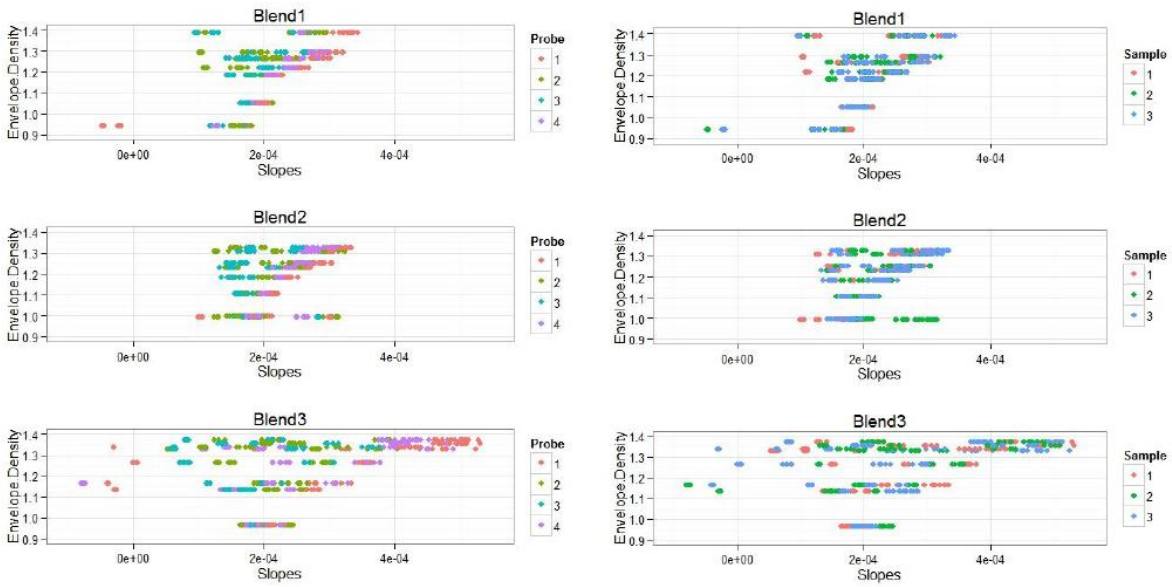


Figure 4. Plots of ribbon envelope density versus NIR spectral slopes. Envelope densities are the average of three ribbon samples produced at each Roll Force/Blend combination. Spectral slopes are determined from individual NIR spectral results for 3 ribbon samples, each scanned 5 times by 4 probes. Left side panel shows spectral slopes colour coded according to the probe number that acquired the NIR spectra and the right side panel shows spectral slopes colour coded according to ribbon sample number. Blends 1-3 differed in the storage conditions of microcrystalline cellulose (MCC) prior to blending with lactose and magnesium stearate; Blend 1 (ambient (41% RH/20^{o}C)), Blend 2 (low humidity (11% RH/20^{o}C)), and Blend 3 (high humidity (75% RH/40^{o}C)).

