RYAN GOSSELIN

ON-LINE QUALITY CONTROL IN POLYMER PROCESSING USING HYPERSPECTRAL IMAGING

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de doctorat en Génie Chimique
pour l'obtention du grade de Philosophiae Doctor (PhD)

DÉPARTEMENT DE GÉNIE CHIMIQUE
FACULTÉ DES SCIENCES ET DE GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

2010

Résumé

L'industrie du plastique se tourne de plus en plus vers les matériaux composites afin d'économiser de la matière et/ou d'utiliser des matières premières à moindres coûts, tout en conservant de bonnes propriétés. L'impressionnante adaptabilité des matériaux composites provient du fait que le manufacturier peut modifier le choix des matériaux utilisés, la proportion selon laquelle ils sont mélangés, ainsi que la méthode de mise en œuvre utilisée.

La principale difficulté associée au développement de ces matériaux est l'hétérogénéité de composition ou de structure, qui entraîne généralement des défaillances mécaniques. La qualité des prototypes est normalement mesurée en laboratoire, à partir de tests destructifs et de méthodes nécessitant la préparation des échantillons. La mesure en-ligne de la qualité permettrait une rétroaction quasi-immédiate sur les conditions d'opération des équipements, en plus d'être directement utilisable pour le contrôle de la qualité dans une situation de production industrielle.

L'objectif de la recherche proposée consiste à développer un outil de contrôle de qualité pour la qualité des matériaux plastiques de tout genre. Quelques sondes de type proche infrarouge ou ultrasons existent présentement pour la mesure de la composition en-ligne, mais celles-ci ne fournissent qu'une valeur ponctuelle à chaque acquisition. Ce type de méthode est donc mal adapté pour identifier la distribution des caractéristiques de surface de la pièce (i.e. homogénéité, orientation, dispersion).

Afin d'atteindre cet objectif, un système d'imagerie hyperspectrale est proposé. À l'aide de cet appareil, il est possible de balayer la surface de la pièce et d'obtenir une image hyperspectrale, c'est-à-dire une image formée de l'intensité lumineuse à des centaines de longueurs d'onde et ce, pour chaque pixel de l'image. L'application de méthodes chimiométriques permettent ensuite d'extraire les caractéristiques spatiales et spectrales de l'échantillon présentes dans ces images. Finalement, les méthodes de régression multivariée permettent d'établir un modèle liant les caractéristiques identifiées aux propriétés de la pièce. La construction d'un modèle mathématique forme donc l'outil d'analyse en-ligne de la qualité des pièces qui peut également prédire et optimiser les conditions de fabrication.

Abstract

The use of plastic composite materials has been increasing in recent years in order to reduce the amount of material used and/or use more economic materials, all of which without compromising the properties. The impressive adaptability of these composite materials comes from the fact that the manufacturer can choose the raw materials, the proportion in which they are blended as well as the processing conditions.

However, these materials tend to suffer from heterogeneous compositions and structures, which lead to mechanical weaknesses. Product quality is generally measured in the laboratory, using destructive tests often requiring extensive sample preparation. On-line quality control would allow near-immediate feedback on the operating conditions and may be transferrable to an industrial production context.

The proposed research consists of developing an on-line quality control tool adaptable to plastic materials of all types. A number of infrared and ultrasound probes presently exist for on-line composition estimation, but only provide single-point values at each acquisition. These methods are therefore less adapted for identifying the spatial distribution of a sample's surface characteristics (e.g. homogeneity, orientation, dispersion).

In order to achieve this objective, a hyperspectral imaging system is proposed. Using this tool, it is possible to scan the surface of a sample and obtain a hyperspectral image, that is to say an image in which each pixel captures the light intensity at hundreds of wavelengths. Chemometrics methods can then be applied to this image in order to extract the relevant spatial and spectral features. Finally, multivariate regression methods are used to build a model between these features and the properties of the sample. This mathematical model forms the backbone of an on-line quality assessment tool used to predict and optimize the operating conditions under which the samples are processed.

Acknowledgements

This doctorate degree has certainly been one of the best moments of my life. Yes, I actually meant that seriously. And there are a few people I would like to thank for making it so.

Most importantly, I would like to express my deepest gratitude to my supervisor, Carl Duchesne, for his expert supervision, continuous guidance and unconditional support. In fact, the only derogatory comments he ever made did not even concern my work, but my performances on the squash court! I would also like to thank him for his genuine kindness and true faith in my abilities throughout the ups and downs of my past four years with him. For he has taught me much, both on and off the court, and for this I will forever be indebted.

I would also like to thank Denis Rodrigue, for his guidance throughout the course of my master's and doctorate degrees. I have learned much from him while under his supervision. However, as I look back on it now, I realize that much of his imparted wisdom came to me during what I shall call *productive disagreements*. To put it simply, I will miss his dry sense of humor, both in the lab and around the dinner table. This thesis would have never been the same without his precious input and I will always remember him as a mentor of the highest class.

Thanks to the graduate students in Chemical Engineering for friendships forged during countless coffee breaks spent discussing research, life and trivia. An exhaustive list is impossible, but Simon Leduc, Dominique Jean, Adrian Villanueva, Luisa Sancen, Jayson Tessier, Eddy Twite, Bora Aydin, Younes Leysi, Myriam Cousineau, Reza Barzegari, Karim Bourai, Amara Aït Aissa, Julien Lauzon, Pierre-Marc Juneau, Zahraa Waly and Vanessa Boudeau, to name a few.

Thanks to Rubén González, Milton Vasquez, Daniel Ramirez and Jorge Robledo for their kindness and hospitality. They made Guadalajara feel like home and Mexico a place I shall forever remember fondly.

I am very thankful to the Department of Chemical Engineering at Université Laval and the Natural Sciences and Engineering Research Council (NSERC) of Canada for their financial support, which relieved me from a lot of worries. I am also thankful to the Centre de Recherche sur les Matériaux Avancés (CERMA) for letting me use their installations and equipment.

I would also like to express my appreciation to my dissertation committee members: André Desbiens, Abdellah Ajji, André Zaccarin, Denis Rodrigue and Carl Duchesne.

Thanks to Susan, Marcel, Alex, Lucille, Ross, Lynda, Gabriel, Laura and Lydhia-Marie for giving me three great places to stay while I finished my thesis. Being homeless wasn't that bad after all!

A special thanks to Marie-Michèle for simply being there. She was always ready and willing to go over my work, to be my own personal cheerleader or to provide any culinary encouragement I may require. You make me glad I took up kickboxing.

And finally, I wish to extend my sincere and deepest gratitude to my parents, Susan and Marcel, for their continued interest, support and friendship in general and in regards to my studies in particular. I will always be grateful for all you have given me. For this, and so much more, I dedicate this thesis to you both. And don't worry if you don't get all the math… it's not that interesting anyway!

To all of you, thanks.

Oh, and Grandpa, thanks for taking me on all those walks in the woods when I was a kid. You always encouraged me to wonder about the trivial things that caught my attention. I learned that asking *"why?"* was a perfectly good way of showing my appreciation for the world around me.

I am more grateful than you could ever know.

Table of contents

List of tables

List of figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of abbreviations

AFM	Atomic force microscopy
CCD	Charged-coupled device
CWT	Continuous wavelet transform
DSC	Differential scanning calorimetry
DWT	Discrete wavelet transform
EVA	Ethyl vinyl acetate
FTIR	Fourier transform infrared
GLCM	Gray-level co-occurrence matrix
HDPE	High density polyethylene
LDV	Laser Doppler velocimetry
LDPE	Low density polyethylene
MIA	Multivariate image analysis
MIR	Multivariate image regression
MLR	Multi-linear regression
MRA	Multiresolutional analysis
MR-MIA	Multiresolutional multivariate image analysis
NMR	Nuclear magnetic resonance
NFC	Natural fiber composite
NIPALS	Non-linear iterative partial least squares
NIR	Near-infrared
ODMT	Oven-dried metric ton (2204 lbs)
PCA	Principal component analysis
PCR	Principal component regression
PE	Polyethylene
PET	Polyethylene terephthalate
PEPT	Positron emission particle tracking
PMMA	Poly(methyl methacrylate)
PLS	Projection to latent structures / Partial least squares
PP	Polypropylene
PS	Polystyrene
PVME	Poly(vinyl methyl ether)
RMSECV	Root mean squared error of cross-validation
SEM	Scanning electron microscopy
SVD	Singular value decomposition
TGA	Thermo-gravimetric analysis
VIP	Variable importance on the projection
WPC	Wood/plastic composite
WTA	Wavelet texture analysis

Chapter 1

Introduction

1.1 Polymers, blends and composites

A recent New York Times editorial crowned the common plastic window frame as the most common architectural feature of the early 21st century (New York Times, 2009). In fact, few would question the ubiquity of this modern *"wonder material"* (CBS News, 2007). The world production of plastics now surpasses that of steel, aluminum and copper combined (Throne, 1996). According to a report published by The Society of the Plastics Industry Inc., the US polymer industry employed more than 1.4 million workers and generated nearly $310 billion US in shipments (All Business, 2004). The report goes on to indicate that with another 770 000 workers indirectly employed by upstream suppliers of the plastics industry, the polymer industry employs about 2% of the US workforce.

The commercialization of novel polymer species, using new monomers and/or new polymerization routes, is restricted to a few specialty applications. The number of new polymer blends, however, has been increased continuously during the past few decades (He et al., 2004). Polymer blending is being viewed as a fast and convenient route for the development of new polymeric materials. It is now commonly used to achieve improved mechanical properties (Salamone, 1996; Chang et al., 2002; Nygard et al., 2005). In fact, polymer blend now constitute over 30% of the total film consumption (Utracki, 1998). In short, polymer blends are revolutionizing the industry due to their great versatility and capability to be tailored for specific end uses (Salamone, 1996).

According to a report by the Composites Fabricators Association, more than 13 000 composites processing facilities presently operate in the United States, employing 236 000 people and contributing $24 billion US to the economy (Stewart, 2002). Furthermore, this report also states that the American composites industry has grown from 113 million kg in 1960 to 1.9 billion kg in 2002, and is expected to continue to grow 4 to 5% annually in

coming years. Closer to home, the Canadian composites industry is estimated at $2.89 billion, 82% of which located in Quebec and Ontario (RICQ, 2001).

According to Rapra (2003) the primary reason for choosing to manufacture an article using polymer composites, rather than simply using neat polymers, is weight saving for a relatively high stiffness and strength. To illustrate this fact, a carbon fiber composite can be five times stronger than 1020 grade steel while having only one fifth the weight. Aluminum (6061 grade) has approximately the same density as the carbon fiber composite, but the composite can have twice the modulus and up to seven times the strength (Rapra, 2003).

Turning specifically to the case of wood/plastic composites (WPC) discussed in much detail in Chapter 7, there is a growing consumer acceptance of wood/plastic composite (WPC) products as replacements for more traditional materials, such as natural wood (Market Research, 2009). This is attributed to their long life span, minimal maintenance requirements, resistance to degradation caused by insect attack and exposure to the elements, and ability to be manipulated like natural wood. Moreover, the perception of these materials as environmentally friendly, due to their incorporation of recycled materials is likely to promote future gains in demand. For these reasons, the production of WPCs quadrupled between 1997 and 2000 (Li and Wolcott, 2004) and is expected to reach $2.4 billion US by 2011 (WRAP, 2003).

1.2 Quality control challenges

Polymer processing plants typically have tightly controlled operating conditions in which automatic feedback control loops on machine variables provide suitable results as long as there are no strict constraints on product quality (Tadmor et al., 1974). Nevertheless, variations in raw material properties may cause production to suffer from quality fluctuations that cannot be tempered by the machine variable controllers alone. To make matters worse, such fluctuations may lead to serious quality control issues if they are not be detected in a timely fashion (Ohshima and Tanigaki, 2000).

In practice, quality control of polymer products is often performed by off-line testing in the laboratory. Such tests involve large investments, long analysis times, sample destruction and handling by qualified personnel. Analytical testing can be thought of as a three-step process: pre-analytic activities (ordering the tests, collecting the samples, transportation, handling, storage, etc.), analytic activities (maintaining the laboratory instruments, managing supplies, ensuring adequate and competent personnel, actual sample testing, etc.), and finally post-analytic activities (timely reporting of the data, sending the report to the appropriate party, taking appropriate action when results exceed critical limits) (Lab Testing, 2009).

By providing unique and relevant information, these off-line tests are necessary and irreplaceable. However, since these analyses are time consuming and destructive, quality control is typically performed on a very low fraction of the production. It is therefore possible that significant amounts of low-grade, or defective, material may be produced before quality issues are detected. This problem may be addressed by estimating the quality on-line (through the use of remote sensing methods and inferential models) and including this information in quality control and monitoring schemes.

1.3 Remote sensing

Various spectroscopic techniques have been proposed for estimating on-line polymer quality since they are rapid and non-intrusive, they require little or no sample preparation, and provide a wealth of information (Lachenal, 1995). However, many of these methods are based on single point or averaged measurements. Estimating sample quality distribution over a larger area using these approaches is difficult to perform on a high-speed production line. At best, a coarse spatial coverage of the surface can be achieved using probe arrays. Nevertheless, fine spatial measurements greatly enhance the probability of identifying local weaknesses that may compromise the quality of the whole. The mechanical, optical and barrier properties of an entire sample may fall below specifications because of local discrepancies that may be overlooked by a coarse spatial analysis. Furthermore, rich spatial

data would also be useful for precise diagnostic of operational problems in the production line, such as non-uniform cooling.

In light of these facts, a quality control methodology based on a visible/near-infrared (VIS-NIR) hyperspectral imaging system is proposed. In essence, a spectral image captures a full spectrum for each pixel of the imaged surface. Such a device opens up a number of possibilities:

- A high-resolution camera with a large charge-coupled device (CCD) can identify local discrepancies, as well as large-scale variations, in the surface properties of the sample.
- The spectral information present in the image (i.e. color) can be extracted through chemometrics methods such as multivariate image analysis (discussed in Section 3.2.6), providing information relating to composition and crystallinity. These in turn can be used to determine the mechanical and barrier properties of the sample.
- The spatial information present in the image (i.e. patterns) can be extracted through chemometrics methods such as wavelet texture analysis (discussed in Section 3.2.7), providing information relating to surface roughness and composition distribution.
- The visible spectral camera can also be used to monitor the esthetic aspect of the sample. Whereas a common RGB camera quantifies the visual intensity of each pixel according to 3 wavelengths (red, green and blue), a spectral camera such as the one used in this study does so using many hundred wavelengths (discussed in Section 3.1), making it much more sensitive to subtle color variations.

1.4 Contributions and outline of the thesis

Contributions

This work contains essentially two main contributions. The first of these is the development of an on-line high-resolution spectral imaging sensor capable of simultaneously mapping sample composition, inferring material properties and monitoring surface quality in a rapid

and non-destructive manner. This tool has been developed specifically for the field of polymer science but can potentially be used in a wide variety of applications in which surface properties can be linked to characteristics of interest. In a context where industrial quality control is often performed using off-line analyses, such an imaging sensor makes it possible to monitor most, if not all, of the production, thus reducing costs, unnecessary delays and variations in quality. This imaging sensor was used to monitor three situations:

- Chapter 4 discusses the possibility of extracting the spectral information contained within a spectral image in order to identify variability patterns in the surface crystallinity of neat polymer films.
- Chapter 5 illustrates the use of both the spatial and spectral information contained in the images to infer the mechanical properties, composition and degree of segregation of polymer film blends.
- Chapter 7 uses both the spatial and spectral information found in the images to infer the crystallinity and mechanical properties of extruded wood/plastic composites as well as to determine the extrusion dynamics.

The second main contribution of this thesis is in the field of chemometrics. The high-resolution spectral imaging sensor used in this thesis has been invaluable in acquiring rich spatial and spectral material information. Nevertheless, the large quantities of data thus generated may be daunting. Moreover, spectral imaging system may be seen as too costly and complex for many industrial settings. In this context, it was realized that a simple wavelength selection methodology could mediate both these issues. In essence, by identifying the spectral intervals that are most highly correlated to the property to be estimated, one increases the interpretability of the model by linking these intervals to the presence, concentration and dispersion of known molecules or to physical states while reducing the complexity of the model by discarding less informative spectral intervals.

Existing methods for relevant variable/wavelength selection essentially fall into two categories: 1) methods based on *individual* wavelengths, and 2) approaches aiming at finding the most informative spectral *intervals*. The individual wavelength methods rank

the importance of the wavelengths, evaluating the correlation between each individual wavelength and a quality variable, and then using a cutoff criterion to segment relevant/irrelevant variables. These methods are often more difficult to interpret since the selected wavelengths are often distributed throughout the spectra instead of within a few well-defined intervals. Interval methods, on the other hand, search for informative groups of wavelengths within the spectra. Such groups, or spectral bands, can then be linked to known functional groups and should lead to more stable predictions. These methods are often very computationally intensive. The method proposed here combines some of the advantages of both individual wavelength and interval selection methods. While remaining simple and computationally efficient, it consistently identifies relevant and contiguous spectral intervals. This methodology is presented in Chapter 6 and is applied to 4 datasets in the polymer, oil and pulp and paper industries. It is also used in Chapter 7 to monitor the production of extrusion of wood/plastic composites.

Outline of the thesis

The experimental results obtained during this thesis are presented in four chapters, each based on articles published, accepted or submitted, to international scientific journals. In addition to these, other aspects of the project were published in scientific journals, proceedings and conference papers. The overarching theme of this Ph.D., including its different facets, is briefly discussed here.

This project concerns the quality control of plastic products using statistical imaging methods and essentially consists of five themes:

1. Raw material mixing. This aspect considers the first step of a blending process: the dry mixing of polymer powders prior to the actual processing step using a textural method. This portion of the work, however, is not presented in the thesis.
2. Crystallinity predictions. Considered here an initial step in the quality control of polymer composites, the determination of spatial crystallinity patterns across the surface of neat polymer films using spectral methods.

3. Polymer blend films. The case study of polymer-polymer blends: extrusion film-blowing of PE/PS blends using a combined spatial/spectral methodology.
4. Wood/plastic composites. The case study of polymer-filler blends: extrusion of wood/plastic composites using independent spatial and spectral methodologies.
5. Wavelength selection. A common problem encountered in spectral analysis is the overabundance of regressor data. In many cases, a significant portion of the spectral information is of little or no relevance to the sample quality. This led to the development of a wavelength selection methodology applicable to a broad range of datasets.

Raw material mixing

In the first part of this project, we investigated the transverse mixing of free flowing particles in a horizontal rotating drum. Such drums are widely implemented industrially as mixers, kilns, granulators, dryers and reactors; but it is principally for their mixing capabilities that they are used in the polymer industry. While the characteristics of any particulate process depend strongly on the dispersion of the raw substances, relatively little is known about the actual mixing dynamics. In fact, the time allotted to reach the required degree of dispersion is often based solely on empiricism and rules of thumb (Berthiaux et al., 2004; Ehrhardt et al., 2005). Better knowledge of this phenomenon would ensure that the chosen mixing time truly corresponds to the required mixing time, therefore saving time and energy.

According to the methodology proposed in this work, the degree of mixing is estimated by analyzing the dispersion of the free flowing particles as captured by an RGB camera located outside the rotating drum. In doing so, this method avoids possible flow disruptions caused by tracers (Woodle and Munro, 1993) or thief probes (Espinosa, 2003) without the added complexity of other non-invasive methods such as positron emission particle tracking (PEPT) and nuclear magnetic resonance imaging (NMR) (Ding et al., 2002; Ingram et al., 2005). In order to avoid this however, the proposed methodology requires the

camera to be in full view of the powder (i.e. the drum must be fitted with an observation window).

While material mixing is a necessary step in the creation of many plastic products, these results are not presented in the thesis. Nevertheless, three publications have been published on this work, the first of which being:

- Gosselin R., Duchesne C. and Rodrigue D., On the characterization of polymer powders mixing dynamics by texture analysis, *Powder Technol.*, 183, 177-188 (2008).

Determining the mixing dynamics of the drum essentially comes down to estimating the powder dispersion at constant sampling intervals throughout the mixing process. This was achieved by first imaging the powder through a glass mixing drum, and then quantifying the textural features of the powder blend at each sampling interval.

After the initial paper cited above, the project was taken over by Amara Aït Aissa, another Ph.D. student. While he was the main experimenter in the subsequent work, I contributed to the computational aspects of his work. This work was also presented in:

- Aït Aissa A., Gosselin R., Duchesne C. and Rodrigue D., Modeling the mixing of two polymer powders of different particle sizes, *ANTEC*, Milwaukee, 0819 (2008).
- Aït Aissa A., Cousineau-Pelletier M., Gosselin R., Duchesne C. and Rodrigue D., Mixing of polymer powders of different particle sizes in a rotating mold, *ANTEC*, Chicago, 0283 (2009).

In the initial paper, I developed the methodology (i.e. experimental setup), performed the experiments, wrote the necessary Matlab codes, developed the powder mixing model and reported the results under the guidance of my supervisors. Subsequent work by the aforementioned student was based on the framework I set up.

Crystallinity predictions

In the second part of this project, we investigated the spatial crystallinity variation across the surface of semi-crystalline polymer films. Widely used industrially, the total

consumption of semi-crystalline polymers such as polyethylene and polypropylene now reaches 90 millions tons annually (Murphy, 2003). Nevertheless, these polymers are sensitive to the conditions in which they are produced; processing pressures, cooling rates and sample geometry are all known to influence their degree of crystallinity, which in turn affects the mechanical, optical and barrier properties of the sample (Albano et al., 2003; Pantani et al., 2005; Shepherd et al., 2006).

The objective of this work is to develop a non-intrusive on-line sensor for monitoring spatio-temporal crystallinity variations across the surface of thin polymer materials. While some NIR probes have already been proposed for monitoring polymer crystallinity, they typically only provide single-point measurements and are therefore poorly suited to identify crystallinity gradients. This work focuses on combining NIR imaging spectroscopy, multivariate image analysis and regression methods to obtain the crystallinity spatial distribution of a sample. This methodology is illustrated using thin semi-crystalline samples (HDPE, LDPE, PP) produced via compression molding under different cooling conditions. Non-isothermal cooling rates throughout the cooling circuit ensure local crystallinity variations within a single sample. Four different methods are compared in this study.

- Regression between the NIR spectra and the crystallinity measurements.
- Regression between the 2nd order derivatives of the NIR spectra and the crystallinity measurements.
- MIA applied to the NIR spectra followed by a regression with the crystallinity measurements.
- MIA applied to the 2nd order derivatives of the NIR spectra followed by a regression with the crystallinity measurements.

This work was presented in:

- Gosselin R., Duchesne C. and Rodrigue D., On-line prediction of crystallinity spatial distribution across polymer films using NIR spectral imaging and chemometrics methods, *Can. J. Chem. Eng.*, 86, 869-878 (2008).

Some of these results were also discussed in:

- Gosselin R., Rodrigue D., González-Nuñez R., Duchesne C., Quality control of polymer products through spectral imaging and chemometrics methods, *ANTEC*, Chicago, 0281 (2009).

No other student participated in this project; I performed all the experimental, computational, modeling and reporting work under the guidance of my supervisors.

Polymer blend films

In the third part of this project, we proposed a real-time non-destructive sensor for monitoring polymer blend films. A number of techniques have recently been developed for on-line measurement of specific film properties (e.g. Raman spectroscopy, birefringence, Fourier transform infrared spectroscopy and X-ray scattering) but these probes either provide a rapid estimation of the properties of a single localized point on the sample or the average properties of a larger area.

Again based on a VIS-NIR spectral imaging sensor, the methodology is based on a multiresolutional multivariate image analysis (MR-MIA) technique was used for extracting both spectral and textural image features that are the most highly correlated with the film properties (i.e., composition distribution and mechanical properties), as well as for mapping them on the surface of the films. The proposed methodology is illustrated using low-density polyethylene/polystyrene (LDPE/PS) films produced by extrusion blowing under different operating conditions.

This work was presented in:

- Gosselin R., Rodrigue D., González-Nuñez R. and Duchesne C., Potential of hyperspectral imaging for quality control of polymer blend films, *Ind. Eng. Chem. Res.*, 48, 3033-3042 (2009).

Results were also discussed in:

- Gosselin R., Rodrigue D., González-Nuñez R., Duchesne C., Quality control of polymer products through spectral imaging and chemometrics methods, *ANTEC*, Chicago, 0281 (2009).

- Gosselin R., Duchesne C., Rodrigue D. and González-Nuñez R., Using Image Texture to Predict the Mechanical Properties of Polymer Films, *CEIMEXCAN*, Québec, 0013 (2007).

No other student participated in this project; I performed all the experimental, computational, modeling and reporting work under the guidance of my supervisors.

Wavelength selection

Finally, in the fifth part of this project, we proposed a wavelength selection methodology. Recent developments in spectral imaging for on-line inspection and quality control in a variety of fields lead to the acquisition of massive amounts of spectral data that must be processed quickly and efficiently. More parsimonious models obtained by relevant wavelength selection may result in enhanced interpretability, increased robustness, and significant reduction in computation time, which is crucial for on-line spectral imaging applications.

The main obstacle in achieving this lies in the high level of collinearity present between adjacent wavelengths throughout the spectra, which render standard statistical methods, such as multilinear regression, virtually useless. Methods for relevant variable selection essentially fall into two categories: approaches based on individual wavelengths vs. approaches aiming at finding the most informative spectral intervals. Whereas the former tends to identify significant wavelengths scattered throughout the spectrum, making interpretations difficult, the latter typically requires the user to arbitrarily split the spectrum into intervals. While effective, these methods also require some form of iterative searching for improving and reducing the number of intervals after the initial spectral splitting. This step can therefore be computationally intensive and time consuming.

In light of this, we proposed a Bootstrap-PLS regression methodology, using the VIPs as a selection criterion for selecting the significant wavelengths that were the most highly correlated with quality response variables. The advantage of this method lies in its

capability of identifying contiguous spectral intervals without the complexities mentioned above. This work was presented in:

- Gosselin R., Duchesne C. and Rodrigue D., A bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications, *Chemometr. Intell. Lab.*, 100, 12-21 (2010).

Results were also discussed in:

- Gosselin R., Rodrigue D., González-Nuñez R., Duchesne C., Quality control of polymer products through spectral imaging and chemometrics methods, *ANTEC*, Chicago, 0281 (2009).

No other student participated in this project; I performed all the experimental, computational, modeling and reporting work under the guidance of my supervisors.

Wood/plastic composites

In the fourth part of this project, we proposed an on-line imaging sensor for quality control of extruded WPCs. In line with the work on crystallinity patterns and polymer blend films, the methodology proposed in this project uses chemometrics methods to extract the spatial and spectral features that are most highly correlated with the material properties, and use this information to predict the properties of samples unknown to the model. For example, McAfee and McNally (2006) estimated the melt viscosity whereas Barnes et al. (2003) estimated the composition). While several sample properties have now been investigated on-line, we propose to simultaneously follow 7 material properties. These are: enthalpy of fusion (i.e. crystallinity), modulus, stress and strain at yield, stress and strain at break, as well as the energy at break. Furthermore, in this portion of the project, the extrudate was studied in both steady-state and transient state flow regimes, thus making it possible to determine process dynamics.

The results show that the proposed methodology can distinguish product quality in known samples, can predict the quality of unknown samples, can be used to estimate the extrusion

dynamics and can even provide an early detection scheme for undesired changes in feedstock or operating conditions. This work was presented in:

- Gosselin R., Duchesne C. and Rodrigue D., A hyperspectral imaging sensor for on-line quality control of extruded polymer composite products, *Comput. Chem. Eng., submitted: August 2009,* CACE-S-09-00415 (2009).

- Gosselin R., Rodrigue D. and Duchesne C., Property prediction of wood/plastic composites through spectral imaging and chemometrics methods, *WCCE 8*, Montréal, 1465 (2009).

Results were also discussed in:

- Gosselin R., Rodrigue D., González-Nuñez R., Duchesne C., Quality control of polymer products through spectral imaging and chemometrics methods, *ANTEC*, Chicago, 0281 (2009).

No other student participated in this project; I performed all the experimental, computational, modeling and reporting work under the guidance of my supervisors.

1.5 References

Albano C., Papa J., Gonzalez E., Navarro O. and Gonzalez R., Temperature and crystallinity profiles in polyolefines isothermal and non-isothermal solidification processes, *Eur. Polym. J.*, 39, 1215-1222 (2003).

All Business (2004).
www.allbusiness.com/manufacturing/chemical-manufacturing-paint/770578-1.html.

Barnes S.E., Sibley M.G., Edwards H.G.M. and Coates P.D., Applications of process spectroscopy to polymer melt processing, *Spectrosc. Eur.*, 15, 22-24 (2003).

Berthiaux H., Marikh K., Mizonov V., Ponomarev D. and Barantzeva E., Modeling continuous powder mixing by means of the theory of Markov chains, *Particul. Sci. Technol.*, 22, 379-389 (2004).

CBS News (2007).
www.cbsnews.com/stories/2007/06/21/sunday/main2962425.shtml.

Chang A.C., Tau L., Hiltner A. and Baer E., Structure of blown film from blends of polyethylene and high melt strength polypropylene, *Polym*er, 43, 4923-4933 (2002).

Ding Y.L., Forster R., Seville J.P.K. and Parker D.J., Segregation of granular flow in the transverse plane of a rolling mode rotating drum, *Int. J. Multiphase Flow*, 28, 635-663 (2002).

Ehrhardt N., Montagne M., Berthiaux H., Dalloz-Dubrujeaud B. and Gatumel C., Assessing the homogeneity of powder mixtures by on-line electrical capacitance, *Chem. Eng. Process.*, 44, 303-313 (2005).

Espinosa K., Thief sampling probe, *US Patent*, 6631650 (2003).

He Y., Zhu B. and Inoue Y., Hydrogen bonds in polymer blends, *Prog. Polym. Sci.*, 29(10), 1021-1051 (2004).

Ingram A., Seville J.P.K., Parker D.J., Fan X. and Forster R.G., Axial and radial dispersion in rolling mode rotating drums, *Powder Technol.*, 158, 76-91 (2005).

Lab Testing (2009).
www.labtestsonline.org.

Lachenal G., Dispersive and Fourier transform near-infrared spectroscopy of polymeric materials, *Vib. Spectrosc.*, 9, 93-100 (1995).

Li T.Q. and Wolcott M.P., Rheology of HDPE-wood composites I. Steady state shear and extensional flow. *Composites Part A.*, 35(3), 303-311 (2004).

Market Research (2009).
www.marketresearch.com/product/display.asp?productid=2401227.

McAfee M. and McNally G., Real-time measurement of melt viscosity in single-screw extrusion, *T. I. Meas. Control*, 28, 481-497 (2006).

Murphy J., *Additives for Plastics Handbook*, Elsevier Advanced Technology: UK (2003).

New York Times (2009).
www.nytimes.com/2009/02/15/arts/15iht-design16.1.20194500.html.

Nygard P.L., Jordy D.W., Knight T.W. and Veddder D.A., Process for extrusion and blowing of polyethylene film, *WIPO*, 005129 (2005).

Ohshima M. and Tanigaki M., Quality control of polymer production processes, *J. Process Contr.*, 10, 135-148 (2000).

Pantani R., Coccorullo I., Speranza V. and Titomanlio G., Modeling of morphology evolution in the injection molding, process of thermoplastic polymers, *Prog. Polym. Sci.*, 30, 1185-1222 (2005).

Rapra (2003).
www.rapra.net/vircon/1_2.asp.

RICQ – Étude sur les retombées économiques de l'industrie des composites au Québec (2001).
www.ricq.ca/File/etudes_economiques/KPMG_etude.pdf.

Salamone J.C., *Polymeric Materials Encyclopedia*, C.R.C. Press: USA (1996).

Shepherd J.E., McDowell D.L. and Jacob K.I., Modeling Morphology Evolution and Mechanical Behavior During Thermo-Mechanical Processing of semi-Crystalline Polymers, *J. Mech. Phys. Solids*, 54, 467-489 (2006).

Stewart R. US composite industry expands despite economy, *Reinf. Plast.*, 46(11), 20-22 (2002).

Tadmor Z., Lipshitz S.D. and Lavie R., Dynamic model of a plasticating extruder, *Polym. Eng. Sci.*, 14, 112-119 (1974).

Throne J.L., *Thermoplastic Foams*, Sherwood Publishers: USA (1996).

Utracki L.A., *Commercial Polymer Blends*, Chapman and Hall: United Kingdom (1998).

Woodle G.R. and Munro J.M., Particle motion and mixing in a rotary kiln, *Powder Technol.*, 76, 241-245 (1993).

WRAP, *Wood Plastic Composites Study Research Report – Technologies and UK Market Opportunities, Waste and Resources Action Programme*, Old Academy: United-Kingdom (2003).

Chapter 2

Background on polymer processing

2.1 Raw materials and processing

2.1.1 Raw materials used in this study

2.1.1.1 Consumer plastics

The term “plastic” is used here to refer to a wide range of polymers of high molecular weight. These can be truly natural molecules (e.g. shellac), chemically modified natural molecules (e.g. collagen) or fully synthetic molecules (e.g. polyethylene). Within this last category, we define thermoplastics (e.g. polypropylene, polyvinyl chloride) and thermosets (e.g. epoxy, polyurethane). If exposed to enough heat, thermoplastics will melt whereas thermosets remain solid until they burn. This difference is in large part due to their respective molecular weights. For thermoplastics, it typically ranges between 20 000 and 500 000, while thermosets can effectively be of “infinite” molecular weight due to irreversible chemical interconnections occurring on a macroscopic scale (Goodman, 2001). Finally, a polyolefin is a thermoplastic produced from a repetition of a single alkene (or olefin), whose general formula is C_nH_{2n} (e.g. polyethylene, polypropylene). This work will focus exclusively on three common consumer polyolefin thermoplastics: low-density polyethylene (LDPE), high-density polyethylene (HDPE) and polypropylene (PP). Nevertheless, the authors believe that the methodology could be extended to other polymer species.

Polyethylene is produced through the polymerization of the monomer ethene (also known as ethylene). As the ethene molecule (C_2H_4) lacks any substituent groups that may influence the stability of the propagation head of the polymer, it can be polymerized into several different geometries. Among these, HDPE essentially consists of long linear molecules with low degrees of short branching whereas LDPE has a high degree of short and long chain branching. This difference strongly influences the spatial organization of the

molecules. High degrees of branching reduce density, crystallinity and molecular interactions, all of which strongly influence the properties of the plastic.

On the one hand, HDPE is defined by a density greater than 0.941 g/cm^3. It is harder and has a greater tensile strength than LDPE, but less ductility. It has a wax-like, lusterless appearance and it is usually opaque. According to Harper (2002), the most common method for processing HDPE is blow molding where it is used to produce hollow sturdy packaging (e.g. detergent bottles, margarine tubs). On the other hand, LDPE is defined by a density range between 0.910-0.940 g/cm^3. It combines high impact strength, toughness and ductility to make it the material of choice of packaging films, one of its largest applications (Harper, 2002). While LDPE is opaque, it can be made transparent in thin films.

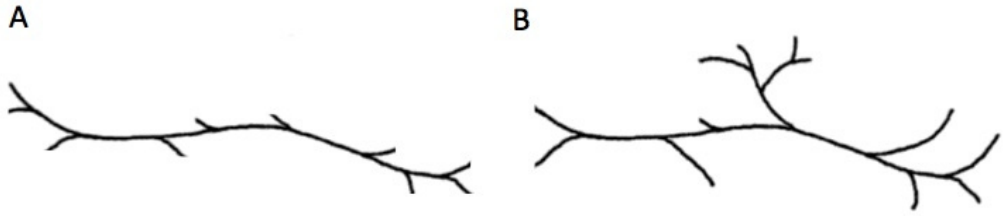

Figure 2.1. Schematic representation of PE branching in HDPE (A) and in LDPE (B).

Polypropylene is produced through the polymerization of the monomer propylene. It has a similar structure to that of polyethylene, except for the substitution of a hydrogen atom by a methyl group on every other carbon. PP is harder and has a higher softening point than PE, but is more prone to high-temperature oxidation (Harper, 2002). Polypropylene is used in a wide variety of applications, including packaging, textiles and a wide range of injection molded automotive components.

The global molded plastics market was estimated to exceed 156 million tons in 2005 (Murphy, 2003). Of this, the overall polyolefin consumption reached 117 million tons in 2007 (Townsend, 2008). Breaking it down further, all polyethylene (PE) resins account for 70 million tons (Townsend, 2008; Global Markets Direct, 2009), of which LDPE and HDPE respectively make up 18 and 29 million tons (Murphy, 2003). Most of the remaining

consumption of polyolefins is taken up by PP with an estimated 46 million tons in 2007 (Ceresana 2008; Townsend, 2008).

2.1.1.2 Polymer blends

Three types of blends are distinguished: miscible blends, partially miscible blends and immiscible blends. According to He et al. (2004), the high molecular weight of polymers makes the entropy of mixing negligible, ensuring the immiscibility of most polymer blends. Nevertheless, the weak adhesion present at polymer interfaces can effectively be improved using common compatibilization agents. Literature can now be found on a wide variety of polymer film blends: PE/PP (Minkova and Nikolova, 1989; Minkova et al., 1992; Rohe et al., 1999), PE/nylon (López-Barrón et al., 2007; Liu et al., 2009), PE/PET (Avila and Duarte, 2003), PS/PMMA (Leukel et al., 1998), PS/PVME (El-Mabrouk et al., 2007) as well as PP/EVA (Minkova and Nikolova, 1989). A number of more complex ternary blends have also been reported (Landry et al., 1991).

Chapter 5 will focus on the use of spectral imaging methods to predict the mechanical properties of PE/PS film blends produced via extrusion film-blowing. Such blends have also been extensively discusses in the literature: Dobreski and Nattinger, 1989; Zhu et al., 1989; Hermes et al., 1998; Nicole et al., 1999; Luzinov et al., 2000; Lorenzo et al., 2006; Díaz et al., 2007.

2.1.1.3 Wood/plastic composites

Natural fibers are biological structures mainly composed of cellulose, hemicellulose and lignin (Jodin, 1994). In much smaller proportions, there are also some extractibles, proteins and a certain number of inorganic products (Rowell et al., 1997). A number of these natural fibers are currently used as reinforcements, or simply as fillers, in polymer matrices; such materials are named natural fiber composites (NFC). These composite materials can be molded using many of the methods used to process common commercial plastics. Such a

symbiotic use of two distinct materials makes it possible to adjust the physical and economic properties according to requirements. Furthermore, NFCs are ecologically friendly and can play an important role in revalorizing household and farming waste products (e.g. plastic bottles and produce stalks). Often degraded and contaminated, these products have little commercial value taken individually but can acquire a certain value when combined. According to Chand and Fahim (2008), the most common natural fibers used industrially in NFCs are: bamboo, cotton, flax, hemp, jute, kenaf, sisal (leaf fiber) and wood. However, within the scope of this project, only wood fiber composites (WPC) are considered here but the authors believe that the methods could be extended to other products if a spectral contrast exists between the filler and the matrix.

Incorporating wood fibers into a polymer matrix greatly influences its mechanical properties. On the one hand, increasing fiber content will increase rigidity, while on the other hand strongly decrease ductility. Nevertheless, WPCs containing very high wood fractions often suffer from poor fiber-matrix cohesion, leading to decreased properties. According to Klason et al. (1984), this transition begins to occurs around 50% loadings, a level at which many polymer processing methods begin to encounter difficulties associated with high melt viscosity and fiber thermo-mechanical degradation. More about polymer processing methods is found in Section 2.1.2.

Wood fibers are common, affordable and light. The density of North American woods varies between 0.28 (corkbark fir) and 1.04 (black ironwood) (Albert Forest Products, 2009) and cost on average 0.22 $ US/kg (Wilkes et al., 2005). In comparison, calcium carbonate, agricultural fibers and fiberglass (all common fillers) cost respectively 0.15, 0.33-0.88 and 1.98 $US/kg (Wilkes et al. 2005). This may help explain increasing global demand for WPCs, which is estimated to reach 2.4 billion $US by 2011 (WRAP, 2003). This demand is mainly generated by 4 markets: gardening (patio, fences), construction (housing), infrastructure (maritime and railway) and finishes (auto industry).

Much research is currently being done on NFCs of all types. In 2009, there have been at least 28 published papers reporting on new development on WPCs alone, with many more reporting on other natural fiber composites. These include:

- 6 papers on HDPE/wood composites (Bouafif et al., 2009; Huang and Zhang, 2009; Sewda and Maiti, 2009; Xiong et al., 2009a,b; Zhang et al., 2009)
- 8 papers on PP/wood composites (Azizi and Ghasemi., 2009; Bouza et al., 2009; Chauhan et al., 2009; El-Sabbagh et al., 2009; Ghasemi et al., 2009; Haque et al., 2009; Nourbakhsh and Kouhpayehzadeh, 2009 ; Soury et al., 2009)
- 2 papers on LDPE/PP/wood composites (Tademr et al., 2009; Younesi and Bahrololoom, 2009)
- 3 papers on PVC/wood composites (Marathe and Joshi, 2009; Rocha et al., 2009; Sadler et al., 2009)
- 5 papers on polyurethane/wood composites (Casado et al., 2009; Gao et al., 2009a,b; Mosiewicki et al., 2009; Yuan and Shi, 2009)
- 3 papers on polylactic acid/wood composites (Gregorova et al., 2009; Noël et al., 2009; Pilla et al., 2009)
- 1 paper on polyamide/wood composites (Alvarez de Arcaya et al., 2009)

In spite of this interest, several problems must be overcome during the molding process. The most important of which is the thermo-mechanical degradation of the fibers. Many natural fibers loose their rigidity around 160°C and suffer from lignin degradation around 200°C (Michaud, 2003). This problem can be avoided by selecting a plastic with a low melting point, reducing processing time and avoiding processing steps involving very high shear rates. The relation between the amount of thermo-mechanical energy delivered to the system and the subsequent fiber degradation is discussed in Chapter 7. Other problems include water retention by the fibers, hydrophilic/hydrophobic incompatibility at the wood/plastic interface as well as heterogeneous composition throughout the structure. While the wood fibers can be dried prior to blending, and coupling agents can be used to strengthen the interface, adequate fiber dispersion is both difficult to quantify and ensure. Poor dispersion ultimately leads to local weaknesses that may compromise the quality of

the whole. Attempts to determine such characteristics are discussed in Section 2.2.3 using soft-sensing measurements.

2.1.2 Common polymer processing methods

2.1.2.1 Overview of polymer processing

According to the Handbook of Industrial Engineering and Operation Management by Salvendy (2001), there are only a limited number of basic processes commonly used in the production of thermoplastic and thermosetting plastics parts:

- Compression molding. An appropriate amount of plastic (often in powder form) is introduced into a mold, which is subsequently closed, heated and put under pressure. The plastic is softened by the heat and formed into a continuous mass of the size and shape of the mold cavity.
- Transfer molding. In this two-step process, the plastic is first melted in an auxiliary chamber before being forced into a closed mold cavity through high pressure. The polymer is allowed to cool under pressure, before being ejected from the mold.
- Extrusion. This process involves the continuous forming of a shape (infinitely long profile) by forcing softened plastic through an orifice. Essentially, an extruder consists of a heated barrel in which a single- or twin-screw system continuously conveys the plastic toward the forming die.
- Injection molding. This semi-batch process is essentially a hybrid between transfer molding and extrusion. To achieve this, the barrel is fitted with a reciprocating screw system, making it possible to store molten polymer near the die of the barrel. When sufficient material has been accumulated, it is forced into a closed mold under pressure.
- Casting. Much like the casting of metals, it involves introducing molten plastic into a mold that has been shaped to contour the piece to be formed.

- Cold molding. In this process, a thermosetting polymer is introduced into a room-temperature mold under pressure. The mold is subsequently opened and the sample is transferred to an oven where it is baked until hardening.
- Calendering. A continuous production of a thin sheet by passing thermoplastics between a series of heated rollers.
- Thermoforming. Restricted to thermoplastic polymers, this is a process in which a plastic sheet is heated to a pliable temperature, formed to a specific shaped using a mold contour, and trimmed to create the finished product.
- Blow molding. In this process, a tube of molted plastic is extruded vertically and incased in a split mold. Air is then injected into this hot section of plastic, forcing it to follow the contour of the mold.
- Rotational molding. In this high-temperature, low-pressure process, biaxial rotation is used to produce hollow single piece parts.

Of these methods, compression molding is used in Chapter 4 to produce thin plastic samples while extrusion is used in Chapters 5 and 7 to produce blown films and wood/plastic composites, respectively. Both methods will be discussed here in greater detail. Nevertheless, the proposed on-line quality control scheme could be applied to any of these approaches since it is based on the product itself, as it comes out of the process.

2.1.2.2 Compression molding

Compression molding is a common industrial process in which a plastic is placed into a heated metal mold, softened by the heat, and forced to conform to the shape of the mold cavity as the mold closes. Pressure is applied to force the material into contact with the entire cavity surface. Both heat and pressure are maintained long enough to remove prior melt history. The mold is cooled under pressure, often through water quenching, before being opened to remove the complete sample (Figure 2.2). Compression molding is known for its ability to mold large complex parts at a relatively low cost. Furthermore, it produces little waste and is capable of molding small quantities, giving it an advantage when

working with expensive compounds. Lastly, it does not create weld lines and causes little or no shear stresses during the molding process.

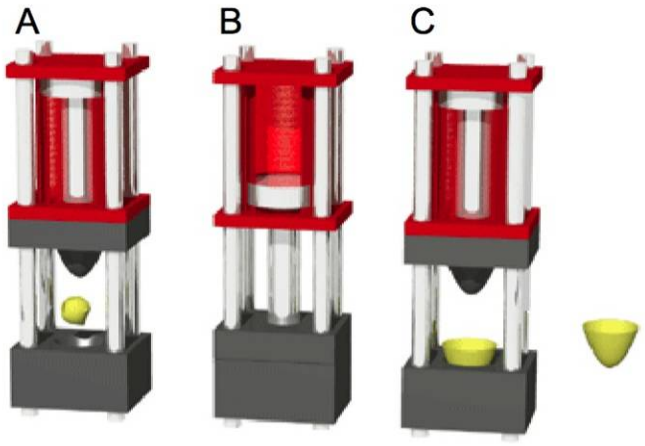

Figure 2.2. Illustration of a compression molding process in which (A) a measured charge is placed between two matching metal molds, (B) the molds close and the sample is subjected to heat and pressure, and (C) the cured part is ejected and trimmed. Adapted from Tecrep Engineering (2009).

2.1.2.3 Extrusion

Plastic extrusion is a high-volume manufacturing process used to produce indefinitely long objects of a fixed cross-section profile such as pipes and window frames. An extruder is basically an Archimedes' screw rotating continuously inside a heated barrel (Figure 2.3). The material enters through the hopper and comes in contact with the screw system, where it is heated and forced forward into the barrel. The screw system itself can vary widely, going from single- to double-screw, co- to counter-rotating, as well as intermeshed or not. Regardless of its geometry, the system is usually divided into three zones: feeding, compression and metering (Figure 2.3). The feeding zone has a wide melt channel (thread gap) in order to receive the plastic pellets, and is used to melt the plastic. As its name indicates, the compression zone constricts the melted plastic, either by reducing the barrel section or the melt channel. This step releases any gases and volatiles that may be present. Finally, the plastic is homogenized in the metering zone. According to Kasliwal and Jones (2004), a proper mixing process must be both dispersive and distributive. In dispersive

mixing, cohesive resistances have to be overcome to achieve finer levels of dispersion; while distributive mixing aims at improving the spatial distribution of the components without cohesive resistance playing a role (Gramann and Rauwendaal, 1998). In order to achieve this, a breaker plate or kneading blocks are often added to increase shear stresses. The plastic is then shaped as it exits the extruder die. A wide range of post-processing steps (e.g. cooling, cutting, thermoforming) is then performed on the extrudate.

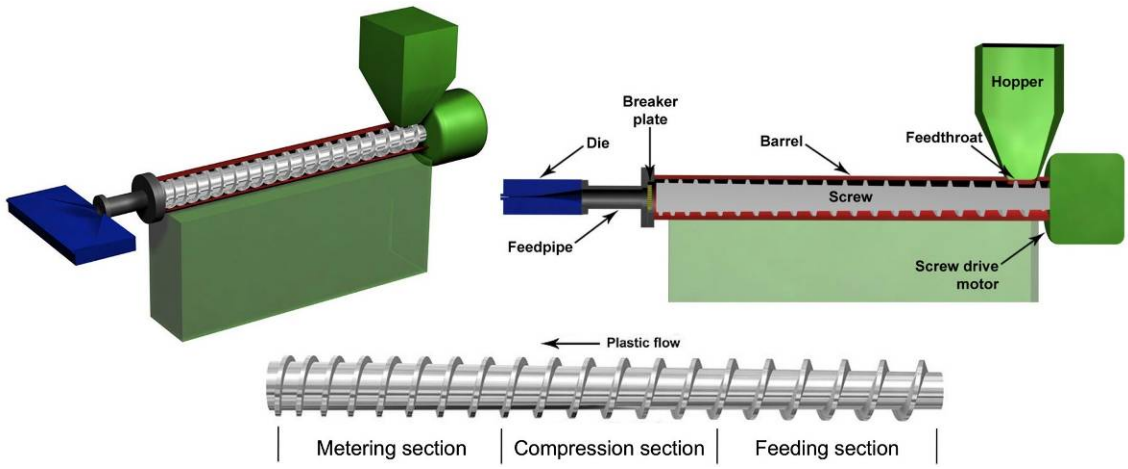

Figure 2.3. Illustration of an extrusion process. Polymer is fed in the hopper, conveyed through the barrel by the screw and shaped by the die. Adapted from Wikipedia Plastics Extrusion (2009).

Extrusion film-blowing

Extrusion film-blowing is widely used to make plastic films for the packaging industry. It is a process of forming thermoplastic film in which a plastic tube is continuously extruded and inflated to several times its initial diameter by internal air pressure, cooled and collapsed by rollers (Figure 2.4). The film is usually extruded vertically upward and air is blown in the center of the extruded tube as it emerges from the die (Figure 2.4, inset). Film thickness is controlled by varying the internal air pressure, as well as the rate of extrusion.

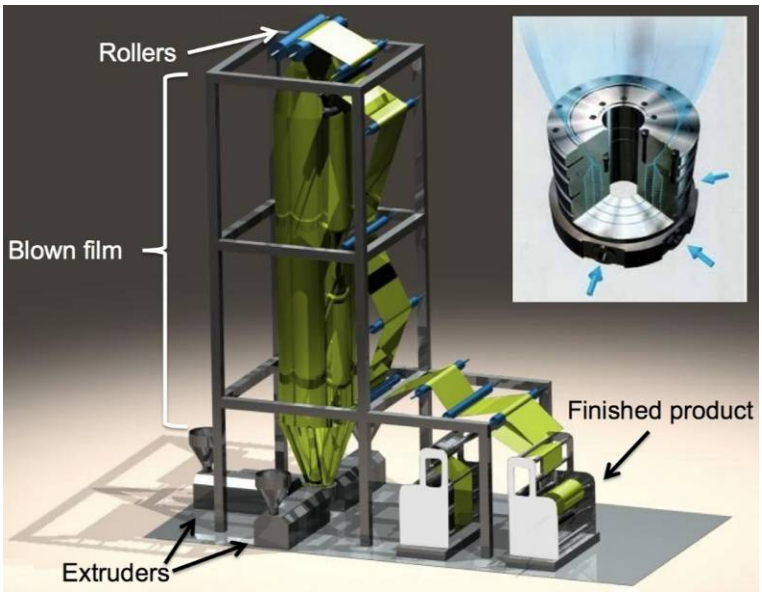

Figure 2.4. Illustration of an extrusion film-blowing process in which the polymer is first extruded, blown into a film (inset) and rolled into the final product. Adapted from Antistaticbar (2009) and Berg (2009).

2.1.3 Problems

According to Gillespie (1988), the compression molding cycle is simple enough that problem solutions are generally obvious. Measuring load and thermal parameters can resolve many difficult problems. Nevertheless, unforeseen problems arise due to processing or material variability, operator related defects, or machine wear. Specifically, variability is typically related to material inconsistencies from the supplier, storage history, or extreme factors such as weather (Davis et al., 2003). These situations lead to poor product consistency and difficulties in controlling waste through mold flashing.

Extrusion is basically a steady state process and most extruders are controlled to maintain relatively constant operating conditions. According to Tadmor et al. (1974), common feedback control loops on the machine variables (i.e. temperature, screw speed) perform well under most circumstances, providing that there are no stringent quality constraints. Nevertheless, the performance of an extruder is mainly determined by the characteristics of the feedstock and less by those of the machine. Material properties are seldom uniform. According to Rauwendaal (2001), the extrusion process may be affected by the feedstock's

bulk properties (e.g. particle size), melt flow properties (e.g. viscosity), as well as thermal properties (e.g. thermal diffusivity). Furthermore, the material samples used for testing may not be representative of the sample as most tests are performed on 0.01 kg or less, while most industrial extruders run at a throughput of several hundred kilograms per hour. This problem may be exacerbated when processing naturally locally heterogeneous materials such as polymer blends or composites.

Beyond the problems commonly associated to the extrusion process, extrusion film-blowing suffers from a host of specific problems. The first of these is a distinct sensitivity to un-melted plastic polymers in the extrudate appearing as "gels" in the final product (Perdikoulias and Vlcek, 2004). Also, fluctuations in the film thickness can be caused by inconsistently shaped pellets in the feedstock (Cantor, 2006), as well as ineffective film cooling (Giles et al., 2005). Lastly, Reifenhauser and Rahlfs (1975) observed that there may be a mismatch in (1) the rate at which the rollers are run, (2) the volumetric extrudate flow and (3) the pressure drop within the film. They suggest that this mismatch may cause unwanted uniaxial or biaxial orientation in the films.

These examples show that feedback and/or feed-forward control loops on the machine variables are not suited to identify all the problems affecting the polymer industry. Unwanted sample variability is still very much present even under the best operating conditions. The next section discusses the notion of quality, as well as means of measuring it, both off- and on-line.

2.2 Quality

2.2.1 Defining quality

Broadly speaking, the term "quality" designates the degree of excellence of an item when measured against other of a similar kind. This definition however can be interpreted in a number of ways; 5 general definitions are discussed here. Customer-based quality (1), often coined the "fitness to use" compares the item to customer expectations (Edwards, 1968;

Gilmore, 1974; Juran and Godfrey, 1999). Manufacturing-based quality (2) compares the item to its design, or specifications, and ensures "conformance to specifications" (Gilmore, 1974; Crosby, 1980). Value-based quality (3) compares an item's usefulness to its price (Broh, 1982; Feigenbaum, 1991). Applying this definition, items qualified as "low-quality" by the two previous definitions may nevertheless be deemed satisfactory if their price is set sufficiently low. Product-based quality (4) refers to an attribute of the item that makes it superior in any way to otherwise comparable items (Abbott, 1956; Leffler, 1982). Finally, a transcendence-based quality (5) is a much harder characteristic to define but is generally linked to a consumer's immediate and irrational interest for a product (Tuchman, 1980). Keeping with the scope of the project, a manufacturing-based definition of quality must necessarily be used here.

Quality control is defined by the American Society for Testing and Materials (ASTM) as a method of implementing the quality specifications for raw materials, intermediate and finished products as established by management, and carries out the operational techniques and the activities that sustain quality of a product or service (ASTM, 1992). Applying this to the polymer industry, sample quality typically refers to thermal properties, mechanical properties and component distribution. According to the Handbook of Industrial Engineering and Operation Management (Salvendy, 2001), the functional designer in the polymer industry is primarily concerned with the specific gravity, hardness, water absorption, outdoor weathering, coefficient of linear thermal expansion, elongation, flexural modulus, IZOD impact, deflect temperature under load and flexural yield, tensile, shear, and compressive strengths.

In the following section, a number of off-line and on-line methods of measuring sample quality are discussed.

2.2.2 Laboratory quality measurements

According to Dumitrescu et al. (2005), quality control in the polymer industry is often limited to visual inspection of the sample, weighing, and, in a relatively small subset of the

production, mechanical and analytical testing may also be carried out. The term "analytical" is used here to refer to any analysis seeking to determine the composition of a sample. While this description is accurate, it may give an over-simplified view of the present state of off-line quality control in the polymer industry. In fact, a wide range of laboratory analyses has been developed. Generally speaking, these tests fall into 7 categories: mechanical, electrical, rheological, analytical, thermal, optical and aging (Intertek Plastics Technology Laboratories, 2009).

