2015

Multipoint Spectroscopy and Stereoscopic Imaging of Pharmaceutical Particles

Hicham Rifai
Technological University Dublin

Institiúid Teicneolaíochta Átha Cliath
Dublin Institute of Technology

MULTIPOINT SPECTROSCOPY AND STEREOSCOPIC IMAGING OF PHARMACEUTICAL PARTICLES

Hicham Rifai
(MSc.)

Thesis submitted to Dublin Institute of Technology, Ireland, in fulfilment of the requirement for the award of the degree of

MASTER OF PHILOSOPHY

School of Food Science and Environmental Health
College of Science
Dublin Institute of Technology
Ireland

Under the supervision of

Dr. Patrick J. Cullen
Dr. Carl Sullivan

2015

Abstract

Particle and granule properties play a key role in the final product quality of pharmaceuticals. Thus the identification and monitoring of key chemical and physical parameters is essential in the production of pharmaceuticals. The existing off-line methods are generally slow and labour intensive. Near infra-red (NIR) multipoint spectroscopy and image analysis are an attractive alternative compared to the traditional methods because they are both non-destructive and non-interfering allowing the analysis in real time of particles physical and chemical properties.

This research is a preliminary study performed at laboratory scale and aims at developing chemometric and imaging algorithms for real time measuring of pharmaceutical chemical and physical properties. These algorithms utilised real time NIR multipoint spectroscopy and a novel imaging system. NIR multipoint spectroscopy followed by a regression technique (such as PLS) was used to build calibration models to quantify a compound in a small size binary granule mixture under both static and dynamic conditions.

The imaging technology provided key physical properties such as size, shape and texture. The Haralick correlation property and the variogram were used to analyse the surface texture of particles. These algorithms allowed the classification of particles by their morphological nature under both static and dynamic conditions.

Declaration

I certify that this thesis which I now submit for the examination for the award of Master of Philosophy (MPhil) is entirely my own work and has not been taken from the work of others, save to the extent that such work has been cited and acknowledged within the text of my work. I also certify that I have adhered to all principles of academics honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in any submission.

This thesis was prepared according to the regulations for postgraduate study by research of Dublin Institute of Technology and has not been submitted in whole or in part for another award in any institute.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for ethics in research.

The Institute has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material or the thesis is duly acknowledged.

Hicham Rifai **Date: 23 June 2015**

"A picture is worth a thousand words"

"tá pictiúr fiú míle focal"

Acknowledgements

I never imagine that I would be so lucky during my MPhil degree to be in such inspiring, supportive, open and friendly environment. I did enjoy every day I have been in the department of food and health science at Dublin Institute of Technology, it has been such a valuable experience in every aspect. I am grateful to my supervisors Dr PJ Cullen and Dr Carl Sullivan for their patience, caring, continuous support and encouragements. I had ideal supervisors since day one Dr PJ Cullen and Dr Carl shared their time and knowledge so that I got a good start. I would also like to thank Dr Denisio Togashi for his encouragement, caring suggestions, and his support for the experimental design. Special thanks to Dr Laura Alvarez for her encouragement and for teaching me how to design experiments. I would like to thank Dr James Curtin for his words of comfort during several occasions. I would like to thank Dr Jesus Frias for the inspiring discussions. I thank Raquel Cama for sharing with me her knowledge about chemometrics. I thank all the researchers, postgraduate student and the lecturers that made the atmosphere at Sackville place open, warm and friendly. I also thank Innopharma Labs for giving me an overview about pharmaceutical manufacturing.

A final thanks to my parents, Mr Rifai Driss and Mrs Fatima Kabbour I dedicate this thesis to you both.

Abbreviations list

APIs	Active Pharmaceutical Ingredients
CI	Chemical Imaging
CQA	Critical Quality Attributes
EMA	European Medicines Agency
FDA	Food and Drug Administration
GLCM	Grey Level Co-Occurrence Matrix
HPLC	High Performance Liquid Chromatography Analysis
LDA	Linear Discriminate Analysis
LNO	Leave-N-Out
LOO	Leave-One-Out
MLR	Multiple Linear Regression
MSC	Multiplicative Scatter Correction
NIR	Near-Infra-Red
NIRS	Near-Infra-Red Spectroscopy
PAT	Process Analytical Technology
PCA	Principle Components Analysis
PLS	Partial Least Square
PLS-LDA	Partial Least Squares Discriminant Analysis
QbD	Quality by Design
RGB	Red, Green, Blue
RMSEP	Root mean squared error of prediction
SEM	Scanning Electron Microscopy
SNV	Standard Normal Variate

Table of Contents

List of Figures

List of Tables

1 Introduction

The pharmaceutical and chemical industry is a vital sector in the Irish economy; it represents 60% of Ireland's manufacturing exports together with the chemical industry[1]. There are over 100 companies specialised in pharmaceutical and chemical production, among them are major players such as Pfizer, GlaxoSmithKline and Merck & Co. Inc [2]. The most common product in pharmaceuticals is solid dosage forms (tablet or capsule) with solid forms accounting for more than 70% of the total pharmaceutical production in the world [3].

Pharmaceutical solid forms can be amorphous particles or crystals with various physical characteristics. They are a mixture of one or more active pharmaceutical ingredients (APIs) and non-active ingredients such as diluents, disintegrants, fillers, binders, lubricants, or other ingredients. Physical and chemical properties of particles and granules can affect the overall performance of the powdered mixture at different stages of pharmaceutical manufacturing. Due to the high degree of regulation imposed on pharmaceutical manufacturing, the overall manufacturing process is divided into processes. Traditionally, quality assessment tests are performed on prepared samples after each manufacturing process in addition to quality tests on samples of the final product.

The United States food and drug administration (FDA) published a guidance document that outlines the use of Process Analytical Technology (PAT) [4]. Unlike traditional methods of manufacturing, PAT achieves quality by design (QbD) instead of accessing the quality at the end-process. This is performed through using in-line measurement tools of critical quality parameters. In-line measurement methods should be fast, non-destructive and require no sample preparation.

For real time measurement of chemical properties, NIR spectroscopy is used extensively in pharmaceutical manufacturing, while near infra-red chemical imaging (NIR-CI) is considered to be a novel method. NIR spectroscopy calculates an average spectrum from a sample area while NIR-CI associates the position of the pixel with its corresponding spectrum producing a hyperspectral data cube. NIR-CI has been used in estimating the uniformity of content during blending and granulation processes [5].

Particle size and shape are critical quality parameters for several processes in pharmaceutical manufacturing like pre-mixing/mixing, granulation, drying, milling, roller compaction, spray-drying, coating and compression [6]. Particle size, shape and surface texture can also affect the flow properties of the powder, the dissolution rate and friability.

There are several real time measurement methods for particle size estimation. One of the most common methods is laser diffraction. Laser diffraction is based on the theory that the particle size is related to the scattering angle of light by the particle. Spatial filtering velocimetry is an example of a chord length measurement tool, it calculates the chord length by making the particles flow between two sapphire windows and applying a laser beam from one window to measure the shadow length of the particle on the other window. Imaging analysis via a scanning electron microscopy (SEM) can also be used to measure the size of particles.

Concerning the measuring of surface morphology characteristics, a small number of studies to date have been dedicated to the analysis of surface morphology during pharmaceutical manufacturing. Laser profilometry and SEM were used to monitor tablet surface roughness during the tablet coating process however both techniques were not conducted in-line [7]. Image analysis using multivariate wavelet texture analysis was used to analyse the surface

texture of tablet cores, wet granules and controlled release tablets. Image analysis of surface texture of bulk powder was also performed to estimate particle size [8].

1.1 Aims and objectives

This work is part of the research project ParticlePro which aims at developing a hybrid real time measurement device capable of obtaining chemical and physical information of pharmaceutical particles. Ideally, the hybrid device should measure key chemical properties such as moisture content, content uniformity and key physical parameters such as size, shape and surface texture. This will allow the monitoring of several pharmaceutical manufacturing operation units such as granulation, drying and milling.

The aim of this work is to develop chemometric and image analysis algorithms in conjunction with a novel multipoint near infra-red spectrometer known as the MultiEye^{TM} and a novel imaging system known as the Eyecon^{TM}. The MultiEye^{TM} was used to measure the chemical information. Particles of different compounds were blended at different ratios before using the multipoint NIR to assess the content ratios. Multivariate models were built to predict chemical content. The predictive power of these models was assessed for static and moving particles.

The Eyecon^{TM} system was used to measure physical properties of particles. It measured the size, shape and surface morphology properties. An image analysis algorithm based on thresholding was developed to extract individual particles before performing size and texture analysis. The measurement was performed on static and moving particles. For the static testing, size analysis was performed on different size particles ranging from 100 to 1000 µm. For dynamic testing particles of three nominal size categories of size 500, 700 and 1000 µm

were fixed on a conveyor belt which was operated at variable speeds. Particle texture analysis was performed on particles of size 700 μm under both static and dynamic conditions.

1.2 Summary of thesis outline

Chapter 2: Literature review, giving a general background of different topics related to this work and also published information in the field. This chapter gives a general outlook of the pharmaceutical manufacturing process, process analytical technology and the theory behind several PAT tools. It gives a brief overview of the mathematical methods used in conjunction with these PAT methods.

Chapter 3: Materials and methods, this section contains a detailed description of the particles used in this study as well as the steps taken to prepare different samples. The section presents the equipment prototypes used in these experiments before illustrating the different experimental set up for both static and dynamic conditions. Finally, it enumerates the different steps of the image analysis algorithm used for particle size and texture analysis algorithms employed to analyse the surface of particles with preliminary results for surface texture analysis.

Chapter 4: Results and discussion, presents the outcomes of all experiments and discusses the different challenges associated with each measurement. First, the results of measuring chemical properties of particles in both static and dynamic conditions using the prototypes equipment are presented. Then the results of particles size analysis and texture analysis in static and dynamic conditions are presented.

Chapter 5: Conclusions, a summary of the novel devices in conjunction with the mathematical methods at predicting both chemical and physical properties and the challenges associated with static and dynamic monitoring of these particles.

Chapter 6: Further work, points out future work that could be explored as part of the research conducted in this thesis.

2 Literature review

2.1 Overview of the pharmaceutical manufacturing of solid forms

Pharmaceutical solid forms commonly consist of powder that contains one or more API and inactive particles known as excipients. Tablets are the most common solid form [9]. Tablet characteristics such as shape, size, weight, hardness, thickness and dissolution profile determine its type and the way they should be taken. To ensure the desirable tablet properties testing is performed after each manufacturing process step.

The process of making tablets starts with blending the raw materials that consist of APIs and non-active materials. Sometimes the raw materials are milled before the mixing step to prevent the formation of powder lumps. Depending on the desired powder flow property required, the blended mixture can be subject to wet or dry granulation.

In wet granulation the particles in the blend are bound to each other with the help of a binding solution forming a larger particle known as a granule. Drying is always performed at the end of the wet granulation step. Typical wet granulation tools are high shear mixers, fluid bed granulators and single or twin screw extruders. In dry granulation, unlike wet granulation the binding of particles is performed through physical compaction forces exerted by a set of counter rotating rolls leading to a compact ribbon. Milling is the next step after granulation and it reduces the size of the granule to the desirable one.

Once the desirable size of the granule is obtained, a conveyor hopper is used to feed a tablet compaction system with a powder mix. Typically the compaction system is composed of the feeder that inputs the powder into the tablet die which is then compressed using a punch. The compacted tablet is then discharged from the die. Then, depending on the type of tablet, the

tablet is coated either with sugar or a thin layer of polymer or gelatine or with enteric coating prior to the packaging step [10].

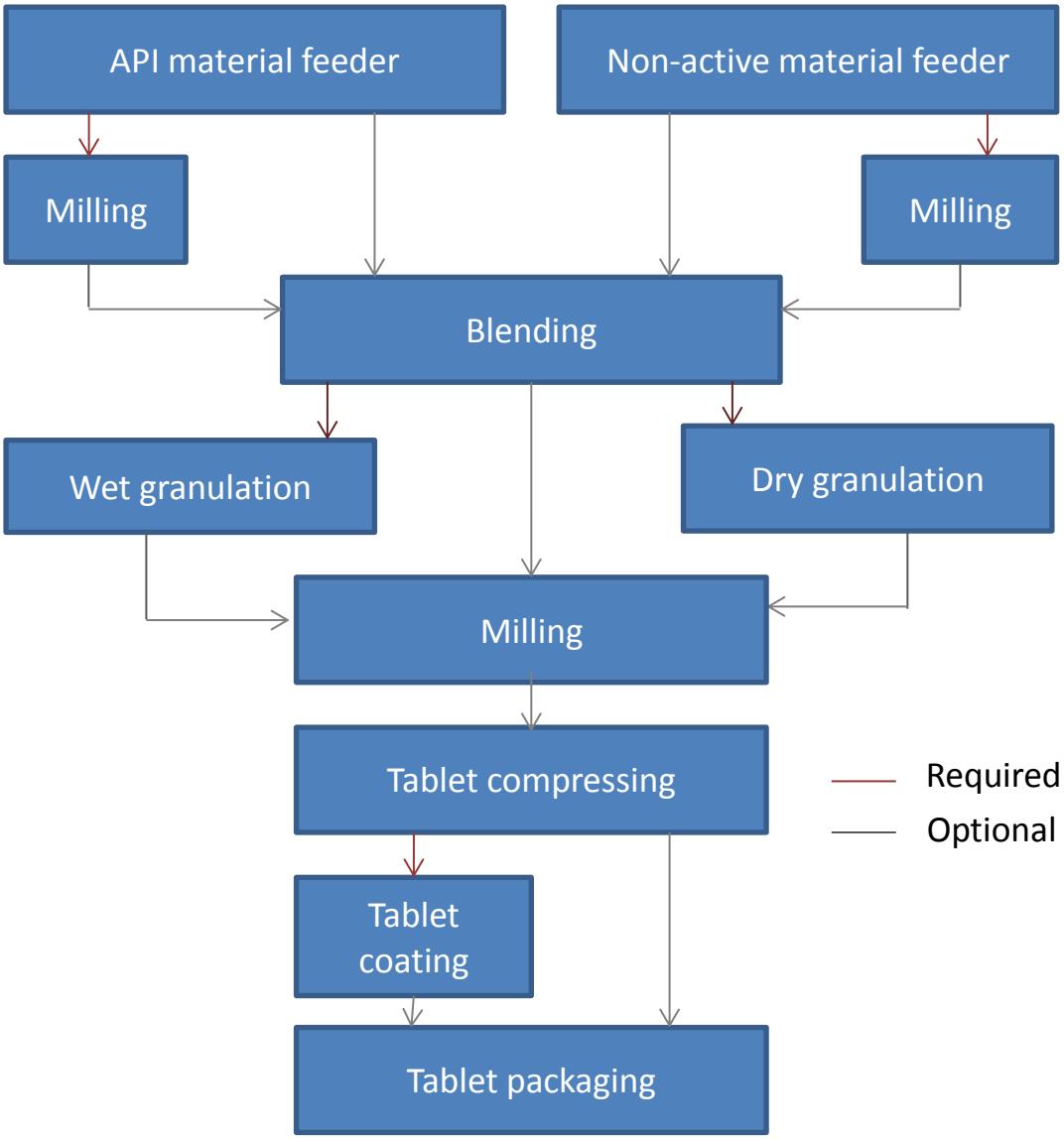


Figure 1 General manufacture procedure for making tablets.

2.1.1 Blending

Blending APIs and other inactive ingredients is an ubiquitous unit operation but at the same time it is a challenging step in the pharmaceutical manufacturing process. Powder blend uniformity refers to the active ingredient content percentage in the final blend, it is classified

as a critical process parameter (CPP). Clearly, the homogeneity of the mixed pharmaceutical powder can affect both the downstream unit operations and also the final product quality.

There are several blending tools used in pharmaceutical manufacturing for mixing different ingredients, the most commonly used are tumbling blenders [11]. A v-shaped blender is an example of tumbling blender where the mixture is continually separated and mixed while the blender is turning around its x-axis. In this type of blender the mixture is added either layer by layer to the blender or side by side. In general the mixing process is performed due to two main factors; the convective mechanism and diffusive mechanism. The convective mechanism is related to movement of groups of particles and thus is affected by the shape of the blender. While, the diffusive mechanism is related to the movement of particles with respect to each other thus it is related to the response of particle to the movement of the blender. The convective mechanism contributes to the blending at macro-level while the diffusive process contributes to the mixing at micro level [12].

The time and speed of the blending process is standardised but because of variations of ingredients' chemical and physical properties, sometimes using a set time may lead to undesirable powder homogeneity levels. Particularly, segregation of particles can occur due to the difference of mobility among particles. This difference in mobility is a direct result of the difference in particle characteristics namely; size, shape, density, and particle interactions. To determine the API content in the mix, the process goes to a standby state to wait for the analysis results which is performed by taking samples using a thief probe from the blender. The samples are analysed in quality control laboratories using in most cases high performance liquid chromatography analysis (HPLC). HPLC is quite slow in determining the content of API in the mix. In addition, the analysis is performed on a limited number of samples which represent a small portion of powder mix that may not reflect the overall homogeneity of the blend. HPLC is performed in laboratories that can be near or far from the

production site, adding more delay time to the overall manufacturing process [13]. Recently, several studies reported the use of NIR spectroscopy and chemical imaging to determine the content uniformity in powder mixes which is faster than HPLC and can be employed to obtain in-line measurement of API content [12-16].

2.2 Process analytical technology and quality by design

Pharmaceutical manufacturing suffers from poor understanding of the different unit manufacturing operations compared to other chemical industries [17]. Traditionally the manufacturing process is performed in a batch mode instead of a continuous mode. To determine the end points of the process, usually the process is halted and tests are performed to determine both if the process reaches the desirable end points and the quality of the blend. Measurement tests can take from a few minutes to days, and sometimes the sample is transported away from the manufacturing facility. In addition, each step is seen as black box with no record of how the powder properties change, so the ideal end points of the process might be reached but the process is not stopped immediately, this can affect the overall quality of tablets. Consequently many pharmaceutical companies spend more money on manufacturing itself than in research and development [17]. So the development of measurement tools that can be integrated into the production line, which enable real time information about the process itself, would reduce cost and time of manufacturing. It will also ease the regulation imposed by different regulation authorities [18]. Real time measurement is classified according to the FDA guideline [4] into three categories; **at-line** in which a sample is taken and analysed in the proximity of the production line without returning to process stream, **on-line** where a sample is deviated from the production stream analysed and can be returned to the process line, and finally **in-line** in which the sample measurement is performed without removing it from the process line. This is the aim of

process analytical technology (PAT) which has emerged from process analytical chemistry (PAC). PAC as the name suggests analyses the process itself using real time measurement tools and chemometrics, it was applied first in the petrochemical industry[19]. PAT was introduced first by FDA to enhance the understanding and control of the manufacturing process in the pharmaceutical and biopharmaceuticals industry which is a highly regulated manufacturing industry[20]. PAT includes a group of instruments that gives more information about the processes. The United States Food and Drug administration (FDA) published guidance for the use of PAT which divided the instruments for real time measurement and control into the following categories [18]:

1. Multivariate data acquisition and analysis tools;
2. Process analysers and sensors;
3. Process end points monitoring and control;
4. Tools for information management for ongoing improvement.

Quality by design (QbD) is often linked to PAT and it aims at using the information provided by instruments to numerate process operation pitfalls that can affect the final product quality in addition to quality risk management principles and quality control strategies to develop a systematic framework for different processes [21]. It is expected that QbD would lead to a better assurance of product quality. The main steps of QbD were described in guideline ICHQ8R2 which described six sequential steps[22]:

- Quality Target Product Profile (QTPP): This is the first step of QbD and it enumerate the desired quality criteria by considering as described in guide line ICHQ8R2 ; The desired product performance/target product profile, container closure system, therapeutic moiety release or delivery and attributes affecting pharmacokinetic

characteristics appropriate to the drug product dosage from being developed, Drug product quality criteria.

- Critical Quality Attributes (CQAs): Describe the permitted limits, ranges, distribution for the chemical , physical biological, microbiological properties or any other feature to ensure the final quality of the drug
- Risk Assessment: Linking material attributes and process parameters to Drug Product CQA and the risk assessment methods are described in ICH Q9 [22].
- Design Space: Describe the complex relationship between different processes and CQAs.
- Control Strategy: Method used to keep the parameters and processes with the permitted range, limits and distribution.
- Product life cycle management and continual improvement: The process can be improved by learning from the accumulated experience and adopting new technologies and methods while keeping the quality within safe zones imposed by the regulatory bodies.

Near infra-red spectroscopy and microscopic imaging are considered to be PAT tools and they may be part of the implementation of QbD principles. They have been used in several studies in conjunction with multivariate methods to monitor several pharmaceutical processes[6, 12, 23-28].

2.3 Near infra-red spectroscopy

Near infra-red spectroscopy (NIRS) is a robust non-intrusive high speed tool that often requires no sample preparation. It is used in a variety of industrial applications. The NIR region is between 780 and 2500 nm and it starts from the visible red light to the mid infrared

region [15]. It was discovered by William Hershel in 1800s while he was trying to establish a relation between heat radiation and the colour of light. Hershel concluded that the temperature increased near and beyond the red colour [29]. The first application of NIRS was in the agricultural field during the late 1960s, Karl Norris of the US department of agriculture conducted experiments using NIR on food. He studied blood in eggs, ripeness of melon, moisture in wheat and the hardness of wheat [29]. By the 1990s there was a major development in software and hardware, subsequently making NIRS an attractive measurement tool in many applications.

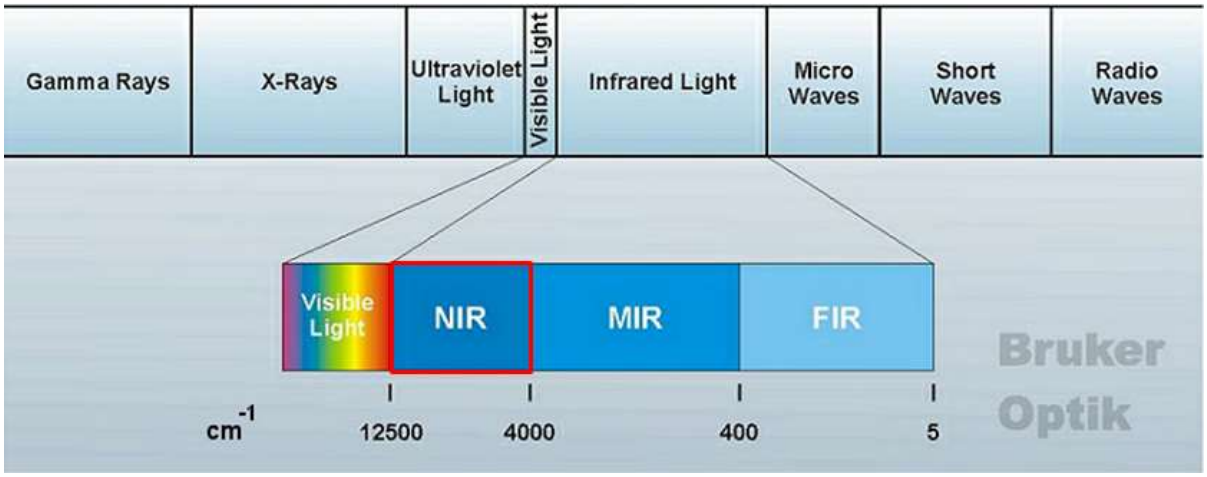


Figure 2 The electromagnetic spectrum [30].

The basic principle of NIR lies within quantum physics. As matter consists of molecules with different strength chemical bonds the energy required to move an electron in a bond to another energy level is proportional to the strength of the chemical bond. This principle is used by NIRS to estimate the chemical content and it works efficiently for samples having bonds C—H,N—H, or O—H [29]. In NIRS, the surface of an object is radiated with near infra-red radiation photons, these photons move the electrons in the atomic bonds to other levels of energy. Depending on the chemical bond strength and physical structure of the surface object, the photon will be absorbed or transmitted or reflected.

Typically a NIRS system is composed of a light source, which is tungsten-halogen in most cases because it is cheap and has a long life span, a monochromator which filters light with certain wavelengths, a sample holder and detectors for reflected or transmitted light [31].

2.3.1 NIRS for pharmaceutical raw material and solid form testing

NIRS has a wide range of applications within the pharmaceutical industry. . NIRS has been used for raw material testing, product quality control and process monitoring. The pharmaceutical industry is highly regulated and requires quality testing during different manufacturing processes starting from raw material to final product. In fact NIRS appears in a number of pharmaceutical guidelines [31]:

- European Pharmacopoeia 5 (2005).
- EMEA – CPMP/QWP/3309/01 and EMEA/CVMP/961/01 Note for Guidance on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations, EMEA, London, 2003.

- USP 29 (2006) page 2979.
- EMEA- CHMP/CVMP/QWP/17760/2009 Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations, EMEA, London 2014.

Good manufacturing process (GMP) guidelines, Chapter 5 state that a pharmaceutical company should “provide suitable procedures or measures to guarantee the identification of the [raw] material contained in each recipient”[32]. Following the later guideline pharmaceutical companies must identify and perform quality tests on the raw material, this is called compliance testing. Laboratory testing of raw materials can be tedious and time

consuming, so many pharmaceutical companies adopted a smart system based on NIRS which is fast and non-destructive.

NIRS can also be used to assert the quality of tablets. It can be used to monitor hardness, content uniformity, coating thickness and dissolution rate. NIRS has been used to detect counterfeit products. In capsules, NIRS can be used to assert the quality of the capsule by estimating the moisture content in an empty shell [14].

2.3.2 Real time monitoring of pharmaceutical processes using NIRS

NIRS can be used as an in-line, at-line and on-line measurement tool since it is non-intrusive and may require no sample preparation. Studies investigated the use of NIRS for in-line and on-line testing of powder mix homogeneity [15]. It has been established that using NIRS in monitoring powder blending is challenging but it is possible and gives a better understanding of the process.

Traditionally, to stop the end point of the drying process, temperature measurement and off-line analysis of moisture content are used. Water is a good absorbent for NIR radiation because of the O—H bond. So NIRS can be used to determine moisture content and the end point of the drying process. In addition, NIRS can give more knowledge about the drying process itself like the desorption rate.

Concerning wet granulation, NIRS can measure the moisture content and particle size [31]. Thus, it can determine the end process point of wet granulation that involves wetting, granule formation and drying. Studies reported that on-line and in-line use of NIRS were successful in determining the moisture content and particle size but required additional calculation for monitoring particle size growth due to the fact that particle growth and wetting affect the

reflection characteristics in a complex manner [15, 31]. Studies showed that NIRS went beyond determining the end point of wet granulation by providing more understanding about the granulation process such as determination of pseudopolymorphic transitions [31].

NIRS can be used at-line or on-line for determining the end point for pelletization and film coating. NIRS has also a potential to monitor the tableting and capsule filling process. It can be used as a quality test before packaging. Recently, a NIRS based system was developed to test the quality of all tablets at the packaging step, the system was developed to handle 12,000 tablets per minute [16].

Table 1 Comparison between NIRS and other spectroscopy methods

Method	Applications	Advantages	disadvantage
NIRS	- Quality of raw material - Blending process - Drying - Granulation - Coating - Packaging[30]	- Non-invasive and no sample preparation - Fast acquisition tool - It is classified as " Green Analytical Technology	- The spectrum includes physical information of the sample. - Require a Calibration model - Low penetration depth
Raman	- Blending process - Coating thickness - Coating uniformity - Polymorphic transformation	- No sample preparation required - Provide information on a complex mixture of component complementary to NIRS	- Interfering fluorescence can affect the accuracy of the result - Can be invasive due to the strong laser source
NMR	- Quality of raw material - Blending process - Drug Design [33]	- Non-invasive - No Calibration is needed for the[34]	- Expensive equipment - Slow method compared to the previous methods - Unable to observe Solid Hydrocarbon material - Can be affected by the presence of paramagnetic materials[34].

2.4 Chemometric methods for NIRS

Chemometrics was defined as "The art of extracting chemically relevant information from data produced in chemical experiments"[35]. The extraction of data is performed through mathematical methods mainly statistics, numerical analysis and sometimes through computer based laboratory. Chemometrics took advantage from the increase of storage capacity and processing speed of computers in recent years. This allowed the instantaneous treatment of data as well as the control of the process [36]. Chemometrics allowed a better understanding of chemical experiments, identification of important parameters that influence the chemical experiments and the prediction of experimental results. It led also to the discovery of new analytical tools [37].

NIRS measurements require the handling of a large amount of data (measuring reflectance at different wavelengths) thus it is an ideal application for chemometric methods. In fact, it is considered to be the major factor in the development of chemometrics. The general chemometrics methods applied on NIRS can be classified into two main categories:

Qualitative analysis methods use NIRS spectra to classify materials. The classification can be performed with or without prior knowledge, in the first case the method is referred to as supervised while in the second case it is referred as unsupervised. Generally, the unsupervised methods require a further classification step. Principal component analysis (PCA) is an example of unsupervised methods, it reduces the number of variables which are then clustered using hierarchical or non-hierarchical clustering methods. Supervised methods are used more often than unsupervised methods and they classify samples with respect to known spectra. Some of the most common methods are correlation based methods, distance based methods, linear discriminate analysis (LDA) and partial least squares discriminant analysis (PLS-DA) [16].

Quantitative methods come after the identification of samples and try to quantify the level of differences in measurement samples. For example, quantitative methods applied to NIRS spectra can be used in the pharmaceutical industry to determine many physical properties

such as particle size or hardness of tablets. Regression methods, artificial neural networks and support vector machines are an example of quantitative methods and they try to establishing a relation between the spectra of the material and a specific physical property [38].

2.4.1 Pre-processing methods for NIRS

The physical properties of samples of the same chemical content will cause a variation of the spectrum. Physical properties can cause a vertical shift of the spectrum and also the change of its original slope (multiplicative effects) [39]. These effects can be removed by pre-processing methods.