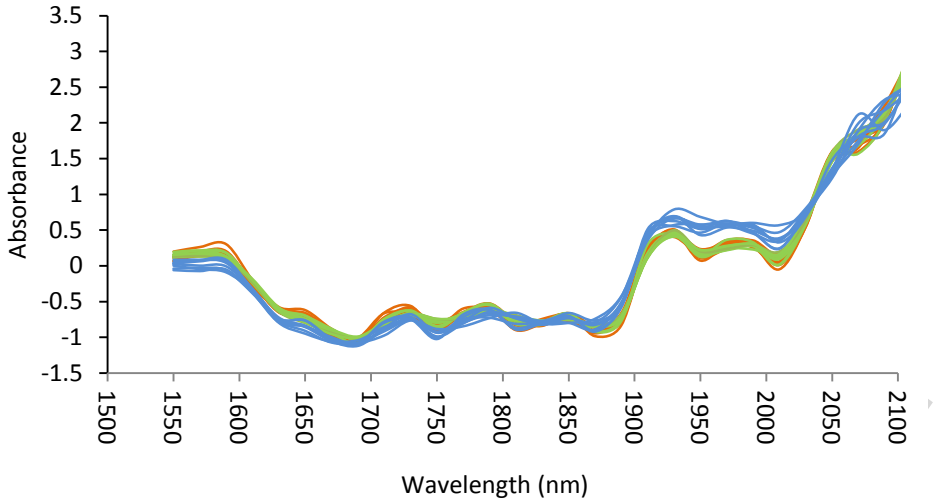


Figure 5. Spectral averages (n=60) of at-line acquired NIR spectra of ribbon samples produced at each of the 24 Roll force/Blend combinations. Blend 1 (orange), blend 2 (green) and blend 3 (blue)**.** Blends 1-3 differed in the storage conditions of microcrystalline cellulose (MCC) prior to blending with lactose and magnesium stearate; Blend 1 (ambient (41% RH/20^{o}C)), Blend 2 (low humidity (11% RH/20^{o}C)), and Blend 3 (high humidity (75% RH/40^{o}C)).

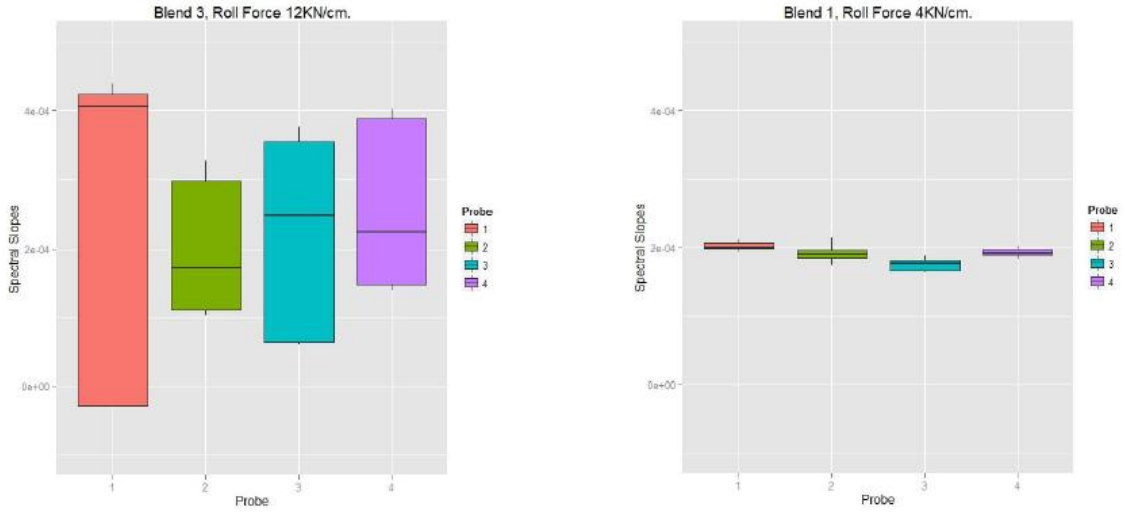


Figure 6. Boxplots of the spectral slopes for the four probes for ribbon samples produced from blend 3 (containing MCC stored at 75% RH/40^{o}C prior to blending) and compacted at a roll force 12 kN/cm (left side panel) and spectral slopes for ribbon samples produced from blend 1 (containing MCC stored at 41% RH/20^{o}C prior to blending) and compacted at a roll force 4 kN/cm.