- Mechanical tests. The mechanical properties usually represent the response of a standardized section of the sample to a known mechanical constraint. This category includes, for example, the tensile modulus ASTM D638, Izod impact resistance ASTM D256 and Rockwell hardness ASTM D785.
- Electrical tests. These tests usually seek to quantify the resistance or capacitance of the heart and surface of the sample. Among these we find the evaluation of the dielectric constant ASTM D150, as well as the electrical resistance of the heart and/or the surface ASTM D257.
- Rheological tests. The study of the rheological properties enables the user to determine behavior, as well as thermal stability, of the polymers during the molding process. Whereas mechanical tests are performed on solid samples, this category only concerns their melt flow properties. Viscosity measurements can be obtained using single-point measurements such as the Melt Flow Index (MFI) ASTM D1238 or through variable shear rates tests like capillary rheometry shear sweeps ASTM D3835.
- Analytical tests. The analytical properties represent the nature and composition of the material. This type of analysis includes differential scanning calorimetry (DSC) ASTM D3417, Fourier Transform Infrared (FTIR) ASTM E125 and Thermo-Gravimetric Analysis (TGA) ASTM E1131. These tests commonly involve very small amounts of material but require highly delicate handling.
- Thermal tests. These methods quantify the behavior of polymers using a wide range of temperatures. Common tests include the thermal expansion coefficient ASTM E831, heat deflection temperature ASTM D648, as well as the determination of

different transition points (i.e. temperature of fusion) by DSC ASTM D3417. These tests are often carried out under a neutral atmosphere to minimize polymer oxidation at high temperatures.

- Optical tests. The most common optical tests concern the gloss ASTM D2457, haze ASTM D751, yellowness index ASTM D313 and RGB color values ASTM D6290. Nevertheless, this category remains relatively unused in the field.
- Aging tests. There exist a number of methods able of accelerating the aging process of polymers by exposing them to harsh conditions (i.e. high temperature, water or electromagnetic radiation). Among these we find the use of a xenon arc ASTM D2565 or fluorescent lighting ASTM D4329. Both require between 500 to 2000 hours to yield complete results.

While these seven categories of tests are well documented for neat polymers, few regulations apply to polymer blends and composites. Nevertheless, the Plastic Lumber Trade Association is working to instate an ASTM norm on HDPE based WPCs. Moreover, the Canadian Construction Materials Centre, a branch of the National Research Council, is also working to allow composite materials, such as WPCs, to be used in construction (Environment and Plastics Industry Council, 2003).

2.2.3 Soft-sensing quality measurements

As seen in Section 2.2.2, the numerous steps involved in off-line laboratory measurements often limit the number of samples tested to a level where one can legitimately ask whether the test set truly represents the whole of the production. This problem may be addressed by taking on-line quality measurements (or estimates) of the samples and including this data in the control mechanism. According to Ohshima and Tanigaki (2000), the lack of on-line sensors for key polymer properties leads to quality control issues in polymer plants. This situation leads to a conundrum: the characterization of a sample must be sufficiently complete to yield relevant information but fast enough to be of any use in automatic feedback control loops. In order to achieve this goal, several rapid and non-destructive so-called soft-sensing methods have been proposed.

According to Zhong and Yu (2000), the basic idea behind remote-sensing, also known as inference estimation, is to predict the output of the variables of interest (i.e. sample purity) by measuring a set of secondary easily measured variables (i.e. spectra), which are nevertheless correlated to the variables of interest. In essence, the core of a soft sensor is the soft sensing model of a plant, which generates a virtual measurement to replace a real sensor measurement (Yan et al., 2004). In recent years, considerable attention has been paid to the development of on-line soft-sensors in order to estimate hard-to-measure quantities or simply avoid laboratory measurement delays (Dai et al., 2006). The following sections discuss a number of soft-sensing methods and their uses in the polymer industry. Two cases of soft-sensing are discussed. Analysis of the melt properties is broadly referred to as "internal probes", whereas analysis of the solid polymer is referred to as "remote sensing".

2.2.3.1 Internal probes

Internal probes have the potential of characterizing the state of a production line without influencing the process or requiring any direct human interaction. A number of authors have studied the viability of electromagnetic and ultrasonic probes on plastic extrusion.

NIR spectroscopic probes have been used successfully to estimate extrusion dynamics (Ohshima and Tanigaki, 2000; Alig et al., 2005), to control the production of complex polymer blends (Xanthos and Patel, 1995; Rohe et al., 1999; Coates et al., 2003; Dumitrescu et al., 2005) as well as natural fiber composites (Fischer et al., 2006). Raman and FTIR spectroscopic probes have also been used on extruded blends (Coates et al., 2003; Alig et al., 2005; Dumitrescu et al., 2005). In general, spectroscopic probes were used to predict the composition, color and humidity levels. For their part, ultrasonic probes, based on the attenuation of ultrasounds inside the polymer, have often been used to estimate blend composition (Gilmore and Tjahjidi, 2002; Coates et al., 2003; Alig et al., 2005; Fischer et al., 2006), melt density (Abu-Zahra and Seth, 2002), as well as the residence time distributions inside the extruder (Sun et al., 2004; Lee et al., 2005). In contrast with ultrasonic probes, which measure the average velocity of the melt (Agemura

et al., 1995), laser Doppler velocimetry (LDV) measure single point velocities inside the extruder (Furness, 1989). Finally, more exotic probe types such as nuclear magnetic probes (NMR) have been proposed to study the molecular mobility via the relaxation times (Gottwald and Scheler, 2005).

These examples show that internal probes can effectively quantify the polymer melt as a whole. Nevertheless, certain limitation must be noted. While electromagnetic probes provide a wealth of information, they require translucent melts, therefore limiting their usefulness when intense coloring and/or particulate fillers are present. On the other hand, ultrasonic probes are not limited to translucent melts but do not provide the same level of detail that electromagnetic probes do.

One must also note that all single-point probes, by their very nature, only provide local or global information on the molten flow (only true for homogeneous flows). This implies that such a probe-based system may not detect local discrepancies or gradients, while the blend is still molten and is not designed to provide information on the final molded sample.

2.2.3.2 Remote sensing

Microscopy

Microscopic methods will broadly be considered as “remote sensing” in spite of the fact that many of these require intensive sample preparation. In fact, microscopy methods vary widely in their capabilities and applications. Of these, atomic force microscopy (AFM), scanning electron microscopy (SEM) and optical microscopy have all been used in micro-structural characterization of polymers (Kantz et al., 1972; Butler et al., 1995; Pantani et al., 2005; Wang and Hahn, 2007). These methods have slow acquisition times and may require surfacing, microtoming or vacuum conditions. However, on-line monitoring of multiphase polymer melts using optical microscopy has nevertheless been published (Xanthos and Patel, 1995; Leukel et al., 1998).

Spectroscopy

Various spectroscopic techniques have also been proposed for estimating polymer properties since they are rapid and non-intrusive. They also require little or no sample preparation and provide a wealth of information (Lachenal, 1995). However, many of these methods are based on single point measurements.

First, among the spectroscopic methods, NIR spectroscopy has been very widely used to analyze polymeric materials. Applications include: the extent of crystallization (Lachenal, 1995; Stuart, 1996; Sato et al., 2003; Pantani et al., 2005; Pelsoci, 2007), orientation (Buffeteau et al., 1995), density (Saeki et al., 2003; Watanabe et al., 2006), film thickness (Heymann et al., 2009), as well as waste identification (Huth-Fehre et al., 1995; van den Broek et al., 1995,1996,1997; Feldhoff et al., 1997; Leitner et al., 2003).

Among these applications, waste identification is of particular interest because it relies on chemometrics methods (i.e. PLS models) to identify plastic waste products using spectroscopic NIR images. These projects essentially tried to identify the nature, shape and size of the plastic objects. For example, van den Broek et al. (1996) were able to isolate a number of waste plastic products (17 polymer types) from other waste products (ceramics, metal, glass, paper and wood). However, the weak spectral resolution used in this study made it impossible to distinguish between the different polymer types. A certain level of success was achieved the very next year by Feldhoff et al. (1997) using uncolored polymer samples. However, an efficient polymer classification scheme would have to wait the work of Leitner et al. (2003), in which they used blob analysis to identify the composition of a number of waste plastic bottles. They effectively distinguished between PET, HDPE, PS and two types of PP with 93% accuracy (Figure 2.5). This method effectively seeks to match the experimental spectrum of each pixel in a spectral image to that of known reference samples. While it can effectively identify the nature of the elements present in the image, it is not suited to take into account intra-class variability.

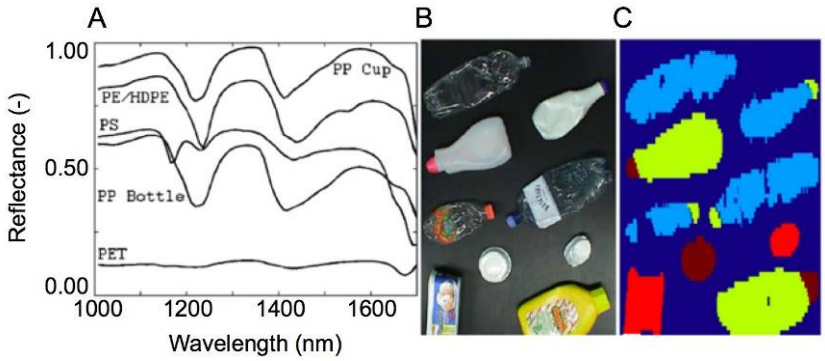

Figure 2.5. The different NIR signatures of common polymer species (A) make it possible to class household plastic waste (B) according to composition (C). Adapted from Leitner et al. (2003).

Second, Raman spectroscopy is another spectroscopic method commonly used in polymer analysis. It uses the inelastic scattering of light to analyze rotational and vibrational modes of molecules. Among recent applications, Alig et al. (2005) used single-point laser Raman spectroscopy for on-line monitoring of an extrusion process. Work has also been reported on real-time crystallinity estimation by single-point laser Raman spectroscopy (Stuart, 1996; Batur et al., 1999; Cherukupalli and Ogale, 2004).

The third spectroscopic method mentioned here is laser-induced breakdown spectroscopy (LIBS). This technique uses laser-induced plasma as to vaporize, atomize and excite the sampled material to determine its composition based on elemental and molecular emission intensities. It has notably been used to identify a number of polymer types (Wang et al., 2008), as well as to perform real-time composition estimation of wood/plastic composites (Rials et al., 2009).

The fourth and last spectroscopic method discussed here is dielectric spectroscopy. Sometimes called impedance spectroscopy, it measures the dielectric properties of a medium as a function of frequency (Kremer et al., 2002). While it is still hardly used for polymer characterization, it has been used successfully by Boersma and van Turnhout (1999) as well as Alig et al. (2005) for on-line monitoring of an extrusion process.

X-ray scattering

X-ray scattering techniques are non-destructive methods based on the scattered intensity of a beam hitting a sample and have been extensive used in polymer analysis (Kantz et al., 1972; Brown et al., 1973; Polizzi et al., 1991; Albano et al., 2003; Alig et al., 2005; Pantani et al., 2005; Pelsoci, 2007; Briskman, 2007). X-ray scattering reveals information about the crystallographic structure, chemical composition, as well as the physical properties of the material. However, as these methods depend on the incident and scattering angles, as well as the polarization and wavelengths of the beams, they are not easily adapted to 2-D surface scans.

Other methods

Non-invasive nuclear magnetic resonance imaging (NMR) and positron emission particle tracking (PEPT) have been used to follow the trajectory of specific particles as they move inside the bulk (Agemura et al., 1995; Nakagawa et al., 1997; Ding et al., 2002; Ingram et al., 2005; Fischer et al., 2006). While these methods provide much physical, chemical and structural information about the polymer melts (Bergmann, 1981; Pelsoci, 2007), their use is typically expensive and time consuming (Vailaya et al., 2001; Saeki et al., 2003).

Ultrasonic probes are usually associated to the monitoring of extruded polymer melts. Nevertheless, Alig et al. (2007) proposed a remote sensing ultrasonic application in which a plastic film's formation was studied. Finally, white light interferometry has also been suggested by Coates (2005) to measure surface roughness across extruded PVC strips.

Closing thoughts on quality measurements

A number of remote sensing techniques used in polymer analysis were discussed in this section. The stated objective of this work was to develop an on-line quality control tool adaptable to plastic materials of all types. First, a discussion was made on laboratory measurements. It can be concluded that they are usually too slow, and require too much

handling, to be used for on-line quality assessment in an industrial setting. Nevertheless, laboratory measurements are essential since they are uniquely capable of quantifying the material properties.

Remote sensing tools are better adapted to the fast response times required by the industry. The real-time methods discussed in this section are both fast and accurate, but typically only provide a single local, or averaged, value at every sampling interval. Internal probes have the added inconvenience of analyzing the polymer melt rather that the final solidified sample, making them oblivious to any post-processing problems (e.g. shrinking). It will be shown that a number of chemometrics methods can effectively be used to extract relevant spectral and spectral features from hyperspectral images of the polymer composites. However, before doing so, both the images and the methods must be discussed in greater detail (Section 3.1 on spectral imaging and Section 3.2 on chemometrics).

2.3 References

Abbott L., Quality and competition: an essay in economic theory, *Ann. Am. Acad. Polit. S. S.*, 304, 158-159 (1956).

Abu-Zahra N.H. and Seth A., In-process density control of extruded foam PVC using wavelet packet analysis of ultrasound waves, *Mechatronics*, 12, 1083-1095 (2002).

Agemura C.K., Kauten R.J. and McCarthy K.L. Flow fields in straight and tapered screw extruders using magnetic resonance imaging, *J. Food Eng.*, 25, 55-72 (1995).

Albano C., Papa J., Gonzalez E., Navarro O. and Gonzalez R., Temperature and crystallinity profiles in polyolefines isothermal and non-isothermal solidification processes, *Eur. Polym. J.*, 39, 1215-1222 (2003).

Albert Forest Products (2009).
www.albertforestproducts.com/density.htm

Alig I., Fischer D., Lellinger D. and Steinhof B., Combination of NIR, Raman, ultrasonic and dielectric spectroscopy for in-line monitoring of the extrusion process, *Macromol. Symp.*, 230, 51-58 (2005).

Alig I., Oehler H., Lellinger D. and Tadjbach S. Monitoring of film formation, curing and ageing of coatings by an ultrasonic reflection method, *Prog. Org. Coat.*, 58, 200-208 (2007).

Alvarez de Arcaya P., Retegi A., Arbelaiz A., Kenny J.M. and Mondragon I., Mechanical properties of natural fibers/polyamides composites, *Polym. Composite*, 30(3), 257-264 (2009).

Antistaticbar (2009).
www.antistaticbar.org/plastic-applications/blown-film-line-application-sheet.html.

ASTM (American Society for Testing and Materials), *The Role of Sensory Analysis in Quality Control*, *ASTM Manual series MNL 14*, J.E. Yantis Editions: USA (1992).

Avila A.F. and Duarte M.V., A mechanical analysis on recycled PET/HDPE composites, *Polym. Degrad. Stabil.*, 80, 373-382 (2003).

Azizi H. and Ghasemi I., Investigation on the dynamic melt rheological properties of polypropylene/wood flour composites, *Polym. Composite*, 30(4), 429-435 (2009).

Batur C., Vhora M.H., Cakmak M. and Serhatkulu T., On-line crystallinity measurement using laser Raman spectrometer and neural network, *ISA Trans.*, 38, 139-148 (1999).

Berg (2009).
www.berg-group.com/blownfilmcoolers.htm.

Bergmann K., Determination of polymer crystallinity from proton solid-echo NMR measurements, *Polym. Bull.*, 5, 355-360 (1981).

Boersma A. and van Turnhout J., Dielectric on-line spectroscopy during extrusion of polymer blends, *Polymer*, 40, 5023-5033 (1999).

Bouafif H., Koubaa A., Perré P., Cloutier A. and Riedl B., Wood particle/high-density polyethylene composites: Thermal sensitivity and nucleating ability of wood particles, *J. Appl. Polym. Sci.*, 113(1), 593-600 (2009).

Bouza R., Pardo S.G., Barral L. and Abad M.J., Design of new polypropylene-wood flour composites: processing and physical characterization, *Polym. Composite*, 30(7), 880-886 (2009).

Briskman B.A., Be careful when using X-ray exposure analysis for polymers, *Nucl. Instrum. Meth. B*, 265, 276-280 (2007).

Broh R.A., *Managing Quality for Higher Profits*, McGraw-Hill: USA (1982).

Brown D.S., Fulcher K.U. and Wetton R.E., Application of small angle X-ray scattering to semi-crystalline polymers. 1. Experimental considerations and analysis of data, *Polymer*, 14, 379-383 (1973).

Buffeteau T., Desbat B. and Bokobza L., The use of near-infra-red spectroscopy coupled to the polarization modulation technique to investigate molecular orientation in uniaxially stretched polymers, *Polymer*, 36(22), 4339-4343 (1995).

Butler J.H., Joy D.C., Bradley G.F. and Krause S.J., Low-voltage scanning electron microscopy of polymers, *Polymer*, 36(9), 1781-1790 (1995).

Cantor K., *Blown Film Extrusion: An Introduction*, Hanser Publishers: USA (2006).

Casado U., Marcovich N.E., Aranguren M.I. and Mosiewicki M.A., High-strength composites based on tung oil polyurethane and wood flour: Effect of the filler concentration on the mechanical properties, *Polym. Eng. Sci.*, 49(4), 713-721 (2009).

Ceresana, *New Study: Strong Growth in the Polypropylene Market*, Ceresana Research: Germany (2008).

Chand N. and Fahim M., *Tribology of Natural Fiber Composites*, C.R.C. Press: USA (2008).

Chauhan S., Karmarkar A. and Aggarwal P., Damping behavior of wood filled polypropylene composites, *J. Appl. Polym. Sci.*, *in press* (2009).

Cherukupalli S.S. and Ogale A.A., Online measurements of crystallinity using Raman spectroscopy during blown film extrusion of a linear low-density polyethylene, *Polym. Eng. Sci.*, 44(8), 1484-1490 (2004).

Coates P.D., Barnes S.E., Sibley M.G., Brown E.C., Edwards H.G.M. and Scowen I.J., In-process vibrational spectroscopy and ultrasound measurements in polymer melt extrusion, *Polymer*, 44, 5937-5949 (2003).

Coates P.D., Seeing processes - process imaging metrology, *PPS-21*, Leigzig (2005).

Crosby P.B., *Quality Is Free*, Penguin Books: Canada (1980).

Dai X., Wang W., Ding Y. and Sun Z., "Assumed inherent sensor" inversion based ANN dynamic soft-sensing method and its application in erythromycin fermentation process, *Computers and Chemical Engineering*, 30, 1203-1225 (2006).

Davis B.A., Gramann P.J. and Osswald T.A. *Compression Molding*, Hanser Gardner Publications: USA (2003).

Díaz M.F., Barbosa S.E. and Capiati N.J., Reactive compatibilization of PE/PS blends. Effect of copolymer chain length on interfacial adhesion and mechanical behavior, *Polymer*, 48(4), 1058-1065 (2007).

Ding Y.L., Forster R., Seville J.P.K. and Parker D.J., Segregation of granular flow in the transverse plane of a rolling mode rotating drum, *Int. J. Multiphase Flow*, 28, 635-663 (2002).

Dobreski D.V. and Nattinger B.E., Plastic bags of blends of linear ethylene polymers and aromatic polymers, *US Patent*, 4804564 (1989).

Dumitrescu O.R., Baker D.C., Foster G.M. and Evans K.E., Near infrared spectroscopy for in-line monitoring during injection moulding, *Polym. Test.*, 24, 367-375 (2005).

Edwards C.D., The meaning of quality, *Qual. Prog.*, 1, 36-39 (1968).

El-Mabrouk K., Belaiche M. and Bousmina M., Phase separation in PS/PVME thin and thick films, *J. Colloid Interf. Sci.*, 306, 354-367 (2007).

El-Sabbagh A., Steuernagel L. and Ziegmann G., Processing and modeling of the mechanical behavior of natural fiber thermoplastic composite: flax/polypropylene, *Polym. Composite*, 30(4), 510-519 (2009).

Environment and Plastics Industry Council (2003).
http://www.cpia.ca/epic/pdfs/special-reports/plastics%20lumber%20special%20report.PDF.

Feigenbaum A.V., *Total Quality Control*, McGraw-Hill: USA (1991).

Feldhoff R., Wienke D., Cammann K. and Fuchs H., On-line post consumer package identification by NIR spectroscopy combined with a fuzzy ARTMAP classifier in an industrial environment, *Appl. Spectrosc.*, 51(3), 362-368 (1997).

Fischer D., Sahre K., Abdelrhim M., Voit B., Sadhu V.B., Pionteck J., Komber H. and Hutschenreuter J., Process monitoring of polymers by in-line ATR-IR, NIR and Raman spectroscopy and ultrasonic measurements, *C.R. Chimie*, 9, 1419-1424 (2006).

Furness R.A., *Fluid Flow Measurement,* Longman Group: UK (1989).

Gao Z., Su W. and Wu D., Polyurethane-solid wood composites. II. Flammability parameters, *J. Appl. Polym. Sci.*, 113(5), 3279-3285 (2009a).

Gao Z., Wu D., Su W. and Ding X., Polyurethane-solid wood composites prepared with various catalysts. I. Mechanical properties and dimensional stabilities, *J. Appl. Polym. Sci.*, 111(3), 1293-1299 (2009b).

Ghasemi I., Azizi H. and Naeimian N., Investigation of the dynamic mechanical behavior of polypropylene/(wood flour)/(kenaf fiber) hybrid composites, *J. Vinyl Addit. Techn.*, 15(2), 113-119 (2009).

Giles H.F. Jr., Wagner J.R. Jr. and Mount E.M. III, *Extrusion: the Definitive Processing Guide and Handbook*, William Andrew Publishing: USA (2005).

Gillespie L.R., *Troubleshooting Manufacturing Processes: A Reference Book for Manufacturing Engineers, Managers, and Technicians*, Society of Manufacturing: USA (1988).

Gilmore H.L., Product conformance cost, *Qual. Prog.*, 7, 16-19 (1974).

Gilmore R.S. and Tjahjidi M., Ultrasonic probe and method for monitoring materials processing in screw driven extruders, *US Patent*, 6415665 (2002).

Global Markets Direct, *Global Polyethylene Market Analysis and Forecasts to 2020*, Aarkstore Enterprise: India (2009).

Goodman S.H., *Handbook of Thermoset Plastics*, William Andrew Publising: USA (2001).

Gottwald A. and Scheler U., Extrusion monitoring of polymer melts using a high-temperature surface-NNMR probe, *Macromol. Mater. Eng.*, 290(5), 438-442 (2005).

Gramann P. and Rauwendaal C., New dispersive and distributive mixers for extrusion and injection molding, *ANTEC*, Atlanta (1998).

Gregorova A., Hrabalova M., Wimmer R., Saake B. and Altaner C., Poly(lactide acid) composites reinforced with fibers obtained from different tissue types of Picea sitchensis, *J. Appl. Polym. Sci.*, *in press* (2009).

Haque M., Hasan M., Islam S. and Ali E., Physico-mechanical properties of chemically treated palm and coir fiber reinforced polypropylene composites, *Bioresource Technol.*, 100, 4903-4906 (2009).

Harper C.A., *Handbook of plastics, elastomers, and composites*, McGraw-Hill: USA (2002).

He Y., Zhu B. and Inoue Y., Hydrogen bonds in polymer blends, *Prog. Polym. Sci.*, 29(10), 1021-1051 (2004).

Hermes H.E., Bucknall D.G., Higgins J.S. and Scherrenberg R.L., The ordering of semi-crystalline PS-b-hPB copolymers at a PS/PE interface and their effects on interfacial strength, *Polymer*, 39(14), 3099-3108 (1998).

Heymann K., Mirschel G., Scherzer T. and Buchmeiser M.R., In-line determination of the thickness of UV-cured coatings on polymer films by NIR spectroscopy, *Vib. Spectrosc.*, *in press* (2009).

Huang H.X. and Zhang J.J., Effects of filler-filler and polymer-filler interactions on rheological and mechanical properties of HDPE-wood composites, *J. Appl. Polym. Sci.*, 111(6), 2806-2812 (2009).

Huth-Fehre T., Feldhoff R., Kantimm T., Quick L., Winter F., Cammann K., van den Broek W., Winke D., Melssen W. and Buydens L., NIR – remote sensing and artifical neural networks for rapid identification of post consumer plastics, *J. Mol. Struct.*, 348, 143-146 (1995).

Ingram A., Seville J.P.K., Parker D.J., Fan X. and Forster R.G., Axial and radial dispersion in rolling mode rotating drums, *Powder Technol.*, 158, 76-91 (2005).

Intertek Plastics Technology Laboratories (2009).
www.ptli.com.

Jodin P., *Le Bois, Matériau d'Ingénierie*, Association pour la recherche sur le bois en Lorraine: France (1994).

Juran J.M. and Godfrey A.B., *Juran's Quality Handbook*, McGraw-Hill: USA (1999).

Kantz M.R., Newman J.R. and Stigale F.H., The skin-core morphology and structure-property relationships in injection-moulded polypropylene, *J. Appl. Polym. Sci.*, 16, 1249-1260 (1972).

Kasliwal S.G. and Jones J.W., Gentle compounding of natural fibre filled PVC composites, *ANTEC*, Chicago, 052 (2004).

Klason C., Kubat J. and Strömvall H.E., The efficiency of cellulosic fillers in common thermoplastics. Part 1. Filling without processings aids or coupling agents. *Inter. J. Polymeric Mater.*, 10, 159-187 (1984).

Kremer F., Schonhals A. and Luck W., *Broadband Dielectric Spectroscopy*, Springer-Verlag: USA (2002).

Lachenal G., Dispersive and Fourier transform near-infrared spectroscopy of polymeric materials, *Vib. Spectrosc.*, 9, 93-100 (1995).

Landry C.J.T., Yang H. and Machell J.S., Miscibility and mechanical properties of a ternary polymer blend: polystyrene/polycarbonte/tetramethyl polycarbonate, *Polymer*, 32(1), 44-52 (1991).

Lee S.M., Park J.C., Lee S.M., Ahn Y.J. and Lee J.W., In-line measurement of residence time distribution in twin-screw extruder using non-destructive ultrasound, *Korea-Aust. Rheol. J.*, 17(2), 87-95 (2005).

Leffler K.B., Ambiguous changes in product quality, *Am. Econ. Rev.*, 72, 956-976 (1982).

Leitner R., Mairer H. and Kercek A., Real-time classification of polymers with NIR spectral imaging and blob analysis, *Real-Time Imaging*, 9, 245-251 (2003).

Leukel J., Weis C., Friedrich C. and Gronski W., Polymer communications: On-line morphology measurement during the extrusion of polymer blends, *Polymer*, 39(25), 6665-6667 (1998).

Liu H., Wu Q. and Zhang Q., Preparation and properties of banana fiber-reinforced composites based on High density polyethylene (HDPE)/Nylon-6 blends, *Bioresource Technol.*, *in press* (2009).

López-Barrón C.R., Robledo-Ortiz J.R., Rodrigue D. and González-Núñez R., Processing, morphology, and properties of polyamide-6/low density polyethylene blends prepared by film blowing, *J. Plast. Film Sheet.*, 23(2), 149-169 (2007).

Lorenzo A.T., Arnal M.L., Müller A.J., Boschetti de Fierro A. and Abetz V., Confinement effects on the crystallization and SSA thermal fractionation of the PE block within PE-b-PS diblock copolymer, *Eur. Polym. J.*, 42(3), 516-533 (2006).

Luzinov I., Pagnoulle C. and Jérôme R., Dependence of phase morphology and mechanical properties of PS/SBR/PE ternary blends on composition: transition from core-shell to triple-phase continuity structures, *Polymer*, 41(9), 3381-3389 (2000).

Marathe D.S. and Joshi P.S., Characterization of highly filled wood flour-PVC composites: morphological and thermal studies, *J. Appl. Polym. Sci.*, 114(1), 90-96 (2009).

Michaud F., *Rhéologie de Panneaux Composites Bois/Thermoplastiques sous Chargement Thermomécanique: Aptitudes au Postformage*, Ph.D. thesis, Université Laval: Canada (2003).

Minkova L. and Nikolova M., Melting of irradiated films prepared from polymer blends, *Polym. Deg. Stab.*, 25, 49-52 (1989).

Minkova L., Lefterova E. and Koleva T., Thermal stability of irradiated LDPE-i-PP films, *Polym. Deg. Stab.*, 37, 247-252 (1992).

Mosiewicki M.A., Casado U., Marcovich N.E. and Aranguren M.I., Polyurethanes from tung oil: Polymer characterization and composites, *Polym. Eng. Sci*., 49(4), 685-692 (2009).

Murphy J., *Additives for Plastics Handbook*, Elsevier Advanced Technology: UK (2003).

Nakagawa M., Altobelli S.A., Caprihan A. and Fukushima E., NMRI study: axial migration of radially segregated core of granular mixtures in a horizontal rotating cylinder, *Chem. Eng. Sci.*, 52(23), 4423-4428 (1997).

Nicole D., Lazo B. and Scott C.E., Morphology development during phase inversion of a PS/PE blend in isothermal, steady shear flow, *Polymer*, 40(20), 5469-5478 (1999).

Noël M., Fredon E., Mougel E., Masson D., Masson E. and Delmotte L., Lactic acid/wood-based composite material. Part 1: synthesis and characterization, *Bioresource Technol.*, 100, 4711-4716 (2009).

Nourbakhsh A. and Kouhpayehzadeh M., Mechanical properties and water absorption of fiber-reinforced polypropylene composites prepared by bagasse and beech fiber, *J. Appl. Polym. Sci.*, 114(1), 653-657 (2009).

Ohshima M. and Tanigaki M., Quality control of polymer production processes, *J. Process Contr.*, 10, 135-148 (2000).

Pantani R., Coccorullo I., Speranza V. and Titomanlio G., Modeling of morphology evolution in the injection molding, process of thermoplastic polymers, *Prog. Polym. Sci.*, 30, 1185-1222 (2005).

Pelsoci T., *ATP-funded green process technologies: improving U.S. industrial competitiveness with applications in packaging, metals recycling, energy, and water treatment*, National Institute of Standards and Technology, U.S. Department of Commerce: USA (2007).

Perdikoulias J. and Vlcek J., *Analysis of film extrusion problems with flow simulation*, Compuplast International: Czech Republic (2004).

Pilla S., Gong S., O'Neill E., Yang L. and Rowell R.M., Polylactide-recycled wood fiber composites, *J. Appl. Polym. Sci.*, 111(1), 37-47 (2009).

Polizzi S., Fagherazzi G., Benedetti A. and Battagliarin M., Crystallinity of polymers by X-ray diffraction: a new fitting approach, *Eur. Polym. J.*, 27(1), 85-87 (1991).

Rauwendaal C., *Polymer Extrusion*, Hanser Gardner Publications: USA (2001).

Reifenhauser F. and Rahlfs H., *Film blowing method*, US Patent, 3926706 (1975).

Rials T., Labbé N., Kelly S. and Gardner D., NIR assessment of wood-polyolefin extruded composites (2009).
www.forestprod.org/woodfiber05rials.pdf.

Rocha N., Coelho J.F.J., Fonseca A.C., Kazlauciunas A., Gil M.H., Gonçalves P.M. and Guthrie J.T., Poly(vinyl chloride) and wood flour press mould composites: New bonding strategies, *J. Appl. Polym. Sci.*, 113(4), 2727-2738 (2009).

Rohe T., Becker W., Kölle S., Eisenreich N. and Eyerer P., Near infrared (NIR) spectroscopy for in-line monitoring of polymer extrusion processes, *Talanta*, 50, 283-290 (1999).

Rowell R.M., Young R.A. and Rowell J.K., *Paper and composites from agro-based resources*, Lewis publishers: USA (1997).

Sadler R.L., Sharpe M., Panduranga R. and Shivakumar K., Water immersion effect on swelling and compression properties of Eco-Core, PVC foam and balsa wood, *Compos. Struct.*, 90, 330-336 (2009).

Saeki K., Tanabe K., Matsumoto T., Uesaka H., Amano T. and Funatsu K., Prediction of polyethylene density by near-infrared spectroscopy combined with neural network analysis, *J. Comput. Chem. Jap*., 2(1), 33-40 (2003).

Salvendy G., *Handbook of Industrial Engineering and Operation Management*, Wiley-IEEE (2001): USA (2001).

Sato H., Shimoyama M., Kamiya T., Amari T., Sasic S., Ninomiya T., Siesler H.W. and Ozaki Y., Near infrared spectra of pellets and thin films of high-density, low-density and linear low-density polyethylenes and prediction of their physical properties by multivariate data analysis, *J. Near Infrared Spec.*, 11, 309-321 (2003).

Sewda K. and Maiti S.N., Mechanical properties of teak wood flour-reinforced HDPE composites, *J. Appl. Polym. Sci.*, 112(3), 1826-1834 (2009).

Soury E., Behravesh A.H., Rouhani Esfahani E. and Zolfaghari A., Design, optimization and manufacturing of wood-plastic composite pellet, *Mater. Des*, 30, 4183-4191 (2009).

Stuart B.H., Polymer crystallinity studied using Raman spectroscopy, *Vib. Spectrosc.*, 10, 79-87 (1996).

Sun Z., Jen C.K., Shih C.K. and Denelsbeck D.A., Application of ultrasound in the determination of fundamental extrusion performance: residence time distribution measurement, *Polym. Eng. Sci.*, 43(1), 102-111 (2004).

Tademr M., Biltekin H. and Caneba G.T., Preparation and characterization of LDPE and PP - Wood fiber composites, *J. Appl. Polym. Sci.*, 112(5), 3095-3102 (2009).

Tadmor Z., Lipshitz S.D. and Lavie R., Dynamic model of a plasticating extruder, *Polym. Eng. Sci.*, 14, 112-119 (1974).

Tecrep Engineering (2009).
www.tecrep.com/rubber_compress.asp.

Townsend, *Global Market Briefs - Polyethylene & Polypropylene*, Townsend Polymers: USA (2008).

Tuchman B.W., The decline of quality, *New York Times*, 2, 38-39 (1980).

Vailaya A., Wang T., Chen Y. and Huffman M., Quantitative analysis of dimethyl titanocene by iodometric titration, gas chromatography and NMR, *J. Pharmaceut. Biomed.*, 25, 577-588 (2001).

van den Broek W.H.A.M, Wiemke D., Melssen W.J., de Crom C.W.A. and Buydens I. Identification of plastics among nonplastics in mixed waste by remote sensing near-infrared imaging spectroscopy. 1. Image improvement and analysis by singular value decomposition, *Anal. Chem.*, 67, 3753-3759 (1995).

van den Broek W.H.A.M., Derks E.P.P.A., van de Ven E.W., Wienke D., Geladi P. and Buydens L.M.C., Plastic Identification by Remote Sensing Spectroscopic NIR Imaging using Kernel Partial Least Squares (KPLS), *Chemometr. Intell. Lab.*, 35, 187-197 (1996).

van den Broek W.H.A.M., Wienke D., Melssen W.J. and Buydens L.M.C., Optimal wavelength range selection by a genetic algorithm for discrimination purposes in spectroscopic infrared imaging, *Appl. Spectros.*, 51(8), 1210-1217 (1997).

Wang Q., Jander P., Fricke-Begemann C. and Noll R., Comparison of 1064 nm and 266 nm excitation of laser-induced plasmas for several types of plastics and one explosive, *Spectrochim. Acta B*, 63 1011-1015 (2008).

Wang Y. and Hahn T.H., AFM characterization of the interfacial properties of carbon fiber reinforced polymer composites subjected to hygrothermal treatments, *Compos. Sci. Technol.*, 67, 92-101 (2007).

Watanabe S., Dybal J., Tashiro K. and Ozaki Y., A near-infrared study of thermally induced structural changes in polyethylene crystal, *Polymer*, 47, 2010-2017 (2006).

Wikipedia Plastics Extrusion (2009).
www.en.wikipedia.org/wiki/plastics_extrusion.

Wilkes C.E., Summers J.W., Daniels C.A. and Berard M.T., *PVC Handbook*, Hanser: Germany (2005).

WRAP, *Wood Plastic Composites Study Research Report – Technologies and UK Market Opportunities, Waste and Resources Action Programme*, Old Academy: United-Kingdom (2003).

Xanthos M. and Patel S.H., Recent developments in in-line spectroscopy/microscopy for monitoring extrusion processes, *Adv. Polym. Tech.*,14(2), 151-157 (1995).

Xiong C., Qi R. and Gong W., The preparation and properties of wood flour/high density polyethylene composites by in-situ reaction extrusion*, Polym. Adv. Technol*., 20(3), 273-279 (2009a).

Xiong C., Qi R. and Wang Y., Wood-thermoplastic composites from wood flour and high-density polyethylene, *J. Appl. Polym. Sci*., 114(2), 1160-1168 (2009b).

Yan W., Shao H. and Wang X., Soft sensing modeling based on support vector machine and Bayesian model selection, *Comput. Chem. Eng.*, 28, 1489-1498 (2004).

Younesi M. and Bahrololoom M.E., Producing toughened PP/HA-LLDPE ternary bio-composite using a two-step blending method, *Mater. Design*, 30, 4253-4259 (2009).

Yuan J. and Shi S.Q., Effect of the addition of wood flours on the properties of rigid polyurethane foam, *J. Appl. Polym. Sci.*, 113(5), 2902-2909 (2009).

Zhang J., Park C.B., Rizvi G.M., Huang H. and Guo Q., Investigation on the uniformity of high-density polyethylene/wood fiber composites in a twin-screw extruder, *J. Appl. Polym. Sci.*, 113(4), 2081-2089 (2009).

Zhong W. and Yu J., MIMO soft sensors for estimating product quality with on-line correction, *Trans. IChemE.*, 78(A), 612-620 (2000).

Zhu C., Umemoto X., Okui S. and Saki T., Craze yielding and fracture mechanisms in PE/PS/PE laminated films. Part 2, *J. Mater. Sci.*, 24(8), 2787-2793 (1989).

Chapter 3

Background on multivariate imaging and chemometrics

3.1 Spectral imaging

3.1.1 Defining spectral imaging

Webster's New World College Dictionary (2005) defines the *electromagnetic spectrum* as: "the complete range of frequencies of electromagnetic waves from the lowest to the highest, including, in order, radio, radar, infrared, visible light, ultraviolet, X-ray, and gamma ray waves". This definition of the spectrum is illustrated in Figure 3.1. While the full electromagnetic spectrum conceptually spans from the Planck length (1.616×10^{-35} m) to the size of the universe, in practice it usually ranges from 10^{-14} to 10^{3} m. The spectral signature of an object is the characteristic distribution of electromagnetic radiation either emitted, absorbed, reflected or transmitted by that object. In this section, we will define three types of imaging sensors (grayscale, RGB and spectral imaging) and relate them to a sample's spectral signature.

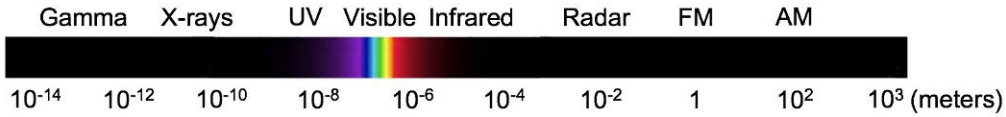

Figure 3.1. Electromagnetic spectrum ranging from gamma rays to radio ways. Adapted from Color Information Resource (2007).

A grayscale, or black-and-white, image is a 2-D matrix in which each pixel only represents intensity information. Such images are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest (Johnson, 2006). These images are in fact monochromatic as they do not contain any chromatic, or color, information. The information provided by the monochrome cameras is in fact the convolution of the spectrum of the incident light and the spectral sensitivity of the CCD. Since no specific wavelength filters are used in front of the CCD photosensitive detectors (as is the case with

RGB/Bayer patters), then the light is not decomposed into different colors, although the CCD itself is sensitive to color. Depending on the sensor, the perceived intensity may either be at a distinct narrow wavelength, or averaged over a wide spectral interval (Figure 3.2A).

Any image containing chromatic information does so by recording each color, or wavelength, separately. According to Bharati et al. (2003), a multivariate image consists of a stack of congruent images, with each image in the stack representing a unique wavelength. Such an image can be represented as a 3-D dataset, where two dimensions (x and y) represent spatial information in the image plane and the third dimension (λ) represents the spectral information. In the case of common color photography, the visible color spectrum is typically represented by multivariate images containing red, green and blue color intensities, forming a ($x \times y \times 3$) 3-D image array (Figure 3.2B). In practice, the blue sensor is especially sensitive to the light between 375 and 500 nm, the green between 500 and 575 nm, and the red between 575 and 850 nm (Nikon, 2009).

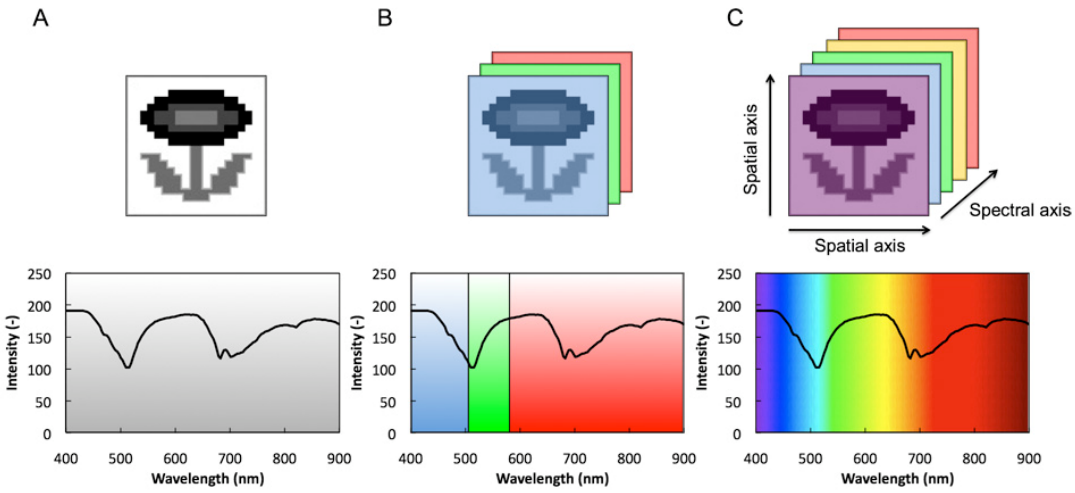

Figure 3.2. A single image as seen using (A) a monochrome (grayscale) camera, (B) a red/green/blue (RGB) color camera, and (C) a spectral imaging system.

Finally, a spectral image can be seen as an extension of an RGB image in that it captures a much greater number of individual wavelengths, usually with an increased spectral range and/or an increased spectral resolution (Figure 3.2C). As early as in the 1960s, multispectral imaging systems with up to 20 wavelength bands were used in remote sensing

applications (Liu and MacGregor, 2007). Later, the term hyperspectral imaging was coined in the fields of satellite and airborne imaging, referring to multispectral images containing more than 100 wavelengths (Grahn and Geladi, 2007). Shippert (2009), however, provides a more specific definition of the term. As the word implies, a hyperspectral image literally contains *an excessive* spectral resolution. These images are over-determined, meaning that they provide ample spectral information to identify and distinguish spectrally unique materials. It is not only the number of measured wavelengths that defines a sensor as hyperspectral but also the narrowness and contiguous nature (i.e. spectral resolution) of the measurements. According to this definition, a sensor capable of capturing over 100 wavelengths may not be hyperspectral if it is sensitive to a very wide spectral range.

3.1.2 Experimental setup

Data acquisition in this study was performed using a line-scan imaging system. There are, however, a number of spectral imaging systems available on the market, falling generally into one of three categories: point-scans, line-scans and area-scans.

Point-scans are conceptually the simplest systems since they acquire a single spectral signature per acquisition period. This system is essentially analogous to a single surface probe. This spectral signature can either represent a very localized area of the sample as was done by Coates et al. (2003) or an average value of a larger surface as in Feldhoff et al. (1997).

Line-scan systems acquire a single line of the sample per acquisition period. For each scanned line of the sample, the transmitted/reflected light enters the spectroscope through a slit, is then dispersed through a grating system, and finally captured by the charge-coupled device (CCD) array of the camera as a spatial/spectral image as illustrated in Figure 3.3. Thus, each image from the spectral camera includes the line pixels in one direction (spatial axis) and spectral distribution (light intensities in spectral elements) in another dimension (spectral axis). The entire sample can be imaged by slightly displacing the sample perpendicularly to the spectroscope's slit between each acquisition; a large image may

require the sample to be imaged hundreds of times. Line-scan systems have widely been used for their unique combination of fast acquisition speed and high spatial resolution (Leitner et al., 2003; Liu et al., 2007; Qiao et al., 2007; ElMasry et al., 2008).

The third category of spectral imaging sensors consists of area-scan systems. In this case, a 2-D area of the sample is captured at every acquisition time, much as the common camera would do. In order to achieve this, the imaging system captures the scene as a succession of grayscale images taken at different wavelengths. Tunable filters are typically used to select the various wavelengths. Doing so greatly increases the capture time. While such a system is suitable for off-line applications, it cannot be used in on-line industrial applications such as scanning samples fed to a process using a rapidly moving conveyor belt. The area scanners require the sample to be static during the complete capture time, which mostly depends on the number of wavelengths captured. Nevertheless, these systems were used successfully by van den Broek et al. (1996) to identify plastic waste products.

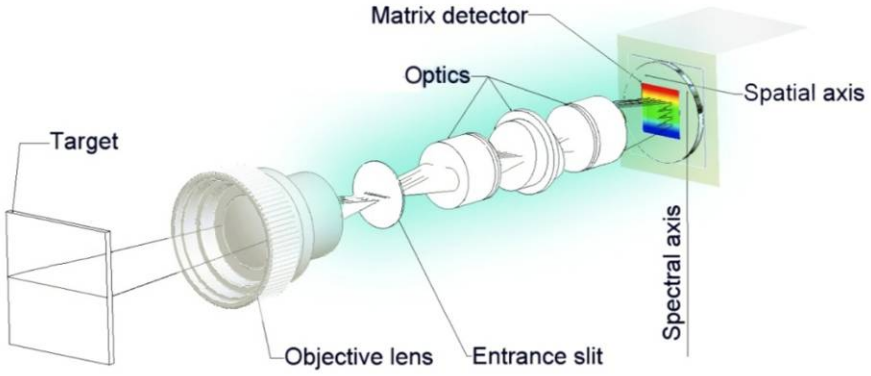

Figure 3.3. Illustration of a line-scan imaging system projecting a single line of the sample onto a 2-D sensor. Adapted from Specim (2009).

A line-scan imaging system (Figure 3.4) was used in this work. It consists of three spectral cameras coregistered on the same area on the sample, effectively covering the ultraviolet (UV), visible (VIS) and near-infrared (NIR) range of the spectrum (200-1700 nm). First, the UV range is captured using a Hamamatsu C8800 camera (1000×1000 pixels) coupled with an ImSpector UV4 grating spectroscope (Specim Spectral Imaging Ltd) effectively sensing the 200-400 nm wavelength range. Second, the visible range is captured using a

Hamamatsu C8484-05G camera (960×1024 pixels) coupled with an ImSpector V10E grating spectroscope (Specim Spectral Imaging Ltd) effectively sensing the 400-1000 nm wavelength range. The infrared range is captured using a XenICs XEVA-USB-FPA camera (256×320 pixels) coupled to a Specim ImSpector N17E grating spectroscope effectively sensing between 900-1700 nm.

The system is equipped with both halogen or xenon lighting, the former for VIS-NIR applications and the latter for UV. These light sources were selected to provide relatively smooth lighting conditions over the entire spectral interval. Basically, light is reflected off the sample (tests were performed in reflection and transflection), through a system of mirrors, divided by 2 dichroic lens into UV, visible and NIR light and directed toward their respective cameras.

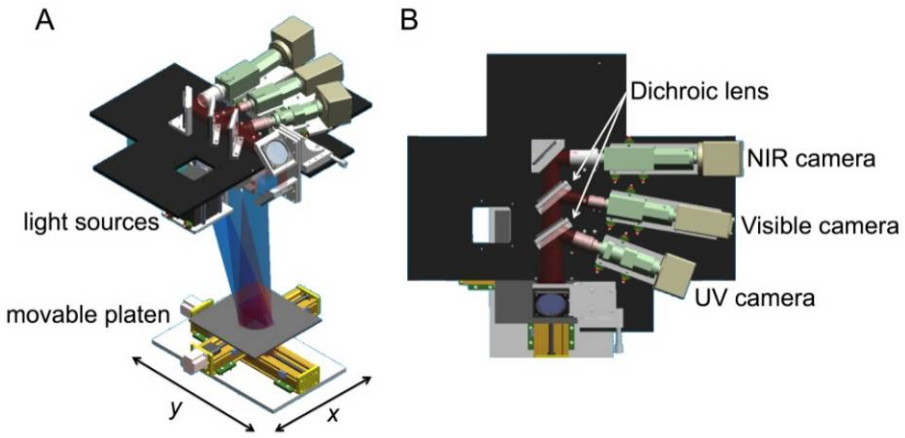

Figure 3.4. Hyperspectral scanning system seen in perspective (A) and from above (B).

Each scanned line of the sample is captured on the CCD array as a spatial/spectral ($x \times \lambda$) image containing hundreds of spectral channels (i.e. wavelengths), rather than the three more commonly obtained in RGB images. A hyperspectral image, or data cube, of the entire sample is achieved by repeatedly displacing the sample perpendicularly to the spectroscope slit. Each hyperspectral image consists of a number of juxtaposed line-scans forming a 3-D array ($x \times y \times \lambda$).

A simple calibration scheme is commonly applied to the raw line-scans in order to remove any linear discrepancies (Leitner et al., 2003; ElMasry et al., 2008). The raw signals of each spectral-spatial image (**R**) were transformed into reflectance units using a flat black (**B**) image and a true white image (**W**):

$$i_{x\lambda} = \frac{r_{x\lambda} - b_{x\lambda}}{w_{x\lambda} - b_{x\lambda}} \tag{3.1}$$

where, $r_{x\lambda}$ is an element of the raw image (**R**), $b_{x\lambda}$ and $w_{x\lambda}$ are elements of the black (**B**) and white (**W**) calibration images, and $i_{x\lambda}$ is an element the standardized reflectance image (**I**). The flat black and true white images were obtained by collecting a spatial-spectral image leaving the lens cap in place (i.e. dark current), and by imaging a 99% reflectance standard sheet (Gigahertz-Optik), respectively. This calibration is performed on each image of a set. The resulting calibration of a narrow sample on a true white background is shown in Figure 3.5.

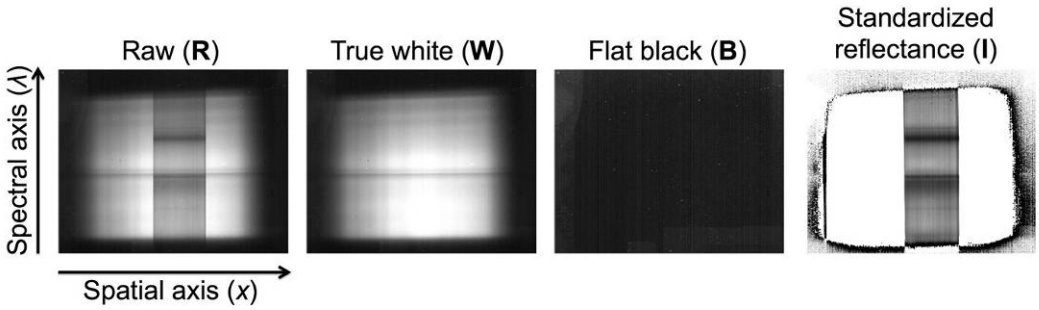

Figure 3.5. Black/white calibration in which a raw image is converted into standardized reflectance units by comparing it to true white and true black images. Figure shows a WPC sample placed on a 99% reflectance standard.

3.2 Chemometrics

3.2.1 Defining chemometrics

According to Zhong and Yu (2000), soft-sensors can be used to obtain a regression model between easily obtained measurements (secondary variables) and quality variables (primary

variables) using statistical (Santen et al., 1997), neural networks (Wang and Shao, 1997) or chemometrics approaches (Wise and Gallagher, 1996). The latter approach was favored in this work for its ability to consider highly correlated data in a straightforward and parsimonious manner. On the one hand, statistical methods may often lead to good results but lack the ability to deal with highly correlated data. On the other hand, neural network are known for the ability to accurately predict the outcomes of complex problems (Berson et al., 1999) but are criticized for lacking interpretability and being susceptible to over-training and over-parameterization (Groth, 2000).

The term "chemometrics" was coined by Svante Wold in 1974 (Wold, 1995) and the International Chemometrics Society (2009) was formed shortly thereafter by Svante Wold and Bruce Kowalski, two pioneers in the field. The International Chemometrics Society (ICS) offers the following definition of chemometrics: *"It is the science of relating measurements made on a chemical system or process to the state of the system via application of mathematical or statistical methods."* Essentially, it is the application of mathematical or statistical methods to chemical data. The ICS states that chemometrics research spans a wide area of different methods that can be applied in chemistry biochemistry and chemical engineering. There are techniques for collecting good data (optimization of experimental parameters, design of experiments, multivariate analysis or multivariate statistics, calibration, signal processing) and for getting information from the data (statistics, pattern recognition, modeling, structure-property-relationship estimations). Some of these techniques were used in this project for image analysis and data regression.

3.2.2 Pre-processing

Pre-processing is an important first step of the data analysis in order to remove non-relevant information (Kohler et al., 2009). Furthermore, it was used extensively throughout this work since both latent variable methods used in this thesis, namely principal component analysis (Section 3.2.3) and projection to latent structures (Section 3.2.4), are scale dependent. It is therefore no surprise that many pre-processing methods are available to the user: mean-centering, autoscaling, direct orthogonal signal correction, standard

normal variate and multiplicative signal correction just to name a few (Bakeev, 2005; Kohler et al., 2009). Nevertheless, Geladi and Grahn (1996) warn that the scaling of spectroscopic data is not always desirable and that trial and error is the only way of finding out which method, if any, works best. Among the pre-processing methods mentioned above, only mean-centering and autoscaling were used in this work and will be discussed in greater detail.

Mean-centering

This method involves the subtraction of the mean of each variable computed using all the samples in the dataset (Bakeev, 2005). Applying this to a data matrix $\mathbf{X}$ ($N \times K$), the mean-centering is achieved by the following equation:

$$\mathbf{X}_{MC} = \mathbf{X} - \mathbf{1}_N^{\mathrm{T}} \bar{\mathbf{x}} \tag{3.2}$$

where $\mathbf{X}_{MC}$ ($N \times K$) is the mean-centered data, $\bar{\mathbf{x}}$ ($1 \times K$) contains the mean response values of each of the variables and $\mathbf{1}_N$ ($1 \times N$) is a unit vector. Mean-centering essentially removes the absolute intensity information (i.e. base-line value) of each variable, focusing on the response variation (Figure 3.6).

Autoscaling

Scaling to unit variance is typically applied after mean-centering by dividing the resulting intensities by the variable-wise standard deviations (Bakeev, 2005):

$$\mathbf{X}_{AS} = \mathbf{X}_{MC} \left(diag\left(\mathbf{S}_M \right) \right)^{-1} \tag{3.3}$$

where $\mathbf{X}_{AS}$ ($N \times K$) is the autoscaled data and $\mathbf{S}_M$ is a square diagonal matrix containing the standard deviations for each of the variables. This method reduces each variable to a mean of zero and a standard deviation of one, effectively removing all scale effects (Figure 3.6).

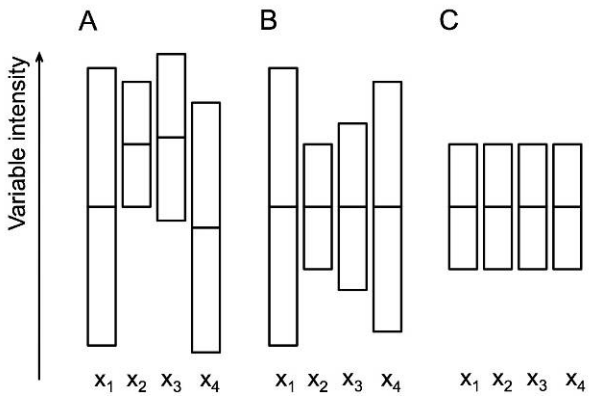

Figure 3.6. Illustration of data pre-processing in which the raw data (A) is either mean-centered (B) or autoscaled (C).