2.4.1.1 Standard Normal variate

Standard Normal Variate (SNV) eliminates multiplicative effects by subtracting the spectrum i by its mean, $\overline{x_i}$ and dividing it by the spectrum's standard deviation (eq.1) [18].

$$x_{SNV\_ij} = \frac{(x_{ij} - \overline{x_i})}{\sqrt{\frac{\sum_{j=1}^{N}(x_{ij} - \overline{x_i})^2}{(N-1)}}} \quad \text{(eq.1)}$$

x_{SNV_ij} represents the corrected absorbance at wavelength j for spectrum i , x_{ij} is the original absorbance of spectrum i at wavelength j. N is the total number of wavelengths in which the absorbance is measured.

2.4.1.2 Multiplicative scatter correction (MSC)

The method is based on the idea that the correct or the reference spectrum endured multiplicative scatter effects so the resulted spectrum is a multiplication of the reference with a multiplicative correction factor plus an additive correction factor (eq.2) where 1_N is a vector of ones having length N which is the number of wavelength measurements. To go back to the reference, MSC estimates the additive and the multiplicative factors a and b using a linear regression method. To estimate the factors a and b, usually the reference is assumed to be the mean spectrum of all spectra. When estimating the factors, (eq.3) is used to calculate the corrected spectrum [18].

$$spectrum_{measured} = a1_N + b\, spectrum_{ref} \quad \text{(eq.2)}$$

$$spectrum_{corrected} = \frac{(spectrum_{measured} - a1_N)}{b} \quad \text{(eq.3)}$$

2.4.1.3 Derivatives

The simplest way to compute the derivative of a spectrum is finite difference which involves taking the difference between two consecutive wavelengths' absorbance. This difference becomes the new spectral data. Derivatives remove the multiplicative effect and also reduce peak overlapping [18].

2.4.1.4 Savitzky-Golay

Savitzky-Golay is a smoothing and differentiation technique based on the least square method. The method fits a set of consecutive points into a polynomial [40]. The idea behind the Savitzky-Golay method is to preserve the shape of the spectrum while removing the noise.

2.4.2 Pattern recognition methods used in NIRS

2.4.2.1 Multiple linear regression

A regression with more than one variable is called multiple linear regression. The relationship between the dependent variable y and the predictors x_i can be expressed in (eq.4) note ε is an error associate with the relationship between independent and dependent variables [18].

$$y = b + a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_ix_i + \varepsilon \quad \text{(eq.4)}$$

2.4.2.2 Principal component analysis

PCA is an unsupervised method that reduces the number of data variables, the resulting new variables are called principal components and are orthogonal to each other. They also reflect the maximum variance in the data. The linear combination between principal components gives back the original variables (eq.5). X is the original data, T contains the new data and P^T is the loadings or the coefficients, ε is the error [18].

$$X = TP^T + \varepsilon \quad \text{(eq.5)}$$

2.4.2.3 Partial least squares regression

When multiple linear regression is used to create models using too many variables, most likely the model will fail to predict new data due to overfitting. Often not all the variables account for the variation in the response. The independent variable that account for the variation are called latent variables, partial least square was developed to extract latent variables. Partial least square is a supervised method, basically it is a multivariate regression method that can predicts either one dependent variable from multiple independent variables in this case the method is called partial least squares 1 or predicts multiple dependent variable

from multiple independent variables in this case the method is labelled as partial least squares 2. Specifically if the independent variables X are decomposed to TP^T where the columns of T are called the latent vector and Y is decomposed to UQ^T. The goal of PLS is to estimate T and U such that $P^TP = 1$ and $T^TT = 1$ and T^TU is maximum [41]. The latent variables are chosen in a way to minimize the total prediction error [42]. A common approach is to divide the data into a training set and a test set. The training set is used to construct the model before applying it to the test set to determine the prediction error this is called cross-validation. There are several types of cross validation methods the most common are Leave-One-Out and repeat (LOO), Leave-N-Objects out and repeat (LNO). In this work LOO cross validation is used, where only one observation is used to test the model constructed from the rest of observations, the procedure is repeated for each observation and at the end prediction error is calculated [43, 44]

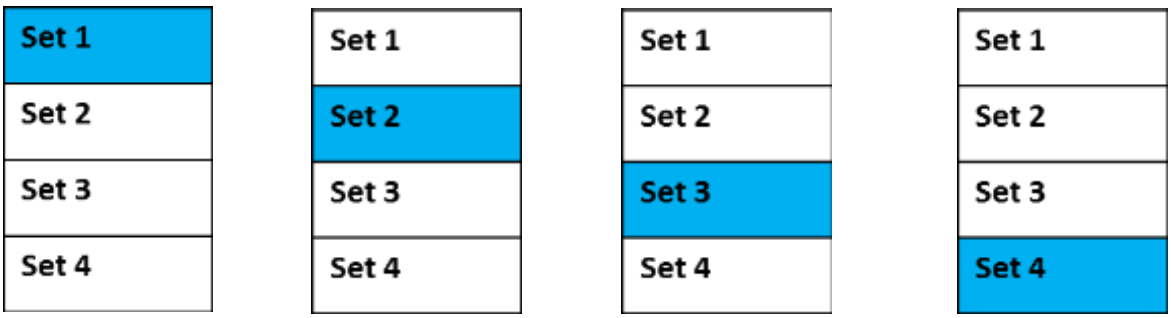


Figure 3 LOO if segment is equal to one sample, LNO if segment constitutes 10-20% of the whole set.

2.5 Image analysis

Image analysis is the use of computer algorithms to analyse digital images. Pixels are the smallest image element and are usually stored in computer memory as matrices. The dimensions of the matrices depend on the type of image. In greyscale images the matrices are two dimensional, where each pixel holds a single value while colour images are stored in

three dimensional matrices where each pixel holds three values corresponding to red, green and blue. Typically there are 256 greyscale levels so each pixel can be stored in 8 bits. For colour images each colour has 256 level intensities so the combination of three colours will give about 256x256x256=16.7 million possible colours. The storage size of an individual pixel is 24 bits.

2.5.1 Image pre-processing

Raw images from the camera can include noise or error resulting from the hardware, light or the alignment of the camera. To improve the quality of the images and remove the noise pre-processing methods are used. Pre-processing methods can be classified into four categories [45]:

- **Pixel brightness transformation** corrects the pixel value depending on the position or applying a greyscale transformation.
- **Geometric transformations** correct the geometric distortion of the image.
- **Local pre-processing methods** correct the image by using a few neighbouring pixels brightness to estimate a new brightness value for a specific pixel location. This method is also called filtration.
- **Image restoration** the changes are applied to the whole image so in this type of pre-processing a convolution is applied to the entire image.

2.5.2 Image Segmentation

An image segmentation step is used to numerate objects that have the same properties in the same image, however it can be challenging. It can be seen as mapping each pixel in the image into an object in the image so it is considered to be a critical step in the image analysis [46]. Segmentation methods can be classified into three categories:

Thresholding based methods which differentiate between pixels depending on their range of intensities for example when identifying the background from individual objects a global threshold is calculated, if the pixel intensity is below the calculated threshold it is set to be zero otherwise it is set to one. The value of the global threshold can be either specified manually by the user by checking the distribution values of pixels' intensity or estimated automatically. A well-known method for automatic tresholding is Otsu's method which will be described in the next chapter.

Edge-based segmentation is the second category and involves classifying pixels into edge and non-edge pixels by applying an edge filter operator. In that way pixels that are located within the same edge are considered to belong to the same object. Examples of edge segmentation methods are the Prewitt filter and Sobel filter.

Region based segmentation group pixels in the same region sharing similar attributes [47].

2.5.3 Image analysis as a measurement tool in pharmaceutical manufacturing

Image analysis is a non-invasive technique, it can allow the analysis of many pharmaceutical sample particles without any sample preparation steps therefore it can give a fast and efficient estimation of solid forms' physical and chemical properties in both powder and tablet forms compared to the classical methods. It can help optimise the formulation development step of solid forms and improve understanding of different manufacturing processes. Image analysis can be used in conjunction with hyperspectral imaging that takes images of a sample at different wavelengths to estimate the chemical properties. Image analysis can also be used in conjunction with a microscopic camera system, which is usually composed of light illumination and a high speed camera to measure physical properties mainly size and shape of particles. For example, during the blending process image analysis of a chemical image of a powder mix was used to determine the content percentage of API [48]. In this case,

multivariate analysis models were used to analyse the different wavelength chemical images and to infer the different compound concentrations [48]. Image analysis of chemical images can also be used to measure the homogeneity of tablets and thus can identify counterfeit tablets. The detection of counterfeit tablets was performed using multivariate analysis of chemical images of tablets at various wavelengths. Image analysis of chemical images was also used to determine the shape and size of particles [48].

For the determination of physical properties of pharmaceutical particles, image analysis is employed on greyscale images in most cases [26]. Several studies also showed the feasibility of in-line measurement of particle physical properties using image analysis during various pharmaceutical manufacturing processes. For example, image analysis of images captured with a high speed camera was used to monitor the twin-screw granulation process, the approach proved to have a quick response to the change of particle size during the granulation process [28]. During in-line measurements, capturing images of moving particles can be challenging due to the complex and fast movement of the particles so in many cases a sampler or chute is attached to the manufacturing devices to sample several batches for measurement. In addition, as it will be shown later in this study image analysis of a bulk powder for direct size determination can be challenging and time consuming. Several studies have used texture analysis of bulk powder together with multivariate analysis to determine particle size [8, 49, 50].

At the tablet manufacturing level, image analysis of digital images was used to determine the end point of tablet coating and to check for the uniformity of coating among tablets [51]. Another study investigated the feasibility of using image analysis for real time quality inspection of gelatine capsules and showed an acceptable accuracy of 95% [52]. Finally, image analysis was also used in inspecting the quality of the print on a tablet [53].

2.6 Particle size

A particle can be defined as the smallest constitute of a substance. Particle size, shape and surface morphology are key physical parameters in pharmaceutical manufacturing. Particle size affects both pharmaceutical unit operational efficiency and the final product quality. Many unit operations such as pre-mixing/mixing, granulation, drying, milling, roller compaction, spray-drying, coating and compression are affected by particle size. Also the final product quality is affected by particle size [6].

2.6.1 Particle size diameter estimation

There are many ways to quantify the size of particles; the simplest one is finding the equivalent circle that has the same projection area as the particle. The size of the particle is then approximated by the diameter of the circle [54]. Other variations involve finding a circle that has the same perimeter as the particle. Particle size can also be quantified by the [55]:

- Diameter sphere that has the maximum or minimum length of the particle.
- Diameter sphere that has the same volume as the particle.
- Diameter sphere that has the same surface area.
- Diameter sphere that has the same weight.
- Diameter sphere that passes the same sieving as the particle.
- Diameter sphere having the same sedimentation rate.

2.6.2 Feret diameter

Feret diameters includes a set of different sizes estimation based on the distance between two parallel plane bounding the particle in different direction (0º, 180º,…). The size of the particle can be the maximum, the minimum or the mean of Feret diameters. Feret diameter 90º to the maximum or minimum Feret diameter can also be considered as particle size [54].

2.6.3 Chord length

Generally chord length is the distance passing through the centre of gravity between two points laying at two extremities of the particle. Similarly to Feret diameters, there are many chord lengths for particles depending on the orientation of the particle. But usually the size of the particle is quantified as the maximum or minimum of the chord lengths in different particle orientation or the mean of chord length in different orientations [54].

2.6.4 Martin diameter

Basically it is defined as the chord length that divides the particle projection area into equal half areas. Martin diameters are also a collection of particle sizes based on the orientation of the particle. So the particle size can be quantified either by the maximum, minimum or mean of the different particle orientations measured [54].

2.7 Particle size analysis in pharmaceuticals

There are a number of methods that can be used in pharmaceutical manufacturing for particle size analysis, they can be classified into off-line and in-process methods [6]. In off-line methods a sample is taken from the production line to perform the analysis while in the in-process the measurement is performed during production.

2.7.1 Sieving analysis

Sieving analysis is a traditional off-line method and it used to be the most common method in particle size analysis. The sieving tool constitutes of many sieves stacked one over the other in a decreasing order of mesh size. When the powder sample is fed into the top sieve, the bottom of the sieve is subjected to agitation to help the particles pass through the sieves [56]. For the particles to pass through the sieve they must have their two largest dimensions smaller than the size of the mesh. At the end, the material in each sieve has particles with the approximate same size, so the sieve stack is disassembled and the proportion of powder in each sieve is weighted to give the distribution of the sample particle size.

2.7.2 Laser diffraction

Laser diffraction (LD) can be used as an off-line or in-process method, currently it is a popular method in pharmaceutical manufacturing. It uses information of laser scattering to deduce the particle size distribution of the sample as large particles scatter light at small angles in comparison with smaller particles which scatter light at bigger angles [57]. The method assumes that particles have a spherical shape [6].

2.7.3 Spatial filtering velocimetry

Spatial filtering velocimetry use chord length as a measure for particle size. Particles cross a laser beam creating shadows over a set of linear arrays of fibres, subsequently a burst signal which is proportional to the velocity of the particle is generated. Simultaneously another

pulse signal is generated and so the chord length is estimated by multiplying the velocity with the time of the pulse [58].

2.7.4 Focused beam reflectance measurements

Focused beam reflectance measurements (FBRM) also use the chord length of a particle as a size measurement. It is a laser based system, the laser is focused on a sapphire window where particles are located (FBRM ignores particles away from the sapphire window). Then the laser is rotated around the sample, in each step the backscattered light from the particle is detected by the laser probe and thus the chord length of the particle is determined. By the end of the measurement, the FBRM constructs a size distribution from the number of measured particles.

2.7.5 Photometric Stereo Imaging

Stereo imaging is based on the change of the illumination angle while keeping viewing at a constant position. So the surface orientation for each pixel is determined and assigned a value between 0 and 255 (greyscale image). Because of the shadow resulted from the changing illumination angle, the high brightness values are considered to be particles so the volume of the particle is estimated by multiplying the area with the diameter [6, 59].

2.8 Surface texture analysis

Surface texture is the repeated fine irregularities in the surface of an object. Usually, surface texture is associated with roughness and waviness of the surface [60]. In pharmaceutical manufacturing processes, surface texture can affect powder flow and compression behaviour in addition to friability, and adhesion properties of manufactured intermediaries and final products [61]. The quantification of surface texture can be made through the measurement of

surface roughness parameters or through image texture analysis. Surface roughness parameters quantify the heights of different small structures and their occurrences on the surface, while image texture analysis applies statistical methods to quantify the difference between pixel intensities. Surface roughness parameters can be measured directly using a variety of profilometry methods which can be classified as contact profilometres and non-contact profilometres. Contact profilometres is much like the old record player, usually a stylus is moved across the studied surface. The stylus records the vertical elevation by transforming its mechanical movement into electric signal which in turn gives roughness surface parameters. Non-contact profilometres as the name suggests analyses the surface without interfering with the object's surface. For image texture analysis usually an image is captured using optical microscopy.

2.8.1 Atomic force microscopy

Atomic force microscopy (AFM) can be operated in a contact or non-contact mode. Generally, it consists of a reflective cantilever attached to a tip. The cantilever tip deflects according to the local surface of the sample and the magnitude of the deflection is registered by changes in direction of a laser beam reflected of the cantilever [62]. In contact mode, the cantilever tip is in contact with the surface sample, while in the tapping mode the cantilever tip is oscillating vertically touching and releasing the surface sample at a constant frequency. In non-contact mode, the tip oscillates above the surface and the interactions of molecular forces between the surface material and the tip, determine the topology of the surface [63]. In all the modes topological maps of the surface texture and surface roughness parameters can be drawn. AFM can measure very fine details down to ~1nm so it is slow compared to the other methods and scans only a small portion of the surface. In addition, AFM in contact mode can change the surface texture of the sample.

2.8.2 Laser profilometry

Laser profilometry is classified as a non-contact profilometry method but it is able to quantify directly roughness parameters. It has sensors that read the reflectance of a laser beam applied on the material surface. The sensors then translate the measured light into amplitude parameters which in turn provide a measure of several roughness parameters [64, 65]. Laser profilometry can cover larger areas than AFM, ranging from ~1mm up to several centimetres, it is also the fastest profiling method [66].

2.8.3 Optical microscopy

The surface of a particle or tablet can be studied using a microscope which captures a magnified image of the surface. Usually, optical microscopy is composed of a light system and magnified lenses. Optical microscopes provide a 2D image of the surface so they do not quantify the surface texture or accurately determine the surface topography. But, it gives a fast general scan of the surface texture. The classification of the surface texture is left to the user or to an image texture analysis method.

2.8.4 Scanning electron microscopy

Scanning electron microscopy (SEM) is similar to optical microscopy in the sense that it takes an image of the material surface but in high resolution. The other difference between SEM and optical microscopy is that SEM uses electrons instead of light. SEM scans the surface morphology by directing a fine beam of electrons onto the surface of the material. The beam of electrons interacts with the atoms on the surface resulting in the emission of electrons which are then collected in the detector to form an image of the surface morphology. So the image describes more accurately the surface texture than optical microscopy. SEM sometimes requires sample preparation for non-conductive material [67].

SEM also gives an accurate 2D image of the surface without quantifying the topology of the surface.

3 Materials and methods

3.1 Particles used in experiments

The particles employed in this study consisted of microcrystalline cellulose spherical pellets known as Cellets® (Pharamatrans Sanaq AG Basel, Switzerland). Table 1 shows the size of different categories of Cellets® while Figure 4 shows an RGB image of different size particles taken by the Eyecon™. In addition, sugar spheres known as pharma-a-spheres™ (Hanns G. Werner GmbH, Tornesch Germany) were used which consists of sucrose and corn starch and their size is between 100 and 2000 µm, Table describes the category sizes of the sugar spheres.

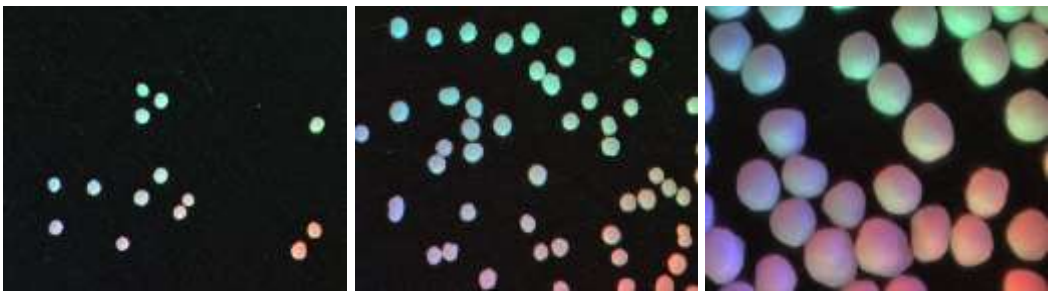
To test the ability of the multipoint NIRS via chemometrics algorithms to measure accurately the proportions of a binary mixture, samples containing blends of different proportions of Cellets® and sugar particles were prepared. The three particles nominal sizes were 100, 500 and 1000 µm. For measuring physical properties, samples were prepared consisting of Cellets® with nominal sizes 100, 200, 350, 500, 700 and 1000 µm to perform size analysis under bench top conditions.

Table 2 Cellets® size categories.

Particle size categories	Size(µm)
Cellets® 100	100-200
Cellets® 200	200-355
Cellets® 350	350-500
Cellets® 500	500-710
Cellets® 700	700-1000
Cellets® 1000	1000-1400

Table 3 Some common size of sugar spheres.

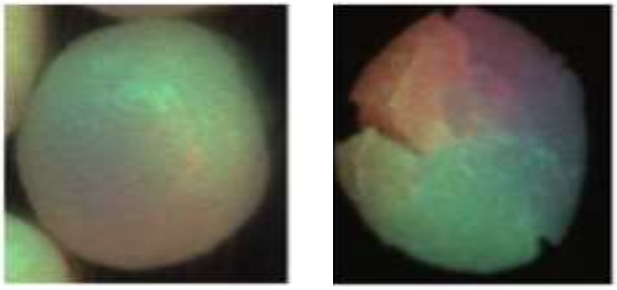
Particle size range (µm)	Particle size range (mesh ASTM)
106-125	100-200
212-250	70-80
250-300	50-60
500-600	30-35
600-710	25-30
850-1000	18-20



(a) (b) (c)

Figure 4 RGB images of different particle categories. (a),(b) and (c) Cellets® of nominal size 200, 350 and 1000 µm.

Particles with a rougher surface morphology were prepared by placing Cellets® 700 in a 12 M hydrochloric acid (Sigma-Aldrich Buchs SG, Switzerland) solution for 10 minutes. Then, the particles are removed from the solution and left in the lab overnight to dry. Figure 5 (b) shows the resulting surface of chemically etched particles using HCL acid. In addition, API particles with rough texture morphology were used (Innopharmalabs, Dublin, Ireland). Due to intellectual property rights, the name and structure of the active pharmaceutical ingredient (API) cannot be shown, it will be denoted as API in this work. Figure 6 illustrate the shape of the API particles which have a spiky surface morphology.



(a) **(b)**

Figure 5 Eyecon™ images of Cellet® 700 (a) non-treated and (b)treated.

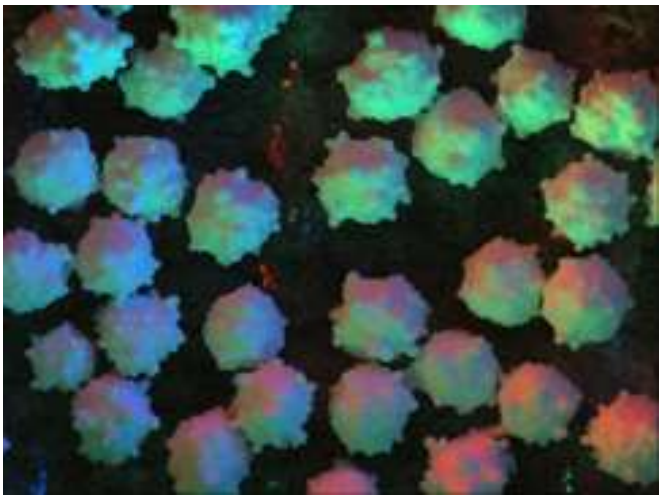


Figure 6 Active pharmaceutical ingredient particles with a spiky surface.

3.2 Instruments

3.2.1 Prototype of the NIRS

A MultiEye™ (Innopharmalabs, Dublin, Ireland) was used, it is a multipoint NIR spectrophotometer. Beside the NIRS software, the system has the following components:

- Illumination unit of type HL-2000 (Ocean Optics, Dunedin, Florida, USA) consisting of a light source (Tungesten halogen), diffusers and illumination optics.

- The MultiEye^{TM} central unit containing the spectrometer is based on the VTT's proprietary piezo-actuated Fabry-Perot interferometer technology and Rikola Ltd's four-channel detector technology.
- The MultiEye^{TM} consists of four probes.
- A probe (Ocean Optics, Dunedin, Florida, USA) is made of a bundle of seven fibres, six fibres transport light from the source to the material and one fibre transports the reflected light from the material to the central unit as depicted in Figure 7.

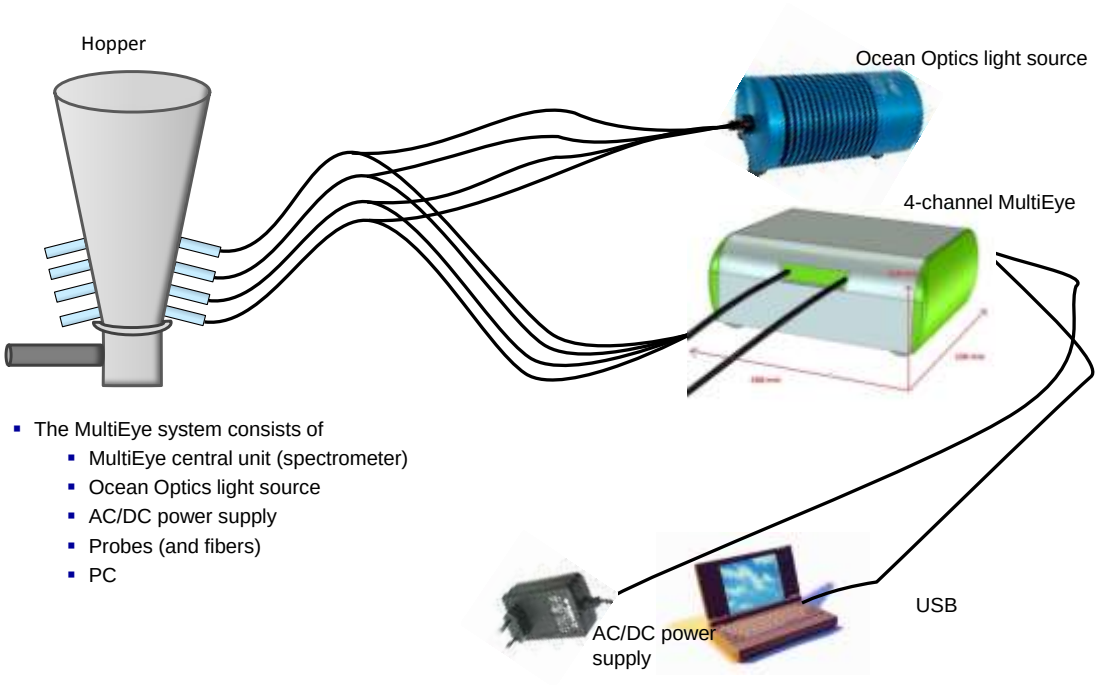


Figure 7 MultiEye^{TM} NIRS prototype system.

Prior to taking any measurement, the light coming from the light source is blocked to measure the dark signal in the full spectral range I_d. Then a white reference (Spectralon, Avion, USA) is measured by taking the full spectrum and is denoted I_0. The measurement of

the reflected light from any sample was collected in the same way as the white reference and the reflectance R values were calculated by equation 6 [68].

$$R = (I - I_d) / (I_0 - I_d) \quad \text{(eq.6)}$$

The diffuse reflectance NIR reflectance spectra obtained was then converted into absorbance, A, by equation 7. The distance of each probe from the sample was kept between 0.5-1 mm and the spot size is about 500 µm.

$$A = \log (1/R) \quad \text{(eq.7)}$$

3.2.2 Imaging system

All the images in this study were captured using an Eyecon™, a novel camera based system that consists of a camera surrounded by a set of 15 LEDs emitting 1 µs red, green and blue light pulses from different angles simultaneously. The speed of the pulses allows the Eyecon™ to capture images of moving particles up to 10 m/s [28]. The illumination is based on the principle of photometric stereo imaging which allows the 3D characterisation of particles. The colour distribution on the sample being illuminated by the Eyecon™ camera provides topological information which in turn permits the identification of particles' edges for size determination. The Eyecon™ camera has a field of view of 9 x 6.5mm and a pixel size of 6x6 µm. It can estimate the size of particles in the range of 50 to 3000 µm [28].

The Eyecon™ system includes image analysis software based on edge detection of particles. The software fits ellipses on particle edges to calculate the two maximum dimensions of the particle, then the aspect ratio is calculated and an average diameter is calculated from the two maximum dimensions. Then the distribution of the diameter is constructed from these measurements. For this study, the Eyecon™ image analysis software will not be used instead

a novel image analysis strategy based on global thresholding was developed to estimate the size of the particles.

Figure 8 Eyecon™ device.

3.3 Samples preparation

In this study different experiments were carried out to test the ability of the MultiEye™ and Eyecon™ to measure the chemical and physical properties of particles mentioned earlier. For chemical properties first a calibration standard containing dysprosium, holmium, and erbium oxides (Avian Technology, Sunapee USA) was used to check the MutiEye™ for measurement in static condition, three samples were prepared containing Cellets® particles, sugar spheres and mixture 50% (w/w) of sugar spheres and Cellets® all having a nominal size of 500µm. The mixture sample was prepared by mixing 5g of sugar and 5g of Cellets® in a plastic container and mixing the containing horizontally and vertically for about one minute in order to obtain a homogenous blend. The same samples were used to take measurements in motion. To perform the in-line measurement of Cellets® content, 25 samples were prepared for building up the calibration model and 10 samples were prepared to perform the in-line measurement. The 25 samples contained 0%, 25%, 50%, 75%, 100% (w/w) of

Cellets® for five different nominal sizes that are 100, 200, 500, 700 and 1000 µm with a total weight of 20g. The mixture was obtained in a similar way as the previous samples by shaking the plastic container in the vertical and horizontal direction for about 1 min. The other 10 samples contained either sugar or Cellets® with five nominal sizes (100, 200, 500, 700 and1000 µm) and their weight were 10g each.

For the physical properties, random samples were taken from Cellets® particles with nominal size 100, 200, 350, 700 and 1000 µm. For dynamic particle size only samples from 100, 500 and 1000 µm were considered. To perform texture analysis six samples were prepared containing Cellets® 700 µm, chemically etched Cellets® 700 µm, a mixture of 50% (w/w) of Cellets® and chemically etched Cellets®, Cellets® 1000 µm, API particles and finally a mixture of 50% (w/w) Cellets® and API.

3.4 Experiments

3.4.1 Measurement of chemical features under static condition

The goal of this experiment is to test the ability of the MultiEye™ to measure chemical properties under static conditions. Prepared samples of Cellets®, sugar spheres and their mixture were placed under the MultiEye™ probes to take measurements. The probes were 0.5~1mm away from the sample and the light spot size was about 500µm.

3.4.2 Measurement of chemical features under dynamic conditions

To investigate the effect of motion on the MultiEye™ ability to measure chemical and physical properties, the movement of particles was simulated via a rotating sample holder. The speed of rotation of the holder was estimated to be 80~90 rpm. A tray containing the sample was placed on the centre of a rotating fan so that it rotates at the same speed as the fan

as shown in Figure 9. For this experiment, the probes were fixed away from the centre of the tray in the moving stream.

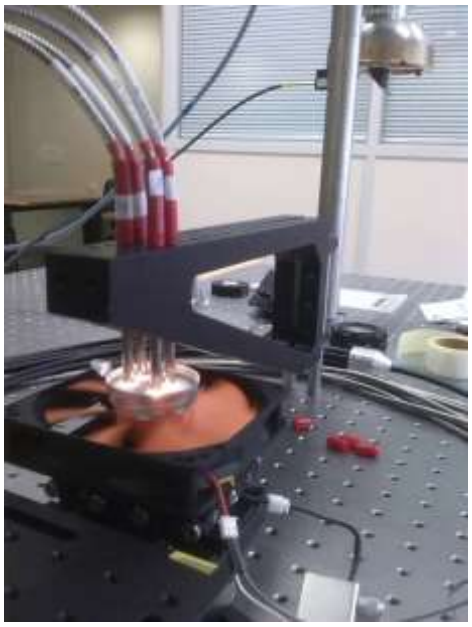


Figure 9 Experiment set up for measuring chemical properties for moving particles.