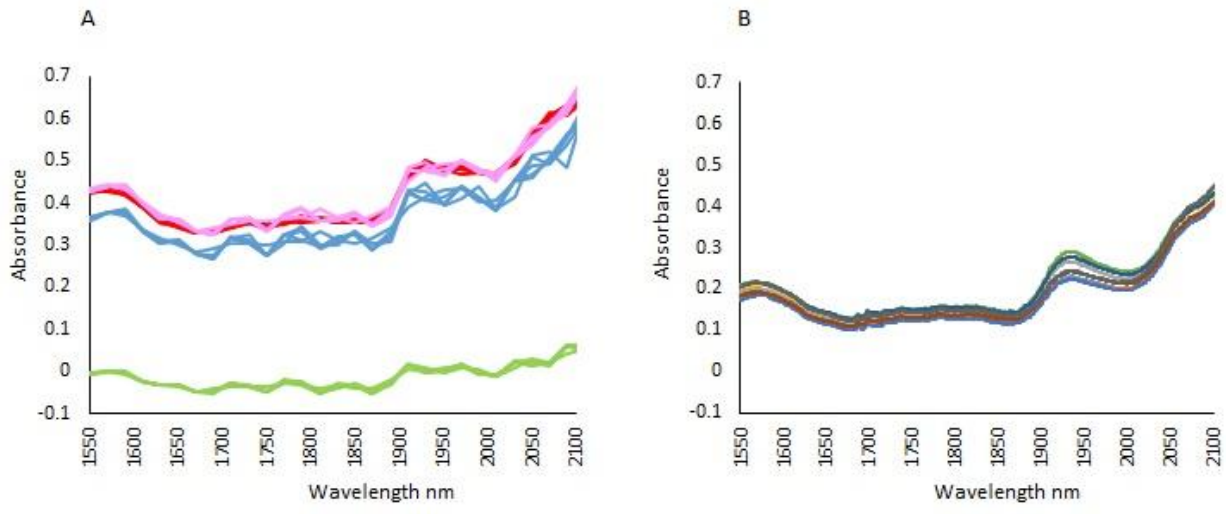


Figure 7 NIR spectral results obtained for one ribbon sample produced from Blend 3 roll force 12kN/cm as scanned by the 4 different probes. Probe 1 (red), Probe 2 (green), Probe 3 (blue) and Probe 4 (pink). Left panel – A: At-line MultiEye™ spectra illustrating the effect of sample presentation issues encountered during spectral data collection due to the production of imperfect ribbon as seen in Figure 5(c-d). Left panel – B NIR spectra as measured by Perkin Elmer off-line, sample presentation off line probe was not an issue.

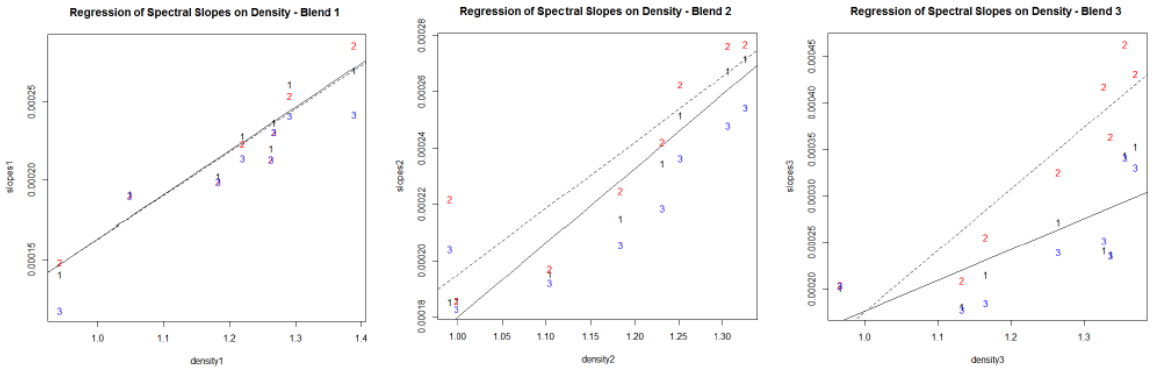


Figure 8. Comparison of the linear regression plots the ribbon density for each of the 3 blends using the NIR spectral slopes data. (1) the 33% trimmed mean (Trim Mean), (2) the average after visual discard (Visual Discard) and (3) the average of all 60 spectral slopes acquired at-line (Full Mean). Solid line shows fitted line for Trim Mean samples, dotted line is for Visual Discard samples.

Table 1. Details of microcrystalline cellulose storage conditions prior to blending and average moisture content following storage determined by thermogravimetric analysis (n=5).