3.2.3 Principal component analysis

Introduction and interpretation

Consider a dataset **X** (*N*×*K*) consisting *N* observations on *K* variables. The observations (i.e. rows) of **X** may represent samples, trials or compounds whereas the variables, or columns of **X**, represent characteristics such as laboratory assays or sensor inputs (e.g. temperature, pressure, spectra). With many such datasets encompassing hundreds of variables quantifying thousands of observations, one may feel the need to extract the important information from **X** in order to display it simply.

Principal component analysis (PCA), the simplest of the eigenvector-based multivariate latent variable method, was first proposed by Karl Pearson (Pearson, 1901). This multivariate statistical method decomposes the variance-covariance structure of a data matrix by finding linear combinations of the original variables in such a way that these combinations explain the greatest amount of variation within the data (Eriksson et al., 2001). PCA is the decomposition of a matrix **X** (*N*×*K*) into simpler matrices $\mathbf{M}_a$ (*N*×*K*) of rank = 1:

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{M}_a \qquad (3.4)$$

In terms of the eigenvector/eigenvalue decomposition, rank is the number of eigenvalues of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ that are greater than zero (Geladi and Grahn, 1996). The rank can also be expressed as the smallest number of outer products of vectors (A) that may be used to represent X exactly. Since each of these $\mathbf{M}_a$ matrices has a rank = 1, they can be represented as the outer product of a single pair of vectors, $\mathbf{t}_a$ and $\mathbf{p}_a$:

$$\mathbf{M}_a = \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} \tag{3.5}$$

The decomposition of the X-matrix can therefore be written as:

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{3.6}$$

where A is the number of orthogonal principal components, $\mathbf{t}_a$ are the score vectors, and $\mathbf{p}_a$ are the loading vectors. The first score vector $\mathbf{t}_1$ is the linear combination of the K variables (defined by the loadings vectors $\mathbf{p}_1$) capturing the largest possible variance within matrix **X**, while the second score vector $\mathbf{t}_2$ represents the second largest source of variance (orthogonal to $\mathbf{t}_1$) and so on. The score vectors can therefore be viewed as a multivariate summary of each **X** observation. If the number of components is lower than the maximum number of components ($A<K$), then **E** ($N \times K$) contains the residuals.

The PCA decomposition is illustrated graphically in Figure 3.7. First, Figure 3.7A represents the relation between the original matrix **X** and the **T** and **P** matrices. This method of representing the matrices is used to illustrate the relation between the data matrix (**X**) and the outer product of the scores and loadings ($\mathbf{TP}^{\mathrm{T}}$). Figure 3.7B illustrates a PCA projection of a 3-D data space (as in the case of an RGB image) down to a 2-D plane. Let us discuss this projection of the data matrix in terms of loadings, scores and residuals (Equation 3.6). The loadings matrix **P** represents the orientation of the latent variable subspace within the original variable space. In essence, this subspace is built by linking each loading vector ($\mathbf{p}_i$) into a single hyperplane. Using this new subspace, the score matrix

T represents the new coordinates of the data points defined by the loadings (Figure 3.7C). In projecting the data onto a 2-D plane as in the case of Figure 3.7, each data point will have a $\mathbf{t_1}$ and $\mathbf{t_2}$ value corresponding to its location on the $\mathbf{p_1}$ and $\mathbf{p_2}$ vectors. In the case where the number of latent variables (A) equals the rank of **X**, the data points are exactly represented by the new system of coordinates. However, if A is chosen inferior to the rank of **X**, the PCA effectively projects data point located in a high-dimensional generic coordinate system onto a low-dimensional latent variable subspace. While doing so captures most of the information in the dataset, a certain amount of information will necessarily be lost. This loss is the orthogonal projection error of the data point and is presented in Equation 3.6 as the residual (**E**). Graphically, it represents the distance of a data point off the plane created by the loadings. Further details on PCA can be found in Wold et al. (1987).

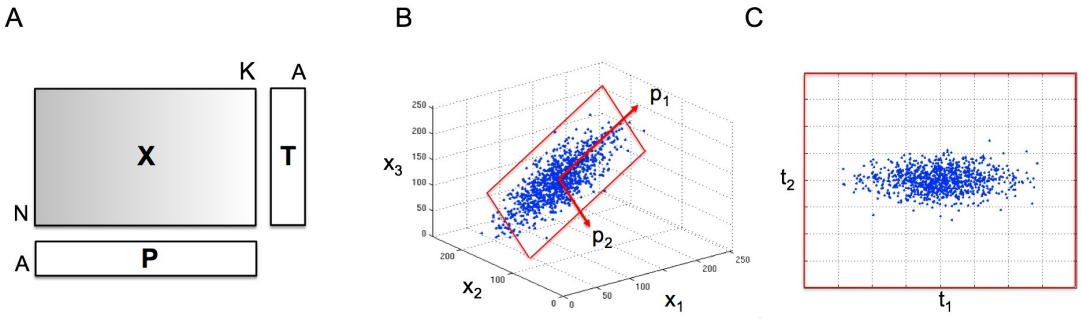

*Figure 3.7. Principal component analysis. In (A) the observations of **X** are represented using a small number of latent variables called scores (**T**) projected onto loadings (**P**). In (B) a 3-variable dataset ($\mathbf{x_1}$, $\mathbf{x_2}$, $\mathbf{x_3}$) projected onto 2 loading ($\mathbf{p_1}$, $\mathbf{p_2}$). In (C) $\mathbf{t_1}$-$\mathbf{t_2}$ score histogram created by $\mathbf{p_1}$ and $\mathbf{p_2}$ onto which each data point is projected.*

According to Geladi and Kowalski (1986), this decomposition is usually achieved by non-linear iterative partial least squares (NIPALS) or by singular value decomposition (SVD). While the NIPALS algorithm is an iterative process that calculates each or the principal component sequentially, the SVD decomposition extracts all the principal components of **X** ($N \times K$) simultaneously:

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathbf{T}} = \mathbf{T}\mathbf{P}^{\mathbf{T}} \qquad (3.7)$$

where the **U**, **D** and **V** matrices have certain specific properties:

- The columns of **U** ($N \times A$) are eigenvectors of $\mathbf{XX^T}$ and are orthonormal.
- The diagonal elements of **D** ($A \times A$) are singular values and the square roots of the eigenvalues of $\mathbf{X^TX}$.
- The columns of **V** ($K \times A$) are eigenvectors of $\mathbf{X^TX}$ and are also orthonormal. **U** and **V** are identical in the case where **X** is a correlation matrix.

In the case where the number of components used in the decomposition (A) is inferior to the rank of **X**, then the score and loading matrices (**T** and **P**) no longer represent the **X** exactly, leading to the presence of an error term in Equation 3.6.

Identifying outliers

The Merriam-Webster Online Dictionary (2009) defines the term *outlier* as: "a statistical observation that is markedly different in value from the others of the sample". While outstanding data points are sometimes relevant and can represent key aspects of the system under study, they may also be signs of flawed data possibly caused by poor acquisition tools or human error. In either case, the presence of such outlying data points can seriously affect the performance of the PCA algorithm (Gabrys et al., 2006) and should be studied with the utmost care. Keeping in mind the projections illustrated in Figure 3.7, two criteria are typically used to identify outliers.

First, the distance to the model in the X-space (*DmodX*) represents the orthogonal distance of an observation from the model plane (Figure 3.8A). As each of the PCA loadings ($\mathbf{p_i}$) seeks to capture the largest possible sources of variation found within **X**, any point falling far from the model plane differs from the bulk of the dataset. *DmodX* is computed by:

$$DmodX = \sqrt{\frac{SPE_n}{DF}} \tag{3.8}$$

where the squared predicted errors (*SPE*) are adjusted for the number of degrees of freedom (*DF*). In doing so, *DmodX* represents the number of standard deviations off the plane. In turn, *SPE* and *DF* are given as:

$$SPE_n = \sum_{k=1}^{K} e_{nk}^2 = e_{n1}^2 + e_{n2}^2 + \ldots + e_{nK}^2 \tag{3.9}$$

$$DF = (N - A - 1)(K - A) \tag{3.10}$$

where SPE_n sums the residual errors over all K variables for each observation n in the PCA decomposition (Equation 3.6) whereas the degrees of freedom are computed using the number of latent variables (A) as well as the dimensions of $\mathbf{X}$ ($N \times K$). In essence, the *DmodX* sums each row of the squared residual $\mathbf{E}$ matrix.

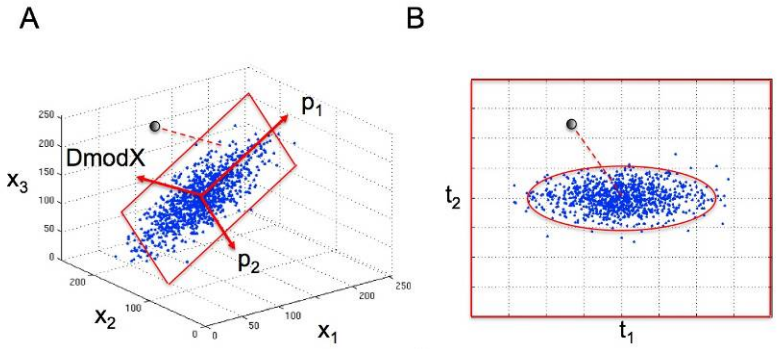

Figure 3.8. Identifying outlying data points in a PCA decomposition according to: (A) perpendicular distance from the plane (DmodX), and (B) distance from the center on the plane (Hotelling's T^2).

The second metric used to identify outliers is the Hotelling's T^2, which is orthogonal to an observation's distance from the model plane. It quantifies how far its projection falls from the center of the $\mathbf{t_1}$-$\mathbf{t_2}$ score plot. While an observation may appear to follow the general pattern of the PCA model (low *DmodX* value), it may nevertheless differ significantly from the bulk of the data by falling very far from the center of mass of the dataset (Figure 3.8B). It can be computed according to:

$$T_n^2 = \sum_{a=1}^{A} \frac{t_{na}^2}{s_a^2} = \frac{t_{n1}^2}{s_1^2} + \frac{t_{n2}^2}{s_2^2} + \ldots + \frac{t_{nA}^2}{s_A^2} \quad (3.11)$$

where A is the number of latent variables used in the model, t_{na}^2 is the score value of any given observation n for the a^{th} latent variable and s_a^2 is the variance of t_a in the model. The weighting by s_a^2 allows all scores to contribute equally to T^2 (MacGregor et al., 2008). As the T^2 values are approximately distributed according to an F-distribution, a 95% confidence interval is often drawn in the score space (Figure 3.8B) forming an ellipse in 2-D space and, in general terms, a hyperellipse in A-dimensional space.

It is important to note that only observations found far outside the confidence intervals of either *DmodX* of Hotelling's T^2 are expected to be outliers.

Quantifying prediction errors

As mentioned, the SPE sums each row of the squared residual matrix **E** ($N \times K$). By summing **E** column-wise, it is possible to compute the prediction errors for each of the K variables ($R_{X,k}^2$). According to MacGregor et al. (2008), it is computed for a single variable using:

$$R_{X,k}^2 = \frac{SS(\mathbf{x}_k) - SS(\mathbf{e}_k)}{SS(\mathbf{x}_k)} \quad (3.12)$$

where $SS(\mathbf{x}_k)$ represents the sum of squared values of **X** for a single column k whereas $SS(\mathbf{e}_k)$ is its counterpart using the **E** matrix. For a single variable k, this ratio is the sum of squares explained by the model divided by the total sum of squares. This notion can also be generalized into a global value in which the prediction errors of all K variables are combined (R_X^2):

$$R_X^2 = \frac{SS(\mathbf{X}) - SS(\mathbf{E})}{SS(\mathbf{X})} \quad (3.13)$$

where $SS(\mathbf{X})$ represents the sum of squared values of the entire **X** matrix whereas $SS(\mathbf{E})$ is its counterpart using the **E** matrix.

The **E** matrix varies with the complexity of the model. When building the model, the prediction error generally decreases in the training set as more components are used. However, when validating the model with new data, the prediction error may increase if too many components are used (Figure 3.9). This is explained by the fact that a very complex model will not only capture the underlying phenomena, but also the small-scale variations present in the training set data (e.g. experimental errors).

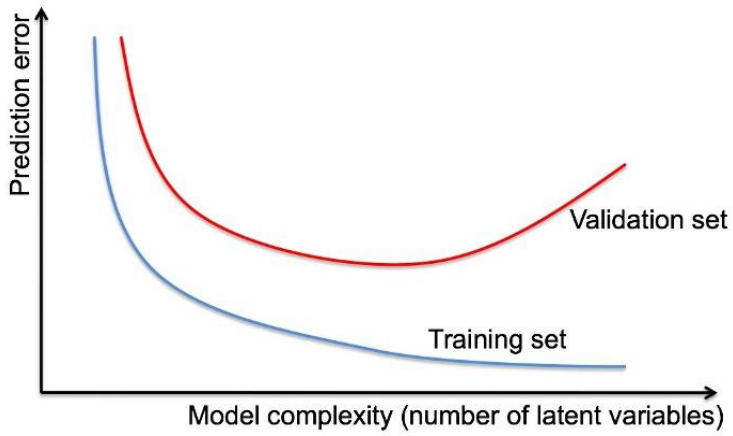

Figure 3.9. Prediction error varying with model complexity.

In order to avoid overfitting, a "goodness of fit" criterion is used to estimate the ability of the model to make predictions, using either internal existing data (i.e. through cross-validation), or an external validation set. It will be referred to as the predicted residual sum of squares (PRESS) and is computed using:

$$PRESS = \sum_{k=1}^{K} e_{nk}^2 \tag{3.14}$$

The resemblance between Equations 3.9 and 3.14 is not coincidental. The expressions for *SPE* and *PRESS* are exactly the same; however the errors used in each come from different

datasets. The former is obtained on the training set whereas the latter is calculated using the internal (i.e. cross-validation) or external validation sets. In practice, *PRESS(A)* is computed with each added latent variable in the model. Overfitting is avoided simply by determining whether *PRESS(A)* is significantly inferior to *PRESS(A-1)*. Finally, the ability of the model to predicted new, or omitted, data is given by:

$$Q^2_{X,cum} = \frac{SS(\mathbf{X}) - PRESS(A)}{SS(\mathbf{X})} \tag{3.15}$$

3.2.4 Regression methods

Phenomenological models, or first principal models, are often not available due to the complexity of industrial processes (Yan et al., 2004). As a result, empirical models are usually preferred when developing soft sensors. The problem of empirical modeling is to find a model with the best generalization and prediction performance, given the empirical data.

Multi-linear regression vs. principal component regression

In multiple linear regression (MLR), the **X** and **Y** matrices are linked directly using a linear relation:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \tag{3.16}$$

where **B** represents the matrix regression coefficients (a vector when a single **y** is used) and **E** consists of the regression residuals, often assumed to be identically and independently normally distributed. Furthermore, this model assumes that **X** is of full rank and is deterministic, that is to say that it is known perfectly without any uncertainties. In this way, all errors are assigned to **Y**. A number of procedures have been developed for parameter estimation in linear regression. The simplest and most commonly used is ordinary least

squares (OLS). The regression parameters can be estimated ($\hat{\mathbf{B}}$) by minimizing the sum of squared residuals (**E**):

$$\mathbf{E} = \mathbf{Y} - \mathbf{XB} \tag{3.17}$$

$$\hat{\mathbf{B}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \tag{3.18}$$

However, when **X** is ill-conditioned, the inversion of the $\mathbf{X}^T\mathbf{X}$ becomes very sensitive and eventually does not exist when **X** is singular (i.e. when **X** is not full rank). Unfortunately, this situation is very typical in industrial process datasets, as well as in spectral data analysis (high spectral resolution implies that many wavelengths are correlated). The properties of PCA can be applied to MLR in order to avoid such collinearity problems. In principal component regression (PCR), the **X** matrix is substituted by the **T** matrix computed by the PCA (only the dimensions corresponding to the largest eigenvalues are kept). By doing so, the multi-linear regression can be expressed as:

$$\mathbf{Y} = \mathbf{TB} + \mathbf{E} \tag{3.19}$$

$$\hat{\mathbf{B}} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{Y} \tag{3.20}$$

As in MLR, **B** and **E** represent the regression parameters and residuals, respectively. The matrix $\mathbf{T}^T\mathbf{T}$ is diagonal since the columns of **T** are orthogonal, so the elements of the inverse are merely the reciprocals of the diagonal elements (Jackson, 1991). In doing so, PCR avoids the problems linked to collinearity while reducing the noise present in the independent variable matrix.

However, in the PCR approach, the decomposition of **X** (i.e. projection onto a lower latent variables subspace) is performed without using information from **Y**, and therefore, this lower dimensional space is not necessarily the most predictive of **Y**. It is also important to note that only selecting the components with the highest variance (large eigenvalues) may also be problematic; Jolliffe (1982) states that low-variance principal components may be more important to the regression than higher variance ones. That is, the dominant variance directions in **X** are not necessarily the most important for predicting **Y**.

Projection to latent structures

The third and last regression method discussed here is the Projection to Latent Structures (PLS), also known as the Partial Least Squares. This method was used throughout this work to build the empirical latent variable models linking the regressor data (e.g. image data) to the quality data (e.g. mechanical properties). PLS was selected in this work since the columns of **X** were highly collinear.

According to Geladi and Kowalski (1986), the PLS analysis consists of simultaneous "internal" and "external" relations in which the **X** and **Y** data blocks are first taken individually (external relation) and then linked (internal relation). The "external" relation is essentially similar to a classic PCA on each of the **X** ($N \times K$) and **Y** ($N \times M$) matrices taken separately.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} + \mathbf{E} \tag{3.21}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^{\mathbf{T}} + \mathbf{F}' \tag{3.22}$$

where the **P** ($K \times A$) and **Q** ($M \times$A) matrices contain the loading vectors that best represent the **X** and **Y** spaces, whereas **T** ($N \times A$) and **U** ($M \times A$) each represent the latent variable space containing a number A of PLS components (i.e. latent variables). The number of PLS components or latent variables A ($A<min(K,N-1)$) is often selected using a cross-validation procedure. The **E** ($N \times K$) and **F'** ($N \times M$) matrices contain the residuals. Using these relations, the "internal" relation essentially links Equations 3.21 and 3.22 by ensuring that **T** and **U** are maximally correlated, thus minimizing the error **H** ($N \times A$) (MacGregor et al., 2008).

$$\mathbf{U} = \mathbf{T} + \mathbf{H} \tag{3.23}$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}^{\mathbf{T}} + \mathbf{F} \tag{3.24}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^{*} \tag{3.25}$$

$$\mathbf{W}^* = \mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1} \tag{3.26}$$

where **W*** ($K \times A$) contains the X-weights defining the common latent variable space **T** ($N \times A$) relating **X** and **Y**. Using this structure, the Y-block residuals are now represented by **F** ($N \times M$). The **W** matrix is the regression coefficients for each latent variable in terms of the "deflated", or dimensionless, X-matrix in the same way **W*** expresses **T** in terms of the original X-matrix (Equation 3.25). Hence, PLS attempts to extract the relationship between regressor variables (**X**) and response variables (**Y**) by assuming that they are generated by a common set of underlying factors **T** (MacGregor et al., 2008). The model structure is graphically illustrated in Figure 3.10.

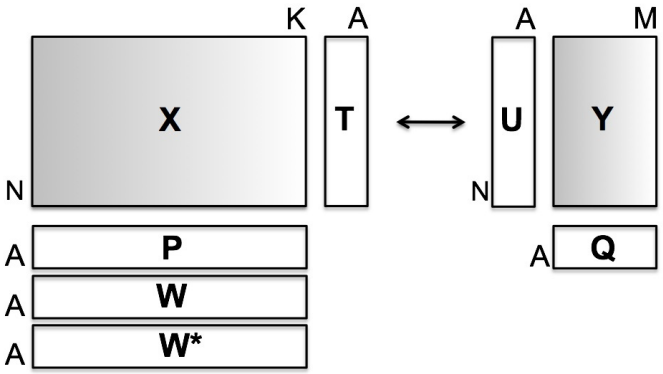

Figure 3.10. Schematic representation of the PLS model.

In PLS as in PCA, the X-data is represented by the outer product $\mathbf{TP}^T$ in which the loadings matrix **P** represents the orientation of the latent variable subspace within the original X-variable space. Using this new subspace, the score matrix **T** represents the new coordinates of the X-data defined by the loadings (**P**). The novelty here comes from the Y-data, modelled as the outer product $\mathbf{TQ}^T$ (Equation 3.24). In fact, the loadings matrix **Q** can be seen as the **P** counterpart, representing the orientation of the latent variable subspace within the original Y-variable space.

An important distinction must be made at this point between **P** and **W***. According to MacGregor et al. (2008), the X-loadings **P** are the coefficients of least squares regression of

T onto **X** but are seldom used in the interpretation of the results. Rather, **W*** is used to interpret how **X** is related to **T** in the same way the loadings are used in the PCA analysis.

By combining Equations 3.24 and 3.25, the PLS latent variable model can be re-organized as a simple prediction equation involving a set of pseudo-regression coefficients $\mathbf{B_{PLS}}$ ($K \times M$), similarly as for multiple linear regression (Wold et al., 2004).

$$\hat{\mathbf{Y}} = \mathbf{TQ}^{\mathbf{T}} = (\mathbf{XW}^*)\mathbf{Q}^{\mathbf{T}} = \mathbf{X}(\mathbf{W}^* \mathbf{Q}^{\mathbf{T}}) = \mathbf{XB_{PLS}} \tag{3.27}$$

The term $\mathbf{W^*Q^T}$ appears in the previous equation as a mean of linking the **X** and **Y** matrices. In fact, relationships between the variables in these matrices are visualized by overlaying the $\mathbf{w}_i^*$ and $\mathbf{q}_i$ plots for each latent variable, forming the $\mathbf{w}_i^*\mathbf{q}_i$ plot, also known as loadings biplot (MacGregor et al., 2008). Use of the biplot will be illustrated in Chapters 5 and 7.

Another commonly used metric for judging the importance of the X-variables on **Y** is the variable importance on the projection (i.e. VIP). The VIP score for the k^{th} variable for a single response **y** is computed according to Eriksson et al., (2001):

$$VIP_k = \sqrt{K \sum_{a=1}^{A} \left[\left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right) \left(w_{ak} / \|\mathbf{w}_k\|^2 \right) \right] \Big/ \sum_{a=1}^{A} \left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right)} \tag{3.28}$$

where a = 1,2…,A, K is the number of columns of **X** (i.e. variables), w_{ak} is the loading weight of the k^{th} variable in the a^{th} component, and $\mathbf{t}_a$, $\mathbf{w}_a$, and q_a are the a^{th} column vectors of **T**, **W**, and **Q**, respectively.

It is generally accepted in practice that variables having a VIP > 1.0 are highly influential, values between 0.8 < VIP < 1.0 indicate moderately influential variables, and VIP < 0.8 represent less important variables (Eriksson et al., 2001). The unity threshold comes from the fact that the average of the squared values of the VIPs is equal to 1.0. The greater-than-one rule is often used as a selection criterion (Eriksson et al., 2001; Olah et al., 2004;

ElMasry et al., 2008; MacGregor et al., 2008), and has a particular meaning when the data is autoscaled to unit variance.

3.2.5 Deterministic vs. stochastic image analysis

Machine vision can be defined as *"interpretation of an image of an object or scene through the use of optical noncontact sensing mechanisms for the purpose of obtaining information and/or controlling machines or processes"* (Marshall and Martin, 1992). According to Liu and MacGregor (2005), the most successful techniques and their applications in machine vision research have been confined to a specific type of environment where certain assumptions can be made about the scene. In typical manufacturing industries, an image provides a scene of objects with predetermined shape, structure and orientation; the primary goal of the inspection is to check whether there are missing objects in the image or whether objects in the image are in desired state (size, location, orientation). Image analysis from such industries will be referred to here as "deterministic". The American Heritage Dictionary (2009) defines the term *deterministic* as: "every state of affairs is the inevitable consequence of antecedent states of affairs". This definition underlines the predictable nature of the sample under scrutiny.

Images collected in the process industries belong to a different class of problems where the major concern is some ill-defined visual appearance of products or processes (Liu, 2004). Examples of this include the state of mineral flotation froth (Bartolacci et al., 2005), the aesthetics of manufactured countertops (Liu and MacGregor, 2006) or injection-molded plastic panels (Liu and MacGregor, 2005) and the grading of softwood lumber (Bharati et al., 2003). Simple assumptions about the scene, such as the size, location or orientation of the features, can no longer be made due to the seemingly random nature of these samples. Machine vision has seldom been applied to those processes and has had little success (Liu, 2004). Image analysis from such industries will be referred to here as "stochastic". The New Oxford American Dictionary (2007) defines the term *stochastic* as: "a random probability distribution that may be analyzed statistically but may not be predicted precisely". Taking the case of softwood lumber grading, this statement may be interpreted

as: it may be possible to determine that pine and spruce have statistically different wood grains, but it is not possible to predict the shape of a bark inclusion in the next sample.

3.2.6 Multivariate image analysis

In general terms, multivariate image analysis (MIA) is the application of PCA to multispectral images. Early work by Ready and Wintz (1973) used PCA for data compression and information extraction in 6-channel multispectral imagery. Later, Donker and Mulder (1977) introduced the score and loading plots in order to classify pixels in Landsat images. However, MIA as it is now known was not developed until Esbensen and Geladi (1989). MIA relies on the assumption that local intensity variations can be extracted by classifying each pixel according to their spectral characteristics without considering their spatial location within the image. Coupled with regression techniques, MIA can be used to extract the relevant features from digital images in order to build models between the sample quality and the visual features of an image. It has been widely used in fields ranging from combustion flame analysis (Yu and MacGregor, 2004; Szatvanyi et al., 2006) to flotation froth (Liu et al., 2005), snack food (Yu and MacGregor, 2003; Yu et al., 2003) as well as softwood lumber grading (Bharati et al., 2003). A complete overview of MIA can be found in Geladi and Grahn (1996) while applications to hyperspectral imaging are discussed in Grahn and Geladi (2007).

The channels of a multivariate image are highly correlated with each other, as they represent congruent images capturing the same spatial information (Bharati et al., 2003). In this way, a multivariate image $\underline{\mathbf{X}}$ ($x \times y \times \lambda$) consists of two spatial dimensions (x and y) and a spectral dimension (λ). In order to analyze such large amounts of highly correlated data, MIA relies on multi-way PCA (MPCA) methods (Esbensen and Geladi, 1989). In the same way PCA decomposes a 2-D matrix $\mathbf{X}$ ($N \times K$) into a lower dimensional subspace defined by the loadings and score vectors ($\mathbf{p}_a$ and $\mathbf{t}_a$), MPCA decomposes a 3-D matrix $\underline{\mathbf{X}}$ ($x \times y \times \lambda$) into a set of score matrices $\mathbf{T}_a$ and loading vectors $\mathbf{p}_a$:

$$\underline{\mathbf{X}} = \sum_{a=1}^{A} \mathbf{T}_a \otimes \mathbf{p}_a + \underline{\mathbf{E}} \qquad (3.29)$$

where $\otimes$ is the Kronecker product, $\underline{\mathbf{E}}$ ($x \times y \times \lambda$) is the array of residuals, and each score matrix $\mathbf{T}_a$ is a representation of the image in terms of the spectral features explained by each component and defined by the loading vectors.

Nevertheless, the above MPCA decomposition is rarely used for practical purposes. Rather, the 3-D array X is first unfolded into a 2-D matrix, which is in turn decomposed using a standard PCA. The term "unfolding" was first coined by Wold et al., (1987) and consists of reorganizing the pixels into a matrix **X** of dimensions ($xy \times \lambda$) as shown in Figure 3.11 (i.e. the spectrum of each pixel are collected row-wise in matrix **X**). In doing so, the spatial organization of the pixels (e.g. texture) is destroyed. Using this unfolding operation, the multi-way PCA method treats each pixel as a separate object independent of its neighbors (Bharati et al., 2003). Such a decomposition of the variance-covariance structure of the spectral matrix allows extracting and classifying the spectral signature of each pixel regardless of their spatial location.

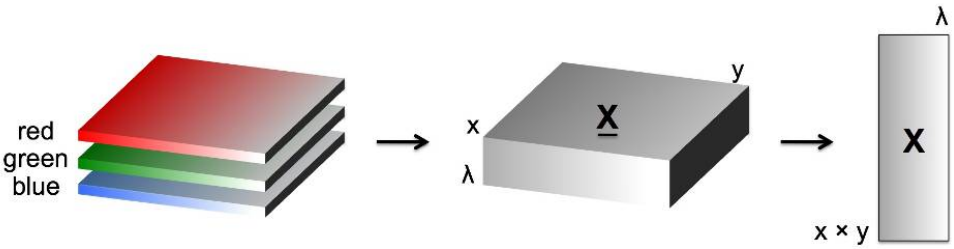

*Figure 3.11. Reorganizing a 3-D image array ($\underline{\mathbf{X}}$) into a long and narrow 2-D matrix (**X**) prior to PCA decomposition.*

While it is possible to directly perform a standard PCA decomposition on the unfolded spectral image, an additional step is often preferred. In image analysis **X** usually has a much larger number of rows than columns (i.e. $xy >> \lambda$). Therefore, it is much less computationally intensive to perform PCA on the kernel matrix **Z**, shown below, which has much smaller dimensions ($\lambda \times \lambda$ vs. $xy \times \lambda$) (Geladi and Grahn, 1996; Liu and MacGregor, 2007):

$$\mathbf{Z} = \mathbf{X}^{\mathrm{T}}\mathbf{X} \tag{3.30}$$

If a PCA model is to be developed on a set of *N* images, rather than on a single one, the kernel matrix is adapted to take into account the entire image set:

$$\mathbf{Z} = \sum_{n=1}^{N} \mathbf{X}_n^{\mathrm{T}}\mathbf{X}_n \tag{3.31}$$

When PCA is applied to a data matrix, the clustering pattern of the extracted features is observed using $\mathbf{t_i}$-$\mathbf{t_j}$ score scatter plots (i≠j). However, when applied to a multivariate image, these score plots are typically shown as color-coded 2-D density histograms in order to ease visual interpretation, since the large number of observations in the score plot is equal to the number of pixels in the image (i.e. *xy* observations for each image). Pixels having similar spectral features in the original image yield similar ($\mathbf{t_1}$, $\mathbf{t_2}$) score combinations, which results in many overlapping points in this scatter plot (Figure 3.12). The number of overlapping pixels represented by a single point in a score plot is called the pixel density. This color-coding scheme uses black to represent bins containing no pixels, and various shades of red, orange, and yellow to represent increasing pixel densities, and ultimately uses white to represent bins having the highest pixel density.

The tendency of similarly colored pixels to project to similar $\mathbf{t_1}$-$\mathbf{t_2}$ locations can be used to analyze the spectral information contained within the image. Thus, segmentation may be performed in the lower dimensional space (i.e. score space) instead of in the original image space. To take a simple example, a region of interest (ROI) drawn in the score histogram in Figure 3.12 as been hand-selected and masked in blue and all the pixels with color characteristics corresponding to those $\mathbf{t_1}$-$\mathbf{t_2}$ coordinates were traced back into the original image. In this case the ROI was selecting the green leaf coloring. By dismissing the spatial organization of the pixels in the image, the MIA analysis could effectively locate any green pixels anywhere in the image. While simple, this methodology has been used successfully

in such applications as softwood lumber grading (Bharati et al., 2003) and flame analysis (Szatvanyi et al., 2006).

Illustrated here using RGB images, this method has the added advantage of strongly simplifying the analysis of spectral images. By identifying the main sources of variance within the spectra, the MIA can help identify important features within the image that may not easily be detected otherwise.

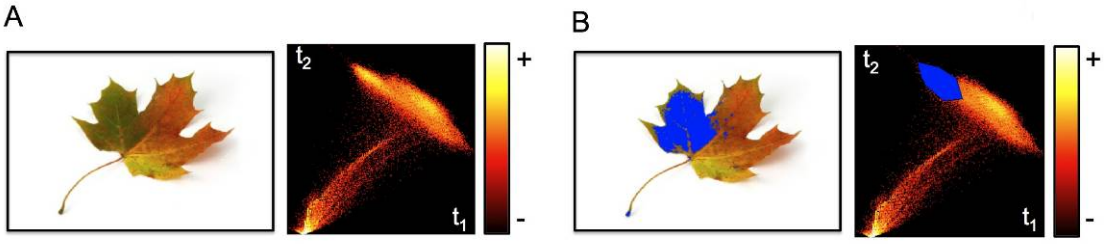

Figure 3.12. Illustration of a MIA t_1-t_2 *score histogram of a maple leaf (A) in which all the pixels having a similar color combination, in this case green, are projected to similar* t_1-t_2 *coordinates (B).*

An alternative way to visualize the image features extracted by each principal component is obtained by refolding the score vectors $\mathbf{t}_a$ ($xy \times 1$) into score matrices $\mathbf{T}_a$ ($x \times y$) according to the original spatial location of each pixel in the image $\underline{\mathbf{X}}$. This step is akin to computing $\mathbf{T}_a$ directly using Equation 3.29. The score matrices $\mathbf{T}_a$ are then displayed as univariate gray-level images after appropriate scaling is applied (Figure 3.13). As mentioned previously, the first component captures the largest possible source of variance within the image while each subsequent score vector captures smaller sources of covariance. In RGB images, $\mathbf{T_1}$ and $\mathbf{T_2}$ typically represent the vast majority of the information (≈99%), and $\mathbf{T_3}$ is often associated with irrelevant information and/or noise.

MIA extracts spectral features independent of their spatial location. It is this characteristic that makes it useful for relating images to product quality properties (Bharati, 2002; Yu, 2003). However, this disregard for the spatial information makes it unsuitable for the analysis of spatial features (i.e. image texture), therefore limiting its application to a variety

of image analysis tasks (Liu, 2004). The next sections discuss a number of methods that take spatial features into account.

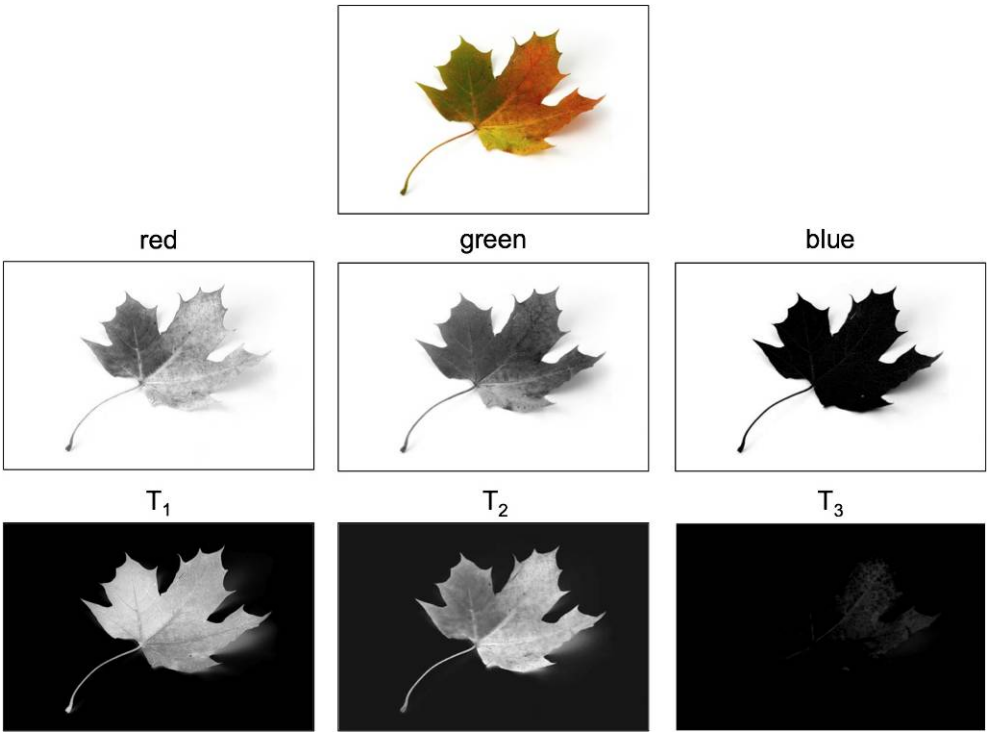

*Figure 3.13. A color image alternately represented as a set of gray-scale images (red, green, and blue) or as a set of gray-scale score images (**T_1**, **T_2** and **T_3**).*

3.2.7 Texture Analysis

The unevenness patterns of the pixels in a region give rise to the visual perception of image texture (Xu and Chen, 2006). Texture reflects the variations of the optical properties of object surfaces and different materials usually produce distinctive surface textures. This makes image texture an important source for discriminating information and has been a subject of intense study by many researchers. Liu (2004) defines 4 classes of methods for texture analysis:

Statistical methods

They describe texture of regions in an image through higher-order moments of their grayscale histograms (Tuceryan and Jain, 1998). First order moments (e.g. mean) do not consider pixel neighbor relationships, second order moments consider the relationship between groups of two pixels and third and higher order moments consider relationships among three or more pixels (Hall-Beyer, 2008). While higher-class methods exist, they are not commonly implemented due to long calculation times and difficulties in interpretation (Akono et al., 2003). The most frequently cited statistical method is the gray-level co-occurrence matrix (GLCM) proposed by Haralick et al. (1973). This method is based on the use of second order statistics of the gray-scale image histograms.

Structural methods

These methods describe the texture as a combination of well-defined texture elements. This may be achieved, for example, through regularly space parallel lines (Tuceryan and Jain, 1998). These methods however appear to be limited to describe very regular textures and are rarely used in practice (Gonzalez and Woods, 1993).

Model-based methods

According to Liu (2004), these methods generate an empirical model of each pixel in the image based on a weighted average of the pixel intensities in its neighborhood. For example, fractal models presented by Keller et al. (1989) create unique fractal patterns that describe the image texture by following a set of well-defined rules. According to Mandelbrot (1983), fractals provide a proper mathematical framework to study the irregular, complex shapes found in nature.

Transform based

These methods convert the image into a new form using the spatial frequency properties of the pixel intensity variations (Liu, 2004). These methods include the Fourier and Wavelet transforms. Unlike the infinite wave functions used in the Fourier transform, a wavelet is a finite-length fast-decaying oscillating wave. This characteristic provides the wavelet transform a distinct advantage over the Fourier transform when representing non-periodic functions containing discontinuities and/or sharp peaks. The finite duration of the wavelet also allows space/time-frequency decomposition of the image signals.

3.2.7.1 Gray-level co-occurrence matrix

Gray-level co-occurrence matrices (GLCM) have shown their worth for quantifying image texture in a number of application such as for classifying images of steel sheets (Bharati et al., 2002, 2004), flotation froths (Bartolacci et al., 2005) and even evergreens (Samal et al., 2006). GLCM is based on second-order statistics (Akono et al., 2003) that compares two neighboring pixels at a time and compiles the frequency for which differing gray-levels can be found within a restricted area. A rough image will have numerous and significant black/white transitions whereas a smoother image will not. The GLCM texture analysis is used in Chapter 7 to identify fiber-to-fiber as well as sample-wide variations in extruded wood/plastic composites.

A gray-scale image represents the entire range of light intensities as a finite number of levels of gray. In practice this number is often 256 (encoding an image in 8 bits = 2^8 gray-levels), but can be much greater in high-end equipment. Using this scheme, each pixel of a gray-scale image is represented by a value between 0-255; allowing the image to be treated as a simple matrix.

The GLCM is a tabulation of how often different combinations of pixel brightness values (gray levels) occur in an image (Hall-Beyer, 2008). According to Haralick et al. (1973), three parameters must be taken into account when computing the GLCM of an image: (1)

the number of gray-levels, which can be lower than the number of gray levels of the image (e.g. 256 for 8 bits), (2) the length of transition and (3) the angle of displacement. As illustrated in Figure 3.14A, the algorithm essentially compares the gray-level of a reference pixel to that of neighboring pixel (in this figure the neighbor is located directly to the right of the reference, noted a [0,1] transitions in a Cartesian-style reference). However, to analyze coarser textures, this second pixel can be taken at a certain distance, which is the length of transition. In Chapter 7, we use horizontal transitions of [0,2], [0,5], [0,10] and vertical transitions of [2,0], [5,0], [10,0]. Combining these results allows the analysis of textures at different resolutions. The result can be considered a roughly diagonal matrix in which increasing distance from the main diagonal indicates increasing roughness (Figure 3.14B).

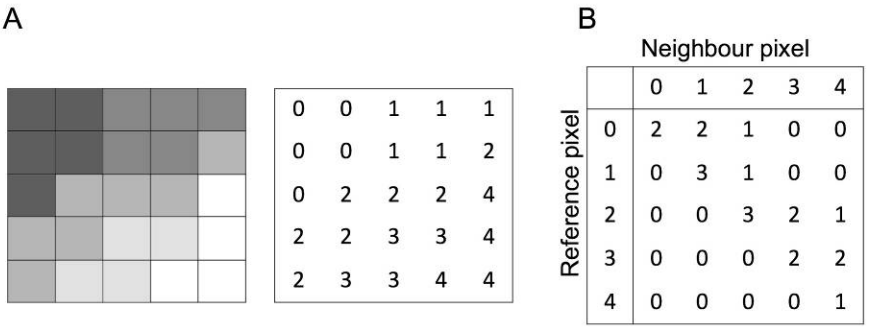

Figure 3.14. Illustration of texture analysis using the GLCM algorithm: (A) converting a gray-scale image into numerical values, and (B) compiling the differences in intensity between two neighboring pixels in the [0,1] *direction.*

A number of scalar textural descriptors (or textural features) have been proposed for quantifying the texture of an image from the GLCM matrices. The most common of these descriptors, often referred to as Haralick features, are shown in Table 3.1. Indices i and j are the row and column numbers in the GLCM matrix, $p_{i,j}$ is the normalized value of an element (i,j) of the matrix and μ and σ are average and standard deviation of $p_{i,j}$ over i and j. This normalization ensures that the sum of the elements of the matrix is equal to 1.

Table 3.1. Common Haralick features computed from the co-occurrence probabilities (Haralick, 1979).

Energy (uniformity)	$\sum_{i,j}\left(p_{i,j}\right)^2$	(3.32)
Entropy	$\sum_{i,j} p_{i,j}\log\left(p_{i,j}\right)$	(3.33)
Maximum probability	$\max_{i,j}\left(p_{i,j}\right)$	(3.34)
Contrast	$\sum_{i,j}\lvert i-j\rvert^k\left(p_{i,j}\right)^2$	(3.35)
Inverse difference moment	$\sum_{i,j}\frac{\left(p_{i,j}\right)^2}{\lvert i-j\rvert^k}$ *where* $i\neq j$	(3.36)
Correlation	$\sum_{i,j}\frac{(i-\mu_i)(j-\mu_j)p_{i,j}}{\sigma_i\sigma_j}$	(3.37)

Optimization of this method requires the fine-tuning of various parameters; the distance of transition depends on the size of the textural features while the angle of displacement depends more on surface isotropy. Raising the number of gray-levels simultaneously increases the sensitivity to low-contrast textural features, as well as the size of the GLCM matrix, occupying more memory and requiring calculation time.

3.2.7.2 Wavelet texture analysis

Wavelet Texture Analysis (WTA) has been used successfully in numerous process applications ranging from analysis of flotation froth, to countertops (Bartolacci et al., 2005; Liu and MacGregor, 2006). WTA is typically suited in applications where data contains both large and small scales of variations, in which case it is capable of isolating the former from the latter. To take a simple example, this is akin to distinguishing between daily, weekly and seasonal variations in temperature. In this thesis, WTA is used in Chapter 5 to determine the degree and direction of stretching of plastic films blends.

In wavelet analysis, a continuous signal $f(x)$ is decomposed in terms of a family of orthonormal bases $\psi_{m,n}(x)$ obtained through dilation and translation of a mother wavelet $\psi(x)$:

$$\psi_{m,n}(x) = 2^{-m/2}\psi\left(2^{-m}x - n\right) \tag{3.38}$$

where m and n are the continuous scale and shift parameters, respectively ($m \in Z^+$ and $n \in Z$). For this task, numerous mother wavelets have been proposed in the literature (*Burke Hubbard, 1998).* According to their orthonormal property, the continuous wavelet transform (CWT) coefficients ($c_{m,n}$) can then be defined as the convolution of the signal with these wavelet bases.

$$c_{m,n} = \int f(x)\psi_{m,n}(x)dx = \left\langle \psi_{m,n}(x), f(x) \right\rangle \tag{3.39}$$

The coefficients are maximized when the frequency of the signal is equivalent to that of the dilated wavelet. The discrete wavelet transform (DWT) approximates the continuous wavelet transform by recursively decomposing the original signal at a dyadic scale: each level of decomposition differing by a factor of two from the preceding resolution (Teppola and Minkkinen, 2000).

$$m = 2^j \tag{3.40}$$

$$n = k2^j \tag{3.41}$$

where $j \in Z+$ and $k \in Z$. In practice, it was shown by Mallat (1989) that orthogonal wavelet analysis could be implemented using the filters traditionally used in the signal processing field. The filter-based implementation makes use of a scaling function $\phi(x)$ used to dilate the mother wavelet $\psi(x)$ in order to have it correspond to the desired level of scrutiny. The sequences $h_0[k]$ and $h_1[k]$ are low- and high-pass discrete filters used in the implementation of the wavelets (Vetterli and Kovacevic, 1995; Liu, 2004):

$$\psi(x) = \sqrt{2}\sum_k h_1[k]\phi(2x - k) \tag{3.42}$$

where

$$\phi(x) = \sqrt{2}\sum_{k} h_0[k]\phi(2x-k) \tag{3.43}$$

$$h_1[k] = (-1)^k h_0[1-k] \tag{3.44}$$

Using these equations, it can be shown that the DWT can be computed for any level j and position l without explicitly using $\phi(x)$ and $\psi(x)$ (Liu, 2004).

$$\phi_{j,l}[k] = 2^{j/2} h_0^{(j)}[k - 2^j l] \tag{3.45}$$

$$\psi_{j,l}[k] = 2^{j/2} h_1^{(j)}[k - 2^j l] \tag{3.46}$$

where

$$h_0^{(j+1)}[k] = [h_0]_{\uparrow 2^j} \times h_0^{(j)}[k] \quad \text{and} \quad h_0^{(0)}[k] = h_0[k] \tag{3.47}$$

$$h_1^{(j+1)}[k] = [h_1]_{\uparrow 2^j} \times h_1^{(j)}[k] \quad \text{and} \quad h_1^{(0)}[k] = h_1[k] \tag{3.48}$$

where $[\cdot]_{\uparrow 2^j}$ denotes up-sampling by a factor of 2^j (dyadic scale), j is the scale index and and l is the translation index (Mallat, 1989; Daubechies, 1992; Rioul and Duhamel, 1992). Two types of DWT coefficients characterize the signal: the detail coefficients ($d_{j,l}$) characterizes the match between the mother wavelet and the original signal at a scale (j) and a position (l) while the approximation coefficients ($a_{j,l}$) corresponds to the remaining information after extraction of the features by the detail coefficients. This rougher representation of the signal is then used as the basis for the $d_{j+1,l}$ coefficients. Recalling the continuous coefficients ($c_{m,n}$), the discrete coefficients of a signal $f(k)$ can be computed as the convolution of the signal expressed in terms of k (a positive integer) and the wavelet or scaling functions (Mallat, 1989; Daubechies, 1992; Liu and MacGregor, 2005, 2006):

$$a_{j,l} = \left\langle f(k), \phi_{j,l}(k) \right\rangle \tag{3.49}$$

$$d_{j,l} = \left\langle f(k), \psi_{j,l}(k) \right\rangle \tag{3.50}$$

The values of these coefficients are maximized when the frequency of the dilated wavelet matches that of the signal. Multiresolution analysis is achieved by associating the dilated wavelet and the wavelet coefficients ($a_{j,l}$ and $d_{j,l}$) to low-pass and high-pass filters, respectively. As shown in Figure 3.15, the first-level of decomposition extracts the highest frequency features $d_1f(x)$ from the original signal, leaving behind a slightly smoother approximation $a_1f(x)$ (not shown here). Level 2 then extracts the highest frequency information $d_2f(x)$ among the remaining details 2^j times larger than the previous, from $a_1f(x)$, leaving an even smoother approximation of the original image. Each successive level then isolates a new level of coarser details from an increasingly simplified low-pass approximation.

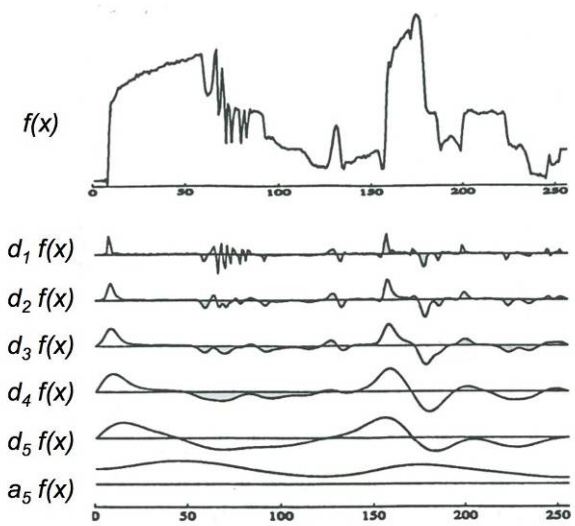

Figure 3.15. A discrete wavelet transform applied to a signal illustrating 5 successive decomposition level coefficients and the remaining 5th level approximation of the original signal. Based on Burke Hubbard (1998).

Analysis of a 2-D image, rather than a 1-D signal, can be achieved through the proper combination of horizontal and vertical 1-D wavelets and their corresponding scaling functions (Vetterli and Kovacevic, 1995; Burke Hubbard, 1998). One way to perform this is through a separable solution, which uses separable filters (i.e. h_0 and h_1) within a multidimensional filter bank as shown in Figure 3.16. A filter bank is an array of band-pass filters that separates an input signal into multiple output components, each one carrying a subband frequency of the original signal.

The algorithm illustrated in Figure 3.16 is basically an extension of the 1-D algorithm discussed earlier in that it converts a low-level input signal (here a gray-scale image A_{j-1}) into higher-level detail signals (horizontal, vertical and diagonal detail images D_j^H, D_j^V and D_j^D) plus a residual approximation signal (approximation image A_j). This is achieved through a two-step process.

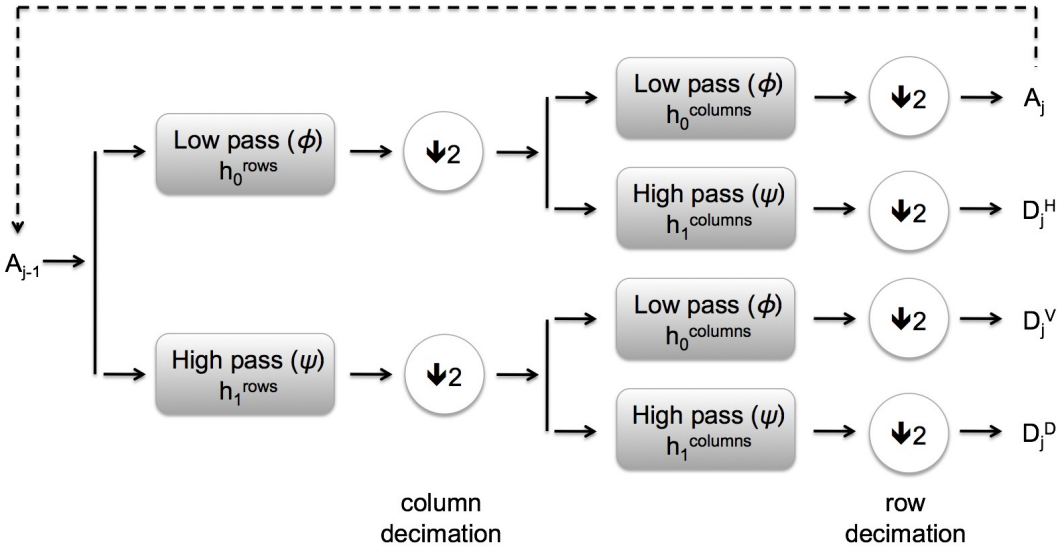

Figure 3.16. A 2-D separable solution for applying a discrete wavelet transform to an image. Adapted from Liu et al. (2005).

First, each row is filtered using the low- and high pass filters (h_0 and h_1), followed by a dyadic column decimation. Then, the resulting low- and high-frequency information along the horizontal direction is filtered vertically (i.e. along each column) using both filters a second time. This second filtering step is followed by a dyadic row decimation. As mentioned, the output of this algorithm consists of 4 images containing the wavelet coefficients:

- Approximation image A_j captures the low-pass information in both directions and is used as an input approximation image for the $(j+1)^{th}$ level WTA decomposition.
- Horizontal detail image D_j^H captures the low-pass horizontal and high-pass vertical features of the image, effectively capturing the high frequency horizontal details (i.e. the vertical edges).

- Vertical detail image D_j^V captures the high-pass horizontal and low-pass vertical features of the image, capturing the high frequency vertical details (i.e. the horizontal edges).
- Diagonal detail image D_j^D captures the high-pass information in both directions. This sensitivity to high frequency horizontal and vertical details essentially conveys information on the high frequency diagonal details.

It is important to note that the j^{th} level images are 4 times smaller than the $(j\text{-}1)^{th}$ level images due to the dyadic column and row decimations. This reduction in resolution between each level allows the algorithm to extract larger and larger textural features.

An example of a 2-level 2-D wavelet decomposition is presented in Figure 3.17 in which progressively larger features are removed from the original image with each level. Further information on image analysis using multidimensional wavelets can be found in Burke Hubbard (1998) and Takei et al. (2004).

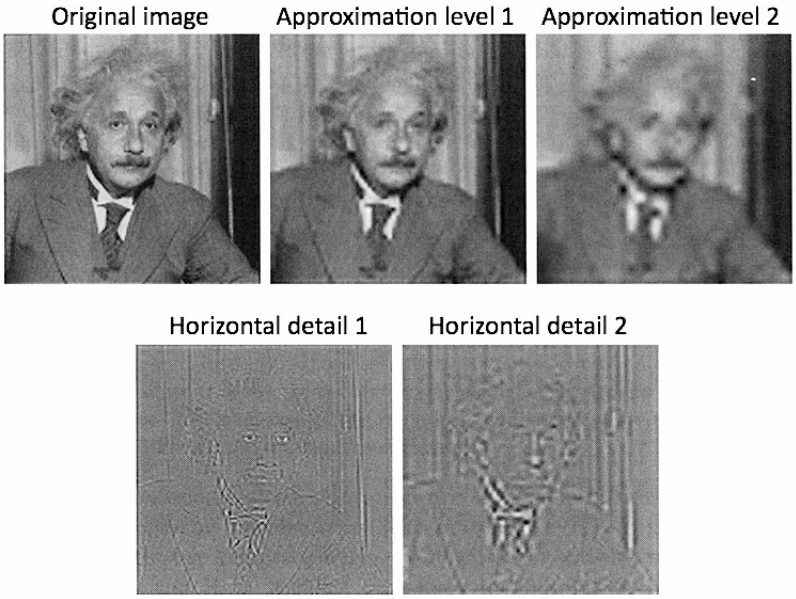

Figure 3.17. Illustration of a 2-D wavelet decomposition. The 1st level approximation image is obtained by removing the 1st level detail coefficients from the original image. This methodology is applied recursively to produce each successive higher-level approximation image. Only the horizontal detail coefficients are shown here. Based on Walker (1999).

3.2.8 Multiresolutional multivariate image analysis

Since MIA is intended to extract the spectral features of a multivariate image independently of their spatial location, it is unsuitable for the analysis of spatial (i.e. textural) features (Liu, 2004). While textural methods such as GLCM and WTA are intended to analyze such spatial features in gray level images, they cannot readily be used in multivariate images. Multiresolutional MIA (MR-MIA) is intended to combine the characteristics of both these classes of methods. According to Liu and MacGregor (2007), as well as Teppola and Minkkinen (2000), this can be achieved in one of two ways:

MR-MIA I

1. Apply multiresolution decomposition on each image gray-scale forming the multivariate image (i.e. images at each wavelength) down to an appropriate level J. In practice, this often implies decomposition levels ranging from 3 to 6. If there are λ variables in the original multivariate image (i.e. an RGB image would have λ=3), multiresolution decomposition yields a new multivariate image with λ variables at each resolution.
2. Apply a PCA to each of the J multivariate images. This gives score images and loading vectors at each resolution.