3.4.3 In-line monitoring of chemical property using NIR

In this experiment the spectra measured from the MultiEye™ was used to determine the proportion of two compounds in a powder mix. The two compound used in this experiment are the sugar sphere and Cellets® both have nominal size 500 µm. The overall procedure is summarised in Figure 10. The first step was to construct a calibration model so five samples were prepared from Cellets® and sugar spheres containing 50% (w/w), 25% of sugar and 75% Cellets®, 75% of sugar and 25% Cellets®,100% Cellets® and 100% sugar spheres. Samples were placed in a stainless steel circular container with an 8 cm diameter and placed on a vortex shaker of type Fisherbrand® (Fisher Scientific, Leicestershire England). Once placed on the shaker the container vibrates in a circular motion, in this work the shaker circular speed is limited to 500 rpm and the MultiEye™ is used to measure the spectrum of each sample. During the experiment the probes are placed above the centre of the container 0.5~1

mm away from the container's content. Partial least square is then used to create a model that correlates the spectra to the proportion of Cellets®. Once the calibration model is constructed, it can be used for the in-line monitoring of proportion measurements. The in-line measurement of proportion experiments is performed by placing, 100% sugar or Cellets® in the vibrator and after some time sugar or Cellets® is added to the vibrator to make proportions of 50% (w/w) or 25% (w/w) or 75% (w/w).

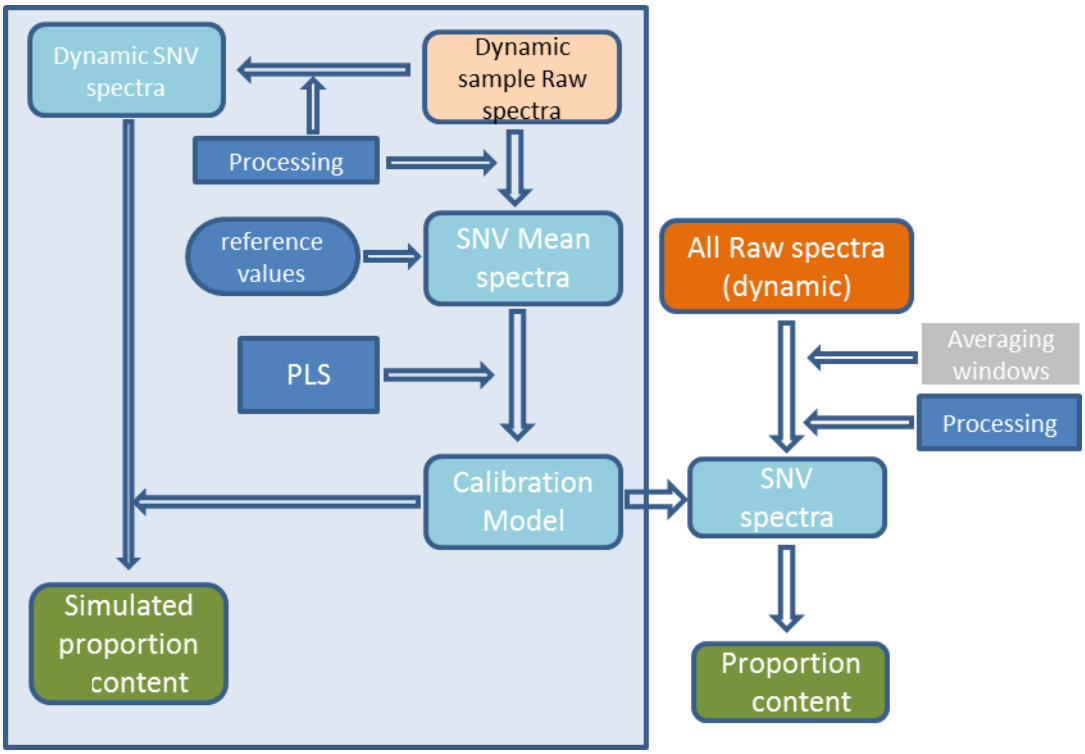


Figure 10 General steps to in-line measurement of mixture proportions.

Figure 11 Measurement of spectrum of mixture particle sample vibrating.

3.4.4 Static particle size and morphology measurements

Samples of Cellets® of different nominal sizes were placed in the field of view of the Eyecon™. Three channel images were captured for size analysis. The particles were placed first in a manner that allowed space between particles and thus the background was apparent in the images and then more particles were added to the tray to hide the background resulting in images of bulk particles.

3.4.5 Dynamic particle size and morphology measurements

For the physical property measurement, the particles were fixed onto a conveyor belt (Dorner® 2200, Hartland, USA) prior to using the Eyecon™ to capture images. The length of the conveyor belt was about 610 mm and width of 44mm, the maximum speed of the conveyor belt is estimated to be 23 cm/s and the minimum speed was estimated to be 3 cm/s. The focal plane of the Eyecon™ was adjusted to the surface of the belt. Particles were fixed to the surface centre of the conveyor belt in the same field of view as the Eyecon™.

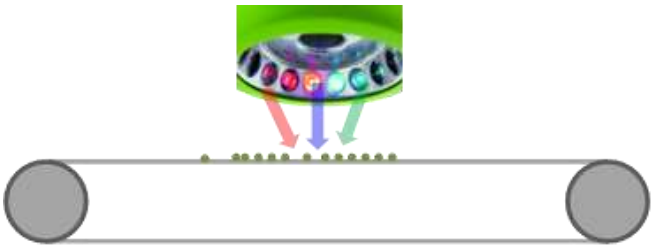


Figure 12 Experiment set up for simulating particle movements using conveyor belt.

Another experiment was conducted to capture images of particles during a free fall state as shown in Figure 13. To ensure that the maximum of particles were within the field of view of the Eyecon^{TM}, a feeder together with a glass window were used to restrict the movement of particles.

Figure 13 Free fall state experiment set up.

3.5 Particle size using image analysis

3.5.1 Pre-processing

The image analysis algorithms were implemented using the R software [69]. Steps conducted to pre-process and segment the three channel images are displayed in Figure 14. The three channel images were converted into single greyscale images. A 9x9 median filter was then applied on the greyscale images to remove potential outlying pixels replacing them with interpolations of the intensity values of the surrounding eight pixels.

3.5.2 Image segmentation

The segmentation step was then conducted to isolate objects of interest in each image. Typically, the first step in image analysis is thresholding where the foreground and the background are labelled resulting in a binary image [70].

3.5.2.1 Otsu's thresholding algorithm

In this study an automatic thresholding method known as Otsu's method was employed [52]. The method starts with calculating the histogram of the greyscale image, then iterates through all possible greyscale values to find the optimal threshold that minimises the within class variance. In each iteration, a greyscale value is used to partition the pixels into two groups. The variance in each group is calculated as well as the probabilities associated with each group before calculating the within class variance. The within class variance is the weighted sum of the two groups variance, the weights are taken as the probabilities of each group. The optimal threshold is the threshold that yields the smallest value of the within class variance.

3.5.2.2 Watershed algorithm

The objects in the foreground were then subsequently segmented using the watershed method [71]. The watershed algorithm considers the greyscale image as a topographic surface where usually the intensities reflect the height of the surfaces but in this case the distance to the background reflects the height of the surface. Therefore prior to the watershed algorithm, the distmap function is used to evaluate the distance to the background for each object, so the function output is a contour plot describing the distance to the background and the centre of the object will have the highest elevation value. The watershed algorithm starts flooding the objects from their centres with different colours until the colours overlaps or the colour reach the background, the result is identified objects. The objects then can be remunerated as shown in Figure 14 (c).

3.5.3 Particle size measurement and area selection for surface texture analysis

Before measuring the objects' size and selecting surfaces, all objects in contact with the image borders were removed. These objects are not complete and therefore their inclusion can bias the calculated mean particle size. In this work Cellets® or sugar spheres are circular in shape so the diameter of the particle can be measured in one direction this was similar to measuring the diameter of the equivalent sphere [72]. There is an additional step for the texture analysis which is to select regions around each particle's centre before applying the surface texture analysis. The area selected is 80x80 pixels corresponding to a region of 480 µm x 480 µm.

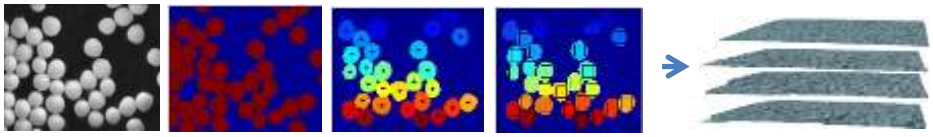


Figure 14 Steps of image segmentation algorithm. (a) Greyscale image. (b) Convert to binary image using Otsu's method. (c)Identify Cellets® centres and estimate size diameter. (d) Select square regions around the centre of each Cellets®.

3.6 Texture analysis

Statistical surface texture analysis can quantify the spatial distribution of pixel intensities in an image. The statistical methods can be classified into first, second or higher order statistics. First order statistics analyse the distribution of pixel intensities without taking into consideration their spatial location. The mean and standard deviation of pixels intensities can be considered as first order statistics. Second or higher order statistics derive statistics from the intensity distributions of two or more pixels at specific locations in relation to each other [73].

3.6.1 Grey level co-occurrence matrix correlation property

Grey level co-occurrence matrix (GLCM) is used extensively to quantify the texture of a surface [74] . In pharmaceutics it has been used to analyse bulk powder surfaces to determine particle size and shape and to analyse the surface morphology of tablets [8, 75]. The method calculates first the co-occurrence matrix from greyscale images by counting how often the grey level intensities occur at a specific distance (d) and orientation (θ). Figure 15 is an image consisting of only of two colours: black and white where black is represented by ones and white with zeroes. So if the co-occurrence matrix is constructed for one step in the -45º direction it will end up with the co-occurrence matrix as shown in the same figure.

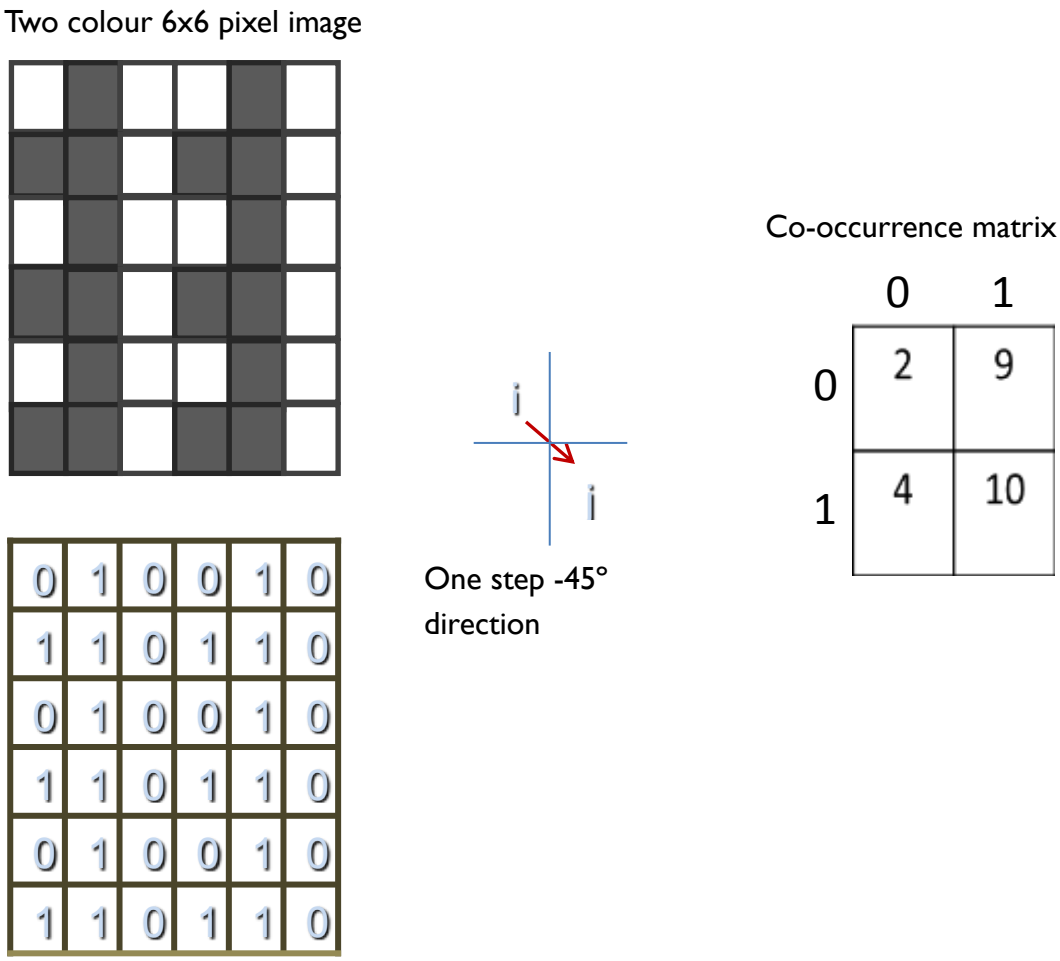


Figure 15 Co-occurrence matrix construction from two colour image.

The co-occurrence matrix is translated to a joint probability distribution $P_{d,\theta}$. The number of rows or columns corresponds to the number of grey-levels (intensities) in the image under study [74]. From the co-occurrence matrix 14 different properties can be calculated [76] however typically only the angular second momentum, contrast, correlation, and entropy properties are used [74]. In this study, only the spatial correlation property was calculated (equation 8)

$$correlation = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} i \cdot j P_{d,\theta}(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \qquad \text{(eq.8)}$$

where i and j are indices of the rows and columns of the co-occurrence matrix representing the intensities levels. Note μ_x, μ_y, σ_x and σ_y are the mean of the rows and columns, the

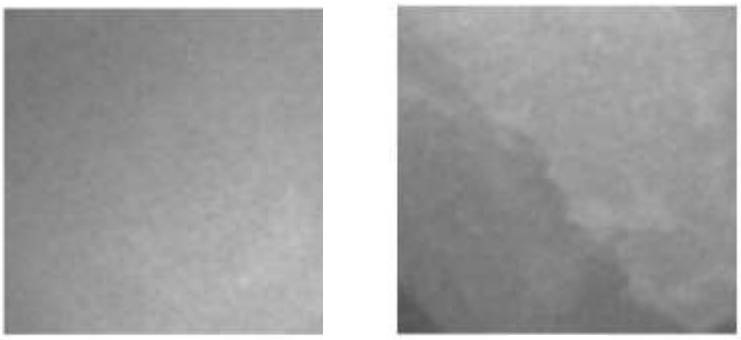
standard deviation of the rows and standard deviation of the columns respectively of the co-occurrence matrix. The plot of the correlation property against the distance step reflects the structure of the surface where the slope increases with an increase in the roughness of the surface.

3.6.2 Variogram

Variograms give a description of how the data are related over distance k and direction θ. Generally, variograms are characterised by the sill which represents the maximum value of the variogram, the range which is the distance where the variogram value becomes constant and the nugget which represents the variability at small distances. Variograms can be defined as the expectation, denoted E[…], of the difference value squared between two locations x_i and x_{i+k} separated by a distance of k in a given direction (eq.9) [77].

$$Variogram = \frac{1}{2}E[(x_i - x_{i+k})^2] \quad \text{(eq.9)}$$

As far as surface texture is concerned, it is indicated by the variogram at the origin, rough texture will have a higher slope at the origin while a smoother surface will have a lower gradient at the origin. In this study the variogram function for different surface textures is calculated using the **variog** function which is a built in function in the GeoR library in the R package [78].



(a) **(b)**

Figure 16 Surface morphology of (a) Smooth particle and (b) Rough particles.

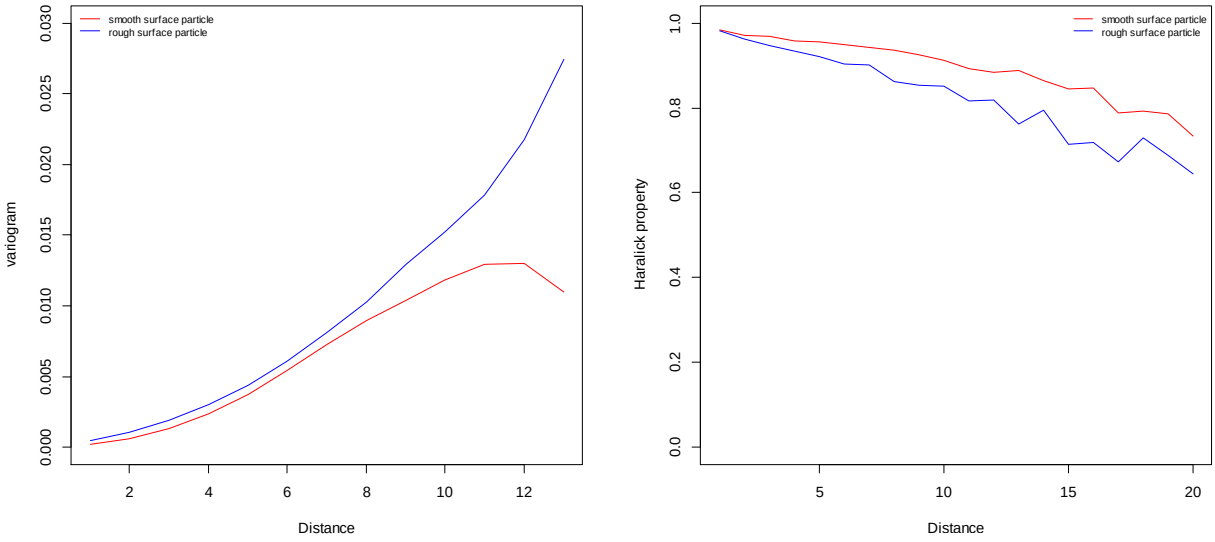


Figure 17 (a) Variogram function and (b) Haralick correlation property of two particle surface texture shown in the previous figure.

Figure 16 is two greyscale images of two different surface textures of two particles. The first one has smooth surface and the second one has rough surface. Figure 17 shows the variogram function and Haralick correlation property corresponding to the two surfaces, the variogram of the smooth surface are smaller than the variogram of the rough surface in each lag distance. Haralick correlation property graph of both smooth and rough texture starts at the same point and decreases at different rates (the slope of the rough texture is higher than the smooth texture).

3.6.3 Autocorrelation function

The autocorrelation function described in equation 10 can assess the amount of periodicity as well as finesse/coarseness of the image by analysing and correlating all the pixel intensities (n). The autocorrelation function starts at one and it decreases at different rates depending on the structure of the surface. It decreases at a slow rate with the increase of pixels distance for an object image with a smooth texture. On the other hand, for object images with rough texture, the slope rate increases because of the difference of intensities over a short distance.

$$\rho(k) = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \quad \text{(eq.10)}$$

3.6.4 Sobel operator and texture analysis

To isolate details of surface structure of different particles, an edge based method can be applied. Sobel operator is an edge detection segmentation method and it can be employed to extract the structure of the surface morphology. The Sobel operator is based on a two matrix operators; one is used to perform a convolution operation with the greyscale image in the horizontal direction and the other in the vertical direction. The two convolution operators give two images showing the edge in the horizontal and vertical directions. A final image is reconstructed from the two using equation 11.

$$Final\_image = \sqrt{image_{horizontal}^2 + image_{vertical}^2} \quad \text{(eq.11)}$$

Figure 18 (a) and (b) shows surface texture of a smooth Cellets® and it corresponding texture features extracted using the Sobel operator and surface texture of a chemically etched Cellets® obtained from the sample preparation in section 3.1. The variograms of the edge based image of a smooth Cellets® and chemically etched Cellets® are shown in Figure 19, in this case the difference between the two variograms is more apparent than the variogram of the original surface, this suggests that analysing the textural feature of surfaces leads to a better classification than analysing the original surface.

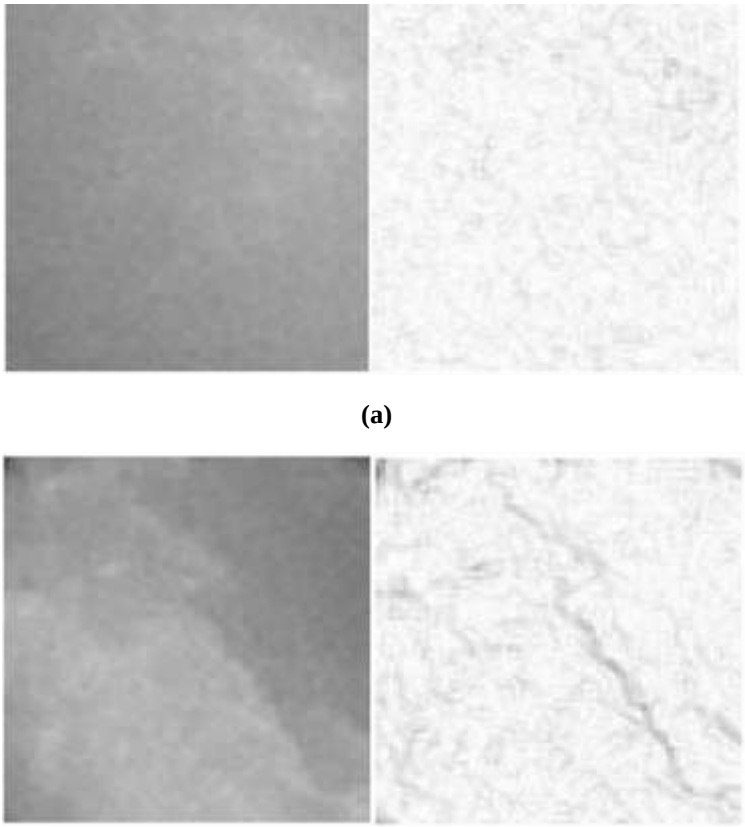


Figure 18 Surface texture after applying the Sobel operator on (a) Smooth Cellets® and (b) chemically etched Cellets®.

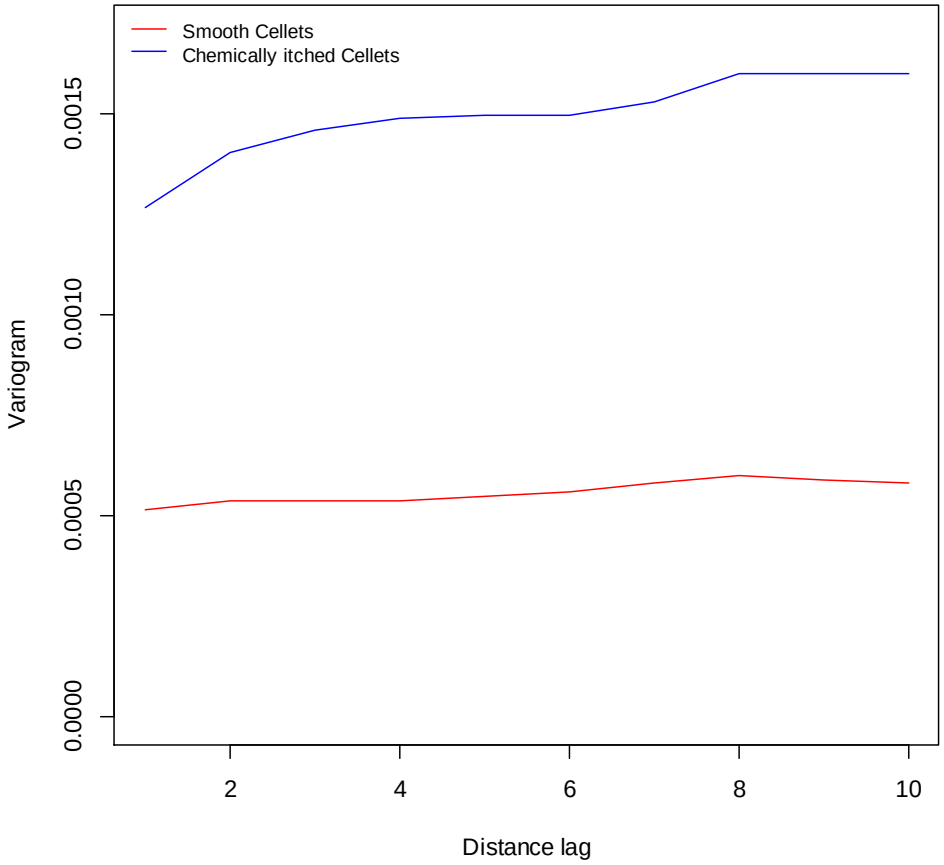


Figure 19 Variogram of edge based image in the previous figure.

4 Results and discussions

4.1 Chemical properties

4.1.1 NIR Calibration trial

For this study, the MultiEye™ prototype was tested using a calibration standard containing dysprosium, holmium, and erbium oxides (Avian Technology, Sunapee USA). These experiments consisted of four probes attached to the MultiEye™. The probes were not in contact with the calibration standard and were approximately ~0.5-1 mm away from the sample. Table 2 presents the settings of this experiment. Note the wavelength range of 1515 to 2295 nm with a measurement taken every 5 nm thus a total of 118 measurements were recorded. The spectrum generated for each probe is an average of five complete wavelength scans. In each scan the absorbance measurement was performed sequentially by averaging 1000 measurements. Hence the total number of measurements was 5000. .

Table 2 Settings for the NIR calibration trial.

Wavelength range	1515 nm-2295 nm
Step size	5 nm
Sampling per wavelengths	1000
Number of scans	5

Figure 20 shows the resulting spectra of the four probes of the MultiEye™ prototype, all the spectra have peaks at the same wavelength as the reference spectrum but with different values

of absorbance because of baseline shifts. Note Figure 20 includes the reference spectrum provided with the calibration standard.

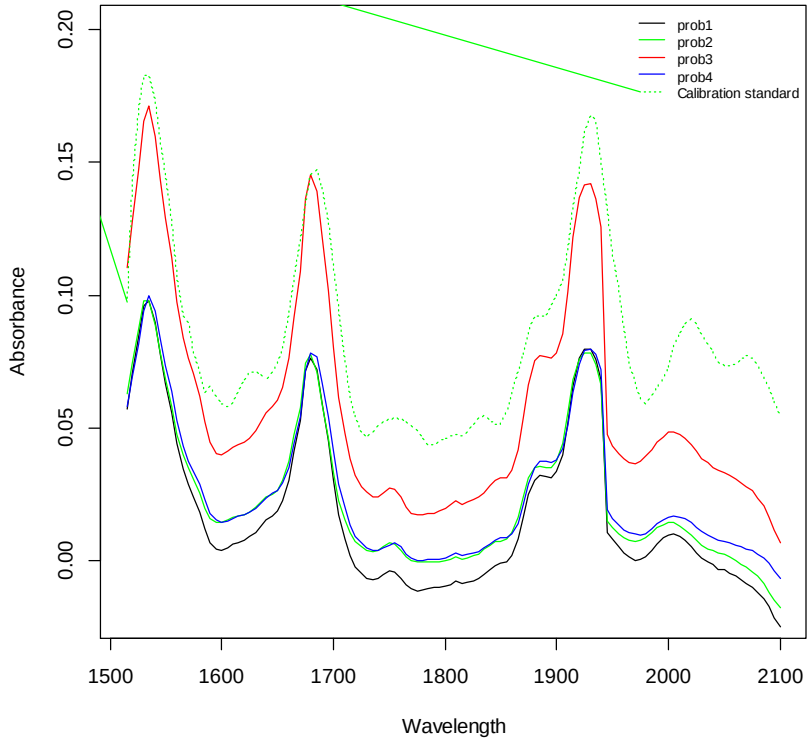


Figure 20 Spectra of the four probes in addition to the calibration standard.

Figure 21 is the spectral data from the calibration standard after using the standard normal variate pre-processing method. The MultiEye^{TM} prototype produces approximately the same spectrum as the reference spectrum under these experimental conditions. All major peaks are correctly identified.

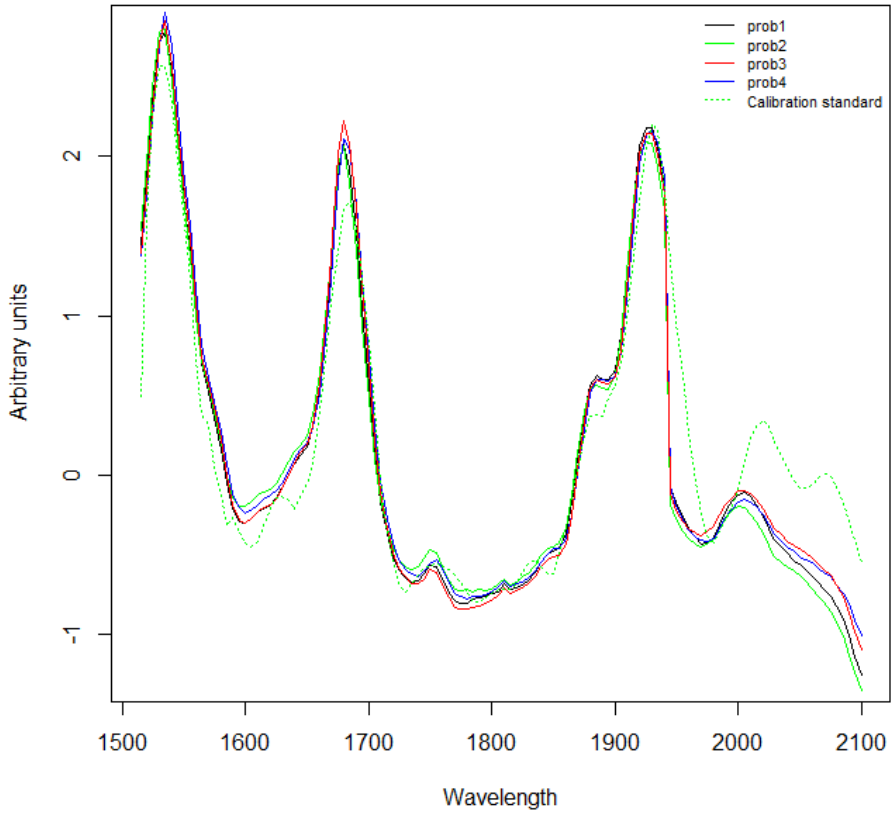


Figure 21 Spectra of the four probes and the calibration standard after SNV pre-processing.

