MULTIRESOLUTIONAL PARTIAL LEAST SQUARES AND PRINCIPAL COMPONENT ANALYSIS OF FLUIDIZED BED DRYING

A Thesis Submitted to the College of

Graduate studies and Research

in Partial Fulfilment of the Requirements

for a Degree of Master of Science

in the Department of Chemical Engineering

University of Saskatchewan

Saskatoon, Saskatchewan, Canada

By

Gerald Michael Frey

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make if freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Chemical Engineering

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5C5

ABSTRACT

Fluidized bed dryers are used in the pharmaceutical industry for the batch drying of pharmaceutical granulate. Maintaining optimal hydrodynamic conditions throughout the drying process is essential to product quality. Due to the complex interactions inherent in the fluidized bed drying process, mechanistic models capable of identifying these optimal modes of operation are either unavailable or limited in their capabilities. Therefore, empirical models based on experimentally generated data are relied upon to study these systems.

Principal Component Analysis (PCA) and Partial Least Squares (PLS) are multivariate statistical techniques that project data onto linear subspaces that are the most descriptive of variance in a dataset. By modeling data in terms of these subspaces, a more parsimonious representation of the system is possible. In this study, PCA and PLS are applied to data collected from a fluidized bed dryer containing pharmaceutical granulate.

System hydrodynamics were quantified in the models using high frequency pressure fluctuation measurements. These pressure fluctuations have previously been identified as a characteristic variable of hydrodynamics in fluidized bed systems. As such, contributions from the macroscale, mesoscale, and microscales of motion are encoded into the signals. A multiresolutional decomposition using a discrete wavelet transformation was used to resolve these signals into components more representative of these individual scales before modeling the data.

The combination of multiresolutional analysis with PCA and PLS was shown to be an effective approach for modeling the conditions in the fluidized bed dryer. In

this study, datasets from both steady state and transient operation of the dryer were analyzed. The steady state dataset contained measurements made on a bed of dry granulate and the transient dataset consisted of measurements taken during the batch drying of granulate from approximately 33 wt.% moisture to 5 wt.%. Correlations involving several scales of motion were identified in both studies.

In the steady state study, deterministic behavior related to superficial velocity, pressure sensor position, and granulate particle size distribution was observed in PCA model parameters. It was determined that these properties could be characterized solely with the use of the high frequency pressure fluctuation data. Macroscopic hydrodynamic characteristics such as bubbling frequency and fluidization regime were identified in the low frequency components of the pressure signals and the particle scale interactions of the microscale were shown to be correlated to the highest frequency signal components. PLS models were able to characterize the effects of superficial velocity, pressure sensor position, and granulate particle size distribution in terms of the pressure signal components. Additionally, it was determined that statistical process control charts capable of monitoring the fluid bed hydrodynamics could be constructed using PCA.

In the transient drying experiments, deterministic behaviors related to inlet air temperature, pressure sensor position, and initial bed mass were observed in PCA and PLS model parameters. The lowest frequency component of the pressure signal was found to be correlated to the overall temperature effects during the drying cycle. As in the steady state study, bubbling behavior was also observed in the low frequency components of the pressure signal. PLS was used to construct an inferential model

of granulate moisture content. The model was found to be capable of predicting the moisture throughout the drying cycle. Preliminary statistical process control models were constructed to monitor the fluid bed hydrodynamics throughout the drying process. These models show promise but will require further investigation to better determine sensitivity to process upsets.

In addition to PCA and PLS analyses, Multiway Principal Component Analysis (MPCA) was used to model the drying process. Several key states related to the mass transfer of moisture and changes in temperature throughout the drying cycle were identified in the MPCA model parameters. It was determined that the mass transfer of moisture throughout the drying process affects all scales of motion and overshadows other hydrodynamic behaviors found in the pressure signals.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Aaron Phoenix, for his guidance, time, and extraordinary support throughout the course of this research. I would also like to sincerely thank my supervisor, Dr. Todd Pugsley, for providing an interesting and challenging research project and for his guidance, time, and encouragement.

Thank you to the faculty and staff of Department of Chemical Engineering for providing a positive atmosphere throughout my university education. In particular, I am grateful to Dr. Richard Evitts and Dr. Gordon Hill for their time and participation in my advisory committee.

This thesis is based upon data generated in an experimental campaign conducted by Mr. Gareth Chaplin at Merck-Frosst Laboratories in Kirkland, Quebec. I would like to thank Gareth for sharing his work, helping me to understand the data, and answering my many questions. My thanks are also extended to Merck-Frosst for their support of the original experimental work and for use of the data.

Thank you to Dr. Wey Lee for introducing me to the field of multivariate data analysis. These concepts provided the foundation for this work.

Finally, I am forever grateful to my parents for impressing upon me the importance of education. Without their support, there is no doubt that I could not have reached this level of achievement.

DEDICATION

To my beautiful wife Jennifer,

Thank you for your love and support. Your patience and understanding made this accomplishment possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

A5, A4, A3…	approximation function at level 5, 4, 3, …
b	PLS regression vector
B	PLS regression matrix
C	correlation matrix
d5, d4, d3,…	detail function at level 5, 4, 3, ….
$\mathbf{e}_i$	X-block error vector
E	X-block error matrix
F	Y-block error matrix
$F_{k,m-k,\alpha}$	F distribution at an 100 - α confidence level
$\mathbf{G}_n$	high-pass wavelet filter convolution matrix
$\mathbf{H}_n$	low-pass wavelet filter convolution matrix
i	index variable
$\hat{\mathbf{i}}$	unit vector in the x-direction
I	identity matrix
$\hat{\mathbf{j}}$	unit vector in the y-direction
k	number of principal components retained in a model
m	number of rows or observations in data matrix, **X**
n	number of variables in data matrix, **X**
o	batches in three-dimensional data matrix, $\underline{\mathbf{X}}$
$\mathbf{p}_i$	eigenvector corresponding to eigenvalue *i*
P	eigenvectors of the correlation matrix, **C**
q	Y-block load vector

$\mathbf{Q}$	Y-block load matrix
$Q_{residual}$	$Q_{residual}$ statistic, Q-statistic, SPE statistic
$\mathbf{q}_{con}$	Q contribution statistic equal to elements of **E**
$\mathbf{S}$	XY correlation matrix
$S(n)$	a discrete signal or function
t_i	Student's t-statistic for PC i
$t_{\upsilon,\alpha/2}$	Student t-statistic for υ d.o.f. at 100(1-α)%
$\mathbf{t}$	X-block score vector for a single sample
$\mathbf{t}_i$	Score vector for sample or observation i
$\mathbf{T}$	X-block score matrix
T^2	Hotelling's T^2 statistic
$\mathbf{t}_{con}$	Hotelling's T^2 contribution statistic
$\mathbf{u}$	Y-block score vector for a single sample
$\mathbf{U}$	Y-block score matrix
$\mathbf{v}$	normalized loading vectors
$\mathbf{V}$	Matrix to maintain orthonormal basis of **P**
$\mathbf{w}$	PLS weight vector
$\mathbf{W}$	PLS weight matrix
$\mathbf{x}_i$	ith column of data set **X**
$\mathbf{X}$	m x n data set consisting of m observations on n variables
$\underline{\mathbf{X}}$	dataset of m observations on n variables over o batches
y	output or response
$\hat{y}$	predicted output or response

$\mathbf{Y}$	output variable data matrix
z_α	standard normal deviate for 1 - α confidence level

GREEK SYMBOLS

α	level of confidence
$\hat{\xi}$	PLS confidence interval for $\hat{y}$ for a single sample
ϕ_i	angle to axis representing the *i*th variable
$\varphi(t)$	scaling function
λ_i	the *i*th eigenvalue of the correlation matrix
μ_i	mean value of column *i*
σ_i	standard deviation of column *i*
$\psi(t)$	mother wavelet function
$\psi_{j,k}(t)$	wavelet function set

ABBREVIATIONS

db1, db2,…	Daubechies 1 wavelet, Daubechies 2 wavelet,….
domfreq	dominant frequency
DWT	Discrete Wavelet Transform
FBD	Fluid Bed Dryer
FFT	Fast Fourier Transform
inT	Fluid Bed Dryer inlet air temperature
kurt	kurtosis

LV	Latent Variable
medn	median
moist	granulate moisture
MPCA	Multiway Principal Component Analysis
MRA	Multiresolutional Analysis
MSECV	Mean Squared Error of Cross Validation
MSPC	Multivariate Statistical Process Control
MSPCA	Multiscale Principal Component Analysis
mu	mean
NIPALS	non-iterative partial least squares algorithm
outT	Fluid Bed Dryer exhaust stream temperature
PC	Principal Component
PC*i*	Principal Component *i*
PCA	Principal Component Analysis
PLS	Partial Least Squares or Projection to Latent Structures
posn	pressure transducer position
PRESS	Predicted Error Sum of Squares
prodT	Fluid Bed Dryer bed temperature
PSD	Particle Size Distribution
RMSECV	Root Mean Squared Error of Cross Validation
sgn	particle size distribution
SIMPLS	statistically inspired modification of PLS
stdev	standard deviation

skew	skewness
STFT	Short-time Fourier Transform
SVD	Singular Value Decomposition
vel	superficial velocity

1. INTRODUCTION

Fluidized bed dryers (FBD) are used in industry for the drying of a variety of products. In most cases, moisture content is critical to final product quality. Fluid bed technology is typically employed because it is characterized by easy solids handling, efficient mixing, a relatively uniform bed temperature, and excellent heat and mass transfer. Although the use of fluidized bed dryers is common, the process is not always devoid of problems.

In the pharmaceutical industry, fluidized bed dryers are used for the batch drying of pharmaceutical granulate. The process is typically controlled by monitoring the temperatures of the bed and outlet air as an indirect measure of product moisture. A drawback of this approach is that it provides no information regarding the hydrodynamic state of the bed. Uneven distribution of conditions in the bed can lead to hot spots or, alternatively, regions containing excessive moisture. Hot spots can reach excessive temperatures that damage the drug properties.

Changes in particle size distribution (PSD) can also be a major concern. Agglomeration of particles early in the process can lead to defluidization in part or all of the bed. In the late stages of drying, attrition of particles can lead to the entrainment of fines in the air stream as well as product quality issues. A means of

monitoring and controlling the hydrodynamic state would help to minimize these adverse effects and result in increased productivity and higher product quality.

Determination of optimal hydrodynamic conditions is a challenging proposition due to the complex interactions inherent in the fluidized bed drying process. This undertaking is further complicated by transient conditions resulting from mass transfer. As a result, mechanistic models capable of identifying these optimal modes of operation are either unavailable or limited in their capabilities. Therefore, empirical models based on experimentally generated data are relied upon to study these systems.

Experimentally based quantification of fluidized bed hydrodynamics begins with physical measurements on the systems under study. Several approaches have been employed for this purpose including fibre optical probes, temperature readings, differential pressure measurements, pressure fluctuation measurements, and various tomography techniques (Werther, 1999). While each approach has its advantages, high frequency pressure fluctuation measurements have been the most practical technique for providing information about fluidized bed hydrodynamics to date.

Absolute pressure fluctuations obtained with high frequency pressure transducers are comprised of compression waves moving both upwards and downwards through the bed (van der Schaaf, Schouten, and van den Bleek, 1998). Consequently, a single high frequency pressure transducer can provide information about the dynamics of the entire bed. These transducers are non-invasive, inexpensive, and can be easily installed in laboratory, pilot, and commercial scale fluidized bed systems.

High frequency pressure fluctuations are a characteristic variable of system hydrodynamics with information from several scales of motion contributing to the signal features (Ren and Li, 1998; Johnsson et al., 2000). The scales of motion in gas-solid fluidized beds can be divided into macroscale, mesoscale, and microscale categories. The definitions of each of these scales can vary and the convention used by Li (2000) has been adopted in this study. Using this convention, the macroscale motion refers to the bulk motion of the particle-fluid suspension within the boundaries of the apparatus and encompasses bubbling phenomenon and apparatus-scale effects. The mesoscale, also known as the cluster scale, involves interaction between the dilute broth phase and the dense cluster phase. Finally, the microscale refers to particle-scale phenomenon involving particle-particle and particle-apparatus interactions.

The traditional approach for analyzing pressure fluctuation data has been to use dominant frequencies and descriptive statistics such as variance, skewness, and kurtosis to summarize the signal features (Chong et al., 1987; Johnsson et al., 2000). This has been somewhat successful in classifying fluidization regimes but tends to provide little information beyond the macroscale behavior of the bed. The failure of these methods to identify mesoscale and microscale behaviors can be attributed to limitations in the resolution of the analysis techniques as well as insufficient data collection rates.

While the macroscale behavior in fluidized beds can be identified at frequencies of 1 to 9 Hz, more recent studies have considered that data collection rates as high as 400 Hz are required to capture microscale phenomenon (van Ommen,

Schouten, and van den Bleek, 1999). As such, data analysis techniques capable of differentiating between the macroscale and the higher frequency behaviors have become necessary. Techniques such as Hurst analysis (Ellis et al., 2003, Zhao and Yang, 2003), chaos analysis (van Ommen et al., 2000), and wavelet transformations (Ren and Li, 1998) have successfully identified higher frequency phenomenon.

Multiresolutional decomposition of pressure fluctuation signals using wavelet transformations is arguably the most promising of these techniques for analyzing the scale-specific behavior of fluidized bed systems. Using this approach, signals can be resolved into components more representative of the individual scales of motion (Lu and Li, 1999; Guo, Yue, and Werther, 2002; Ellis et al., 2003; Zhao and Yang, 2003). However, this approach is still in its infancy and, unlike chaos analysis (van Ommen et al, 2000; Chaplin, Pugsley, and Winters, 2004a-b), there have been no documented attempts of this technique being applied for process monitoring in fluidized bed systems.

Principal Component Analysis (PCA) and Partial Least Squares (PLS) are multivariate statistical data analysis techniques which have successfully been used for modeling, analysis, and multivariate statistical process control (MSPC) of complex industrial processes (Kourti and MacGregor, 1996; Teppola et al., 1998; Wise et al., 1999). The strength of these techniques lie in their ability to transform highly correlated, redundant, and noisy data into a model whose components often offer insight into the underlying physical phenomena and relationships involved in the process. Additionally, when used for statistical process control, the methods not only have the ability to detect faults but also allow for diagnosis of the most likely causes

of the disturbance. This diagnosis, along with some knowledge of the process, allows process operators to make any necessary corrections.

1.1 Objectives

Despite the success of PCA and PLS in other applications, there have been no documented attempts to use these techniques for analyzing the hydrodynamics of fluidized bed systems. Furthermore, with the exception of Chaplin, Pugsley, and Winters (2004a-b), attempts to monitor the hydrodynamics in a fluidized bed dryer are also absent from the literature. This thesis attempts to address these omissions by assessing the suitability of PCA and PLS for the modeling, analysis, and statistical control of the fluidized bed drying process. More specifically, the objectives of the current work are:

(i) The first objective of this study is to determine whether PCA and PLS modeling techniques can be applied to data collected from a fluidized bed drying process to analyze the effects of several manipulated variables. PCA and PLS are techniques that are known for being suitable for the analysis of noisy, highly correlated, complex, multivariate data. The fluidized bed drying process can be characterized by these same attributes.

(ii) Pressure fluctuations collected at high sampling rates in fluidized beds are said to contain contributions from several scales of motion. Multiresolutional analysis using discrete wavelet transformations has shown promise for exploiting scale specific behaviours in these pressure signals. The second objective of this study is to determine whether combining a multiresolutional representation of the pressure signals with PCA and PLS is capable of

extracting additional information regarding the hydrodynamic behaviours of the fluidized bed drying process. If this is the case, correlations involving individual scales of motion may be identified. Intuitively, it is easy to see that understanding behaviours at the individual scales as well as the correlations between scales is necessary for developing a more complete understanding of the fluidized bed drying process.