Blend	Description	% Relative Humidity	Temperature	Moisture Content (% w/w) +/- standard deviation
1	Ambient %RH	41	20 ± 1°C	4.86 ± 0.06
2	Low %RH	11	20 ± 1°C	4.12 ± 0.34
3	High % RH	75	40 ± 1°C	6.96 ± 0.33

Table 2. Comparison of Null and alternate models (equations 2 and 3) for the 96 roll force/blend/ probe combinations using the **exactRLRT** command in the RLRSim package in R which tests the (exact) restricted likelihood ratios for the 2 models. The structural difference between these two models is that the alternative model contains a variance term allowing for random differences between samples. The null model does not allow for this source of variation.

	Probe 1		Probe 2		Probe 3		Probe 4	
Blend (Roll Force)	Sample SD/ Residual SD	RLRT p-value	Sample SD/ Residual SD	RLRT p-value	Sample SD/ Residual SD	RLRT p-value	Sample SD/ Residual SD	RLRT p-value
Blend 1 (2 kN/cm)	56.22	≤0.001	1.29	≤0.01	2.99	≤0.001	4.37	≤0.001
Blend 1 (4 kN/cm)	0.86	≤0.05	1.93	≤0.001	1.21	≤0.01	0.85	≤0.05
Blend 1 (6 kN/cm)	0.00	1.00	1.8	≤0.001	2.91	≤0.001	0.86	≤0.05
Blend 1 (8 kN/cm)	1.02	≤0.05	3.63	≤0.001	1.75	≤0.001	0.00	≤0.001
Blend 1 (10 kN/cm)	1.74	≤0.001	5.69	≤0.001	1.75	≤0.001	0.05	0.39
Blend 1 (12 kN/cm)	0.70	0.06	1.17	≤0.01	1.51	≤0.01	1.33	≤0.01
Blend 1 (14 kN/cm)	3.58	≤0.001	9.98	≤0.001	7.86	≤0.001	3.29	≤0.001
Blend 1 (16 kN/cm)	2.77	≤0.001	10.97	≤0.001	9.25	≤0.001	1.62	≤0.01
Blend 2 (2 kN/cm)	37.13	≤0.001	8.63	≤0.001	6.12	≤0.001	7.31	≤0.001
Blend 2 (4 kN/cm)	1.25	≤0.01	3.99	≤0.001	1.49	≤0.01	0.83	0.03
Blend 2 (6 kN/cm)	0.72	0.06	1.60	≤0.001	0.29	0.29	0.00	1.00
Blend 2 (8 kN/cm)	0.52	0.14	5.44	≤0.001	3.08	≤0.001	0.45	0.18
Blend 2 (10 kN/cm)	1.17	≤0.01	6.88	≤0.001	5.12	≤0.001	0.73	0.06
Blend 2 (12 kN/cm)	0.00	0.40	3.18	≤0.001	1.8	≤0.001	1.64	≤0.001
Blend 2 (14 kN/cm)	0.96	≤0.05	8.46	≤0.001	4.69	≤0.001	0.32	0.26
Blend 2 (16 kN/cm)	1.84	≤0.001	8.38	≤0.001	7.43	≤0.001	2.12	≤0.001
Blend 3 (2 kN/cm)	4.75	≤0.001	6.45	≤0.001	3.92	≤0.001	5.52	≤0.001
Blend 3 (4 kN/cm)	29.00	≤0.001	8.43	≤0.001	3.46	≤0.001	11.27	≤0.001
Blend 3 (6 kN/cm)	32.56	≤0.001	5.42	≤0.001	8.74	≤0.001	27.42	≤0.001
Blend 3 (8 kN/cm)	56.16	≤0.001	15.14	≤0.001	23.39	≤0.001	7.92	≤0.001
Blend 3 (10 kN/cm)	10.07	≤0.001	16.08	≤0.001	16.23	≤0.001	15.6	≤0.001
Blend 3 (12 kN/cm)	22.52	≤0.001	10.46	≤0.001	17.31	≤0.001	13.33	≤0.001
Blend 3 (14 kN/cm)	1.44	≤0.01	13.87	≤0.001	12.44	≤0.001	2.14	≤0.001
Blend 3 (16 kN/cm)	2.23	≤0.001	3.37	≤0.001	14.42	≤0.001	4.11	≤0.001

Table 3. The estimated values of Model 2 (equation 5) for the random effects (samples and sample-probe interaction) for two representative ribbon types.