4.1.2 NIR spectrum of cellulose and sugar spheres in static conditions

The spectra of three samples containing pure cellulose particles of nominal size 500 µm, sugar spheres particles of nominal size 500 µm, and mixtures of equal weight Cellets® and sugar spheres were measured using the four probes of the MultiEye™. The mixture sample was prepared by weighting approximately 15g of Cellets® and 15g of sugar spheres before mixing them. The mixing was performed until a uniform mixture was obtained. The four probes of the MultiEye™ were fixed at a vertical position 0.5~1mm away from the samples' surface. Table 3 presents the settings of the MuliEye™ in this experiment. Note the

wavelength range was from 1520 to 2290 nm with a measurement taken every 10 nm thus a total of 79 measurements were recorded. The spectrum generated for each probe is an average of five complete wavelength scans. In each scan the absorbance measurement was performed sequentially by averaging 1000 measurements. Hence the total number of measurements was 5000.

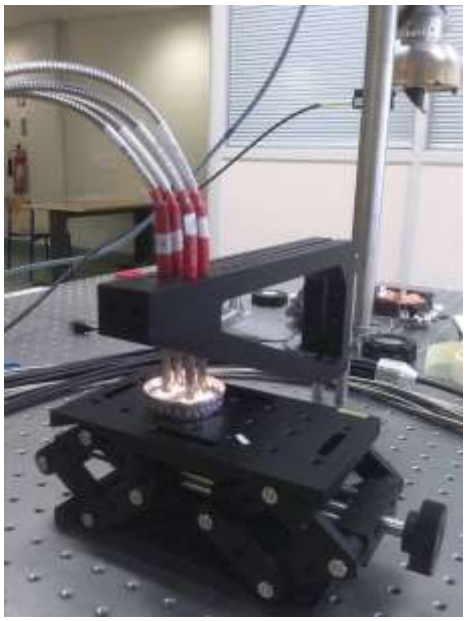


Figure 22 Experiment setup for the NIR.

Table 3 Setting of the Cellets®, sugar spheres and mixtures for NIR experiments.

Wavelength range	1520 nm-2290 nm
Step size	10 nm
Sampling per wavelengths	1000
Number of scans	5

Figure 23 (a) shows the resulting spectra of the Cellets®, sugar and mixture sample obtained by the four probes. Figure 23 (b) shows the spectra after applying the standard normal variate pre-processing method. SNV removed the baseline shifts. The spectra obtained by

the probes were validated in the NIR literature [79]. The mixture spectra are between the spectra of the pure Cellets® sample and the spectra of the pure sugar sample at wavelength 1930. It has also more variation than the pure samples.

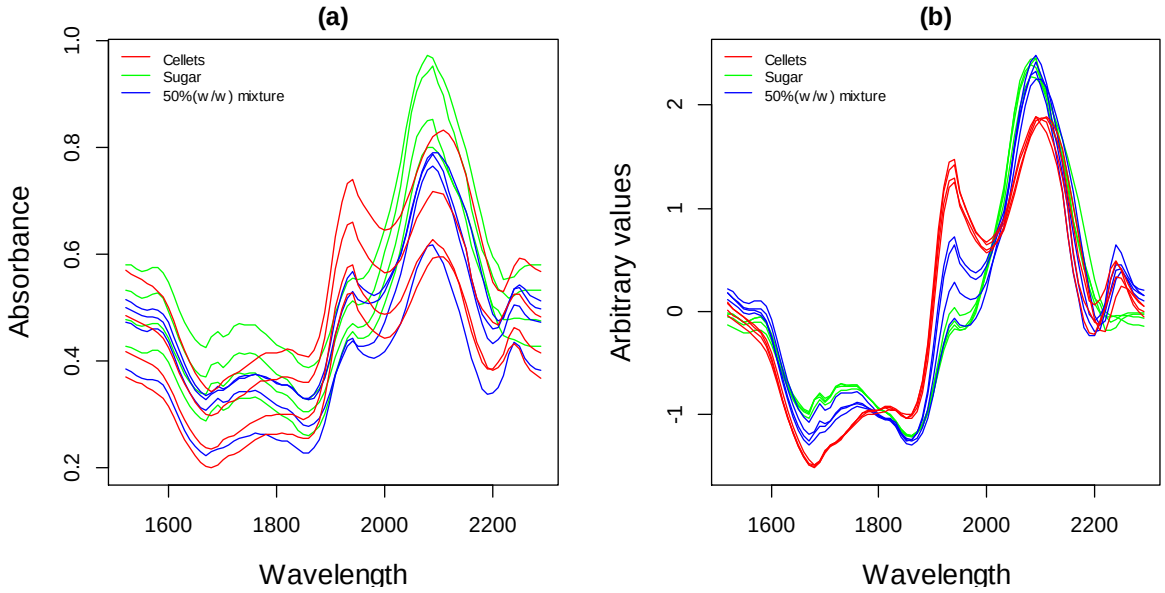


Figure 23 (a) Raw spectra of the four probes (b) Spectra data of the four probes after applying SNV pre-processing methods.

4.1.3 Content prediction models in static conditions

Figure 24 (a) shows the percentage content of cellulose plotted against the pre-processed spectral data at wavelength 1930nm. In the same figure, a linear relationship between the content and the pre-processed absorbance values at 1930 nm is established using a simple regression line. PCA was applied on the spectra of the three samples. Figure 24 (b) shows principal component one plotted against principal two, the figure shows that PCA was able to cluster the samples based on their spectra.

All the spectra data can be used to predict the cellulose content instead of using only spectral data at wavelength 1930 nm. This is performed using partial least square on the pre-

processed data of the three samples. Figure 24 (c) and (d) show the results after applying PLS, validation was carried out using the leave-one-out option, these figures show the prediction of the cellulose content using two and five components. Using two components PLS can predict the content with a root mean square error of prediction (RMSEP) of 14%. Using five components the error decreases to 10%.

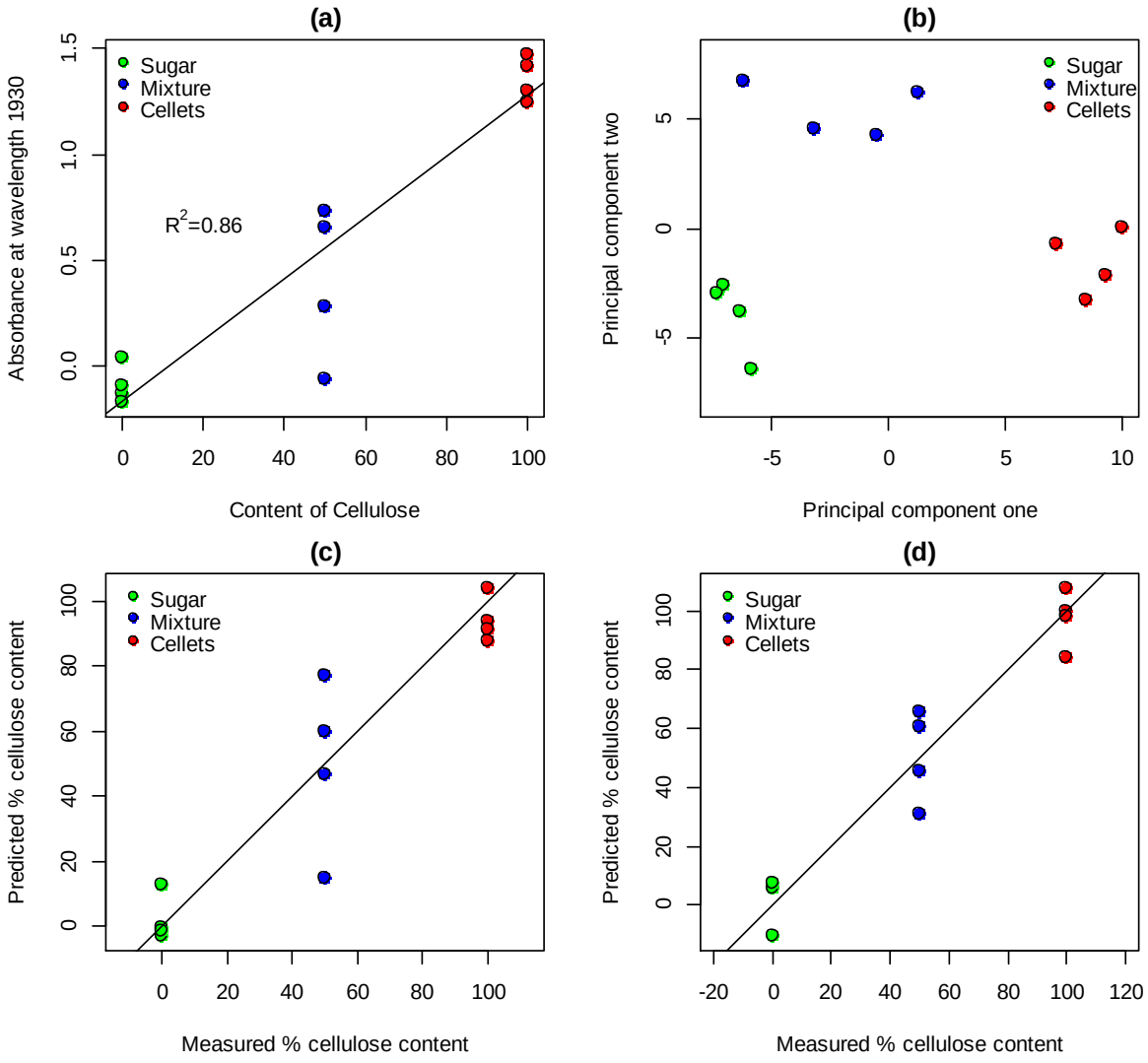


Figure 24 (a) Arbitrary values of the three samples at wavelength 1930 in static conditions; **(b)** Principal component analysis applied on the whole spectrum; **(c)** PLS using two components RMSEP=14%; and**(d)** PLS using five components RMSEP=10%.

4.1.4 NIR spectrum of cellulose and sugar spheres in dynamic conditions

To evaluate the ability of the MultiEye^{TM} to measure accurately the spectrum during particle movement, the same samples were placed on a rotating device which rotated with a speed of 80~90 rpm, as described in section 3.4. The four probes of the MultiEye^{TM} were fixed above the rotating device at about 0.5~1 mm from the sample. The raw spectra of Cellets®, sugar spheres and 50% (w/w) mixture obtained from the four probes are shown in Figure 25 (a). The spectra are noisy compared to the spectra under static conditions. It should be noted that the raw spectra of the four probes is approximately equal to each other.

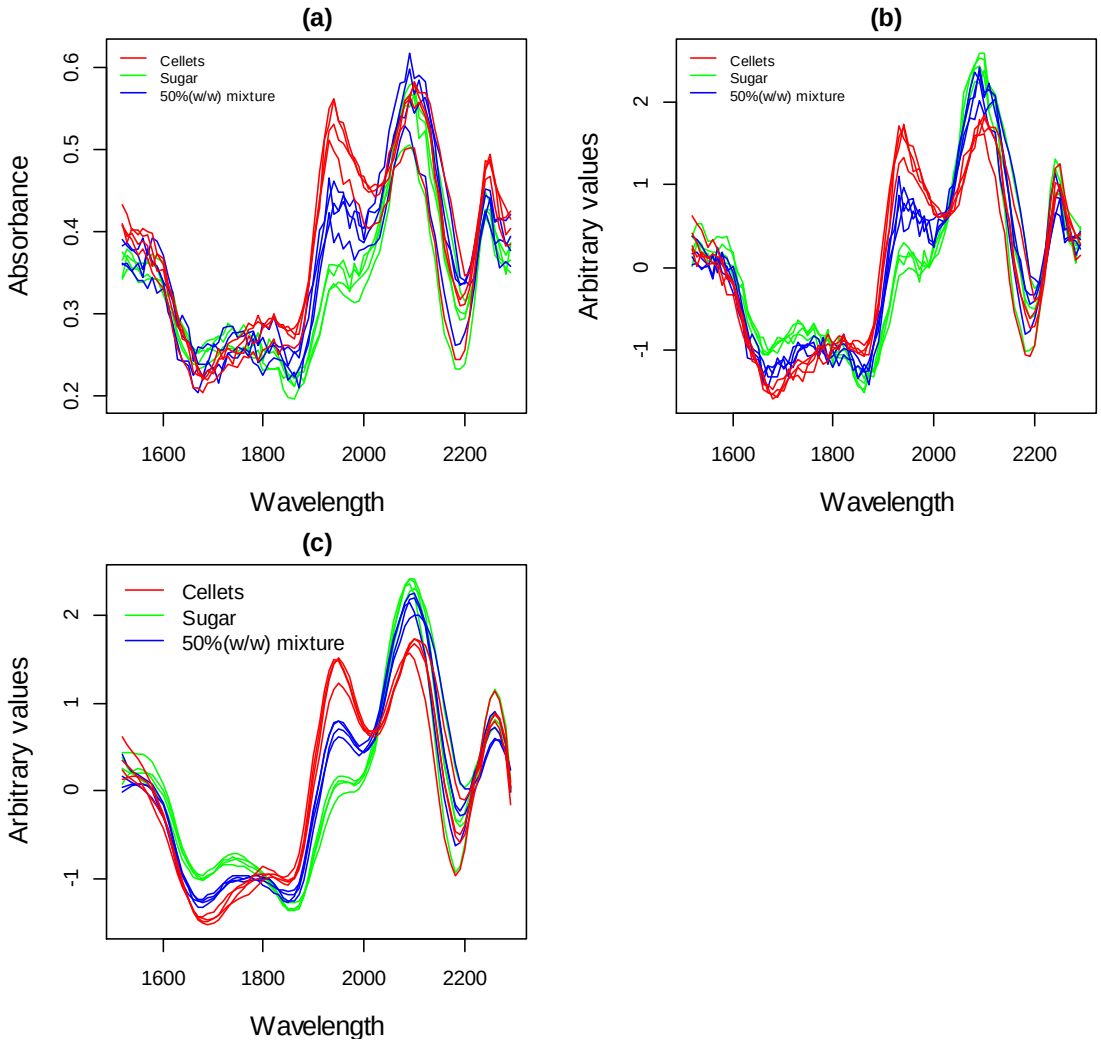


Figure 25 (a) The raw spectra of Cellets® under dynamic conditions obtained from MultiEye™ four probes (b) Spectra data of the four probes after applying SNV pre-processing methods and **(c)** Spectra data of the four probes after applying Savitzky-Golay method.

Figure 25 (b) shows the spectra of the four probes after applying the standard normal variate pre-processing method. In the case of particle movement the baseline shifts is less apparent that the static case so applying SNV corrects slightly the spectra of the probes.

The same figure shows the mixture spectra is between the spectra of the pure Cellets® sample and the spectra of the pure sugar sample. Standard normal variate removed the baseline shift but it did not smooth the spectra. To smooth the spectra, Savitzky-Golay with window length equal to 13 is used, Figure 25 (c) shows the resulting smoothed spectra. The same figure shows that all the samples spectra have a varying peak at wavelength 1930 nm.

4.1.5 Content prediction models in dynamic conditions

The content of cellulose is plotted against the absorbance of the spectra at wavelength 1930 nm obtained from the four probes as shown in Figure 26 (a). Similar to the static condition, a linear relationship between the content and the pre-processed absorbance values at 1930 nm is established using a simple regression line.

Partial least square was applied on the SNV and Savitzky-Golay smoothed data of the three samples. Figure 26 (a) and (b) show the result after applying PLS and validation was carried out using leave one out option, these figures show the prediction of the cellulose content using two and five component models respectively. In the two components model, PLS predict the content with a root mean square error of prediction equal to 6% while for five components the error is equal to 4.3%.

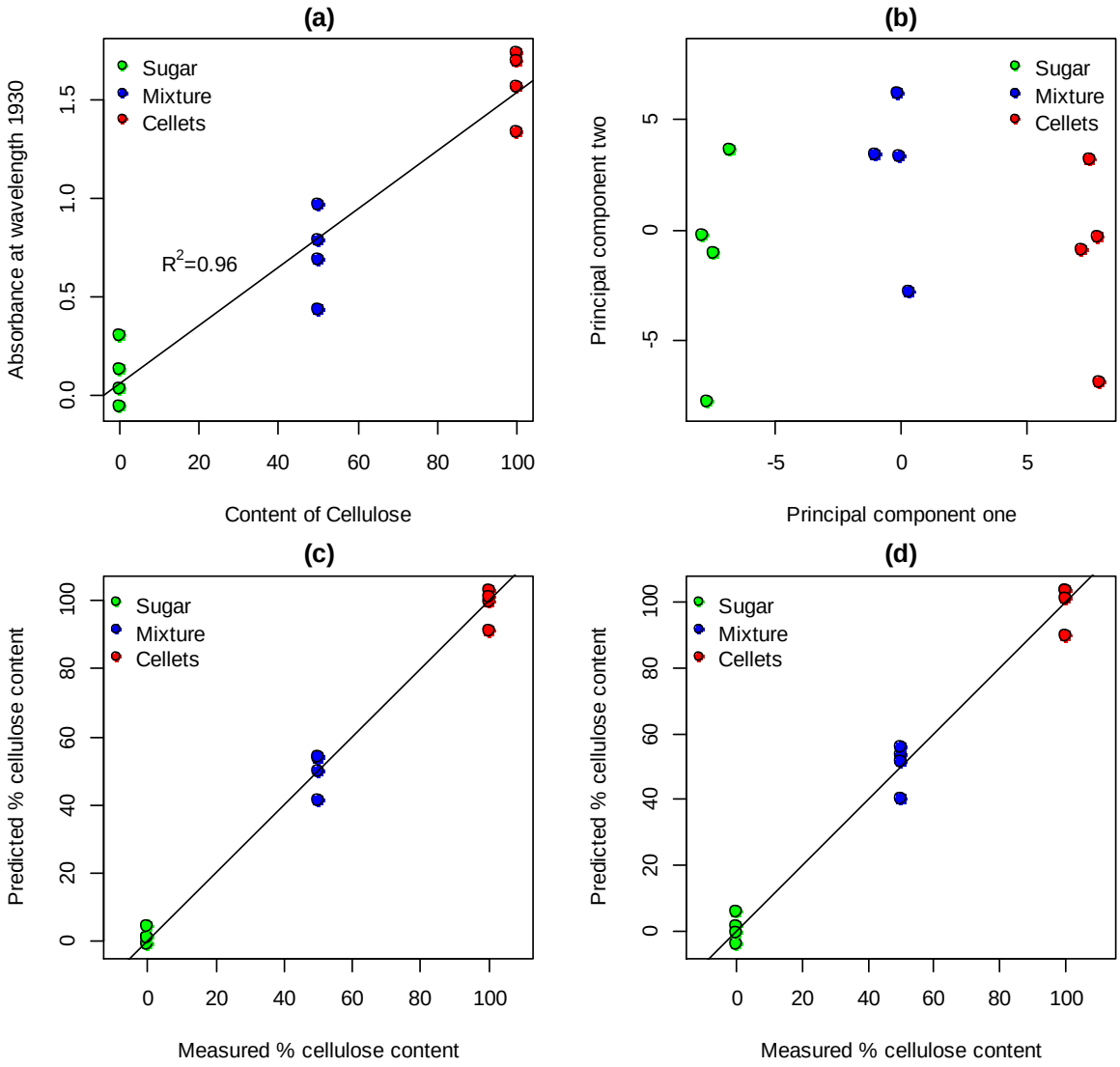


Figure 26 (a) Arbitrary values of the three samples at wavelength 1930 in static conditions; (b) Principal component analysis applied on the whole spectrum; **(c**) PLS using two components RMSEP=6%; and (d) PLS using five components RMSEP=4.3%.

PCA was applied also on the spectra of the three samples. Plotting principal component one against principal component two yields the plot in Figure 26 (b). PCA was able to classify the samples based on their spectra which reflect the cellulose content.

4.1.6 Real time measurement of blend homogeneity

As explained in section 3.4.3 a calibration model is constructed from the measurement of spectra of five samples containing different proportion of sugar and Cellets®. The plot in

Figure 27 indicates that four components are sufficient to explain most of the variation in the data. So for the in-line measurement four components were used to predict the proportion of Cellets®. To test the calibration model 10g of sugar were placed initially in the vortex shaker device that vibrate with circular motion that has a speed of 500 rpm then 10 g of Cellets® were added after 1000 seconds to the vibrating unit to create 50% (w/w) mixture.

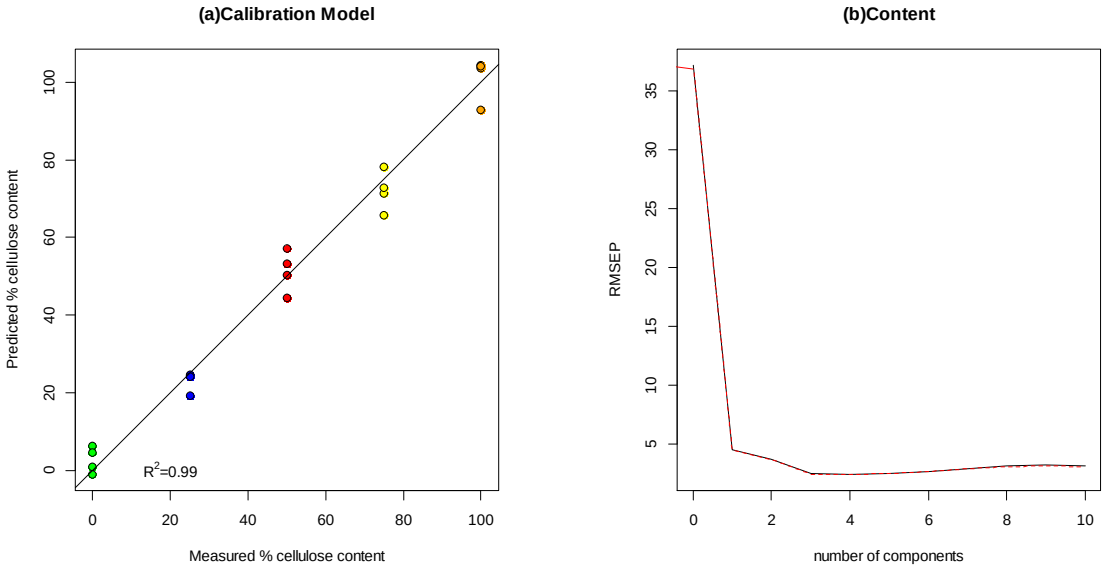


Figure 27 (a) Calibration model resulted from PLS using ten components; (b) The number of component vs the RMSEP.

Figure 28 shows the online interface that contains three plots, the first plot is a plot of the real time measurement of the four probes' spectra and the second are the spectra after applying SNV with smoothing and the last plot shows the inferred content of Cellets®, based on the calibration model constructed earlier, over the whole experimental period. The model performed well in this case as it indicates that there is no Cellets® in the sample. As the 10 g of Cellets® is added, the model responded immediately and the measured Cellets® predicted content by the four probes jumping to a level higher than 50% before it stabilise between 30%-60%.

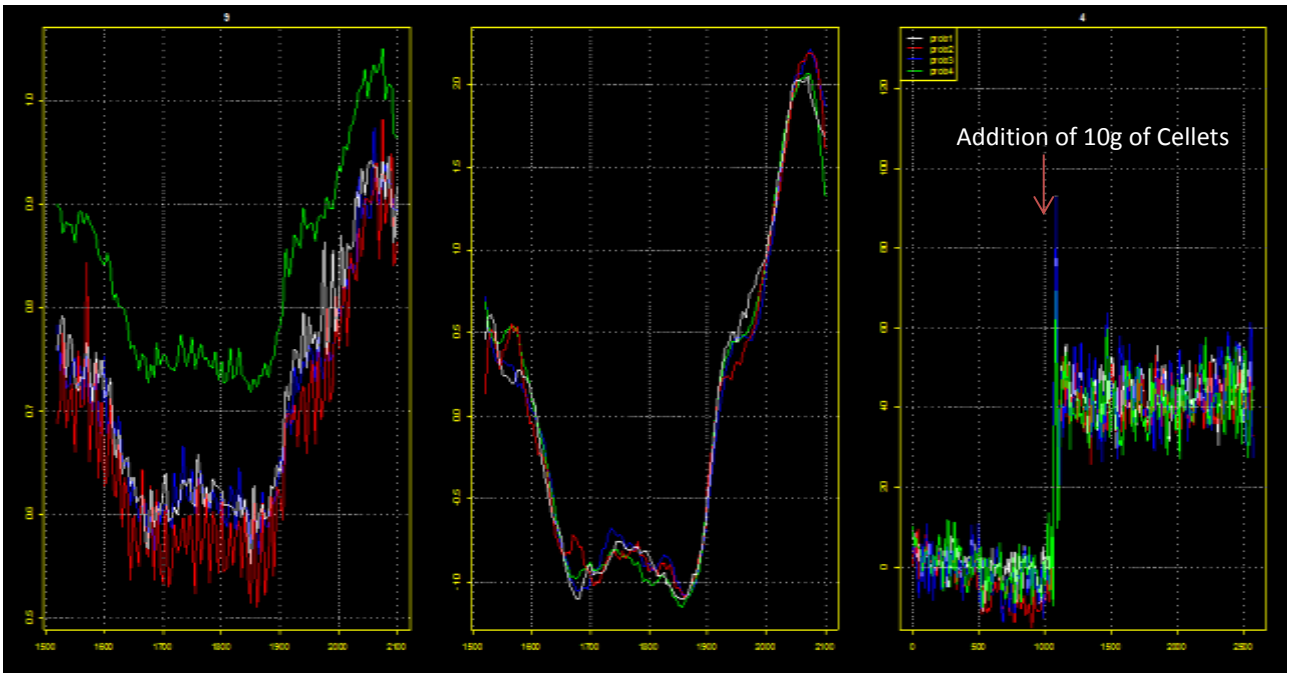


Figure 28 Proportion content tests for 50% (w/w) mixture of sugar spheres and Cellets® of size 500 µm.

The second experiment was performed by placing 15g of sugar in the stainless steel container shaking with circular motion of 500 rpm with the previous PLS model used to predict the content of Cellets®. Figure 29 shows the evolution of the estimated content of Cellets® in the sample over 2000 seconds. The first plot in Figure 29 represents the real time measurement of the four probes' spectra so in this case it represents the measurement at 3000 seconds. The spectra are noisy due to the movement of particles, SNV was used to remove the baseline shift and then the noise was reduced using the Savitzky-Golay method and the resulting spectra are displayed in the second plot. The smoothed spectrum was used to predict the content of Cellets® in the sample for each probe before updating the predictions in the last four probes. The predictions are displayed in the last four plots. Initially, between time 0 and time 700 seconds, all the probes indicate a level of content between 15-20%. At time 700 seconds 5g of Cellets® is introduced to the container to make a 75% sugar and 25% Cellets® mixture.

This change in the sample chemical composition is sensed by the probes after 50 seconds and the prediction of Cellets® content in each probe oscillates between 20%-30%.

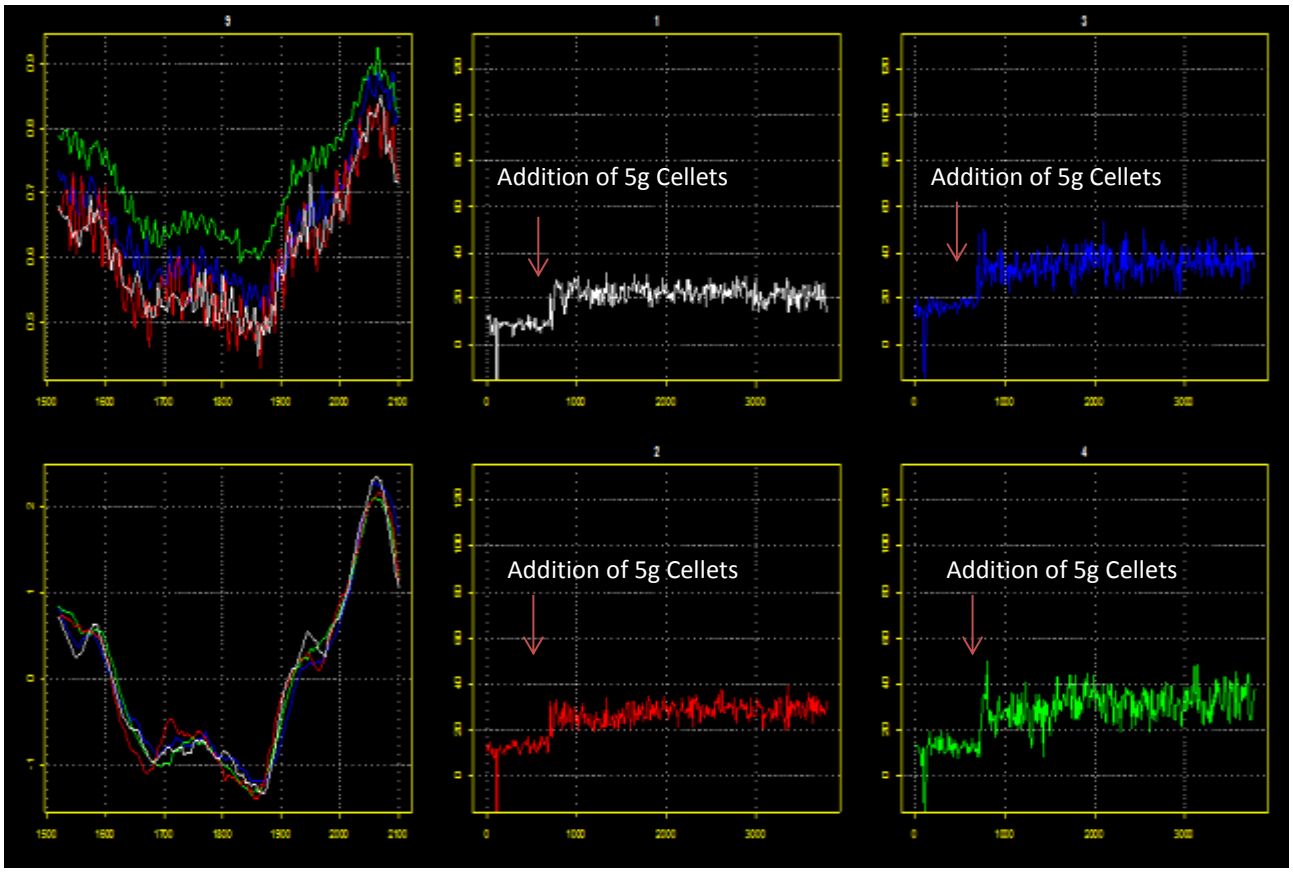


Figure 29 Proportion content tests for 25% Cellets® and 75% sugar spheres, the left plots are the instantaneous measured spectra the top is raw spectra and bottom plot is the spectra after pre-processing the other four plots show the evolution of inferred Cellets® percentage content from the four probes during the entire experiment

The third experiment was performed by placing 15g of Cellets® in the stainless steel container shaking with circular motion of 500 rpm then the PLS model is used to predict the content of Cellets®. Here also, SNV was used to remove the baseline shift and then the noise was removed using the Savitzky-Golay method before estimating the levels of Cellets® content at each time step. The predictions of Cellets® content of the four probes over 3000 seconds are shown in Figure 30, initially all the probes indicates a level of content between 15-20%. At time 700 second 5g of sugar spheres is introduced to the container to make 75%

Cellets® and 25% sugar. This change in the sample's chemical composition is sensed by the probes immediately. In this case there was a difference between the probes predictions resulted from the PLS model. Probes have similar prediction oscillating between 75%-80% while the other two probes oscillate between 70%-90%.