(iii) Multiway Principal Component Analysis (MPCA) is an extension of PCA that is often used to model transient behavior in batch data. These models typically contain information about the transitions involved from the start to the end of a process. Another objective of this study is to create an MPCA model of the fluidized bed drying process and determine whether stable states or key points in the process can be identified in the model parameters. Furthermore, if key states are apparent in the model it is desired to determine what factors or phenomena characterize them.

(iv) PCA has successfully been used in other chemical engineering applications for the multivariate statistical process control of complex processes. It is differentiated from other approaches by its ability to diagnose the probable causes of faulty operating conditions. The final objective of this study is to determine whether PCA can be used for the statistical process monitoring of the hydrodynamic conditions in a fluidized bed dryer.

The current study was confined to the analysis of previously existing data generated by Mr. Gareth Chaplin at Merck-Frosst Laboratories at Kirkland, Quebec in 2002 (Chaplin, Pugsley, and Winters, 2004a-b). Chaplin's investigation consisted

of two distinct components. The first component, referred to as the dry bed study, examined the steady state behaviour of dry pharmaceutical granulate in a clinical-scale fluidized bed dryer. The other component, referred to as the batch drying study, involved the batch drying of pharmaceutical granulate in the same fluidized bed dryer. In both cases, hydrodynamics were quantified with high frequency pressure fluctuations. The manipulated variables in the dry bed study were superficial velocity, pressure sensor position, and granulate particle size distribution. For the batch drying study, the variables under investigation were inlet air temperature, initial bed mass, and pressure sensor position.

2. LITERATURE REVIEW

2.1 Principal Component Analysis

Principal Component Analysis (PCA) is a linear projection modeling technique first developed by Pearson (1901). PCA projects data onto the linear hyperplanes or subspaces that best describe the variance in a data set. This may also be viewed as a coordinate system transformation. It is often found that the dimensionality of the subspaces required to describe the data is far less than the original number of variables (Jackson, 1980). Figure 2.1 illustrates this concept with the simple example of a two-dimensional subspace defined within a three-dimensional coordinate system. This data reduction quality of PCA makes it very attractive for problems with a large number of variables, redundant measurements, or strong correlations between many process variables.

The correlation and redundancy in process data is often due to changes in many variables at the same time due to a common underlying physical event. These physical events manifest themselves as a systematic shift of the data points in multivariate space. The coordinate axes of the new subspaces formed in PCA are defined by the direction of the largest shifts or maximum variances in multivariate space. Thus, the subspaces reflect the most significant events or states in a process.

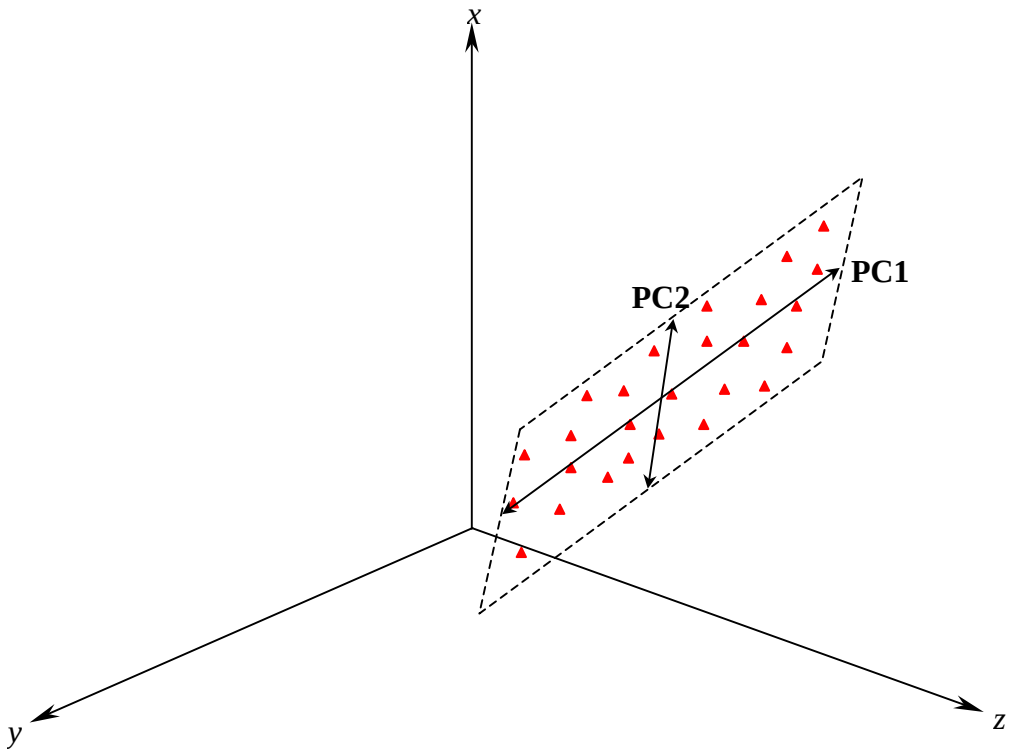

Figure 2.1: Illustration of a two-dimensional subspace defined by principal component 1 (PC1) and principal component 2 (PC2) within a three-dimensional coordinate system defined by x, y, and z

Linear combinations of the measured variables make up new pseudo- or latent variables describing the directions of the subspaces. In PCA, these latent variables are referred to as principal components. After the major events or states have been accounted for by the dominant principal components, the remaining components are interpreted as being the result of random errors or noise in the data. These components are grouped together in an error term. As such, PCA acts as a filtering technique.

Chemical engineering applications involving PCA have been well documented in the literature (Kresta, MacGregor, and Marlin, 1991; Kresta, Marlin, and MacGregor, 1994; Kourti and MacGregor, 1996; Teppola et al., 1998; Wise et

al., 1999). This use of PCA for process analysis has, to a certain extent, been a response to large volumes of process data resulting from technological advances in instrumentation and computers. However, the method continues to gain popularity with engineers and scientists as a tool for analyzing complex multivariate systems as new applications appear in the literature.

The formulation of a PCA model is relatively simple. Consider a data set, $\mathbf{X}_{mxn}$, consisting of *m* observations on *n* measured variables. These observations may be the conditions or results of a particular experiment or alternatively, time series measurements on several process variables in a chemical process. To begin, the data is typically 'autoscaled' or preprocessed by mean centering the *n* variables in the columns and dividing by the standard deviation as:

$$\mathbf{X} = \frac{\mathbf{x}_i - \mu_i}{\sigma_i} \tag{2.1}$$

This is done to prevent variables from appearing dominant in the analysis due solely to their relative magnitude with respect to the other variables. Once the data has been autoscaled, the *n* x *n* data correlation matrix, **C**, can be calculated as:

$$\mathbf{C} = \frac{\mathbf{X}^T\mathbf{X}}{m-1} \tag{2.2}$$

The data correlation matrix gives a measure of the correlation or relationship of the data set elements to one another. This matrix is symmetric, square, dense, and populated exclusively with real values. The combination of these attributes is important for defining the subspaces in PCA. Mathematically, these subspaces are defined by the eigenvectors and eigenvalues of the correlation matrix as:

$$\mathbf{CP} = \lambda_i\mathbf{P} \tag{2.3}$$

Eigenvectors of real symmetric matrices are orthonormal. This allows the eigenvectors of the correlation matrix to be used as the basis set (ie. coordinate system) for the new subspaces. Furthermore, the eigenvalues of a real symmetric matrix will also be real. The magnitude of each eigenvalue describes the variance in the direction of its corresponding eigenvector. The largest eigenvector describing the largest variation in the data is referred to as principal component 1 (PC1), the next largest principal component 2 (PC2), and so on.

Equation 2.3 represents a standard eigenvalue problem. To solve this eigenvalue problem, several techniques can be used. The complete solution set of eigenvectors and eigenvalues is desired in PCA, therefore transformation techniques such as the Jacobi diagonalization method or QR algorithm are suitable. In particular, a variant of the QR algorithm called the singular value decomposition (SVD) is convenient because it returns the eigenvalues in order of decreasing magnitude as the elements of a diagonal matrix and the corresponding eigenvectors as columns in a matrix. Solution of the eigenvalue problem in this study was performed using the SVD function in Matlab®. Results of the Matlab® function were further validated with an algorithm written in C++.

The general form of the PCA model relates the subspaces to the original scaled data set:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + ... + \mathbf{t}_k\mathbf{p}_k^T + \mathbf{t}_n\mathbf{p}_n^T \qquad (2.4)$$

Here $\mathbf{t}_i$ are known as score vectors and $\mathbf{p}_i$ are the loading vectors (ie. eigenvectors or principal components) of the correlation matrix. The scores describe where the data points project in the PCA subspaces, while the loading vectors describe the linear

combinations of the original variables that constitute each principal component. Once the eigenvalues, λ_i, and the eigenvectors, $\mathbf{p}_i$, have been calculated the scores, $\mathbf{t}_i$, can be easily determined from:

$$\mathbf{X}\mathbf{p}_i = \mathbf{t}_i \qquad (2.5)$$

The geometric interpretations of the loadings and scores are depicted in Figures 2.2 and 2.3, respectively. Each component of an eigenvector is determined by the cosine of the angle from each original coordinate axis. Therefore, each eigenvector will have *n* components. The overall contribution or weighting of a variable to a loading vector increases as the included angle approaches zero. The scores vectors simply define the projection of the sample points onto each eigenvector (loading vector) and thus describe the location of each observation in the PCA subspaces.

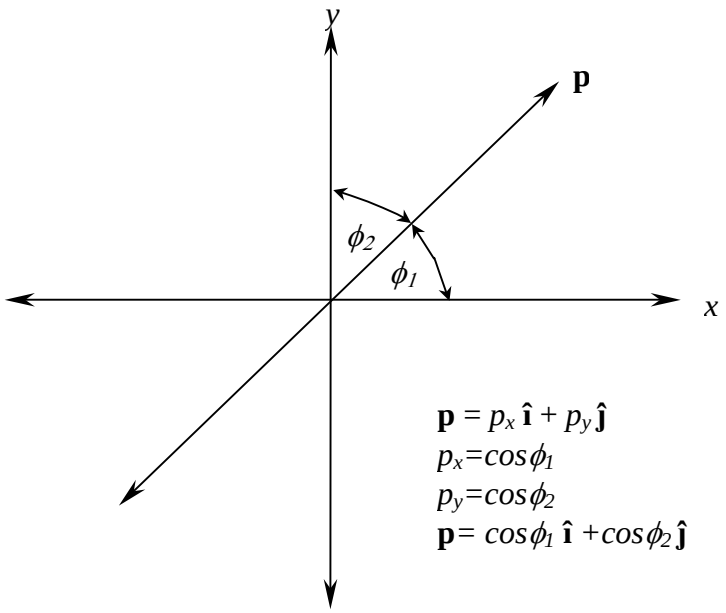

Figure 2.2: Geometric interpretation of an eigenvector, p (Geladi and Kowalski, 1986)

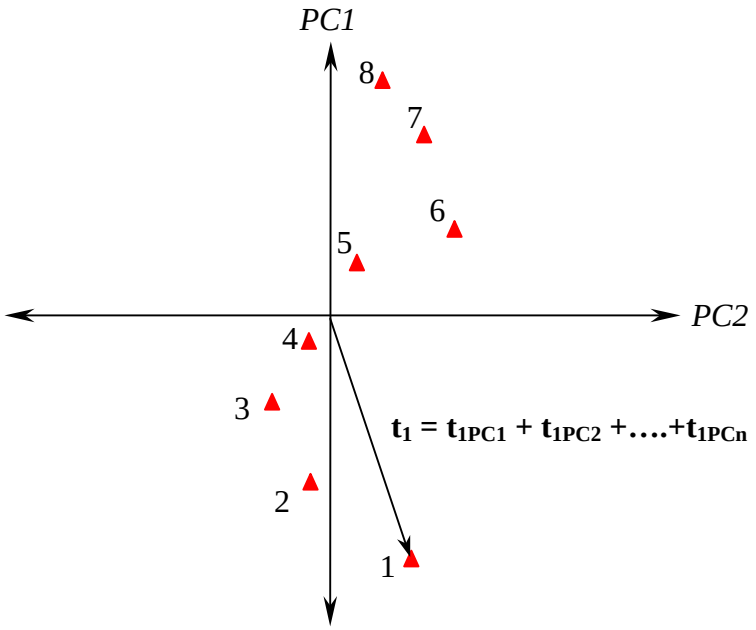

Figure 2.3: Geometric interpretation of a score vector, t_1, for data point 1 projected in the subspace defined by PC1 and PC2 (Geladi and Kowalski, 1986)

2.1.1 Data Selection

PCA models are data-driven and validity of the models is strongly dependent on the selection of data used. An underlying assumption made in PCA is that the data is multivariate normally distributed. This is a fair assumption for random processes. Furthermore, data projected onto the subspaces defined by the eigenvalues and eigenvectors will tend to be normally distributed due to the central value theorem which states that the sum of several normally distributed variables will tend to be normally distributed. This allows confidence limits to be applied on the data.

Another important consideration is that the data should adequately cover the range of values to be modeled. Extrapolation beyond these ranges is dangerous and will result in data points exceeding confidence intervals. Additionally, the data should be taken from the state of operation or interval of interest. This will ensure the

model contains the properties desired for comparison or evaluation. The model will only be as good as the data from which it is constructed.

2.1.2 Model Selection

In a PCA model, a decision is made on the number of components to be retained as significant contributors to the model. The remaining components are considered to be due to noise and errors and are grouped together as an error term, **E**. The reduced form of a PCA model containing *k* retained principal components can be expressed as:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + ... + \mathbf{t}_k\mathbf{p}^T{}_k + \mathbf{E} \qquad (2.6)$$

There is no standard convention for determining the number of components to retain in a PCA model. The choice is generally made based on a few rules of thumb and some knowledge of the data. A good first step is to calculate the eigenvalues and construct a table similar to Table 2.1.

Table 2.1: Sample eigenvalue distribution table

Principal Component Number	Eigenvalue	Variance Described (%)	Cumulative Variance (%)
1	3.63	36.3	36.3
2	2.33	23.3	56.9
3	1.72	17.2	76.8
4	1.66	16.6	93.4
5	0.21	2.1	95.5
6	0.17	1.7	97.2
7	0.13	1.3	98.5
8	0.08	0.8	99.3
9	0.04	0.4	99.7
10	0.03	0.3	100.0

If the eigenvalues have been calculated from the correlation matrix, they will sum up to the original number of variables. For example, a correlation matrix formed from a data set containing 10 variables will have 10 eigenvalues whose sum will total 10. The magnitudes of the eigenvalues themselves give an indication of the number of original variables each eigenvalue is "worth". Eigenvalues with magnitudes less than unity are probably not describing any systematic variation in the data set and are generally added to the error term.

The next feature to look for in the tabulated data is a sharp drop in values. This can be an indication of where terms describing errors begin to appear. Another way to visualize this is to plot the eigenvalues versus the principal components as in Figure 2.4. Here you are looking for the "knee" or point in the data where the slope abruptly tapers off. In Figure 2.4, this is seen after the fourth eigenvalue. If one were combining this approach along with the criteria for eigenvalues to be a minimum of unity, four principal components would be chosen (Table 2.1).

A less attractive approach sometimes used is to select some predetermined variance which the model must account for. If a model describing 95% of the variance in the data were desired, 8 components would be retained from the example in Table 2.1. This method does not make efficient use of the data reduction qualities of PCA. For instance, consider a data set with 20 measured variables. Let's say the eigenvector decomposition shows the first 2 principal components describe 91% of the variance, the next 8 principal components describe 4 %, and the remaining variance is described by the last 10 components. With a 95% criterion for the model

variance, 10 components would be retained where 2 components clearly describe the systematic variance in the model.