Source of variation	**Blend 3 Roll force 12 kN**	**Blend 1 Roll force 4 kN**
Sample:Probe	167×10^{-10}	2.37×10^{-11}
Sample	28×10^{-10}	1.21×10^{-11}
Residual	1×10^{-10}	2.77×10^{-11}
Total	196×10^{-10}	6.35×10^{-11}

Table 4. The estimated values for the fixed effects (probes) of Model 2 (equation 5) for two representative ribbon types. Probe 1 is set as the intercept.

	Blend 3 Roll force 12 kN	**Blend 1 Roll force 4 kN**
Intercept	26.9×10^{-5}	202.0×10^{-6}
Probe 2	-7.29×10^{-5}	-10.5×10^{-6}
Probe 3	-4.41×10^{-5}	-26.8×10^{-6}
Probe 4	-1.70×10^{-5}	-9.53×10^{-6}

Table 5. Confidence Limits x (10^3) on linear contrasts for all ribbons produced from each of the 24 Roll Force/Blend combinations. A linear contrasts method was used to investigate the magnitude of any difference in variance between the end probes (probes 1 and 4) and the central probes (probes 2 and 3).

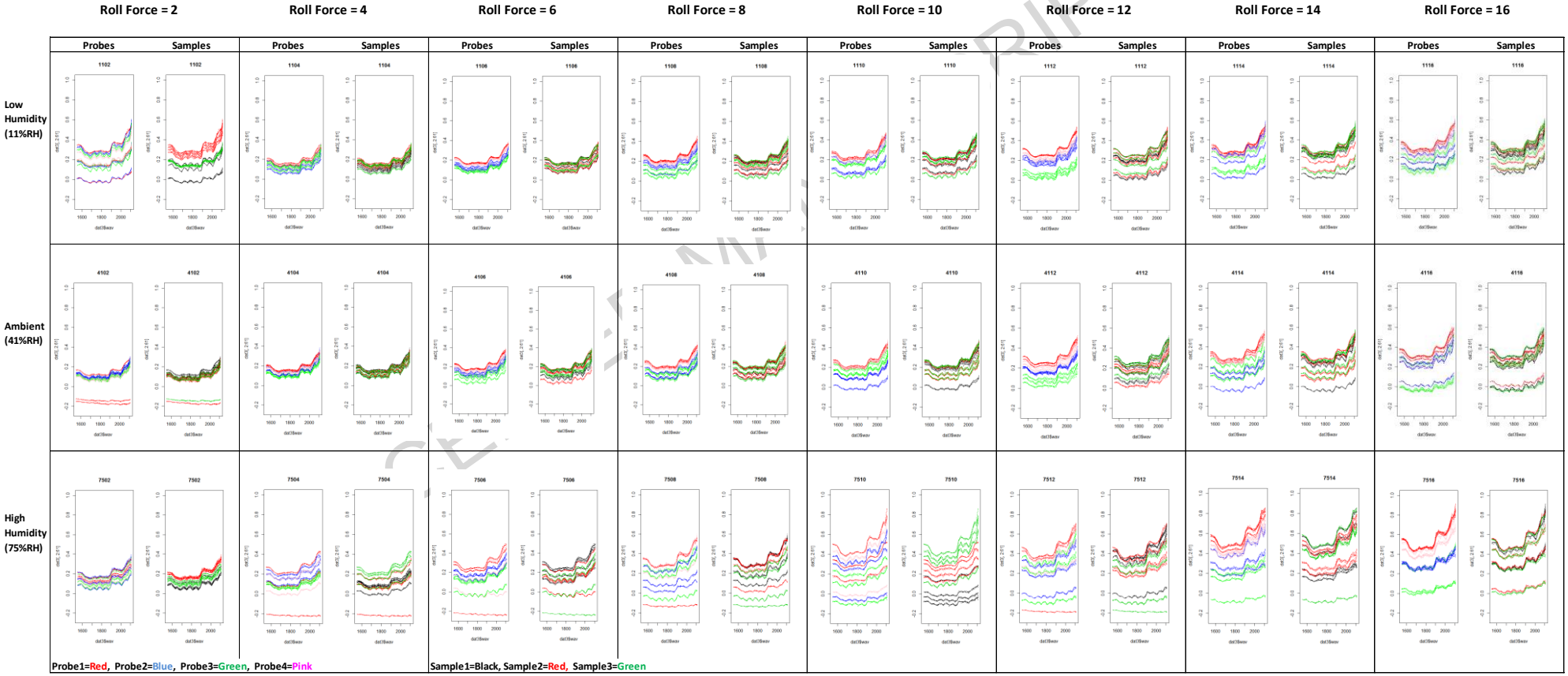
Ribbon Blend 1	Confidence Limits ('000s)	Ribbon Blend 2	Confidence Limits ('000s)	Ribbon Blend 3	Confidence Limits ('000s)
2 kN	(-0.656,-0.595)	2 kN	(-0.098,-0.018)	2 kN	(0.040, 0.096)
4 kN	(0.112, 0.166)	4 kN	(0.104, 0.179)	4 kN	(-0.390,-0.341)
6 kN	(0.333, 0.399)	6 kN	(0.294, 0.363)	6 kN	(-0.392,-0.324)
8 kN	(0.499, 0.574)	8 kN	(0.522, 0.590)	8 kN	(0.694, 0.760)
10 kN	(0.668, 0.735)	10 kN	(0.575, 0.656)	10 kN	(0.767, 0.881)
12 kN	(0.764, 0.857)	12 kN	(0.714, 0.798)	12 kN	(0.448, 0.551)
14 kN	(0.997, 1.043)	14 kN	(0.721, 0.814)	14 kN	(2.409, 2.513)
16 kN	(1.047, 1.133)	16 kN	(0.846, 0.920)	16 kN	(2.358, 2.444)