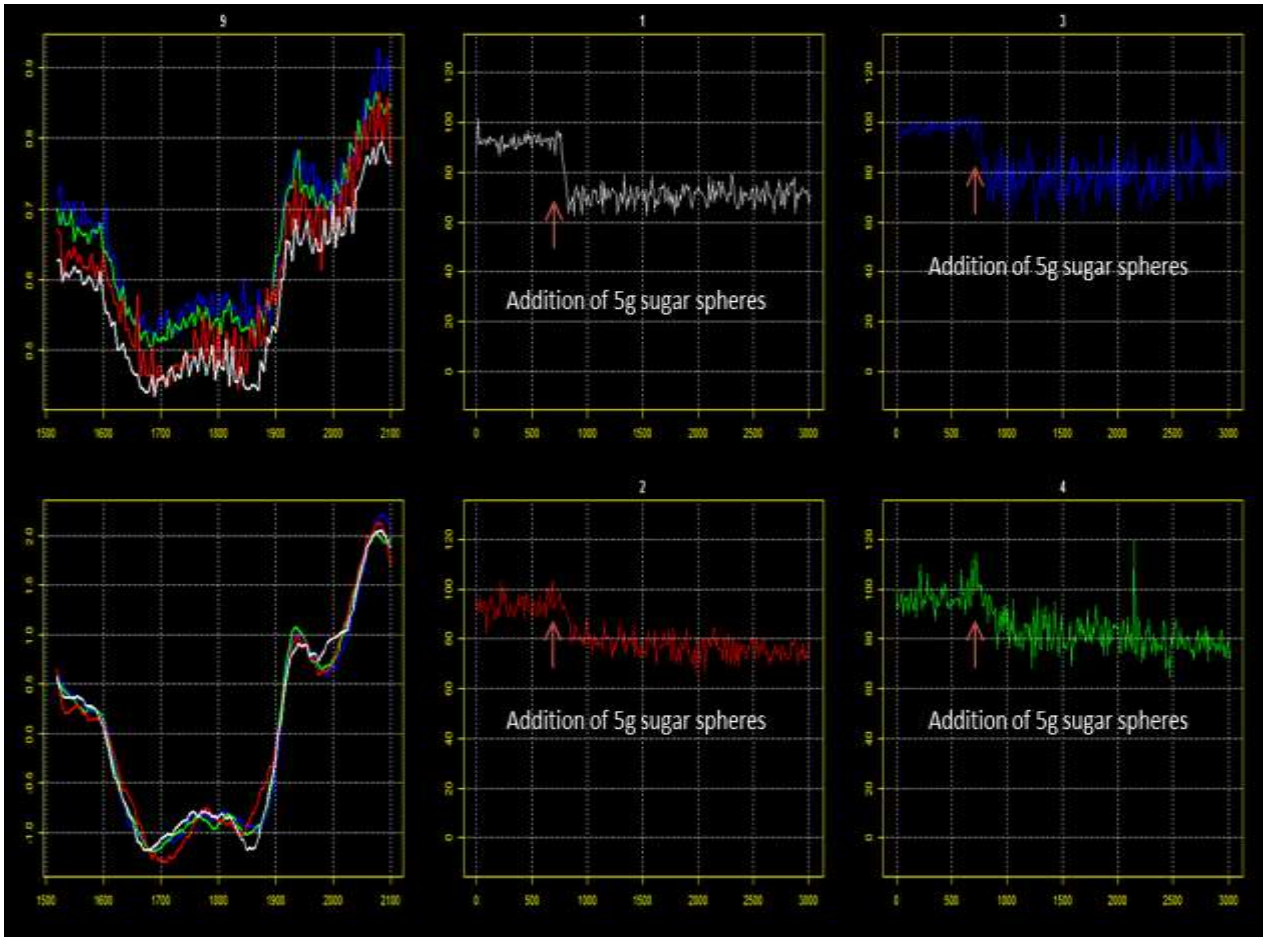


Figure 30 Proportion content tests for 75% Cellets® and 25% sugar spheres. , The left plots are the instantaneous measured spectra the top is raw spectra and bottom plot is the spectra after pre-processing the other four plots show the evolution of inferred Cellets® percentage content from the four probes during the entire experiment

4.1.7 Effect of particle size on the blending process

The in-line measurement of Cellets® content during the mixing process with sugar spheres is repeated for 50% (w/w) proportion for different particle size ranging from 100 µm to 1000 µm. As described previously, the first step is to construct a calibration model by taking

measurement of different spectra of the following proportion of Cellets® content 0%, 25%, 50%, 75% and 100% (w/w) with a total weight of 20g. The validation was performed using the leave-one-out method. Figure 32 shows the different calibration models using four components for different size of particles.

Prior to the in-line measurement of Cellets® content, 10g of Cellets® was placed in one end of a cylindrical container connected to a vortex mixer and 10g of sugar spheres were placed on the other end as shown in Figure 31.

Figure 31 mixture of 50% (w/w) Cellets® and sugar before the mixing.

For the in-line measurement of 100 µm the probes were placed above the middle of the cylindrical container about 1-3mm far from the particle samples. The speed of the vortex shaker was limited to 500 rpm. Initially, the predicted Cellets® content was noisy and fluctuated between 0% and 100%, but after 500s the inferred content of Cellets® fluctuation interval was between 55% to 75%, 62% to 78%, 40% to 65%, and 50% to 70% for probe

1,probe 2, probe3, and probe 4 respectively (Figure 33).

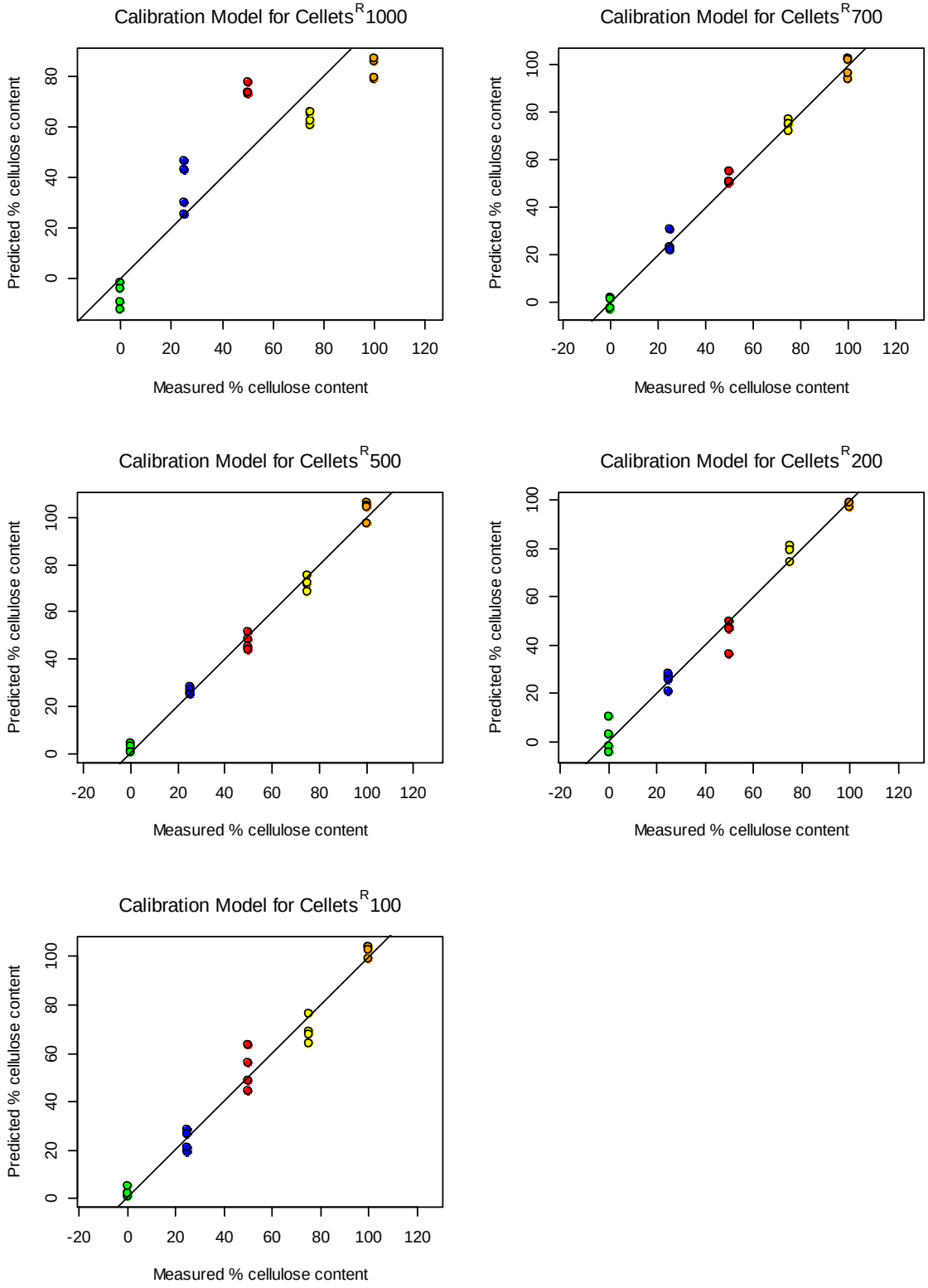


Figure 32 Calibration models for (a) Cellets® 1000 µm (b) Cellets® 700 µm (c) Cellets® 500 µm (d) Cellets® 200 µm and (e) Cellets® 100.

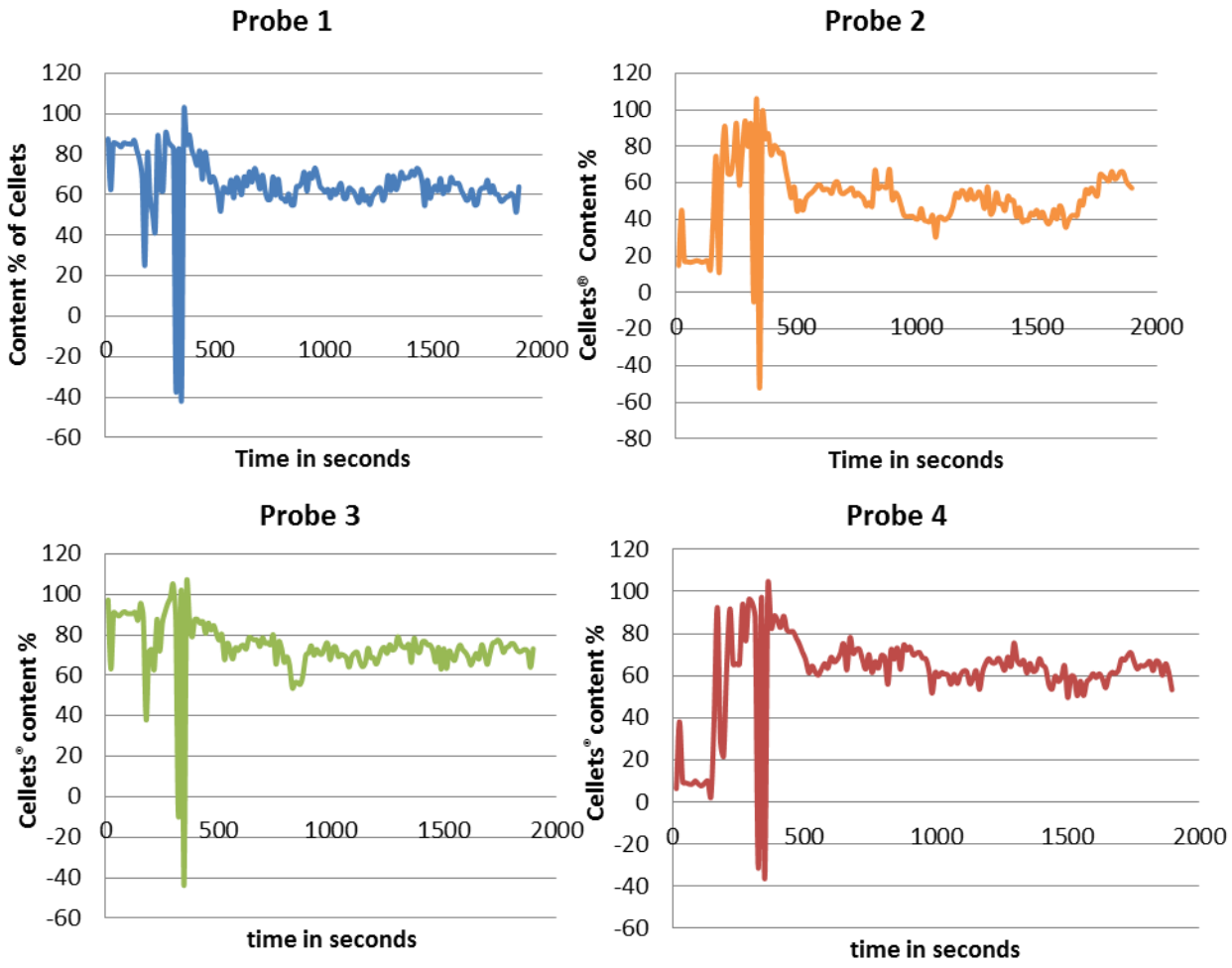


Figure 33 Summary of in-line experiment of 100 µm particle size with content proportions 50% (w/w).

For the second experiment 10 g of Cellets® and 10 g of sugar spheres having nominal particle size of 200 µm were placed in the cylindrical container connected to the vortex shaker at the same manner as the previous size. This time the probes initially where above sugar sample near edge of the cylindrical container while the vortex shaker is turning at speed of 100 rpm. At this speed the sugar and Cellets® did not blend, so all the probes initially indicate a low level of Cellets®, after 100 second the probes where moved to above the centre of the container and the speed of the vortex is increased to 500 rpm, the Cellets® start to mix with the sugar spheres. The model responded by indicating an increase in Cellets® content. The

fluctuations interval of the inferred Cellets® content were 10% to 70%, 0% to 40%, 10% to 55%, and 2% to 36% for probe 1, probe 2, probe 3 and probe 4 respectively. After 500 seconds the fluctuations interval width decreases to between 0% and 20% for all the probes except for the fourth probe that detect a maximum of 30% in Cellets® content.

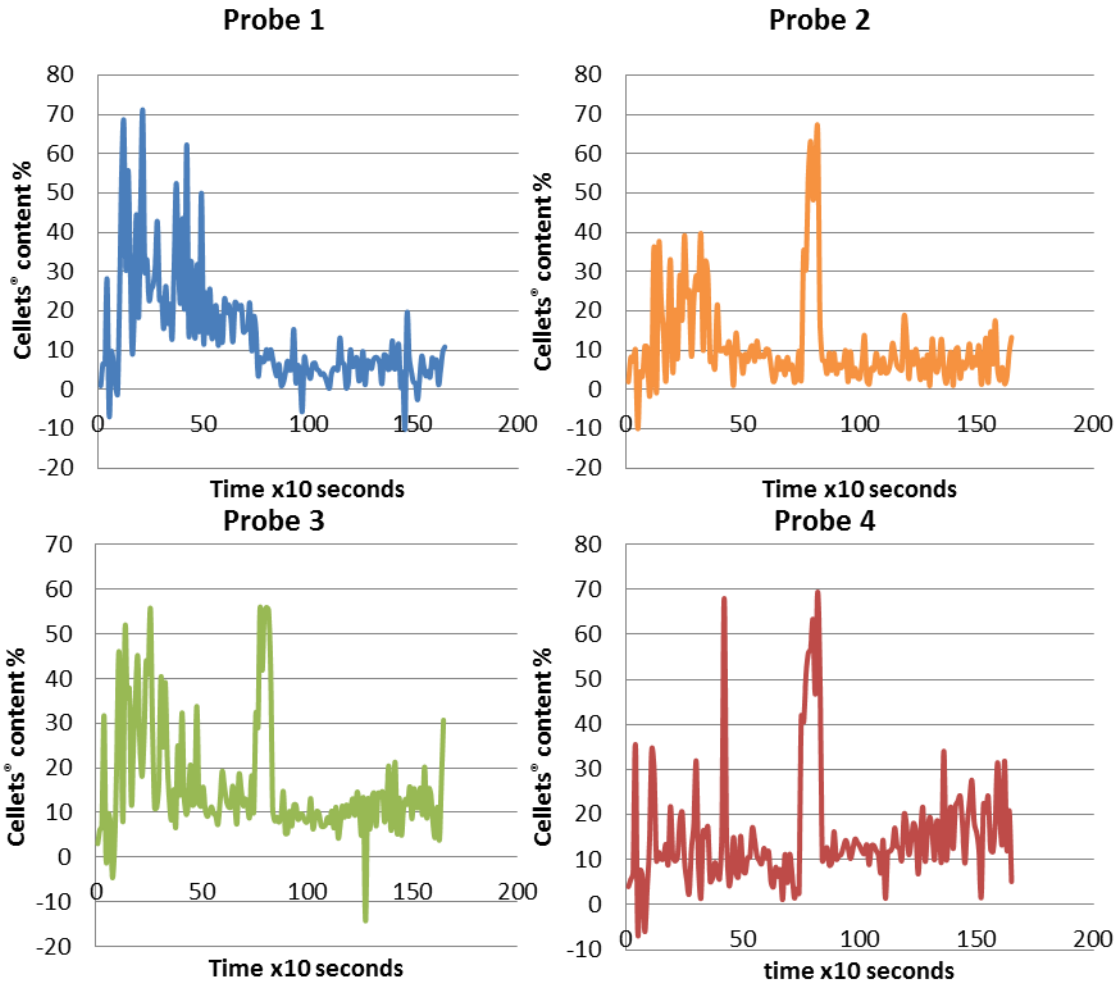


Figure 34 Summary of in-line experiment of 200 µm particle size with content proportions 50% (w/w).

An image of the mix after the blending process is shown in Figure 35, the sugar sphere colour is whiter than Cellets® so the image shows a whiter circle in the centre of the container and darker colour on the edge of the container suggesting that sugar sphere segregated into the centre of the container while Cellets® segregated on edge of the container. This may be due

to slight particle size differences or other properties. Normally Cellets® 200 µm have wider size interval ranging between 200 and 355 µm while for sugar spheres the range is between 212 and 250 µm. So even slight difference in particle size can affect the blending process.

Figure 35 Mixture of 50% (w/w) sugar spheres and Cellets® of nominal size 200 µm at the end the of the blending.

The next mixture to blend was Cellets® and sugar sphere with nominal size 500 µm, it was mixed in the same way as the previous particles by placing 10g of sugar spheres side by side with 10g of Cellets®. The probes were initially placed at the same side as the sugar sphere.

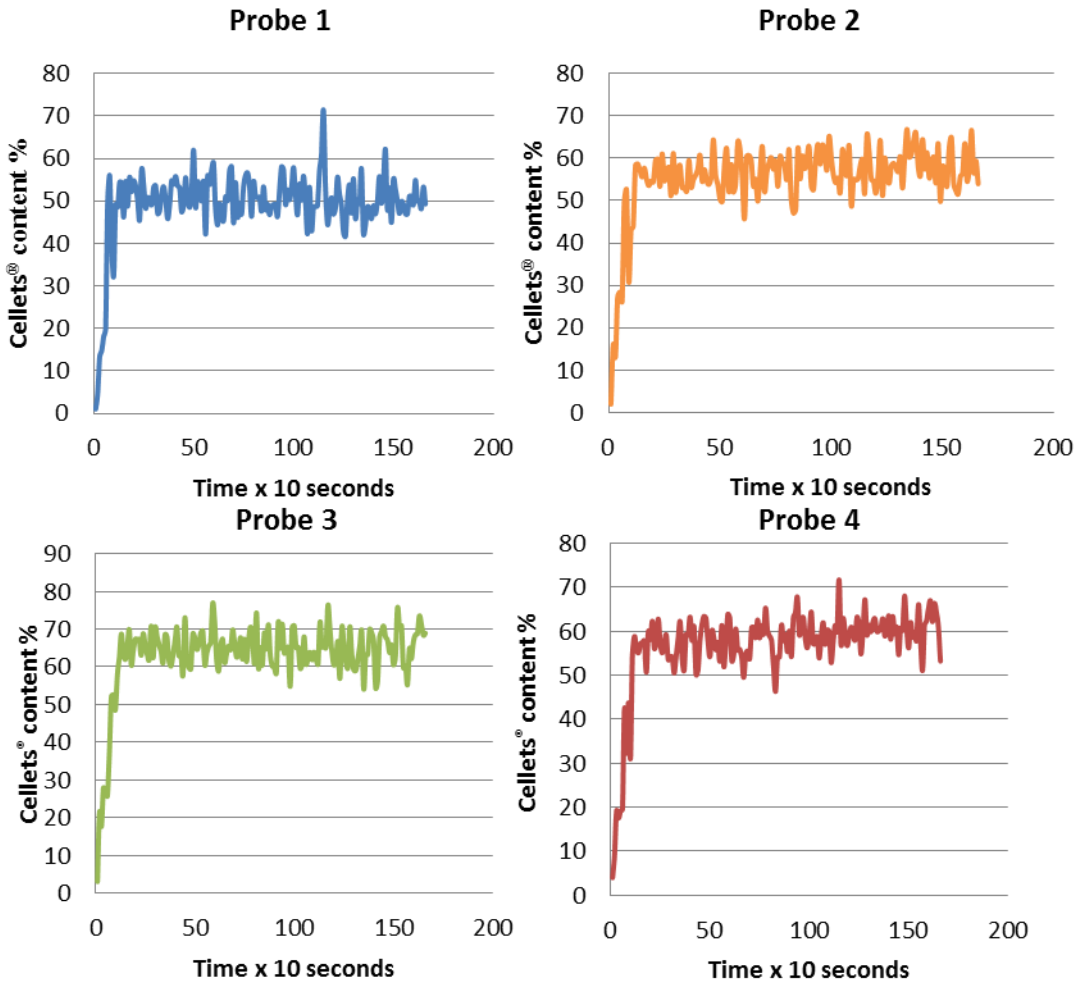


Figure 36 Summary of in-line experiment of 500 µm particle size with content proportions 50% (w/w).

In the case of mixing the blend of particle nominal size 500 µm, the predicted content of Cellets® stabilised quickly compared to the previous particle size. The predicted content of Cellets® stabilises in about 100 seconds at a concentration of around 50%-70% depending on the probe, as shown in Figure 36.

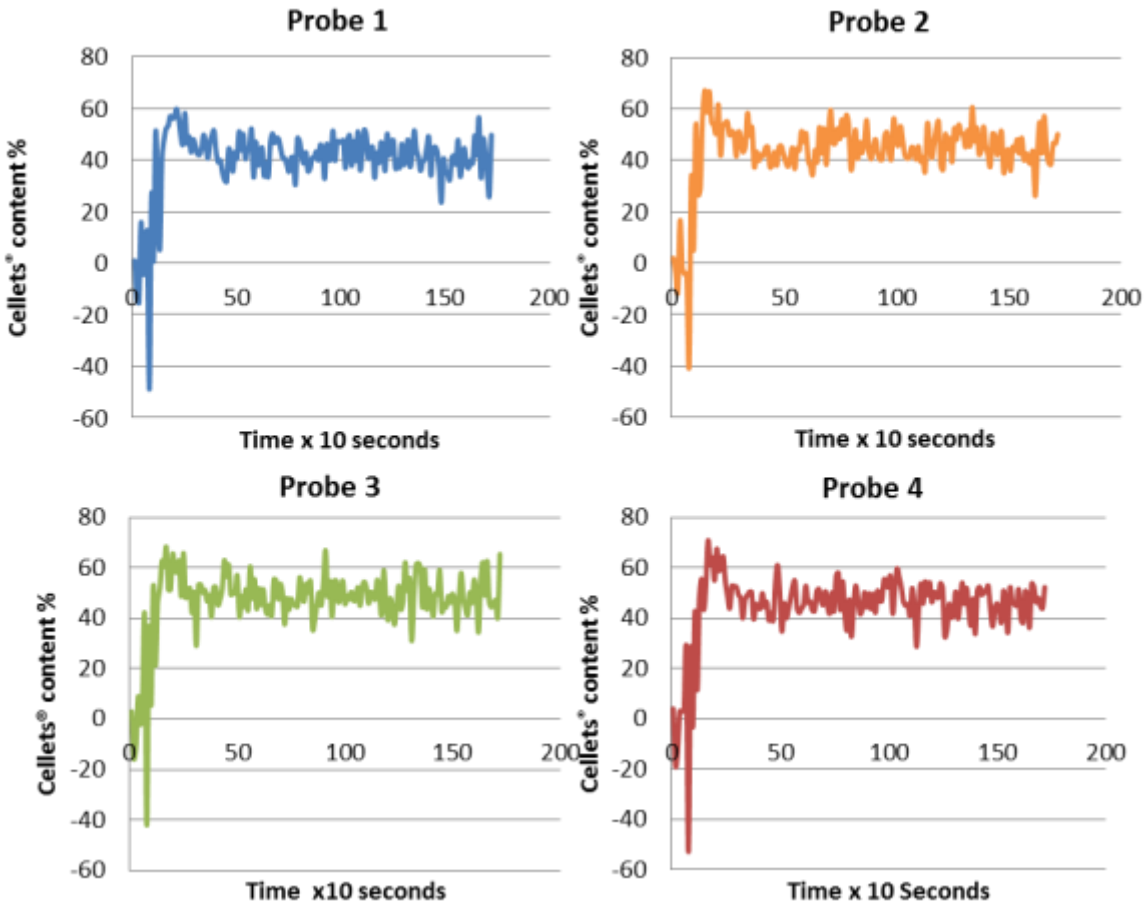


Figure 37 Summary of in-line experiment of 700 µm particle size with content proportions 50% (w/w).

The experiment was repeated for the particle nominal size 700 µm, the summary of the experiment is presented in Figure 37. As the probes moved to the centre of the container and the speed of the vortex shaker was increased to 500 rpm to start the blending process, the predicted Cellets® content increased immediately. The interval of the fluctuations of the predicted Cellets® content stabilise after a 100 seconds from the start of the blending process.

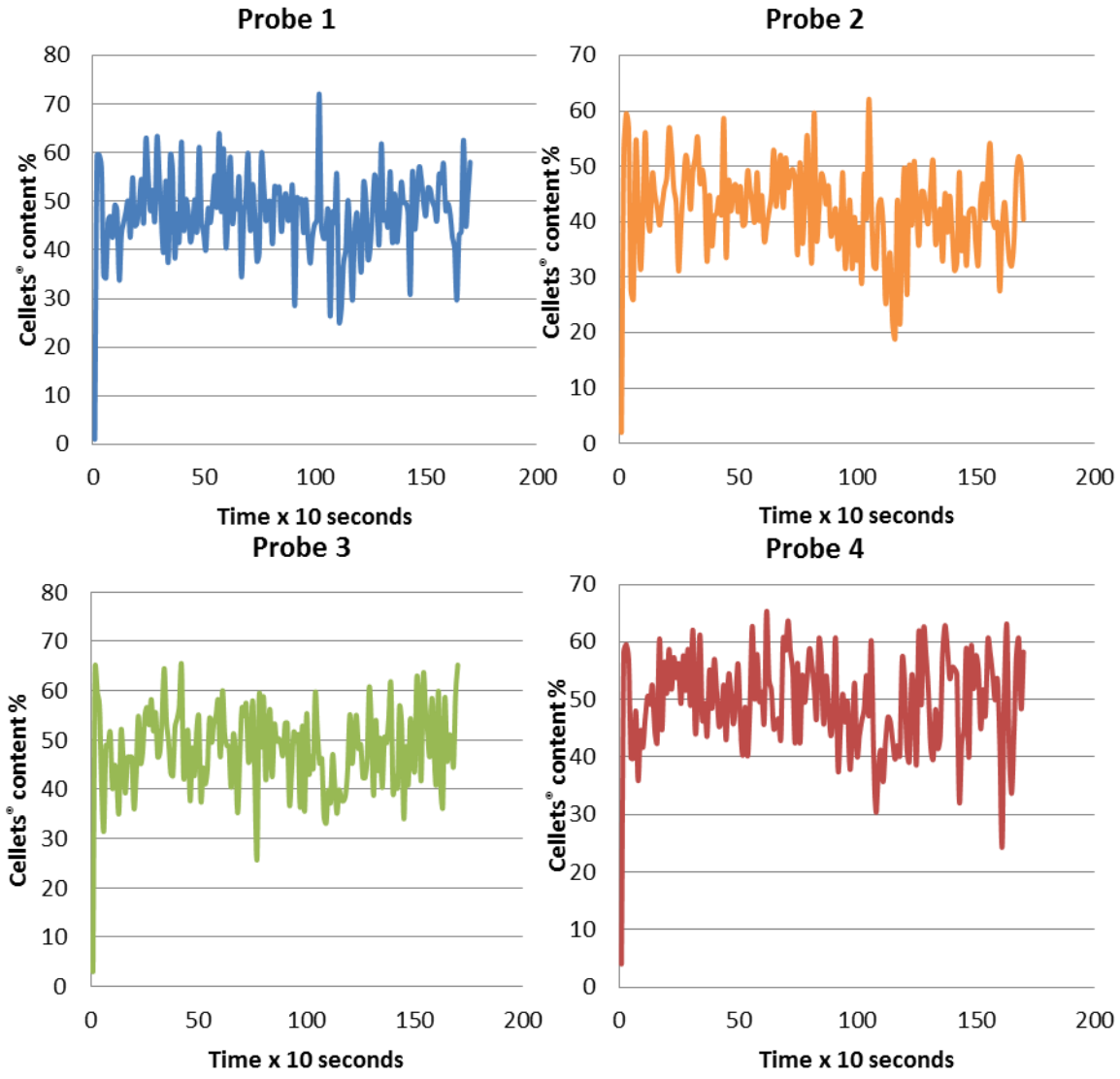


Figure 38 Summary of in-line experiment of 1000 µm particle size with content proportions 50% (w/w)

Finally, the same experiment is performed for Cellets®1000 and the results are presented in Figure 38. The figure shows that all the probes detected the change in the content; the predicted Cellets® content fluctuate between 30% and 65% for all the probes which is larger when compared to the previous experiments.