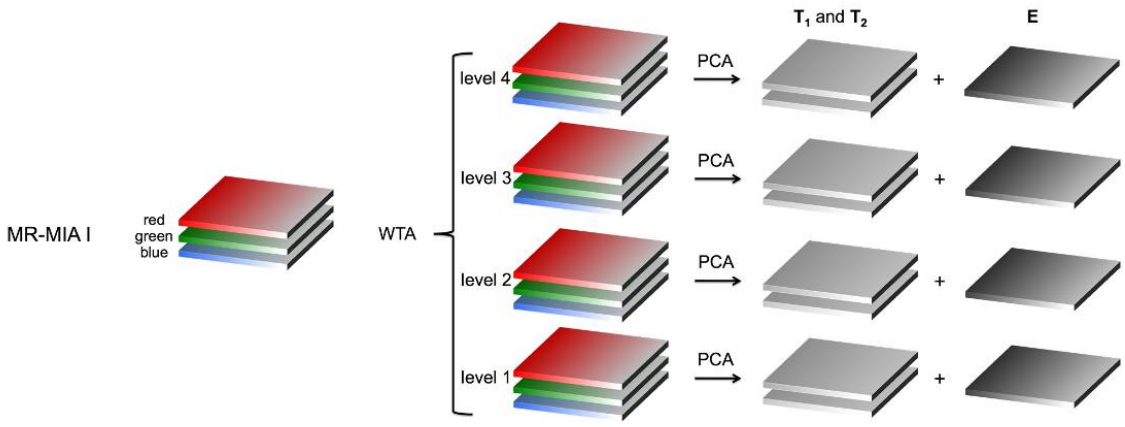

Figure 3.18. Illustration of the MR-MIA I algorithm for an RGB image (λ=3) using J=4 decomposition levels in the wavelet texture analysis (WTA) and 2 principal components in the principal component analysis (PCA).

MR-MIA I makes it possible to determine the important factors present at each decomposition level and at each wavelength. In doing so, it characterizes the surface with a very high level of precision. This algorithm is best used when there are important interactions between the spectral and spectral features, for example when some textural features are only observed at specific wavelengths.

However, MR-MIA I implies a full multiresolutional decomposition for each observed wavelength in the image. While this may be feasible for RGB imaging, it quickly becomes overwhelming in hyperspectral imaging. It would imply hundreds of parallel texture analyses, each of which followed by a PCA decomposition. Doing so leads to very lengthy calculations and large amounts of data to interpret. In such a case, the user would be well advised to perform an initial wavelenght selection on the spectral image. Chapter 6 deals exclusively with the subject of wavelenght selection.

MR-MIA II

1. Apply a PCA to the original multivariate image in order to obtain a single set of score images and loading vectors.
2. Apply a multiresolution decomposition on each score image. If there are A principal components in the PCA, there will be A multiresolution decompositions, each performed down to J levels.

By minimizing the required computations, MR-MIA II also simplifies the interpretation of the results. In practice, the PCA reduces the information down to a very low number of significant components before performing the texture analyses. Nevertheless, MR-MIA II relies on the hypothesis that there is little or no interaction between the spectral and spatial features. Doing so effectively smoothes over local variations when they are strongly correlated. In practice however, weak interaction between spectral and spatial features is not uncommon: spatial features captured by an RGB camera, for example, often appear simultaneously in the red, green of blue channels.

In spite of these shortcomings, MR-MIA has been used successfully. Gao and Ren (2005) used MR-MIA I to differentiate between nickel, cadmium, copper and zing surfaces whereas Teppola and Minkkinen (2000) and Liu et al. (2005) both used MR-MIA II to study a water purification plant and a zinc flotation column, respectively. In this study, MR-MIA II will be used to analyze an extrusion film-blowing process in Chapter 5.

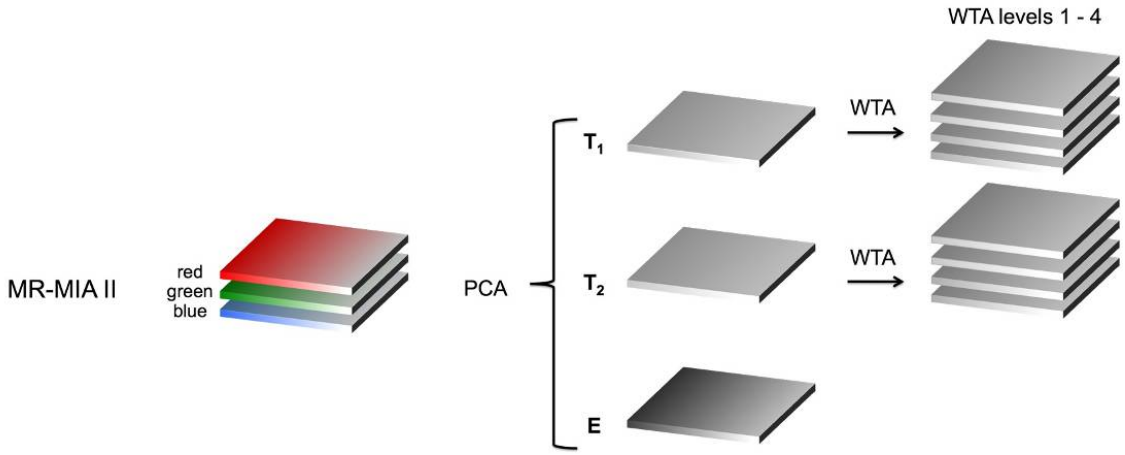

Figure 3.19. Illustration of the MR-MIA II algorithm for an RGB image using 4 decomposition levels in the wavelet texture analysis (WTA) and 2 principal components in the principal component analysis (PCA).

3.3 References

Akono A., Tonyé E. and Nyoungui A.N., Nouvelle méthodologie d'évaluation des paramètres de texture d'ordre trois, *Int. J. Remote Sens.*, 24(9), 1957-1967 (2003).

American Heritage Dictionary, Houghton Mifflin Harcourt: USA (2009).

Bakeev K., *Process Analytical Technology : Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, Wiley-Blackwell: UK (2005).

Bartolacci G., Pelletier P., Tessier J., Duchesne C., Bossé P.A. and Fournier J., Application of numerical image analysis to process diagnosis and physical parameter measurement in mineral processes - part I: flotation control based on froth textural characteristics, *Centenary of Flotation Symposium*, 73-84 (2005).

Berson A., Smith S. and Thearling K., *Building Data Mining Applications for CRM*, McGraw-Hill: USA (1999).

Bharati M.H., *Multivariate Image Analysis and Regression for Industrial Process Monitoring and Product Quality Control*, Ph.D. thesis, McMaster University: Canada (2002).

Bharati M.H., Tropper W. and MacGregor J.F., Softwood lumber grading through on-line multivariate analysis techniques, *Ind. Eng. Chem. Res.*, 42, 5345-5353 (2003).

Bharati M.H., Liu J.J. and MacGregor J.F., Image texture analysis: methods and comparisons, *Chemometr. Intell. Lab.*, 72, 57-71 (2004).

Brás L.P., Lopes M., Ferreira A.P. and Menezes J.C., A bootstrap-based strategy for spectral interval selection in PLS regression, *J. Chemom.*, 22, 695-700 (2008).

Burke Hubbard B., *The world according to wavelets, the story of a mathematical technique in the making,* A.K. Peters: USA (1998).

Coates P.D., Barnes S.E., Sibley M.G., Brown E.C., Edwards H.G.M. and Scowen I.J., In-process vibrational spectroscopy and ultrasound measurements in polymer melt extrusion, *Polymer*, 44, 5937-5949 (2003).

Color Information Resource (2007).
http://www.colourtherapyhealing.com/colour/electromagnetic_spectrum.php.

Daubechies I., *Ten Lectures on Wavelets*, SIAM: USA (1992).

Donker N. and Mulder N., Analysis of MSS digital imagery with the aid of principal component transform, *ITC Journal*, 3, 434-466 (1977).

ElMasry G., Wanga N., Vigneault C., Qiao J. and ElSayed A., Early detection of apple bruises on different background colors using hyperspectral imaging, *LWT*, 41, 337-345 (2008).

Eriksson L., Johansson, E., Kettaneh-Wold, N. and Wold S., *Multi- and Megavariate Data Analysis*, Umetrics Academy: Sweden (2001).

Esbensen K.H. and Geladi P., Strategy of multivariate image analysis (MIA), *Chemometr. Intell. Lab.*, 7, 67-86 (1989).

Feldhoff R., Wienke D., Cammann K. and Fuchs H., On-line post consumer package identification by NIR spectroscopy combined with a fuzzy ARTMAP classifier in an industrial environment, *Appl. Spectrosc.*, 51(3), 362-368 (1997).

Gabrys B., Baruque B. and Corchado E., *Lecture Notes in Computer Science*, Springer-Verlag: Germany (2006).

Gao L. and Ren S., Simultaneous spectroscopic determination of four metals by two kinds of partial least squares methods, *Spectrochim. Acta, Part A*, 61, 3013-3019 (2005).

Geladi P. and Grahn H., *Multivariate Image Analysis*, Wiley: United Kingdom (1996).

Geladi P. and Kowalski B.R., Partial least-squares regression: a tutorial, *Anal. Chim. Acta*, 186, 1-17 (1986).

Gonzalez R.C. and Woods R.E., *Digital Image Processing*, Addison-Wesley: USA (1993).

Grahn H. and Geladi P., *Techniques and Applications of Hyperspectral Image Analysis*, Wiley: USA (2007).

Groth R.*, Data Mining: Building Competitive Advantage,* Prentice Hall PTR: USA (2000).

Hall-Beyer M., The GLCM texture tutorial (2008).
http://www.fp.ucalgary.ca/mhallbey/.

Haralick R.M., Shanmugam K. and Dinstein I., Textural features for image classification, *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, 6, 610-621 (1973).

Haralick R.M., Statistical and structural approaches to texture, *Proceedings IEEE*, 67(5), 786-804 (1979).

International Chemometrics Society (2009).
http://www.namics.nysaes.cornell.edu/.

Jackson J.E., A User's Guide to Principal Components, Wiley and sons: Canada (1991).

Johnson S., *Stephen Johnson on Digital Photography*, O'Reilly: USA (2006).

Jolliffe T., A note on the use of principal components in regression, *J. Roy. Stat. Soc. C-App.*, 31(3), 300-303 (1982).

Keller J.M., Chen S.S. and Crownover R.M., Texture description and segmentation through fractal geometry, *CRGIP*, 45(2), 150-166 (1989).

Kohler A., Zimonja M.S., Segtnan V.H. and Martens H., Data preprocessing: SNV, MSC and EMSC pre-processing in biospectroscopy, *Comprehensive Chemometrics, Four-Volume Set: Chemical and Biochemical Data Analysis*, 2, 139-162 (2009).

Leitner R., Mairer H. and Kercek A., Real-time classification of polymers with NIR spectral imaging and blob analysis, *Real-Time Imaging*, 9, 245-251 (2003).

Liu J.J., *Machine Vision for Process Industries: Monitoring, Control, and Optimization of Visual Quality of Processes and Products*, Ph.D. thesis, McMaster: Canada (2004).

Liu J.J., MacGregor J.F., Duchesne C. and Bartolacci G., Flotation Froth Monitoring using Multiresolutional Multivariate Image Analysis, *Miner. Eng.*, 18, 65-76 (2005).

Liu J.J. and MacGregor J.F., Modeling and optimization of product appearance: application to injection-molded plastic panels, *Ind. Eng. Chem. Res.*, 44, 4687-4696 (2005).

Liu J.J. and MacGregor J.F., Estimation and monitoring of product aesthetics: application to manufacturing of "engineered stone" countertops, *Mach. Vision Appl.*, 16(6), 374-383 (2006).

Liu J.J. and MacGregor J.F., On the extraction of spectral and spatial information from images, *Chemometr. Intell. Lab.*, 85, 119-130 (2007).

Liu Z., Yu H. and MacGregor J.F., Standardization of line-scan NIR imaging systems, *J. Chemometr.*, 21, 88-95 (2007).

MacGregor J.F., Bruwer M.J., Cardin M., Wallace D. and Liu Z., *Multivariate Methods for Process Analysis, Monitoring and Quality Improvement*, ProSensus Multivariate Data Course: Canada (2008).

Mallat S.G., A theory for multiresolution signal decomposition: the wavelet representation, *IEEE T. Pattern Anal.*, 11(7), 674-693 (1989).

Mandelbrot B.B., *The Fractal Geometry of Nature*, Freeman: USA (1983).

Marshall A.D. and Martin R.R., *Computer Vision, Models and Inspection*, World Scientific Publishing: Singapore (1992).

Merriam-Webster Online Dictionary (2009).
www.merriam-webster.com/dictionary/outlier

New Oxford American Dictionary, Apple: USA (2007).

Nikon (2009).
www.maxmax.com/nikon_d200_study.htm.

Olah M., Bologa C. and Oprea T.I., An automated PLS search for biologically relevant QSAR descriptors, *J. Comput. Aided Mol. Des.*, 18, 437-449 (2004).

Pearson K., On line and planes of closest fit to systems of points in space, *Philos. Mag.*, 2(6), 559-572 (1901).

Qiao J., Ngadi M.O., Wang N., Gariépy C. and Prasher S.O., Pork quality and marbling level assessment using a hyperspectral imaging system, *J. Food Eng.*, 83, 10-16 (2007).

Ready P. and Wintz P., Information extraction, SNR improvement, and data compression in multispectral imagery, *IEEE Trans. Commun.*, COM-21, 1123-1131 (1973).

Rioul O. and Duhamel P., Fast algorithms for discrete and continuous wavelet transforms. *IEEE Trans. Inform. Theory*, 38(2), 569-586 (1992).

Samal A., Brandle J.R. and Zhang D., Texture as the basis for individual tree identification, *Inform. Sci.*, 176, 565-576 (2006).

Santen A., Koot G.L.M. and Zullo L.C., Statistical data analysis of a chemical plant, *Comp. Chem. Eng.*, 21, 1123-1129 (1997).

Shippert P., Introduction to Hyperspectral Image Analysis, Research Systems Inc (2009). http://satjournal.tcom.ohiou.edu/pdf/shippert.pdf.

Specim (2009).
www.specim.fi.

Szatvanyi G., Duchesne C. and Bartolacci G., Multivariate Image Analysis of Flames for Product Quality and Combustion Control in Rotary Kilns, *J. Am. Chem. Soc.*, 5(13), 4706-4715 (2006).

Takei M., Ochi M. and Saito Y., Image extraction of particle concentration at the plug front using 3D wavelet and comparison with LDV, *Powder Tech.*, 142, 70-78 (2004).

Teppola P. and Minkkinen P., Wavelet-PLS regression models for both exploratory data analysis and process monitoring, *J. Chemometr.*, 14, 383-399 (2000).

Tuceryan M. and Jain A.K., *Handbook of Pattern Recognition and Computer Vision,* World Scientific Publishing: USA (1998).

van den Broek W.H.A.M., Derks E.P.P.A., van de Ven E.W., Wienke D., Geladi P. and Buydens L.M.C., Plastic Identification by Remote Sensing Spectroscopic NIR Imaging using Kernel Partial Least Squares (KPLS), *Chemometr. Intell. Lab.*, 35, 187-197 (1996).

Vetterli M. and Kovacevic J., *Wavelets and Subband Coding*, Prentice Hall PTR: USA (1995).

Walker J.S., *A Primer On Wavelets And Their Scientific Applications*, C.R.C. Press: USA (1999).

Wang X.D. and Shao H.H., Non-linear system modeling using the radial basis function neural networks, *Lect. Notes Contr. Inf.*, 14(1), 59-66 (1997).

Webster's New World College Dictionary, Wiley: USA (2005).

Wise B.M. and Gallagher N.B., The process chemometrics approach to process monitoring and fault detection, *J. Proc. Control*, 6(6), 329-348 (1996).

Wold S., Geladi P., Esbensen K. and Ohman J., Multi-way principal component and PLS analysis, *J. Chemom.*, 1, 41-56 (1987).

Wold S., Chemometrics; what do we mean with it and what do we want from it, *Chemometr. Intell. Lab.*, 30, 109-115 (1995).

Wold S., Eriksson L., Trygg J. and Kattaneh N., The PLS method - partial least squares projections to latent structures - and its applications in industrial RDP (research, development, and production), *Compstat*, Prague (2004).

Xu C.L. and Chen Y.Q., Classifying image texture with statistical landscape features, *Pattern Anal. Appl.*, 8(4), 321-331 (2006).

Yan W., Shao H. and Wang X., Soft sensing modeling based on support vector machine and Bayesian model selection, *Comput. Chem. Eng.*, 28, 1489-1498 (2004).

Yu H., *Development of Vision-based Inferential Sensors for Process Monitoring and Control*, Ph.D. thesis, McMaster University: Canada (2003).

Yu H., MacGregor J.F., Haarsma G. and Bourg W., Digital Imaging for Online Monitoring and Control of Industrial Snack Food Processes, *Ind. Eng. Chem. Res.*, 42, 3036-3044 (2003).

Yu H. and MacGregor J.F., Multivariate Image Analysis and Regression for Prediction of Coating Content and Distribution in the Production of Snack Foods, *Chemometr. Intell. Lab.*, 67, 125-144 (2003).

Yu H. and MacGregor J.F., Monitoring Flames in an Industrial Boiler Using Multivariate Image Analysis, *AIChE J.*, 50, 1474-1483 (2004).

Zhong W. and Yu J., MIMO soft sensors for estimating product quality with on-line correction, *Trans. IChemE.*, 78(A), 612-620 (2000).

Chapter 4

On-line prediction of crystallinity spatial distribution across polymer films using NIR spectral imaging and chemometrics methods*

Résumé

Un capteur basé sur l'imagerie en proche infrarouge (NIR) rapide et non intrusif a été mis au point dans cet article afin de faire le suivi des variations spatio-temporelles de cristallinité à la surface des films de polymères. Une approche par régression et analyse d'images multivariées (MIA/MIR) a été proposée et comparée aux techniques de calibration NIR utilisant des spectres moyens ou des dérivées du second ordre. Chacune des approches a permis d'obtenir des prédictions des variations de cristallinité locales et globales d'échantillons de polymères HDPE, LDPE et PP. Nos résultats montrent que de faibles variations dans la cristallinité introduites par des changements dans les vitesses de refroidissement peuvent être prédites à l'intérieur des erreurs expérimentales. Les distributions spatiales de cristallinité ont également été validées et s'avèrent cohérentes avec les conditions de traitement.

Abstract

A rapid non-intrusive on-line NIR imaging sensor is developed for monitoring spatio-temporal crystallinity variations across the surface of polymer films, which strongly affects the mechanical, optical and barrier properties. A multivariate image analysis and regression approach (MIA/MIR) is proposed in this work, and compared with three other techniques, for predicting both the local and global crystallinity variations of HDPE, LDPE, and PP polymer samples. Subtle variations in crystallinity introduced by changes in cooling rates were predicted within experimental errors. Predictions of crystallinity spatial distribution were also validated and found consistent with polymer processing knowledge.

* Gosselin R., Rodrigue D. and Duchesne C., On-line prediction of crystallinity spatial distribution across polymer films using NIR spectral imaging and chemometrics methods, *Can. J. Chem. Eng.*, 86, 869-878 (2008).

4.1 Introduction

The global molded plastics market was estimated to exceed 156 million tons in 2005, a large portion of which was taken up by common semi-crystalline polymers; the consumption of low-density polyethylene (LDPE), high-density polyethylene (HDPE) and polypropylene (PP) reaching 18, 29 and 40 million tons respectively (Murphy, 2003). These materials are very sensitive to the conditions in which they were produced, especially the geometry, pressure and cooling rates are known to influence their crystallinity, which in turn strongly affects their mechanical, optical and barrier properties (Rastogi et al., 1991; Apichartpattanasiri et al., 2001; Albano et al., 2003; Pantani et al., 2005; Shepherd et al., 2006; Dusunceli and Colak, 2008). Controlling the amount of crystallinity is thus essential when working with semi-crystalline polymers (Albano et al., 2003). However, theoretically establishing a relationship between the processing conditions and the final microstructure of a polymer is complex and time consuming (Pantani et al., 2005). On-line quality control of plastic film, either cast or blown, is typically limited to shape and clarity monitoring as well as detecting other irregularities like impurities, holes and scratches. While efficiently detecting macroscopic defects, these methods are somewhat ineffective at identifying smaller characteristics such as crystallinity variations, leading to film yields between 90-95% typically (Osborn and Jenkins, 1992). Rapid and non-intrusive tools for on-line crystallinity measurement would help improve quality control of thin polymer products.

Several methods are currently being used to determine polymer crystallinity, but most of these provide off-line measurements, are destructive, and provide single point measurements. They often rely on small samples sizes and require significant handling times. The most common off-line method is differential scanning calorimetry (DSC), which relies on the heat associated with melting of small samples, typically less than 20 mg. DSC has been widely used in polymer analysis (Batur et al., 1999; Saeki et al., 2003; Sato et al., 2003; Albano et al., 2003; Watanabe et al., 2006; Pelsoci, 2007). Numerous microscopy methods also exist. Of these, atomic force microscopy (AFM), scanning electron microscopy (SEM) and optical microscopy have all been used in micro-structural characterization of polymers (Kantz et al., 1972; Butler et al., 1995; Pantani et al., 2005).

These methods have a slow acquisition time and may require surfacing and/or vacuum conditions. Nuclear magnetic resonance (NMR) spectroscopy has also been used to obtain physical, chemical and structural information about polymer phases (Bergmann, 1981; Pelsoci, 2007). Application of this method is typically expensive, time consuming and destructive (Vailaya et al., 2001; Saeki et al., 2003). X-ray scattering methods are a family of non-destructive methods based on the scattered intensity of a beam hitting a sample and providing information about the average structure and composition (Brown et al., 1973; Polizzi et al., 1991; Albano et al., 2003; Pantani et al., 2005; Pelsoci, 2007; Briskman, 2007). As these methods depend on the incident and scattering angles, the polarization and wavelengths of the beams, they are not easily adapted to 2-D surface scans.

Various vibrational spectroscopic techniques, such as near-infrared (NIR), have also been proposed for estimating polymer crystallinity since they are rapid and non-intrusive, they require little or no sample preparation, and provide a wealth of information (Lachenal, 1995). However, most of the methods published in the past are based on single point measurements (i.e. using probes) and result in a separate calibration model for each type of polymer to be monitored. Estimating crystallinity over a larger area using these approaches require either moving the sample or the probe, which is difficult to perform on a high-speed production line, or using probe arrays, which generally result in coarse spatial coverage of the surface. However, fine spatial measurements are key in identifying local crystallinity gradients that may compromise the whole. The mechanical, optical and barrier properties of an entire sample may fall below specifications because of local discrepancies that may be overlooked by a coarse spatial analysis. Furthermore, rich spatial data would also be useful for precisely diagnosing operational problems in the production line, such as non-uniform cooling of the polymer products.

Nevertheless, NIR spectroscopy has been used to analyze the extent of crystallization (Lachenal, 1994; Stuart, 1996; Sato et al., 2003; Pantani et al., 2005; Pelsoci, 2007), the orientation (Buffeteau et al., 1995) and the density (Saeki et al., 2003; Watanabe et al., 2006) of semi-crystalline polymers. Since these spectroscopic techniques are based on absorbance, they are usually applied to films or thin samples (< 2 mm) in which case the

average degree of crystallinity is known to remain constant along the thickness (Pantani et al., 2005). While different methods have been reported, FTIR transmission spectroscopy remains the standard technique when studying polymer crystallinity when applied to microtomed slices (Pantani et al., 2005), blown films (Buffeteau et al., 1995) or thicker samples (0.5 - 2 mm) (Lachenal, 1995). Infrared reflection spectroscopy has also been used to study polymer crystallinity; Sato et al. (2003) notably collected diffusely reflected radiation, or average spectra, of plastic pellets by placing them on rotating cells. Work has also been reported on real-time crystallinity estimation by single-point laser Raman spectroscopy (Batur et al., 1999; Cherukupalli and Ogale, 2004).

The objective of this work is to develop a non-intrusive on-line sensor for monitoring spatio-temporal crystallinity variations across the surface of thin polymer materials. While single point NIR vibrational spectroscopy has proven useful in the study of polymer crystallinity, this work focuses on combining NIR imaging spectroscopy and multivariate image analysis and regression methods (MIA and MIR) to efficiently obtain the crystallinity spatial distribution of a sample. The proposed methodology is illustrated using thin polyolefin samples produced via compression molding under different cooling conditions. Non-isothermal cooling rates throughout the cooling circuit ensure local crystallinity variations within a single sample. A total of four different methods are compared in this study: PLS regression between the full averaged NIR spectra or 2^{nd} order spectra derivatives and crystallinity measurements (method 1 and 2), and MIA applied to the NIR spectra or 2^{nd} order spectra derivatives followed by regression (method 3 and 4). The results obtained for 3 polymers (HDPE, LDPE and PP) are incorporated into a single model, rather than producing a model for each, in order to highlight their capability to cope with complex datasets as may be encountered with polymer blends or composites. Local crystallinity estimates may be used in process monitoring, fault detection and diagnosis (i.e. identification of non-uniform cooling rates) or for predicting both the local or overall polymer properties (i.e. mechanical, optical, orientation, etc.) of thin samples (< 2 mm) for automatic process and quality control applications. As with any surface analysis, it may also be used with caution on thicker samples, monitoring operating conditions rather than inferring the properties of the product.

4.2 Experimental

4.2.1 Materials and processes

Polymers used in this study were commercial PP (Pro-fax PF-814, ρ = 0.902 g/cm^3; MFI = 3.0 g/10 min) supplied by Basell, HDPE (HDPE 3000, ρ = 0.965 g/cm^3; MFI = 8.8 g/10 min) by Pretromont and LDPE (Novapol LA-0219-A, ρ = 0.919 g/cm^3; MFI = 2.3 g/10 min) by Nova Chemicals.

Sample molding

A compression press was used to mould the polymers into thin films (100 × 40 × 1.5 mm). The batch process consists in molding 6 g of resin at 180°C under 3 tons of pressure for 10 minutes, long enough to remove any previous thermal history (Chew et al., 1989; Albano et al., 2003). The mould was then water cooled to release the sample. A total of 18 samples were produced, 6 for each polymer, using average cooling rates of 1, 2, 4, 5, 8 and 16°C/min. These changes in cooling rates will produce samples of different overall crystallinity. Moreover, the water-cooling circuit produced in-sample variability within the films. In other words, while each sample is characterized by its own average crystallinity, it also presents crystallinity patterns, or mosaics, across is surface.

Differential scanning calorimetry

DSC is a thermoanalytical technique measuring heat flow into a material as a function of temperature. The degree of crystallinity (ϕ) of a semi-crystalline polymer can be determined from the endothermic heat flow due to melting (ΔH_m), the exothermic heat flow due to crystallisation (ΔH_c) and the heat of fusion of the 100% crystalline phase (ΔH_∞):

$$\phi = \frac{\Delta H_m - \Delta H_c}{\Delta H_\infty} \tag{4.1}$$

Here, DSC thermograms were obtained using temperature ramps between 30-200°C at 10°C/min under a nitrogen atmosphere. The samples were approximately 7.5 mg and the tests were performed on a Perkin Elmer DSC 6. The fusion peak were found to be located at 110°C (LDPE), 138°C (HDPE) and 160°C (PP) and range respectively between 60-125°C, 90-155°C and 100-180°C. The heat of fusion of the crystalline phases (ΔH_∞) used was 290 J/g (LDPE), 293 J/g (HDPE) or 207 J/g (PP).

4.2.2 Image acquisition

The line-scan NIR imaging system used in this work is composed of a XenICs XEVA-USB-FPA camera (256 × 320 pixels) coupled to a Specim ImSpector N17E grating spectroscope (slit: 30 μm × 14.3 mm) sensing between 900-1700 nm. Each scanned line of the sample is captured on the CCD array as a spatial/spectral image (Figure 4.1) consisting of 256 × 320 pixels, providing a 0.5 mm/pixel spatial resolution and 2 nm/pixel spectral resolution. A hyperspectral image (**R**) of the entire sample is achieved by repeatedly displacing the sample perpendicularly to the spectroscope slit (along the *y*-axis). Each hyperspectral image consists of 500 juxtaposed line-scans, at 0.25 mm intervals.

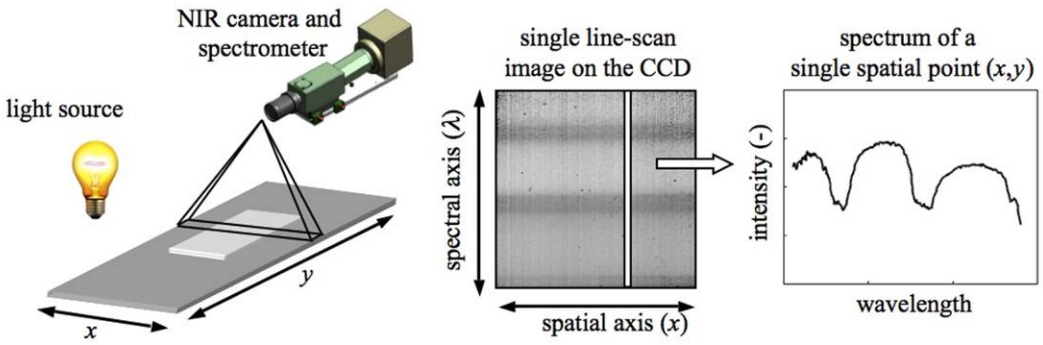

Figure 4.1. Line-scan NIR imaging system used to collect spatial/spectral reflectance images.

Diffused light from a tungsten-halogen lamp was used as the light source. The raw signals of each spectral-spatial image (**R**) were transformed into reflectance units using a flat black image (**B**) and a true white image (**W**):

$$i_{xy\lambda} = \frac{r_{xy\lambda} - b_{x\lambda}}{w_{x\lambda} - b_{x\lambda}} \tag{4.2}$$

where $r_{xy\lambda}$ is an element of the raw signal image, $b_{x\lambda}$ and $w_{x\lambda}$ are elements of the black and white calibration images, and $i_{xy\lambda}$ is the standardized reflectance image (**I**). The flat black and true white images were obtained by collecting a spatial-spectral image leaving the lens cap in place (i.e. dark current), and by imaging a 99% reflectance standard sheet (Gigahertz-Optik), respectively. This simple calibration scheme was used since it successfully removed most the linear discrepancies encountered in the NIR images collected in this work. All images were captured using a single imaging system, under controlled lighting conditions, and over a short period of time. More advanced calibration techniques should be considered for long-term, industrial applications (Geladi et al., 2004; Burger and Geladi, 2005; Liu et al., 2007). Data acquisition was performed using a LabView 8.0 interface (National Instruments) and data analysis was performed using custom scripts developed in the Matlab R14 (MathWorks) environment.

4.3 Methodology

Predicting the polymer crystallinity spatial distribution requires building a regression model (i.e. correlations) between the information contained within NIR spectral images of the samples (**X**) and analytical crystallinity measurements (**y**) obtained using DSC. A total of four different methods were tested and compared in this work, differing mainly in the way NIR spectra are processed before applying appropriate regression techniques. Conventional NIR calibration methods involving averaged spectra (method 1) or 2nd order derivatives (method 2) and PLS regression are first tested, followed by a multivariate image regression (MIR) approach involving either the raw spectra (method 3) or its 2nd order derivatives

(method 4). All spectra were initially submitted to an initial smoothing step using a symmetric 5-point smoothing window in order to stabilize the 2nd derivative computations.

4.3.1 Collection of training dataset

Initially, each of the 18 samples was fully scanned and calibrated into reflectance images $\underline{\mathbf{I}}_k$ according to Equation 4.2, where k is the sample number (k = 1, 2, …18). A training data set was obtained by selecting a smaller region of interest (2.5 × 5 mm or 10 × 10 pixels) for each polymer sample as shown in Figure 4.2. These regions correspond to a multivariate sub-image $\underline{\tilde{\mathbf{I}}}_k$ of size (10 × 10 × λ), where λ is the number of spectral channels (λ = 256 in this work). These arrays were then unfolded into matrices $\tilde{\mathbf{I}}_k$ of size (100 × λ) by simply storing each spectrum row-wise. The $\tilde{\mathbf{I}}_k$ matrices are the common input of all four methods described in this section. Finally, each region of interest corresponding to the $\tilde{\mathbf{I}}_k$ sub-images were cut out of the samples and sent for DSC analysis. The resulting crystallinity measurements (y_k) were stored in a response matrix **y** (18 × 1).

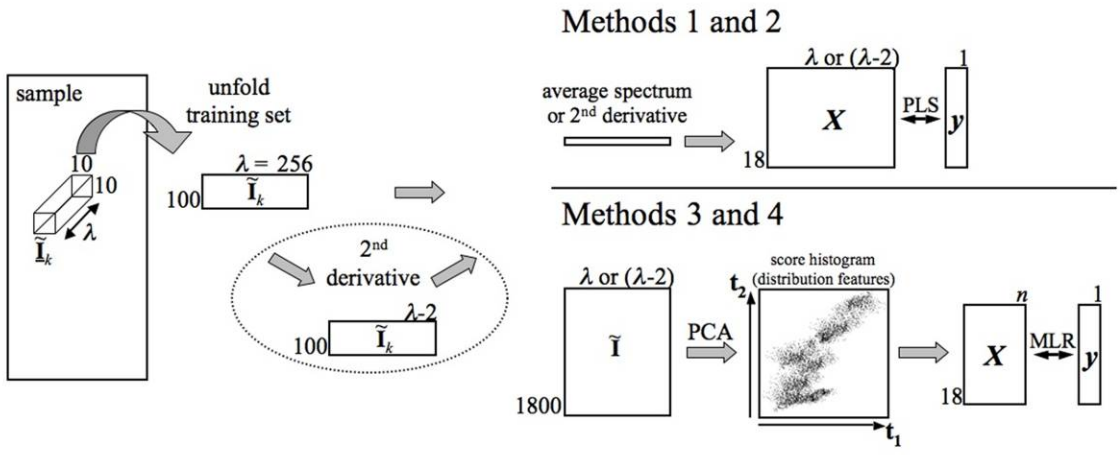

Figure 4.2. Schematic representation of the 4 methods: averaged spectra (method 1), 2nd order derivatives (method 2), MIR of the spectra (method 3) or MIR of 2nd order derivatives (method 4).

4.3.2 PLS models based on average spectra or 2nd order derivatives

Method 1: averaged NIR spectra

The first method simply consists of averaging all available reflectance spectrum obtained from the region of interest of a sample (i.e. column-wise averaging of each $\tilde{\mathbf{I}}_k$ matrix) and storing the averaged spectra of each sample in a regressor matrix **X** (18 × λ), as shown in Figure 4.2. A latent variable PLS regression model is then built between the averaged spectra matrix **X** and the corresponding crystallinity measurements **Y** as:

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{4.3}$$

$$\mathbf{Y} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{q}_a^{\mathrm{T}} + \mathbf{F} = \mathbf{T}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \tag{4.4}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^{*} \tag{4.5}$$

where the **P** and **Q** matrices contain the loading vectors that best represent the **X** and **Y** spaces respectively, whereas **W*** contain the loading vectors defining the common latent variable space **T** relating **X** and **Y**. The **E** and **F** matrices contain the residuals. The number of PLS components or latent variables (*A*) was selected using a standard leave-one-out cross-validation procedure. PLS was selected in this case since the columns of **X** are highly collinear.

Method 2: 2nd order derivative of the NIR spectra

Method 2 uses the 2nd order derivative of the spectra $\tilde{\mathbf{I}}_k$ rather than the spectra themselves. The spectra derivatives were also averaged over the region of interest and then collected in regressor matrix **X** (18 × (λ–2)) for each sample before building a PLS regression model with **y** (Figure 4.2). Taking second order derivatives is a commonly used pretreatment applied to NIR spectra (Chau et al., 1990). This offers a distinct advantage when sharp

absorption bands are present in the spectrum (Whitbeck, 1981). It is preferred over the 1st derivative because it does not shift the peaks, allowing better interpretability. While the 2nd order derivative emphasizes spectral transitions, it remains insensitive to systematic shifts in spectra intensity. This can be seen with the 3 spectral bands emphasized in the original spectrum (centered around 1100nm, 1300nm and 1600nm) (Figure 4.3). Numerical approximation of the 2nd order derivative were used (Gerald and Wheatley, 1994) after smoothing of the line-scan in the spectral direction using a window of 5 pixels in order to reduce the effect of bad pixels (Savitzky and Golay, 1964). Numerical differentiation leads to the loss of two columns (i.e. spectral channels) in matrix **X** compared to using the spectra directly.

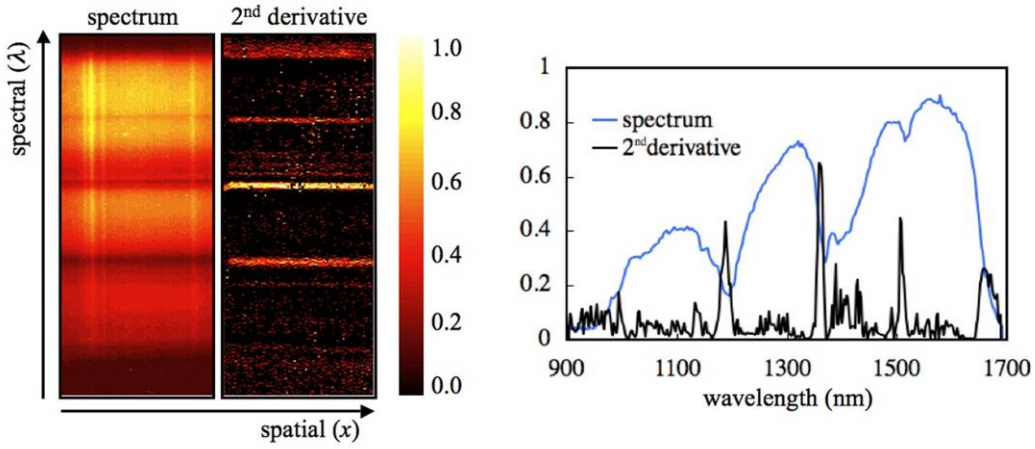

Figure 4.3. Spectrum vs. 2nd derivative: 2-D line-scan of a HDPE sample (left) and 1-D spectra of a single spatial position (right). Dimensionless scale used.

4.3.3 Multivariate image regression applied to spectra or 2nd order derivatives

Multivariate image regression (MIR) consists of a family of latent variable techniques used for regressing quality or response variables onto features extracted from a set of digital images. Image regression problems can be formulated in several ways, depending on the nature of the features extracted from the images, ranging from simple statistics computed for each spectral channel (i.e. mean, variance, etc.) to distribution features. The reader is referred to Esbensen et al. (1992) and Yu and MacGregor (2003) for more details on MIR

problem formulation and methods. In this work, distribution features were extracted from NIR spectral images. These were obtained through a decomposition of spectral image data cubes (i.e. reflectance spectra - method 3, or 2nd order derivatives - method 4, see Figure 4.2) using multi-way principal component analysis (MPCA), which is the first step of a method known as multivariate image analysis (MIA) (Geladi and Grahn, 1996).

Multivariate image analysis

Originally introduced by Esbensen and Geladi (1989), it has been widely used in fields ranging from flame analysis (Yu and MacGregor, 2004; Szatvanyi et al., 2006) to flotation froth (Liu et al., 2005), snack food (Yu and MacGregor, 2003; Yu et al., 2003) as well as softwood lumber grading (Bharati et al., 2003). A complete overview of MIA can be found in Geladi and Grahn (1996). MIA relies on the basis that local intensity variations can be extracted by classifying each image pixel according to its spectral characteristics regardless of its spatial location within the image. Coupled to regression techniques, MIA can be used to extract the relevant features from digital images in order to build models between the sample quality and the visual features of an image.

The unfolded spectra matrices $\tilde{\mathbf{I}}_k$ corresponding to the regions of interest selected for each polymer sample are first collected together in a single matrix $\tilde{\mathbf{I}}$ of size (1800 × λ) (18 samples × 100 spectrum/sample). Principal component analysis is then used to decompose the image information into a set of A orthogonal loadings vectors $\mathbf{p}_a$ (1 × λ) and score vectors $\mathbf{t}_a$ (1800 × 1) as shown below and schematically in Figure 4.4:

$$\tilde{\mathbf{I}} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} \qquad (4.6)$$

where $\mathbf{E}$ (1800 × λ) contains the projection residuals (non zero when A < λ). The loadings vectors ($\mathbf{p}_a$) are usually obtained using singular value decomposition (SVD) of a kernel matrix ($\tilde{\mathbf{I}}^{\mathrm{T}}\tilde{\mathbf{I}}$) (Geladi and Grahn, 1996) of much smaller dimensions (i.e. 256 × 256). The

score vectors are computed using $\mathbf{t}_a = \tilde{\mathbf{I}}\mathbf{p}_a$. The first score vector $\mathbf{t_1}$ is the linear combination of the 256 spectral channels (defined by the loadings vectors $\mathbf{p_1}$) capturing the largest possible variance within the spectral matrix $\tilde{\mathbf{I}}$, while the second score vector $\mathbf{t_2}$ represents the second largest source of variance, and so on. The score vectors can therefore be viewed as a multivariate summary of each spectrum.

A small number of components (*A*) are often found sufficient in practice to extract most of the relevant information within multivariate image data. The few score vectors can therefore be used as representative distribution features of the multivariate image. These are typically displayed using scatter plots of score vectors as shown in Figure 4.4 ($\mathbf{t_1}$-$\mathbf{t_2}$ score plot is illustrated) or as 2-D density histograms (Geladi and Grahn, 1996; Yu and MacGregor, 2003; Szatvanyi et al., 2006).

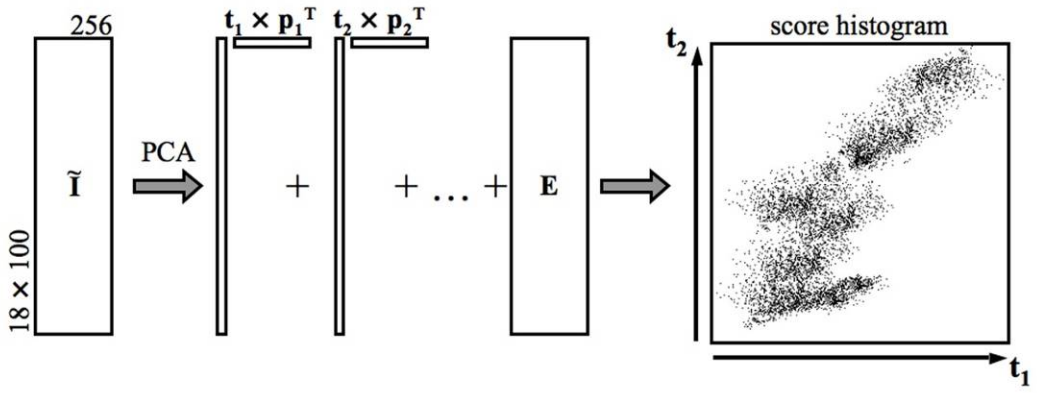

Figure 4.4. MPCA decomposition of a hyperspectral image.

Multivariate image regression

Formulating a regression problem between such score scatter plot (or 2-D density histograms) and response variables was investigated by Yu and MacGregor (2003). It involves extracting a certain number of features (n) from the score plots (or histograms) obtained from each of the K images, collecting them in a regressor matrix **X** (K × n), and building a regression model with the response variable of interest (i.e. crystallinity) **y** (K × 1), as shown in Figure 4.2.

The formulation used specifically for relating NIR spectral images to polymer crystallinity measurements is based on observing the clustering pattern of the NIR spectra of the three polymer types shown in Figure 4.5A.

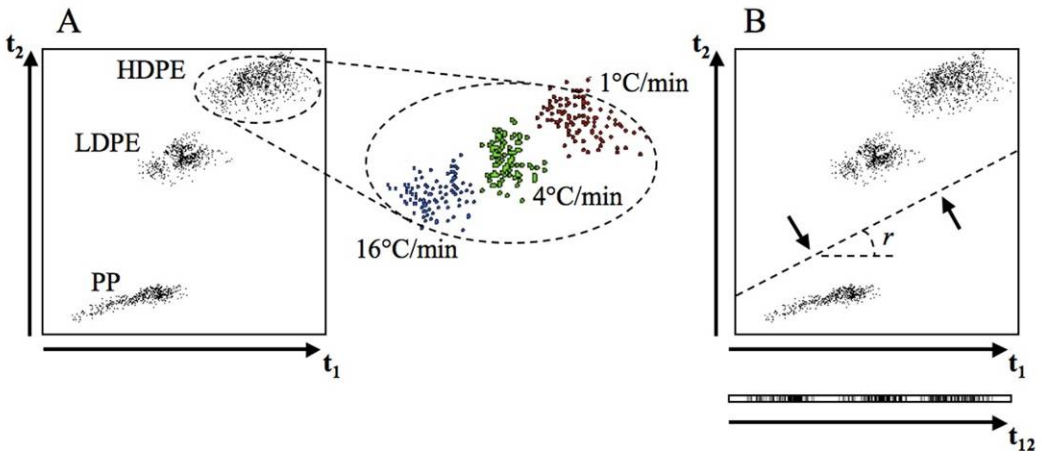

Figure 4.5. Score histogram allows both polymer and cooling rate identification (A). Projection of the dataset onto a single vector ($\mathbf{t_{12}}$) according to an angle (r) allows dimensional reduction (B).

This $\mathbf{t_1}$-$\mathbf{t_2}$ scatter plot was obtained by PCA decomposition of the spectral matrix $\tilde{\mathbf{I}}$ (1800 × 256). The two first score vectors explained 95.8% and 3.2% of the variance of $\tilde{\mathbf{I}}$ for method 3, and 77.7% and 11.1% for method 4, respectively. The spectra corresponding to each of the three polymer species appear as very distinct cluster as expected; NIR spectroscopy is often used for polymer identification. Furthermore, the spectral data corresponding to each polymer type also cluster according to cooling rates with a clear spatial orientation (Figure 4.5A, HDPE cluster zoom-in). Furthermore, this zoom-in illustrates that samples cooled at 4°C/min appear equidistant from samples cooled at either 1 or 16°C/min, indicating that these samples followed asymptotical crystallization rates. While the experimental equipment used in this study could not allow for faster or slower cooling rates than 1 to 16°C/min, it is expected that similar results could be obtained for a broader range of cooling rates, as long as the relationship between the spectral information and crystallinity remain unchanged. Thus, a crystallinity related vector $\mathbf{t_{12}}$ could therefore be obtained by projecting the spectral data shown in Figure 4.5A onto a linear combination

of the first two score vectors (Figure 4.5B), and then building a simple linear regression model between $\mathbf{t_{12}}$ and **y** using ordinary least squares:

$$\mathbf{t_{12}} = r\mathbf{t_1} + (1-r)\mathbf{t_2} \quad (4.7)$$

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{t_{12}} + \varepsilon \quad (4.8)$$

The linear combination or, alternatively, the angle (r) of the $\mathbf{t_{12}}$ vector should be selected in such a way to maximize the correlation between $\mathbf{t_{12}}$ and **y**. A similar approach was discussed in Yu and MacGregor (2003) for score density histogram segmentation as one possible formulation of multivariate image regression problems.

4.4 Results and discussion

4.4.1 Prediction model building using the training dataset

Summary statistics for the PLS models built on the training set are shown in Table 4.1 for methods 1 and 2. The number of PLS components (*A*) of each model were selected by a standard leave-one-out cross-validation procedure using the SIMCA-P+ V10 software (Umetrics, Inc.). Three cumulative multiple correlation coefficients were computed for quantifying the model predictive ability: $R^2_{X,cum}$ is the cumulative percentage of the total variance in **X** used to explain **y**, $R^2_{y,cum}$ gives the cumulative percentage of the total variance of **y** explained by the model, and Q^2_{cum} is the cumulative percentage of the total variance of **y** that can be predicted by the models using the cross-validation procedure. A total of 2 components for each model were selected through cross-validation. Model 2 based on 2nd order derivatives seems to provide a better fit of the data compared to using the averaged spectra (i.e. method 1).

The results obtained for methods 3 and 4 are presented in Table 4.2 and Figure 4.6. The linear combination of the first two score vectors obtained using MIA ($\mathbf{t_{12}}$) was selected to maximize the predictive ability (i.e. Q^2_{cum}) of the models as shown in Figure 4.6. A *r*-ratio of 0.8 provides the best results for method 3 ($R^2_{y,cum}$ = 0.991 and Q^2_{cum} = 0.923). For

method 4, a *r*-ratio of 0.1 was found to provide the best MIR model using 2nd order derivatives ($R^2_{y,cum}$ = 0.861 and Q^2_{cum} = 0.252). The lower predictive power (Q^2_{cum}) of this model may be explained by the fact that a low *r*-ratio implies that crystallinity is hardly related to the first score ($\mathbf{t_1}$) and much more to the second ($\mathbf{t_2}$). Since $\mathbf{t_1}$ represents the largest source of variation in the data (77.7%), it appears that the bulk of the information contained in the 2nd order derivatives hardly correlate to crystallinity. Additional MIA score vectors (i.e. $\mathbf{t_3}$, $\mathbf{t_4}$ and $\mathbf{t_5}$) were investigated for model 4, but these were found to be uncorrelated to crystallinity.

Table 4.1 Summary statistics of the predictive ability of methods 1 and 2.

Method	A	$R^2_{X,cum}$ (%)	$R^2_{y,cum}$ (%)	Q^2_{cum} (%)
1	1	0.907	0.640	0.605
	2	0.967	0.756	0.681
2	1	0.631	0.512	0.472
	2	0.764	0.957	0.932

Table 4.2 Model parameters and predictive ability for method 3 and 4.

Method	r (-)	$\hat{\beta}_0$	$\hat{\beta}_1$	$R^2_{y,cum}$ (%)	Q^2_{cum} (%)
3	0.8	-64.41	0.90	0.991	0.923
4	0.1	-8.13	0.40	0.861	0.252

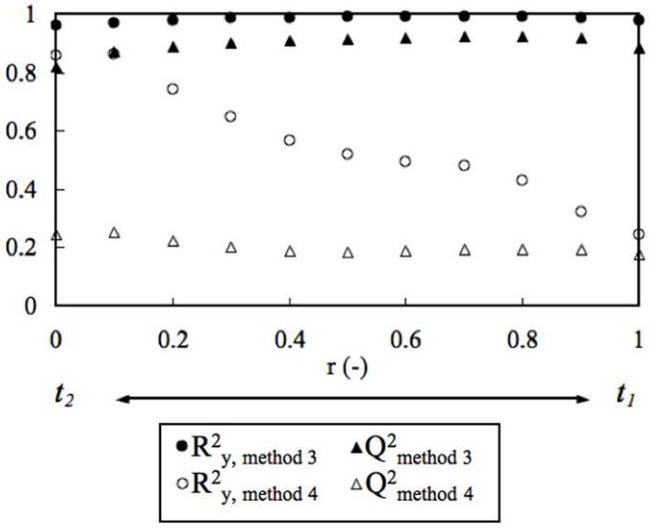

Figure 4.6. Selection of the projection angle r yielding the crystallinity related vector $\mathbf{t_{12}}$ for methods 3 and 4.

The performance of the four methods in fit are compared in Figure 4.7. This figure shows the experimentally measured crystallinity by DSC compared to model predictions obtained by cross-validation. The original reflectance spectrum (method 1) provides the largest errors (root mean square error of estimation RMSEE = 9.31) while the 2nd derivative (method 2) provides much better results (RMSEE = 3.88). The MIR methods applied to the reflectance spectra (method 3) and to the 2nd order derivatives (method 4) provided RMSEE values of 1.78 and 6.83, respectively. It is therefore clear that performing MIA on the image data in order to extract the relevant features does improve predictive ability when applied to the spectra directly, but is detrimental when using 2nd order derivatives.

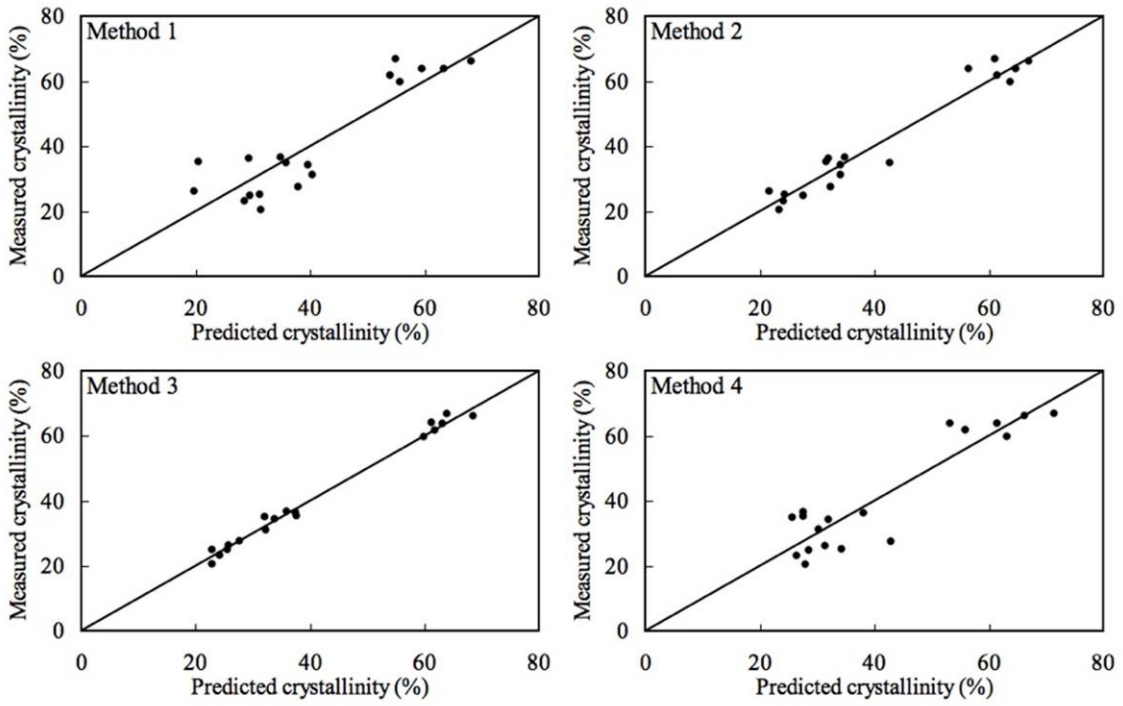

Figure 4.7. Measured crystallinity vs. cross-validated predictions obtained for the training dataset using the 4 methods.

Generally speaking, this may be explained by the degree of correlation between the image data (spectrum or 2nd derivative) and the quality parameter to be predicted (crystallinity). In method 3, the crystallinity is predominantly correlated to $\mathbf{t_1}$, which represents 95.8% of the variance of $\tilde{\mathbf{I}}$ whereas method 4 essentially correlates crystallinity to $\mathbf{t_2}$ and only 11.1% of the variance in $\tilde{\mathbf{I}}$. In other words, method 3 relies on nearly all the spectral data while

method 4 only relies on a small fraction since most of the spectral data contained in the 2nd derivatives is irrelevant to the prediction of polymer crystallinity. This characteristic of the 2nd derivative may explain the relative inability of methods 2 and 4 to explain fine surface variations.

4.4.2 Model validation over the entire samples

The preceding discussion (Section 4.4.1) was centered on the predictive capabilities of the 4 methods within the training sets themselves. In the present section they will be used to predict the remaining areas of the samples that were not used to construct the models. As previously mentioned, the cooling system design is known to produce localized gradients in the cooling rates that should produce crystallinity patterns throughout the surface of the samples. If successful, it is expected that these methods should distinguish such patterns.

Using the four models developed on the training set, one can predict crystallinity for any point on the surface of any sample (i.e. for every pixel or groups of pixels of spectral images $\underline{\tilde{\mathbf{I}}}_k$). One sample of each polymer type was retained for this purpose. Figure 4.8 shows the predicted crystallinity distribution for one such sample (HDPE: 4°C/min). To validate these predictions four regions of interest dispersed throughout the samples were removed, analyzed using DSC in triplicate, and compared to the predicted values. The location of these regions is also shown in Figure 4.8.

As illustrated in Figure 4.8, the predicted crystallinity differs greatly according to the method used. The prediction made by method 1 seems to show an area of slightly lower crystallinity in the lower right section of the sample. Method 2 predicts a vertical gradient while method 4 predicts a homogeneous surface. Last, method 3 predicts a complex pattern in which the right hand side is significantly less crystalline. Comparing these predictions to measured crystallinity values (see DSC measurements on top of Figure 4.8), it can be seen that method 1 slightly overestimates crystallinity, and that both models based on 2nd order derivative (methods 2 and 4) predict an average value throughout the sample. Only method

3 based on MIR of the reflectance spectra provides a reasonable map of crystallinity distribution.