Table 6. Comparison of the linear regression models to predict the ribbon density for each of the 3 blends using the NIR spectral slopes data. (i) the average of all 60 spectral slopes acquired at-line (Full Mean) (ii) the average after visual discard (Visual Discard), (iii) the 33% trimmed mean (Trim Mean), and (iv) off-line spectral data, as the dependent variables.

Data discard method	**Blend**	**Linear regression models**	**r**
MultiEye™ at-line Full data set (no spectra discarded)	1	Env Density = 0.420 + 3788 x Average Spectral Slope	0.93
	2	Env Density = -0.041 + 5587 x Average Spectral Slope	0.89
	3	Env Density = 0.474 + 3115 x Average Spectral Slope	0.74
Visual Discard	1	Env Density = 0.409 + 3628 x Average Spectral Slope	0.95
	2	Env Density = 0.170 + 4257 x Average Spectral Slope	0.88
	3	Env Density = 0.734 + 1515 x Average Spectral Slope	0.92
Trimmed Mean	1	Env Density = 0.416 + 3584 x Average Spectral Slope	0.96
	2	Env Density = 0.319 + 3788 x Average Spectral Slope	0.97
	3	Env Density = 0.465 + 3021 x Average Spectral Slope	0.74
Spotlight 400 FT-IR off-line (no spectra discarded)	1	Env Density = 0.246 + 3212 x Average Spectral Slope	0.93
	2	Env Density= 0.164 + 3392 x Average Spectral Slope	0.79
	3	Env Density = 0.374 + 2906 x Average Spectral Slope	0.74