4.2 Physical properties

4.2.1 Particle size in static conditions

4.2.1.1 One layer of particles

Six samples of different size Cellets® were prepared. Each sample was placed in a sample holder as a monolayer of particles prior to the imaging step. In total, six images were taken by the Eyecon™ corresponding for each of the six size categories. The size was estimated using the image segmentation analysis algorithm described in section 3.5.2. Table 4 presents a summary of the estimated average per nominal size category.

Table 4 Particle size categories versus average diameter estimates.

Nominal size of particles (µm)	Estimated average diameter(µm)
1000	1022
700	845
500	541
350	403
200	316
100	164

A linear regression model was applied to the data in Table 4 and a R^2 value of 0.98 was obtained. Figure 39 shows a good linearity between the nominal size and the estimated average diameter.

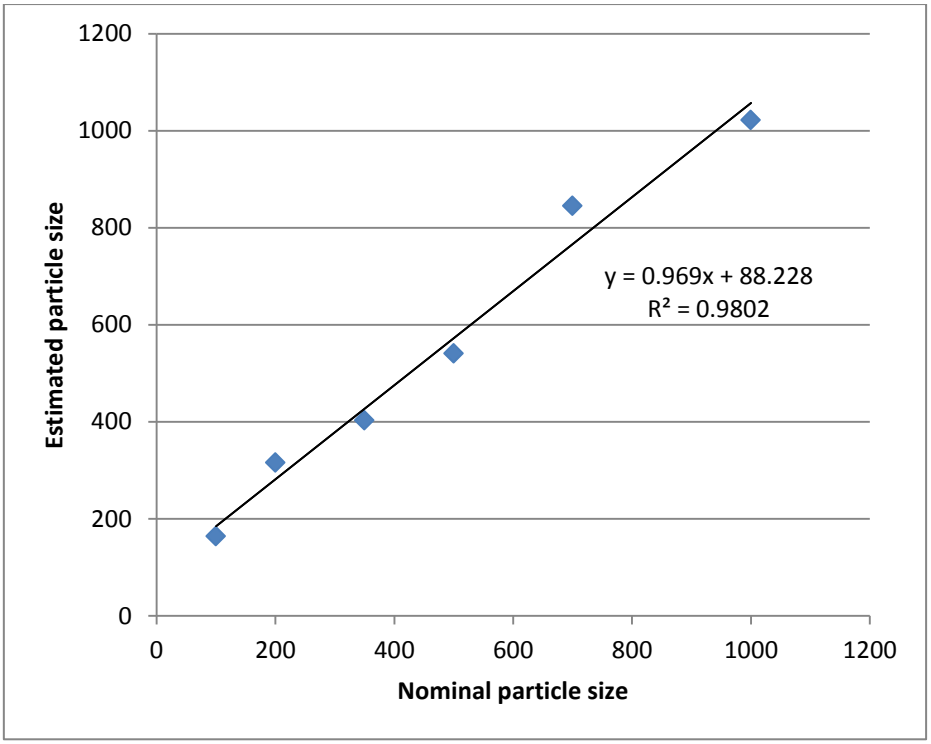
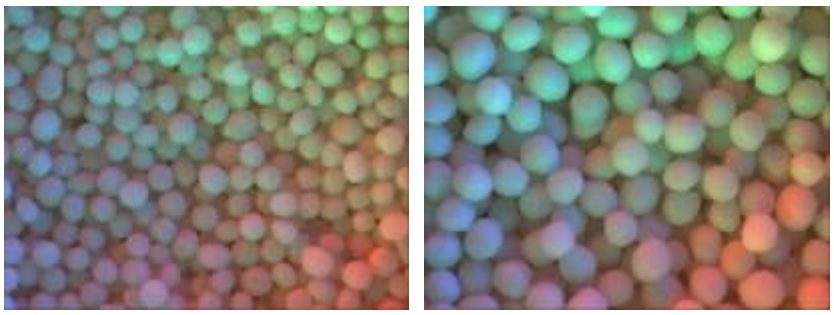


Figure 39 Six particle size categories versus the average diameter established via image analysis of the Eyecon™ images.

4.2.1.2 Bulk powder particle size analysis

To analyse the size of particles in a bulk powder bed, five samples were taken from five different nominal size categories of Cellets®. The samples were placed in a tray in such a way as to have images of the bulk powder. Figure 40 shows the bulk particle images taken by the Eyecon™. The size analysis of bulk powder using image segmentation algorithm is time consuming due to the overlapping particles and also due to the high number of particles in each image.



(a) **(b)**

Figure 40 Eyecon™ images of bulk powder particles of nominal size: (a) 200μm and (b) 700μm.

Since there is no background in this case, particles with low light intensities, usually particle at the bottom of the container, are considered as background by the algorithm so this will increase the value of the threshold. A high threshold means that the algorithm will underestimate the size of particles because the extremities of a particle have relatively lower light intensity than its centre, so the algorithm will erode particles at their extremities leading to under estimation of particle size. This can be noticed in Figure 41, even though several particles that are located at the top of the bulk powder had high intensity pixels, the algorithm removed several pixels from those particles to include them in the background at the extremities and thus reducing the particle size. So, instead of measuring the particle size directly from image segmentation, texture analysis of bulk powder can be studied to infer the texture of a bulk particle surface to a specific size.

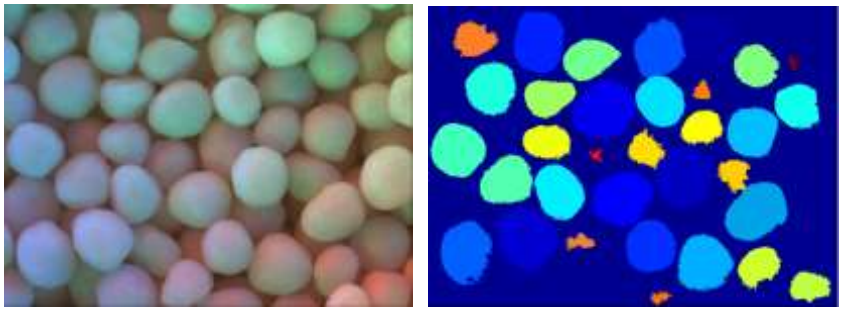


Figure 41 The resulting image after applying the image segmentation algorithm.

Texture analysis using the variogram function was obtained from each image and was analysed to establish a relationship between texture of the image and size of the particle. Unfortunately due to limitation on the memory size of the R package, it is not possible to use the built in variogram function for calculating the variogram on the image directly. So a pre-processing technique is used in order to reduce the size of the image from 1037x1387 to 129x174 pixels before the texture analysis step.

Figure 42 shows the variogram functions of the five nominal size categories Cellets®. The variogram functions of 700 and 1000 μm have higher values than the variogram functions of the other nominal size of Cellets®. The experiment was repeated for ten images for each nominal size. It can be observed that the values of the variogram function after lag four are ranked in accordance to the size of the particles. To establish a relation between the size of the Cellets® categories and the variogram function, theoretical exponential variogram parameters were estimated from each variogram function using the built in function in the GeoR package in R [78]. Figure 43 shows that the average of ten images sills for the five nominal sizes plotted against the corresponding nominal size. Applying least square regression to find a linear relationship between the sill of the theoretical variogram and the nominal size of the Cellets® and a value of R^2=0.95 suggest a positive linear correlation between the theoretical sill of the variogram and the nominal size of the Cellets®.

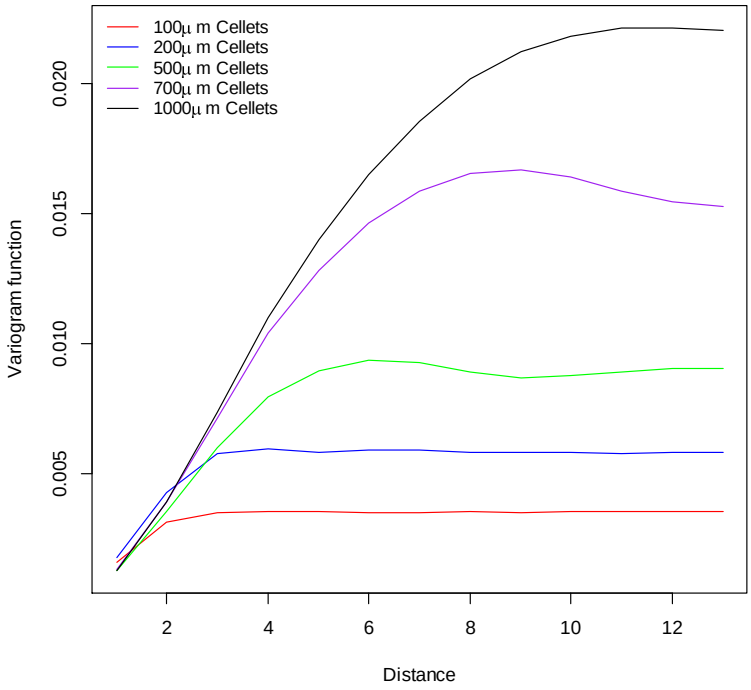


Figure 42 Variogram functions of bulk particle images corresponding to five different nominal Cellets® sizes.

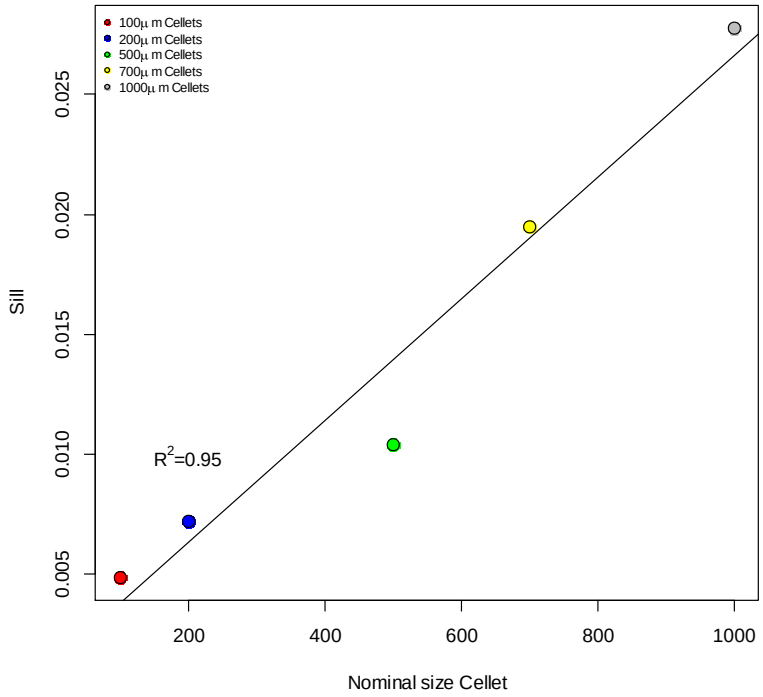


Figure 43 Linear relation between Nominal size of Cellets® and variogram sill

4.2.2 Particle size estimation in dynamic conditions

4.2.2.1 Conveyor belt

This experiment examined the Eyecon™ in conjunction with the image analysis methods presented in section 3.3 for determining the size of particles under motion. Three different size particles of nominal values 500, 700 and 1000 µm were used in this experiment. Ten samples each of the three nominal size categories were prepared hence a total of thirty images were acquired by the Eyecon™. This experiment was repeated under static, slow and fast speeds see Table .

Table 7 Summary of experiments for particle size estimation under different conditions.

	Static	Slow (3 cm/s)	Fast (23 cm/s)
Cellets® 500 µm	10 samples	10 samples	10 samples
Cellets® 700 µm	10 samples	10 samples	10 samples
Cellets® 1000 µm	10 samples	10 samples	10 samples

Figure 44 shows the results of calculating the mean particle diameters for the experiemental design in Table . It shows that the estimated mean diameter for each of the three nominal size Cellets® is approximately the same regardless of the state of the Cellets®. A one-way ANOVA was performed for each of the three nominal sizes to compare the estimated mean size showing no significant difference ($p<0.05$) between static, slow and fast conditions.

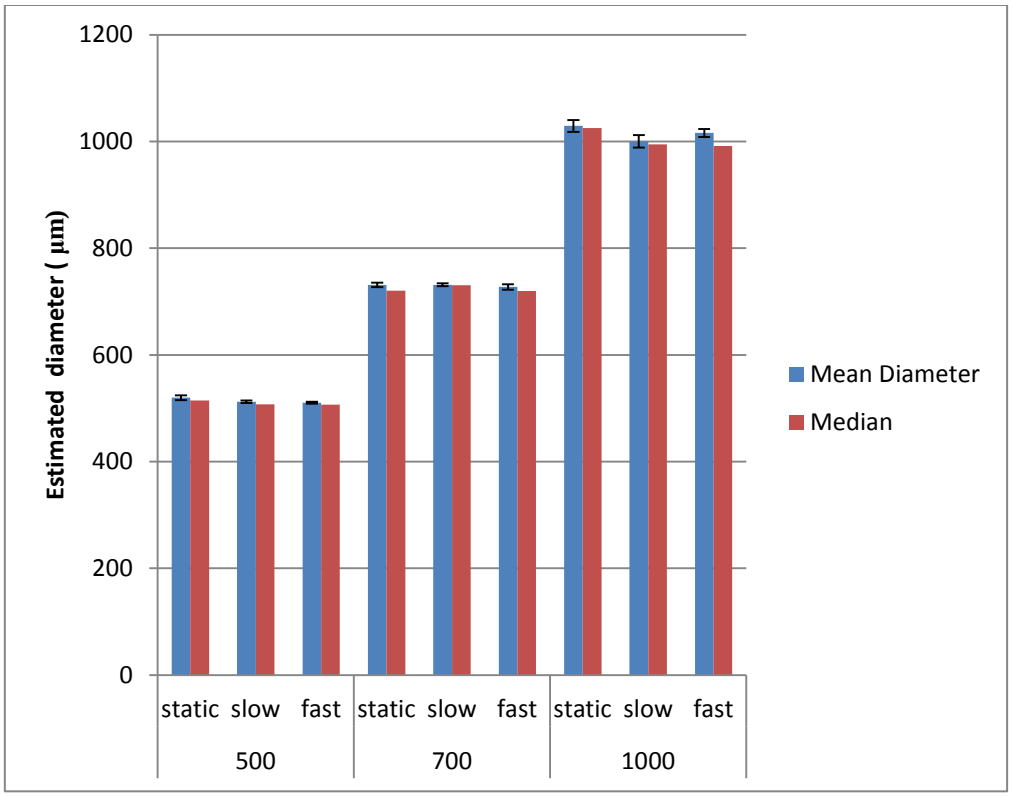


Figure 44 Estimated mean diameter for Cellets® of nominal size 500, 700 and 1000 µm in static and dynamic conditions.

Table 8 Estimated means of Cellets® in static and dynamic conditions, and the one-way ANOVA results between the means in static and dynamic condition for different size particles.

Cellets® Nominal Size µm	**Estimated Mean diameter at 0 cm/s**	**Estimated Mean diameter 3 cm/ s**	**Estimated Mean diameter 23 cm/s**	**one-way ANOVA (0 and 3cm/s)**	**one-way ANOVA (0 and 23 cm/s)**
500 µm	494.1	503.0	496.2	P>0.05	P>0.05
700 µm	731.2	731.4	727.4	P>0.05	P>0.05
1000 µm	1035.5	991.8	996.8	P>0.05	P>0.05

4.2.2.2 Free Fall

Using the experimental setup described in section 3.4, different size particles were fed between two glass windows and subsequently images were captured using the Eyecon™. Only one image per nominal size particle is considered. The challenging part in the free fall is that not all the particles were in the camera's focal plane, here the size of particles not in the focal plan might be either overestimated or underestimated depending on the position of particles with respect to the focal plan of the camera (Figure 45). Also particles can be partially hiding each other if they are in the same horizontal line as the camera which can mislead the image analysis algorithm that might consider the two particles as one and thus overestimate the average size of particles (Figure 45).

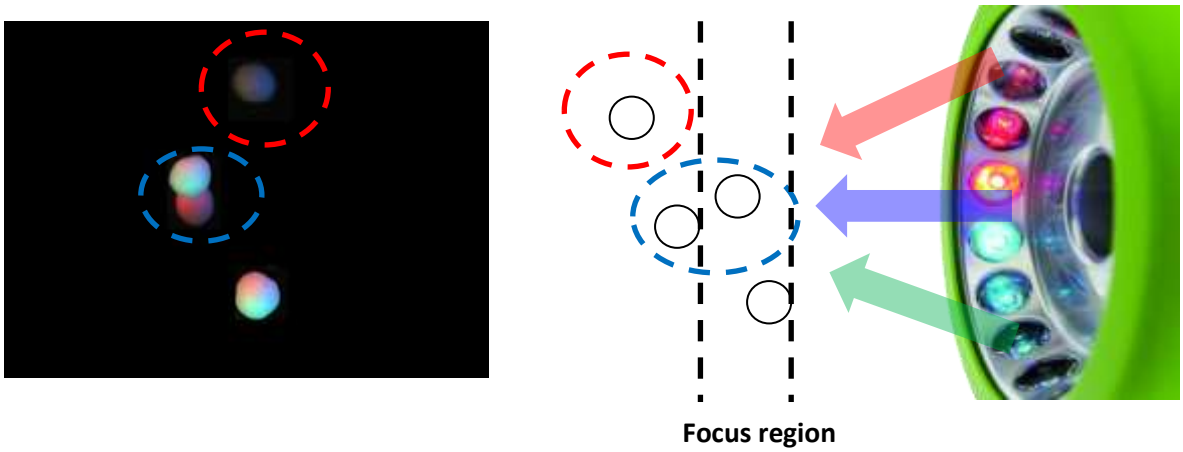


Figure 45 Challenges associated with free fall imaging.

The intensity of particles in focus and out of focus was compared to each other as shown by plotting the intensity level of the red line in Figure 46. The resulting plot shows that the maximum intensity of the particle in focus is higher than the particle out of focus. In addition, the intensity of the in focus particle has a spiky peak while the intensity of the blurry particle changes at a slower rate around the maximum intensity.

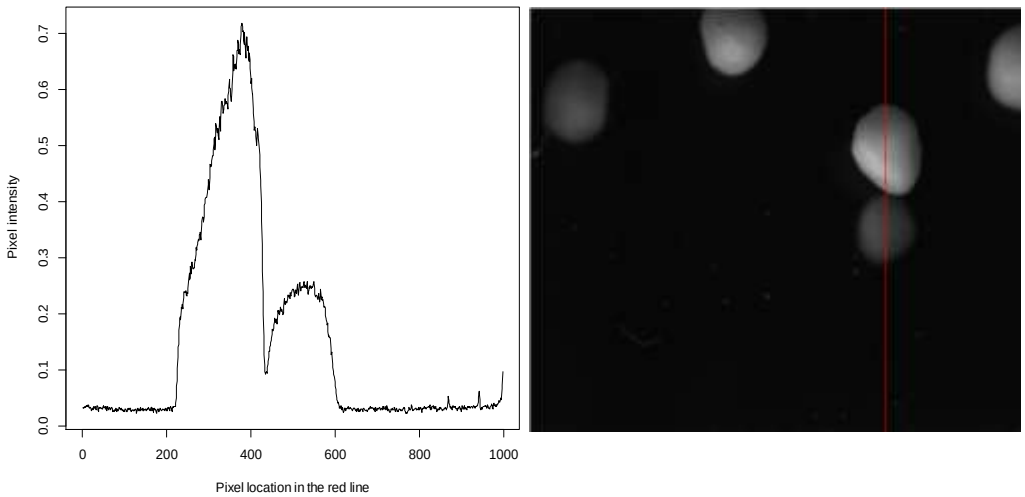


Figure 46 Plot of pixels' intensity of a vertical line in free fall greyscale image.

Size analysis was performed in similar manner as in the previous section. A linear regression model was applied and a R^2 value of 0.99 was obtained (Figure 47). The result shows a similar trend to the static case, although in the free fall state the estimated size is slightly smaller than the nominal size which may be due to the out of focus particles.

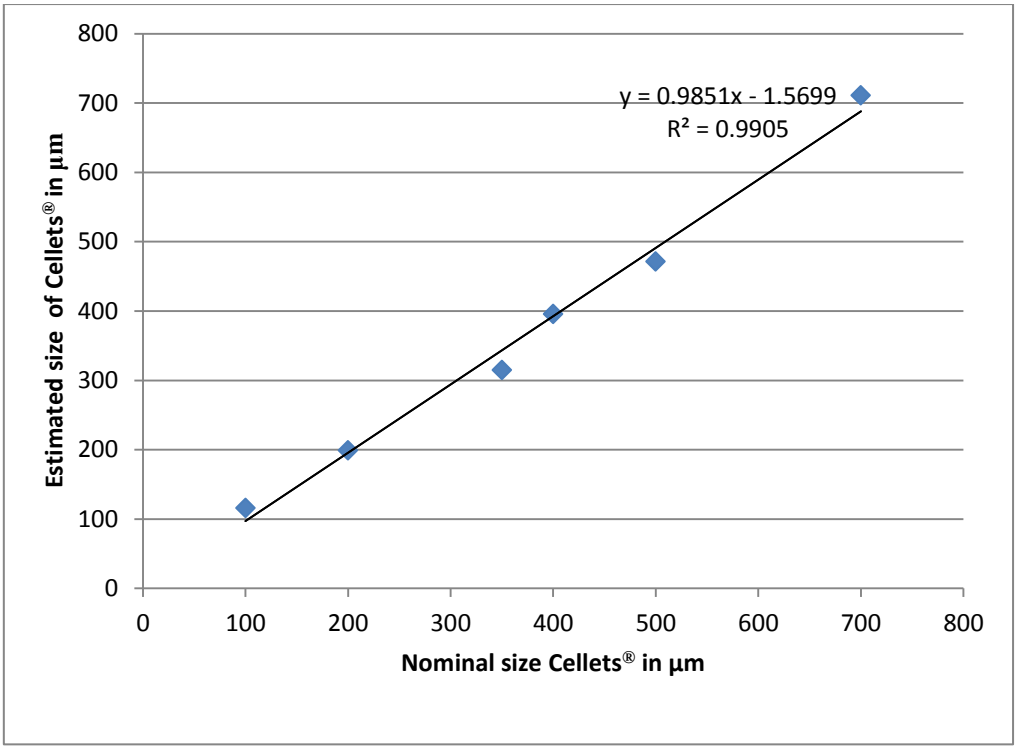


Figure 47 The average diameter of particles in free fall state established via image analysis by the Eyecon™ images. The slope of the correlation plot is not significantly different ($p<0.05$) from 1.

4.2.3 Surface texture

Prior to texture analysis, image segmentation algorithms described in the previous section were applied. At the end of the image segmentation step, surface areas around the particles' centres were selected. The Haralick correlation property and variogram function were calculated for all particles in one image before obtaining an image average for both the Haralick correlation property and the variogram function between distances of lag 1 and 20.

4.2.3.1 Static conditions

The following experiment examined the Eyecon™, in conjunction with surface texture analysis techniques presented in section 3.6, at classifying particles based on their surface morphologies under static conditions. Three different samples were used containing 700 µm category smooth Cellets®, treated Cellets® and a mixture of 50% (w/w) treated and untreated

Cellets®. Ten samples were obtained from each of the three categories so a total of thirty images were captured before applying the surface texture analysis.

Table 9 Summary of experiment related to static surface texture estimation.

	Static
Untreated 700 µm Cellets® sample	10 samples
Mixture sample	10 samples
Treated Cellets® sample	10 samples

Figure 48 shows the resulting means of the variogram and Haralick correlation property of pixel intensity scores of neighbouring pixels at distances (arbitrary unit for distance between pixels) of between 1 and 20. The smooth particles' mean variogram function has a smaller slope than the blended sample and the treated particles. For the treated particles, the variogram means also show higher variability than the other samples. The means of the Haralick correlation function of the smooth particles have smaller slopes compared to the Haralick correlation means of the other samples. The behaviour of the Haralick mean correlation property and variogram mean can be explained by the greater correlation of pixels separated by a distance k for the smooth particles. It would be expected that pixels in treated particles would be less correlated due to sudden spatial changes in pixel intensities due to its different morphological properties. There is also a slightly higher variability in the treated particle samples suggesting the spatial variability between particles within the treated group is larger than the smoother particles.

In the same figure, applying PCA on the mean variogram and Haralick correlation feature showed two distinguishable groups. The first cluster contains smooth Cellets® and mixed particles while the second group contains treated Cellets®.

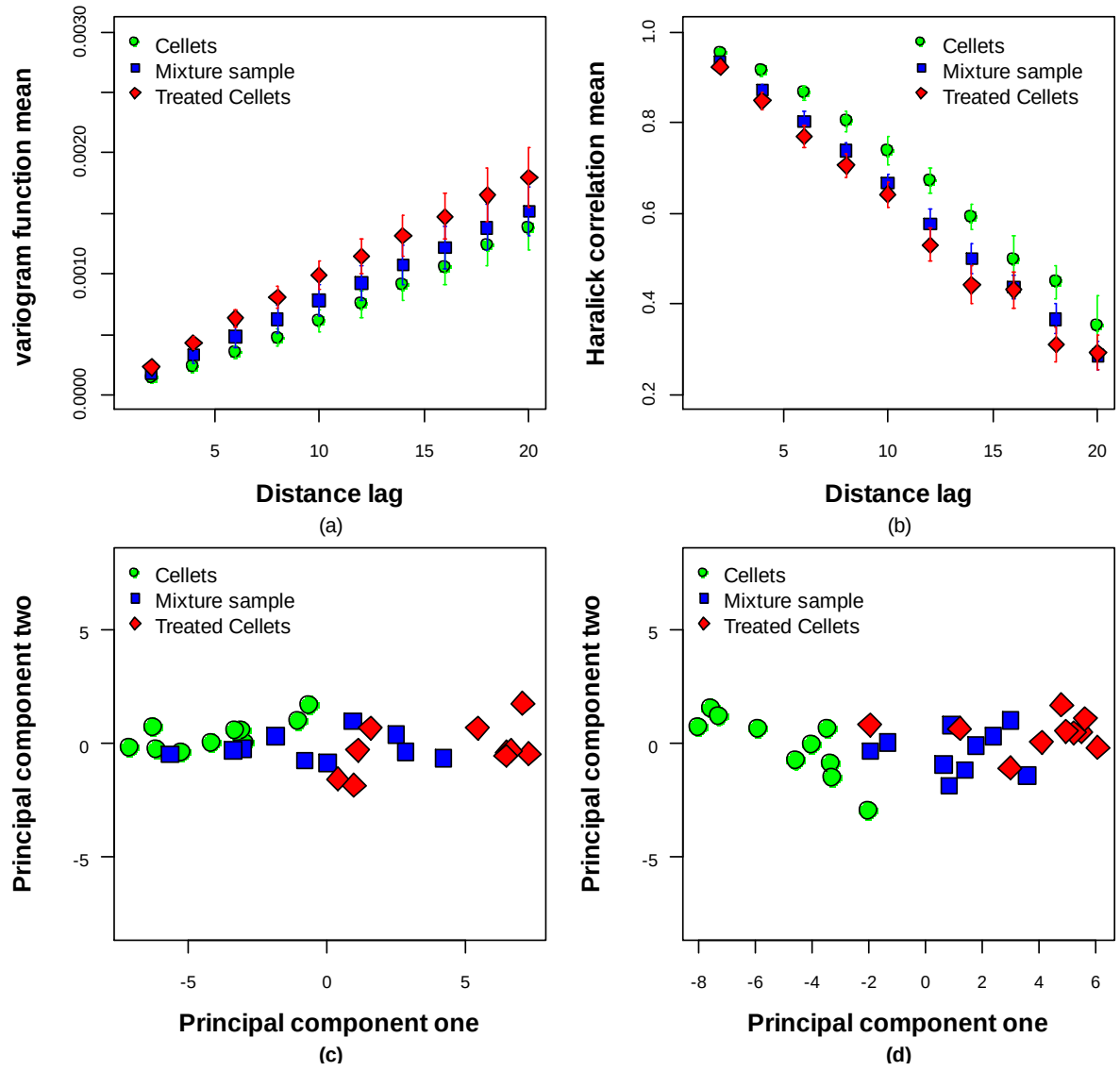


Figure 48 Texture analysis for the first sets of particles under static conditions: (a) Mean of variogram function; (b) Mean of Haralick function; (c) PCA classification based on the mean variogram function; (d) PCA classification based on the mean of Haralick function.

4.2.3.2 Dynamic conditions

The following experiment examined the Eyecon™, in conjunction with surface texture analysis techniques presented in section 3.6, at classifying particles based on their surface morphologies under dynamic conditions. Particles were adhered to a conveyor belt as described in section 3.4. The conveyor belt was operated at two speeds: slow (3cm/s) and fast (23 cm/s). For each speed the experimental design consisted of three different samples containing smooth Cellets®, treated Cellets® and a mixture of 50% (w/w) treated and

untreated Cellets®. Ten samples were obtained for each of the three categories so a total of thirty images were captured before applying the surface texture analysis at each speed Table .