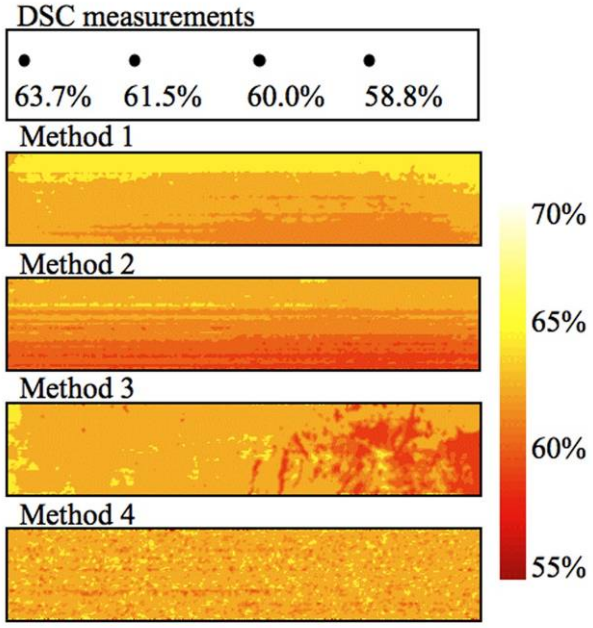

Figure 4.8. Predicted crystallinity distribution for the HDPE sample cooled at 4°C/min. Values are given in % crystallinity. Sample size: 100 × 40 mm.

Together, Table 4.3 and Figure 4.9 illustrate the complete results for the 3 tested samples (1 for each polymer type) including the predicted crystallinity and the absolute mean prediction error. DSC standard deviations are based on triplicate tests. Model predictions were averaged over the region sent for DSC analyses (equivalent surface area of 20×20 pixels since more material was necessary for triplicates). These results show that the averaged experimental error associated with DSC measurements is 0.7%, independently of the polymer tested. Method 3 provided the lowest mean prediction errors (0.8%), falling practically within the DSC experimental error (i.e. method 3 yields a statistically adequate model). Finally, method 3 and 4 (MIR models) provide more consistent prediction errors for all polymer species since the angle (r) of vector $\mathbf{t_{12}}$ was selected to maximize correlation with crystallinity, irrespectively of polymer type. This may help cope with more complex materials such as polymer blends or composites.

Table 4.3. Experimental (DSC) and model prediction errors for all three polymer species. Experimental errors are reported as one standard deviation. Prediction errors are reported in absolute values (i.e. % crystallinity).

	HDPE	PP	LDPE	Mean
DSC	0.6	0.7	0.7	0.7
Method 1	2.6	1.8	1.2	1.8
Method 2	1.7	1.0	0.9	1.1
Method 3	0.8	0.9	0.7	0.8
Method 4	1.5	1.3	1.6	1.5

Since prediction results for both method 2 and 3 on the training set (i.e. see Figure 4.7) containing all three polymer species are quite similar, although method 3 has a slightly better fit (i.e. less scattering around 45° line), it may seem surprising that both methods predict such a different crystallinity distribution across the surface a particular polymer type as shown in Figure 4.8. This may be explained by recognizing that the variance of the cystallinity vector (**y**) has two main sources: (1) each polymer type has a different mean crystallinity, and (2) changes introduced by the cooling rate (i.e. scattering around each the mean crystallinity of each polymer type). Method 2 and 3 differ mainly in the source of variance each of them explains. Indeed, by enhancing spectral transitions (i.e. peaks in the spectra), the method based on 2^{nd} order derivatives (method 2) seems better at predicting the mean crystallinity of each polymer specie or, alternatively, at identifying polymer type, than predicting more subtle variations in crystallinity due to changes in cooling rates, what we are interested in from a process operation perspective.

To confirm this, the crystallinity measurements within the training set (**y**) were recalculated in deviation from the mean crystallinity values of each polymer type: 63.8% (HDPE), 35.9% (PP) and 24.7% (LDPE). Removing the mean crystallinity of each polymer type should enhance crystallinity variations caused by changes in the cooling rates. Prediction results shown in Figure 4.10 were obtained after building new regression models for method 2 and 3 using the new centered crystallinity vector ($\mathbf{y}^C$). In accordance with previous results, method 3 explains a greater amount of the local crystallinity variations ($R^2_{y,cum}$ = 0.757) compared to method 2 ($R^2_{y,cum}$ = 0.542). This confirms the higher predictive power of the MIR model based on the original reflectance spectrum. Note that

this method could also be used for polymer type identification by segmenting vector $\mathbf{t_{12}}$ according to the mean crystallinity of each polymer type.

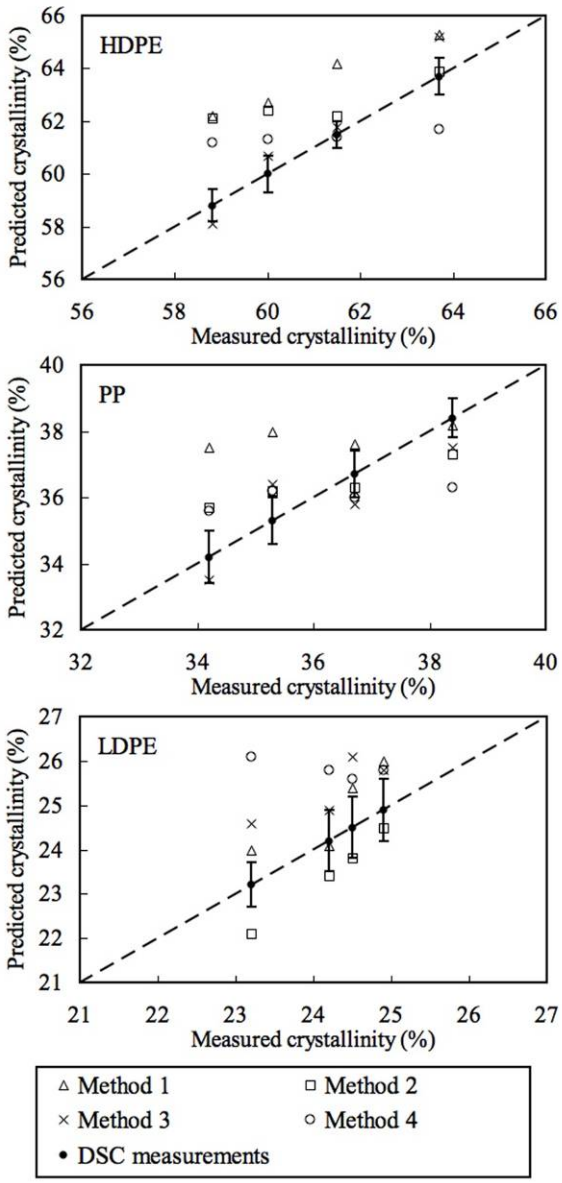

Figure 4.9. Crystallinity predictions obtained by the 4 methods for the three polymer species (cooling rate = 4 ℃/min) compared to DSC values (including standard deviation). Values are given in % crystallinity.

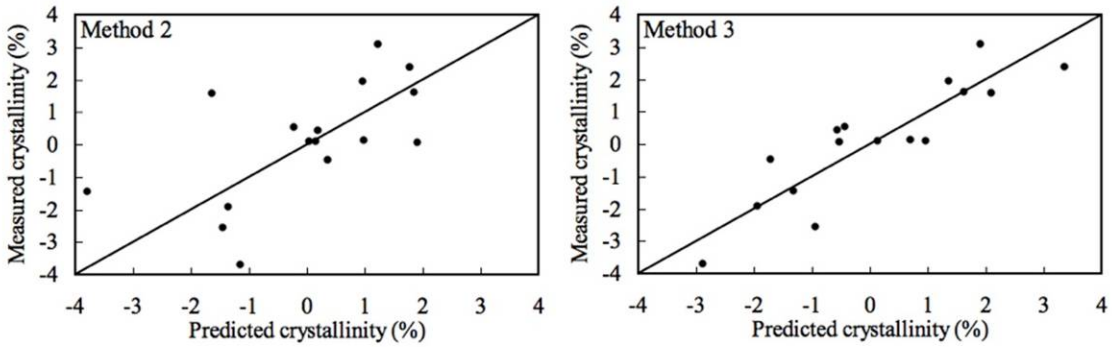

Figure 4.10. Measured crystallinity vs. predicted centered crystallinity (y^C) values for methods 2 and 3 using the data from the training set.

Finally, Figure 4.11 illustrates the predictions obtained for the LDPE and PP samples (both cooled at a rate of 4°C/min) using method 2 and 3. The general patterns and conclusions concur with those previously discussed with the HDPE sample. The 2nd derivative efficiently predicts the average crystallinity of the sample but remains insensitive to local variations. On the other hand, method 3 adequately predicts both local and global variations in crystallinity.

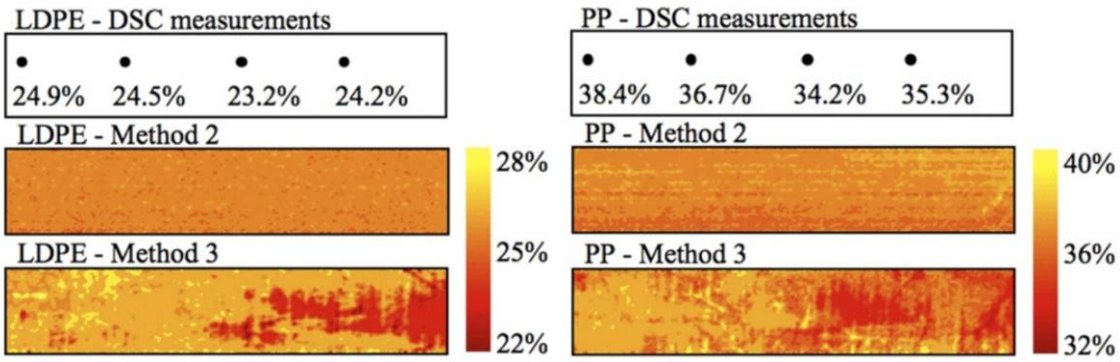

Figure 4.11. Predicted crystallinity distribution for LDPE and a PP sample (4°C/min) using both method 2 and 3. Values are given in % crystallinity. Samples: 100 × 40 mm.

4.5 Conclusion

The objective of this work was to develop a rapid and non-intrusive on-line NIR imaging sensor for monitoring spatio-temporal variations in crystallinity across the surface of thin polymer materials. Although NIR spectroscopy has already been investigated in the past for

estimating polymer crystallinity, most methods were based on probes, hence providing measurements for a single point (or line) on a sample. Since most of the overall end-user properties of thin polymer materials, such as mechanical, optical and barrier properties, depend upon the local microstructure (i.e. weakest point on the sample), fine spatial measurements are key in identifying local crystallinity gradients that may compromise the quality of whole. Furthermore, rich spatial data would also be useful for precisely diagnosing operational problems in the production line, such as non-uniform cooling of the polymer products.

Thin samples of LDPE, HDPE and PP were produced via compression molding and cooled at six different rates in order to induce a wide range of crystallinity. Each sample was then imaged using an NIR imaging spectroscope, and a certain number of regions of interest were selected for each sample, cut-off, and sent for analytical measurement of crystallinity using Differential Scanning Calorimetry (DSC). Four methods were tested and compared for extracting features from the NIR images that are correlated to polymer crystallinity obtained by DSC: 1) a conventional NIR calibration methods involving averaged spectra and PLS regression with crystallinity; this approach was used by Sato et al. (2003) for estimating crystallinity of polymer pellets, 2) a similar approach but using 2^{nd} order spectra derivatives for calibration against crystallinity (i.e. a standard pre-treatment used in spectroscopy for enhancing the peaks), 3) a multivariate image analysis and regression approach (MIA/MIR) using either the raw reflectance spectra, or 4) its 2^{nd} order derivatives.

The results show that the models obtained using method 2 and 3 provide the best fit of crystallinity measurements when predicting over a wide range of crystallinity, such as the one obtained using the three polymer species (LDPE, HDPE, and PP). However, only method 3 based on MIA/MIR analysis of the original NIR spectra was successful at predicting the more subtle variations in crystallinity introduced by changes in cooling rates, with prediction errors falling practically within DSC experimental errors. It was also shown that this method provides predictions of crystallinity spatial distribution that are consistent with polymer processing knowledge, and partially validated using additional DSC

measurements. Finally, such predictive power was also consistently high for all three polymer species using a single model. This feature of the proposed approach shows an interesting potential for application to more complex materials such as polymer blends or composites.

4.6 References

Albano C., Papa J., Gonzalez E., Navarro O. and Gonzalez R., Temperature and crystallinity profiles in polyolefines isothermal and non-isothermal solidification Processes, *Eur. Polym. J.*, 39, 1215-1222 (2003).

Apichartpattanasiri S., Hay J.N. and Kukureka S.N., A study of the tribological behaviour of polyamide 66 with varying injection-moulding parameters, *Wear*, 251, 1557-1566 (2001).

Batur C., Vhora M.H., Cakmak M. and Serhatkulu T., On-line crystallinity measurement using laser Raman spectrometer and neural network, *ISA T.*, 38, 139-148 (1999).

Bergmann K., Determination of Polymer Crystallinity from proton solid-echo NMR measurements, *Polym. Bull.*, 5, 355-360 (1981).

Bharati M., MacGregor J.F. and Tropper W., Softwood lumber grading through on-line multivariate image analysis, *Ind. Eng. Chem. Res.*, 42(21), 5345-5353 (2003).

Buffeteau T., Desbat B. and Bokobza L., The use of near-infra-red spectroscopy coupled to the polarization modulation technique to investigate molecular orientation in uniaxially stretched polymers*, Polymer*, 36(22), 4339-4343 (1995).

Butler J.H., Joy D.C., Bradley G.F. and Krause S.J., Low-voltage scanning electron microscopy of polymers, *Polymer*, 36(9), 1781-1790 (1995).

Burger J. and Geladi P., Hyperspectral image regression part I: calibration and correction, *J. Chemometr.*, 19, 355-363 (2005).

Briskman B.A., Be careful when using X-ray exposure analysis for polymers, *Nucl. Instrum. Meth. B*., 265, 276-280 (2007).

Brown D.S., Fulcher K.U. and Wetton R.E., Application of Small Angle X-ray Scattering to semi-Crystalline Polymers: 1. Experimental considerations and analysis of data, *Polymer*, 14, 379-383 (1973).

Chau F., Liang Y., Gao J. and Shao X., *Chemometrics, from Basics to Wavelet Transmform*, Wiley: United Kingdom (1990).

Cherukupalli S.S. and Ogale A.A., Online Measurements of Crystallinity using Raman Spectroscopy during Blown Film Extrusion of a Linear low-density Polyethylene, *Polym. Eng. Sci.*, 44(8), 1484-1490 (2004).

Chew S., Griffiths J.R. and Stachurski Z.H., The Crystallization Kinetics of Polyethylene under Isothermal and non-Isothermal Conditions, *Polymer*, 30, 874-881 (1989).

Dusunceli N. and Colak O.U., Modelling Effects of Degree of Crystallinity on Mechanical Behavior of Semicrystalline Polymers, *Int. J. Plasticity*, 24(7), 1224-1242 (2008).

Esbensen K.H., Geladi P. and Grahn H., Strategies for Multivariate Image Regression, *Chemom. Intell. Lab. Syst.*, 14(1-3), 357-374 (1992).

Esbensen K.H. and Geladi P., Strategy of Multivariate Image Analysis (MIA), *Chemometr. Intell. Lab.*, 7, 67-86 (1989).

Geladi P., Burger J. and Lestander T., Hypespectral Imaging: Calibration Problems and Solutions, *Chemometr. Intell. Lab.*, 72, 209-217 (2004).

Geladi P. and Grahn H., *Multivariate Image Analysis*, Wiley: United Kingdom (1996).

Gerald C.F. and Wheatley P.O., *Applied Numerical Analysis*, Addison-Wesley Publishing Company: New-York, USA (1994).

Kantz M.R., Newman J.R. and Stigale F.H., The Skin-core Morphology and Structure–property Relationships in Injection-moulded Polypropylene, *J. Appl. Polym. Sci.*, 16, 1249-1260 (1972).

Lachenal G., Dispersive and Fourier Transform near-Infrared Spectroscopy of Polymeric Materials, *Vib. Spectrosc.*, 9, 93-100 (1995).

Liu J.J., MacGregor J.F., Duchesne C. and Bartolacci G., Flotation Froth Monitoring using Multiresolutional Multivariate Image Analysis, *Miner. Eng.*, 18, 65-76 (2005).

Liu Z., Yu H. and MacGregor J.F., Standardization of line-scan NIR Imaging Systems, *J. Chemometr.*, 21, 88-95 (2007).

Murphy J., *Additives for Plastics Handbook*, Elsevier Advanced Technology: Great Britain (2003).

Osborn K.R. and Jenkins W.A., *Plastic Films Technology and Packaging Applications*, Technomic Publishing: Switzerland (1992).

Pantani R., Coccorullo I., Speranza V. and Titomanlio G., Modeling of Morphology Evolution in the Injection Molding Process of Thermoplastic Polymers, *Prog. Polym. Sci.*, 30, 1185-1222 (2005).

Pelsoci T., *ATP-Funded Green Process Technologies: Improving U.S. Industrial Competitiveness with Applications in Packaging, Metals Recycling, Energy, and Water Treatment*, National Institute of Standards and Technology: U.S. Department of Commerce (2007).

Polizzi S., Fagherazzi G., Benedetti A. and Battagliarin M., Crystallinity of Polymers by X-ray Diffraction: a new Fitting Approach, *Eur. Polym. J.*, 27(1), 85-87 (1991).

Rastogi S., Newman M. and Keller A., Pressure-induced Amorphization and Disordering on Cooling in a Crystalline Polymer, *Nature*, 353, 55-57 (1991).

Saeki K., Tanabe K., Matsumoto T., Uesaka H., Amano T. and Funatsu K., Prediction of Polyethylene Density by near-Infrared Spectroscopy Combined with Neural Network Analysis, *J. Comput. Chem. Jpn.*, 2(1), 33-40 (2003).

Sato H., Shimoyama M., Kamiya T., Amari T., Sasic S., Ninomiya T., Siesler H.W. and Ozaki Y., Near Infrared Spectra of Pellets and thin Films of High-density, Low-density and Linear Low-density Polyethylenes and Prediction of their Physical Properties by Multivariate Data Analysis, *J. Near Infrared Spec.*, 11, 309-321 (2003).

Savitzky A. and Golay M., Smoothing and Differentiation of Data by Simplified Least Squares Procedures, *Anal. Chem.*, 36: 1627-1639 (1964).

Shepherd J.E., McDowell D.L. and Jacob K.I., Modeling Morphology Evolution and Mechanical Behavior During Thermo-Mechanical Processing of semi-Crystalline Polymers, *J. Mech. Phys. Solids*, 54, 467-489 (2006).

Stuart B.H., Polymer Crystallinity Studied using Raman Spectroscopy, *Vib. Spectrosc.*, 10, 79-87 (1996).

Szatvanyi G., Duchesne C. and Bartolacci G., Multivariate Image Analysis of Flames for Product Quality and Combustion Control in Rotary Kilns, *J. Am. Chem. Soc.*, 5(13), 4706-4715 (2006).

Vailaya A., Wang T., Chen Y., Huffman M., Quantitative Analysis of Dimethyl Titanocene by Iodometric Titration, gas Chromatography and NMR, *J. Pharmaceut. Biomed.*, 25, 577-588 (2001).

Watanabe S., J. Dybal, K. Tashiro, Y. Ozaki, "A near-Infrared study of Thermally Induced Structural changes in Polyethylene Crystal," Polymer, 47 2010-2017 (2006).

Whitbeck M.R., Second Derivative Infrared Spectroscopy, *Appl. Spectrosc.*, 35(1), 93-95 (1981).

Yu H., MacGregor J.F., Haarsma G. and Bourg W., Digital Imaging for Online Monitoring and Control of Industrial Snack Food Processes, *Ind. Eng. Chem. Res.*, 42, 3036-3044 (2003).

Yu H. and MacGregor J.F., Multivariate Image Analysis and Regression for Prediction of Coating Content and Distribution in the Production of Snack Foods, *Chemometr. Intell. Lab.*, 67, 125-144 (2003).

Yu H. and MacGregor J.F., Monitoring Flames in an Industrial Boiler Using Multivariate Image Analysis, *AIChE J.*, 50, 1474-1483 (2004).

Chapter 5

Potential of hyperspectral imaging for quality control of polymer blend films*

Résumé

Un capteur basé sur l'imagerie visible et proche infrarouge (VIS-NIR) rapide et non intrusif est proposé dans cet article afin de faire le suivi de films composites produits par extrusion soufflage. Un tel capteur a une résolution spatiale nettement supérieure à un système basé sur les sondes. Cette caractéristique lui permet d'identifier de petits défauts localisés pouvant compromettre la qualité (mécanique, optique, propriétés de barrière) de l'échantillon entier. L'analyse d'image multirésolution et multivariée (MR-MIA) a été utilisée pour extraire les caractéristiques spectrales et spatiales des images qui sont le plus fortement corrélées à la distribution de composition des composantes ainsi qu'aux propriétés mécaniques des films. Appliqué aux mélanges de LDPE/PS produits sous différentes conditions d'opération, le capteur a été en mesure de faire le suivi des effets dominants influençant les propriétés mécaniques: (1) l'effet de la composition (matières premières) et (2) l'étirement causé par les conditions post-extrusion (procédé).

Abstract

A VIS-NIR hyperspectral imaging sensor is proposed for on-line non-destructive monitoring of polymer film blends produced by extrusion blowing. Such sensor provides a much higher spatial resolution of analysis compared to probes or probe arrays, which should help detecting smaller localized defects inducing poor film quality (mechanical, optical, barrier properties) at a macroscopic scale. Multiresolutional multivariate image analysis (MR-MIA) was used to extract those spectral and textural features from images that are the most highly correlated to film composition distribution and mechanical properties. When applied to a LDPE/PS blend produced under various operating conditions, the imaging sensor was able to monitor the two dominant effects on mechanical properties: (1) the effect of film composition (raw materials) and (2) the stretching effect caused by post-extrusion conditions (process).

* Gosselin R., Rodrigue D., González-Núñez R. and Duchesne C., Potential of hyperspectral imaging for quality control of polymer blend films, *Ind. Eng. Chem. Res.*, 48, 3033-3048 (2009).

5.1 Introduction

Extrusion film blowing is a continuous industrial process widely used to manufacture thin biaxially stretched polymeric materials like plastic bags and other packaging materials. Although virgin polyethylene (PE) and polypropylene (PP) resins are the most extensively used polymers in this process, films made of polymer blends now constitute over 30% of the total polymer consumption (Utracki, 1998). Film blends have revolutionized the polymer industry due to their great versatility and capability to be tailored for specific end uses (Salamone, 1996). While polymer blending is commonly used to achieve improved mechanical properties (Salamone, 1996; Chang et al., 2002; Nygard et al., 2005), it is currently used in plastic bags to improve biodegradability and recycling (Getlichermann and David, 1994; Goldstein and Olivares, 2007). Literature can now be found on a wide variety of polymer film blends: LLDPE/PS (Dobreski and Nattinger, 1989), PS/PVME (El-Mabrouk et al., 2007), LDPE/PP (Minkova and Nikolova, 1989; Minkova et al., 1992), EVA/PP (Minkova and Nikolova, 1989) as well as PE/Nylon (López-Barrón et al., 2007).

In industrial practice, film quality control is often performed off-line in spite of the high production rates used to ensure film stability (typically 0.75 - 1.25 m/s) and the sensitivity of the film's mechanical properties to physical defects and distribution of the various polymer species within the blend. According to Martyn et al. (2005), operators typically focus on problems such as production instability (i.e. helical, draw resonance, oscillating frost line, bubble sagging, tearing, fluttering, and breathing) as well as others caused by poor dispersion, lack of cohesion and inadequate viscosity paring when dealing with polymer blends. Laboratory quality assessment such as tensile tests (ASTM D882), Elmendorf tear strength (ASTM D1922), density (Lahrouni and Armani, 1995) and crystalline morphology (lamellar, spherulitic, etc.) using microscopy (Ajji et al., 2005) are performed at a low frequency on small samples cut-off from the rolls. Since these analyses are time consuming and destructive, film quality control is performed on a very low fraction of the production. In order to improve film's assessment, correlations have been proposed to link tensile properties to processing conditions (Simpson and Harrison, 1994; Babel and Campbell, 1995; Ghaneh-Fard, 1999). Unfortunately, structure development in

film blowing is very complex (Ghaneh-Fard, 1999). Therefore, on-line quality analysis and control is required to quickly identify any defect and to ensure consistent high quality without loss of production (Ali et al, 1994; Michie et al., 2004; Martyn et al., 2005).

A number of techniques were recently developed for on-line measurement of specific film properties, but those are mostly based on single-point probes (or probe arrays). Therefore, they have a limited spatial coverage of the film surface. Raman spectroscopy is based on the inelastic scattering of a monochromatic (often laser) light source and has been used to acquire real-time crystallinity measurements in blown film (Giriprasath et al., 2004). Birefringence, or the decomposition of a light ray as it passes through anisotropic materials, may be used in the determination of the average total (both crystalline and amorphous) orientation of the film (Mackley et al., 1998; Zhang et al., 2001; Ajji et al., 2005). Fourier transform infrared spectroscopy (FTIR) was used for determining the specific orientation factors of the crystalline and amorphous phases (Ajji et al., 2005; Patel et al., 2006). X-ray scattering through a thin sample reveals information about the structure, composition and properties. In polymer film analysis, X-ray pole figures yield details of crystalline phase orientation (Lahrouni and Armani, 1995; Ajji et al., 2005) while a variation of this method, the wide-angle X-ray diffraction (WAXD), refers specifically to sub-nanometer sized crystalline structures (Patel et al., 2006). These probes either provide a rapid estimation of the properties of a single localized point on the sample, or the average properties of a larger area. They are not intended to analyze a production line presenting both spatial and temporal variations. However, fine spatial measurements are the key in identifying local composition gradients and flow patterns that may compromise global quality of the films. The mechanical, optical and barrier properties of an entire sample may fall below specifications because of local discrepancies easily overlooked by a coarser spatial analysis. Recent developments in spectral imaging cameras may help developing rapid and non-destructive on-line film quality sensors, capturing spatial (both cross and machine directions) and temporal variations in film properties with much higher spatial resolution than using probe arrays.

The objective of this work is therefore to investigate the potential of VIS-NIR spectral imaging for on-line, non-destructive monitoring of spatial as well as temporal variations in film blend properties. The proposed methodology is illustrated using low-density polyethylene/polystyrene (LDPE/PS) films produced by extrusion blowing under different operating conditions. The multiresolutional multivariate image analysis (MR-MIA) technique was used for extracting both spectral and textural image features that are the most highly correlated with the film properties (i.e. composition distribution and mechanical properties), as well as for mapping them on the surface of the films.

5.2 Experimental

5.2.1 Materials and processes

Low-density polyethylene (PEMEX PX-20020X) having a melt-flow index of 2.0 g/10 min was used throughout this work as the continuous phase for the films. For the dispersed phase, three polystyrenes, all provided by Resistol, differing mainly in melt-flow index (MFI) were used: HH-103 (1.7 g/10 min), HF-777 (8.0 g/10 min) and HF-555 (16 g/10 min).

Film production was performed on the same experimental setup as reported in previous studies (Robledo-Ortiz et al., 2006; López-Barrón et al., 2007). A single-screw Haake extruder (Rheomix 254, L/D = 27) coupled with a film-blowing device was used to prepare tubular films as described in Figure 5.1. Process specifications are typically given in terms of film thickness (T), radius (R), and velocity (V) at two positions: the initial properties (I) are taken at the die; while the final properties (F) are measured after solidification of the polymer (freeze line). The pressure drop within the bubble (ΔP) was measured using a water manometer fixed to the air feed. A bypass line was used to compensate for pressure variations within the bubble.

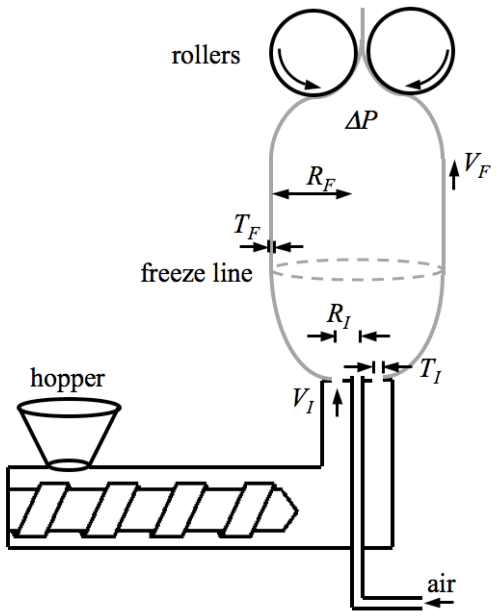

Figure 5.1. Schematic representation of the film blowing process indicating the locations where the thickness (T), radius (R) and velocity (V) of the film are measured.

From the operating conditions, three dimensionless parameters are commonly reported to characterize the films: the blow-up ratio is the ratio of bubble radius to die radius ($BUR=R_F/R_I$), the thickness ratio is the ratio of die slit to bubble thickness ($TR=T_I/T_F$), and the draw ratio is the ratio of the take-off roller velocity to initial film velocity ($DR=V_F/V_I$).

Experimental conditions were chosen in order to span a wide range of variations in end-user film properties introduced by both the processing conditions and the raw material characteristics. The experiments were carried out by varying a single condition at a time around nominal processing conditions (die temperature: 200°C, pressure drop: 30 mm H_2O, screw speed: 50 rpm, polymer throughput: 17.7 g/min, PS melt-flow index: 1.5 g/min, PS content: 1%, premixing: no). Investigated process conditions were: die temperature (190, 200 and 210°C), pressure drop within the bubble (20, 30 and 40 mm H_2O), film drawing speed (1.50, 1.65 and 1.80 m/min) and screw speed (40, 50 and 60 rpm). The latter allowed changing the polymer throughput in flood-feed conditions (16.2, 17.7 and 19 g/min). Raw material characteristics include polystyrene (dispersed phase) melt-flow index (1.5, 8 and 17 g/10 min) and its content within the blend (0, 1, 2, 3, 4, 5, 6, 8 and 10 wt%). Blend premixing (yes or no) was also considered. Each of these conditions were varied one at a

time from a nominal set of operating conditions defined as follows: screw speed of 50 rpm, temperature of 200°C, pressure drop of 30 mm H_2O, drawing speed of 1.65 m/s, no initial premixing and 1% polystyrene (MFI: 1.5 g/10 min). A total of 26 experimental conditions were therefore tested including a few replicates of the nominal conditions. About 15-20 m of film was produced at each condition after reaching steady-state.

To calibrate a regression model between the spectral images and the film properties, each film sample was characterized using standard methods. Tensile properties are by far the most commonly available mechanical tests for films (Hong et al., 2000; Bobovitch et al., 2006). These were measured in triplicate with a United SFM Universal Testing Machine according to ASTM D882 (gauge length 5.08 cm). Measured properties include Young's modulus, toughness, tensile strength, stress at break, as well as the strain at which the latter two occurred (Figure 5.2). All properties were tested solely in the machine direction.

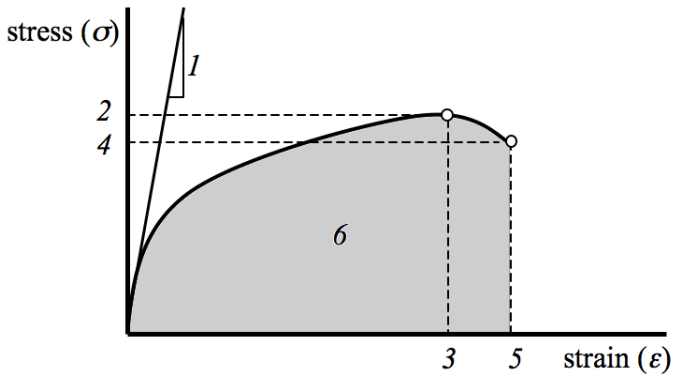

Figure 5.2. Typical stress-strain curve for LDPE/PS film with the parameters studied: Young's modulus (1), the strength and associated strain (2 and 3), the stress and strain at rupture (4 and 5) as well as the sample toughness (6).

5.2.2 Hyperspectral imaging system

The system consists of two line-scan imaging spectroscopes covering the VIS-NIR wavelength ranges, which fields of view are centered on the same area of a sample. However, since the visible range was sufficient to extract the desired information from the LDPE/PS blend, only the VIS spectral scanner is described in this section. Indeed, as

shown later in this paper, both polymers have a different level of clarity; PS is more amorphous due to its aromatic ring whereas LDPE is more crystalline. Other polymer blends may require the use of the NIR or a combination of the VIS-NIR spectral scanners, depending on which wavelength range allow discriminating the polymer species. The proposed method for correlating spectral images to film properties is general and could be used with any spectral scanner.

The line-scan visible imaging system consists of a Hamamatsu C8484-05G camera (960×1024 pixels) coupled with an ImSpector V10E grating spectroscope (Specim Spectral Imaging Ltd) having a slit size of 30 μm × 14.3 mm and sensing the 400-1000 nm wavelength range. For each scanned line of the sample, the transmitted light is dispersed through the spectroscope and captured by the CCD array of the camera as a spatial/spectral image as shown in Figure 5.3. That is, for each pixel along the cross-direction of the sample (i.e. spatial axis x) corresponds a spectrum of light intensities at various wavelengths (i.e. spectral axis λ). The spatial and spectral resolutions (i.e. number of pixels along the x axis and number of intensity values along the λ axis) are determined by the resolution of the CCD camera. In this work, each line-scan has an effective resolution of 300 pixels along x and 600 pixels along λ since it was necessary to crop the images; the film samples did not cover the entire field of view and the sensitivity of the scanner at both ends of the spectral range was low. On the sample, the effective resolution of a line-scan translates into a 0.14 mm/pixel spatial resolution and 0.75 nm/pixel spectral resolution.

After capturing and collecting several line-scans on a moving sample, such as a polymer film drawn in the machine direction (the y axis in Figure 5.3), one obtains a three-way array of data $\underline{\mathbf{R}}$ (or data cube) – a digital hyperspectral image. To investigate the potential of spectral imagery, the films were not scanned directly on the extrusion line but rather off-line in the lab. Representative film samples were therefore cut from the rolls and 200 line-scans were collected using a Velmex motorized translation stage to move the film samples along the y axis. The scans were collected every 375 micron intervals to provide a composite image of true (x,y) spatial proportions. The $(x \times y \times \lambda)$ dimensions of each $\underline{\mathbf{R}}$

spectral image were therefore (300×200×600). A total of 26 spectral images of films $\underline{\mathbf{R}}_i$ (i=1,2,…26) were obtained, one for each experimental condition.

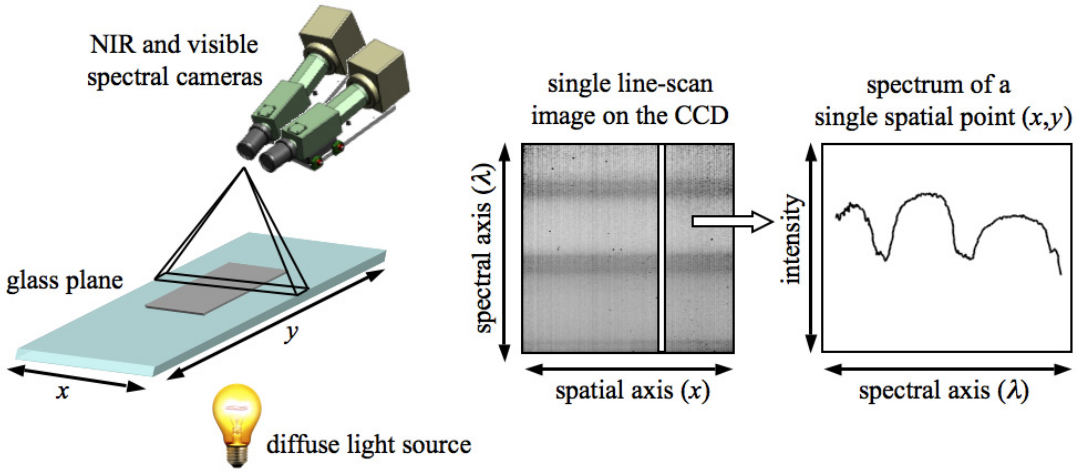

Figure 5.3. Schematic representation of the line-scan system used in transmission and spatial/spectral image as captured by the CCD.

All transmission images were obtained by placing the film on a glass plate above a diffuse light source. A tungsten-halogen lamp was chosen for its smooth flat emission spectrum over the required range. The raw spectral images $\underline{\mathbf{R}}_i$ were first submitted to a linear pixel-by-pixel calibration between a flat black image (**B**), obtained by collecting an image with the lens cap in place; and a true white image (**W**), collected using 99% reflectance standard (Gigahertz-Optik):

$$x_{xy\lambda} = \frac{r_{xy\lambda} - b_{x\lambda}}{w_{x\lambda} - b_{x\lambda}} \qquad (5.1)$$

where $r_{xy\lambda}$ is an element of the raw signal image $\underline{\mathbf{R}}$, $b_{x\lambda}$ and $w_{x\lambda}$ are elements of the black and white calibration images (**B** and **W**), and $x_{xy\lambda}$ is an element of the calibrated transmittance image $\underline{\mathbf{X}}$. Such calibration was performed periodically in order to minimize the impact of variations in lighting conditions over time. Note that more advanced calibration techniques should be considered for long-term industrial applications (Geladi et al., 2004; Burger and Geladi, 2005; Liu et al., 2007). Image acquisition was controlled

through a Labview 8.0 software interface (National Instruments). Data analysis was performed using custom scripts developed in the Matlab R14 (MathWorks) environment unless otherwise stated.

5.3 Multivariate methods for extracting image features

5.3.1 Multivariate image analysis

The MIA imaging technique was originally introduced by Esbensen and Geladi (1989), a detailed description of the method is provided by Geladi and Grahn (1996), and applications to hyperspectral imaging are discussed in Grahn and Geladi (2007). It basically consists of first unfolding a single spectral image $\underline{\mathbf{X}}$ ($x \times y \times \lambda$) into a matrix $\mathbf{X}$ of dimensions ($xy \times \lambda$) as shown in Figure 5.4 (i.e. the spectrum of each pixel are collected row-wise in matrix $\mathbf{X}$), and then performing a variance-covariance decomposition of the spectral matrix $\mathbf{X}$ using principal component analysis (PCA):

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} \qquad (5.2)$$

This decomposition yields a set of A score vectors $\mathbf{t}_a$ ($xy \times 1$) and loading vectors $\mathbf{p}_a$ ($1 \times \lambda$), where A is the number of principal components kept in the model. A non-zero matrix $\mathbf{E}$ ($xy \times \lambda$), containing the projection residuals, is obtained when A is smaller than the number of columns of $\mathbf{X}$. The first score vector $\mathbf{t_1}$ is the linear combination of wavelengths capturing the greatest amount of variations within the spectral image, whereas the second score vector $\mathbf{t_2}$ explains the second greatest source of variations, and so on. The linear combinations are defined by the corresponding loading vectors (i.e. $\mathbf{p_1}$, $\mathbf{p_2}$, etc.). Such a decomposition of the variance-covariance structure of the spectral matrix allows extracting and classifying the spectral signature of each pixel regardless of their spatial location.

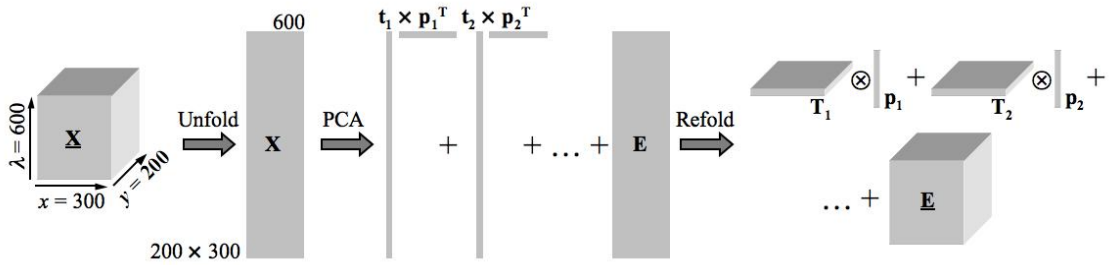

Figure 5.4. Applying standard PCA on a hyperspectral image requires initial unfolding of the data into a 2-D matrix as well as subsequent folding to regain the spatial disposition.

In image analysis **X** usually has a much larger number of rows than columns (i.e. $xy >> \lambda$). Therefore, it is much less computationally intensive to perform PCA (or alternatively singular value decomposition SVD) on the kernel matrix $\mathbf{Z} = \mathbf{X}^{\mathrm{T}}\mathbf{X}$ which is of much smaller dimensions ($\lambda \times \lambda$ rather than $xy \times \lambda$). When a global PCA model is to be developed on a set of N images, the kernel matrix is given by: $\mathbf{Z} = \sum_{a=1}^{N} \mathbf{X}_a^T \mathbf{X}_a$.

As for PCA applied to a database, the clustering pattern of the extracted features is observed using $\mathbf{t_i}$-$\mathbf{t_j}$ score scatter plots (i≠j). However, when applied to a multivariate image, these score plots are typically shown as 2-D density histograms using a pre-defined color map to enhance visual interpretation, since the number of observations in the score plot is equal to the number of pixels of the image (i.e. $x \times y$ or 300×200=60000 observations for each image). An alternative way to visualize the image features extracted by each principal component is obtained by refolding the score vectors $\mathbf{t}_a$ ($xy \times 1$) into score matrices $\mathbf{T}_a$ ($x \times y$) according to the original spatial location of each pixel in the image $\underline{\mathbf{X}}$ as shown in Figure 5.4 (i.e. $\underline{\mathbf{X}} = \sum_{a=1}^{A} \mathbf{T}_a \otimes \mathbf{p}_a + \underline{\mathbf{E}}$ where ⊗ is the Kronecker product). The score matrices $\mathbf{T}_a$ are then displayed as univariate gray-level images after appropriate scaling is applied. Figure 5.5 shows the $\mathbf{T_1}$ and $\mathbf{T_2}$ images of a film sample produced under the nominal process conditions described in the previous section. These images not only show intensity variations (i.e. spectral variations), but also show some textural variations across the images. As shown later in Section 5.4, both spectral and spatial (i.e. textural) information are necessary for predicting the mechanical properties of the polymer films.

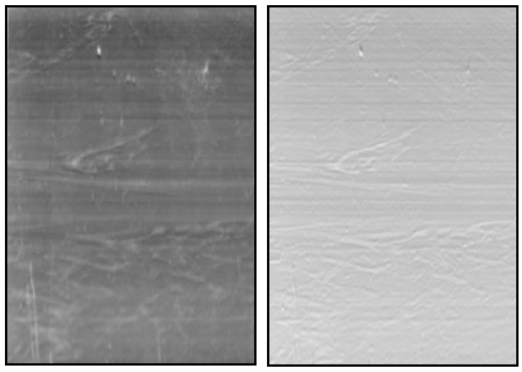

Figure 5.5. Textural features of a film sample produced under the nominal conditions as seen from the $\boldsymbol{T_1}$ *(left) and* $\boldsymbol{T_2}$ *(right) score image.*

5.3.2. Multiresolutional multivariate image analysis

Spectral information of an image is efficiently extracted using MIA. However, unfolding the spectral array $\underline{\mathbf{X}}$ as shown in Figure 5.4 leads to spatial information loss. The spatial organization of pixel intensities or *textural information* of the $\mathbf{T_1}$ and $\mathbf{T_2}$ images shown in Figure 5.5 cannot be extracted using MIA, but requires the use of texture analysis methods. Several textural methods were overviewed and compared in Bharati et al. (2004) using images of steel sheets. One conclusion drawn in this work was that the Wavelet Texture Analysis (WTA) method appeared to be the most powerful for image texture analysis owing to its space-frequency decomposition of the image information. Later, Liu and MacGregor (2007) proposed the multiresolutional multivariate image analysis (MR-MIA) framework, combining the strengths of both MIA and WTA for spatial-spectral analysis of multispectral digital images. Two algorithms were proposed depending upon the degree of interaction between the spatial and spectral information. First, MR-MIA (I) consists of extracting textural information from each spectral channel first using WTA resulting in a set of congruent texture images, one at each resolution (i.e. frequency), and direction of analysis (horizontal, vertical and diagonal), and then using MIA on these texture images for extracting the spatial-spectral information. This approach is general, widely applicable and particularly well suited for interacting spatial-spectral information. Second, MR-MIA (II) performs spectral analysis first using MIA, and then extracts textural information using

WTA applied to score images $\mathbf{T}_a$. This algorithm should be used when little interaction exist between spatial-spectral information. The latter algorithm is used in this work since the textural features seen on films (Figure 5.5) are observable in most of the wavelengths across the VIS range, hence in the $\mathbf{T_1}$ and $\mathbf{T_2}$ images (i.e. there is little interaction between spectral and spatial information). The MR-MIA (II) algorithm is briefly described below, but the reader is referred to Liu and Macgregor (2005, 2006, 2007) for more details on this algorithm, and to Chau et al. (2004) and Burke Hubbard (1998) for a more general overview of spatial-spectral information extraction.

Multiresolution analysis (MRA) is basically a succession of approximations of a signal at different resolutions (scales or frequencies), which can be obtained using methods such as orthogonal wavelet decomposition. Wavelet texture analysis (WTA) is typically suited in applications where the data contains both large and small scales variation in which case it is capable of isolating the former from the latter. WTA is akin to the windowed Fourier transform (WFT) using a non-periodic function, or wavelet. When applied to textural analysis of univariate digital images such as a gray-scale image of a scene of a score image $\mathbf{T}_a$ (i.e. 2-dimensional discrete signals), the separable 2-D discrete wavelet transform (DWT) algorithm is used. It approximates the image by repeatedly dilating and shifting a single mother wavelet, each level of dilation representing larger or smaller scale textural features. Two types of coefficients characterize the signal: the detail coefficients ($d_{j,l}$) characterizes the match between the mother wavelet and the original image signal at a level (j) and a position (l), while the approximation coefficients ($a_{j,l}$) corresponds to the remaining information after extraction of the features by the detail coefficients. This rougher representation of the signal is then used as the basis for the $d_{j+1,l}$ coefficients. The signal $\mathbf{x}$ (time series of a row or a column of pixel intensities) is decomposed through simultaneous filtering using a high-pass filter $\psi(\mathbf{x})$, called the wavelet function, and a low pass filter $\varphi(\mathbf{x})$ called the scaling function:

$$a_{j,l} = \left\langle f(k), \phi_{j,l}(k) \right\rangle \tag{5.3}$$

$$d_{j,l} = \left\langle f(k), \psi_{j,l}(k) \right\rangle \tag{5.4}$$

First level decomposition is achieved by applying the DWT algorithm to the input image (in this case $\mathbf{T}_1$). This produces three detail coefficients matrices ($\mathbf{D}_{Hj}$, $\mathbf{D}_{Vj}$, $\mathbf{D}_{Dj}$) corresponding to the horizontal, vertical and diagonal matches between the original image and the mother wavelet (Figure 5.6). The remaining information unexplained by the detail coefficients matrices is stored in the first level approximation coefficient matrix ($\mathbf{A}_j$). As each successive level isolates a new level of detail from an increasingly rough low-pass approximation, it halves the frequency interval according to Nyquist's law. In 2-D DWT, it also halves the spatial resolution in each direction, making the approximation coefficient matrix a quarter the size of the original image. Higher-order decompositions (level j+1) are achieved by applying the aforementioned algorithm on the preceding approximation coefficient matrix ($\mathbf{A}_j$), reducing the size of the image at each level.

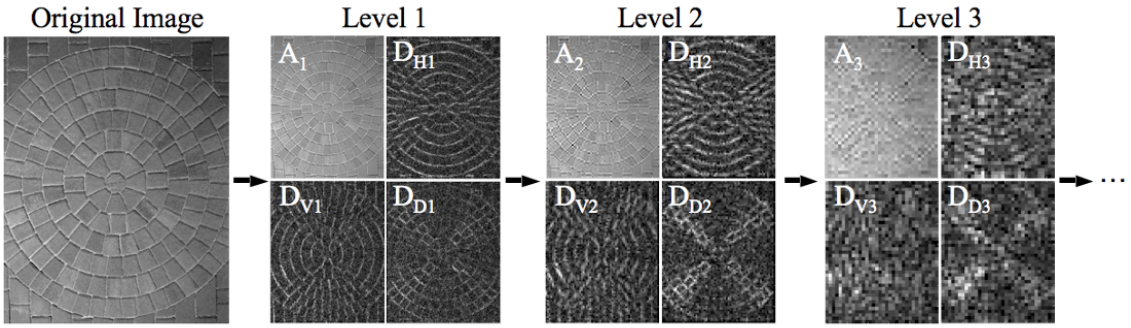

Figure 5.6. Illustration of a 2-D wavelet decomposition of a manufactured stone walkway in which the horizontal, vertical and diagonal detail coefficient matrices (D_{Hj}, D_{Vj} and D_{Dj}) as well as the approximation coefficient matrix (A_j) are shown for each level j.

5.4 Results and discussion

5.4.1 Causal relationships between materials, process conditions and mechanical properties

The experimental design discussed in Section 5.2 generated a dataset of input variables $\mathbf{X}_{\mathbf{data}}$ (26×10) containing the raw material properties and process conditions for each experiment as well as a corresponding response matrix $\mathbf{Y}$ (26×6) containing the mechanical properties of the films (Figure 5.7). These operating parameters can be divided into three sets: dispersed phase effect (PS MFI and content), processing parameters (screw speed, die

temperature, take-off speed, bubble pressure drop and premixing) and dimensionless parameters (*BUR*, *TR* and *DR*). This rich database was used to build causal relationships between the raw material properties, the processing conditions and the resulting film mechanical properties, which were found to provide important insights for developing the imaging sensor.

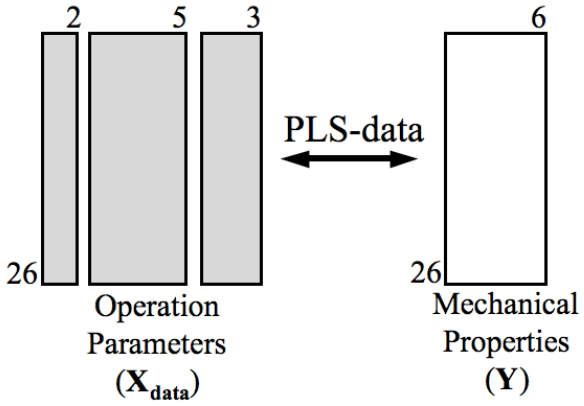

Figure 5.7. The PLS-data model is based on the operation conditions (2 PS characteristics, 5 processing parameters and 3 dimensionless ratios) to predict the mechanical properties.

A latent variable PLS regression model was built between $\mathbf{X}_{\mathbf{data}}$ and $\mathbf{Y}$. The PLS model structure is given below:

$$\mathbf{X}_{data} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} = \mathbf{T}_A \mathbf{P}_A^{\mathrm{T}} + \mathbf{E} \tag{5.5}$$

$$\mathbf{Y} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{q}_a^{\mathrm{T}} + \mathbf{F} = \mathbf{T}_A \mathbf{Q}_A^{\mathrm{T}} + \mathbf{F} \tag{5.6}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^{*} \tag{5.7}$$

The **P** and **Q** matrices contain the loading vectors that best represent the **X** and **Y** spaces respectively, whereas **W*** contain the loading vectors defining the common latent variable space **T** relating **X** and **Y**. The **E** and **F** matrices contain the PLS model residuals (i.e. projection distance from the latent variable space of the model), and A is the total number of PLS components. Even if the design matrix $\mathbf{X}_{\mathbf{data}}$ is very well conditioned, PLS was selected here since the property matrix **Y** is highly collinear (i.e. effective rank=2 using

PCA and cross-validation). The first three principal components (PC) were found significant by leave-one-out cross-validation. The cumulative **Y** variance explained and predicted (cross-validation) by the model were $R^2_{Y,cum}$=69% and Q^2_{cum}=38%. These statistics are relatively low since they combine all 6 mechanical properties, and some of them have a higher degree of uncertainty.

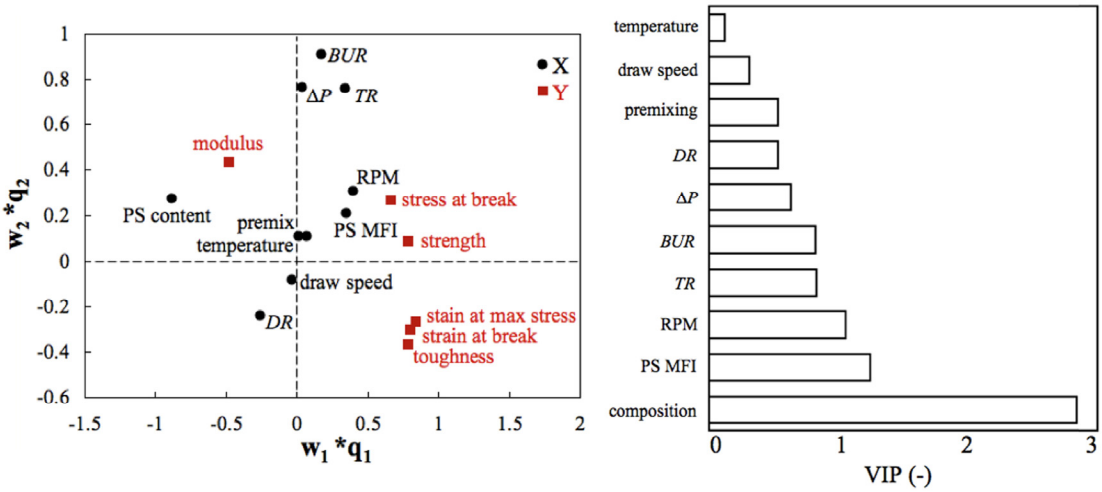

Figure 5.8. Loading biplot for the first two components (left) and the VIP plot (right) of the PLS-data model.

Relationships between raw materials properties, processing conditions and film mechanical properties are interpreted using the two plots shown in Figure 5.8. The first (left) is the PLS-data loading biplot for the first two PCs. It is used to interpret the inter-correlations of the variables both within the $\mathbf{X_{data}}$ and **Y** blocks as well as between them. Variables having similar coordinates in this plot are positively correlated, those falling in opposite quadrants are negatively correlated, and those falling close to the origin (0,0) have small or no correlation with other variables. The second plot in Figure 5.8 (right) is the variable importance (VIP) plot. It indicates the relative importance of the 10 regressor variables for explaining variations in **Y** based on the PLS-data model.

The PLS-data model strongly relies on polystyrene characteristics (1st and 2nd VIPs), followed by the throughput rate (3rd VIP) and the bubble geometry (4th - 7th VIP). However, blend premixing, draw ratio and extrusion temperature were much less important in

modeling the tensile properties. The PS content is the single most important factor in determining the mechanical properties of the films and as such is expressed along the first component of the biplot ($\mathbf{w_1}$*$\mathbf{q_1}$). In accordance with previous studies (Barentsen and Heikens, 1973; Wang et al., 1998; Guo et al., 1998), Figure 5.8 shows that high PS content increases the modulus while decreasing all other properties, especially affecting the elongational properties (i.e. strain at break, strain at maximum stress and the toughness). In order to increase these three properties in the machine direction it is necessary to promote molecular orientation along that same axis (Simpson and Harrison, 1994; Fatahi et al., 2007). In practice, this can be achieved by reducing the bubble diameter and increasing the stretching in the machine direction. Graphically, bubble geometry is represented along the $\mathbf{w_2}$*$\mathbf{q_2}$ axis of the PLS-data biplot in which case the elongational properties are negatively correlated to *BUR, TR* and ΔP, and positively correlated to *DR* and the draw speed.

It is important to note that the PLS-data model was developed using some variables that are generally not available on-line in practice, such as film composition (i.e. PS content), thickness and draw ratios (TR and DR) as well as the mechanical properties. Nevertheless, this model was developed and used 1) for verifying that the observed relationships between raw materials properties, processing conditions and film mechanical properties were consistent with polymer processing knowledge, and 2) as a benchmark for the spectral imaging sensor. Indeed, this model showed that two dominant sources of variations affect mechanical properties: film composition and stretching. The spectral imaging sensor should be capable of detecting both of these parameters and of predicting the mechanical properties of the films solely based on image information.