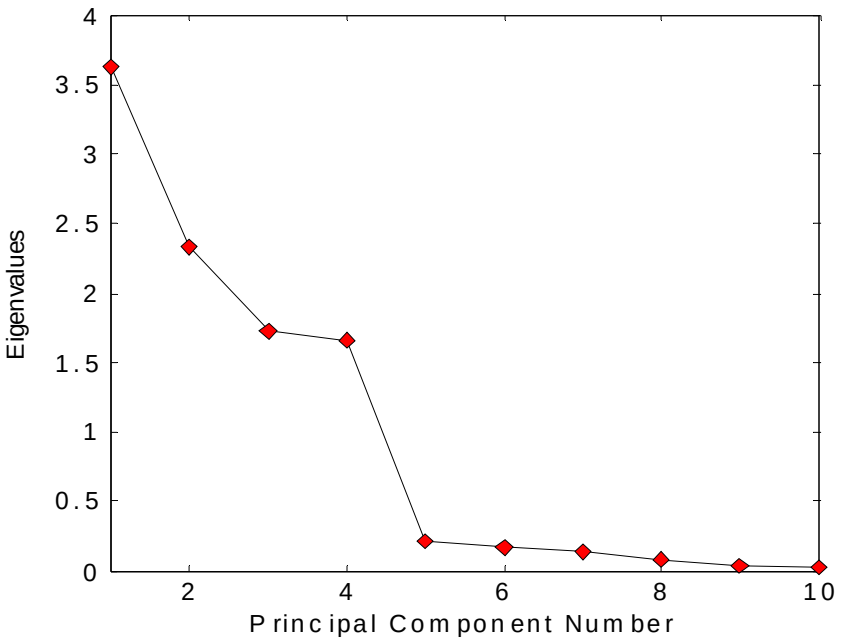

Figure 2.4: Sample eigenvalue curve for a PCA model

Cross-validation techniques provide another means to logically determine the number of principal components to retain. Several cross validation algorithms can be found in the literature, however some are dependent on the algorithm used to solve the eigenvalue problem. Wold's algorithm (1978) required the eigenvectors to be calculated using a deflation algorithm. The cross-validation method presented by Eastment and Krazanowski (1982) allowed the SVD to be used for solution of the eigenvalue problem. Regardless of the algorithm used, the basic concepts are the same. The data is divided into an equal number of segments and a PCA model is built on all but one segment. This model is then used to estimate the components left out.

The error of reconstruction, usually expressed as the cumulative predicted error sum of squares (Cumulative PRESS) or root mean squared error of cross validation (RMSECV), is then plotted as a function of the principal components as in Figure 2.5. As components describing systematic variance are added to the model, the error will decrease. Principal components describing noise variance will begin to increase the error. In Figure 2.5, four principal components describe the systematic variance.

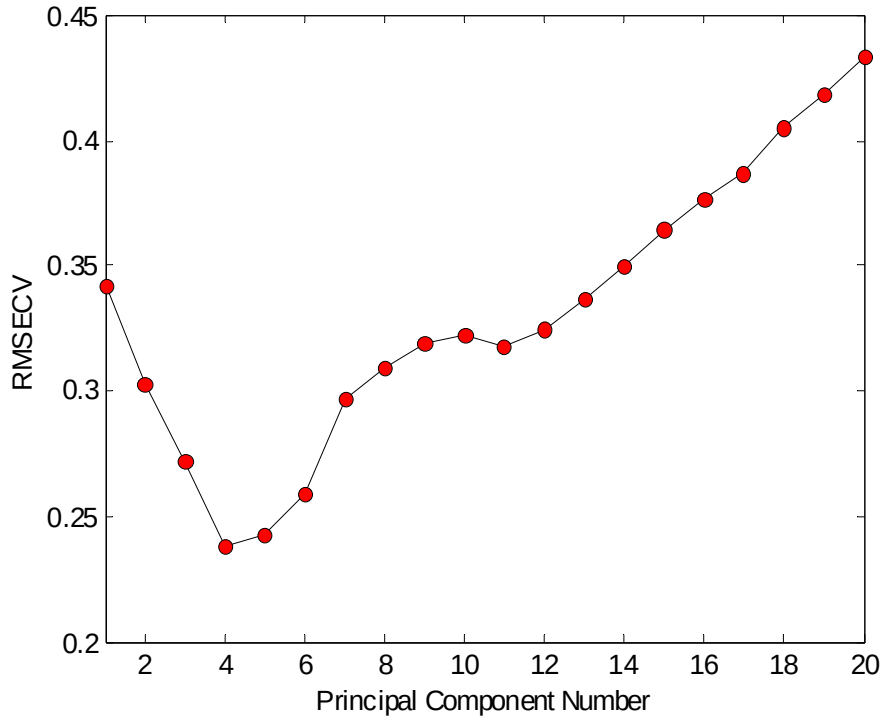

Figure 2.5: Sample cross validation results showing cumulative RMSECV vs. principal component

Even with these methods and rules of thumb, the decision of whether to retain or discard a component can be difficult. Looking at the correlations in each loading vector is sometimes helpful when making this decision. Often a physical significance or relationship can be seen in these correlations. If some physical relationship is

apparent in a loading vector, it should be kept in the model. The loading vectors consisting of random noise will be random.

Another point that should be made is that like sized loading vectors (ie. those with eigenvalues of similar magnitude) should either all be retained or discarded together. Principal components are directionally ambiguous. If the variances in two components are nearly equal but extend in different directions, then each component provides an equally valid representation of the data. Throwing one of them out would be an arbitrary act. For the example in Table 2.1, principal components 3 and 4 should be retained or discarded together.

The final decision in choosing the number of components is ultimately up to the modeler and their knowledge of the data. Common sense should always be exercised. Estimating the error contained in a data set will help establish some confidence in the model selected (ie. Is the random error likely to be 50%, 25%, or 10%?). Different models can be constructed and tested until a model appropriate for the analysis is found.

2.1.3 Data Analysis

Once a model has been formulated using the approaches discussed, the data can be analyzed to determine underlying physical relationships in the system and dominant states of operation. This is usually accomplished using plots known as scores plots and loadings plots (Geladi and Kowalski, 1986; Wold, Esbensen, and Geladi, 1987). Scores plots provide information on how the samples contained in the *n* original observations relate to one another. Loadings plots show how the original

variables combine to form each principal component. This provides some detail on the relationships of the original variables relative to one another.

2.1.3.1 Scores Plots

Score plots show the projection of the sample points onto the principal components. The magnitude and direction of each score vector is given by $\mathbf{t}_i$. Scores plots show how the samples relate to one another. Samples which are in close proximity to one another have similar characteristics. That is, they come from similar modes of operation in the process or experiment. A scores plot is shown in Figure 2.6.

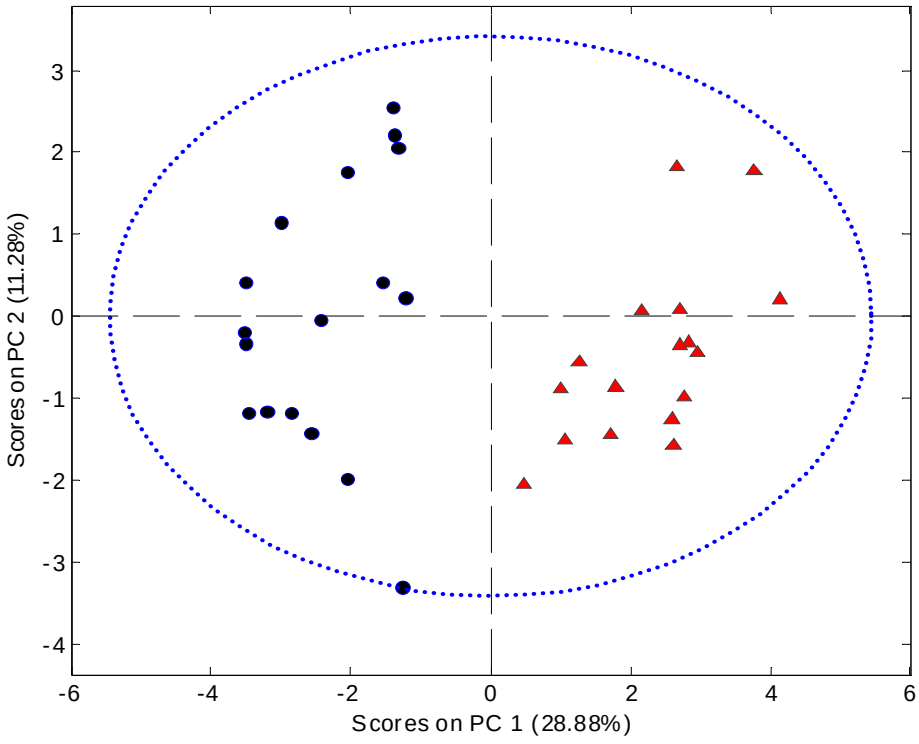

Figure 2.6: Sample two-dimensional scores plot with 95% confidence limits from PCA model containing data from 2 modes of operation (circles and triangles)

Scores often show up in clusters that are seen to represent these different modes or states in the process. In some cases, separate models are formed from the samples comprising each of these clusters. If the state of operation depicted by a cluster is a known fault that occurs in the plant, modeling this cluster and comparing new signals to it provide a means of detecting and diagnosing the problem quickly. To aid in this task, confidence limits for a model can be formed by applying the Student t-statistic as:

$$t_i = \text{sqrt}(\lambda_i) * t_{m-1,\alpha/2} \quad (2.7)$$

The validity of this statistic is dependent on the assumption that the data is multivariate normal. New data are projected on existing models as:

$$\mathbf{X}_{NEW}\mathbf{p}_i = \mathbf{t}_{i,NEW} \quad (2.8)$$

The location of a sample or observation in a scores plot is dependent on its elements and the composition of the principal components. For example, if fictitious variables A and B are highly weighted in a principal component, then the values of variables A and B in the observation will have the most effect on where that sample is projected in the subspace. The significance of this is that if a principal component can be shown to represent a particular physical phenomenon such as conversion rate in a chemical reaction, then negative scores may represent samples with low conversions and positive scores may represent the samples with high conversions.

2.1.3.2 Loadings Plots

Loadings plots, also referred to as load plots, give a measure of how the variables relate to one another. Variables that are loaded on a principal component in the same direction are positively correlated. Variables loaded in different directions

are negatively correlated. The relative magnitudes of the variables will give an indication of the strength of this correlation in the context of the principal component. This is a powerful tool and the relationships of the variables loadings are often seen to show relationships in the physical phenomena driving the process. An example of a loadings plot is shown in Figure 2.7.

While loadings plots provide an excellent means for examining the relationships between variables, caution is advised in the interpretation. Loads plots show correlations and do not imply causality (ie. increasing A does not necessarily cause an increase in B, although they may be positively correlated). Causality can only be determined through designed experiments.

2.1.4 Classification and Statistical Control

The successful implementation of PCA for multivariate statistical process control applications (MSPC) and classification has been extensively reported in the literature. Kresta, MacGregor, and Marlin (1991) cover this topic in detail. PCA can be used for statistical control applications by applying two different statistics, Hotelling's T^2 statistic and the Q-statistic also known as the Q_{residual} statistic or SPE statistic (Jackson and Mudholkar, 1979; Westerhuis, Gurden, and Smilde, 2000a-b). These statistics can be used to construct multivariate control charts with confidence limits for the entire PCA model.

Scores plots generally cannot be used for this purpose because each scores plot only represents one or two specific dimensions (principal components) of a PCA model. Therefore, a point which projects inside the confidence limits on a

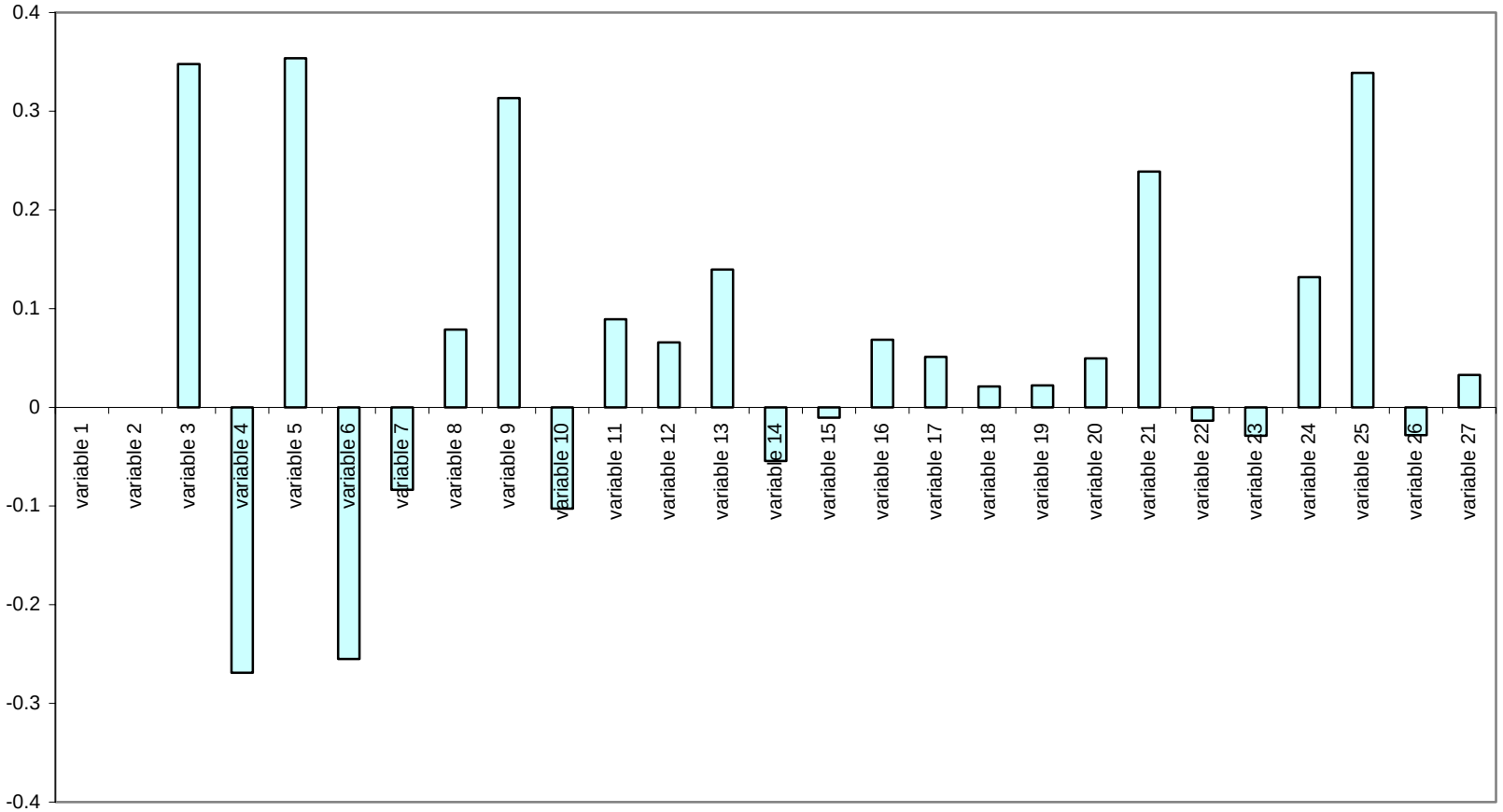

Figure 2.7: Sample loadings plot showing principal component 1 from a PCA model

scores plot may not conform to the model within the constraints of its remaining dimensions. Scores plots confidence limits should only be used for determining whether a sample point belongs in the subspace defined by a particular principal component. Such is the case where a cluster of points exists in a particular scores plot.

The application of multivariate models for statistical process control is inherently superior to a univariate approach as most processes are multivariate in nature. The ability of PCA to identify and capture process operating states through the measured variables and their interactions makes it a very powerful approach to statistical control. Additionally, this multivariate approach provides a means for assigning probable causes to any detected faults. This allows process operators to correct problems and bring the process back into a desired operating region.