Table 10 Summary of experiment related to dynamic surface texture estimation

	Slow (3 cm/s)	**Fast (23 cm/s)**
Untreated Cellets® sample	10 samples	10 samples
Mixture sample	10 samples	10 samples
Treated Cellets® sample	10 samples	10 samples

The results for the dynamic conditions show a relatively similar behaviour as for the static conditions. Figure 49 shows the mean of the variogram and the Haralick correlation property respectively for particle surfaces under slow speed. Under slow speed conditions the Haralick correlation property and the variogram function of the smooth particles had smaller slopes than both the blended and treated samples. There is less variability among surface texture functions for the three sample types. Although the variance at distance one is very small, it increases gradually with distance. It is suggested that the movement of the particle might affect the resolution of the captured image with the details of the surface becoming less clear. Applying PCA on the Haralick correlation property gave two distinct clusters, one is the smooth particles while the second is the treated samples and the blended particles overlapping.

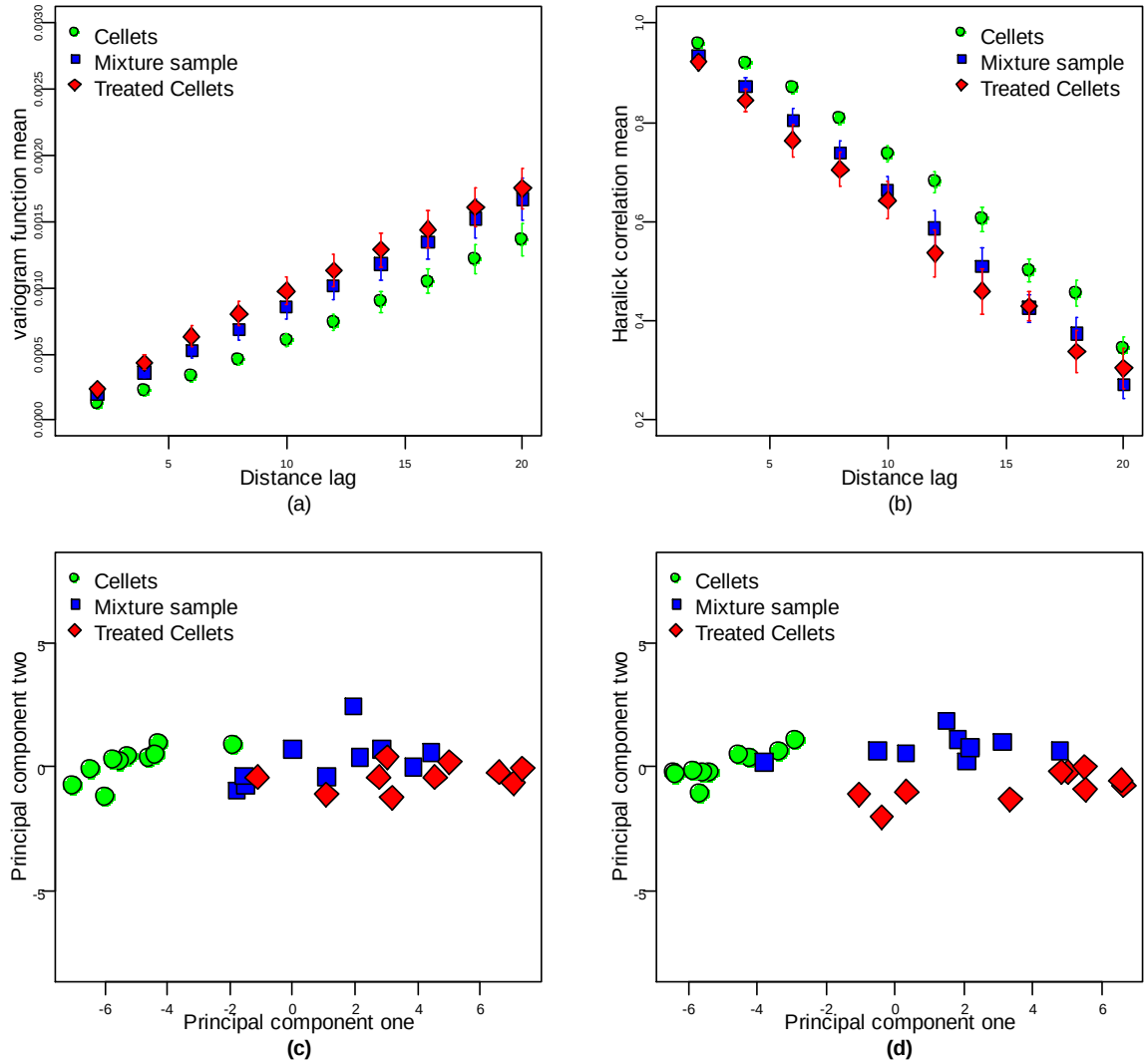


Figure 49 Results for the first set of particles at slow speed: (a) Mean of variogram function; (b) Mean of Haralick function; (c) PCA classification based on the mean variogram function and (d) PCA classification based on the mean of Haralick correlation feature.

As the speed of the conveyor belt increased to 23 cm/s, the mean of the variogram and the mean of the Haralick correlation property function of the three sample types become closer to each other as showed Figure 50. The variogram function showed more variability than the Haralick correlation property under this dynamic condition. Applying PCA using either the variogram or Haralick correlation function can identify two groups: one is the treated particles while the other one is a smooth particle group and the blended group which are overlapping.

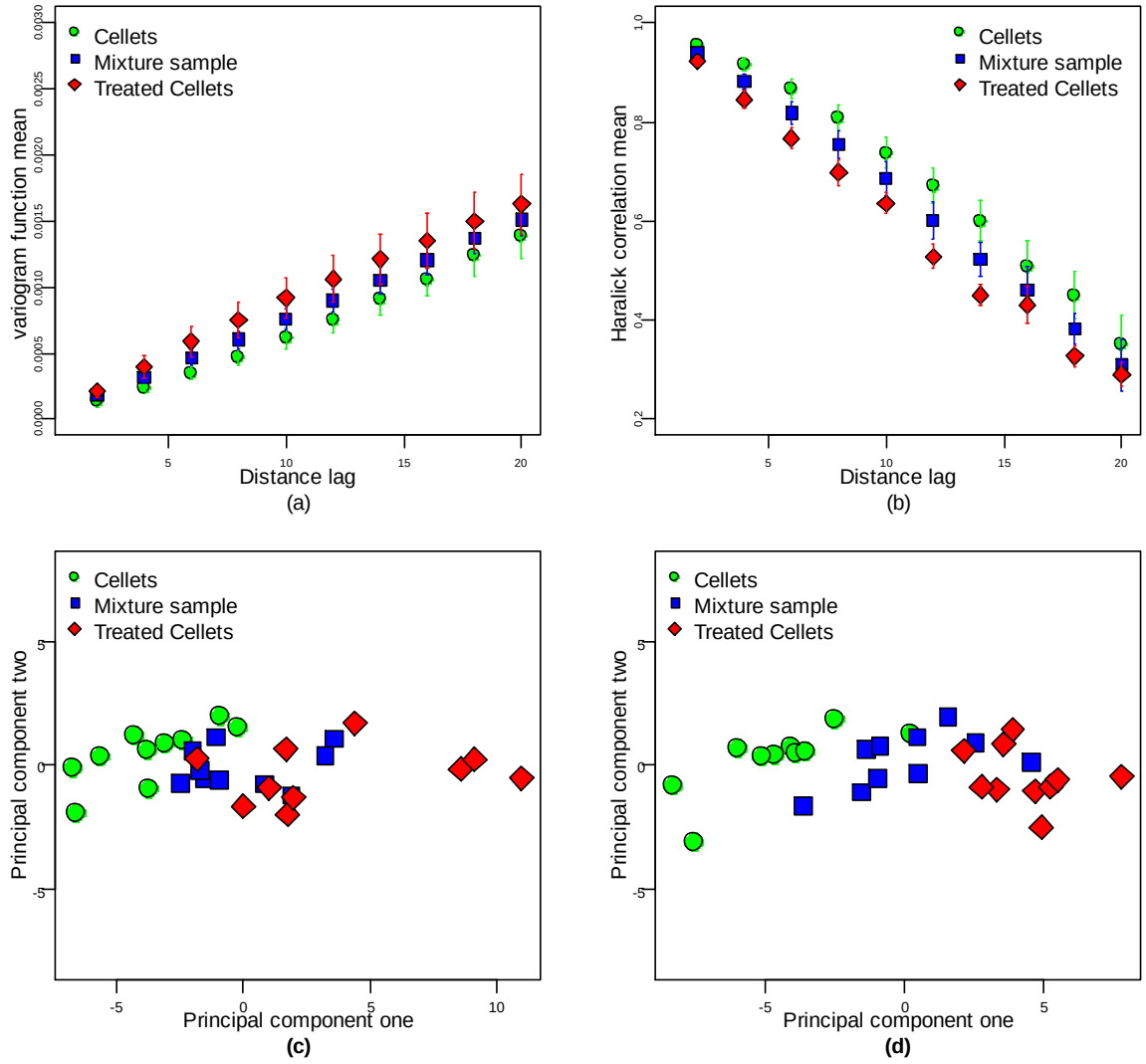


Figure 50 Texture analysis for the first set of particles at faster speed 23cm/s: (a) Mean of variogram function; (b) Mean of Haralick correlation feature; (c) PCA classification based on the mean variogram function; and (d) PCA classification based on the mean of Haralick correlation feature.

4.2.4 Texture analysis on the API sample

In this experiment the Eyecon™ was used in conjunction with texture analysis algorithms described in section 3.6 on three different particle samples containing smooth Cellets® with nominal size 1000 µm, API with nominal size 1000 µm and a blend of 50% (w/w) API and Cellets®. Ten samples were obtained from each of the three categories so a total of thirty images were captured before applying the surface texture analysis under static conditions.

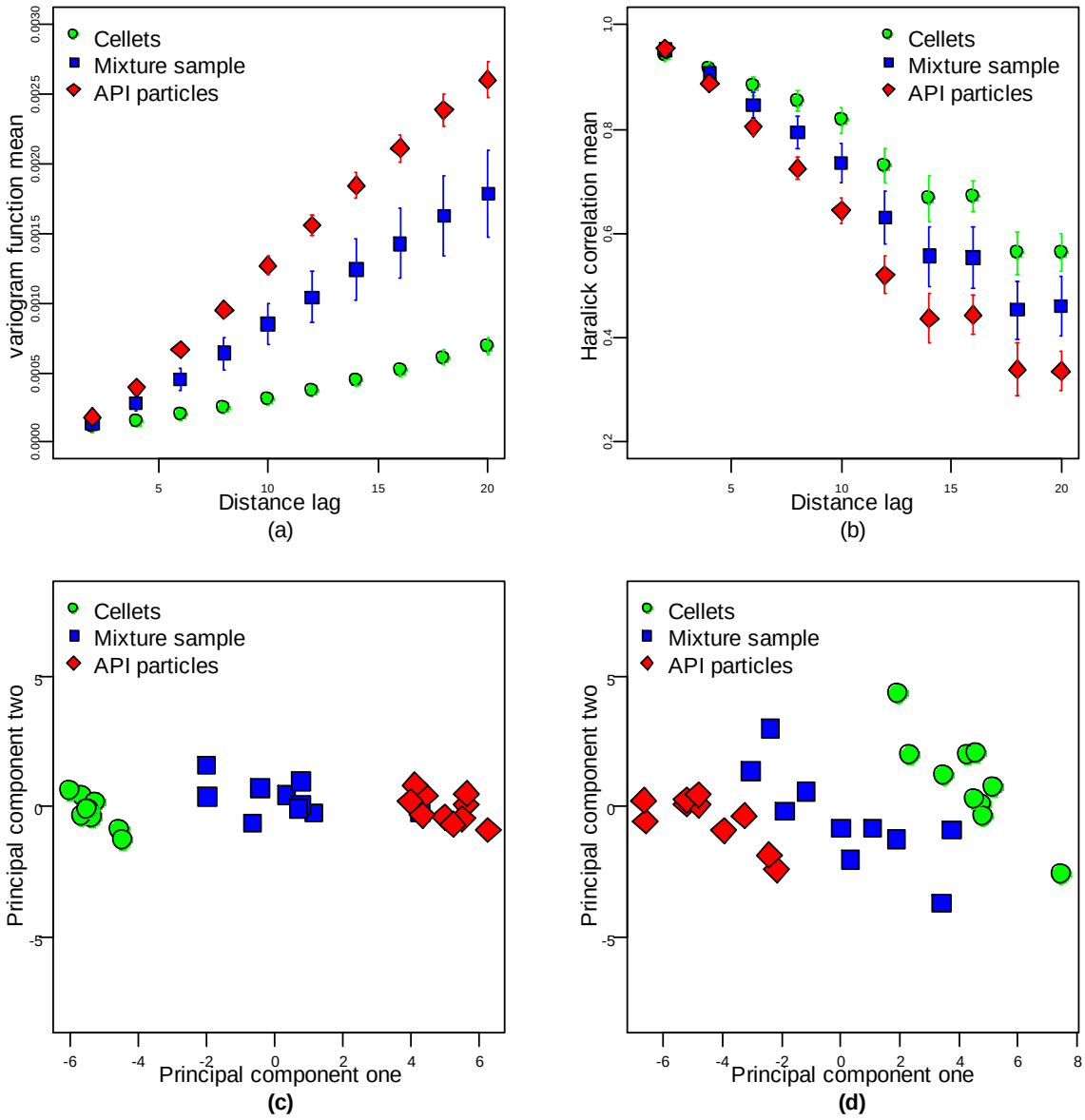


Figure 51 Results for the second set of particles under static conditions; (a) Mean of variogram function (b) Mean of Haralick function; (c) PCA classification based on the mean variogram function and (d) PCA classification based on the mean of Haralick function.

Figure 51 shows the mean variogram and Haralick correlation functions for the three categories of particles in the static state. Although all the plots started at the same point at distance one, the rate of change in the variogram and Haralick correlation functions was slower in the case of Cellets® and faster in case of the API. In the case of the mean variogram the variance at each distance step of the blended sample is the highest, this is due to the fact that images of the blended sample may contain different ratio of Cellets® and API particles. Again, the behaviour of the Haralick mean correlation property and variogram mean can be

explained by the greater correlation of pixels separated by a distance k for the smooth particles. And the pixels in the API would be less correlated. In this case the mean variogram and Haralick correlation were successful in reflecting the morphology of the particles as three clusters of points were obtained after applying PCA.

4.2.4.1 Dynamic conditions

This is similar to the previous experiment but with different samples namely smooth Cellets® with nominal size 1000 µm, API with nominal size 1000 µm and a blend of 50% (w/w) API. Ten samples were obtained for each of the three categories so a total of thirty images were captured before applying the surface texture analysis at each speed. At 3cm/s the mean variogram and Haralick correlation property function have similar behaviour as for the static case as illustrated in Figure 52; the rate of change in both mean functions is slower in case of Cellets® and higher in case of API. The difference from the static state is noticed in the higher variance in the variogram mean of the blended particles and also this time the API and the blend mean variogram and Haralick correlation property are close to each other at each distance step. Consequently, applying PCA on both the variogram and the Haralick correlation property gave two clusters of points in each case, one corresponding to smooth particles while the other correspond to both the blend and API.

Increasing the speed of the conveyor belt to 23 cm/s, the mean variogram and Haralick correlation property showed a similar result to the texture analysis at 3cm/s, the only difference was that the variance of the mean variogram and Haralick correlation function of the API particles are higher than for the previous particle speed (Figure 53).

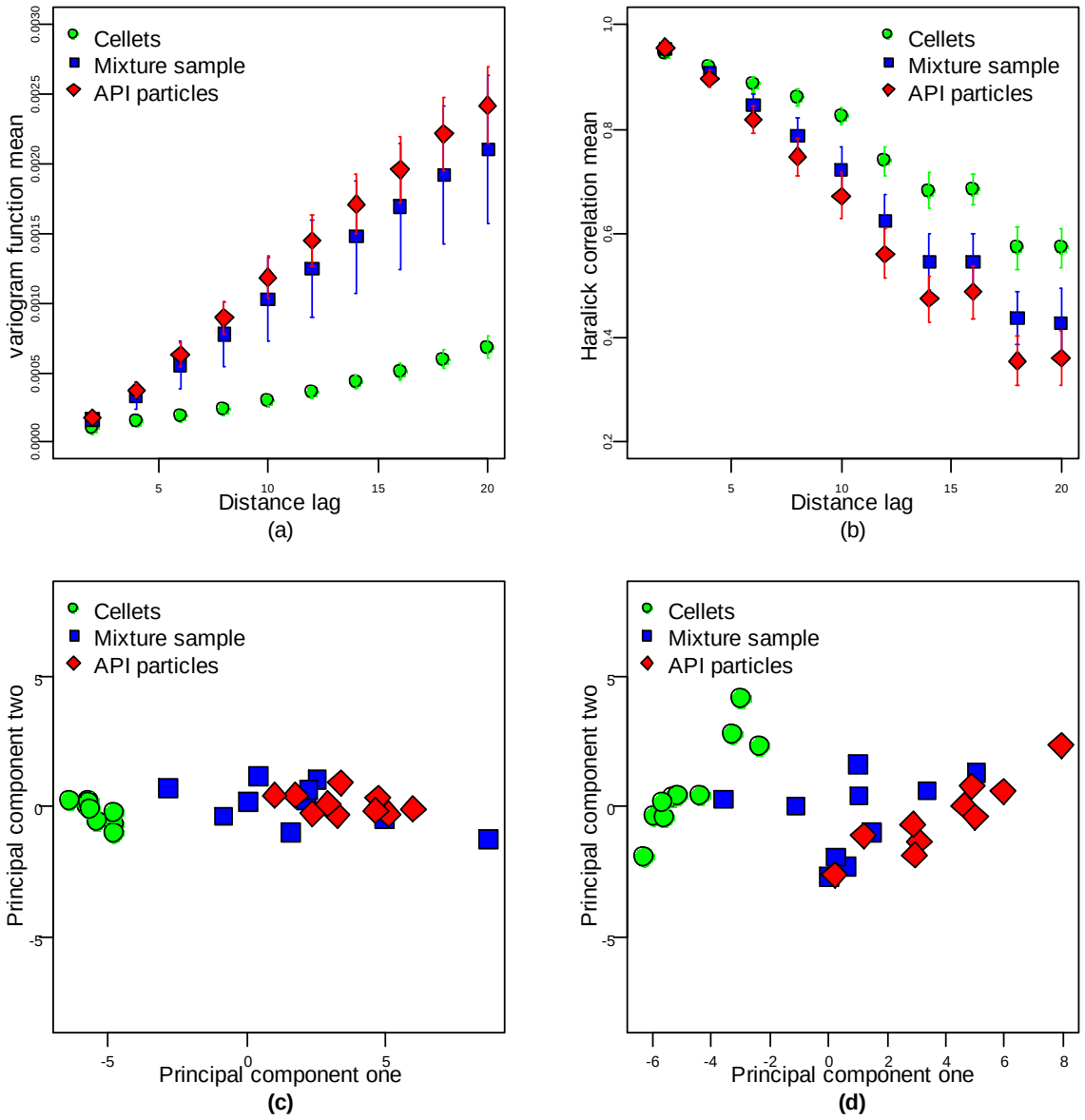


Figure 52 Texture analysis for the second set of particles at slower speed (3cm/s): (a) Mean of variogram function; (b) Mean of Haralick function; (c) PCA classification based on the mean variogram function; and (d) PCA classification based on the mean of Haralick correlation property.

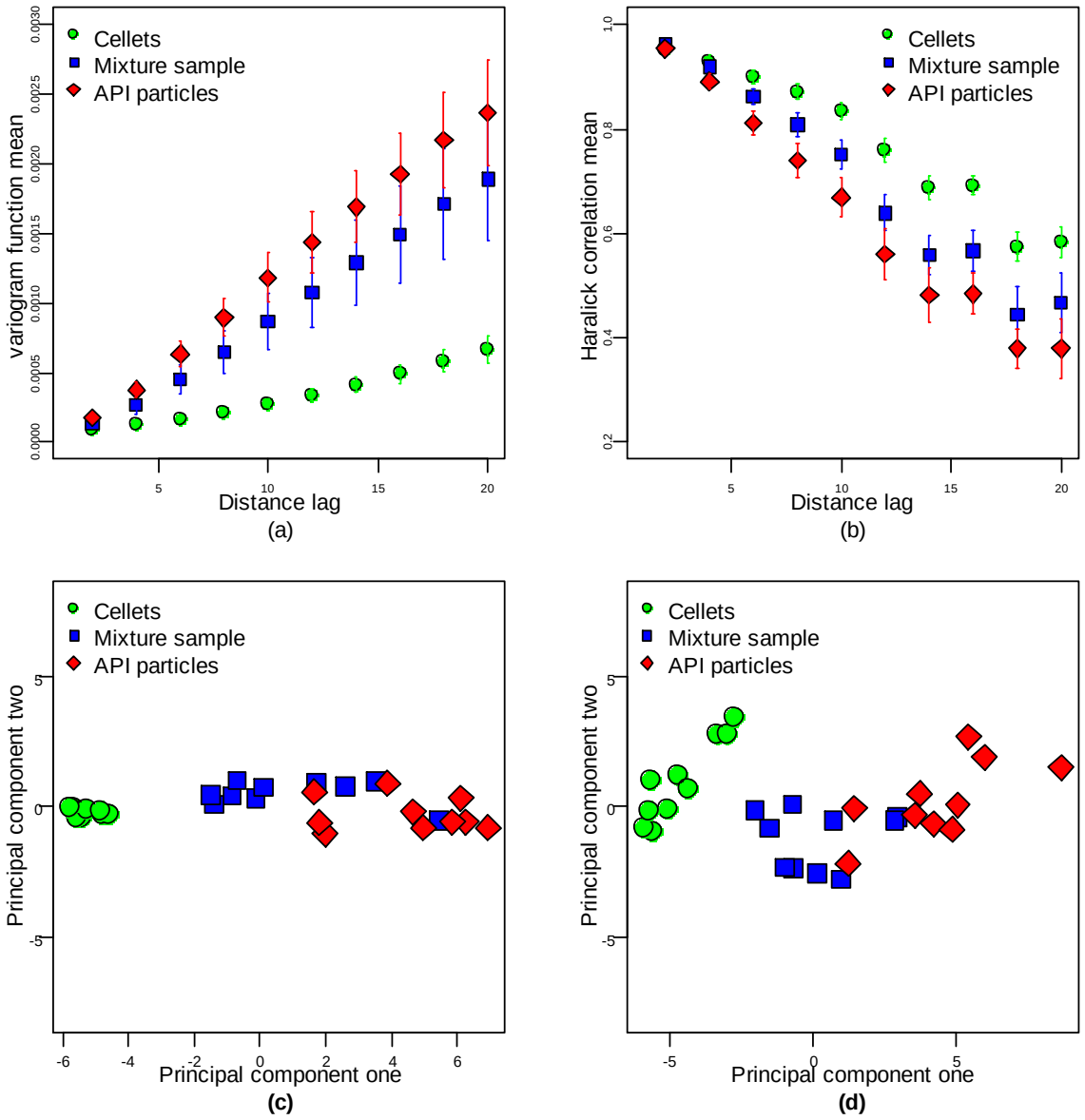


Figure 53 Texture analysis for the second set of particles at faster speed (23cm/s): (a) Mean of variogram function; (b) Mean of Haralick correlation feature; (c) PCA classification based on the mean variogram function; and (d) PCA classification based on the mean of Haralick correlation feature.

4.2.5 Texture analysis using Auto-correlation function

The auto-correlation function is computationally cheap compared to the previous surface texture analysis methods. In this work, it is computed in the horizontal and the vertical directions and then an average between the two directions is calculated. Similarly, to the previous texture analysis functions the texture analysis is performed on pixels that are apart between 1 and 20 pixels (distance lags).

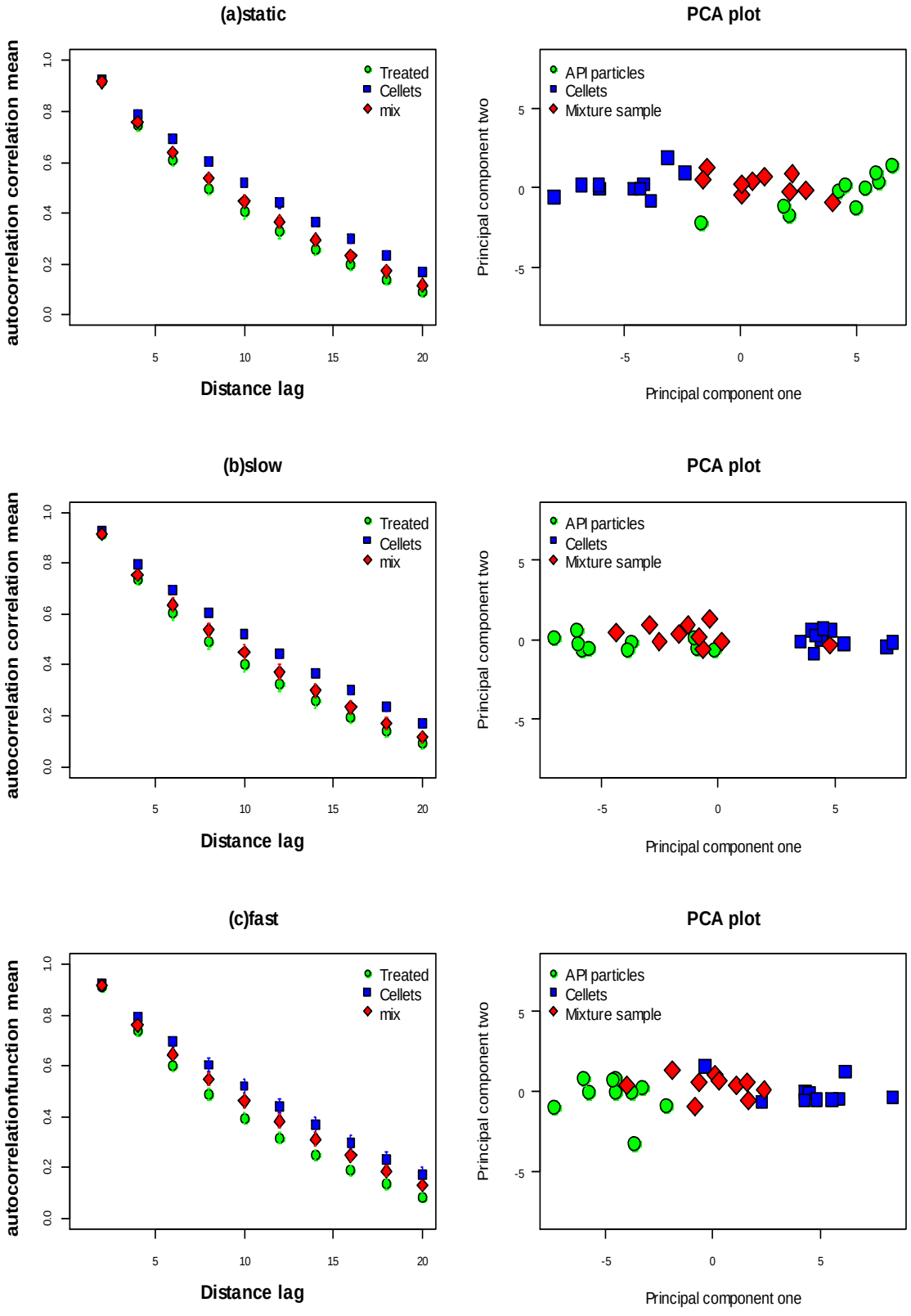


Figure 54 Autocorrelation function for the three particle state: (a)static; (b)slow and (c)fast.

An average of the auto-correlation function corresponding to particles from ten images is determined for each particle state (static, slow and fast). Figure 54 shows the resulting plots from the first set of particles, the autocorrelation function shows a similar behaviour as the Haralick correlation property function. In all particle states the average autocorrelation function ranked the three speed categories of particles as Cellets®. Autocorrelation mean of the ten images is bigger than both the mixed sample and treated Cellets®. The figure also shows a low variation for the surface texture of Cellets® in both static and slow motion states having a smaller standard deviation compared to both the mixed and treated particle samples. This is captured in the PCA plot of the first component against the second principal component where Cellets® are more clustered than the other two particle sample categories. The source of variation of the mixed sample arises from the fact that it contains two types of surface texture while in the case of treated Cellets® variation is coming from the resulted etched Cellets® surface structure. In static and slow motion states PCA classified particles into two clustered groups, the first group was composed of Cellets® and the second group is composed of the mixed sample and treated Cellets®.

For the second set of particles, Figure 55 shows the resulting plots from the autocorrelation function means of 10 images and the PCA plot from principal component analysis. The autocorrelation function in this case does not indicate any change in the surface texture between distance 1 and 5, but after distance 5 the function average of Cellets® is higher than the mix sample and the API particles for all particle states. In the classification, the auto-correlation function gave three groups corresponding to each particle category.