5.4.2 Spectral analysis of the films for estimating blend composition

The goal of the image analysis was to build a model based on the optical aspect of the films with no prior knowledge of the experimental conditions or raw material composition. Both composition and orientation must therefore be inferred from the images. Figure 5.9 shows the average transmission spectra of thin films made of pure LDPE and PS scaled according to Equation 5.1.

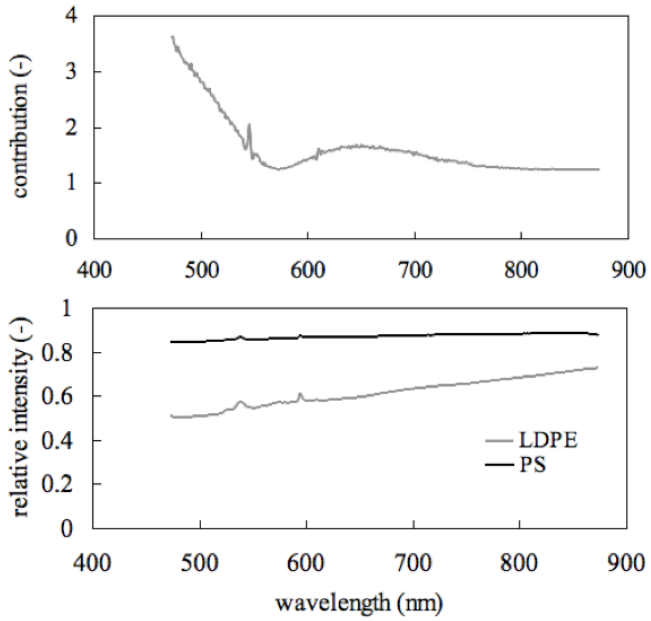

Figure 5.9. Average visible spectra of thin films made of LDPE and PS.

A mostly flat response was obtained for PS over the entire range whereas LDPE had a steadily increasing response implying a greater transparence at longer wavelengths. Considering that both polymers can clearly be distinguished based on their transmission spectra it should therefore be possible to estimate blend composition using image spectral analysis (e.g. MIA). Furthermore, score images (later used in the texture analysis) should draw more information from shorter wavelengths (near 500 nm) than from longer ones (above 700 nm) due to the greater spectral contrast in the lower range. The localized glitches in the curves around 546 and 612 nm correspond to the typical peaks of room fluorescent lighting.

Film composition was estimated using the first two score vectors ($\mathbf{t_1}$ and $\mathbf{t_2}$) obtained using MIA as shown in Figure 5.10. The clustering pattern of the spectra obtained for all 26 film samples is shown in Figure 5.10A using a 2-D density histogram. It shows the classification of the pixels of each film image according to their spectral characteristics. Figure 5.10B provides a closer look at this clustering pattern using the films produced with the lowest and the highest PS content. Finally, Figure 5.10C shows the regression line within the $\mathbf{t_1}$-$\mathbf{t_2}$

scatter plot that is the most highly correlated with PS content. As expected, it is strongly oriented towards $\mathbf{t_1}$.

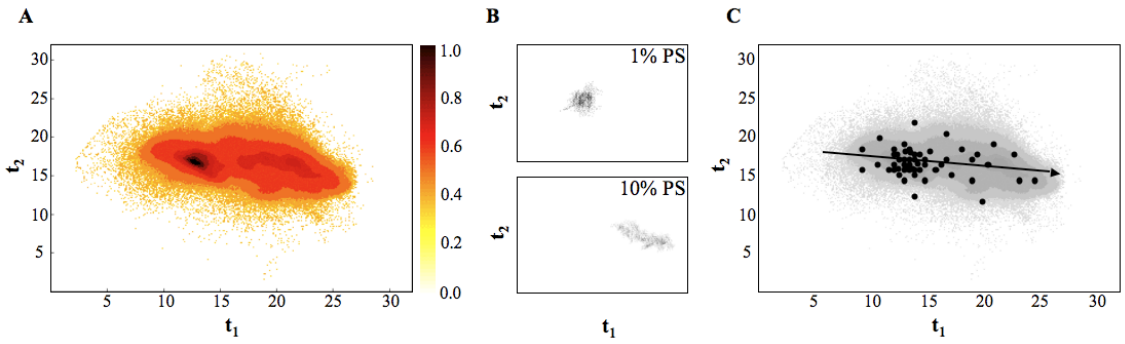

Figure 5.10. Score histogram of the film illustrating the global dataset (A), individual samples within this dataset (B) and the average location of all the samples (C). Dimensionless intensity scale given.

The data illustrated in Figure 5.10A fall into two basic areas: samples containing 1% PS ($10 < \mathbf{t_1} < 15$) and the samples containing 2-10% PS ($15 < \mathbf{t_1} < 25$). Within this latter group there is also a strong correlation between PS content and the first score vector, indicating that it accounts for a large faction of the variation. On average, pixels taken from samples containing higher levels of polystyrene have higher $\mathbf{t_1}$ values than pixels taken from samples containing less polystyrene. To illustrate this, samples containing 1 and 10% PS are shown in Figure 5.10B. However, this figure also shows that the pixels for 1% PS are tightly clustered, indicating a uniform composition, while those for 10% PS are scattered over a wider range of $\mathbf{t_1}$ values, indicating a more heterogeneous composition. This local variability at higher PS contents may be increased by the natural incompatibility between the PE and PS phases.

The $\mathbf{t_1}$-$\mathbf{t_2}$ scatter plot can be used to predict the local composition of the films. This can be achieved by regressing the average PS content (known by design) of the samples against their corresponding average $\mathbf{t_1}$-$\mathbf{t_2}$ values using least squares regression (Figure 5.10C). This model was then used to predict PS concentration distribution within a single film sample based on the $\mathbf{t_1}$ and $\mathbf{t_2}$ values of each image pixel. This local information can be used to identify non-uniformities in the polymer distribution (Figure 5.11). Furthermore, this local

composition estimation can be extended to a global value simply by summing this information for each pixel of an image.

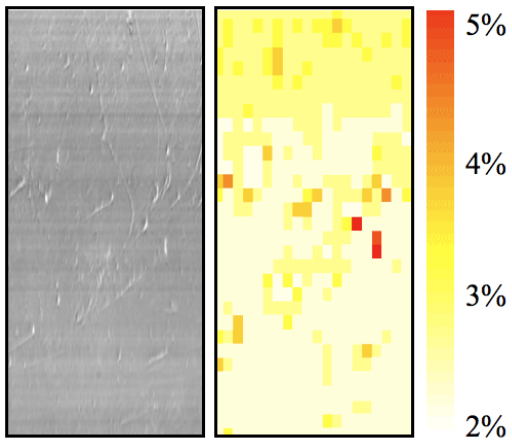

Figure 5.11. Determination of local PS content within a sample through spectral analysis. Right: score image of a sample containing defects. Left: illustration of estimated composition showing PS in red, PE in white and intermediates in shades of yellow – color scale is arbitrary. Sample produced at nominal conditions and 3% PS.

It is important to note that Figure 5.11 illustrates such a classification in which the resolution was lowered to make small PS inclusions clearly visible. This analysis shows that apparently similar defects present in the score image are found to contain different concentrations of PS. In a more subtle way, it also becomes apparent that the upper section of this figure contains intermediate levels of PS dispersed throughout the matrix while the bottom of the figure does not. It is therefore expected that this method could also be used to distinguish the local concentrations of miscible polymers that do not lead to visible defects.

Local composition gradients can cause weaknesses in the film and influence its mechanical properties. Composition mapping can therefore be used to quantify the PS dispersion throughout the PE matrix and link it to the properties of the films. The degree of composition homogeneity can be quantified using Danckwerts' segregation intensity (I_{PS}) for binary systems (Danckwerts et al., 1952):

$$I_{PS} = \frac{\sigma_{PS}^2}{\bar{x}_{PS}(1-\bar{x}_{PS})} \qquad (5.8)$$

where $\bar{x}_{PS}$ and σ^2_{PS} are the average and variance of the estimated PS content within the image. I_{PS} does not reflect the size of the PS domains, but the extent to which the concentration in the clumps departs from the mean. The predicted composition, as well as its segregation intensity, can be used to predict the tensile properties.

5.4.3 Textural analysis for modeling the stretching effect

Application of the MR-MIA II algorithm went as follows. A global MIA was first performed on the image set in order to extract two components ($\mathbf{T_1}$ and $\mathbf{T_2}$), explaining 96% and 3% of the variance, respectively. While each loading is independent, the first two score images are essentially mirror images of one another (Figure 5.5). Quantitatively, the relative textural features extracted from $\mathbf{T_1}$ and $\mathbf{T_2}$ by WTA do not differ significantly from each other. Therefore, only the first score images were used in the following analysis. A 3-level WTA was performed on the images using bior3.7, a bi-orthogonal wavelet available in the Matlab Wavelet Toolbox. The textural features being relatively small, higher-level decompositions were not found to contain meaningful information. In order to quantify image texture, the mean energies of the horizontal and vertical detail coefficients matrices ($\mathbf{D}_{Hj}$, $\mathbf{D}_{Vj}$) were calculated horizontally and vertically for each score image according to:

$$H_j = \sqrt{\sum diag\left(\mathbf{D}_{Hj}{}^{\mathrm{T}}\mathbf{D}_{Hj}\right)} \tag{5.9}$$

$$V_j = \sqrt{\sum diag\left(\mathbf{D}_{Vj}{}^{\mathrm{T}}\mathbf{D}_{Vj}\right)} \tag{5.10}$$

The horizontal and vertical energy details (H_j or V_j) are scalars based on the Frobenius norm of the coefficient matrices ($\mathbf{D}_{Hj}$ or $\mathbf{D}_{Vj}$) at decomposition level j. Ratios of horizontal (H: machine direction) and vertical (V: cross direction) values were also used to emphasize certain features: H_i/V_i ratios provide information on relative horizontal/vertical orientation of the films at each WTA level i while H_i/H_j and V_i/V_j ratios describe the level at which the features occur. That is to say; a large H_1/H_3 ratio would either imply that the sample is globally uniform (low H_3) or locally rough (high H_1). To summarize, 15 textural features

(H_i, V_i, H_i/V_i, H_i/H_j and V_i/V_j for i,j = [1, 2, 3] where $i \neq j$ and $i<j$) describe each image according to orientation and level as quantified by energy of the wavelet coefficients.

5.4.4 Relationships between visual features and mechanical properties

A second model, referred to as PLS-image, was built in which visual dataset of input variables $\mathbf{X}_{\mathbf{image}}$ (26×17) was used as a predictor for the mechanical properties **Y** (26×6) as shown in Figure 5.12. Of the 17 visual features, 15 were drawn from the textural analysis (Section 5.3) while the other two estimated the polystyrene content and polystyrene segregation intensity (Section 5.2).

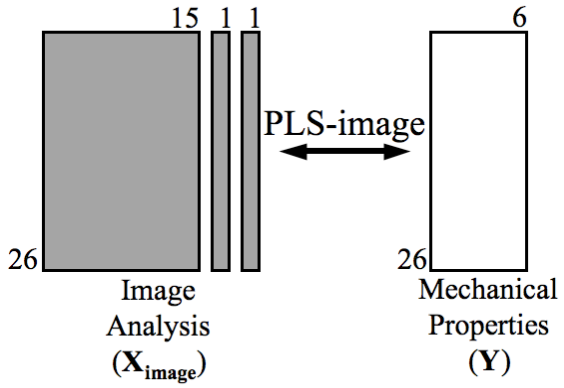

Figure 5.12. The PLS-image model is based on the visual aspects of the films (15 textural features, estimated composition and segregation intensity) to predict the mechanical properties.

A VIP plot for the PLS-image model, based on the 17 visual features, reveals that this model also strongly relies on the presence of polystyrene (composition: 1st VIP) as well as its distribution throughout the film (segregation: 8th VIP) to predict the tensile properties (Figure 5.13). Furthermore, PLS-image model is based on the large-scale horizontal/vertical orientation of the films (H_3/V_3: 2nd VIP), followed by the horizontal components at each decomposition level (H_3, H_2 and H_1), confirming that the image analysis uses both local and global information to model the properties.

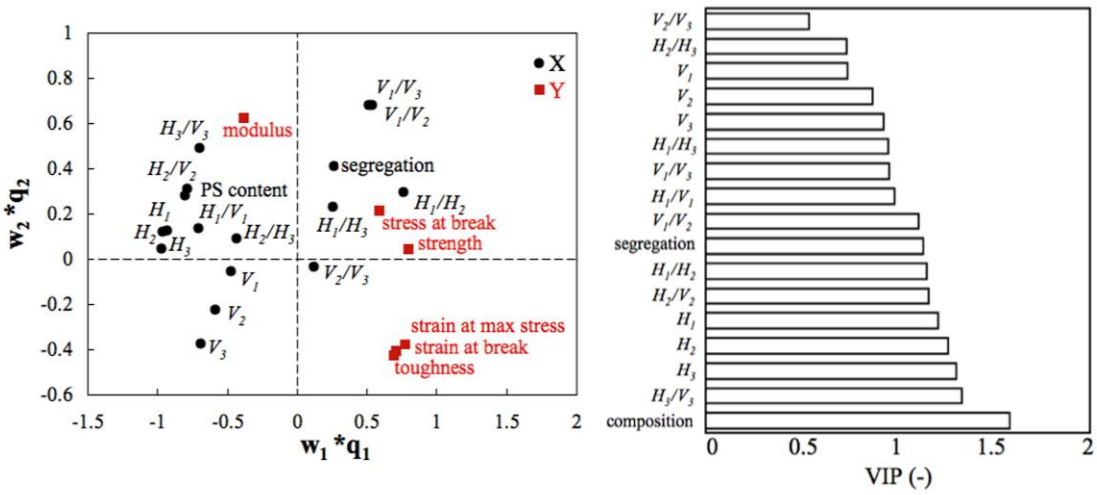

Figure 5.13. Loading biplot for the first two components (left) and the VIP plot (right) of the PLS-image model.

Similar conclusions can be drawn from the PLS-image biplot. As LDPE and PS are incompatible, higher PS contents will increase the intensity of the textural features in both the machine (*H*) and transverse (*V*) directions along the $\mathbf{w_1*q_1}$ axis. Here again, the secondary $\mathbf{w_2*q_2}$ axis represents the bubble geometry (stretching) as shown by the strong negative correlation between opposing textural features (i.e. V_3 vs. H_3/V_3).

5.4.5 Prediction of tensile properties

In order to judge the adequacy of this second PLS-image model, it must be compared to the previous PLS-data model (Section 5.1). To do so, two cumulative multiple correlation coefficients were computed for quantifying the model predictive ability: $R^2_{Y,cum}$ gives the cumulative percentage of the total variance of **Y** explained by the model, and Q^2_{cum} is the cumulative percentage of the total variance of **Y** that can be predicted by the models using the cross-validation procedure. Both PLS models were constructed keeping only the first three components. The PLS-data model had an $R^2_{Y,cum}$ of 0.69 and Q^2_{cum} of 0.38 while the PLS-image model had an $R^2_{Y,cum}$ of 0.64 and Q^2_{cum} of 0.41. Once constructed, these models were then used to predict the six mechanical properties. Relative errors, expressed as the ratio of the prediction error to the measured value, are compiled for the 26 experimental conditions. Mean values are presented in Table 5.1.

Table 5.1. Relative error in prediction by 3-component PLS models; comparing image analysis (PLS-image) to operation parameters (PLS-data). Values are in percentage.

Mechanical property	PLS-image	PLS-data
Young's modulus	5 (3)	5 (4)
Tensile strength	6 (4)	5 (4)
Stress at break	8 (10)	7 (9)
Strain at max stress	17 (14)	16 (13)
Strain at break	14 (13)	14 (10)
Toughness	14 (13)	12 (11)
Average error	11 (11)	10 (10)

Table 5.1 shows that both PLS models have similar prediction capability. The properties that do not depend on the breaking point (Young's modulus and tensile strength) are well represented by the models, with relative errors less than 5-6%; while the properties at break (i.e. the strain at break, strain at maximum stress and toughness) are not as well represented. This was expected since ultimate properties (at break) usually have high data scatter (Ghaneh-Fard, 1999). In our case, scatter was especially high for samples containing higher PS contents (6-10%). Despite being included in the models, these samples differ significantly from the remaining samples. The strain at which the maximum stress and break occur is much less that would predict the models. Since toughness is strongly dependent on ultimate properties, it is also difficult to predict with high precision.

The performance of the models in both situations (i.e. properties depending on the breaking point and those that do not) is illustrated in Figure 5.14. This figure shows the experimentally measured toughness and tensile strength compared with the predictions obtained by cross-validation. Accordingly, the predicted tensile strength is closer to the experimental values ($R^2_{Y,cum}$: PLS-image = 0.785 and PLS-data = 0.826) than for toughness ($R^2_{Y,cum}$: PLS-image = 0.562 and PLS-data = 0.639).

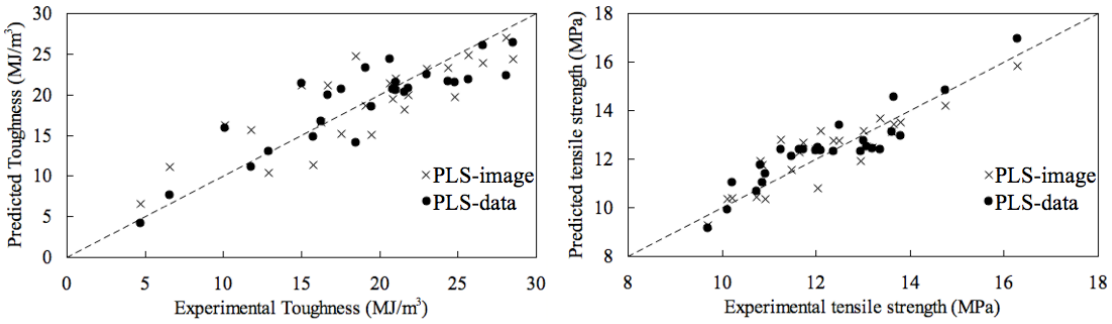

Figure 5.14. Prediction capabilities for toughness and tensile strength.

Despite the fact that the PLS-data model provides marginally better predictions than PLS-image, it is important to note that it relies on more information; namely the film composition, geometry and process conditions. However, the exact polystyrene content throughout the blend at all times, the present state of the operating conditions, as well as less accessible data such as bubble thickness, diameter and draw ratio are not usually available in real time and can only be obtained through and an array of sensors or estimated by mass balance calculations (Ito et al., 2004); i.e. the PLS-data model should be seen as a best-case scenario for which the entire process is well instrumented.

The proposed methodology, through the PLS-image model, only relies on the visual aspect of the films through a spectral estimation of the composition and a textural estimation of the film geometry. Moreover, it serves also as a final quality control by identifying defects (such as PS inclusions) as well as composition gradients throughout the sample.

5.5 Conclusion

The objective of this work was to develop a rapid and non-intrusive on-line imaging sensor for monitoring spatio-temporal variations across the surface of polymer film blends. Although NIR spectroscopy has previously been investigated for estimating some film properties, most methods were based on probes, hence providing measurements for a single point (or line) on a sample. Since most of the overall end-user properties of polymer films, such as mechanical, optical and barrier properties, depend upon the local microstructure,

fine spatial measurements are key in identifying local composition gradients that may compromise the quality as a whole. Rich spatial data would also be useful for precise diagnostics of operational problems in the production line such as non-uniform cooling or poor distribution of the dispersed phase.

As a proof of concept, LDPE/PS film blends were produced via extrusion film blowing in a wide range of processing conditions including the rotational speed of the extruder screw (flow rate), the die temperature, the pressure drop within the film bubble, the drawing speed of the film as well as blend premixing. Polymer characteristics include the melt-flow index of the polystyrene dispersed phase as well as its content within the blend. Each sample was then imaged using a visible imaging spectroscope (400-1000 nm), and sent for mechanical testing. Samples were characterized according to 6 mechanical properties: Young's modulus, the strength and associated strain, the stress and strain at rupture as well as toughness.

A first PLS model was constructed using all the process data typically available (true composition, the operating conditions and the bubble geometry) to predict the properties of the film. In doing so, two major sources of variation were observed: the composition and stretching of the film. In light of this, a second PLS model, based on multiresolutional multivariate image analysis (MR-MIA), was constructed. This method relies on multivariate image analysis (MIA) coupled with wavelet texture analysis (WTA) to estimate the texture, composition and phase dispersion.

Results show that both PLS models have similar predictive capabilities. However, the first model, based on process data, does so by relying on estimated process information that cannot always be easily obtained in real time (i.e. film thickness and true PS content). Knowledge of these parameters requires multiple sensors throughout the extrusion line. On the other hand, the second model avoids this limitation altogether by relying solely on image data obtained with a single sensor. Beyond predicting the mechanical properties, image analysis can provide real-time 2-D composition estimates, therefore identifying all potential local defects and composition gradients. The ability of the proposed approach to

combine both local and global characteristics shows an interesting potential for application to complex materials requiring fine spatial measurements.

5.6 References

Ajji A., Zhang X. and Elkoun S., Biaxial Orientation in HDPE Films: Comparison of infrared spectroscopy, X-ray pole figures and birefringence techniques, *Polymer*, 46, 3838-3846 (2005).

Ali A.H., Hsieh J.T.T., Kauffman K.J., Kissin Y.V., Ong S.C., Prasad G.N., Pruden A.L. and Schregenberger S.D., Producing blown film and blends from bimodal high density high molecular weight film resin using magnesium oxide-supported Ziegler catalyst, *US Patent*, 5284613 (1994).

Babel A.K. and Campbell G.A., Model linking process variables to the strength of blown films produced from LDPE and LLDPE, *TAPPI J.*, 78(5), 199-206 (1995).

Barentsen W.M. and Heikens D., Mechanical properties of polystyrene/low density polyethylene blends, *Polymer*, 14, 579-585 (1973).

Bharati M.H., Liu J.J. and MacGregor J.F., Image texture analysis: methods and comparisons. *Chemom. Intell. Lab. Syst*., 72, 57-71 (2004).

Bobovitch A.L., Tkach R., Ajji A., Elkoun S., Nir Y., Unigovski Y. and Gutman E.M., Mechanical properties, stress-relaxation, and orientation of double bubble biaxially oriented polyethylene films, *J. Appl. Polym. Sci.*, 100, 3545-3553 (2006).

Burger J. and Geladi P., Hyperspectral image regression part I: calibration and correction, *J. Chemometr.*, 19, 355-363 (2005).

Burke Hubbard B., *The world according to wavelets, the story of a mathematical technique in the making,* A.K. Peters: USA (1998).

Chang A.C., Tau L., Hiltner A. and Baer E., Structure of blown film from blends of polyethylene and high melt strength polypropylene, *Polymer*, 43, 4923-4933 (2002).

Chau F., Liang Y., Gao J. and Shao X., *Chemometrics from basics to wavelet transform*, Wiley & Sons: USA (2004).

Danckwerts P.V., The definition and measurement of some characteristic of mixtures, *Appl. Sci. Res.*, 3A, 279-296 (1952).

Dobreski D.V. and Nattinger B.E., Plastic bags of blends of linear ethylene polymers and aromatic polymers, *US Patent*, 4804564 (1989).

El-Mabrouk K., Belaiche M. and Bousmina M., Phase separation in PS/PVME thin and thick films, *J. Colloid Interf. S.*, 306, 354-367 (2007).

Esbensen K.H. and Geladi P., Strategy of multivariate image analysis (MIA). *Chemom. Intell. Lab. Syst*., 7, 67-86 (1989).

Fatahi S., Ajji A. and Lafleur P.G., Invertigation on the structure and properties of different PE blown film, *Int. Polym. Proc.*, 22, 334-345 (2007).

Geladi P., Burger J. and Lestander T., Hypespectral imaging: calibration problems and solutions, *Chemometr. Intell. Lab.*, 72, 209-217 (2004).

Geladi P. and Grahn H. *Multivariate Image Analysis*, Wiley: United Kingdom (1996).

Getlichermann M. and David C., Morphology of polyethylene blends: 3. Blend films of polyethylene with styrene-butadiene or styrene isoprene copolymers obtained by extrusion-blowing, *Polymer*, 35(12), 2542-2548 (1994).

Ghaneh-Fard A., Effects of film blowing conditions on molecular orientation and mechanical properties of polyethylene films, *J. Plast. Film Sheet.*, 15(3), 194-218 (1999).

Giriprasath G., Cherukupalli S. and Ogale A., Real-time crystallinity measurements in a multilayer blown film, *ANTEC*, Chicago (Il), 200-204 (2004).

Goldstein N. and Olivares C., What's what and who's who in compostable products, *BioCycle*, 48(7), 38-47 (2007).

Grahn H. and Geladi P., *Techniques and Applications of Hyperspectral Image Analysis*, Wiley: USA (2007).

Guo H.F., Packirisamy S., Mani R.S., Aronson C.L., Gvozdic N.V. and Meier D.J. Compatibilizing effects of block copolymers in low-density polyethylene/polystyrene blends, *Polymer*, 39(12), 2495-2502 (1998).

Hong Y., Coombs S.J., Cooper-White J.J., Mackay M.E., Hawker C.J., Malmström E. and Rehnberg N., Film blowing of linear low-density polyethylene blended with a novel hyperbranched polymer processing aid, *Polymer*, 41, 7705-7713 (2000).

Ito H., Suzuki K., Kikutani T., Kang H. and Kanai T., On-line birefringence measurements in film blowing process, *ANTEC*, Chicago (Il), 1856-1859 (2004).

Lahrouni A. and Armani J., Étude de mélanges de polyéthylène et de polystyrène par diffraction de rayons X et par densimétrie, *Eur. Polym. J.*, 31(4), 347-352 (1995).

Liu Z., Yu H. and MacGregor J.F., Standardization of line-scan NIR Imaging Systems, *J. Chemometr.*, 21, 88-95 (2007).

Liu J.J. and MacGregor J.F., Modeling and optimization of product appearance: application to injection-molded plastic panels, *Ind. Eng. Chem. Res.*, 44, 4687-4696 (2005).

Liu J.J. and MacGregor J.F., Estimation and monitoring of product aesthetics: application to manufacturing of "engineered stone" countertops, *Mach. Vision Appl.*, 16(6), 374-383 (2006).

Liu J.J. and MacGregor J.F., On the extraction of spectral and spatial information from images, *Chemometr. Intell. Lab.*, 85, 119-130 (2007).

López-Barrón C.R., Robledo-Ortiz J.R., Rodrigue D. and González-Núñez R., Processing, morphology, and properties of polyamide-6/low density polyethylene blends prepared by film blowing, *J. Plast. Film Sheet.*, 23(2), 149-169 (2007).

Mackley M.R., Rutgers R.P.G. and Gilbert D.G., Surface instabilities during the extrusion of linear low density polyethylene, *J. Non-Newton. Fluid*, 76, 281-297 (1998).

Martyn M.T., Smith G.D. and Coates P.D. Effect of screw design on polyethylene blown film processing, *ANTEC,* Boston (NH), 286-295 (2005).

Michie W.J. Jr., Neubauer A.C., Cobler B.A. and Baker C.F., Polymer composition and process to manufacture high molecular weight-high density polyethylene and film therefrom, *US Patent*, 101674 (2004).

Minkova L. and Nikolova M., Melting of irradiated films prepared from polymer blends, *Polym. Deg. Stab.*, 25, 49-52 (1989).

Minkova L., Lefterova E. and Koleva T., Thermal stability of irradiated LDPE-i-PP films, *Polym. Deg. Stab.*, 37, 247-252 (1992).

Nygard P.L., Jordy D.W., Knight T.W. and Vedder D.A., Process for extrusion and blowing of polyethylene film, *US Patent*, 005129 (2005).

Patel R., Hermel-Davidock T., Paradkar R., Landes B., Liu L., Demirors M. and Anderson K., Characterization of amorphous and crystalline orientation in polyethylene films, *ANTEC,* Charlotte (NC), 612 (2006).

Robledo-Ortiz J.R., Ramirez-Arreola D., Gonzalez-Nuñez R. and Rodrigue D., Effect of freeze-line position and stretching force on the morphology of LDPE-PA6 blown films, *J. Plast. Film Sheet.*, 22(4), 287-314 (2006).

Salamone, J.C. *Polymeric Materials Encyclopedia*, CRC Press: USA (1996).

Simpson D.M. and Harrison I.R. Study of the effects of processing parameters on the morphologies and tensile modulus of HDPE blown films: application of composite theories on a molecular level to characterize tensile modulus, *J. Plast. Film Sheet.*, 10, 302-325 (1994).

Utracki L.A. *Commercial Polymer Blends*, Chapman and Hall: UK (1998).

Wang Z., Chan C., Zhu S.H. and Shen J., Compatibilization of polystyrene and low density polyethylene blends by a two-step crosslinking process, *Polymer*, 39(26), 6801-6806 (1998).

Zhang X., Ajji A. and Verilhac J.M., Processing-structure-properties relationship of multilayer films. 1. Structure characterization, *Polymer*, 42, 8179-8195 (2001).

Chapter 6

A Bootstrap-VIP Approach for Selecting Wavelength Intervals in Spectral Imaging Applications*

Résumé

Une méthode de ré-échantillonnage basée sur la régression par variables latentes, nommée ici PLS Bootstrap-VIP, est proposée comme outil de sélection de longueurs d'ondes. Cette méthode a l'avantage de combiner simplicité et rapidité de calcul, ce qui la rend particulièrement utile pour l'identification d'intervalles spectraux pertinentes au sein des images hyperspectrales. Elle a été testée sur quatre jeux de données (captés par des caméras visible et proche infrarouge) provenant des industries du plastique, de l'huile, ainsi que des pâtes et papier. Les résultats obtenus en utilisant les VIP ont été comparés à ceux obtenus en utilisant les coefficients de régression PLS, ainsi qu'à ceux obtenus par deux méthodes plus sophistiquées tirées de la littérature. Nos travaux montrent que la méthode PLS Bootstrap-VIP concorde avec ces méthodes plus complexes en identifiant un faible nombre d'intervalles spectrales congruentes, ce qui facilite l'interprétation des données tout en réduisant la forte corrélation présente entre les longueurs d'onde. Cette méthodologie parcimonieuse requiert parfois des ajustements mineurs améliorant légèrement les capacités prédictives du modèle.

Abstract

A PLS Bootstrap-VIP approach is proposed as a simple wavelength selection method, yet having the ability to identify relevant spectral intervals. This approach is particularly attractive for wavelength selection within hyperspectral images due to its simplicity and relatively low computational cost compared to more sophisticated interval search methods. The method was tested on four visible-NIR spectral imaging datasets taken from the polymer, oil and pulp and paper industries. The results were compared with those obtained using PLS regression coefficients as well as two more sophisticated methods involving several metrics or the search for wavelengths intervals. It is shown that a small number of well-defined relevant spectral intervals are identified with the proposed approach, providing easy spectral interpretation in agreement with more complex interval search methods. Before final use, fine adjustments to the VIP threshold may be tested to verify whether predictive power can be improved.

* Gosselin R., Duchesne C. and Rodrigue D., A bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications, *Chemometr. Intell. Lab.*, 100, 12-21 (2010).

6.1 Introduction

Methods for relevant variable selection in PLS regression models have drawn considerable attention in the Chemometrics literature. These can be traced back to the work of Frank (1987) and Martens and Naes (1989) based on truncating the PLS weight vectors (**w**) or setting the smallest weights to zero and re-estimating the models in an iterative fashion. Since then, multivariate calibration problems had a significant contribution to new developments in variable, or more specifically, in wavelength selection algorithms. In the last two years, at least 10 new papers appeared on the subject (Ye et al., 2008; Cai et al., 2008; Krier et al., 2008; Cramer et al., 2008; ElMasry et al., 2008; Brás et al., 2008; Liu et al., 2008; Teófilo et al., 2009; Fang et al., 2009; Xu et al., 2009). The growing fields of applications of vibrational spectroscopy in chemistry and chemical engineering certainly stimulate interest in this area. In particular, recent developments in spectral imaging for on-line inspection and quality control in a variety of fields ranging from food, to agricultural, pharmaceutical and polymer processing (Kramer and Ebel, 2000; Kleynen et al., 2003; Yan et al., 2005; Cai et al., 2008; Ye et al., 2008), to name a few, led to massive amounts of spatial-spectral data (i.e. a complete spectrum for each pixel of the image) in a fraction of a second. More parsimonious models obtained by relevant wavelength selection may result in enhanced interpretability, increased robustness, and significant reduction in computation time, which is crucial for on-line spectral imaging applications.

The wavelength selection problem arises mainly from the fact that vibrational spectroscopic methods (i.e. NIR, IR, and Raman) and other techniques sharing similar features yields complex spectra made of overlapping absorption bands, interference from diffuse light scatter, and instrumental noise (Martens and Naes, 1989; Osborne et al., 1997; Cai et al., 2008). Hence, the response of functional groups or molecules is often dispersed over several adjacent wavelengths, leading to strong collinearity in some regions of the spectra, while other regions may be corrupted by noise, or more generally, may contain irrelevant information. A theoretical justification for wavelength selection in PLS regression models was provided by Spiegelman et al. (1998), who showed that removing irrelevant information is desirable regardless of the spectral noise distribution. Experienced

spectroscopists often have a good idea of the most important wavelength ranges that should be kept in calibration models. However, in the development of new products, when less theoretical work was carried out, or when a large number of chemical species are present in a sample, including some a priori unknown components, a systematic method for identifying relevant wavelengths or spectral regions may be of great help to the spectroscopist.

Methods for relevant variable/wavelength selection essentially fall into two categories: 1) methods based on *individual* wavelengths, and 2) approaches aiming at finding the most informative spectral *intervals*. The individual wavelength methods rank the importance of the *individual* wavelengths according to one or several metrics, directly or indirectly evaluating the correlation between each individual wavelength and one or many responses **Y**, and then using a cutoff criterion to segment relevant/irrelevant variables. The metrics used for ranking wavelengths include the magnitude of the PLS regression coefficients (Frenich et al., 1995; Centner et al., 1996, Osborne et al., 1997; Spiegelman et al., 1998; Lazraq et al., 2003; Ferreira et al., 2005; Preisner et al., 2007; Cai et al., 2008; Liu et al., 2008; Xu et al., 2009), the PLS weight vectors **W** (Frank, 1987; Lindgren et al. 1994; Jouan-Rimbaud et al., 1995; Liu et al., 2008), the correlation coefficients or the covariance (Jouan-Rimbaud et al., 1995), and more recently, the variable importance on the projection VIP (Lazraq et al., 2003; Chong and Jun, 2005; ElMasry et al., 2008). Several of these metrics were also jointly used by Teófilo et al. (2009). The cutoff criteria used are either determined arbitrarily, based on past experience, or through statistical assessment of uncertainties in the parameters using the Bootstrap, the Jackknife or the Monte-Carlo re-sampling methods (Centner et al., 1996; Lazraq et al., 2003; Ferreira et al., 2005; Preisner et al., 2007; Cai et al., 2008). Prediction ability criteria were proposed and assessed using variants of stepwise regression (Osborne et al., 1997; Spiegelman et al., 1998). Global search methods, such as the generalized simulated annealing (GSA) (Kalivas et al., 1989) and genetic algorithms (GA) (Ferreira et al., 2005; Preisner et al., 2007) were also used to identify the most informative wavelengths using some PLS model predictive ability criteria.

Interval methods, on the other hand, aim at finding the most informative group of wavelengths (i.e. spectral bands). This is motivated by the fact that functional groups absorb within a relatively short wavelength bands, which are easier to interpret (Nørgaard et al., 2000; Rossi et al., 2007), and should lead to more stable predictions (Höskuldsson, 2001). Intervals are essentially obtained by splitting the spectra into a certain number of intervals or through iterative construction of these intervals using, for example, moving windows. The most informative intervals are also assessed using various metrics and some cutoff criterion. One of the first interval approaches was proposed by Nørgaard et al. (2000). It consists of dividing the spectra into a certain number of equal width intervals (typically 20-30), and building a local PLS model on each (i.e. iPLS). The most important intervals to maintain good prediction performance are kept in the final PLS model. Finding the best combinations of these spectral intervals was investigated by Xu et al. (2007) using Monte Carlo cross-validation (MCCV) stacked regression. Brás et al. (2008) modified the iPLS approach by first selecting the most important regression coefficients using bootstrap intervals, followed by the construction of intervals based on a cutoff distance between selected coefficients. Orthogonal signal correction (OSC) approaches to wavelength selection, as well as more efficient approaches for defining the iPLS intervals using the most highly correlated wavelengths with **Y**, were also investigated by Höskuldsson (2001). Moving windows of tunable size were proposed as a method for searching the most informative combinations of wavelength intervals (Du et al., 2004; Cramer et al., 2008; Fang et al., 2009). The use of B-splines was also investigated for defining spectral intervals (Rossi et al., 2007; Krier et al., 2008), which involves careful selection of the number of nodes or width of the B-splines. A curve-resolution approach not relying on iterative variable search was proposed by Bogomolov and Hachey (2007). Intervals were defined as a fixed width around the known peaks of pure components spectra. Finally, a simple method based on the correlation between a given wavelength and its neighbors was also proposed for delimiting spectral regions (Ye et al., 2008). A representative wavelength (i.e. showing the highest correlation coefficient) was selected for each interval.

Both families of approaches have advantages and drawbacks. The individual wavelength ranking methods make use of simple metrics (regression coefficients, **W**, VIP, etc.), most

of which are parameters of the PLS model structure and are readily available in commercial software. Since most of these approaches are not formulated as direct search methods, they are generally less computationally intensive. However, it is argued that the results are more difficult to interpret since selected wavelengths are often distributed across the complete spectra instead of within a few confined intervals (Ferreira et al., 2005; Rossi et al., 2007). This issue will also be demonstrated in this paper using ranking of regression coefficients. Interval methods, on the other hand, directly identify the most informative wavelength bands and should lead to easier interpretation and more consistent results. However, intensive iterative search is often required.

In this paper, a Bootstrap-VIP (variable importance on the projection) approach is investigated for variable/wavelength selection. It consists of computing the VIP scores of each wavelength, obtained from a PLS model minimizing the root mean squared error in cross-validation (RMSECV). Bootstrap re-sampling is then applied to generate consistency intervals around VIP estimates. Those wavelengths significantly greater than a threshold (e.g. 1.0) are identified as relevant. This approach combines some of the advantages of both individual wavelength and interval selection methods and could, when more precise results are required, provide a very good starting point for interval search methods. Indeed, the VIP metric and the Bootstrap or Jackknife re-sampling methods are simple and already available in Chemometrics software, and as will be shown in this paper, the approach naturally leads to the identification of relevant spectral intervals. The choice of the VIP metric was partly motivated by the simulation study performed by Chong and Jun (2005), showing that using VIP scores for selecting the most informative wavelengths was more sensitive and specific than using regression coefficients across a wide range of data structures.

There is certainly no unique “best” approach to variable/wavelength selection for any data structure, and the authors fully recognize the difficulty to generalize conclusions obtained with a single method tested on a few datasets. However, the proposed approach is found to work well on four datasets collected in the polymer, oil and pulp and paper industries. A comparison is made with selection by ranking of PLS regression coefficients, one of the

most commonly used selection technique. It is clearly shown that for comparable RMSECV, Bootstrap-VIP identifies relevant intervals whereas regression coefficients tend to spread relevant individual wavelengths across the entire spectral range. Moreover, the proposed approach is found to provide very similar results to more complex interval selection methods on one of the datasets, the diesel shootout example, which was used in a few other papers (Rossi et al., 2007; Teófilo et al., 2009).

6.2 Methods

6.2.1 Projection to latent structures

Consider two data matrices, **X** (*N*×*K*) and **Y** (*N*×*M*), containing *K* spectral signatures from some spectroscopic method and *M* response variables (e.g. chemical concentration, mechanical properties, quality characteristics, etc.) for a set of *N* samples. The structure of the PLS calibration model defining a common latent variable space relating both data blocks is given by:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{6.1}$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \tag{6.2}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^{*} \tag{6.3}$$

$$\mathbf{W}^{*} = \mathbf{W}\left(\mathbf{P}^{\mathrm{T}}\mathbf{W}\right)^{-1} \tag{6.4}$$

where **W*** (*K*×*A*) contains the X-weights defining the common latent variable space **T** (*N*×*A*) relating **X** and **Y**, and *A* is the number of PLS components (i.e. latent variables). Cross-validation using 10 random subsets was used in this work for selecting the number of PLS components (*A*). The **P** (*K*×*A*) and **Q** (*M*×*A*) matrices contain the loading vectors that best represent the **X** and **Y** spaces, respectively. The **E** (*N*×*K*) and **F** (*N*×*M*) matrices contain the projection residuals. The PLS latent variable model can be re-organized as a simple prediction equation involving a set of regression coefficients $\mathbf{B}_{\mathbf{PLS}}$ (*K*×*M*), similarly as for multiple linear regression (Wold et al., 2004).

$$\hat{\mathbf{Y}} = \mathbf{X}\mathbf{B}_{\mathbf{PLS}} \tag{6.5}$$

$$\mathbf{B}_{\mathbf{PLS}} = \mathbf{W} * \mathbf{Q}^{\mathrm{T}} \tag{6.6}$$

The PLS regression coefficient $\mathbf{B}_{\mathbf{PLS}}$ has been one of the most frequently used metric for selecting individual wavelengths in a similar way as if the model between **X** and **Y** was a multi-linear regression model. Another commonly used metric for judging the importance of the X-variables on **Y** is the variable importance on the projection (i.e. VIP). The VIP score for the k^{th} variable (wavelength) for a single response **y** is computed according to Eriksson et al., (2001):

$$VIP_k = \sqrt{K \sum_{a=1}^{A} \left[\left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right) \left(w_{ak} / \left\| \mathbf{w}_a \right\|^2 \right) \right] \Big/ \sum_{a=1}^{A} \left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right)} \tag{6.7}$$

where $a = 1,2\ldots,A$, K is the number of columns of **X** (i.e. variables or wavelengths), w_{ak} is the loading weight of the k^{th} variable in the a^{th} component, and $\mathbf{t}_a$, $\mathbf{w}_a$, and q_a are the a^{th} column vectors of **T**, **W**, and **Q**, respectively.

The main difference between the $\mathbf{B}_{\mathbf{PLS}}$ and the VIP metrics, in terms of assessing the importance of each variable or wavelength in predicting **y**, lies in the way the VIP metric weights the contribution of each variable according to the variance explained by each PLS component. Indeed, the $q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a$ term in the numerator of Equation 6.7 is the variance of y explained by the a^{th} PLS component, whereas the summation of these terms over all A components (i.e. denominator in Equation 6.7) is the total variance explained by the model. The $w_{ak} / \left\| \mathbf{w}_a \right\|^2$ expression represents the importance of the k^{th} variable in the a^{th} component. As the number of PLS components increases, the variance explained by each component generally decreases, and the importance of the variables is progressively down-weighted. This limits the noise propagation (associated with smaller eigenvalue components) into the VIP metric. The latter should therefore be a more stable and consistent indicator of variable importance.

It is generally accepted in practice that variables having a VIP > 1.0 are highly influential, values between 0.8 < VIP < 1.0 indicate moderately influential variables, and VIP < 0.8 represent less important variables (Eriksson et al., 2001). The unity threshold comes from the fact that the average of the squared values of the VIPs is equal to 1.0. The greater-than-one rule is often used as a selection criterion (Eriksson et al., 2001; Olah et al., 2004; ElMasry et al., 2008). A systematic simulation study conducted by Chong and Jun (2005) indicates that this cutoff criterion should be a function of the data structure, and that proper cutoff values, maximizing sensitivity and specificity in finding the truly important variables, range between 0.83-1.21.

6.2.2 Bootstrap and wavelength selection

Bootstrap is a well-known re-sampling method used in statistics to estimate bias and standard errors of parameter estimates. Similarly to Monte Carlo simulations, the dataset is randomly re-sampled *b* times with replacement leading to *b* estimates of the parameters of interest after building a regression model on each of the *b* datasets. This method assumes that both the original dataset and the bootstrap subsamples are representative of the population. Moreover, the bootstrap subsamples should be balanced; i.e. that the subsets are to be evenly collected throughout the dataset (Miller, 1974). Larger values of *b* more efficiently treat unsmooth and non-linear data at the expense of increased computation time (Efron and Tibshirani, 1993). More information on the bootstrap method can be found in Shao and Tu (1995).

In this work, bootstrap was used to assess the importance (i.e. VIP) of each wavelength band of spectral images on the predictions of sample quality using a subsample size of $0.4N$, where N is the size of the original dataset. This falls within the $0.27N - 0.5N$ range recommended by Babu (1991). Re-sampling b=500 times was found to yield very consistent results. Such a large number of subsamples was selected to ensure that stable bootstrap uncertainty intervals were obtained for all four datasets. Wavelength selection was achieved through a PLS-bootstrap algorithm based on the variable importance on the

projection metric (VIP). The PLS-bootstrap approach for regression coefficients was also used for comparison since regression coefficients are commonly used for variable/wavelength selection.

The typical threshold for VIP based variable selection is the greater-than-one rule (Eriksson et al., 2001; ElMasry et al., 2008). This criterion is based on the assumption that variables with VIPs above the average (1.0 as previously mentioned), should be retained in the final prediction model. The PLS-bootstrap algorithm was used to assess the uncertainty in the VIP metrics, and a wavelength was considered relevant when its average VIP value, along with its one standard deviation error bar (obtained from the bootstrap), was above 1.0 as shown in Figure 6.1. A similar methodology was applied to regression coefficients ($\mathbf{B_{PLS}}$) i.e. they were considered significant when their one standard deviation error bar did not include 0.0. Bootstrap results for one dataset are presented in Figure 6.1 to illustrate the typical behavior of the VIP and regression coefficients when used for wavelength selection. The VIP plot shows a fairly smooth spectrum with a certain number of well-defined peaks, while the $\mathbf{B_{PLS}}$ curve is much noisier with a few peaks clearly standing out. The smoother VIP results are likely related to the weighing of each variable contribution according to the variance explained by each PLS.

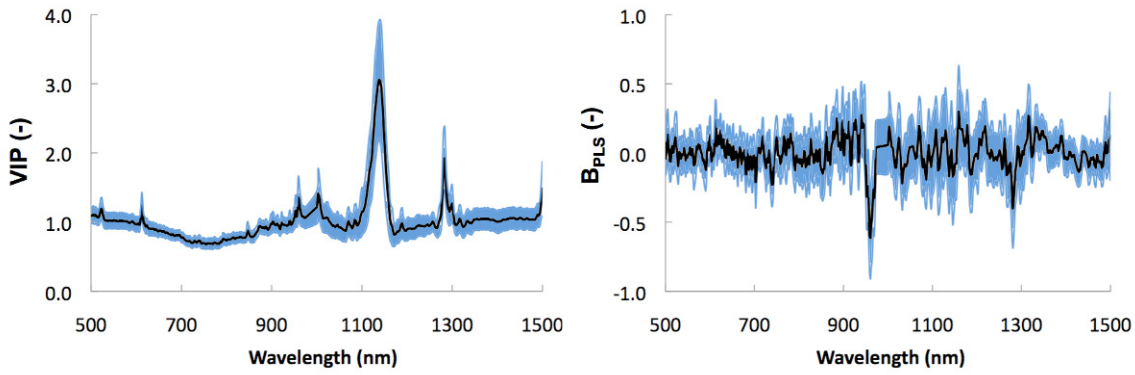

Figure 6.1. Bootstrap results at each wavelength illustrating the bootstrap average value and standard deviation for VIP and B_{PLS} metrics. The PLS model built on dataset 2 with 10 latent variables is shown.

6.3 Experimental

6.3.1 Spectral image acquisition

The line-scan imaging system used in this work is composed of two cameras centered on the same area of a sample, each coupled with a grating spectroscope, effectively covering the visible and near-infrared (NIR) range of the spectrum (475-1500 nm). The visible range is captured using a Hamamatsu C8484-05G camera (480×512 pixels) coupled with an ImSpector V10E grating spectroscope (Specim Spectral Imaging Ltd) effectively sensing the 475-1000 nm wavelength range. The infrared range is captured using a XenICs XEVA-USB-FPA camera (256×320 pixels) coupled to a Specim ImSpector N17E grating spectroscope effectively sensing between 1000-1500 nm. Each scanned line of the sample is captured on the CCD array as a spatial/spectral ($x \times \lambda$) image containing hundreds of spectral channels (i.e. wavelengths), rather than the 3 more commonly obtained in RGB images. A hyperspectral image, or data cube, of the entire sample is achieved by repeatedly displacing the sample perpendicularly to the spectroscope slit (along the y-axis). Each hyperspectral image consists of a number of juxtaposed line-scans, at 1 mm intervals, forming a 3-D array ($x \times y \times \lambda$).

6.3.2 Description of the datasets

Four datasets of spectral images and corresponding quality measurements were used to test the proposed method on different materials: (1) polymer films, (2) wood/plastic composites, (3) Kraft liquors and (4) diesel samples. Each dataset is explained in details below. These four examples were chosen to verify the behavior of the method in various fields (polymer processing, pulp and paper, and energy). This allowed verifying whether the results obtained with the VIP-bootstrap approach, and the comparison with other metrics (i.e. $\mathbf{B_{PLS}}$), yields consistent conclusions across different fields and data structures. More specific points may also be addressed, such as whether the same wavelengths are recognized as significant in both polymer datasets, as one may expect.

Dataset 1: polymer films

A compression press was used to mould thin polymer films (100×40×1.5 mm) of either polypropylene (PP), high-density polyethylene (HDPE) or low-density polyethylene (LDPE). The batch process consists in molding 6 g of resin at 180°C under 3 tons of pressure for 10 minutes, long enough to remove any previous thermal history (Chew et al., 1989; Albano et al., 2003). The mould was then water cooled to release the sample. A total of 18 samples were produced, 6 for each polymer, using average cooling rates of 1, 2, 4, 5, 8 and 16°C/min. These changes in cooling rates produced samples of different overall crystallinity, measured using Differential Scanning Calorimetry (DSC). The samples were imaged using the NIR line-scanner (185 wavelengths). A total of 10 line-scans were captured for each of the 18 samples (3 polymer types and 6 cooling rates) and were averaged along one spatial axis leading to a spectral matrix (**X**) of dimensions (180×185). Degree of crystallinity was measured for each sample and these measurements were repeated 10 times to obtain a congruent response vector **y** (180×1). More details on this study are available in Gosselin et al. (2008).

Dataset 2: wood/plastic composites

Thin HDPE/wood composites were produced using a twin-screw extruder and tested experimentally under different processing conditions and wood contents. The continuously extruded samples measured 2 mm in thickness and 80 mm in width. They were made of sieved birch wood fibers (45 and 180 μm) and commercial HDPE (Fortiflex A60-70-162). Wood fibers were obtained from a local sawmill. The following process conditions were investigated in a complete 2^3 factorial design: screw speed (70 and 90 rpm), die temperature (160 and 180°C), polymer throughput (500 and 600 g/h) with a fixed wood content of 20 wt.%. Wood content (main feed disturbance) was also varied in a separate set of experiments at three levels (10, 20 and 30 wt.%) under constant nominal process conditions (i.e. center point of 2^3 design). Steady-states were allowed to run for about 10 minutes for each condition tested. The samples were scanned using both the visible (485 wavelengths) and near-infrared (185 wavelengths) imaging spectrometers. In this case 25

line-scans were collected for each of the 11 samples. These were again averaged along the machine cross-direction to yield the spectral matrix **X** (275×485). Sample crystallinity is again used as the quality variable **y** (275×1) measured by DSC. Further details are a provided in Gosselin et al. (2009).

Dataset 3: Kraft liquors

Kraft pulp and paper mills use the so-called white liquor, a mixture of sodium hydroxide and sodium sulfide, to extract lignin from wood fibers. The mixture of white liquor and lignin is known as weak black liquor. Recycling this liquor is achieved by concentrating the black liquor and burning off the excess solids, leaving behind a solution of sodium carbonate, sodium sulfate and sodium sulfide. At this stage, the blend is known as green liquor. The cycle back to white liquor is completed by processing the green liquor in a lime kiln. A preliminary study was performed to test whether spectral NIR images of liquor samples could be used to predict their total sulfide (S^{2-}) concentration (**y**) at any point during the cycle. A total of 15 samples were obtained from Irving Pulp & Paper (St-John, New-Brunswick, Canada), 5 of each liquor type (weak black, clarified green, and clarified white liquors). These were selected to span a wide range of sulfide concentration (i.e. S^{2-}). Previous works by Leclerc (1997) and Mustonen and Laszlo (2007) have shown that the sulfidity, or S^{2-} concentration, of Kraft liquors can be predicted using near-infrared spectra. Thus, 10 line-scans were captured using the NIR spectrometer for each of the 15 liquor samples and were averaged along one direction leading to the input matrix **X** (150×185). Sulfidity (**y**) was measured for each sample in the analytical chemistry lab of the company.

Dataset 4: diesel

In addition to the 3 datasets obtained in our laboratory, a fourth dataset was tested in order to directly compare the results of the proposed wavelength selection strategy to those obtained by other groups using different, but more complex methods. This dataset consists of NIR spectra of diesel samples measured at the Southwest Research Institute (SWRI) and was obtained from the Eigenvector Research Incorporated website (2005). Using cetane

number (equivalent to octane number for gasoline) as a quality variable, the proposed method will be compared to the results published by Rossi et al. (2007) and Teófilo et al. (2008) on the same dataset. The spectral range captured in this dataset was from 750 to 1550 nm (discretized into 401 wavelengths) and had previously been mean centered and reduced to unit variance. After removing a few outliers from the original set, the spectral matrix **X** (107×401) was obtained. Cetane number (**y**) was measured for each sample.

Finally, each dataset was mean-centered and scaled to unit variance (i.e. autoscaled) prior to be used in PLS modeling.

6.4 Results

6.4.1 Number of PLS components and wavelength selection consistency

The number of PLS components used in the calibration model have an important impact on both the number and the consistency of selected wavelengths. The number of relevant wavelengths based on the Bootstrap-VIP and the Bootstrap-B (i.e. regression coefficients) methods as a function of the number of PLS components is shown in Table 6.1 for dataset 1. The number of intervals is also provided. An interval is defined by consecutive significant wavelengths. A single wavelength was also considered an interval if it was both identified as significant and isolated from other significant wavelengths. These results show that both the number of wavelengths and spectral intervals seem to stabilize when a sufficient number of latent variables (LV), or PLS components is used. In dataset 1, the VIPs appear to stabilize with at least 5 latent variables, while this number tends toward 10 for the regression coefficients. Similar results were observed for the other datasets. It therefore appears that the VIPs require fewer LVs to yield stable results. This was expected due to the down-weighing of the latent variables associated with smaller eigenvalues (see Equation 6.7).

Table 6.1. Number of significant spectral bands and wavelengths obtained for dataset 1 using increasing numbers of latent variables.

Method	LV	Intervals	Wavelengths
VIP	1	5	58
	3	4	40
	5	7	42
	10	6	41
	20	6	41
B_{PLS}	1	2	160
	3	7	139
	5	12	115
	10	16	74
	20	14	68

While the number of significant wavelengths and spectral intervals seems to stabilize when enough components are used, this does not indicate whether the same wavelengths are being selected in each case. To investigate the consistency of the results provided by each selection method, 15 PLS models were built on dataset 1 with different number of components (1-10, 12, 15, 20, 25, and 30) and, for each of them, relevant wavelengths were selected using the bootstrap-VIP and bootstrap-B approaches. The ratio of number of times each individual wavelength was found relevant by the VIP and B approaches for the 15 PLS models is shown in Figure 6.2. A ratio near 1.0 for a given wavelength indicates that all 15 models retained it as significant, while a ratio close to 0.0 indicates that none of the models did. The VIP selection show drastic changes along the spectra; some wavelengths are nearly always identified as significant while others, even neighboring wavelengths, are always rejected. Moreover, the selected wavelengths often occur in bands, or spectral intervals, as one would expect from correlated spectral data. This indicates that Bootstrap-VIP selection consistently identifies similar spectral regions, with little variations as the number of latent variables is changed. On the other hand, selection based on the regression coefficients follows a very different behavior: very few well-defined spectral intervals are identified, or even clear patterns which could lead to the identification of significant spectral regions. The selected wavelengths almost appear randomly distributed within the spectral range, and vary widely as a function of the number of PLS components since most of the wavelengths were found relevant about 40-60% of the time (i.e. ratios between 0.4-

0.6). Very similar results were obtained on other datasets. The consistency of the Bootstrap-VIP approach was established in this section. Interpretability is investigated next.