Hotelling's T^2 and the Q-residual statistics are not limited solely to online statistical control applications. In fact, they can also be used in offline analyses for classification and comparison of data. As in the statistical control cases, this approach has advantages over traditional approaches because it considers the correlation structure of the data.

2.1.4.1 Hotelling's T^2 Statistic

Hotelling's T^2 statistic, equation (2.9), provides a measure of variation within a PCA model. This variation can be thought of as being in the same plane as the subspace. Control charts can be formed by plotting T^2 for each sample. The 95% confidence limits representing the upper "in-control" or normal operating limit is

given by equation (2.10). Large variations in T^2 indicate large variations in the model. Unusual variations or "out-of-control" samples in Hotelling's T^2 plots (Figure 2.8) may indicate that the process is operating outside the ranges of the data used to construct the model (ie. beginning to extrapolate outside the limits of the model).

$$T_i^2 = \mathbf{t}_i \lambda^{-1} \mathbf{t}_i^T = \mathbf{x}_i \mathbf{P} \lambda^{-1} \mathbf{P}^T \mathbf{x}_i^T \tag{2.9}$$

$$T_{k,m,\alpha}^2 = \frac{k(m-1)}{m-k} F_{k,m-k,\alpha} \tag{2.10}$$

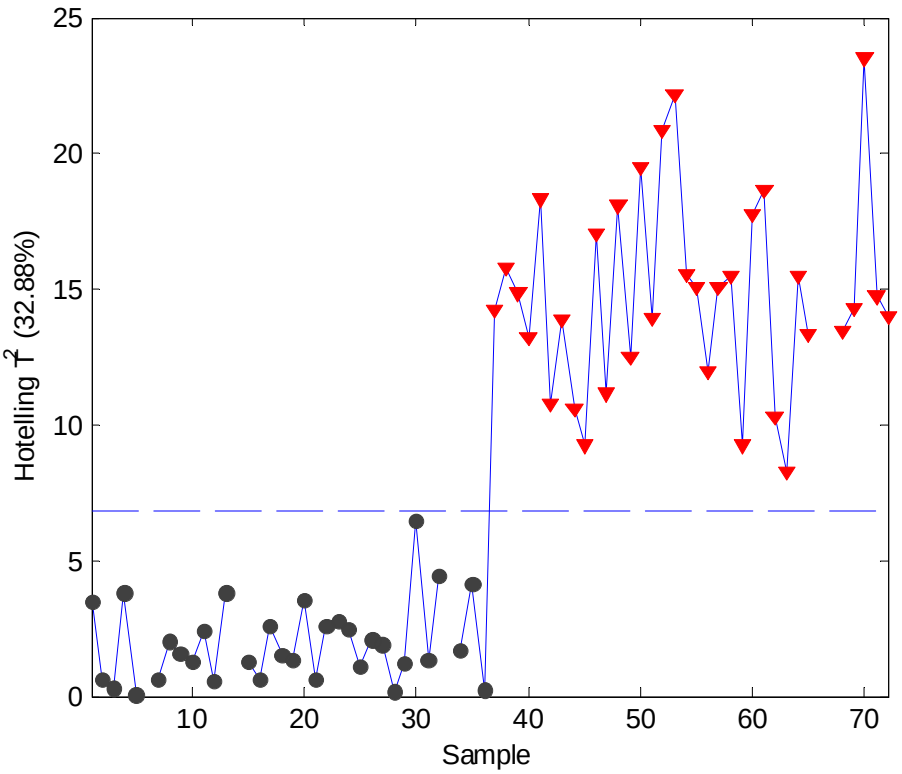

Figure 2.8: Hotelling's T^2 plot showing new data (triangles) exceeding the 95% confidence limits when projected against an existing PCA model (circles)

2.1.4.2 $Q_{residual}$ Statistic

The $Q_{residual}$ statistic, also known as the *Q*-statistic, is calculated from the error term, **E**, in the reduced PCA model, equation (2.6), and is simply the sum of squares of its elements as given in equation (2.11). The *Q*-statistic provides a measure of distance of a sample point outside the PCA model. Mathematically, it gives a measure of the (Euclidian) distance of the point off the hyperplane or subspace formed by the principal components. Confidence limits for the *Q*-statistic can be calculated with equations (2.12) to (2.14).

As with Hotelling's T^2 statistic, control charts can be formed by plotting Q_i versus sample number along with the 95% confidence limits. Here the 95% confidence limits give the control limits for "in-control" operation of the process. This is a very effective multivariate statistical process control technique. In practice, the *Q*-statistic is found to be a much more powerful tool for detecting faults than its counterpart, T^2. This is because the *Q*-statistic can identify conditions where the correlation structure has changed before faults have manifested themselves in the individual variables by causing them to deviate outside of their normal operating ranges.

$$Q_i = \mathbf{e}_i \mathbf{e}_i^T = \mathbf{x}_i (\mathbf{I} - \mathbf{P}_k \mathbf{P}_k^T) \mathbf{x}_i^T \tag{2.11}$$

$$Q_\alpha = \theta_1 \left[\frac{z_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \tag{2.12}$$

$$\theta_i = \sum_{j=k+1}^{n} \lambda_j^i \quad for\ i = 1, 2, 3 \tag{2.13}$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \qquad (2.14)$$

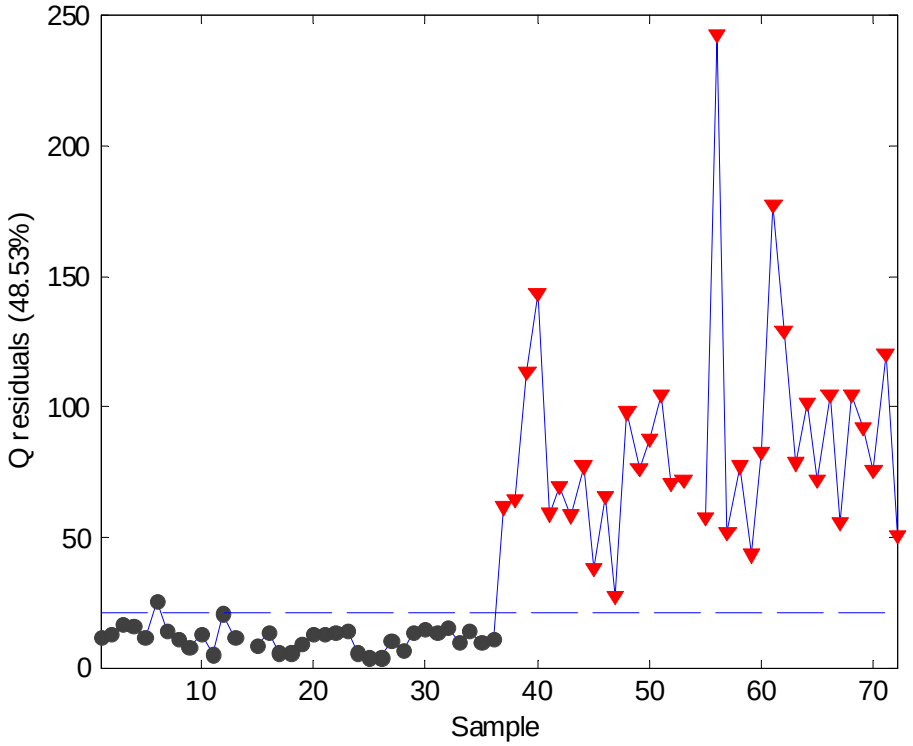

Figure 2.9: Sample $Q_{residual}$ plot showing new data (triangles) exceeding the 95% confidence limits when projected against an existing PCA model (circles)

2.1.4.3 Contribution Plots

Once "out-of-control" or faulty operating conditions have been identified, faults must be identified and corrected to bring the process back to a more desirable state. Contribution plots, as described in the paper by Westerhuis et al (2000a), provide a means for identifying the variable or groups of variables responsible for out-of-control signals. Contribution plots can be formed for both Hotelling's T^2 statistic and the Q-statistic.

The contribution statistic for T^2 is denoted as T_{con} and is given by equation (2.15). A T_{con} plot can be constructed for each sample which exceeds the 95% confidence limits of the T^2 statistic (Figure 2.10). This plot gives the relative contribution of each of the variables in producing the T^2 value. Variables that are greater contributors (positively or negatively) are more likely to have caused the event or are more likely to be a direct result of the fault. This gives engineers and operators a reference point in the process to begin looking for problems.

$$\mathbf{t}_{con} = \mathbf{t}\sqrt{\lambda^{-1}}\,\mathbf{P}^T \qquad (2.15)$$

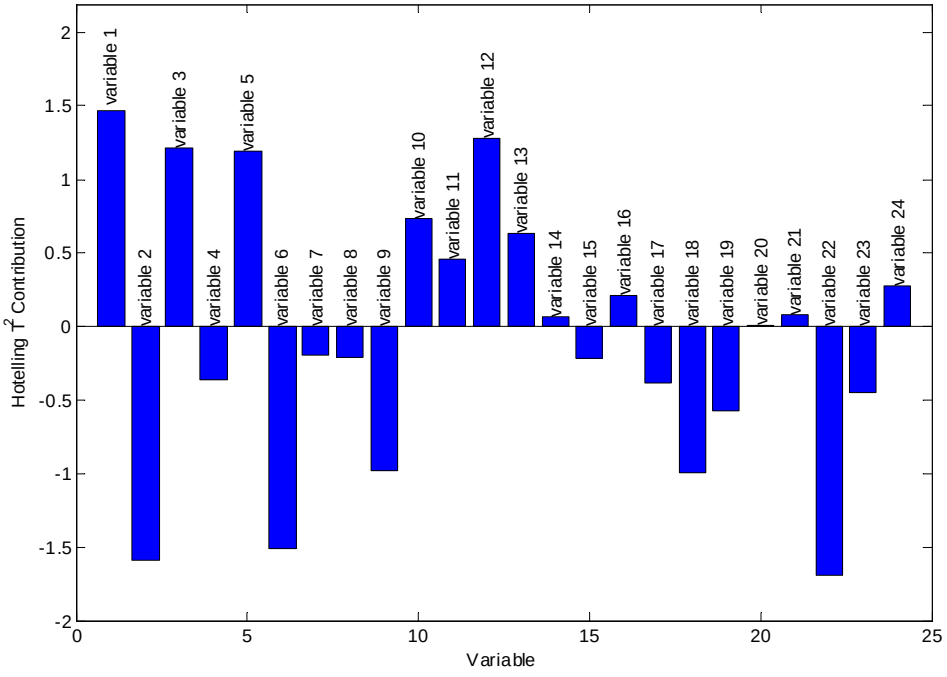

Figure 2.10: Sample t_{con} plot for a sample exceeding the 95% confidence limits

The contribution plots for the *Q*-statistic are derived directly from the error matrix in equation (2.6). The rows of this matrix give the contributions of each variable to the distance a sample lies outside the model. Therefore, the more a

variable or group of variables contribute to these error terms, the larger the probability that the variable or variables are related to the fault. Figure 2.11 shows a *Q*-statistic contribution plot for an out-of-control sample.

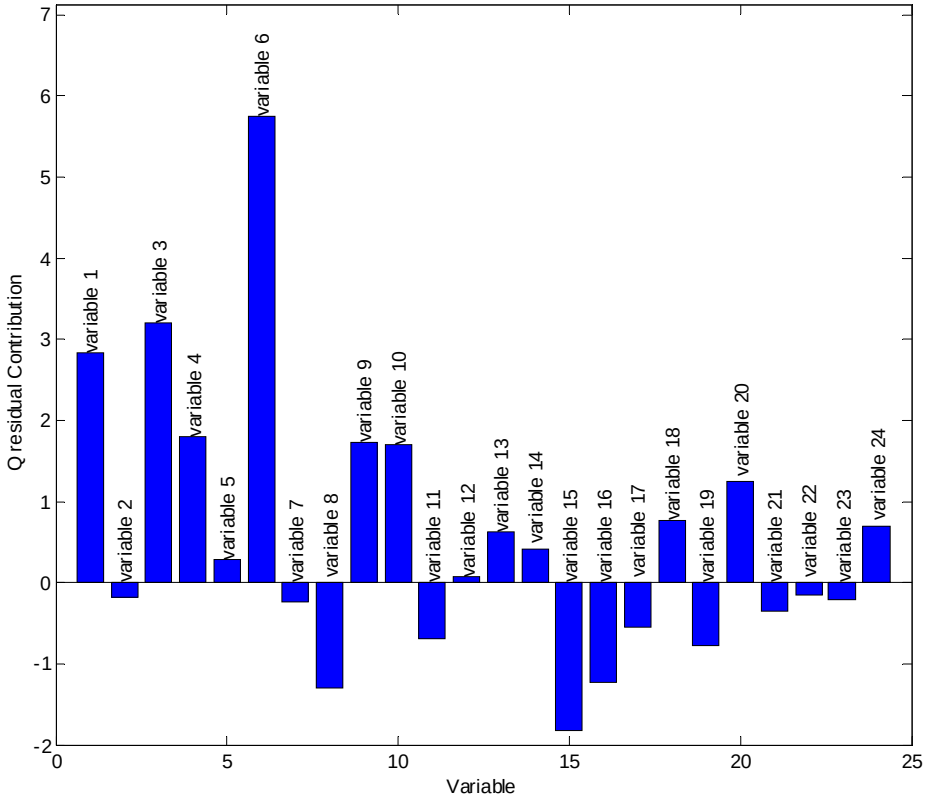

Figure 2.11: Sample *Q*-statistic contribution plot for a sample exceeding the 95% confidence limits

2.2 Multiway Principal Component Analysis

Multiway Principal Component Analysis (MPCA) is an extension of PCA that was first proposed by Wold et al. (1987). It was developed for the purpose of being able to apply PCA to data sets consisting of three-dimensional arrays, as is typically the case in applications such as image analysis. Subsequently, Nomokis and MacGregor (1994) adapted the technique to analysis and monitoring of batch process data. Since this adaptation, MPCA has been successfully applied in a number of

applications involving batch processes (Nomikos and MacGregor, 1995a; Chen and McAvoy, 1998; Wise et al., 1999; Westerhuis, Kourti, and MacGregor, 1999).

MPCA is algorithmically consistent with PCA. The key difference is a preprocessing step known as "unfolding" required to convert a three-dimensional data matrix to a large two-dimensional matrix. Unfolding simply involves a rearrangement of the data. Once a dataset has been rearranged, a PCA model is formulated in the usual fashion. Loadings plots, scores plots, Hotelling's T^2 statistic, the Q-statistic, and contribution plots can be calculated for MPCA models in the same manner as for regular PCA models.

2.2.1 MPCA for Batch Processes

Batch processes are typically dynamic in nature. They start with a set of measured variables at some initial condition and each variable changes as the process proceeds to a desired final state. The trajectory these individual process variables follow is typically predictable. However, as in PCA, more information about the state of the process can be exploited by examining the auto-correlation and cross-correlation of the measured variables than can be obtained by examining the individual trajectories separately. Furthermore, when conditions differ slightly between each batch it is desirable to account for batch-to-batch variability in the models as well. This is the goal of MPCA.

MPCA of batch process data begins with a three-dimensional data matrix. In any batch process, a single data set, $\mathbf{X}_{mxn}$, can be formulated by taking measurements on n variables at m different time intervals. For illustrative purposes, let us suppose that in a particular batch experiment, measurements are made every minute for 20

minutes on 4 variables: temperature, pressure, flow, and mass. The resulting dataset would have the dimensions 20 x 4. A PCA model could be calculated directly from this data but it would lack information regarding batch-to-batch covariance. In an attempt to quantify this variability, let us supposed the experiment was repeated six times. These 6 data sets, $\mathbf{X}_{20x4}$, can be grouped together as a single three-dimensional data array, $\underline{\mathbf{X}}_{mxnxo}$, with dimensions m x n x o. In the case of our example, $\underline{\mathbf{X}}$ would have dimensions of 20 x 4 x 6.