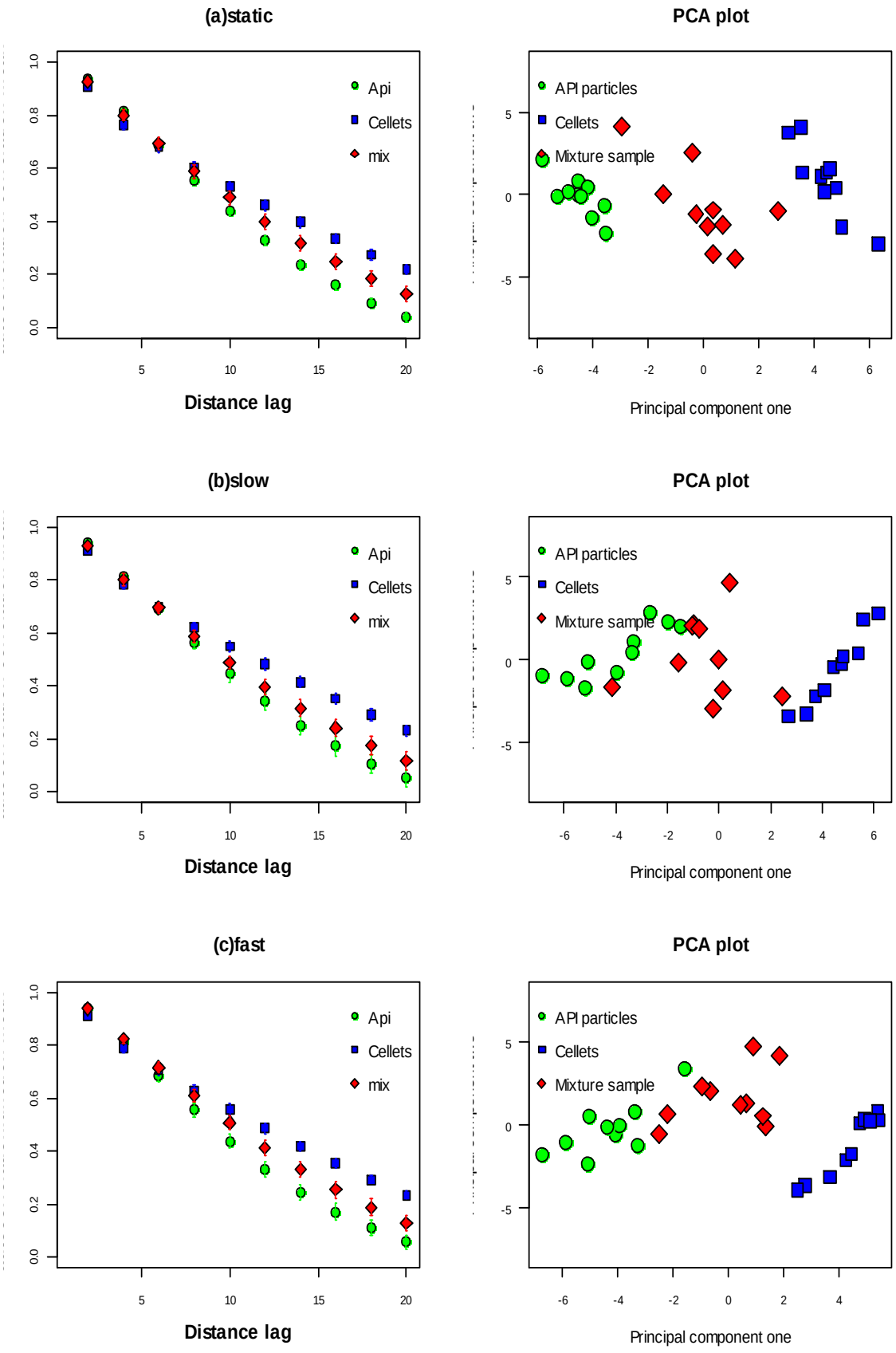


Figure 55 Autocorrelation function for the second set or particles three particle state: (a)static; (b)slow and and (c)fast.

4.2.6 Texture analysis on Edge based image

The images from the previous section of the API, Cellets® and their 50% (w/w) mixture were used to test the edge based texture analysis approach. The Sobel operator was applied on the surface of the greyscale surface of each particle before performing the texture analysis using the variogram function.

The texture analysis results are shown in Figure 56, for all states the variograms were able to rank the particles in accordance with their surface morphology. The mixture contains two types of surface morphology, explaining the higher variations observed in the figure for all particles and also the higher spread of mixture points in the principal components graph for the three states. The variograms of all the particles have the same range but they have different sills so in PCA all the points were approximately in one line which means that only one component is needed to explain most of the variation of the data, thus the sill is the only measure that can differentiate between the particle surface morphology in this case. However, compared to the variograms obtained in the previous section using the surface of the grayscale image, the variogram of the edges showed more difference between Cellets® and API.

The state of the particles also affected the texture analysis for Cellets, as the particle speed increased the morphology of Cellets® appeared to be smoother thus variation was detected in the fast speed case compared to the previous states. For the other particles the variation was about the same.

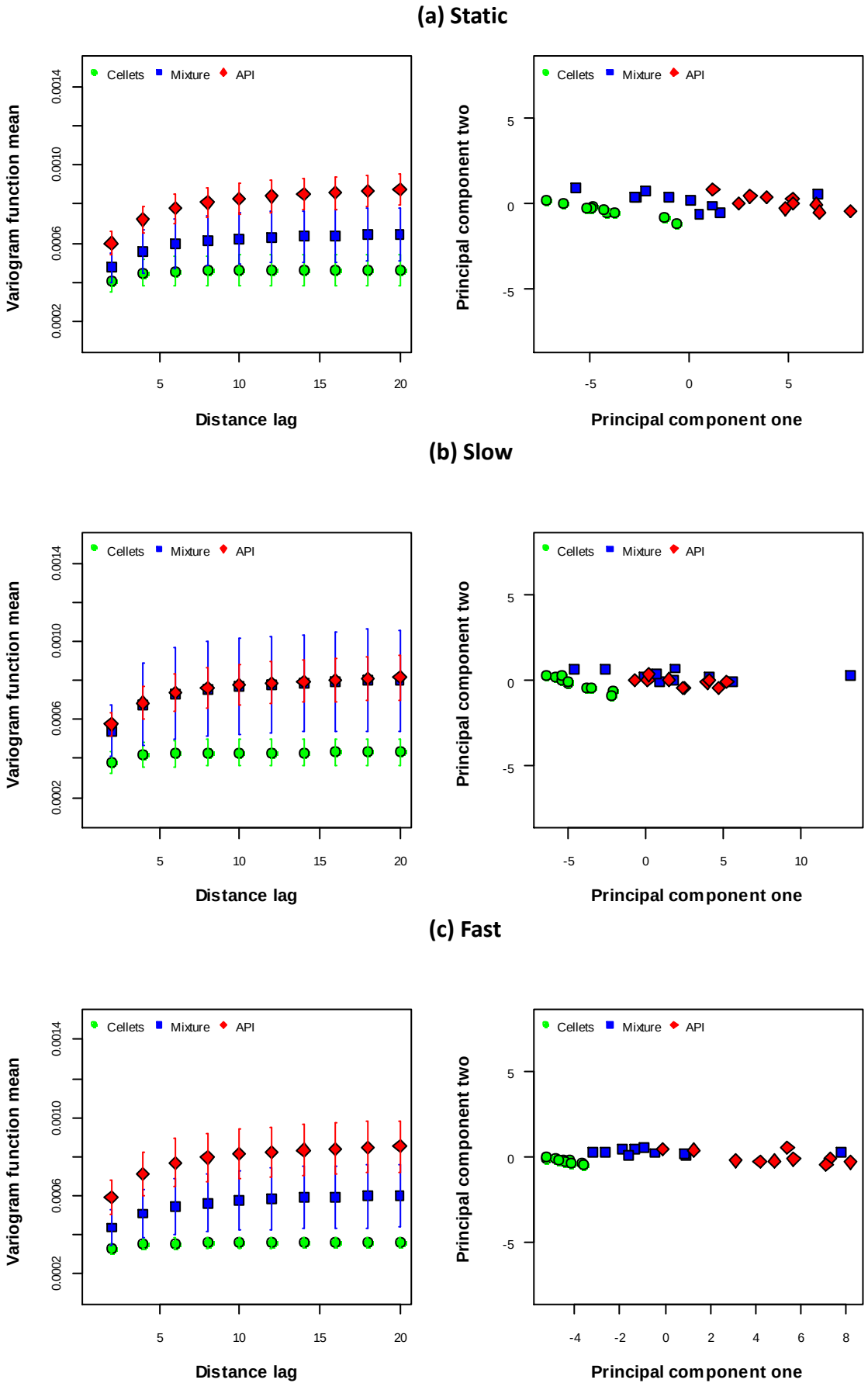


Figure 56 Variogram obtained using edge based images of Cellets, API and mixture sample in three states (a)static; (b)slow and (c)fast.

5 Conclusions

5.1 Chemical Properties and NIRS

- The MultiEye™ can, in conjunction with the multivariate techniques explored in this thesis, be used to correctly identify the chemical composition of pharmaceutical particles in both static and moving particle samples. As the particles move it becomes more challenging to measure the spectrum of these particles because the MultiEye™ probes require some time to detect the reflectance of the samples.
- The spectrum of moving particles is noisy, applying smoothing algorithm reduces the noise and produces a spectrum similar to the spectrum of the particles in static state.
- The MultiEye™ can, in conjunction with multivariate techniques be used as in-line tool of the pharmaceutical blending process to determine the homogeneity of the blend at a lab scale.

5.2 Physical Properties and Image Analysis

- During blending, particle size affected the time required for the blend to reach a stable mixture; this study showed that for small particles more time is required to reach blend stability. It was shown in this thesis that a small difference of size among particles in the blend can lead to segregation of particles. At the other end particles bigger than 500 µm required less time to reach a stable mixture but for 1000µm the blend appears to constantly mix and segregate, this might be due to the inaccurate measurement of the spectra of bigger particles.
- Capturing images requires a fraction of a second, but the image analysis might be costly in terms of computation. Image analysis using the thresholding and

segmentation techniques, described in this thesis, were able to estimate particle size in both static and dynamic conditions.

- Particle size analysis on bulk powder using image segmentation is challenging since there is no background in the image thus the algorithm considers particles with low light intensity as a background therefore the particle size is underestimated. In this case particle size can be estimated by inferring the obtained texture analysis property to size. In this study the variogram sill is correlated to the nominal size of the particle.
- The image texture analysis techniques explored in this thesis can classify particles based on their surface texture under static conditions. However, in dynamic conditions the quality of the surface image of the particles becomes lower. This is mainly due to the blurriness of pixels caused by particle movement.

6 Recommendation for further work

- NIRS can be investigated more in depth for samples under dynamic conditions, for example in a free fall situation The experiment performed in this work were at lab scale a further step is to apply this system on factory scale using a real pharmaceutical powder.
- Several studies reported the use of NIRS to determine the surface texture of the material. Future works can investigate the use of NIRS to classify particles based on their surface texture.
- This study investigated the blending process of a binary mixture, in pharmaceutical blending process often has more than two components thus the model developed in this study can be further developed to take into account more components. In addition, the method presented in this work can be extended to estimate limits of detection.
- In this work spherical particles were used to perform the size analysis, analysis for non-spherical should be considered in future works.
- The accuracy of the size analysis under free fall state can be improved by either removing the blurriness from the particles using de-blurring methods or simply removing the blurry particles from the image analysis.
- Particles in a pharmaceutical process might be in a bulk form, the use of image segmentation is challenging and time consuming which will not permit the in-line calculation of the particle size. In this work particle size analysis on bulk particles was performed indirectly by inferring the obtained texture analysis property to size. The texture analysis used in this thesis gives one overall average particle ignoring the variation within particle size. The current method can be improved by dividing the image into regions and estimate the size in each region. This might give an estimation

of different particle size in the sample. In addition, edge based segmentation can be used to estimate the particle size in bulk powder. The next step is to develop edge segmentation algorithms.

- Particles can be classified by studying their microscopic surface texture so their proportion in pharmaceutical powder can be estimated. An attempt to do that is shown in Figure 57, where the surface texture of an API and Cellets® is studied using the variogram function and then their mean is calculated. This method can be improved by including other physical information such as the size and shape of particle or by using some statistical models. Further work can be continued to implement the adoption of this technology by the food and pharmaceutical industry following the challenges described in Cullen, O'Donnell and Fagan [8080].

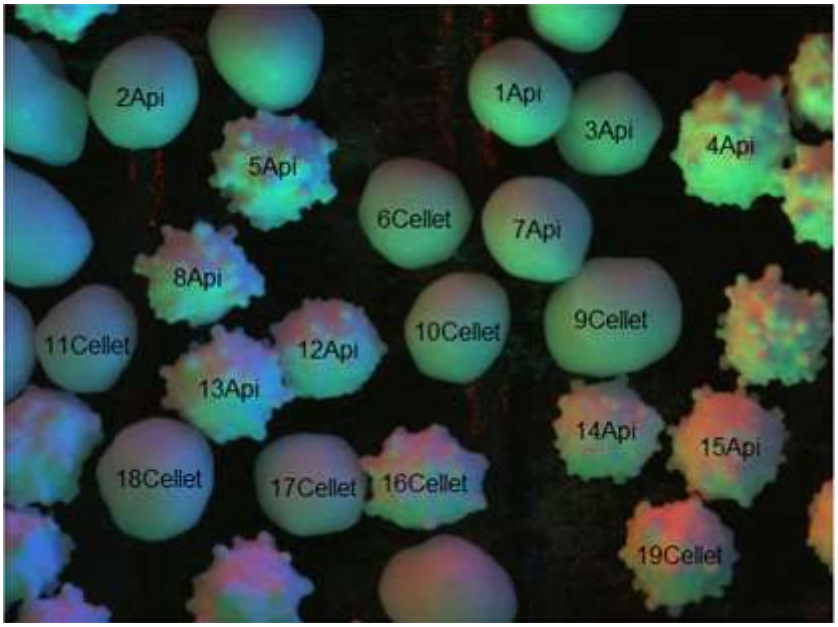


Figure 57 Identifying content percentages of different types of particles

7 References

1. Martin H. Thelle TJ, Jori Veng Pinje, Anders Kjøller-Hansen, Ronald B. Davies, Joseph F. Francois. TTIP impact in Ireland. 2015.

2. Business W. Pharmaceuticals in Ireland. 2010. http://www.een-ireland.ie/eei/assets/documents/uploaded/general/Pharmaceuticals%20Fact%20sheet.pdf. Accessed 14-03 2014.

3. Jones D. Pharmaceutical Applications of Polymers for Drug Delivery. Rapra Technology Limited; 2004.

4. Food, Administration D. Guidance for industry PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance. DHHS, Rockville, MD. 2004.

5. Palou A, Cruz J, Blanco M, Tomàs J, de los Ríos J, Alcalà M. Determination of drug, excipients and coating distribution in pharmaceutical tablets using NIR-CI. Journal of Pharmaceutical Analysis. 2012;2(2):90-7.

6. Silva AFT, Burggraeve A, Denon Q, Van der Meeren P, Sandler N, Van Den Kerkhof T et al. Particle sizing measurements in pharmaceutical applications: Comparison of in-process methods versus off-line methods. Eur J Pharm Biopharm. 2013;85(3):1006-18. doi:DOI 10.1016/j.ejpb.2013.03.032.

7. Seitavuopio P, Heinamaki J, Rantanen J, Yliruusi J. Monitoring tablet surface roughness during the film coating process. Aaps Pharmscitech. 2006;7(2).

8. Ryan Gosselin M-OL, Nicolas Abatzoglou. Determining Bulk Powder Particle Size through Texture Analysis. American Pharmaceutical Review. 2013.

9. Helliwell M, Taylor D. Solid oral dosage forms. Professional nurse (London, England). 1993;8(5):313-7.

10. Augsburger LL, Hoag SW. Pharmaceutical Dosage Forms - Tablets: Manufacture and Process Control. Taylor & Francis; 2008.

11. Levin M. Pharmaceutical Process Scale-Up. Taylor & Francis; 2001.

12. Scheibelhofer O, Balak N, Wahl PR, Koller DM, Glasser BJ, Khinast JG. Monitoring Blending of Pharmaceutical Powders with Multipoint NIR Spectroscopy. Aaps Pharmscitech. 2013;14(1):234-44. doi:DOI 10.1208/s12249-012-9910-4.

13. Skibsted ETS, Boelens HFM, Westerhuis JA, Witte DT, Smilde AK. Simple assessment of homogeneity in pharmaceutical mixing processes using a near-infrared reflectance probe and control charts. J Pharmaceut Biomed. 2006;41(1):26-35. doi:DOI 10.1016/j.jpba.2005.10.009.

14. Jamrogiewicz M. Application of the near-infrared spectroscopy in the pharmaceutical technology. J Pharmaceut Biomed. 2012;66:1-10. doi:DOI 10.1016/j.jpba.2012.03.009.

15. Luypaert J, Massart DL, Heyden YV. Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta. 2007;72(3):865-83. doi:DOI 10.1016/j.talanta.2006.12.023.

16. Roggo Y, Chalus P, Maurer L, Lema-Martinez C, Edmond A, Jent N. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J Pharmaceut Biomed. 2007;44(3):683-700. doi:DOI 10.1016/j.jpba.2007.03.023.

17. Amanda J. Rogers AHaMGI. Modeling of Particulate Processes for the Continuous Manufacture of Solid-Based Pharmaceutical Dosage Forms. processes. 2013.

18. Bakeev KA. Process analytical technology: spectroscopic tools and implementation strategies for the chemical and pharmaceutical industries. John Wiley & Sons; 2010.

19. Baughman E. Process analytical chemistry: introduction and historical perspective. Analytical Technology. 2008:1.

20. Eriksson L, Byrne T, Johansson E, Trygg J, Vikström C. Multi-and megavariate data analysis basic principles and applications. Umetrics Academy; 2013.

21. Riley BS, Li XH. Quality by Design and Process Analytical Technology for Sterile Products-Where Are We Now? Aaps Pharmscitech. 2011;12(1):114-8. doi:DOI 10.1208/s12249-010-9566-x.

22. ICH. Guidance for Industry, Q8 (R2) Pharmaceutical Development. 2009.

23. El-Hagrasy AS, Morris HR, D'Amico F, Lodder RA, Drennen JK. Near-infrared spectroscopy and imaging for the monitoring of powder blend homogeneity. J Pharm Sci-Us. 2001;90(9):1298-307. doi:Doi 10.1002/Jps.1082.

24. Naervanen T, Seppaelae K, Antikainen O, Yliruusi J. A new rapid on-line imaging method to determine particle size distribution of granules. Aaps Pharmscitech. 2008;9(1):282-7. doi:DOI 10.1208/s12249-008-9043-y.

25. Schaefer C, Lecomte C, Clicq D, Merschaert A, Norrant E, Fotiadu F. On-line near infrared spectroscopy as a Process Analytical Technology (PAT) tool to control an industrial seeded API crystallization. J Pharmaceut Biomed. 2013;83:194-201. doi:DOI 10.1016/j.jpba.2013.05.015.

26. Soppela I, Airaksinen S, Hatara J, Raikkonen H, Antikainen O, Yliruusi J et al. Rapid Particle Size Measurement Using 3D Surface Imaging. Aaps Pharmscitech. 2011;12(2):476-84. doi:DOI 10.1208/s12249-011-9607-0.

27. Wu YJ, Jin Y, Li YR, Sun D, Liu XS, Chen Y. NIR spectroscopy as a process analytical technology (PAT) tool for on-line and real-time monitoring of an extraction process. Vib Spectrosc. 2012;58:109-18. doi:DOI 10.1016/j.vibspec.2011.10.006.

28. El Hagrasy AS, Cruise P, Jones I, Litster JD. In-line Size Monitoring of a Twin Screw Granulation Process Using High-Speed Imaging. J Pharm Innov. 2013;8(2):90-8. doi:DOI 10.1007/s12247-013-9149-y.

29. Newgard EC. Near-Infrared Spectroscopy for Analysis of Agricultural Material. Final Report for Physics 598 OS (Optical Spectroscopy) Fall 2004. 2004.

30. Mantanus J. New pharmaceutical applications involving Near Infrared spectroscopy as a PAT compliant process analyzer. 2012.

31. Reich G. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. Adv Drug Deliver Rev. 2005;57(8):1109-43. doi:DOI 10.1016/j.addr.2005.01.020.

32. EudraLex. The Rules Governing Me dicinal Products in the European Union. In: DIRECTORATE-GENERAL HAC, editor. Good Manufacturing Practice Medicinal Products for Human and Veterinary Use Brussel

33. Sarma A, Kunwar A. Pharmaceutical Applications of NMR Spectroscopy. PROCEEDINGS-INDIAN NATIONAL SCIENCE ACADEMY PART A. 1999;65(6):655-82.

34. Maiwald M, Steinhof O, Sleigh C, Bernstein M, Hasse H. Quantitative high-resolution online NMR spectroscopy in pharmaceutical reaction and process monitoring. Elsevier, Oxford; 2008.

35. Wold S. Chemometrics; what do we mean with it, and what do we want from it? Chemometrics and Intelligent Laboratory Systems. 1995;30(1):109-15.

36. Miller JN, Miller JC. Statistics and Chemometrics for Analytical Chemistry. Pearson/Prentice Hall; 2005.

37. Frank LE, Friedman JH. A statistical view of some chemometrics regression tools. Technometrics. 1993;35(2):109-35.

38. Massart DL, Vandeginste B, Deming S, Michotte Y, Kaufman L. Chemometrics: a textbook. 1988.

39. Huang J, Romero-Torres S, Moshgbar M. RAMAN: Practical Considerations in Data Pre-treatment for NIR and Raman Spectroscopy. American Pharmaceutical Review. 2010;13(6):116.

40. Savitzky A, Golay MJE. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. Anal Chem. 1964;36(8):1627-&. doi:Doi 10.1021/Ac60214a047.

41. Abdi H. Partial least squares regression (PLS-regression). Thousand Oaks, CA: Sage; 2003. p. 792-5.

42. Tobias RD, editor. An introduction to partial least squares regression. Proc. Ann. SAS Users Group Int. Conf., 20th, Orlando, FL; 1995.

43. Mahobia NK, Patel RD, Sheikh NW, Singh SK, Mishra A, Dhardubey R. Validation method used in quantitative structure activity relationship. Der Pharma Chemica. 2010;2(5).

44. Mevik B-H, Wehrens R. The pls package: principal component and partial least squares regression in R. Journal of Statistical Software. 2007;18(2):1-24.

45. Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision. Cengage Learning; 2014.

46. Glasbey CA, Horgan GW. Image analysis for the biological sciences. John Wiley & Sons New York; 1995.

47. Adamek T, O'Connor NE, Murphy N. Region-based segmentation of images using syntactic visual features. 2005.

48. Gendrin C, Roggo Y, Collet C. Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review. J Pharmaceut Biomed. 2008;48(3):533-53. doi:DOI 10.1016/j.jpba.2008.08.014.

49. Laitinen N, Antikainen O, Yliruusi J. Characterization of particle sizes in bulk pharmaceutical solids using digital image information. Aaps Pharmscitech. 2003;4(4):383-91.

50. Garcia-Munoz S, Carmody A. Multivariate wavelet texture analysis for pharmaceutical solid product characterization. Int J Pharm. 2010;398(1-2):97-106. doi:DOI 10.1016/j.ijpharm.2010.07.032.

51. Garcia-Munoz S, Gierer DS. Coating uniformity assessment for colored immediate release tablets using multivariate image analysis. Int J Pharm. 2010;395(1-2):104-13. doi:DOI 10.1016/j.ijpharm.2010.05.026.

52. Islam MJ, Ahmadi M, Sid-Ahmed MA, editors. Image processing techniques for quality inspection of gelatin capsules in pharmaceutical applications. Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on; 2008: IEEE.

53. Špiclin Ž, Bukovec M, Pernuš F, Likar B. Image registration for visual inspection of imprinted pharmaceutical tablets. Machine Vision and Applications. 2011;22(1):197-206.

54. GmbH S. Fundamentals Particle Size and Shape Calculation by Image Analysis. http://www.sympatec.com/EN/ImageAnalysis/Fundamentals.html. Accessed 28/03/2014 2014.

55. Worldwide MI. A Basic Guide to Particle Characterization2012.

56. Brittain HG. Particle-size distribution, part III: determination by analytical sieving. Pharmaceutical technology. 2002;26(12):56-64.

57. Stojanovic Z, Markovic S. Determination of Particle Size Distributions by Laser Diffraction'. New Materials. 2012(21):11-20.

58. Burggraeve A, Van Den Kerkhof T, Hellings M, Remon JP, Vervaet C, De Beer T. Evaluation of in-line spatial filter velocimetry as PAT monitoring tool for particle growth during fluid bed granulation. Eur J Pharm Biopharm. 2010;76(1):138-46.

59. Sandler N. Photometric imaging in particle size measurement and surface visualization. Int J Pharm. 2011;417(1–2):227-34. doi:http://dx.doi.org/10.1016/j.ijpharm.2010.11.007.

60. Song J, Ma L, Whitenton E, Vorburger T. 2D and 3D surface texture comparisons using autocorrelation functions. Key Eng Mat. 2005;295-296:437-40.

61. Burnett DJ, Heng JYY, Thielmann F, Garcia AR, Naderi M, Acharya M. Measuring Surface Roughness of Pharmaceutical Powders Using Vapor Sorption Methods. Aaps Pharmscitech. 2011;12(1):56-61. doi:DOI 10.1208/s12249-010-9571-0.

62. Edwardson JM, Henderson RM. Atomic force microscopy and drug discovery. Drug discovery today. 2004;9(2):64-71.

63. Moreno-Herrero F, Colchero J, Gomez-Herrero J, Baro AM. Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. Phys Rev E. 2004;69(3). doi:Artn 031915 Doi 10.1103/Physreve.69.031915.

64. Bashaiwoldu AB, Podczeck F, Newton JM. The application of non-contact laser profilometry to the determination of permanent structural change induced by compaction of pellets - II. Pellets dried by different techniques. Eur J Pharm Sci. 2004;22(1):55-61. doi:DOI 10.1016/j.ejps.2004.02.005.

65. Bashaiwoldu AB, Podczeck F, Newton JM. The application of non-contact laser profilometry to the determination of permanent structural changes induced by compaction of pellets I. Pellets of different composition. Eur J Pharm Sci. 2004;21(2-3):143-54. doi:DOI 10.1016/j.ejps.2003.09.012.

66. Seitavuopio P. The roughness and imaging characterisation of different pharmaceutical surfaces. 2006.

67. Company F, Optics PE. All You Wanted to Know about Electron Microscopy: But Didn't Dare to Ask! FEI Electron Optics.

68. Alvarez-Jubete L, Mishra J, Jones I, Cullen P, Sullivan C. Feasibility of near infrared chemical imaging for pharmaceutical cleaning verification. Journal of Near Infrared Spectroscopy. 2013;21(3):173-82.

69. Team RDC. R: A Language and Environment for Statistical Computing. 2012.

70. Sezgin M. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic imaging. 2004;13(1):146-68.

71. Gonzalez MA, Cuadrado TR, Ballarin VL. Comparing marker definition algorithms for Watershed segmentation in microscopy images. Journal of Computer Science & Technology (JCS&T). 2008;8(3).

72. Merkus HG. Particle size measurements: fundamentals Practice, Quality. 2009:1-7.

73. Srinivasan G, Shobha G. Statistical Texture Analysis. Proceedings of World Academy of Science: Engineering & Technology. 2008;48.

74. Bharati MH, Liu JJ, MacGregor JF. Image texture analysis: methods and comparisons. Chemometrics and intelligent laboratory systems. 2004;72(1):57-71.

75. Mollereau G, Mazel V, Busignies V, Tchoreloff P, Mouveaux F, Rivière P. Image Analysis Quantification of Sticking and Picking Events of Pharmaceutical Powders Compressed on a Rotary Tablet Press Simulator. Pharmaceutical research. 2013;30(9):2303-14.

76. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. Systems, Man and Cybernetics, IEEE Transactions on. 1973(6):610-21.

77. Barnes RJ. The variogram sill and the sample variance. Mathematical geology. 1991;23(4):673-8.

78. Diggle PJRJPJ. Model Based Geostatistics. In: R-NEWS, editor. geoR: a package for geostatistical analysis: Springer; 2007.

79. Jay M. Rapid Near-Infrared Qualification of Microcrystalline Cellulose and Sodium Caprate Minitablets Through Intact Enteric Coated Capsules.

80. O'Donnell CP, Fagan C, Cullen PJ. Process Analytical Technology for the Food Industry. Springer; 2014.

8 List of conferences, poster presentations and publications

- Paper submitted to of **Journal Near Infrared Spectroscopy** (2014) "Evaluation of Diffuse Reflectance Near Infrared Fiber Optical Sensors in Measurements for Chemical Identification and Quantification for Binary Granule Blends" D. Togashi, L. Alvarez-Jubete, H. Rifai,R. Cama-Moncunill, P. Cruise, C. Sullivan and P.J. Cullen.

- **Food science and technology** journal,"Multipoint Near Infrared spectrometry for real-time monitoring of protein conformational stability in powdered infant formula" Ritesh Pabari, Denisio Togashi, Raquel Cama Moncunill, Toufic El Arnaout, Hicham Rifai, Paul Cruise, Patrick J Cullen, and Carl Sullivan

- Oral presentation at the Sixth pan-European QbD & PAT Science Conference, EuPAT6, being held in Porto Portugal, 22-24 September 2013 on the topic "Towards a PAT solution of measuring particle size and roughness", H. Rifai, L**.** Alvarez-Jubete, D. Togashi, C. Sullivan and P.J. Cullen.

- Poster on "Validation of a compact Near Infrared Spectrophotometer with Independent Fiber Optic Multiprobes" D. Togashi, L. Alvarez-Jubete, H. Rifai, C. Sullivan and P.J. Cullen; at the Sixth pan-European QbD & PAT Science Conference, EuPAT6, being held in Porto Portugal, 22-24 September 2013.

- Oral presentation at the Irish Medical Board during the Opex session, being held in Dublin, Ireland, 14 January 2014, on the topic "Real time monitoring of pharmaceutical particles during production with respect to size, shape and texture" H. Rifai, L**.** Alvarez-Jubete , D. Togashi, C. Sullivan and P.J. Cullen.

- Oral presentation at the International Forum Process Analytical Technology IFPAC 2014, being held in Arlington, VA (Washington DC), 25-28 January 2014 on the topic "Feasibility of Multipoint NIR Spectroscopy to characterise powder and granule blends" R. Cama, D. Togashi, L. Alvarez-Jubete, H. Rifai, C. Sullivan and P.J. Cullen

- Poster on "Classification of surface texture of particles using image surface analysis methods", H. Rifai, L**.** Alvarez-Jubete , R. Cama, D. Togashi, C. Sullivan and P.J. Cullen; at the International Forum Process Analytical Technology IFPAC 2014, being held at Arlington, VA (Washington DC), 25-28 January 2014