Appendix 1 –NIR spectra data set

The dataset consists of 60 NIR spectra for ribbons produced at 24 different Roll Force/Blend combinations. For each of these Roll Force/Blend combinations, a sample of 3 ribbon sections were selected and scanned 5 times under MultiEye™ (NIR Spectrometer) using 4 probes at 29 wavelengths (1550−2110 nm in steps of 20 nm). The 3 blends differed only in the storage conditions of microcrystalline cellulose (MCC) prior to blending (low (11% RH/20°C), ambient (41% RH/20°C), and high (75% RH/40°C)) relative humidity. Blends were compacted at 8 roll force settings which increased in intervals of 2 kN/cm between 2 – 16 kN/cm.
Left panel shows spectra by acquisition probes: Probe 1=Red, Probe 2=Green, Probe 3=Blue, Probe 4=Pink.
Right panels shows spectra by ribbon sample: Sample 1 = Black, Sample 2= Red, Sample 3= Green



Appendix 2 Comparison of Model 1 (equation 4) and Model 2 (equation 5) using the likelihood ratio test.

Blend	Roll Force kN/cm		AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
1	2	Model 1	-1030	-1017	520.8	-1042			
		Model 2	-1192	-1177	602.8	-1206	164.1	1	≤0.001
1	4	Model 1	-1239	-1226	625.5	-1251			
		Model 2	-1251	-1237	632.7	-1265	14.46	1	≤0.001
1	6	Model 1	-1184	-1171	597.9	-1196			
		Model 2	-1222	-1207	617.8	-1236	39.63	1	≤0.001
1	8	Model 1	-1156	-1143	584	-1168			
		Model 2	-1203	-1188	608.6	-1217	49.13	1	≤0.001
1	10	Model 1	-1154	-1142	583.1	-1166			
		Model 2	-1212	-1197	613	-1226	59.79	1	≤0.001
1	12	Model 1	-1175	-1163	593.7	-1187			
		Model 2	-1190	-1175	602	-1204	16.54	1	≤0.001
1	14	Model 1	-1078	-1065	544.9	-1090			
		Model 2	-1230	-1216	622.2	-1244	154.7	1	≤0.001
1	16	Model 1	-997.1	-984.6	504.6	-1009			
		Model 2	-1154.8	-1140.2	584.4	-1169	159.7	1	≤0.001
2	2	Model 1	-1073	-1060	542.5	-1085			
		Model 2	-1169	-1155	591.6	-1183	98.24	1	≤0.001
2	4	Model 1	-1182	-1169	596.8	-1194			
		Model 2	-1206	-1191	609.9	-1220	26.05	1	≤0.001
2	6	Model 1	-1225	-1213	618.6	-1237			
		Model 2	-1231	-1216	622.4	-1245	7.693	1	≤0.01
2	8	Model 1	-1135	-1122	573.4	-1147			
		Model 2	-1206	-1192	610.2	-1220	73.47	1	≤0.001
2	10	Model 1	-1067	-1054	539.5	-1079			
		Model 2	-1175	-1160	594.4	-1189	109.7	1	≤0.001
2	12	Model 1	-1144	-1131	577.8	-1156			
		Model 2	-1189	-1174	601.3	-1203	47.04	1	≤0.001
2	14	Model 1	-1030	-1018	521.2	-1042			
		Model 2	-1154	-1140	584.2	-1168	126	1	≤0.001
2	16	Model 1	-1041	-1029	526.6	-1053			
		Model 2	-1178	-1164	596.2	-1192	139.3	1	≤0.001
3	2	Model 1	-1216	-1204	614.2	-1228			
		Model 2	-1235	-1220	624.4	-1249	20.52	1	≤0.001
3	4	Model 1	-993.5	-981	502.8	-1006			
		Model 2	-1205.1	-1190	609.6	-1219	213.6	1	≤0.001
3	6	Model 1	-913.5	-900.9	462.7	-925.5			
		Model 2	-1159.8	-1145.1	586.9	-1173.8	248.3	1	≤0.001
3	8	Model 1	-929.3	-916.8	470.7	-941.3			
		Model 2	-1164.8	-1150.1	589.4	-1178.8	237.4	1	≤0.001
3	10	Model 1	-1010	-997.4	511	-1022			
		Model 2	-1121	-1106.6	567.6	-1135	113.3	1	≤0.001
3	12	Model 1	-898.7	-886.2	455.4	-910.7			
		Model 2	-1115.1	-1100.4	564.5	-1129.1	218.4	1	≤0.001
3	14	Model 1	-977.5	-964.9	494.7	-989.5			
		Model 2	-1133.9	-1119.3	574	-1147.9	158.4	1	≤0.001
3	16	Model 1	-1064	-1052	538.3	-1076			
		Model 2	-1165	-1150	589.4	-1179	102.2	1	≤0.001