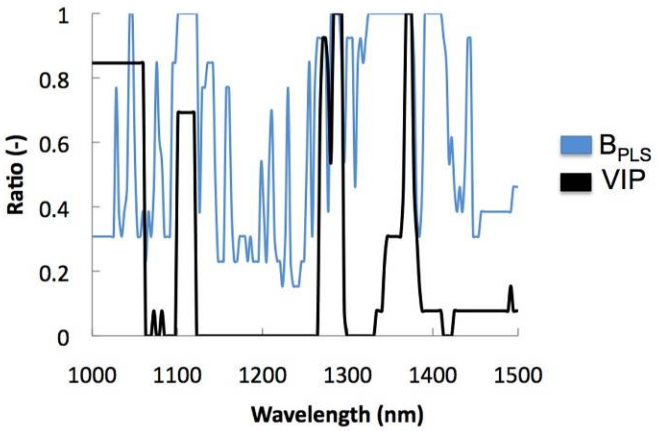

Figure 6.2. Influence of the number of PLS latent variables on wavelength selection consistency illustrated with dataset 1. Shown is the ratio of the number of times each wavelength was found significant by the VIP and regression coefficients (B_{PLS}).

6.4.2 Wavelength selection and interpretation

As shown in Table 6.1, the number of wavelengths and spectral intervals predicted by the models stabilizes when a sufficient number of latent variables are used. Nevertheless, keeping too many latent variables may lead to increased prediction errors. Several criteria for selecting the optimal number of latent variables have been published: the normalized mean squared error (*NMSE*) (Rossi et al., 2007), the relative root mean squared error (*RRMSE*) (Ye et al., 2008) and the root mean squared error in cross-validation (*RMSECV*) (Brás et al., 2008; Cai et al., 2008; Cramer et al., 2008; Teófilo et al., 2008). The latter criterion was chosen in this work and is defined as:

$$RMSECV = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(y_n - \hat{y}_n\right)^2} \qquad (6.8)$$

where y_n and $\hat{y}_n$ are the measured value and predicted response value in cross-validation of the n^{th} sample in a dataset containing N samples.

In practice, wavelength selection was achieved through a two-step procedure. First, a PLS model was built using the complete spectral matrix **X** and **y** (Section 6.3.2) in which the number of latent variables (LV_1) was computed to minimize the *RMSECV* achieved by the PLS model. For all datasets used in this study, the *RMSECV* first decreased with an increasing number of components, reached a minimum with less than 10 components, and then increased monotonously. Applying this method to the datasets, 8 latent variables are used in set 1 while sets 2, 3 and 4 required 7, 9 and 5 latent variables, respectively. Once the number of components was determined, the PLS-bootstrap method was applied to the data using the number of latent variables obtained in the first step (LV_1).

One objective of this study was to demonstrate that the stability and consistency of the VIP approach naturally leads to the identification of contiguous spectral intervals rather than spreading selected wavelengths throughout the spectra as Bootstrap-B does. Given the continuity of the spectra, these intervals are not only used by spectroscopists, but they are also much easier to interpret qualitatively (Brás et al., 2008). The relevant wavelengths identified by both PLS-bootstrap methods (i.e. VIP and B) with LV_1 latent variables are shown in Figure 6.3 for all four datasets. For convenience, the wavelengths found significant by the VIP metric are shown below the mean spectra whereas those selected by the regression coefficients are identified above the spectra. As noted previously, the VIPs identify a relatively small number of spectral intervals, while the regression coefficients B_{PLS} leads to the selection of a large number of individual wavelengths, distributed throughout the entire spectral range.

As illustrated in Figure 6.3, the PLS-bootstrap method based on the VIPs identified a number of spectral intervals. Four such intervals were identified in dataset 1 (polymer films). The first two intervals (1000-1060 nm and 1100-1120 nm) represent the C-H stretching second overtone band and are linked to CH_2 and CH_3 vibrations (Watari et al., 2004). Polymer crystallinity refers to the degree of structural organization within the material. High crystallinity levels reduce the freedom of the macromolecules and have been linked to decreased CH_2 vibrations in the C-H stretching second overtone band (Bower and

Maddams, 1992). The third and fourth intervals (1270-1290 nm and 1365-1375 nm) represent the first overtone combination band (Ren et al., 1999; Watari et al., 2004). This band represents combined stretching and vibrational movements linked to polyethylene crystallinity (Watari et al., 2004).

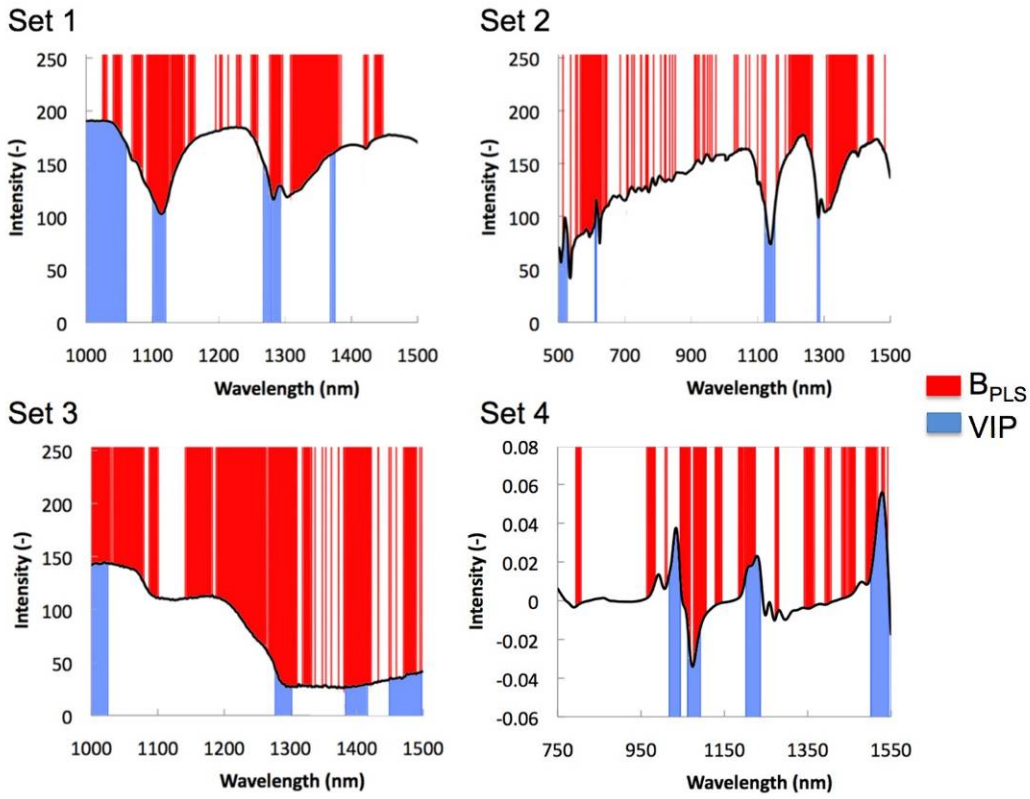

Figure 6.3. Relevant wavelength selection results for all datasets. Black line: average spectra. Blue bars (below spectra): VIP selection. Red bars (above spectra): B_{PLS} selection (regression coefficients).

Four intervals were identified in dataset 2 (wood/plastic composites). The first two (475-525 nm and 600-615 nm) represent blue-green and orange wavelengths, respectively. Note that the small but sharp peaks located around 546 and 612 nm are associated with room fluorescent lighting. Nevertheless, further tests discussed in Section 6.4.3 indicate that the spectral range between 500-620 nm may have some significance. This spectral interval is known to be linked to a number of wood species (Pandey et al., 1998; Pastore et al., 2004) and appears to be associated to the presence of cellulose (Ilharco et al., 1997). The two remaining intervals of importance (1110-1155 nm and 1275-1300 nm) essentially

correspond to the second and third intervals discussed in the 1st dataset (1100-1120 nm and 1270-1290 nm). It is important to note that very similar spectral bands around 1110 and 1275 nm were classified as being significant in both dataset 1 and 2 as one would expect when similar chemical compounds are present. This indicates once again that the VIPs lead to consistent spectral intervals identification.

Four intervals were identified in dataset 3 (Kraft liquors). In this case however, even if the VIP selection led to very well defined intervals, interpretation of the first three intervals (1000-1025 nm, 1280-1300 nm, and 1380-1420 nm) was not straightforward due to the very large number of chemical species present in the liquor samples. Both lignin and sodium sulfide, for example, strongly absorb in the 1000-1400 nm range (Huang et al., 2008; Muttiah, 2002). On the other hand, the fourth interval (1450-1500 nm) seems to be linked with sodium hydroxide concentration (Heiman and Licht, 1999). This component is known to vary with the liquor composition.

Four spectral intervals were also identified in dataset 4 (diesel). These are located in the following ranges: 1018-1044, 1062-1092, 1202-1236, and 1502-1544 nm. According to Rossi et al. (2007), these intervals correspond to absorption bands of hydrocarbons whose combustibility influences the cetane number. At this point, we remind the reader that the diesel dataset is part of the public domain and was chosen to compare the proposed wavelength selection method to more complex techniques previously published. However, one must keep in mind that the results will depend on the initial pre-processing of the dataset as well as criterion with which outlying samples are excluded. Accordingly, Rossi et al. (2007) identified intervals located at 974-1086, 1200-1262 and 1486-1550 nm as the most relevant ones using a functional representation of the spectra (i.e. B-splines), followed by forward-backward procedure using the mutual information metric for selecting the B-splines coefficients that are the most highly correlated with **Y** and, finally, building a non-linear neural network model between the selected coefficients and the responses **Y**. A second research group also identified the following 3 intervals as the most relevant ones: 1000-1125, 1225-1325, and 1475-1550 nm (Teófilo et al., 2008). These results were obtained by constructing a set of informative vectors based on 7 metrics (regression

coefficients, VIPs, signal/noise ratios, etc.), taken separately or in pairs, followed by an iterative search for the best combination minimizing the *RMSECV* of the PLS calibration model. The close agreement with some of the more complex methods confirms the ability of the Bootstrap-VIP algorithm to clearly identify meaningful spectral intervals. The predictive ability of the resulting more parsimonious PLS models will be investigated in the next section.

6.4.3 Predictions with reduced spectra

The predictive ability of the more parsimonious PLS models, obtained after wavelength selection, is investigated in this section. Ideally, the reduced model should maintain or improve the predictive power of the model using the complete spectra. The proposed wavelength selection is a two-steps process involving: 1) building a PLS model on the complete spectra with LV_1 latent variables, selected to minimize *RMSECV*, followed by relevant wavelength selection using the Bootstrap-VIP metric, and 2) re-estimation of the PLS model on the reduced spectral matrix. The number of latent variables of the final PLS model (LV_2) is again determined to minimize *RMSECV*. The number of latent variables selected in each step as well as the predictive power of the final PLS model, quantified by *RMSECV* and Q^2 statistics, are shown in Table 6.2 for all four datasets. Note that selection results using the VIP metric without bootstrapping are also presented in Table 6.2, but are discussed later in Section 6.4.3.

Comparing the results obtained by using the complete spectra and the bootstrapped VIP and B_{PLS} metrics (see boot-VIP and boot-B in Table 6.2) it is observed that: 1) the boot-VIP approach always selects much fewer wavelengths (or spectral intervals), 2) wavelength selection using bootstrap based approaches leads to a reduction of up to 2 components in the final PLS model (LV_2) compared to the model based on the complete spectra (LV_1), and 3) boot-VIP have a slightly lower predictive ability. As discussed earlier, the smaller number of boot-VIP selected wavelengths gathered into well-defined spectral bands leads to easier interpretation. However such a high degree of parsimony seems to be at the limit of loosing predictive power, but still maintains the **y**-relevant latent structure contained

within the complete spectra (i.e. LV_2 close to LV_1). This approach might in fact establish a lower bound on the number of wavelength intervals or, when considering their location within the spectral range, establish the minimal amount of information required for interpretation and prediction of **y**.

Table 6.2. Predictive ability of the final PLS models obtained after wavelength selection by VIP (with and without bootstrapping) and B_{PLS} metrics, compared against the performance obtained with the complete spectra. The number of selected wavelengths and intervals as well as the number of latent variables in the selection (LV_1) and final (LV_2) steps are also shown. Note that LV_1 corresponds to the number of components of the PLS model on the complete spectra.

Set	Spectral matrix (**X**)	Wavelengths (Intervals)	LV_2	*RMSECV*	Q^2_{cum} (%)
1	Complete	185(-)	8	3.038	97.1
	Boot-VIP	41(4)	7	3.471	96.1
	Boot-B	81(14)	7	2.613	97.8
	VIP	59(6)	6	3.405	96.2
2	Complete	485(-)	7	2.570	84.5
	Boot-VIP	51(4)	6	2.686	84.3
	Boot-B	233(32)	7	2.491	85.4
	VIP	196(10)	7	2.600	84.4
3	Complete	185(-)	9	0.229	88.0
	Boot-VIP	53(4)	7	0.259	85.2
	Boot-B	130(18)	8	0.229	87.8
	VIP	78(4)	7	0.250	85.6
4	Complete	401(-)	5	2.011	68.0
	Boot-VIP	71(4)	6	2.057	66.6
	Boot-B	114(16)	4	2.163	63.7
	VIP	82(3)	6	2.041	67.1

This suggests that after the interpretation phase is completed, the boot-VIP reduced model might require some fine-tuning in order to achieve the lowest RMSE. This could be done in a number of ways. Some of them are listed below:

1. Progressively increase the width of the boot-VIP selected spectral bands to verify whether any improvements could be achieved without adding other intervals.

2. Reduce the VIP cut-off criterion from 1.0 to 0.8 or lower if necessary. This will increase the number of selected wavelength intervals. Some of them will likely be

located in other section of the spectral range. This operation does not require redoing the bootstrap procedure, but only verifying which wavelengths have their bootstrap intervals above the new cutoff values.

3. Alternatively, one could take advantage of the boot-B results which was found to spread relevant wavelengths across the entire spectral range (i.e. interpretation not as straightforward), but achieved better prediction results (see Table 6.2). The latter is obtained with virtually no additional computation cost when the boot-VIP procedure is performed. Wavelengths or intervals selected by boot-B, but not by boot-VIP, could be sequentially added to the boot-VIP model in a step-wise procedure. One could also cast this procedure into a multivariate design of experiments problem (Wold et al., 1986) to verify which additional intervals or combinations of intervals could be added to the boot-VIP model.

4. The spectral intervals selected by boot-VIP could also be used as a very good starting point for more sophisticated interval search methods (Nørgaard et al., 2000; Du et al., 2004; Rossi et al., 2007; Cramer et al., 2008; Fang et al., 2009).

To further validate that the boot-VIP is likely to select the minimum number of wavelength for prediction purposes, a randomization experiment was conducted. Given the high degree of redundancy within the spectra (i.e. collinearity), if boot-VIP selects the minimum number of wavelength intervals (λ_{min}), then randomly selecting λ_{min} wavelengths across the entire spectral range should yields relatively similar predictive ability. It should be emphasized here that this holds true only due to the high degree of collinearity within the spectra. Indeed, it is likely that random selection would pick wavelengths within some of the relevant intervals if a sufficient number of wavelengths (i.e. λ_{min}) is used. For each dataset, a total of 10 random selections were made of 41, 51, 53, and 71 wavelengths for datasets 1-4, respectively. A PLS model was built on each selection (i.e. 10 random selections × 4 datasets) and their number of latent variables was again selected to minimize *RMSECV*. The mean *RMSECV* and Q^2_{cum} statistics, as well as their standard deviations, are reported in Table 6.3 with the boot-VIP results (taken from Table 6.2).

Table 6.3. Predictive capabilities of purely random wavelength selection using the same number of wavelengths as PLS-bootstrap methods based on VIPs. Ten random replicates were performed for standard deviation.

Dataset	Wavelengths	Selection method	*LV*	*RMSECV*	Q^2_{cum}
1	41	Boot-VIP	7	3.471	96.1
		Random	7-10	3.703 (0.276)	95.6 (0.6)
2	51	Boot-VIP	6	2.686	84.3
		Random	6-7	2.651 (0.035)	83.5 (0.4)
3	53	Boot-VIP	7	0.259	85.2
		Random	7-9	0.283 (0.017)	81.4 (2.2)
4	71	Boot-VIP	6	2.010	66.6
		Random	6-8	2.156 (0.024)	65.0 (0.6)

Prediction results for random selections are very similar to those obtained through the boot-VIP selection method. In many cases, the number of latent variables, the Q^2_{cum} and the *RMSECV* of the proposed methods fall within one standard deviation of the random selection results. This supports the idea that boot-VIP might very well identify the minimum number of wavelengths below which predictive ability degrades. The randomization experiment was used here only to provide additional support for this idea. Of course, random selection of the wavelength would not systematically lead to clear interpretations such as those provided by the boot-VIP approach.

6.4.4 Advantages of bootstrapping

The Bootstrap-VIP approach uses a re-sampling algorithm to assess the consistency of the VIP metrics, which comes at a computational cost. To justify this added layer of complexity, a comparison was made with the VIP selection using the greater-than-one rule, but without bootstrapping. These results are shown in Table 6.2 (see rows labeled Boot-VIP and VIP). It is clear that using the bootstrapped VIP metrics further reduces the number of selected wavelengths and leads to slightly lower prediction power than without bootstrap (VIP greater-than-one rule only). The VIP-selected wavelength intervals with and without bootstrapping are presented in Figure 6.4.

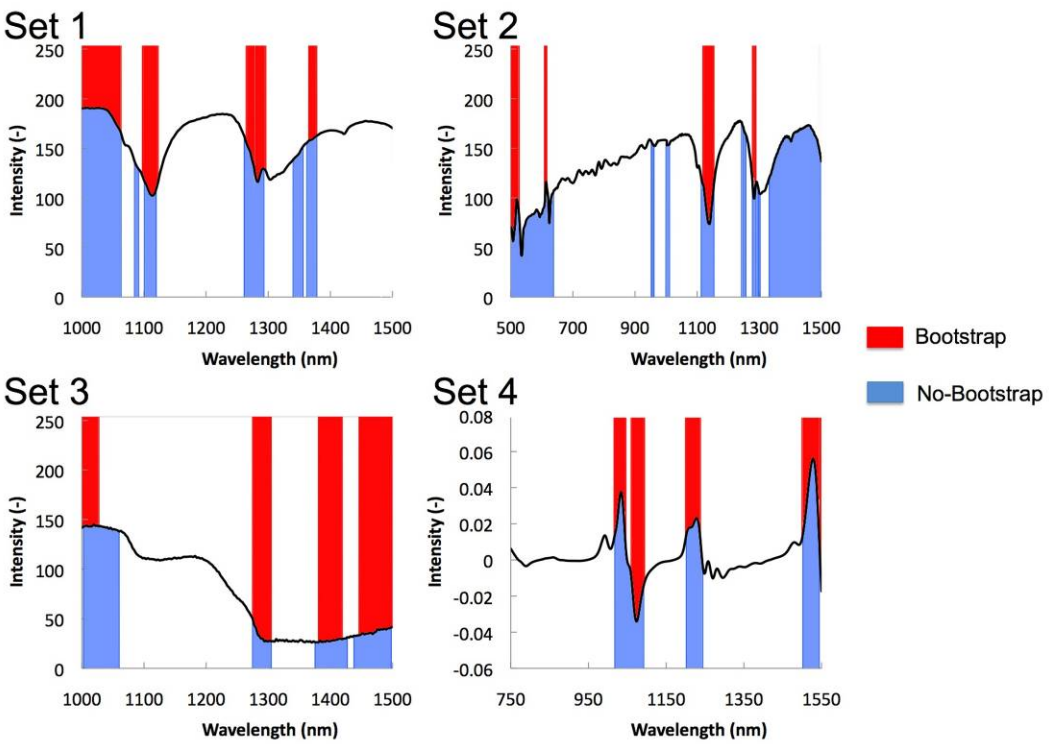

Figure 6.4. Impact of bootstrapping on wavelength selection using the VIP metric. Black line: average spectra. Blue bars (below spectra): No-bootstrap selection. Red bars (above spectra): Bootstrap selection.

The wavelength bands selected by the VIP metric with and without bootstrapping method differ to a certain extent. Overall, the boot-VIP method identifies both fewer and narrower spectral intervals. This difference is due to the fact that the bootstrap standard error of these spectral intervals included the threshold values of 1.0. However, fine-tuning the model by decreasing the cutoff values towards 0.8, with bootstrapping, would have led to similar results. The bootstrap consistency intervals around the VIP estimates allow discriminating wavelengths of borderline significance that may slightly improve prediction power, but with the advantage of providing results that are easier to interpret due to fewer, but more consistent wavelength bands. As mentioned in the previous section, it is recommended that the Boot-VIP approach should be used first for interpretation purposes, and then fine-tuned for improving prediction ability.

6.4.4 Discussion on the VIP threshold

Until now, the usual VIP threshold of 1.0 was used for selecting wavelengths. This cutoff value is not absolute as the zero cutoff (i.e. $\mathbf{B_{PLS}}$ significantly different from 0) used to assess the significance of regression coefficients. In fact, it has often been suggested that the proper cutoff may vary between 0.8 and 1.2 (Eriksson et al., 2001; Olah et al., 2004; Chong and Jun, 2005). The impact of the VIP cutoff value on the PLS model predictive ability is investigated for each dataset. A comparison is provided in Table 6.4 between the Boot-VIP selection with a threshold of 1.0 and the optimal cutoff value leading to the lowest *RMSECV*, identified by trial and error. Even if minimizing *RMSECV* is in fact a 2-dimensional optimization problem, requiring to select both the number of latent variables of the reduced model (LV_2) and the optimal VIP cutoff, the number of latent variables was kept the same (i.e. for each dataset) to test the individual impact of the VIP cutoff on selected wavelengths.

Table 6.4. Influence of VIP cutoff on the reduced PLS model obtained after wavelength selection. The achieved RMSECV statistics for the optimal VIP cutoff (minimizing RMSECV) is compared to the greater-than-one rule.

Dataset	Optimal cutoff value	*RMSECV* at optimal cutoff	*RMSECV* Cutoff = 1
1	0.95	3.389	3.471
2	0.90	2.600	2.686
3	0.85	0.251	0.259
4	1.10	2.010	2.011

The results presented in Table 6.4 show that the optimal VIP cutoff values vary between 0.85 and 1.10, providing marginal improvements on *RMSECV* values compared to the 1.0 threshold (reduction of 2.2% on average). These observations indicate that a VIP cutoff value of 1.0 is a good starting point, but can later be fine-tuned if necessary.

The dependence of the VIP cutoff on the scaling of spectral data needs to be investigated. The greater-than-one rule has a particular meaning for autoscaled data which may not hold for other scaling options. However, after repeating the analysis of datasets 1-3 using mean-centering only, it was found that the VIP=1.0 threshold was still a good initial value.

Similar results were obtained in terms of both the predictive power and the selected spectral ranges. Since these conclusions can hardly be generalized, it is recommended, in any case, to test other threshold values to verify whether a better model can be obtained.

6.5 Conclusion

This study explored the ability of the Bootstrap-VIP method for wavelength selection in spectral imaging applications and spectroscopy in general. The proposed two-steps approach involves building a first PLS model using the complete spectra database, with as many components as necessary for minimizing *RMSECV*. The bootstrap procedure is applied to that model in order to compute consistency intervals around the VIP metrics, followed by selection of the wavelengths having their intervals above a certain cutoff value. The greater-than-one rule was used in this paper. A final PLS model is then built on the reduced spectra. The proposed approach was tested on four different datasets, most of them using the NIR spectral range. Comparisons were made with the ranking of PLS regression coefficients, one of the most commonly used approach for selecting wavelengths individually (Centner et al., 1996; Spiegelman et al, 1998; Lazraq et al., 2003; Brás et al., 2008), and with some of the more sophisticated methods involving several metrics in informative vectors (Teófilo et al., 2009) or searching for wavelengths intervals (Rossi et al., 2007).

The main conclusions drawn from this work are as follows: 1) even if the bootstrap-VIP approach is essentially an individual wavelength selection method, it naturally leads to the identification of the most relevant spectral intervals related to **Y**, whereas selection by ranking of regression coefficients spreads relevant wavelengths across the spectra; 2) the reduced spectra obtained by the bootstrap-VIP are therefore much easier to interpret and this was demonstrated on the four datasets used in this work (i.e. two from polymer processing, a Kraft liquor database taken from a pulp and paper mill, and the diesel shootout database); 3) the results obtained on the diesel shootout database were found to be in very close agreement with those published with more sophisticated methods (Rossi et al., 2007; Teófilo et al., 2008); 4) the very parsimonious PLS models obtained after selection

have a slightly lower predictive power compared to the complete spectra or to the selection by regression coefficients; 5) the bootstrap-VIP approach seems to select the minimal amount of information (i.e. wavelength intervals) below which predictive capabilities starts degrading. It is therefore recommended to use the proposed selection method for the interpretation phase, after which some fine-tuning of the reduced PLS models may be performed to slightly improve the predictive power.

6.6 References

Albano C., Papa J., Gonzalez E., Navarro O. and Gonzalez R., Temperature and crystallinity profiles in polyolefin isothermal and non-isothermal solidification processes, *Eur. Polym. J.*, 39,1215-1222 (2003).

Babu G.J., Discussion of P.K. Sen's paper "Nonparametrics: retrospectives and perspectives", *J. Nonparam. Statist.*, 1, 33-35 (1991).

Bogomolov A. and Hachey M., Application of SIMPLISMA purity function for variable selection in multivariate regression analysis: A case study of protein secondary structure determination from infrared spectra, *Chemom. Intell. Lab. Syst.*, 88, 132-142 (2007).

Bower D.I. and Maddams W.F., *The Vibrational Spectroscopy of Polymers*, Cambridge University Press: Great Britain (1992).

Brás L.P., Lopes M., Ferreira A.P. and Menezes J.C., A bootstrap-based strategy for spectral interval selection in PLS regression, *J. Chemom.*, 22, 695-700 (2008).

Cai W., Li Y. and Shao X., A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra, *Chemom. Intell. Lab. Syst.*, 90, 188-194 (2008).

Centner V., Massart D.L., de Noord O.E., de Jong S., Vandeginste B.M. and Sterna C., Elimination of uninformative variables for multivariate calibration, *Anal. Chem.*, 68, 3851-3858 (1996).

Chew C., Griffiths J.R., Stachurski Z.H., The crystallization kinetics of polyethylene under isothermal and non-isothermal conditions, *Polymer*, 30, 874-881 (1989).

Chong I. and Jun C., Performance of some variable selection methods when multicollinearity is present, *Chemom. Intell. Lab. Syst.*, 78, 103-112 (2005).

Cramer J.A., Kramer K.E., Johnson K.J., Morris R.E. and Rose-Pehrsson S.L., Automated wavelength selection for spectroscopic fuel models by symmetrically contracting repeated unmoving window partial least squares, *Chemom. Intell. Lab. Syst.*, 92, 13-21 (2008).

Du Y.P., Liang Y.Z., Jiang J.H., Berry R.J. and Ozaki Y., Spectral regions selection to improve prediction ability of PLS models by changeable size window partial least squares and searching combination moving window partial least squares, *Anal. Chim. Acta*, 501, 183-191 (2004).

Efron B. and Tibshirani R.J., *An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability*, Chapman and Hall: USA (1993).

Eigenvector Research Incorporated (2005)
http://software.eigenvector.com/Data/SWRI.

ElMasry G., Wanga N., Vigneault C., Qiao J. and ElSayed A., Early detection of apple bruises on different background colors using hyperspectral imaging, *LWT*, 41, 337-345 (2008).

Eriksson L., Johansson E., Kettaneh-Wold N. and Wold S., *Multi- and Megavariate Data Analysis. Principles and Applications*, Umetrics Academy: Sweden (2001).

Fang S., Zhua M.-Q. and He C.-H., Moving window as a variable selection method in potentiometric titration multivariate calibration and its application to the simultaneous determination of ions in Raschig synthesis mixtures, *J. Chemom.*, 23, 117-123 (2009).

Ferreira A.P., Alves T.P. and Menezes J.C., Monitoring complex media fermentations with near-infrared spectroscopy: comparison of different variable selection methods, *Biotechnol. Bioeng.*, 91, 474-481 (2005).

Frank I.E., Intermediate least squares regression method, *Chemom. Intell. Lab. Syst.*, 1, 233-242 (1987).

Frenich A.G., Jouan-Rimbaud D., Massart D.L., Kutta-Tharmmakul S., Martinez Galera M. and Matinez Vidal J.L., Wavelength selection method for multicomponent spectrophotometric determinations using partial least squares, *Analyst*, 120, 2787-2792 (1995).

Gosselin R., Rodrigue D. and Duchesne C., On-line prediction of crystallinity spatial distribution across polymer films using NIR spectral imaging and chemometrics methods, *Can. J. Chem. Eng.*, 86, 869-878 (2008).

Gosselin R., Rodrigue D., González-Nuñez R. and Duchesne C., Quality control of polymer products through spectral imaging and chemometrics methods, *ANTEC*, Chicago, 1919-1924 (2009).

Heiman A. and Licht S., Fundamental baseline variations in aqueous near-infrared analysis, *Anal. Chim. Acta*, 394, 135-147 (1999).

Höskuldsson A., Variable and subset selection in PLS regression, *Chemom. Intell. Lab. Syst.*, 55, 23-38 (2001).

Huang A., Gaiyun L., Fu F. and Fei B., Use of visible and near infrared spectroscopy to predict the Klason lignin content of bamboo, chinese fir, paulownia and poplar, *J. Wood Chem. Technol.*, 28, 194-206 (2008).

Ilharco L.M., Garcia A.R., da Silva J.L., Lemos M.J. and Vieira Ferreira L.F., Ultraviolet-visible and Fourier transform infrared diffuse reflectance studies of benzophenone and fluorenone adsorbed onto microcrystalline cellulose, *Langmuir*, 13, 3787-3793 (1997).

Jouan-Rimbaud D., Walczak B., Massart D.-L., Last I.R. and Prebble K.A., Comparison of multivariate methods based on latent vectors and methods based on wavelength selection for the analysis of near-infrared spectroscopic data, *Anal. Chim. Acta*, 304, 285-295 (1995).

Kalivas J.H., Variance-decomposition of pure-component spectra as a measure of selectivity, *J. Chemom.*, 3, 409-418 (1989).

Kleynen O., Leemans V. and Destain M.F., Selection of the most efficient wavelength bands for "Jonagold" apple sorting, *Postharvest Biol. Tec.*, 30, 221-232 (2003).

Kramer K. and Ebel S., Application of NIR reflectance spectroscopy for the identification of pharmaceutical excipients, *Anal. Chim. Acta*, 420, 155-161 (2000).

Krier C., Rossi F., François D. and Verleysen M., A data-driven functional projection approach for the selection of feature ranges in spectra with ICA or cluster analysis, *Chemom. Intell. Lab. Syst.*, 91, 43-53 (2008).

Lazraq A., Cleroux R. and Gauchi J.P., Selecting both latent and explanatory variables in the PLS1 regression model, *Chemom. Intell. Lab. Syst.*, 66, 117-126 (2003).

Leclerc D.F., Determination of sodium sulfide and sulfidity in green liquors and smelt solutions, *WIPO*, 010501 (1997).

Lindgren F., Geladi P., Rännar S., Wold S., Interactive variable selection (IVS) for PLS. Part 1: Theory and algorithms, *J. Chemom.*, 8, 349-363 (1994).

Liu F., He Y. and Wang L., Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis, *Anal. Chim. Acta*, 615, 10-17 (2008).

Martens H. and Naes T., *Multivariate Calibration*, Wiley: United Kingdom (1989).

Miller R.G., An unbalanced jackknife, *Ann. Statist.*, 2, 880-891 (1974).

Mustonen M. and Laszlo J., NIR (near infrared) based liquor analyzer in the pulping and recovery area, *61st Appita Annual Conference and Exhibition*, Gold Coast, Australia (2007).

Muttiah R.S., *From Laboratory Spectroscopy to Remotely Sensed of Terrestrial Ecosystems*, Springer: USA (2002).

Nørgaard L., Saudland A., Wagner J., Nielsen J.P., Munck L. and Engelsen S.B., Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. *Appl. Spectrosc.*, 54, 413-419 (2000).

Olah M., Bologa C. and Oprea T.I., An automated PLS search for biologically relevant QSAR descriptors, *J. Comput. Aided Mol. Des.*, 18, 437-449 (2004).

Osborne S.D., Jordan R.B. and Künnemeyer R., Method of wavelength selection for partial least squares, *Analyst*, 122, 1531-1537 (1997).

Pandey K.K., Upreti N.K. and Srinivasan V.V., A fluorescence spectroscopic study on wood, *Wood Sci. Technol.*, 32, 309-315 (1998).

Pastore T.C.M., Santos K.O. and Rubim J.C., A spectrocolorimetric study on the effect of ultraviolet irradiation of four tropical hardwoods, *Bioresource Technol.*, 93, 37-42 (2004).

Preisner O., Lopes J.A., Guiomar R., Machado J. and Menezes J.C., Fourier transform infrared (FT-IR) spectroscopy in bacteriology: towards a reference method for bacteria discrimination, *Anal. Bioanal. Chem.*, 387, 1739-1748 (2007).

Ren Y., Shimoyama M., Ninomiya T., Matsukawa K., Inoue H., Noda I. and Osaka Y., Two-dimensional near-infrared correlation spectroscopy studies on composition-dependent spectral variations in ethylene/vinyl acetate copolymers: assignments of bands due to ethylene units in amorphous, disordered, and orthorhombic crystalline phases, *Appl. Spectrosc.*, 53, 919-926 (1999).

Rossi F., Francois D., Wertz V., Meurens M. and Verleysen M., Fast selection of spectral variables with B-spline compression, *Chemom. Intell. Lab. Syst.*, 86, 208-218 (2007).

Shao J. and Tu D., *The Jackknife and Bootstrap*, Springer: USA (1995).

Spiegelman C.H., McShane M.J., Goetz M.J., Motamedi M., Yue Q.L., Cot G.L., Theoretical justification of wavelength selection in PLS calibration: development of a new algorithm, *Anal. Chem.*, 70, 35-44 (1998).

Teófilo R.F., Martins J.P.A. and Ferreira M.M.C., Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression, *J. Chemom.*, 23, 32-48 (2009).

Watari M., Higashiyama H., Mitsui N., Tomo M. and Ozaki Y., On-line monitoring of the density of linear low-density polyethylene in a real plant by near-infrared spectroscopy and chemometrics, *Appl. Spectro.*, 58, 248-255 (2004).

Wold S., Sjöström M., Carlson R., Lundstedt T., Hellberg S., Skagerberg B., Wiksström C. and Öhman J., Multivariate design, *Anal. Chim. Acta*, 191, 17-32 (1986).

Wold S., Eriksson L., Trygg J. and Kattaneh N., The PLS method - partial least squares projections to latent structures - and its applications in industrial RDP (research, development, and production), *Compstat*, Prague (2004).

Xu L., Jiang J.H., Zhou Y.P., Wu H.L., Shen G.L. and Yu R.Q., MCCV stacked regression for model combination and fast spectral interval selection in multivariate calibration, *Chemom. Intell. Lab. Syst.*, 87, 226-230 (2007).

Xu H., Liu Z., Cai W. and Shao X., A wavelength selection method based on randomization test for near-infrared spectral analysis, *Chemom. Intell. Lab. Syst.*, 97, 189-193 (2009).

Yan Y.L., Zhao L.L., Han D.H. and Yang S.M., *The Foundation and Application of Near Infrared Spectroscopy Analysis*, China Light Industry Press: China (2005).

Ye X., Sakai K., Sasao A. and Asada S., Potential of airborne hyperspectral imagery to estimate fruit yield in citrus, *Chemom. Intell. Lab. Syst.*, 90, 132-144 (2008).

Chapter 7

A Hyperspectral Imaging Sensor for On-line Quality Control of Extruded Polymer Composite Products*

Résumé

Cette étude porte sur la capacité des méthodes chimiométriques à extraire de l'information pertinente à partir d'images spectrales afin de prédire les variations spatio-temporelles des propriétés de matériaux composite bois/plastique extrudés. Les échantillons ont été produits en variant les conditions opératoires, ainsi que l'alimentation selon un plan d'expériences. Les propriétés mécaniques des échantillons ont été mesurées à la fois pendant les périodes de régime stationnaire et les périodes de transition. D'abord, une méthode de régression Bootstrap-PLS a permis de sélectionner les intervalles spectraux les plus fortement corrélées aux propriétés mécaniques. Ensuite, un modèle PLS liant les caractéristiques spatiales et spectrales extraites des images à dimensionnalité réduite à la qualité des échantillons a été construit. Ce capteur peut simultanément estimer 7 propriétés mécaniques de ces matériaux composites à tout moment durant la production.

Abstract

This study examines the ability of chemometrics methods to extract meaningful information from visible and near-infrared spectral images of extruded wood/plastic composite materials for predicting spatio-temporal variations in their properties. The samples were produced under varying process and feed conditions according to designed experiments. Mechanical properties of the samples were measured using standard analytical methods both during steady-state and dynamic transition periods. A Bootstrap-PLS regression technique was first used for selecting the spectral bands (i.e. wavelengths) that were the most highly correlated with the material properties. In a second step, a more parsimonious PLS regression model was built between the spectral and textural features extracted from the lower dimensional spectral images and the corresponding quality properties of each sample. The imaging sensor was able to simultaneously monitor 7 properties in both steady-state operation and during transitions.

* Gosselin R., Duchesne C. and Rodrigue D., A hyperspectral imaging sensor for on-line quality control of extruded polymer composite products, *Comput. Chem. Eng., submitted: August 2009,* CACE-S-09-00415 (2009).

7.1 Introduction

Polymer extrusion is essentially a well instrumented process with tightly controlled operating conditions. Common automatic feedback control loops on machine variables (i.e. temperature, screw speed, etc.) provide satisfactory results as long as there are no stringent constraints on product quality (Tadmor et al., 1974). On the other hand, raw material properties, such as melt flow index and composition, are seldom uniform in complex polymeric flows. The resulting extrudate may suffer from quality fluctuations that cannot be tempered by the machine variable controllers. Moreover, these fluctuations may never be detected in a timely fashion due to the lack of on-line sensors for key polymer properties, which leads to quality control issues in polymer processing plants (Ohshima and Tanigaki, 2000). In practice, quality control of polymer products is mainly performed by destructive testing in the laboratory using a very small proportion of the production. It is therefore likely that large amounts of abnormal products may be generated before quality issues are detected. This problem may be addressed by taking on-line quality measurements (or estimates) of the samples and including this information in quality control and monitoring schemes. Specifically, an efficient system should rely on in-process analysis of polymer characteristics during manufacturing providing real-time assessment of sample properties such as structure, morphology and composition (Dumoulin et al., 1996).

Spectroscopic soft sensing techniques have been successfully applied in a number of fields such as food analysis (Mezreb et al., 2003; De Temmerman et al., 2007), pharmaceuticals (Côté and Abatzoglou, 2006; Cournoyer et al. 2008) and polymer science (Hansen and Vedula, 1998; Rohe et al., 1999; Barnes et al., 2003). The appeal of these methods lies in their ability to provide a great amount of information on the nature and structure of the sample through the location of the peaks and the intensity of the spectrum. Conventional machine variable control methods are simply not suited for maintaining molecular-specific properties to set-point, especially in complex or multi-component polymer systems.

Analytical soft-sensing methods such as UV-visible, near-IR, mid-IR and Raman spectroscopy can be applied in four modes of operation depending on sample nature,

geometry and property of interest: transmittance, reflectance, transflectance (light passing through the sample onto a mirror and reflected back through the sample) and diffuse reflectance (light penetrating partway into a moderately opaque sample) (Reshadat et al., 1999). Other on-line soft sensing methods include ultrasound transit time measurements (Chen et al., 1999; Barnes et al., 2003), X-ray scattering (Albano et al., 2003; Pantani et al., 2005; Stark and Matuana, 2007), fluorescence spectroscopy (Barnes et al., 2003), and nuclear magnetic resonance (NMR) spectroscopy (Bergmann, 1981; Pelsoci, 2007). According to Barnes et al. (2003), the polymer industry has so far mainly applied spectroscopic methods in either off-line or near-line (e.g. pellet composition) applications. Nevertheless, several sample properties have now been investigated on-line. These include estimation of the melt viscosity (Broadhead et al., 1993; Hansen and Vedula, 1998; Merikoski et al., 2001; McAfee and McNally, 2006), yield stress (Broadhead et al., 1993), polymer density (Watari et al., 1996), composition (Hansen and Vedula, 1998; Rohe et al., 1999; Barnes et al. 2003), extrudate temperature (Chen et al., 1999) and surface chemistry (Gupta et al., 2007; Stark and Matuana, 2007).

However, most soft sensing methods are based on single point or averaged measurements. Estimating sample quality distribution over a larger area using these approaches is difficult to perform on a high-speed production line. At best, a coarse spatial coverage of the surface can be achieved using probe arrays. Nevertheless, fine spatial measurements greatly enhance the probability of identifying local weaknesses that may compromise the quality of the whole. The mechanical, optical and barrier properties of an entire sample may fall below specifications because of local discrepancies that may be overlooked by a coarse spatial analysis.

Historically, modeling the open-loop response of a plasticating extrusion process can be traced back to the 1970s. During this period, research aimed at proposing relationships between screw speed, die pressure, extrudate flowrate and melt temperature (Reber et al., 1973; Tadmor et al., 1974). Later, work was done on the dynamics and control of the extruder pressure in cases where variations in feed quality and feed rates were present (Costin et al., 1982; Chan et al., 1986; Akdogan and Rumsey, 1996). More recently,

Mudalamane and Bigio (2003) studied the transmission of external disturbances through the extruder whereas McAfee and McNally (2006) determined the extruder dynamics using the melt viscosity. More complex modeling techniques were also applied to the extrusion process, for example using finite element methods (Yuan et al., 1994; Bajimaya et al., 2007; Ha et al., 2008), artificial neural networks (Shihani et al., 2006) and fuzzy logic (Fodil-Pacha et al., 2007).

This work investigates a new machine vision approach for on-line monitoring of product quality obtained by an extrusion process. On-line visible-NIR spectral imaging is used for rapid and non-destructive detection of spatio-temporal variations in the polymer properties in far greater detail than spectroscopic probes or probe arrays may achieve. Furthermore, rich spatial data would also be useful for precise diagnostic of operational problems in the production line. The proposed methodology uses chemometrics methods for extracting the spatial and spectral features that are most highly correlated with the material properties, and use this information to predict the properties of samples unknown to the model. To illustrate the concepts, the proposed techniques will be applied to the production of extruded wood/plastic composites (WPC). Nevertheless, these methods are not specific to WPCs. They are intended for any product that may be distinguished by its spectral and/or spatial surface characteristics.

The rationale behind this approach is twofold: 1) WPCs are complex multiphase systems that are not as well understood and controlled as neat polymers and 2) the production of WPCs has been increasing steadily in response to growing concerns over waste plastic management. In fact, the production of WPCs quadrupled between 1997 and 2000 (Li and Wolcott, 2004) and is expected to reach $2.4 billion US by 2011 (WRAP, 2003).

7.2 Materials and methods

7.2.1 Sample production and testing

Extruded high density polyethylene (HDPE)/wood composites were produced and tested under different processing conditions and wood contents. The polymer used was commercial HDPE (Fortiflex A60-70-162) and birch wood fibers were obtained from a local sawmill. These fibers were sieved between 45 and 180 μm and dried in a vacuum oven at 104°C overnight in order to minimize the amount of volatiles. Compounding of the polymer and wood was achieved using a Haake Rheomix TW-100 dual-feeder counter-rotating twin-screw extruder. This allows proper distribution of fibers in the polymer blend. In this first step the barrel temperatures of the extruder were controlled at 150 – 160 – 160°C for zones 1, 2 and 3, respectively, and 160°C for the die. The compound was extruded through a cylindrical 5 mm die, air cooled, and pelletized. After this blending phase, the composites were extruded using a flat horizontal 3 × 80 mm die. The following process conditions were investigated in a complete 2^3 factorial design: die temperature (160 and 180°C), screw speed (70 and 90 rpm), and polymer throughput (500 and 600 g/h) with a fixed wood content of 20 wt.%. Wood content was also varied in a separate set of experiments at three levels (10, 20 and 30 wt.%) under constant nominal process conditions (i.e. center point of 2^3 design). Steady-states were allowed to run for about 10 minutes for each condition tested. These processing conditions are presented in Table 7.1.

Sample quality consists of 6 tensile properties (modulus, stress and strain at yield, stress and strain at break, as well as the energy at break), measured according to ASTM D638 using an Instron 5565 in the direction of extrusion at a crosshead speed of 1 mm/min, measured over a 25 mm gage length. The measurements also include 1 thermal property (enthalpy of fusion) obtained using a Perkin Elmer DSC 7. While enthalpy of fusion and crystallinity are closely linked (the former is used to calculate the latter), crystallinity also requires an estimation of the sample composition. In order to avoid this complication during DSC testing (test sample typically weight 10 mg), only enthalpy of fusion was used. Five replicates of each composite formulation were fully tested. The quality of the extruded parts was studied both during steady-states and in the transitions between them (i.e.

dynamics). The steady-states were used as a training set for model building, while the transitions were used for model validation.

Table 7.1. Randomized 2^3 full factorial design for machine variables at 20% wood (samples 1-8) combined with 3 additional experiments in which wood content was varied under nominal processing conditions (samples 9-11).

Sample	Wood (%)	Temperature (°C)	Throughput (g/h)	Screw Speed (rpm)
1	20	160	600	70
2	20	160	500	90
3	20	180	500	70
4	20	180	600	90
5	20	160	600	90
6	20	160	500	70
7	20	180	600	70
8	20	180	500	90
9	10	170	550	80
10	20	170	550	80
11	30	170	550	80

7.2.2 Spectral image acquisition

Spectral imaging is well suited for monitoring spatio-temporal variations across a surface. A line-scan imaging system was used in this study since it rapidly captures light intensities simultaneously at many wavelengths for a single row of pixels ($x \times \lambda$) which is better suited for moving samples such as those found on production lines (Figure 7.1). A full image is obtained after collecting several line-scans along the y-axis. The imaging system used in this work is composed of two cameras, each coupled with a grating spectroscope, effectively covering the visible and infrared range of the spectrum (475-1500 nm). Further information on the optical setup can be found in Gosselin et al. (2008, 2009). For each scanned line of the samples, the reflected light is dispersed through the spectroscope and captured by the CCD array of the camera as a spatial/spectral image. A hyperspectral image of the entire sample is achieved by repeatedly displacing the sample perpendicularly to the spectroscope slit, along the y-axis, forming a 3-D array ($x \times y \times \lambda$) of data. Each of these images contains 500 spectral channels (i.e. wavelengths), rather than the 3 more commonly obtained in RGB images.

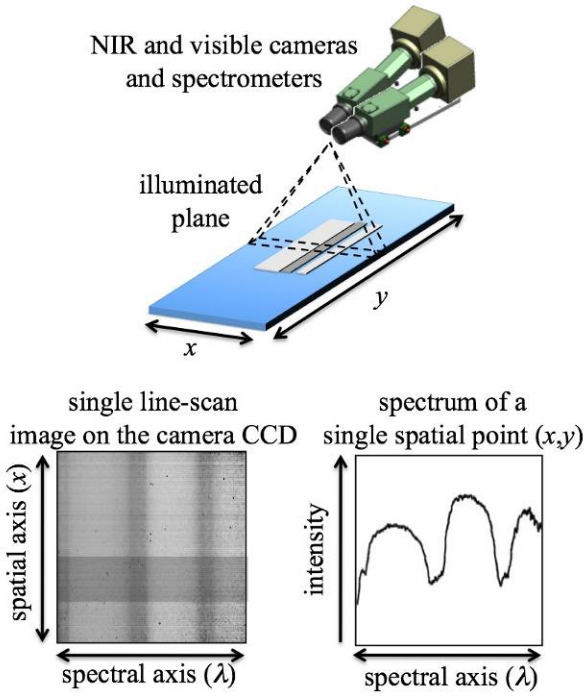

Figure 7.1. Line-scan VIS-NIR imaging system used to collect spatial/spectral reflectance images.

A diffuse tungsten-halogen lamp was chosen for its smooth flat emission spectrum over the required range. The raw spectral-spatial image was transformed into reflectance units using a flat black image and a true white image (Valadez-Blanco et al., 2007; Gosselin et al., 2008, 2009). While NIR spectra are commonly used to study neat polymers, little work has been done on wood/plastic composites (Hansen and Vedula, 1998; Barnes et al., 2003). The samples were therefore imaged using both visible and NIR cameras (λ=500 wavelengths). Figure 7.2 illustrates the average spectral signature for samples produced under the 11 steady-state conditions shown in Table 7.1. Averaging was performed along both spatial directions (x and y). The spectra appear to have similar peaks but differ in intensity, as one would expect in a dataset where the processing conditions were varied but the components remained the same.

A training set of spectral data was constructed using the images collected during steady-state operation. Each condition was imaged 25 times (y=25) at 1 mm space intervals. As a result, 11 spectral images of size (25×25×500) were obtained for describing product quality

in steady-state. As for the validation set, spectral images of the same size were captured every about 30 cm (or about every minute) during the transitions from one steady-state to the other. Product quality during steady-state and transitions was tested in the lab using samples taken within the field of view of each spectral image to ascertain the validity of the subsequent predictions.

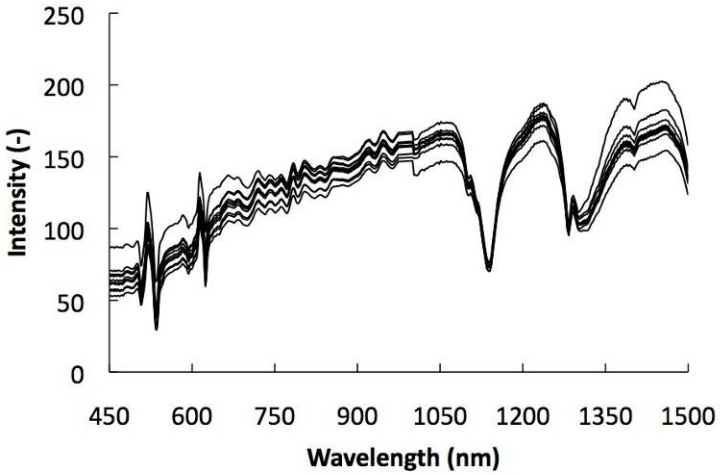

Figure 7.2. Averaged spectrum for each of the 11 steady-state samples (8-bit images).

7.2.3 Spectral and textural analysis

Spectral imaging typically contains both spectral and spatial information. In the case of WPCs, the spectral signature essentially reflects the chemical composition and small-scale physical structure (i.e. crystallinity) of the composites whereas the spatial information represents large-scale variations throughout the sample (i.e. fiber orientation and distribution) and can be thought of as surface smoothness.

Spectral analysis

Despite the fact that spectral images may contain large quantities of information relative to sample nature (peak location) and structure (spectrum intensity), there are often high levels of collinearity between neighbor wavelengths. To increase the robustness of the prediction models, numerous methods have been proposed to reduce the number of wavelengths from

spectral data including PLS-based methods (Cai et al., 2008; Brás et al., 2008), Genetic Algorithms (Reynes et al., 2006), Neural-Network (Boger, 2003), Support Vector Machines (Liu et al., 2008) as well as Fuzzy Logic methods (Setarehdan, 2006). In this study, a PLS-bootstrap based method previously developed and presented in Gosselin et al. (2010) is used. Prior to wavelength selection, the spectral images were averaged along the machine cross direction since no systematic spectral variation were observed along the sample width. The spectral data matrices **X** obtained for steady-state (i.e. training set) and transitions (i.e. validation set) were reduced to (275×500) and (6750×500) dimensional, respectively. The first dimension corresponds to the total number of line-scans, and the second to the number of wavelengths.

Wavelength selection was achieved through bootstrapping of a PLS regression model between spectral data (**X**) and corresponding quality measurements (**Y**). The bootstrap is a well-known re-sampling method used in statistics to estimate bias and standard errors of parameter estimates. Through generating a population of randomly selected subsets with replacement from the complete dataset, the uncertainty interval of any statistical parameter of interest describing the importance of each variable (i.e. wavelength) in the model can be obtained. Variable deemed uninformative can then be discarded. The PLS-bootstrap method is applied by first computing the PLS variable importance metric (VIP) for each wavelength using the complete dataset. Then, the uncertainties in each wavelength VIP was assessed using the bootstrap. A wavelength was considered important for predicting sample quality when the one standard deviation error bar around each VIP estimate were above 1.0 (Eriksson et al., 2001; ElMasry et al., 2008). More information on PLS models is given in Section 7.2.4.

Spatial analysis

Image texture consists of the spatial organization of pixel intensities and may be used to quantify the aesthetics of the sample (i.e. smoothness, orientation, etc.) or may linked to the sample properties through the identification of local features, such as component segregation, that may compromise sample quality as a whole. Liu and MacGregor (2005)

used this concept of image texture to optimize the visual appearance of industrial injection-molded plastic panels. The gray-level co-occurrence matrix (GLCM), a state-of-the-art statistical texture analysis method first proposed by Haralick et al. (1973), was used to quantify the visual texture of the samples. This was achieved based on a representative gray-level image taken from the original spectral images (25×25×500). Which wavelength was selected for texture analysis is discussed later. The GLCM compares the gray-levels of neighboring pixels, two-by-two, at different distance and orientation from each other, throughout the image. In essence, the image is found to be smooth when the gray-levels are similar, and rough, when they are not. Three user-defined parameters need to be selected for texture quantification: the distance between any pair of pixels, the angle between them, and the number of gray-levels in the image. The result is a roughly diagonal matrix, named the GLCM matrix, in which the distance from the main diagonal indicates the roughness of the sample. Such a matrix is obtained for each combination of the 3 GLCM parameters (i.e. distance, angle, and gray-levels). Several scalar features computed from these GLCM matrices were proposed by Haralick (1979) (e.g. correlation, entropy, energy, etc.) in order to allow texture comparison between different images.

Distances of 2, 5 and 10 pixels were selected in both horizontal and vertical directions (i.e. 0 and 90°) in order to quantify textural features at different scales and orientation. As the line-scans were acquired at 1 mm intervals, the distances between pixels correspond to variations on the order of magnitude of the wood fibers (2 pixels) up to sample-wide variations (10 pixels). Texture descriptors for each GLCM matrix, computed using 32 gray-levels, were the correlation and entropy Haralick's scalar features. Entropy is a measure of randomness of the pixel intensities over the entire image and correlation as a measure of how correlated each pixel is to its neighbor. These are computed as follows:

$$entropy = -\sum_{i,j} p_{i,j} \log\left(p_{i,j}\right) \tag{7.1}$$

$$correlation = \sum_{i,j} \frac{ijp_{i,j} - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{7.2}$$

where the indices i and j are the row and column numbers in the GLCM matrices, $p_{i,j}$ is the normalized value of an element (i,j) of the GLCM matrices and μ and σ are the gray-level average and standard deviation over i and j. Using this method, image texture is quantified at 3 scales, in 2 directions and using 2 features, for a total of 12 texture descriptors.

7.2.4 Multivariate latent variable methods

Principal component analysis

Principal component analysis (PCA) is a well-known multivariate statistical method that decomposes the variance-covariance structure of a data matrix by finding linear combinations of the original variables in such a way that these combinations explain the greatest amount of variation within the data (Eriksson et al., 2001). In the case of a dataset $\mathbf{X}$ ($N \times K$) consisting of N observations described by K variables, this decomposition is written as:

$$\mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \qquad (7.3)$$

where a is the number of principal components, $\mathbf{t}_a$ are the score vectors, and $\mathbf{p}_a$ are the loading vectors. The first score vector $\mathbf{t_1}$ is the linear combination of the K variables (defined by the loadings vectors $\mathbf{p_1}$) capturing the largest possible variance within the matrix $\mathbf{X}$, while the second score vector $\mathbf{t_2}$ represents the second largest source of variance, and so on. The score vectors can therefore be viewed as a multivariate summary of each $\mathbf{X}$ observation. If the number of components is lower than the maximum number of components ($A<K$), then $\mathbf{E}$ ($N \times K$) contains the residuals. Further information on PCA can be found in Wold et al. (1987).