The first step in the MPCA algorithm is to unfold this three-dimensional matrix. It can be seen that there are 3 unique ways to unfold this matrix: time-wise, batch-wise, or variable-wise. The decision on how to unfold the data depends on the particular objective of the model or the analysis. In practice, the time-wise and batch-wise options are most common.

Time-wise unfolding simply involves placing each dataset from the individual batches adjacent to one another to create a large two-dimensional dataset (Figure 2.12). This new dataset will have dimensions m x (n x o) = m x no. In the case of our example the new dataset would have dimensions 20 x (6 x 4) = 20 x 24. A second time-wise unfolding option exists with dimensions m x on, but it is equivalent to the first. The newly unfolded dataset, $\mathbf{X}$, can be analyzed using the PCA algorithm, equation (2.1) – (2.6).

There are several implications that arise from time-wise unfolding. Firstly, the resulting dataset has m observations. Therefore, PCA models based on this dataset will have m points in each scores plot. Similarly, there will be n x o = no variables in each loadings plot (ie. each variable repeated over each batch is

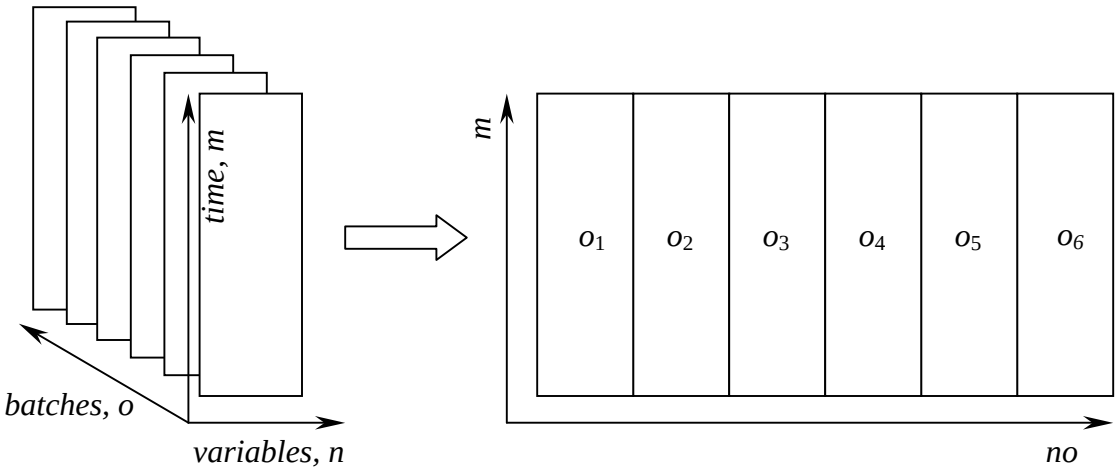

Figure 2.12: Time-wise unfolding of a three-dimensional dataset containing information from 6 experiments of a batch process (Kourti, 2003)

considered a separate variable). Our earlier example would have 20 scores projected on latent variables with a dimension of 24. The time-wise unfolding of batch datasets provides an analysis of the dynamic process behavior about the midpoint in time of the experiments. This may also be considered an examination of the multivariate trajectories with time.

The batch-wise unfolding of the data matrix follows the same principles. In this case, the unfolding results in an o x (m x n) = o x mn dataset. Once again this is equivalent to the o x nm unfolding. For our example, the dimensions of the new two-dimensional dataset would be 6 x 80. PCA models derived from batch-wise unfolding examine the dynamic behavior of the process about the overall mean between batches. An alternative way to think about this is that the models are a measure of the batch-to-batch variances. Thus, the time trajectory is effectively removed. For this reason, batch-wise unfolding is usually selected when the MPCA is to be used for statistical process control.

The last option for unfolding a three-dimensional dataset is the variable-wise approach that results in a *n* x *mo* two-dimensional data matrix. The resultant matrix in our example would be 4 x 120 in size. Analysis of these datasets would provide information about the dynamic behavior of the process with respect to the overall mean for each variable. While there may be specific instances where this type of analysis is desired, the first 2 unfolding alternatives presented are more commonly presented in the literature.

2.2.1.1 Trajectory Alignment

The main challenge in working with batch data arises when batches have variable time durations. This occurs due to differences in process conditions (ie. differences in temperatures, differences between initial bed masses, impurities in reactants, etc.). In some instances, these differences also cause the shape of the trajectories to differ. This complicates the issue. If the shape of the trajectories were the same but duration differed, the variables could be linearly scaled for the purpose of forming the model. However, when the shape of the trajectory differs, the covariance structure also differs and linear scaling cannot be applied.

The most common solution to this problem is to use a variable other than time to align the batches (Nomokis and MacGregor, 1994). If the data set has a variable which increases monotonically in time and has similar start and finish values for each batch, it can be used as an "indicator variable". The remaining variables would then be aligned according to the indicator variable rather than time. In practice, this approach works very well.

Another cause of misalignment of trajectories between datasets results from measurements on the variables occurring at different time periods. For instance, if the variables in one batch from our example were recorded at 45-second intervals we would have a dataset measuring 26 x 4. This would pose a problem when the data from this batch is combined with the 20 x 4 datasets from the remaining experiments. This must be rectified before the three-dimensional data matrix can be constructed. However, this issue can be easily resolved by interpolating the variables in each dataset to correspond to a common time increment (ie. every minute in our example). The choice of interpolation algorithm depends on the nature of the individual variable trajectories and the desired precision.

2.2.1.2 MPCA Statistical Process Control

A problem arises when MPCA is to be used online for multivariate statistical process control. In most cases, a batch-wise model is formulated for these applications. In these instances, appropriate data for comparison to a model will not be available until completion of the batch. Consider our batch-wise unfolded example, over the first 19 minutes the dataset will be incomplete, that is it will have dimensions of 6 x $4m$, $1 \leq m \leq 19$. This poses a problem because the model was constructed with dimensions of 6 x 80. The simplest solution to this problem would be to have m individual models for comparison (ie. one for each point in time). This is not always practical.

Other approaches to deal with this have been suggested in the literature by Nomokis and MacGregor (1995a). They include filling in future observations with values from an in-control reference series or adding the variance of the current

measurement to all future values in a similar reference series. Regardless of the approach chosen, this is one aspect of MPCA statistical control that should not be overlooked. Simply filling in future values in the dataset with zeros will cause points to exceed 95% confidence limits.

2.3 Partial Least Squares (PLS)

Partial Least Squares (PLS), also known as Projection to Latent Structures, is a latent variable modeling method derived from similar principles to PCA. The main conceptual difference between the two methods is that PLS is a regression-based technique where a block of input variables, **X**, is used to predict a block of output variables, **Y**. PLS was first developed by Herman Wold in the late 1960's and evolved to its present form by 1982 (Wold, 1982). Since that time, PLS has been successfully applied in several applications for inferential modeling, process monitoring, and data analysis (Kresta, MacGregor, and Marlin, 1991; Kresta, Marlin, and MacGregor, 1994; Kourti and MacGregor, 1996).

There are two categories of PLS models known as PLS1 and PLS2. The primary goal of both types of model is to provide a linear regression between a m x n block of input data, **X**, and an m x p block of output or response data, **Y**, in the form:

$$\mathbf{Y} = \mathbf{XB} \tag{2.17}$$

In equation (2.17), **B** is a matrix of regression vectors relating each response variable in **Y** to the original **X** data. PLS1 is simplified version of PLS2 where **Y** consists of a single response variable, y. The current study does not consider the case where there are multiple responses. Therefore, general references to PLS should be interpreted as references to PLS1 unless otherwise explicitly stated. Regardless, the

differentiation between PLS1 and PLS2 is inconsequential in the context of this work because the same algorithm is used for PLS1 models as for PLS2 models.

PLS modeling combines the powerful approach of latent variable modeling with regression modeling. This combination results in several advantages over traditional approaches. One such advantage is that PLS can be used to perform regressions in the case where there are a far greater number of variables than available observations. The other main advantage lies in the representation of the data. Like its counterpart, PCA, PLS is very good at handling noisy, redundant, highly correlated data. PLS models present the data in terms of latent variables such that the **X** data and **Y** data can be presented as:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \tag{2.18}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} \tag{2.19}$$

Equations (2.18) and (2.19) are referred to as the outer-relationships of the PLS model and are seen to have the same structure as found in PCA models. However, PLS models are not a simple regression between the two PCA models. In practice, a regression of this sort does not provide the best predictive model. Instead, PLS models attempt to maximize the covariance of **X** with **Y**. In other words, PLS models determine the maximum variance in **X** that describes **Y**.

There are three widely used algorithms for calculating a PLS model. These include the NIPALS algorithm (Geladi and Kowalski, 1986; Höskuldsson, 1988), the kernel algorithm (Lindgren, Geladi, and Wold, 1993), and the SIMPLS algorithm (de Jong, 1993). The NIPALS algorithm is the most closely related of these algorithms to the original form of the model presented by Wold (1982). However, it suffers from a

computationally intensive iterative algorithm. The kernel algorithm and SIMPLS are comparable in computational requirements depending on the particular application (de Jong, 1993). It should be noted that the model produced by the SIMPLS algorithm is identical to the other algorithms for PLS1 models but slightly different for PLS2 models.

SIMPLS produces a different result for PLS2 models as a result of an additional orthogonality constraint introduced to ensure that the scores between the outer-relationships are independent. NIPALS and the kernel algorithm do not contain this requirement and, consequentially, produce a different solution in the case of PLS2 models. The SIMPLS algorithm will be used here to describe the calculation of a PLS model because, unlike the kernel algorithm, scores for the outer-relationships are calculated at each step in the regression. A detailed description of the NIPALS and the kernel algorithms are presented in Höskuldsson (1988).

SIMPLS is a stepwise algorithm that determines the orthogonal scores of **X** that have the maximum possible covariance with the corresponding scores of **Y** (ie. $\mathbf{u}^T\mathbf{t}$ = maximum). These scores are derived from the n x p autoscaled **XY** covariance matrix, **S**:

$$\mathbf{S} = \mathbf{X}^T\mathbf{Y} \tag{2.20}$$

After each set of score and loading vector are determined, the covariance matrix is "deflated" by removing the orthogonal components that define the *X*-block variance from **S**. This process is repeated m times until the full solution is defined. A correlation matrix, C, that provides a measure of the **XY** relationship is calculated as:

$$\mathbf{C} = \mathbf{S}^T\mathbf{S} \tag{2.21}$$

The dominant eigenvector, **q**, of the correlation matrix, **C**, determines the maximum covariance between the **X** data and the **Y** data. This eigenvector can be obtained as the dominant eigenvector from a singular value decomposition (SVD) of **C**. In order to relate the covariance matrix, **S**, to the **X**-block scores, PLS algorithms require an additional set of latent variables known as weights, **w**. These weights are calculated directly from the eigenvectors of the correlation matrix as:

$$\mathbf{w} = \mathbf{Sq} \tag{2.22}$$

Once the weights have been determined, the **X**-block scores can be calculated and normalized:

$$\mathbf{t} = \mathbf{Xw} \tag{2.23}$$

$$\mathbf{t} = \frac{\mathbf{t}}{\|\mathbf{t}\|} \tag{2.24}$$

The weights are normalized as well such that $\mathbf{w}^T\mathbf{w} = 1$:

$$\mathbf{w} = \frac{\mathbf{w}}{\|\mathbf{w}\|} \tag{2.25}$$

Subsequently, the **X**-block and **Y**-block loading vectors are calculated from the **X**-block scores:

$$\mathbf{p} = \mathbf{X}^T\mathbf{t} \tag{2.26}$$

$$\mathbf{q} = \mathbf{Y}^T\mathbf{t} \tag{2.27}$$

Scores for the **Y**-block data are then obtained simply as:

$$\mathbf{u} = \mathbf{Yq} \tag{2.28}$$

Another set of vectors, **v**, representing the orthogonal basis of the **X**-block loading vectors must be introduced into the calculation. For the initial step of the SIMPLS algorithm, this vector is taken to be equal to the **X**-block loading vector (ie. **v** = **p**). In

determination of subsequent latent variables, **v** must be computed to ensure it is orthogonal to all previously calculated **X**-block loading vectors:

$$\mathbf{v} = \mathbf{v} - \mathbf{V}(\mathbf{V}^T\mathbf{p}) \tag{2.29}$$

In a similar manner, the **Y**-block scores must be scaled so that they remain orthogonal to all previous **X**-block scores:

$$\mathbf{u} = \mathbf{u} - \mathbf{T}(\mathbf{T}^T\mathbf{u}) \tag{2.30}$$

To calculate the next set of scores and latent variables, the covariance matrix is deflated by removing the normalized loading vector basis variable, **v**:

$$\mathbf{v} = \frac{\mathbf{v}}{\|\mathbf{v}\|} \tag{2.31}$$

$$\mathbf{S} = \mathbf{S} - \mathbf{v}(\mathbf{v}^T\mathbf{S}) \tag{2.32}$$

After each step through the algorithm **w**, **t**, **p**, **q**, **u**, and **v** are stored as columns in **W**, **T**, **P**, **Q**, **U**, and **V**, respectively, before stepping through the algorithm again beginning at equation (2.22). Once the entire solution consisting of *m* latent variables for each of **X** and **Y** has been determined, a matrix of regression coefficients can be calculated directly as:

$$\mathbf{B} = \mathbf{W}\mathbf{Q}^T \tag{2.33}$$

This regression coefficient is the same coefficient as in equation (2.17) and represents the solution to the multivariate regression between **X** and **Y**. Predictions can be obtained directly from equation (2.17) but must consider the fact that the data has been autoscaled and scale the predictions appropriately by multiplying by σ and adding the mean value of *y*.

It should be pointed out that in a PLS model the loadings are referred to as latent variables (LV) rather than principal components (PC). The *X*-block latent

variables are calculated to determine the maximum variance in **X** that is predictive of **Y**. This differs from the structure of a PCA model where the sole goal is to describe the maximum variance or principal directions in a single block of data. As such, successive latent variables in the **X**-block will not necessarily be in descending order of variance in the **X**-block data but rather in descending order of the variance that is indicative of the **XY** covariance. This is illustrated in Table 2.2 for latent variables 3 and 4.

Table 2.2: Sample cumulative eigenvalue distributions for X-block and Y-block

Latent Variable	X-Block		Y-Block	
	This LV (%)	Cumulative Total (%)	This LV (%)	Cumulative Total (%)
1	32.48	32.48	94.50	94.50
2	10.49	42.97	2.60	97.10
3	5.44	48.40	0.79	97.89
4	7.18	55.58	0.20	98.09
5	6.24	61.82	0.18	98.26
6	3.95	65.77	0.19	98.45
7	3.90	69.67	0.13	98.58
8	3.61	73.28	0.09	98.67
9	5.12	78.40	0.08	98.76
10	2.70	81.10	0.13	98.88

In a similar manner to PCA modeling, once the PLS covariance matrix has been completely deflated, the modeler selects a number of components to retain in the model. The outer-relationships are then reduced to the form given by equations (2.18) and (2.19). The number of latent variables to be retained in a PLS model is usually determined using a cross-validation technique to minimize the prediction error and the modeler's discretion (Wold, 1978; Eastment and Krazanowski, 1982). The guidelines for retaining components in PCA modeling can be applied here as well. In

this investigation, the minimum RMSECV was used along with the **X**-block eigenvalues curves to determine the number of retained components. In Table 2.2, the choice to retain 2 latent variables in the model was optimal.

As in PCA modeling, scores plots can be used to visualize the data (Figure 2.13). Loadings plots can also be constructed to view the linear combinations of the original variables that make up the weight, loading, and regression vectors. The weight vectors present a normalized version of the loading vectors and the regression vectors show the variables and correlations that best relate the **X** data to the **Y** data. Hotelling's T^2 and Q-statistic plots can also be constructed, however they are not as useful for monitoring the overall system behavior as those derived from PCA models.