Graphical abstract

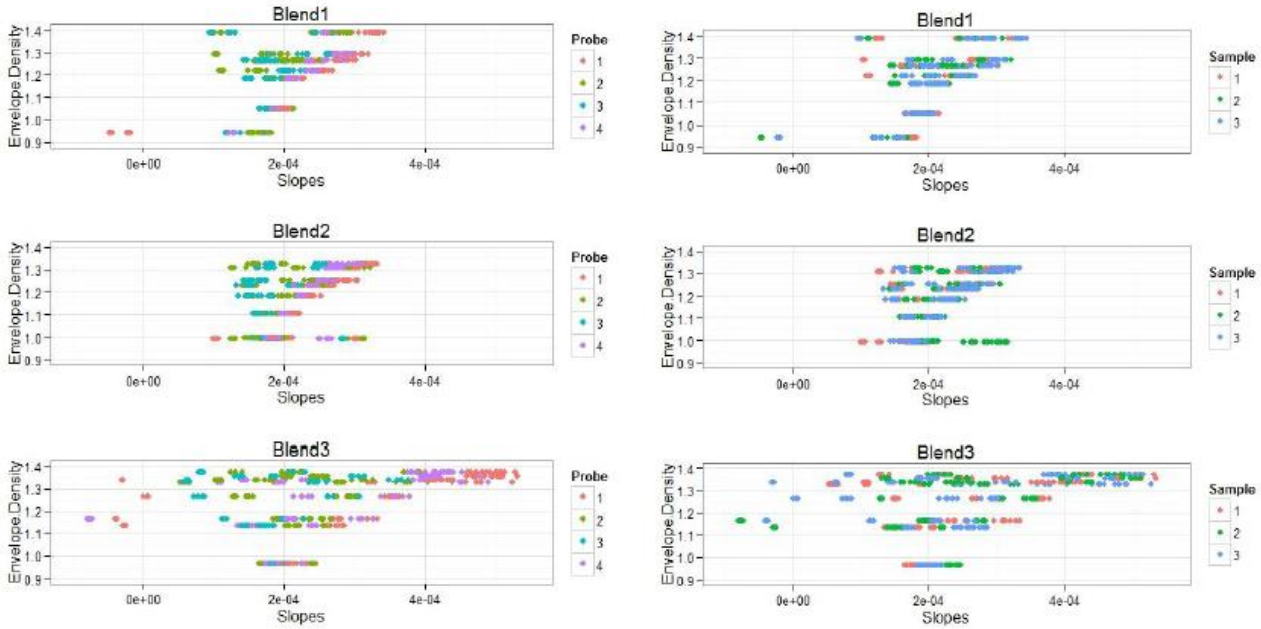


Figure 3. Plots of ribbon envelope density measurement versus NIR spectral slopes. Envelope densities are average of three ribbon samples produced at each Roll Force/Blend combination. Spectral slopes are determined from individual NIR spectral results for 3 ribbon samples, each scanned 5 times by 4 probes. Left side panel shows spectral slopes colour coded according to the probe number that acquired the NIR spectra and the right side panel shows spectral slopes colour coded according to ribbon sample number. Blends 1-3 differed in the storage conditions of microcrystalline cellulose (MCC) prior to blending with lactose and magnesium stearate; Blend 1 (ambient (41% RH/20^{o}C)), Blend 2 (low humidity (11% RH/20^{o}C)), and Blend 3 (high humidity (75% RH/40^{o}C))