Projection to latent structures

The aim of this work is to build a model between the spectral and spatial features extracted from each image and the corresponding quality variables. According to Hansen and Vedula (1998), considerable overlap between the spectral bands in the NIR spectrum makes it difficult to build such a model using simple multiple linear regression techniques. Projection to Latent Structures (PLS) model may be used in such a situation. The PLS algorithm extracts the relationship between regressor variables (**X**) and response variables (**Y**) by assuming that they are generated by a common set of underlying factors (Brás et al., 2008). This characteristic enables the model to handle highly collinear datasets. The structure of the latent variable PLS regression model between the image features (*K* variables) of the samples **X** (*N*×*K*) and their corresponding quality matrix **Y** (*N*×*M*) is given below:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{7.4}$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \tag{7.5}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^{*} \tag{7.6}$$

$$\mathbf{W}^{*} = \mathbf{W}\left(\mathbf{P}^{\mathrm{T}}\mathbf{W}\right)^{-1} \tag{7.7}$$

where the **P** (*K*×*A*) and **Q** (*M*×*A*) matrices contain the loading vectors of the **X** and **Y** spaces, respectively. The **E** (*N*×*K*) and **F** (*N*×*M*) matrices contain the residuals. The orthogonality of the PLS components ($a=1,2,\ldots A$) make it possible to use a relatively small number of latent variables ($a<K$). The PLS model can be expressed in such a way as to link **T** to **X** through **W*** (*K*×*A*), containing the X-weights. Finally, the VIP score of a predictor variable (i.e. wavelength) quantifies the importance of this variable in the projection model (Eriksson et al., 2001; Chong and Jun, 2005). The VIP score for the k^{th} variable (wavelength) for a single response **y** is computed according to Eriksson et al. (2001):

$$VIP_k = \sqrt{K \sum_{a=1}^{A} \left[\left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right) \left(w_{ak} / \left\| \mathbf{w}_a \right\|^2 \right) \right] \Big/ \sum_{a=1}^{A} \left(q_a^2 \mathbf{t}_a^{\mathrm{T}} \mathbf{t}_a \right)} \tag{7.8}$$

where a = 1,2…,A, K is the number of columns of **X** (i.e. variables or wavelengths), w_{ak} is the loading weight of the k^{th} variable in the a^{th} component, and $\boldsymbol{t}_a$, $\boldsymbol{w}_a$, and q_a are the a^{th} column vectors of **T**, **W**, and **Q**, respectively. Since wavelengths with higher VIP scores are more relevant in prediction and the average of the squared values of the VIPs is equal to 1, a greater-than-one rule is often used as a selection criterion (Olah et al., 2004; Chong and Jun, 2005). More details on PLS can be found in Eriksson et al. (2001).

7.3 Results

7.3.1 PCA on quality data for relevant wavelength selection

The high level of collinearity (i.e. redundancy) present in spectral data allows reducing the number of wavelengths, resulting in simpler and potentially more robust models. The PLS-bootstrap wavelength selection method (Gosselin et al., 2010) was designed for a single quality variable (**y**). However, 7 properties were chosen in this study: enthalpy of fusion, modulus, stress and strain at yield, stress and strain at break as well as the energy at break. Rather than selecting appropriate wavelengths for each quality variable individually, PCA was first performed on the quality matrix (**Y**) of the training set (extrusion steady-states) in order to determine possible correlations between the material properties. Two components were found sufficient for explaining most of the information contained in the quality data, each explaining 63.1% and 22.6% of the variance, respectively, for a total of 85.7%. The $\mathbf{p_1}$-$\mathbf{p_2}$ loading plot of the PCA-Y decomposition shown in Figure 7.3 is used to interpret the relationships between the 7 properties. The following conclusions are drawn: (1) modulus is strongly negatively correlated to most other properties, (2) enthalpy of fusion, energy at break, and the strains (both at yield and at break) are positively correlated among each other, and (3) properties at break are highly correlated to those at yield.

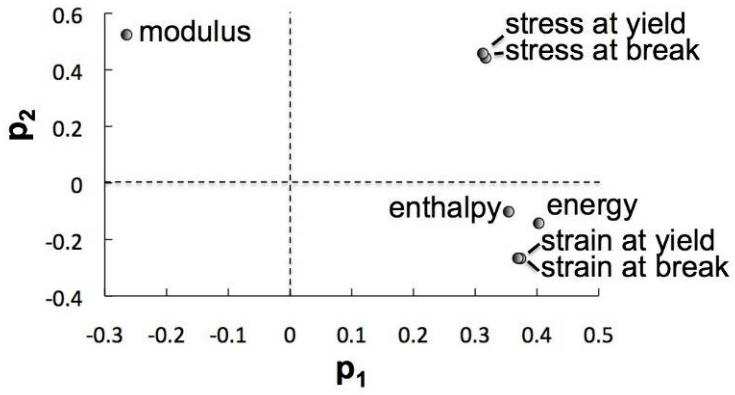

Figure 7.3. Loadings plot of the PCA-Y performed on the quality variables.

As the first two components represent most of the variance in **Y**, these will be used as representative new variables to perform the wavelength selection. Rather than identifying the relevant wavelengths for each of the 7 y-variables individually, this selection is performed using the corresponding PCA-Y score vectors, referred to as $\mathbf{y}_{PCA1}$ and $\mathbf{y}_{PCA2}$. Simplifying the Y-matrix through PCA both lightens the wavelength reduction scheme and increases interpretability. The PLS-bootstrap method starts with building a PLS model using all **X** and the [$\mathbf{y}_{PCA1}$ $\mathbf{y}_{PCA2}$] data. The number of PLS components was chosen by random-subset cross-validation on entire training set. The importance of each wavelength (i.e. VIP metric) was computed based on this model. Wavelengths having a VIP significantly greater than one were considered to be the most important for predicting $\mathbf{y}_{PCA1}$ and $\mathbf{y}_{PCA2}$, and were kept in the final prediction model. The relevant wavelengths for either $\mathbf{y}_{PCA1}$ and $\mathbf{y}_{PCA2}$ based on the bootstrap one standard deviation error are shown in Figure 7.4A (see regions in blue). A limited number of wavelengths around 1100 nm are particularly significant, while the interval between 500-1000 nm is not (Figures 7.4B and C).

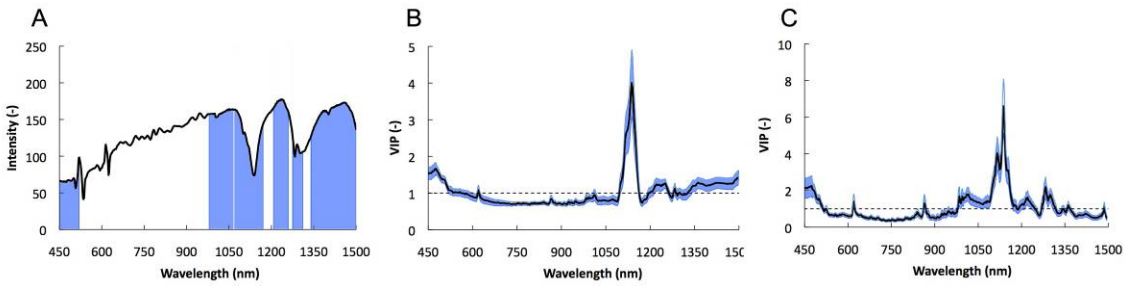

Figure 7.4. Wavelength selection achieved using the Bootstrap-PLS algorithm. The 6 relevant wavelengths bands are shown by the blue areas in A). The VIP values and one standard deviation for y_{PCA1} and y_{PCA2} are shown in B and C, respectively. The VIP=1.0 threshold is identified using the dash line.

As illustrated in Figure 7.4A, the PLS-bootstrap based method identified 6 quality-relevant spectral bands. Inter-correlation analysis between these bands varies from 37% (bands 1 and 3) to 95% (bands 2 and 4) with an average of 75%, confirming a certain level of independence between the selected wavelengths. The first of these (450-515 nm) represents the blue-green wavelengths. This spectral interval has previously been linked to a number of wood species (Pandey et al., 1998; Pastore et al., 2004) and may be associated to the cellulose in the wood (Ilharco et al., 1997). The second and third intervals (960-1065 nm and 1070-1165 nm) represent the C-H stretching second overtone band and are linked to CH_2 vibrations (Watari et al., 2004). According to Bower and Maddams (1992), high crystallinity levels reduce the freedom of the macromolecules to vibrations and can be observed through this band. The remaining intervals (1210-1260 nm, 1280-1310 nm and 1340-1500 nm) appear to be associated to the first overtone combination band. They represent combined stretching and vibrational movements linked to polyethylene crystallinity (Ren et al., 1999; Watari et al., 2004).

The 6 spectral bands selected account for 202 of the 500 wavelengths present in the original images. In order to simplify the remaining spectral information, only the average value of each of these 6 spectral bands was kept since the wavelengths within each band are highly correlated to one another. This further reduced the size of the spectral data from the original 500 wavelengths to 202 significant wavelengths grouped together into 6 equivalent spectral bands. The hyperspectral images now contain 6 wavelengths (i.e. after averaging) in the same way that RGB images contain 3 wavelengths.

7.3.2 PLS model building and interpretation for steady-state data

Formulating a regression problem between multivariate images and response variables involves extracting a certain number of features (K) from each of the N images, collecting them in a regressor matrix **X** ($N \times K$), and building a PLS regression model against the response variables of interest **Y** ($N \times 7$). In other words, the relevant information contained within each image must be condensed into a single vector of length K. This is achieved in two steps. First, the spectral information included in the model was averaged over each line-scan (extrusion machine cross-direction). As the 500 wavelengths captured by the imaging system had previously been reduced to 6 significant wavelength bands, the spectral information in each line-scan is reduced into 6 spectral features. As mentioned in Section 7.2.2, the training set consists of 11 steady-states, each imaged 25 times, for a total of 275 line-scans. Then, the GLCM texture analysis was performed on images corresponding to these 6 wavelength bands. Preliminary tests indicated that the textural information does not differ significantly from one wavelength band to another. The textural features extracted from the 4th spectral band (1210-1260 nm) had the strongest reflectance and provided a good signal-to-noise ratio. As previously mentioned, the GLCM analysis led to 12 textural variables. Combined, the matrix **X** contains a total of 18 image features (Figure 7.5).

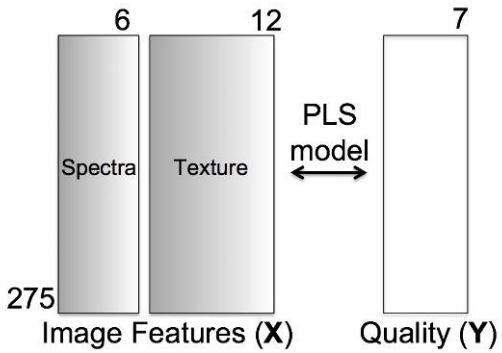

Figure 7.5. The PLS model is based on 18 image features (6 averaged wavelengths and 12 spatial characteristics) to predict 7 quality properties of the WPC polymer composites.

A 6-component PLS model was then built using the PLS-toolbox software (Eigenvector Research Inc.) and a standard random-subsets cross-validation procedure. Overall, the cumulative **X** and **Y** variances explained by the model were $R^2_{X,cum}$= 96.2% and $R^2_{Y,cum}$= 88.3%. Nevertheless, as much as 53.7% of the total Y-variance was captured using only the first 2 components of the model. To verify whether this model is able to discriminate the quality of the samples produced in 11 steady-state conditions using both spectral and textural image information, the 3 score scatter plots ($\mathbf{t_1}$-$\mathbf{t_2}$, $\mathbf{t_3}$-$\mathbf{t_4}$ and $\mathbf{t_5}$-$\mathbf{t_6}$) shown in Figure 7.6 are used. To enhance visual interpretation, only the average location of each steady-state sample (numbered circles), as well as the dispersion of the data within each steady-state (3-σ error bars), are presented.

The first score vector $\mathbf{t_1}$ captures the largest possible covariance between **X** and **Y** while each subsequent score vector captures smaller sources of covariance. This is illustrated in Figure 7.6 by the reduction in axis length occurring between $\mathbf{t_1}$-$\mathbf{t_2}$, $\mathbf{t_3}$-$\mathbf{t_4}$ and $\mathbf{t_5}$-$\mathbf{t_6}$. Consequently, the first components contain more relevant information, leading to clearer clustering of the data, and will be used in illustrations throughout the remainder of this paper. Nevertheless, the latter components make it possible to distinguish between similar samples; while some samples can easily be identified using $\mathbf{t_1}$-$\mathbf{t_2}$ alone (i.e. 4, 9, 11), others require more subtle information (i.e. samples 1 and 8 may be difficult to distinguish in $\mathbf{t_1}$-$\mathbf{t_2}$ but differ greatly in $\mathbf{t_3}$-$\mathbf{t_4}$ and $\mathbf{t_5}$-$\mathbf{t_6}$).

In order to determine whether the 11 sample types differ significantly from one another, the Hotelling's T^2 two-sample multivariate statistic was used (Hotelling, 1931; Mardia et al., 1979; Shaw and Lee, 2007). This generalization of the univarite Student's t statistic may be used in multivariate hypothesis testing to compare two populations. The results show that, when the 6 components are used, all 11 sample types differ significantly from each other ($p < 0.001$). To illustrate this point, samples 3 and 5 were taken as their intervals overlap in all three score plots of Figure 7.6. In this case, $p = 0.1621$ when a single component is used ($\mathbf{t_1}$), $p = 0.0780$ with two components ($\mathbf{t_1}$-$\mathbf{t_2}$) and $p=0.0002$ with 6 components ($\mathbf{t_1}$-$\mathbf{t_6}$). The proposed hyperspectral sensor can therefore discriminate the quality of the polymer product

obtained at each of the 11 sets of steady-state extrusion conditions. The transitions between them will be examined later.

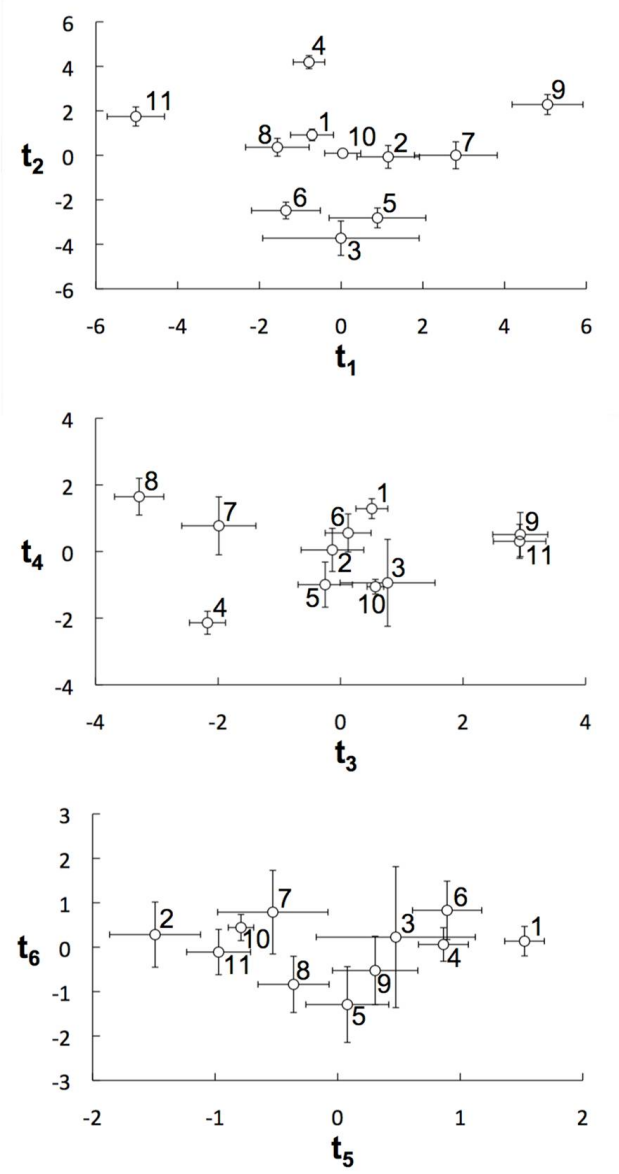

Figure 7.6. Score plots illustrating the 6 components of the PLS model (**t_1-t_2**, **t_3-t_4** *and* **t_5-t_6**) *in which each of the 11 steady-states can be distinguished. Error bars depict 3 standard deviations around steady-states along each component.*

The relationships between the spectral/textural features and the materials properties are interpreted using the PLS loading biplot for the first two principal components (Figure 7.7A). This biplot shows the correlation structure of the variables both within each of the **X** and **Y** blocks, as well as between them. Variables with similar coordinates in the biplot are

positively correlated, those falling in opposite quadrants are negatively correlated, and those near the origin (0,0) have little or no correlation with other variables. Analyzing the material properties in Figure 7.7A, one finds the same type of pattern found in Figure 7.3, namely that the modulus is negatively correlated to all the other properties and that both the stress at yield and at break fall apart from the remaining 4 properties. In fact, being strongly correlated to the 1st component, this distinction makes up the largest source of variance in the dataset. Meanwhile, the 2nd component distinguishes between the enthalpy of fusion, strain at yield and at break as well as the energy at break. Components 3-6 only serve to enhance these distinctions.

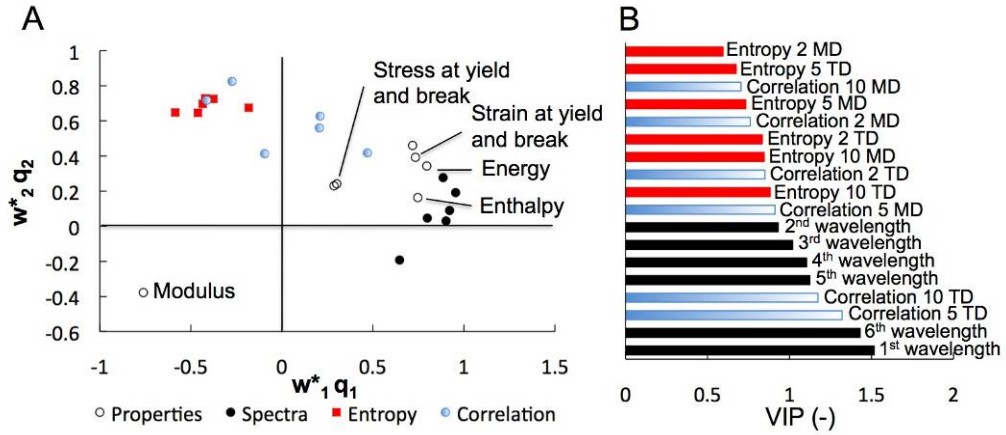

Figure 7.7. Loadings biplot of the first 2 components of the PLS model (A) depicting the relation between spectral variables, spatial variables and the properties of the samples. VIP plot depicting the relative importance of each of the regressor variable. Spatial variables are denoted here according to entropy and correlation, the distance between neighboring pixels involved in the GLCM analysis (2, 5 or 10) and the direction of which the texture analysis was carried out (machine direction: MD vs. transverse direction: TD).

The joint analysis of the loadings of both the regressor and quality variables in Figure 7.7A shows that the first component (i.e. horizontal axis) mainly explains variations in wood content. Polymer processing knowledge shows that when wood content increases, modulus also increases while other properties such as the strains and enthalpy of fusion are reduced. Product coloration becomes darker and its surface texture is rougher. These relationships are clearly shown in Figure 7.7A. Modulus is positively correlated with the entropy features (i.e. surface roughness) but negatively correlated with the spectral intensities and other material properties. Indeed, lower intensities in the 1st wavelength band means coloration is

darker whereas the other NIR bands are related with lower polymer concentration and crystallinity. Therefore, based on these correlations, a material having a higher modulus and a rougher surface texture but lower values for all other variables should fall in the negative $\mathbf{t_1}$ quadrants, and vice-versa. This is verified in Figure 7.6 since steady-state 11 (30% wood) fall on the left hand side in the $\mathbf{t_1}$-$\mathbf{t_2}$ score plot whereas the material produced with 10% wood (steady-state 9) fall on the right hand side. The spectral sensor could therefore detect fluctuations in the feed composition.

The second component (vertical axis in Figure 7.7A) mainly involves textural features and modulus. These are negatively correlated with one another in the $\mathbf{t_2}$ direction. This was found to be related with the severity of the thermo-mechanical energy delivered to the system. As more energy is supplied to the melt by increasing barrel temperature set-point (thermal energy) and/or throughput or screw speed (mechanical energy) wood fiber and polymer degradation is progressively observed. Damaged fibers increase material surface roughness (entropy features) and decrease some mechanical properties such as modulus. This is confirmed in the score plot of Figure 7.6 where the most severe conditions, achieved in steady-state 4, have the highest $\mathbf{t_2}$ values whereas the less severe conditions of steady state 6 project in the negative $\mathbf{t_2}$ region. The $\mathbf{t_1}$-$\mathbf{t_2}$ score plot could also be used to warn operators that extrusion conditions may lead to material degradation.

The importance of each of the 18 X-regressor variables in the 6-components PLS model (i.e. VIP) is reported in Figure 7.7B. The model strongly relies on the spectral information (i.e. the 1st and 6th wavelengths) as well as the textural information (i.e. 5 and 10-step correlation in the transverse direction). However, the entropy textural features are less important taken individually but are important as a group.

In order to judge the adequacy of the model, two cumulative multiple correlation coefficients were computed to quantify the model predictive ability: $R^2_{Y,cum}$ represents the cumulative percentage of the total variance of the **Y** matrix used to build the model while Q^2_{cum} represents the cumulative percentage of the total variance of a **Y** matrix that has not been used to build the model. The performance of the method in training and validation are

compared in Table 7.2. The results obtained for the training set show that all of the properties are well predicted by the model, but there are some distinctions to be made. First, enthalpy of fusion shows a lower $R^2_{Y,cum}$ compared to other properties. This was expected since enthalpy of fusion is obtained by measuring the heat of fusion of the polymer phase within a sample and comparing it to the sample total weight. However, laboratory testing is typically performed on small samples of the composite (10 mg), making it virtually impossible to estimate the true polymer fraction within any given sample.

Second, as is typically the case, the performance of the model in validation (Q^2_{cum}) are inferior to those in fit ($R^2_{Y,cum}$). The magnitude of this drop however shows a distinct pattern. The properties at yield (both stress and strain) have higher Q^2_{cum} values relative to their $R^2_{Y,cum}$ than do the properties at break. In other words, the yield properties are more easily predicted than the break properties. This was expected since ultimate properties (at break) usually have high data scatter (Ghaneh-Fard, 1999). Further inquiry into the subject indicated that the full spectra (500 wavelengths) could not predict these properties with any greater accuracy. Nevertheless, it is clear in the results (Figures 7.3 and 7.7) that the properties at break and at yield have similar trends.

Table 7.2. PLS model ability in training ($R^2_{Y,cum}$) and validation (Q^2_{cum}).

Property	Steady states Training set ($R^2_{Y,cum}$)	Transitions Validation set (Q^2_{cum})
Enthalpy of fusion	0.853	0.851
Modulus	0.884	0.880
Stress at yield	0.893	0.869
Strain at yield	0.891	0.870
Stress at break	0.870	0.818
Strain at break	0.880	0.809
Energy	0.912	0.890
Overall	0.883	0.855

7.3.3 Transitions in latent variable space of the PLS model

Product quality and spectral images were also collected during the transitions between the 11 steady-states. These are used to validate the quality sensor by applying the PLS model developed in the previous section to the images collected in the transients. This will allow investigating the process dynamics between machine variables, the raw materials, and the final product quality.

Figure 7.8A is an adaptation of the $\mathbf{t_1}$-$\mathbf{t_2}$ score plot shown in Figure 7.6 in which only the center points of the 11 steady-states are shown. Two transitions are superimposed on this plot: between steady-states 5 and 6, as well as between 9 and 10. These were chosen to represent simple transitions in which a single parameter was varied (9 to 10: step in wood content) as well as more complex transitions in which 2 step changes occurred at once (5 to 6: simultaneous steps in throughput rate and screw speed). Both transitions begin near the center-point of the initial steady-state, cross the $\mathbf{t_1}$-$\mathbf{t_2}$ score plot in a relatively straight line and stabilize near the center-point of the final steady-state. While not shown here, transitions in the $\mathbf{t_3}$-$\mathbf{t_4}$ and $\mathbf{t_5}$-$\mathbf{t_6}$ score plots also occur linearly but follow noisier paths, as the signal-to-noise ratios of components 3-6 are inferior to those of components 1 and 2. It is important to note that the PLS model developed earlier is a steady state model relating the spectral image to the properties of the material within the field view of the camera. Therefore, this model can also be applied for monitoring transitions since there are no dynamics between the image and the material itself.

Figure 7.8B illustrates transitions 5 to 6 and 9 to 10 as they occur through time, rather than across the score space. They are calculated using Hotelling's T^2 single-sample multivariate statistic (Hotelling, 1931) which essentially measures the Mahalanobis distance in the score plot of the current operating conditions from the initial steady-state; large T^2 values therefore indicate that the samples imaged no longer belong to the initial steady-state conditions. Mirroring their representation in Figure 7.8A, the transitions begin very near their initial steady-states (5 or 9) and stabilize a certain distance (measured in T^2) from it, implying that they have reached their final steady-states (6 or 10). Between these two

points, both transitions appear to approximately follow a 1st order with dead time dynamic response. In fact, most of the transitions studied in this work followed similar dynamics.

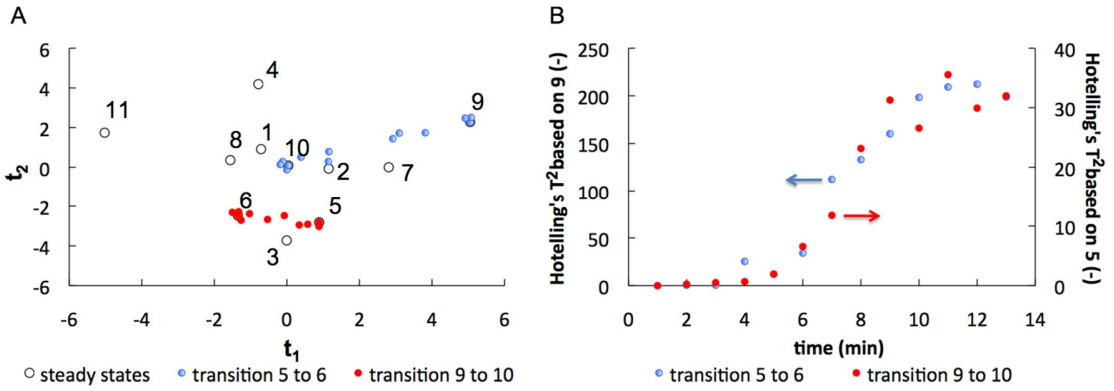

Figure 7.8. Typical transitions as seen in the PLS score plot (A), and in time (B). Average steady-states are shown as open symbols. Transition 5 to 6: simultaneous -100 g/h throughput rate and +20 rpm screw speed (closed red dots). Transition 9 to 10: +10% step in wood content (light blue dots).

To quantify extrusion dynamics, multiple input – single output (MISO) transfer function models were built between machine variables (inputs) and the measurements provided by the hyperspectral sensor (i.e. scores of the PLS model) using the Matlab's System Identification Toolbox (R14). The first PLS component was selected for illustration, since it explains most of the variance in the material properties (e.g. wood content and surface esthetics). The four input variables were barrel temperature, throughput rate, screw-speed and wood content in the recipe. Based on Figure 7.8B, it was concluded that first-order with dead-time models adequately capture the dominant process dynamics and it was found that second-order models contribute little to improving the fit. These results corroborate earlier results (Reber et al., 1973; Cayot et al., 1995; Haley and Mulvaney, 2000; McAfee and McNally 2006). First-order time constant and dead-time (T_p and T_d), along with their respective confidence intervals, were computed for transfer functions between each of the 4 machine variables and the first score of the PLS model. The results are presented in Table 7.3. While the system identification was performed using $\mathbf{t_1}$, similar results were obtained using the other PLS components.

Table 7.3. First order dynamic time constant and dead-time for each of the four transfer functions between machine variables and the 1st component of the PLS model. Confidence intervals are indicated between parentheses.

Variable	Time constant T_p (min)	Dead-time T_d (min)
Temperature (°C)	6.55 (1.78)	5.58 (2.48)
Throughput (g/h)	3.75 (1.04)	2.61 (0.77)
Screw speed (rpm)	3.92 (1.46)	1.70 (1.56)
Wood content (%)	4.31 (0.33)	3.62 (0.22)

Analyzing the dynamic constants in Table 7.3, it can be seen first that temperature variations lead to slow dynamics preceded by long delays, confirming previous observations (Reber et al., 1973; McAfee and McNally, 2006). Second, variations in throughput rates and wood content follow relatively similar dynamics. This may be expected since both are feed variations that need to propagate from the feed hopper to the die exit. However, variations in wood content do appear slightly slower than variations in throughput rates. This may be caused by the fact that changing the throughput rate only needs to affect the fill ratio of the screw to impact the extrudate whereas changes in wood content must also purge the entire screw (i.e. in a plug flow system, all the 20% wood composite needs to exit the die before the 30% wood composite can do so). Last, variations in screw speed lead to very short delays (T_d). Unlike variations in feed conditions, this manipulated parameter influences the entire length of the extruder at once.

A single transition does not appear to follow such a 1st order dynamic. The transition between samples 4 and 5 (-20°C temperature step). Figure 7.9A illustrates a significant excursion towards more severe thermo-mechanical energy transfer (i.e. high $\mathbf{t_2}$ values). This was caused by a temporary failure of the barrel temperature controller. The excursion along the positive $\mathbf{t_2}$ axis is associated with an abnormal increase in the entropy textural features for the transition samples. Looking at the samples themselves (Figure 7.9B), the presence of very distinct local anomalies in the composites is observed during this transitional state. Local agglomerations of wood fibers within the polymeric matrix can be seen. Such anomalies were not found elsewhere in the dataset. This illustrates the advantage of spectral imaging over single-point probes by their capability to identify localized defects.

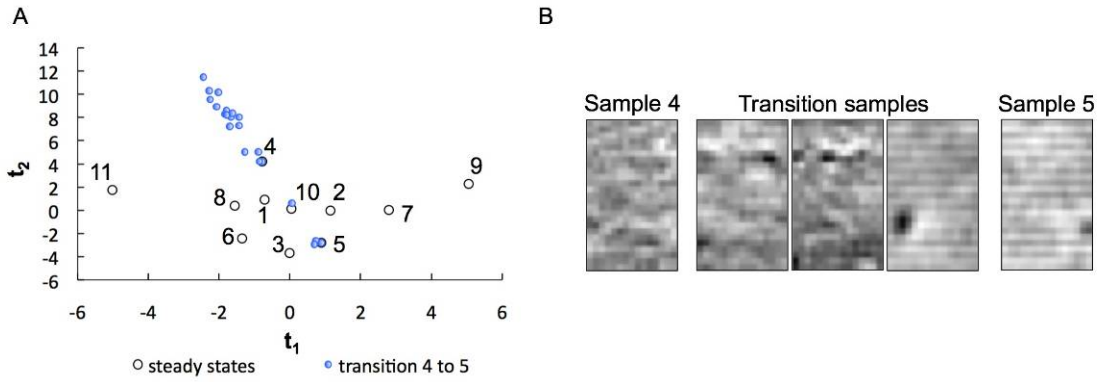

Figure 7.9. Abnormal transition 4 to 5 (-20°C temperature step) as seen in the PLS score plot (A), and a few images taken within the transitions indicating the presence of distinct local anomalies in the composite.

Whereas a sample's location within the score plot may be useful to determine its general characteristics (i.e. $\mathbf{t_1}$ is linked to spectral information, $\mathbf{t_2}$ to texture), it is also necessary to determine its true properties. It was possible to predict each of 7 measured properties for all the samples within the training and validation sets by adapting Equation 7.5:

$$\hat{\mathbf{Y}} = \mathbf{TQ}^{\mathrm{T}} \tag{7.9}$$

In accordance with the experimental plan (Section 7.2.1), the entire run between samples 1 and 8 was considered as a single run consisting of 20% wood samples in which the temperature, throughput rate and screw speed were varied (run 1). Likewise, samples 9-11 also consist of a single run in which the wood content was varied under nominal conditions (run 2). Beyond the 11th sample type (30% wood), the experimental conditions were returned to the 10th and 9th conditions (20 and 10% wood, respectively) in order to further validate the model. Figure 7.10 illustrates the predicted modulus and enthalpy of fusion for both experimental runs (run 1: 1-180 minutes run 2: 180-300 minutes). In this figure, the experimental dataset is shown in black, the values predicted by the training (steady states) set are shown in blue and the values predicted for the validation set (transitions) are shown in red. Results show that, in general, there is a good agreement between the measured and the predicted data. However, at times, the training set shows distinct biases for some samples: the modulus of sample 6 (≈ 120 minutes) is overestimated while the enthalpy of fusion of sample 2 is underestimated (≈ 30 minutes). This appears to be attributable to the

fact that the PLS model only accounted for 88.3% of the Y-variance ($R^2_{Y,cum}$), leaving 11.7% of the information unexplained. It is therefore no surprise that the validation data, which is obtained using the same model, shows the same biases the training set does. The transition results generally appear to fall between their respective steady-states as one would expect with first-order transitions.

Nevertheless, the problematic transition between samples 4 and 5 (≈ 90 minutes) clearly underestimates both properties. This difference may be caused by the fact that the laboratory tests did not happen to include some of the wood fiber agglomerations identified by the camera. This is not unexpected; a local anomaly in the image will influence the textural features of the image as a whole while the laboratory tests (especially DSC) only measure a portion of the sample. Last, the samples produced after 250 minutes were used solely for validation purposes. These samples show that the model could clearly identify both steady states (≈ 260 and 290 minutes) even though validation data within these steady states is clearly noisier than the equivalent steady states used to build the training set (sample 9 ≈ 190 minutes and sample 10 ≈ 210 minutes). Similar results were obtained for the other properties.

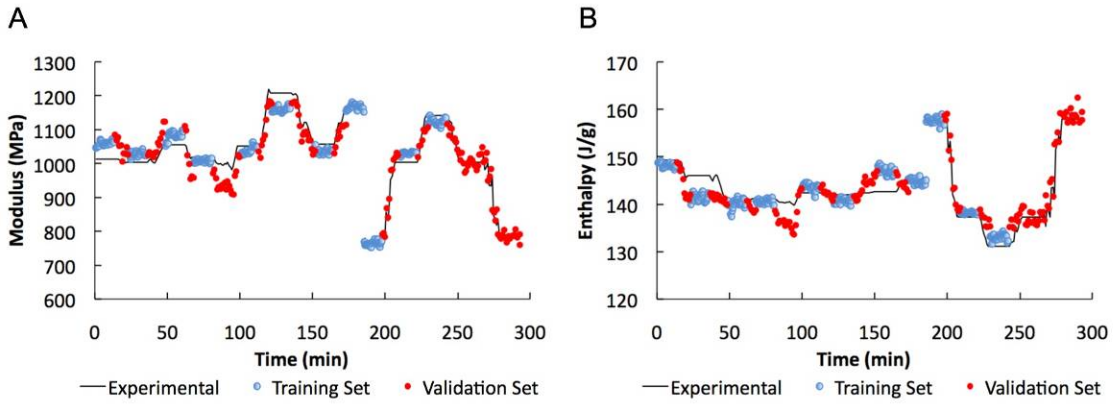

Figure 7.10. Modulus and enthalpy of fusion transitions comparing the experimental values (black line), to the values predicted by the training set (blue dots) and the validation set (red dots).

Multivariate statistical process monitoring charts based on the PLS scores and projection residuals could be directly used for quality control of extruded composite products as was

achieved in a wide variety of fields (Kourti, 2002, 2005; Liu and MacGregor, 2008). This would enable the operator to quickly identify products that are off-specification due to disturbances such as variations in the feed (e.g. wood content) and operational problems such as severity of thermo-mechanical energy transfer, a faulty barrel temperature controller, etc. Appropriate corrective actions could then be applied with the help of the process understanding provided by the PLS model (i.e. loadings biplots presented in Figure 7.7).

7.4 Conclusion

The objective of this work was to develop a rapid and non-intrusive on-line imaging sensor for monitoring spatio-temporal variations in quality across the surface of thin polymer composite materials. This paper discusses the ability of chemometrics methods, to extract relevant spectral and textural information from the visible/NIR spectral images and link it to sample quality. On one hand, the spectral signature of a sample reflects its chemical composition and small-scale physical structure. The significant wavelength intervals included in the model were computed using a PLS-bootstrap based method previously developed by Gosselin et al. (2010). On the other hand, the spatial information (visual texture) represents large-scale variations throughout the sample. It was quantified through the gray-level co-occurrence matrix (GLCM) algorithm, a common, simple and effective texture analysis method.

The proposed approach was tested on extruded wood plastic composites. These samples were produced under conditions varying in barrel temperature, screw speed, throughput rate and wood content. Several properties were quantified: enthalpy of fusion, modulus, stress at yield and at break, strain at yield and at break as well as the energy at break. Overall, 11 steady-state samples were produced to build the PLS regression model. The model was subsequently used to predict the quality of samples produced while changing from one steady-state condition to another (transient state).

The results show that the proposed methodology can (1) distinguish between the quality of the product obtained under the 11 sets of steady-state operating conditions, (2) predict the 7 quality variables of the samples in both the steady-states and the transitions, (3) estimate the extrusion dynamics, and (4) provide an early detection scheme for undesired changes in feed composition (e.g. wood content) as well as in the severity of the thermo-mechanical energy delivered to the systems, which may lead to degradation in product quality. It is believed that such a spectral imaging sensor can provided valuable information for establishing an on-line product quality control strategy for the polymer processing industry.

7.5 References

Akdogan H. and Rumsey T.R., Dynamic response of a twin-screw lab-size extruder to changes in operating variables, *Lebensm.-Wiss. Technol.*, 29, 691-701 (1996).

Albano C., Papa J., Gonzalez E., Navarro O. and Gonzalez R., Temperature and crystallinity profiles in polyolefines isothermal and non-isothermal solidification processes, *Eur. Polym. J.*, 39, 1215-1222 (2003).

Bajimaya S.M., Park S. and Wang, G., Predicting extrusion process parameters using neural networks, *PWASET*, 26, 429-433 (2007).

Barnes S.E., Sibley M.G., Edwards H.G.M. and Coates P.D., Applications of process spectroscopy to polymer melt processing, *Spectrosc. Eur.*, 15, 22-24 (2003).

Bergmann K., Determination of polymer crystallinity from proton solid-echo NMR measurements, *Polym. Bull.*, 5, 355-360 (1981).

Boger, Z., Selection of quasi-optimal inputs in chemometrics modeling by artificial neural network analysis, *Anal. Chim. Acta*, 490, 31-40 (2003).

Bower, D.I., & Maddams, W.F., *The Vibrational Spectroscopy of Polymers*, Cambridge University Press: USA (1992).

Brás L.P., Lopes M., Ferreira A.P. and Menezes J.C., A bootstrap-based strategy for spectral interval selection in PLS regression, *J. Chemom.*, 22, 695-700 (2008).

Broadhead T.O., Nelson B.I. and Dealy J.M., An in-line rheometer for molten plastics-design and steady state performance characteristics, *Int. Polym. Proc.*, 8, 104-112 (1993).

Cai W., Li Y. and Shao, X., A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra, *Chemom. Intell. Lab. Syst.*, 90, 188-194 (2008).

Cayot N., Bounie D. and Baussartz H., Dynamic modelling for a twin screw food extruder: analysis of the dynamic behaviour through process variables, *Int. J. Food Eng.*, 25, 245-260 (1995).

Chan D., Nelson R.W. and Lee J., Dynamic behavior of a single screw plasticating extruder part II: dynamic modeling, *Polym. Eng. Sci.*, 26, 152-161 (1986).

Chen T., Nguyen K.T., Wen S.L. and Jen C., Temperature measurement of polymer extrusion by ultrasonic techniques, *Meas. Sci. Technol.*, 10, 139-145 (1999).

Chong I. and Jun C., Performance of some variable selection methods when multicollinearity is present, *Chemom. Intell. Lab. Syst.*, 78, 103-112 (2005).

Costin M.H., Taylor P.A. and Wright, J.D., On the dynamics and control of a plasticating extruder, *Polym. Eng. Sci.*, 22, 1095-1106 (1982).

Côté P. and Abatzoglou N., Powder and other divided solids mixing. Scale-up and parametric study of a ribbon blender used in pharmaceutical powders mixing, *Pharma. Dev. Technol.*, 11, 29-45 (2006).

Cournoyer A., Simard J.S., Cartilier L. and Abatzoglou N., Quality control of multi-component, intact pharmaceutical tablets with three different near-infrared apparatuses, *Pharma. Dev. Technol*., 13, 333-343 (2008).

De Temmerman J., Saeys W., Nicolaï B. and Ramon H., Near infrared reflectance spectroscopy as a tool for the in-line determination of the moisture concentration in extruded semolina pasta, *Biosystems Eng.*, 97, 313-321 (2007).

Dumoulin M.M., Gendron R. and Cole K.C., Techniques for real-time monitoring of polymer processing, *Trends Polym. Sci.*, 4, 109-114 (1996).

ElMasry G., Wang N., Vigneault C., Qiao J. and ElSayed A., Early detection of apple bruises on different background colors using hyperspectral imaging, *LWT*, 41, 337-345 (2008).

Eriksson L., Johansson, E., Kettaneh-Wold, N. and Wold S., *Multi- and Megavariate Data Analysis*, Umetrics Academy: Sweden (2001).

Fodil-Pacha F., Arhaliass A., Aït-Ahmed N., Boillereaux L. and Legrand J., Fuzzy control of the start-up phase of the food extrusion process, *Food Control*, 18, 1143-1148 (2007).

Ghaneh-Fard A., Effects of film blowing conditions on molecular orientation and mechanical properties of polyethylene films. *J. Plast. Film Sheet.*, 15, 194-218 (1999).

Gosselin R., Rodrigue D. and Duchesne C., On-line prediction of crystallinity spatial distribution across polymer films using NIR spectral imaging and chemometrics methods, *Can. J. Chem. Eng.*, 86, 869-878 (2008).

Gosselin R., Rodrigue D., González-Núñez R. and Duchesne C., Potential of hyperspectral imaging for quality control of polymer blend films, *Ind. Eng. Chem. Res.*, 48, 3033-3042 (2009).

Gosselin R., Rodrigue D. and Duchesne C., A bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications, *Chemom. Intell. Lab. Syst.*, 100, 12-21 (2010).

Gupta B.S., Reiniati I. and Laborie M.G., Surface properties and adhesion of wood fiber reinforced thermoplastic composites, *Colloids Surf., A*, 302, 388-395 (2007).

Ha Y.S., Cho J.R., Kim T.H. and Kim J.H., Finite element analysis of rubber extrusion forming process for automobile weather strip, *J. Mater. Process. Technol.*, 201, 168-173 (2008).

Haley T.A. and Mulvaney S.J., On-line system identification and control design of an extrusion cooking process: Part I. System identification, *Food Control*, 11, 103-120 (2000).

Hansen M.G. and Vedula S., In-line fiber-optic near-infrared spectroscopy: monitoring of rheological properties in an extrusion process. Part I, *J. Appl. Polym. Sci.*, 68, 859-872 (1998).

Haralick R.M., Statistical and structural approaches to texture, *Proceedings IEEE*, 67(5), 786-804 (1979).

Haralick R.M., Shanmugam K. and Dinstein I., Textural features for image classification, *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, 6, 610-621 (1973).

Hotelling H., The generalization of Student's ratio, *Ann. Math. Statist.*, 2, 360-378 (1931).

Ilharco L.M., Garcia A.R., da Silva J.L., Lemos M.J. and Vieira Ferreira L.F., Ultraviolet-visible and Fourier transform infrared diffuse reflectance studies of benzophenone and fluorenone adsorbed onto microcrystalline cellulose, *Langmuir*, 13, 3787-3793 (1997).

Kourti T., Process analysis and abnormal situation detection: from theory to practice, *IEEE Control Syst. Mag.*, 22(5), 10-25 (2002).

Kourti T., Application of latent variable methods to process control and multivariate statistical process control in industry, *Int. J. Adapt Control Signal Process.*, 19, 213-246 (2005).

Li T.Q. and Wolcott M.P., Rheology of HDPE-wood composites I. Steady state shear and extensional flow. *Composites Part A.*, 35(3), 303-311 (2004).

Liu F., He Y. and Wang L., Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis, *Anal. Chim. Acta*, 615, 10-17 (2008).

Liu J.J. and MacGregor J.F., Modeling and optimization of product appearance: application to injection-molded plastic panels, *Ind. Eng. Chem. Res.*, 44, 4687-4696 (2005).

Liu J.J. and MacGregor J.F., Froth-based modeling and control of flotation processes, *Miner. Eng*., 21, 642-651 (2008).

Mardia K.V., Kent J.T. and Bibby J.M., *Multivariate Analysis*, Academic Press: Great Britain (1979).

McAfee M. and McNally G., Real-time measurement of melt viscosity in single-screw extrusion, *T. I. Meas. Control*, 28, 481-497 (2006).

Merikoski S., Laurikkala M. and Koivisto H., Modelling viscosity in rubber mixing process using an adaptive neuro-fuzzy inference system (ANFIS), *Automaatio2001* (2006).

Mezreb K., Goullieux A., Ralainirina R. and Queneudec M., Application of image analysis to measure screw speed influence on physical properties of corn and wheat extrudates, *J. Food Eng.*, 57, 145-152 (2003).

Mudalamane R. and Bigio D.I., Process variations and the transient behavior of extruders, *AIChE J.*, 49, 3150-3160 (2003).

Ohshima M. and Tanigaki M., Quality control of polymer production process, *J. Process Contr.*, 10, 135-148 (2000).

Olah M., Bologa C. and Oprea T.I., An automated PLS search for biologically relevant QSAR descriptors, *J. Comput. Aid. Mol. Des.*, 18, 437-449 (2004).

Pantani R., Coccorullo I., Speranza V. and Titomanlio G., Modeling of morphology evolution in the injection molding, process of thermoplastic polymers, *Prog. Polym. Sci.*, 30, 1185-1222 (2005).

Pandey K.K., Upreti N.K. and Srinivasan V.V., A fluorescence spectroscopic study on wood, *Wood Sci. Technol.*, 32, 309-315 (1998).

Pastore T.C.M., Santos K.O. and Rubim, J.C., A spectrocolorimetric study on the effect of ultraviolet irradiation of four tropical hardwoods, *Bioresource Technol.*, 93, 37-42 (2004).

Pelsoci T., *ATP-Funded Green Process Technologies: Improving U.S. Industrial Competitiveness with Applications in Packaging, Metals Recycling, Energy, and Water*

Treatment, National Institute of Standards and Technology, U.S. Department of Commerce: USA (2007).

Reber D.H., Lynn R.E. Jr. and Freeh E.J., A mathematical model for predicting dynamic behavior of a plasticating extruder, *Polym. Eng. Sci.*, 13, 346-356 (1973).

Ren Y., Shimoyama M., Ninomiya T., Matsukawa K., Inoue H., Noda I. and Ozaki Y., Two-dimensional near-infrared correlation spectroscopy studies on composition-dependent spectral variations in ethylene/vinyl acetate copolymers: assignments of bands due to ethylene units in amorphous, disordered, and orthorhombic crystalline phases, *Appl. Spectrosc.*, 53(8), 919-916 (1999).

Reshadat R., Desa S., Joseph S., Mehra M., Stoev N. and Balke T., In-line near-infrared monitoring of polymer processing. Part I: process/monitor interface development, *Appl. Spectrosc.*, 53(11), 1412-1418 (1999).

Reynes C., de Souza S., Sabatier R., Figueres G. and Vidal B., Selection of discriminant wavelength intervals in NIR spectrometry with genetic algorithms, *J. Chemom*., 20, 136-145 (2006).

Rohe T., Becker W., Kölle S., Eisenreich N. and Eyerer P., Near infrared (NIR) spectroscopy for in-line monitoring of polymer extrusion processes, *Talanta*, 50, 283-290 (1999).

Setarehdan S.K., Fuzzy-based purest wavelength selection from spectral data, *J. Chemom.*, 20, 239-246 (2006).

Shaw W.T. and Lee K.T.A., Copula methods vs. canonical multivariate distributions: the multivariate Student T distribution with general degrees of freedom, *KCL Working Paper*, 1-25 (2007).

Shihani N., Kumbhar B.K. and Kulshreshtha M., Modeling of extrusion process using response surface methodology and artificial neural networks, *J. Eng. Sci. Technol.*, 1, 31-40 (2006).

Stark N.M. and Matuana L.M., Characterization of weathered wood-plastic composite surfaces using FTIR spectroscopy, contact angle, and XPS, *Polym. Degrad. Stabil.*, 92, 1883-1890 (2007).

Tadmor Z., Lipshitz S.D. and Lavie R., Dynamic model of a plasticating extruder, *Polym. Eng. Sci.*, 14, 112-119 (1974).

Valadez-Blanco R., Virdi A.I.S., Balke S.T. and Diosady L.L., In-line colour monitoring during food extrusion: sensitivity and correlation with product colour, *Food Res. Int.*, 40, 1129-1139 (2007).

Watari M., Mitsui N., Higashiyama H. and Tomo M., On-line molten polymer measurement using near infrared Fourier transfer spectroscopy, *Tech. Info. Yokogawa*, TI11VOA1-11 (1996).

Watari M., Higashiyama H., Mitsui N., Tomo M. and Ozaki Y., On-line monitoring of the density of linear low-density polyethylene in a real plant by near-infrared spectroscopy and chemometrics, *Appl. Spectr.*, 58(2), 248-255 (2004).

Wold S., Geladi P., Esbensen K., Ohman J., Multi-way principal component and PLS analysis, *J. Chemom.*, 1, 41-56 (1987).

WRAP, *Wood Plastic Composites Study Research Report – Technologies and UK Market Opportunities, Waste and Resources Action Programme*, Old Academy: United-Kingdom (2003).

Yuan X.F., Ball R.C. and Edwards S.F., Dynamical modeling of viscoelastic extrusion flows, *J. Non-Newt. Fluid Mech.*, 54, 423-435 (1994).

Chapter 8

Thesis conclusion

The title of this thesis, *"On-line Quality Control in Polymer Processing using Hyperspectral Imaging"*, is intended to reflect the major themes discussed in this work.

- The quality issues associated with the production of polymer materials and the limits of feedback control loops to address these quality issues.
- The potential of spectral imaging sensors to provide a wealth of information that most common probe sensors cannot.
- The notion that combining spectral imaging cameras and chemometrics methods can lead to better understanding and better quality control of complex composite materials.

The stated objective of this work was to develop an on-line quality control tool adaptable to plastic materials of all types. Specifically, one with high spatial resolution able of detecting quality gradients and/or localized defects in the finished product. In order to achieve this, a visible and infrared (VIS-NIR) spectral imaging sensor was proposed. As our results have shown, this methodology made it possible to monitor spatio-temporal variations in plastic products (both neat and composite plastics) with a degree of precision not yet achieved in the field by remote sensing probes.

The experimental results obtained during this thesis are presented in four chapters, each based on articles either published, accepted or submitted to international scientific journals. As mentioned in the outline of the thesis, these are:

1. Crystallinity predictions. Determination of spatial crystallinity patterns across the surface of neat polymer films using spectral methods.
2. Polymer blend films. Monitoring of PE/PS blends produced via extrusion film-blowing using a combined spatial/spectral methodology.

3. Wood/plastic composites. Monitoring of extruded wood/plastic composites using independent spatial and spectral methodologies.
4. Wavelength selection. Reducing the number of wavelengths in the spectral imaging data using a latent variable methodology in order to produce more parsimonious remote sensing models.

These four chapters are linked to one another and can be interpreted sequentially. First, the determination of crystallinity patterns across the surface of neat polymer films was used as an initial step in the monitoring of polymer composites. In essence, one must determine the quality of a simple system before moving on to a more complex one. In this work, subtle variations in crystallinity were introduced in the films (LDPE, HDPE and PP) by controlling the cooling rates at which they were produced. These variations were then identified using an image analysis and regression approach (MIA/MIR) based on the near infrared reflectance spectra of the samples. By doing so, these latent variable techniques were shown to outperform other methods based on the 2^{nd} derivative of NIR spectra. Finally, such predictive power was also consistently high for all three polymer species. This feature of the proposed approach shows an interesting potential for application to more complex materials such as polymer blends or composites.

Second, the monitoring of polymer film blend produced via extrusion film-blowing was used as a case study of polymer-polymer blends. In this work, a multiresolutional multivariate image analysis (MR-MIA) methodology was used to extract the spatial and spectral features that are most highly correlated to the film properties (i.e. composition distribution and mechanical properties). Applied to PE/PS blends produced under various operating conditions, the imaging sensor was able to monitor the two dominant effects on mechanical properties: the effect of film composition (raw materials) and the stretching effect caused by post-extrusion conditions (process).

Third, the monitoring of extruded wood/plastic composites was used as a case study of polymer-filler blends. In this work, a two-step method was used to extract the significant spatial and spectral image features and link them to the material properties. Unlike the film

blowing dataset, the seven properties of interest were measured not only during the steady-state operating conditions, but also during the transition periods. The results show that the proposed methodology can distinguish between the quality of the product obtained under steady-state operating conditions, predict the quality of the samples in both the steady-states and the transition states, estimate the extrusion dynamics, and provide early detection of abnormal operation when unforeseen situations occur.

Fourth, a simple wavelength selection method having the ability to identify relevant spectral intervals was developed for spectral imaging applications and spectroscopy in general. While the work described in this chapter is not specific to composite materials, it plays an important role in increasing the robustness and interpretability of any multivariate image regression model. The proposed approach combined PLS modeling of the complete spectral dataset and a bootstrap procedure, making it possible to compute consistency intervals around the PLS VIP metrics of each wavelength. Wavelengths with VIP values above a certain cutoff value were deemed as important and kept for further modeling. Using four different datasets, it was shown that a small number of well-defined relevant spectral intervals were identified with the proposed approach and that these intervals were in agreement with more complex interval search methods.

Concluding remarks

These four chapters encompass much of the work carried to fruition during this project. They represent, however, only an introduction to what promises to be a complex and rewarding field of inquiry. Even in the short term, many more interesting projects could, and should, be tackled:

- In all of these tests, the imaging system has been physically decoupled from the extruder, implying that the image analysis be done off-line. This need not to be the case. Mounting the sensors onto the extruder would make it possible to test the remote-sensing model in real-time.

- Only work on flat-surfaced samples was done in order to help with the image capture. However, there is no reason why, with proper consideration of the lighting patterns, such a methodology could not be extended to more complex shapes. A cylindrical extrudate may be the perfect candidate to test the principle.
- Also, only thin samples were analyzed, reducing possible discrepancies between surface and core properties. A study based on thicker samples may lead to new insights relating to core/surface properties. Moreover, preliminary tests indicate that NIR waves travel significant distances inside WPCs (of the order of millimeters) and that and different wavelengths have different penetration capabilities. This suggests that a spectral NIR sensor may be used as a non-destructive means of obtaining property profiles inside thicker composite samples. While these observations appear promising, they remain highly case-specific.

It is also important to consider that these tests were performed in academic settings under controlled laboratory conditions and should be considered as a proof of concept. While they show promise, they need to be put to the test in industrial settings. A number of points must be taken into account.

- First, it is of the utmost importance to identify the proper remote sensing device for the case at hand. While it should be very sensitive to the material properties found to influence the quality of the product, its performance should be balanced with other criteria such as price and ease of use if it is to be industrially viable. In other words, the costs of highly technological solutions may be found to outweigh the benefits.
- Second, it is important to considerer the nature of the production line in question. While some go virtually unchanged over long periods, others undergo constant changes (e.g. product shape or composition). In light of this, the remote sensing device must be sufficiently adaptable to suit the situation under consideration. Beyond this, it is necessary to ascertain model validity through routine maintenance if accuracy is to be maintained over time.

- Third, if the chosen device is to be successfully implemented, it must be seen by both the administration and the production line operator to provide a clear and obvious advantage over the preceding situation. While it is up to the plant administration to decide to implement such a system, it is up ultimately to the operator to determine its usefulness. In this way, the system must be considered simple, effective and reliable by all parties concerned.
- Finally, several technological difficulties must also be considered. It is of paramount importance that the system copes with its intended environment. Such limitations may come in many forms. Among these, let us mention high throughput rates (i.e. short acquisition times), extreme operation temperatures, soiling (through dust or other contaminants), uneven lighting conditions, mechanical stresses (e.g. vibrations or impacts), as well as any limitations due to hardware maintenance.

In closing, these image analysis techniques are not specific to any material, process or field. At their most general, they can be adapted to any product that may be distinguished by its spectral and/or spatial surface characteristics. As with all types of surface analysis, these techniques should only be used with caution on thicker samples. Nevertheless, even this restriction does not prevent them from being used to monitor the operating conditions rather than inferring product properties. By doing so, such a sensor could provide valuable information in a wide range of on-line quality control strategies.