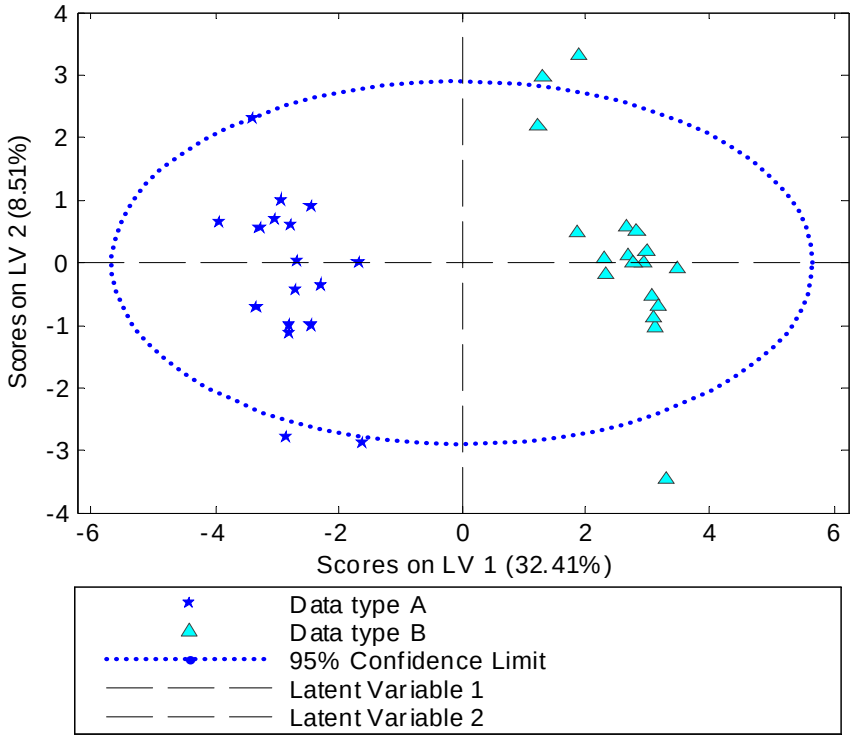

Figure 2.13: Sample scores plot for X-block data in a PLS model

2.3.1 PLS Confidence Intervals

In most PLS studies, the model accuracy is evaluated by examining the prediction residuals. However, the literature does contain attempts to derive statistical confidence intervals for PLS predictions (Nomikos and MacGregor, 1995b; Denham, 1997; Faber and Kowalski, 1997; Faber, 2000; Duchesne and MacGregor, 2001; Baffi, Martin, and Morris, 2002). Approaches include cross-validation techniques (jackknifing) and using the Jacobian matrix to predict error limits.

The PLS models in this study used cross validation to help determine the number of latent variables to retain in each model. Therefore, it would make sense to use this information to determine the confidence intervals of the predictions. The cross validation algorithm in this study divided the data into 10 random subsets. Models were calculated leaving each subset out once. The number of components that minimized the predicted root mean squared error (RMSECV) was considered to be optimal unless the contribution of the component was less than unity.

The cross-validation approach to determining confidence limits uses the errors calculated in the sub-models to determine a standard error (Figure 2.14). The Student *t*-distribution is then used to give confidence intervals. This method yields good estimates because PLS parameters are linear combinations of the original variables and are therefore approximately normally distributed. The confidence intervals of Nomikos and MacGregor (1995) were used here to estimate the standard error, $\hat{\xi}$:

$$\hat{\xi} = \hat{y} \pm t_{n-1,\alpha/2}(MSECV)^{\frac{1}{2}}(1+\mathbf{t}(\mathbf{T}^T\mathbf{T})^{-1}\mathbf{t}^T)^{\frac{1}{2}} \qquad (2.34)$$

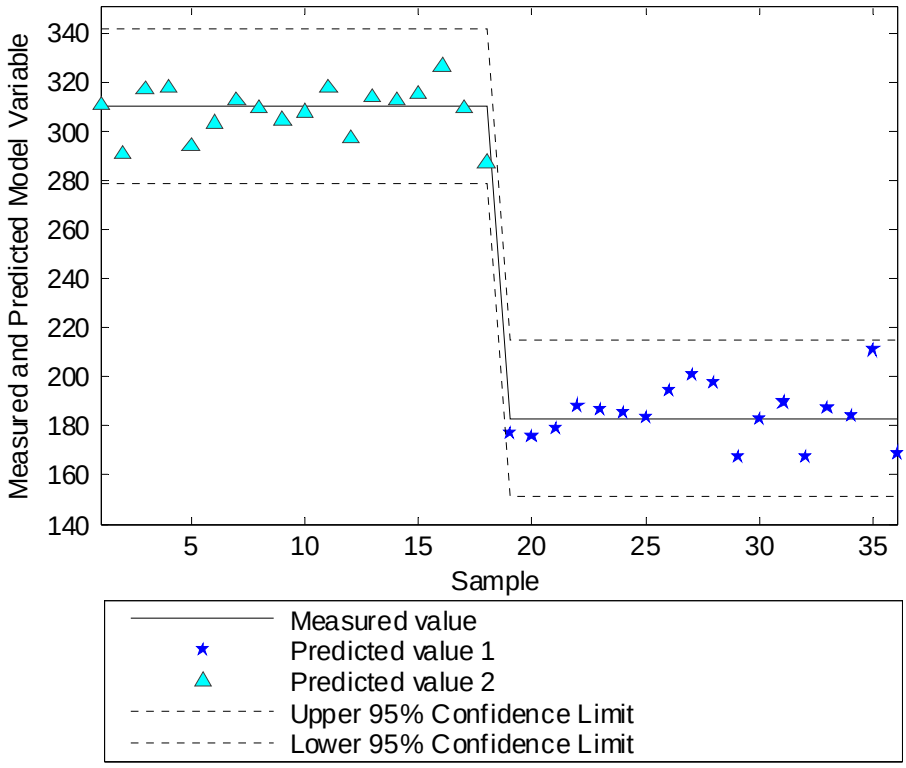

Figure 2.14: Sample PLS prediction plot with confidence limits

2.4 Wavelet Transformations

Wavelet transforms are a relatively new mathematical tool that have found a wide range of applications in signal processing, image analysis, and numerical modeling. Although based on a concept originally presented by Alfred Haar in 1910, wavelets in their present form have only been in existence for the past 20 years. The application of wavelets to problems outside the field of pure mathematics is an active area of research.

Wavelet analysis involves using a family of basis functions (wavelets) to represent features of a signal or function. Wavelet functions are localized in both time and frequency. This fact coupled with a dyadic scaling convention allows

discrete wavelet transformations to present the low-frequency components of a signal over long time intervals and higher frequency components at a higher resolution over short time intervals (Figure 2.15). This contrasts Fourier analysis where a signal is represented solely in the frequency domain. A variation of the Fourier Transform, known as the Short-Time Fourier Transform (STFT) or windowed Fourier Transform, was developed by Gabor in 1946 in order to retain time domain information. However, due to the use of sinusoids as basis functions there is usually a trade off in resolution between the time and frequency domains (Alsberg et al., 1997). The mathematical features of wavelets overcome this deficiency and can provide a superior representation of most signals.

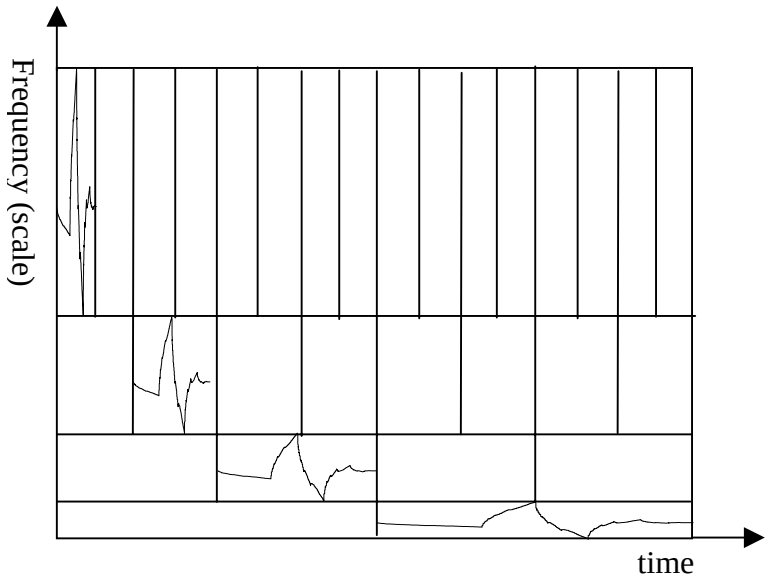

Figure 2.15: Scaling and translation of a wavelet over the discrete time-frequency domain

2.4.1 Discrete Wavelet Transform (DWT)

In the DWT, a basis function known as the mother wavelet, $\psi(t)$, is scaled and translated to approximate a discrete signal. The 'wavelets' are sets of the function $\psi_{j,k}(t)$ formed by these scalings and translations as shown in equation (2.35). The translation parameter, k, determines the location of the wavelet in the time domain while the dilation parameter, j, determines the scale or extent of the time-frequency location.

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{2.35}$$

There are several mother wavelets that can be used as a basis function. In this study the Daubechies wavelet family was chosen for their features of orthogonality, smoothness, and compact support (Daubechies, 1988). This decision was also influenced by previous researchers who chose Daubechies wavelets to model fluidized bed conditions (Lu and Li, 1999; Guo, Yue, and Werther, 2002; Ellis et al., 2003; Zhao and Yang, 2003).

The mother wavelet, $\psi(t)$ as in equation (2.36), has a companion function known as the scaling function, $\varphi(t)$, shown in equation (2.37). Together, these functions provide a two-scaling relationship that is used to transform or "decompose" a signal. The discrete low-pass filter, H_n, and the discrete high-pass filter, G_n, are unique for each mother wavelet.

$$\psi(t) = \sqrt{2}\sum_{n=1-N}^{N} \mathbf{G}_n \varphi(2t - n) \tag{2.36}$$

$$\varphi(t) = \sqrt{2}\sum_{n=0}^{N} \mathbf{H}_n \varphi(2t - n) \tag{2.37}$$

2.4.2 Multiresolution Analysis (MRA)

In multiresolution analysis the DWT is used to decompose a discrete signal or function, S*(n)*, into an orthogonal set of approximation and detail functions, A*j(n)* and d*j(n)*, respectively (Mallat, 1989). The approximation functions are the low-frequency components of the signal containing the coarse features and are obtained using the low-pass filter, $\mathbf{H}_n$ (equation 2.38).

$$\mathrm{A}(j+1)(n) = \sum_m \mathbf{H}_{m-2n} \mathrm{A}j(m) \tag{2.38}$$

The detail functions are the high-frequency components of the signal and are extracted from the signal A*j* using the high-pass filter, $\mathbf{G}_n$ (equation 2.39).

$$\mathrm{d}(j+1)(n) = \sum_m \mathbf{G}_{m-2n} \mathrm{A}j(m) \tag{2.39}$$

Initially, the original signal is decomposed into an approximation signal and a detail signal using the low pass and high pass filters, respectively. This is followed by successive decompositions of the *jth* level approximation signal. These decompositions are repeated until a *J*-level decomposition has been reached. The level of decomposition is up to the discretion of the user, however at some point there will be no features left in the approximation signal to extract. A 5-level wavelet decomposition is depicted in Figure 2.16. The original signal can be reconstituted as the summation of the detail signals with the *Jth* level approximation signal (equation 2.40).

$$\mathrm{S}(n) = \mathrm{A}J(n) + \sum_{j=1}^{J} \mathrm{d}j(n) \tag{2.40}$$

Further details of wavelets, the discrete wavelet transform, and multiresolution analysis can be found in Daubechies (1992) and Burris, Gopinath, and Guo (1998).

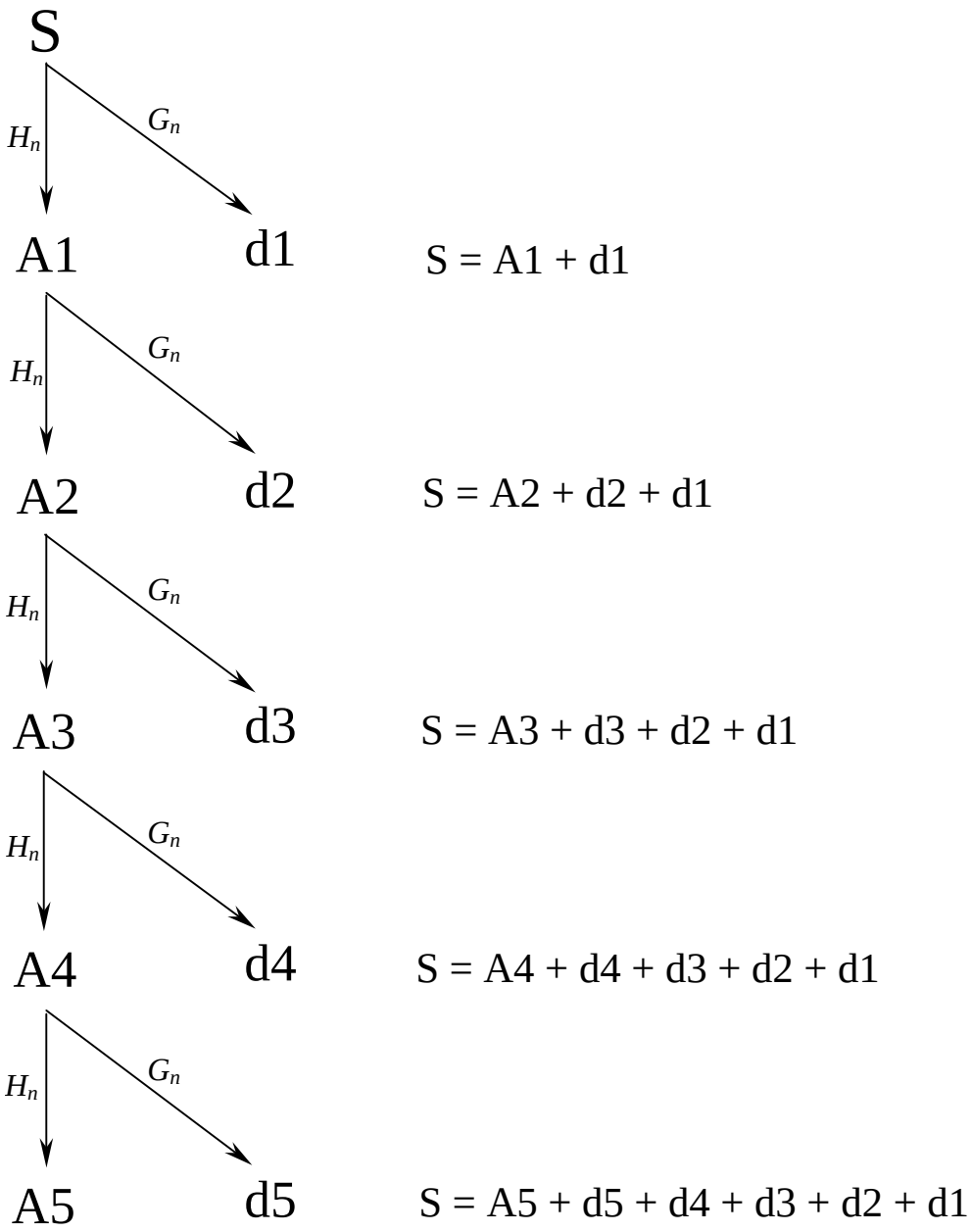

Figure 2.16: 5-level wavelet decomposition of a pressure fluctuation signal using Daubechies 2 mother wavelet (Mallet, 1989)

2.5 Wavelet Analysis of Fluidized Bed Systems

While the application of wavelet analysis to the fluid bed drying process is absent from the literature, there have been recent attempts to analyze the multiscale behavior of other fluid bed systems using wavelets. Wavelet transforms have been used to analyze signals from fiber optical probes (Ren et al., 2001; Ellis et al., 2003), Phase Doppler Particle Analyzers (Ren and Li, 1998), differential pressure measurements (Ellis et al., 2003), and high-frequency pressure fluctuation measurements (Lu and Li, 1999; Guo, Yue, and Werther, 2002; Guo et al., 2003; Zhao and Yang (2003)) in fluidized beds. Regardless of the type of measurement, this multiresolution approach has shown great promise for analyzing scale-specific behaviors and many parallels in the analyses are beginning to emerge.

The use of wavelet transforms for the analysis of fluidized bed systems is a novel concept. As such, there are no standard conventions or procedures for performing the analysis. The selection of mother wavelet and level of decomposition varies considerably in the literature. Wavelets from the Daubechies family (Daubechies, 1988) have been used in the most recent attempts (Guo, Yue, and Werther, 2002; Ellis et al., 2003; Guo et al., 2003; Zhao and Yang, 2003). Zhao and Yang (2003) provided a methodology for the choice of wavelet and level of decomposition used in their study. Their selection was based on the minimum absolute error between the original signal and the signals reconstructed from the approximation and detail coefficients at each level.

Most researchers agree that three scales of motion are adequate for classifying the behavior in fluidized beds. Ren and Li (1998) labeled these scales as the

macroscale, mesoscale, and microscale. The definitions of each of these scales can vary and the convention used by Li (2000) has been adopted in this study. Using this convention, the macroscale motion refers to the bulk motion of the particle-fluid suspension within the boundaries of the apparatus and encompasses bubbling phenomenon and apparatus-scale effects. The mesoscale, also known as the cluster scale, involves interaction between the dilute broth phase and the dense cluster phase. Finally, the microscale refers to particle-scale phenomenon involving particle-particle and particle-apparatus interactions.

Multiscale behaviors in high-frequency pressure fluctuations have been suggested by several researchers using wavelet analysis. However, with the exception of the macroscale behaviors, a definitive correlation between the decomposed pressure signals and scales of motion has not been fully demonstrated. Guo, Yue, and Werther (2002) used a 5-level decomposition using Daubechies 3 (db3) wavelet to analyze a pressure fluctuation signal from a bubbling fluidized bed of ash. They found that the dominant frequency of the d5 signal was equal to the dominant frequency of the original signal and was, therefore, an indicator of bubbling frequency (ie. macroscale behavior). They also suggested that the d3 and d4 components revealed information about particles rising below bubbles and that the d1 and d2 scales corresponded to particle motion. However, in a similar investigation Guo et al. (2003) performed an 8 level decomposition using Daubechies 7 (db7) wavelet, and while they found that the d6 peaks once again corresponded to bubbling frequency, they concluded the high frequency scales d1, d2, and d3 were the result of Gaussian White Noise generated by jetting action above the distributor.

Zhao and Yang (2003) performed a 9-level decomposition using Daubechies 2 (db2) wavelet on a pressure signal from a bubbling fluidized bed. They concluded that the A9 approximation signal depicted the macroscale behavior of the bed. Zhao and Yang interpreted the level 3 to 9 detail signals (d3 to d9) as reflecting the mesoscale behaviors and the level 1 and 2 detail signals (d1 and d2) as exhibiting microscale characteristics but, once again, did not provide any direct evidence of these relationships. Rather, the interpretations were based on the results of a Hurst analysis on the individual scales.

2.6 Wavelets and Latent Variable Modeling

The literature contains several examples of wavelet transformations being integrated with PCA and PLS analyses. Most of these applications involve decomposing the entire data set using wavelet transforms and constructing models using the wavelet coefficients (Kosanovich and Piovoso, 1997). Multiscale Principal Component Analysis (MSPCA) is a well-known variation of this approach (Bakshi, 1998). In MSPCA, the data is transformed using wavelets before individual PCA models are constructed from the coefficients at each scale. This is done to take advantage of the time-frequency localization properties of the wavelet transform. When wavelets are used along with PCA in statistical process control applications, phenomenon previously only detectable in the frequency domain can be identified.

Wavelets have been also been used as a compression tool to reduce large data sets before entering the data into PLS or PCA models (Trygg and Wold, 1998; Trygg, Kettaneh-Wold, and Wallbacks, 2001). In other instances, wavelets have been used to filter high-frequency noise or low-frequency drift effects from data before

constructing latent variable models (Teppola and Minkkinen, 2000). Compression and filtering are typically accomplished by decomposing the data set with wavelets before reconstructing it without the scales believed to be the result of noise, errors, or drift. By excluding these scales, the number of coefficients required to describe the signal is reduced.

The approach of Bjork and Danielsson (2002) differed from previous attempts in that the entire data set was not transformed with wavelets. They recognized the ability of wavelets to extract information from several scales of a multiscale signal. In their study, wavelets were used as a pre-filter to separate an acoustic measurement into various scales. The information from these scales was subsequently summarized with Fast Fourier Transforms (FFT) and then used as input for a PLS model. By extracting more information from the acoustic measurement, they were able to improve their model predictions.

Wavelets have not previously been combined with latent variable modeling techniques for the purpose of modeling the hydrodynamic conditions in any fluidized bed system.

3. DATA ANALYSIS PROCEDURE

The data analyzed in this study was generated in an experimental campaign conducted by Mr. Gareth Chaplin at Merck-Frost Laboratories, Kirkland, Quebec in 2002. Chaplin's study consisted of two distinct components. The first, referred to as the dry bed study, examined the steady-state behavior of dry granulate in the fluid bed dryer. While the other, referred to as the batch drying study, involved the batch drying of wet granulate from an initial moisture content of approximately 33 wt.% to 5 wt.%. Detailed descriptions of the experimental procedures, materials, and apparatus can be found in Chaplin, Pugsley, and Winters (2004a-b).

In this study, PCA and PLS were used to analyze the data from both the dry bed study and the batch drying study. In both cases, the hydrodynamics in the dryer were characterized using pressure fluctuation measurements taken at a sampling rate of 400 Hz. The basic approach to the analysis was similar to that used by Bjork and Danielsson (2002) where the wavelet transformation was used solely as a pre-filter to extract the maximum available information from a multiscale signal. However, in the current research the signal was a pressure fluctuation signal rather than an acoustic measurement.

The procedure for the multiresolutional decomposition and numerical preprocessing of the pressure signals was the same for the both the dry bed study and

the batch drying study. An optimal wavelet and level of resolution were chosen, the signals were decomposed, and statistics describing the new signals at each level were entered into the data sets with the remaining measured variables. Additional preprocessing of the batch drying study data was required prior to performing the latent variable analyses in order to align the trajectories and normalize the data between batches.

3.1 Selection of the Optimal Wavelet

Daubechies wavelets were selected as the mother wavelet for their properties of orthonormality, compact support, and high number of vanishing moments. These properties allow for efficient and concise reconstruction of the original signal, ease of computation, and the storage of a large amount of information in a small number of wavelet coefficients. The decision to use Daubechies wavelets was also influenced by the work of previous researchers who chose Daubechies wavelets to study fluid bed systems (Lu and Li, 1999; Guo, Yue, and Werther, 2002; Ellis et al., 2003; Guo et al., 2003; Zhao and Yang, 2003).

The optimal wavelet and level of decomposition was chosen using the method of Zhao and Yang (2003). Each signal was decomposed using each Daubechies wavelets (db2- db10) at levels 1 to 12 and then reconstructed. The absolute error between the reconstructed signal and original signal was calculated and plotted (Appendix A).

For the dry bed study pressure signals, Daubechies 2 (db2) wavelet produced the lowest reconstruction error at all levels. A 5-level decomposition was chosen as a compromise between reconstruction error and resolution. In this context, resolution

was interpreted as a level of decomposition sufficient to separate the behaviors at each scale of motion. When the procedure was repeated for the pressure signals from the batch drying tests, the choice of wavelet was not so clear. In the end, Daubechies 2 (db2) wavelet was selected in order to be consistent with the dry bed analysis. A 7-level decomposition was selected as optimal based on minimum error (Appendix A).

3.2 Pressure Signal Preprocessing

The 400 Hz pressure signals required several preprocessing steps before they could be used in a PCA or PLS model. This preprocessing included selection of a mother wavelet and level of decomposition as previously discussed. The resulting approximation and detail signals required additional processing in order to provide meaningful information in a format that could be easily integrated with the remaining measured variables in the data sets.

When a signal is decomposed using the DWT algorithm, the approximation and detail signals are initially in a compressed format referred to as the "downsampled" wavelet coefficients. These coefficients must be decompressed or "upsampled" to produce a set of signals in the original time domain. In this study, the pressure signals were decomposed and upsampled to produce a *J*th level approximation signal and *J* detail signals in the time domain.

In the next stage of preprocessing, the features of each upsampled signal were summarized using a more traditional approach for describing the characteristics of a discrete signal. First, a FFT was performed to determine the dominant frequency of each signal. Then the descriptive statistics: standard deviation (stdev), mean (mu), median (medn), skewness (skew), and kurtosis (kurt) were calculated for each

upsampled signal. The dominant frequencies and descriptive statistics were treated as individual variables in the latent variable models. A naming convention was implemented where each of these variables were identified by the level of the decomposition (ie. A5 dominant frequency, A5 standard deviation, A5 mean, A5 median,…..,d1 median ,d1 kurtosis, d1 skewness). For a 5-level decomposition on a single pressure signal, 36 new variables were created that described the characteristics of the A5, d5, d4, d3, d2, and d1 signals. Similarly, 48 calculated variables were used to describe each signal after a 7-level decomposition.

One might argue that transforming the data in this way limits the benefits that wavelet analysis provides for examining events at each scale because the remaining measured variables are left out of the transformations. However, wavelets were used in this analysis simply to extract information from the pressure signals. The methodology followed here allows other process variables (ie. inlet temperature, outlet temperature, product temperature, moisture, bed mass, superficial velocity, and sensor position) to be easily integrated with the pressure signal data in the latent variable analyses. Thus, all variables are represented over the same time intervals and correlations between the fluid bed scales of motion and other measurements are possible.

3.3 Selection of Signal Length

In the dry bed study, signals under each condition (ie. a single superficial velocity, sensor position, and granulate formulation) were recorded for 180 seconds. At a frequency of 400 Hz, this equates to 72000 data points. In an industrial application, computation time may be of concern. Additionally, if the same

information could be extracted from signals of less than 180 seconds, the signals from the dry bed study could be subdivided resulting in more data for the analysis.

To address this issue, each signal was split into 15, 30, and 60-second intervals and a 5-level decomposition using the Daubechies 2 wavelet was performed. The dominant frequency and descriptive statistics were subsequently calculated at each level. The resulting computation times using a computer with a 3.4 GHz processor are summarized in Table 3.1.

PCA models were formulated from the transformed data at each signal length (15, 30, 60, and 180 seconds) and compared. The loads plots were nearly identical between all cases. The scores plots were also comparable. This lead to the conclusion that pressure signals shorter in duration than 180 seconds could be used without altering the model structures. In general, more points are better due to the statistical nature of the analysis. However, when applied to a transient process, too long of an interval may obscure important transient phenomenon.

Table 3.1: Computation time for signal decomposition

Signal Length (seconds)	Number of Data Points	Processing Time (seconds)
180	72000	22.0
60	24000	3.5
30	12000	1.5
15	6000	1.0

It is clear from Table 3.1 that the computational time for decomposing signals and calculating the new statistics would not be a limiting factor. Using a moving window approach, a new data point could be generated on process monitoring charts every couple of seconds (for signals 30 seconds or less).

A 30-second signal length for each point was used in both the dry bed analysis and in the batch drying study. This decision was based on minimizing the computation time, providing a reasonable window for monitoring transient phenomenon, and retaining enough points to maintain statistical confidence.

3.4 Treatment of Batch Drying Data

The data from the batch drying experiments required additional processing before the wavelet decomposition and latent variable modeling could be performed.

3.4.1 Trajectory Alignment

Differences in initial conditions between batch drying runs resulted in different total drying times. For instance, runs where the inlet temperature was 55°C and the initial bed mass was 3.25 kg (33 wt.% H_2O) took close to 90 minutes to complete a drying cycle while runs with an inlet temperature of 75°C and an initial bed mass of 2.75 kg (33 wt.% H_2O) took only half that time. As a result, different phenomenon or key states in the drying process took place at different points in time. This complicates any sort of analysis where the goal is to formulate a model using data from several batch runs such as in MPCA modeling.

One solution to this difficulty is to select an indicator variable to replace time as the independent variable. The remaining variables can then be aligned with respect to the indicator variable. In this study, moisture was chosen as the indicator variable. All batches began with approximately 33 wt.% moisture and ended with 5 wt.% moisture. By using moisture as the new independent variable in the study, an assumption was made that the hydrodynamic behaviors and key states that occur in the drying cycle occur at the same moisture content between batches.

3.4.2 Interpolation of Variables

In the batch drying tests, the pressure fluctuation signal was the only continuously recorded variable throughout the runs. The remaining variables were manually recorded at sporadic intervals throughout the run. To include these variables in the latent variable models, values were required for each increment in the moisture scale. The moisture scale was divided up into 100 evenly spaced points between 30.00% and 5.25% and a linear interpolation was performed to determine the corresponding values for each variable. A linear interpolation was deemed adequate as the measured variables were recorded to the nearest minute.

The experimental time elapsed was also interpolated at each moisture decrement. The linearly interpolated time values were used to select the appropriate portion of the pressure signal for wavelet decomposition. A 30-second section of the pressure signal was transformed at each of the 100 moisture intervals (ie. 15-seconds preceding the interpolated time value and 15-seconds after). In this way, the wavelet decomposed pressure signals corresponded to each of the manually recorded measured variables. Thus, the interpolation of variables was required before the wavelet decompositions.

3.5 PCA Modeling

PCA models were constructed to analyze the data from both the dry bed study and the batch drying study. These PCA models could be classified according to the objectives of the model. The first type of PCA model was constructed to analyze the effects of the variables manipulated in the study. The remaining models were

constructed to determine whether PCA could be used for classification or multivariate statistical control purposes in the fluid bed dryer.

Data sets were constructed with the wavelet decomposed pressure signal variables and the manually recorded measured variables. Variables were autoscaled and the PCA model was calculated. The number of components retained in each model varied and was determined using the rules of thumb previously discussed. In general, the number of components retained was kept to a minimum in order to exclude stochastic variances from the models.

Where the goal of a PCA model was to examine the effect of a particular variable, a data set was constructed such that the only variance between samples was due to that variable. For instance, if the effect of superficial velocity was being investigated, the data used to construct the model would contain a range of velocities and the remaining variables would be fixed. In principle, this should ensure that the variance between samples was due to the variable under investigation (ie. superficial velocity). The scores plots were then examined to see whether there was a systematic variation in how the points projected in the model subspaces. Loadings plots were also scrutinized to determine how the variables, and the pressure signal components in particular, combined to form each principal component.

The approach to building PCA models for classification and statistical control was somewhat different. Models were constructed using the data containing only the desired features. These features could include a single condition such as an acceptable granulate particle size distribution or they could include a range of conditions such as several superficial velocities. The models were then validated by

projecting new data into the model space. When points exceeded Hotelling's T^2 or the *Q*-residual 95% confidence limits, contribution plots were constructed to determine whether an accurate diagnosis of the fault could be determined. This was possible because any differences in the new data were known. Data that was from the same mode of operation as the model was also used for validation. In these cases, the data was projected into the model subspaces to ensure that they were within the 95% confidence limits.

3.6 PLS Modeling

PLS models were constructed to examine the individual effects of the variables under investigation in each study. This proved to be an excellent way to isolate the effects of a single variable. Additionally, the results were used as a check on the results of the PCA models. PLS models were also constructed to determine whether the PLS technique could be used to construct "soft sensors" or predictive models of system properties based on other physical measurements of the system.

For the PLS model construction, a single response or variable under investigation was taken as the **Y**-block and the remaining variables as the **X**-block. A PLS regression was performed and a decision was made on how many components to retain. The **X**-block weight vectors and the regression vector were then examined to determine the structure of the latent variables that best described the **Y**-block variance. This is a valid approach because the **X**-block latent variables are calculated based on the maximum **Y**-block variances. Consequently, the **X**-block subspaces are seen to be those that are the most descriptive of the predicted variable. The **X**-block

latent variables were checked for consistency against the principal components of the PCA modeling.

3.7 Software

The analysis in this study was carried out using the MATLAB® version 7.0 development environment. Wavelet decompositions and FFT were performed using the algorithms from the Wavelet Toolbox version 3.0 and the Signal Processing Toolbox version 6.2, respectively. Wavelet algorithms were validated using the WaveLab Toolbox and dominant frequencies obtained from the FFT were compared with the results obtained by Chaplin (2005). Descriptive statistics were calculated using algorithms in the Statistics Toolbox version 5.0.

PCA, MPCA, and PLS (SIMPLS and NIPALS) algorithms were written in MATLAB® and compared to the results obtained using algorithms from the PLS Toolbox version 3.5 and the Statistics Toolbox version 5.0.

4. RESULTS AND DISCUSSION

PCA, MPCA, and PLS models were constructed to analyze the data generated by Chaplin (Chaplin, Pugsley, and Winters, 2004a-b). Chaplin's investigation consisted of two distinct experimental components. The first study, referred to as the dry bed study, examined the steady-state behaviour of dry granulate in a fluidized bed dryer. The other, referred to as the batch drying study, consisted of the batch drying of granulate from an initial moisture content of 33wt.% to approximately 5wt.%. The results of the modeling are presented according to each respective study.

4.1 Dry Bed Study

The dry bed study examined the effects of three variables on the steady-state behavior of dry pharmaceutical granulate in a laboratory-scale fluidized bed dryer. These variables were superficial gas velocity, pressure sensor position, and particle size distribution of the granulate. In all, 35 different combinations of these variables were tested in triplicate resulting in a total of 105 tests. The details of each experiment are tabulated in Appendix B.

Chaplin (2005) logged the pressure signals at a sampling rate of 400 Hz for 180 seconds at each test condition. In the current study, these signals were parsed into 30-second sections resulting in a total of 630 data points for the analysis. Models constructed using the subdivided signals were found to be nearly identical to those

constructed using the entire 180-second signals. Unless otherwise stated, the results presented are from the analysis conducted using the 30-second pressure signals.

When the pressure signals were decomposed using a 5-level Daubechies 2 (db2) mother wavelet it was obvious that the A5 approximation signal contained the same coarse (low-frequency) features of the original pressure signal (Figure 4.1). The dominant frequencies of the two signals were compared and found to be equal in all cases. The dominant frequency of pressure fluctuations in a fluidized bed is an indicator of bubble frequency. Thus, the A5 approximation signals contain information about the macroscale behavior in the fluid bed dryer. Similar results have been reported in the literature by Lu and Li (1999), Guo, Yue, and Werther (2002), Guo et al. (2003), and Zhao and Yang (2003).

After the wavelet decompositions were completed, the dominant frequency, standard deviation, mean, median, kurtosis, and skewness were calculated for each approximation and detail signal. Each of these variables were plotted and examined for trends. The A5 mean and A5 median variables exhibited an exponential decay-like behavior that appeared to be a function of the order the experiments were performed in at each pressure sensor position and batch of granulate (Figure 4.2). The same statistics were calculated for the original pressure signals and the trend was seen to exist there as well.

The order of the dry bed experiments was not randomized. For each sensor position, a single batch of granulate was placed in the dryer and the velocity was incrementally raised from 1.95 m/s to 2.92 m/s. Each 3-minute pressure signal was recorded at a single superficial velocity three times before it was increased. In

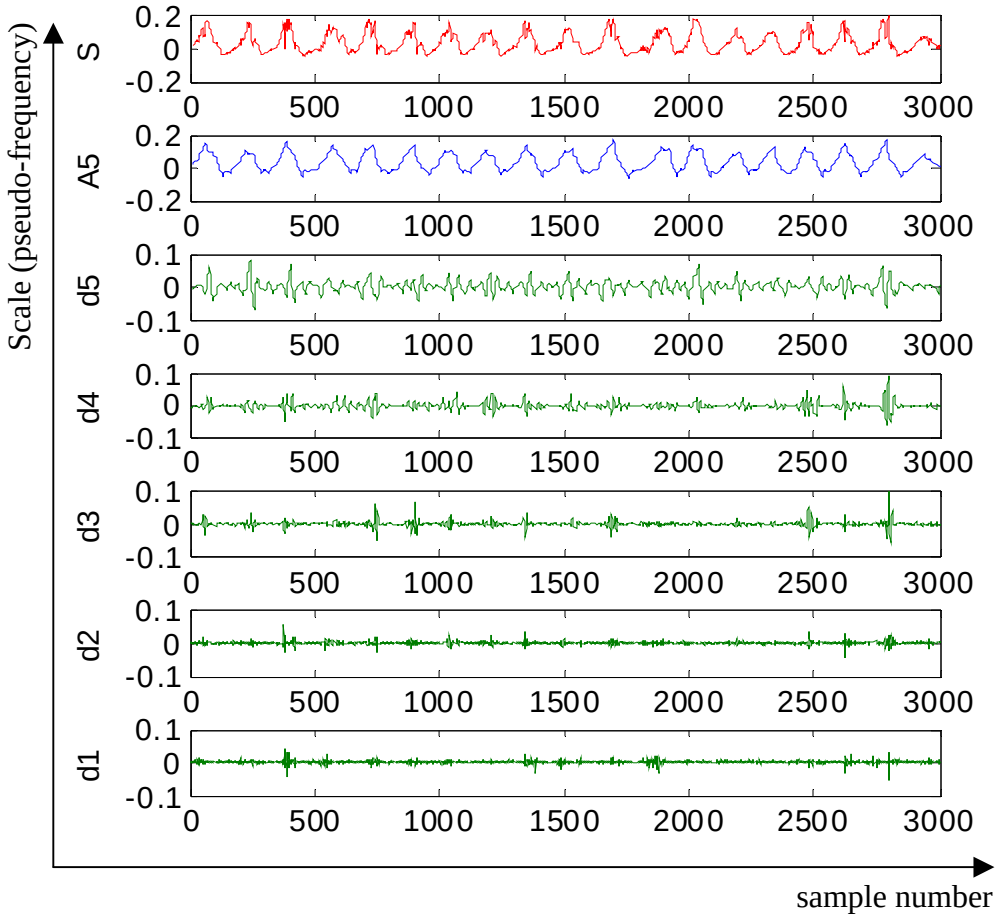

Figure 4.1: 5-level wavelet decomposition of a 400 Hz pressure signal from the dry bed study using Daubechies 2 mother wavelet. A 7.5 second sample (3000 pts. at 400 Hz) is shown for each signal.

other words, the conditions in the dryer were held constant for 9 minutes before increasing the superficial velocity. Consequently, an effect that was the result of experiment order became encoded in the data. Fines entrainment or the segregation of larger granulate particles in the bed might explain this type of behavior. Fines entrainment could block the filter bag and change the pressure drop across the bed while particle segregation may change the characteristics of compression waves within the bed. However, this could not be proven with the existing data set.

Therefore, an attempt was made to remove this hysteresis-like effect by deleting the mean and median statistics from the dry bed data set.

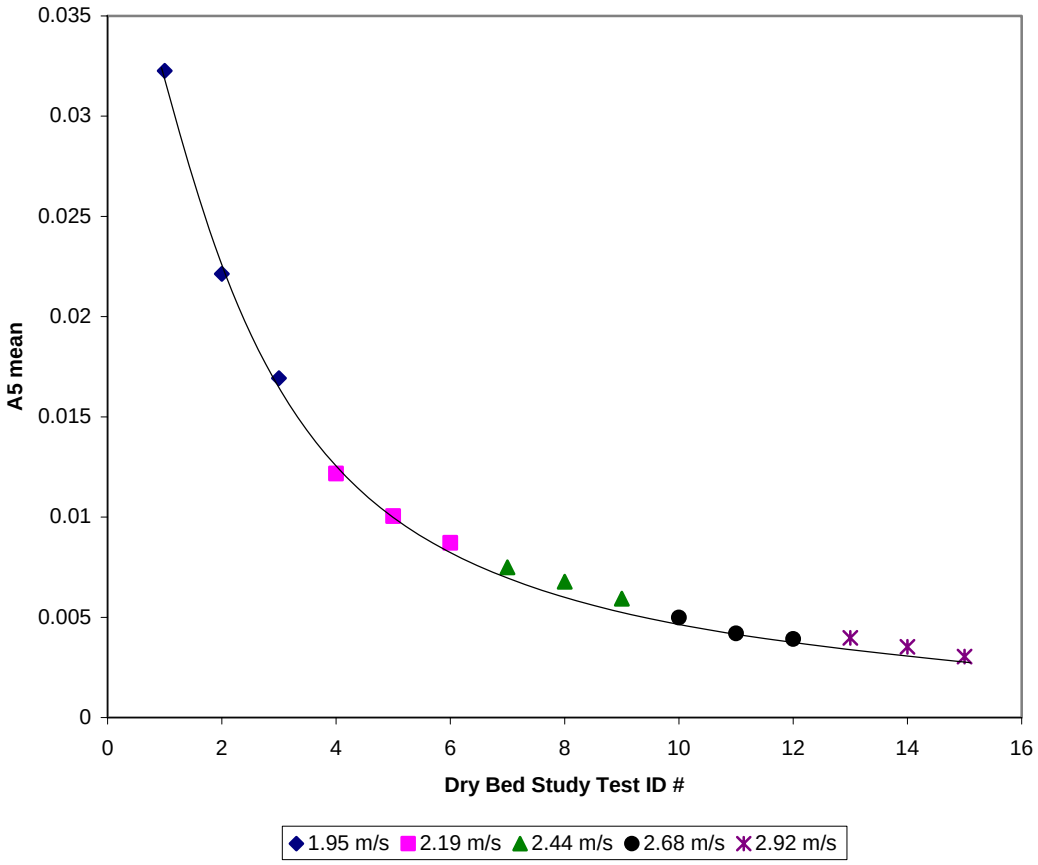

Figure 4.2: A5 mean as a function of experimental order. Both the A5 mean and A5 median statistics showed an exponential decay in value for each test at a single pressure sensor location and batch of granulate.

Before examining the effects of the individual variables, some overall PCA models containing all 630 points were constructed and examined for general trends. The data sets contained all variables describing the approximation and detail signals as well as the superficial velocity, pressure sensor position, and the mean particle size of the granulate used in each test. Scores plots showed systematic variances correlating to each of the 3 manipulated variables (Appendix C). As a result, the

influence of the manipulated variables in the data sets was questioned. To address this issue, the models were reformulated, however this time the data was limited solely to the decomposed pressure signal data. The new scores plots were found to present the same systematic variations (Appendix C). It was therefore concluded that any variance in the data as a result of superficial velocity, pressure sensor position, or granulate particle size distribution could be characterized solely with the use of the pressure fluctuation data. Subsequently, PCA models used to analyze the dry bed conditions were calculated exclusively from pressure fluctuation data.

4.1.1 Effect of Superficial Velocity

Superficial velocity is an important parameter in the operation of a fluidized bed dryer. Changing the superficial velocity has a direct impact on the hydrodynamics and maintaining optimal conditions is critical to product quality. In addition to superficial velocity, there are many other factors that affect these hydrodynamics. However, if the hydrodynamic state in the dryer can be determined using a measurement such as pressure fluctuations, then knowing how the superficial velocity affects that state provides a basis for control.

Individual PCA models were constructed containing the full range of superficial velocities used in the experiments (1.95, 2.19, 2.44, 2.68, and 2.92 m/s). Each model considered only a single sensor position and granulate formulation. Models with 3 retained components that described between 40% and 50% of the variance in the data were found to be optimal. The first principal component, also referred to as the dominant principal component, typically accounted for more than

half the model variance. The projection of scores as a function of superficial velocity was very apparent.

An interesting observation was made from examining the scores projected only on the dominant principal component. Points generated at the lowest superficial velocity, 1.95 m/s, showed a strong negative projection (Figure 4.3). This contrasted the remaining points that began projecting slightly positively with increasing velocity. It was later determined that the fluidization regime in the dryer at 1.95 m/s was on the boundary between the bubbling and the turbulent regimes. Thus, it was determined that PCA models are likely capable of identifying changes in fluidization regimes. Data containing a wider range of superficial velocities would be required to further investigate the full spectrum of possibilities of using this method to classify and quantify the differences between flow regimes.

For the analysis of the dry bed data, a convention was adopted for relating the scales of motion to the decomposed wavelet signals. The A5 - d5 components were taken to be representative of the macroscale behaviors, d4 - d3 represented the mesoscale, and d2 - d1 were assumed to represent the particle level interactions of the microscale. Based on the combined results of the study, this convention was deemed to be suitable.

The loading vector for the dominant principal component showed strong positive correlations in elements throughout all scales of motion (Figure 4.4). These correlations were greatest between the A5 and d5 variables (macroscale) and decreased for the higher frequency components (microscale). In particular, there were strong correlations between the superficial velocity, A5 dominant frequency,

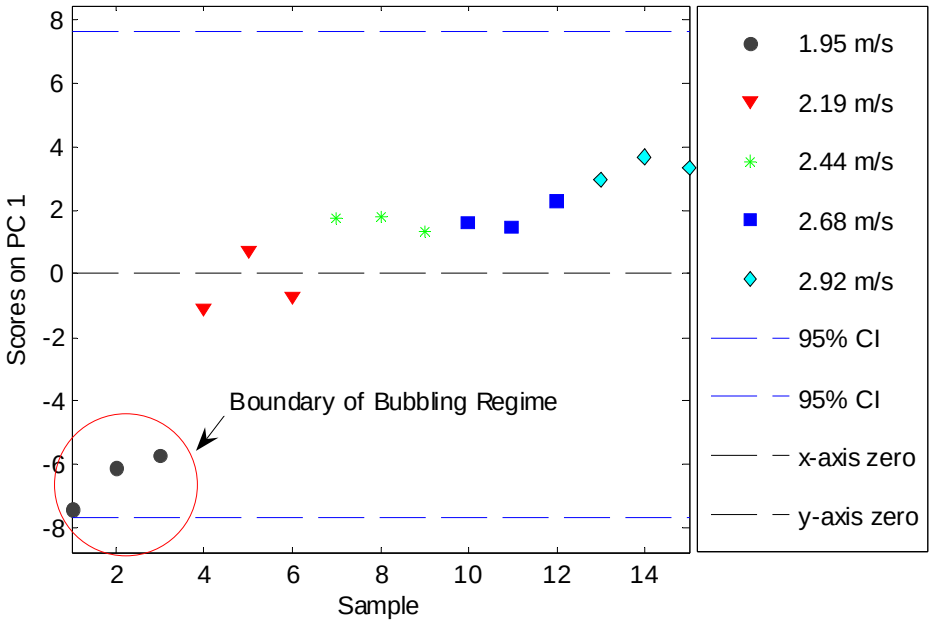

Figure 4.3: Scores plot from a PCA model constructed to analyze the effects of superficial velocity. Model shown here constructed with 180-second pressure signals for clarity of illustration.

macroscale kurtosis, macroscale skewness, and the standard deviations for all of the detail signals. Standard deviations are a measure of the signal amplitude.

PLS models were also constructed to analyze the effect of varying the superficial velocity. Models with 2 latent variables (LV) were found to be optimal. The first and second latent variables accounted for 83% and 5% of the variance in the **Y**-block, respectively. In other words, the 2 latent variable PLS models accounted for 88% of the variance in the superficial velocity.

The dominant weight vector (Figure 4.5) and the regression vector (Figure 4.6) for each PLS model show nearly identical trends to the dominant principal

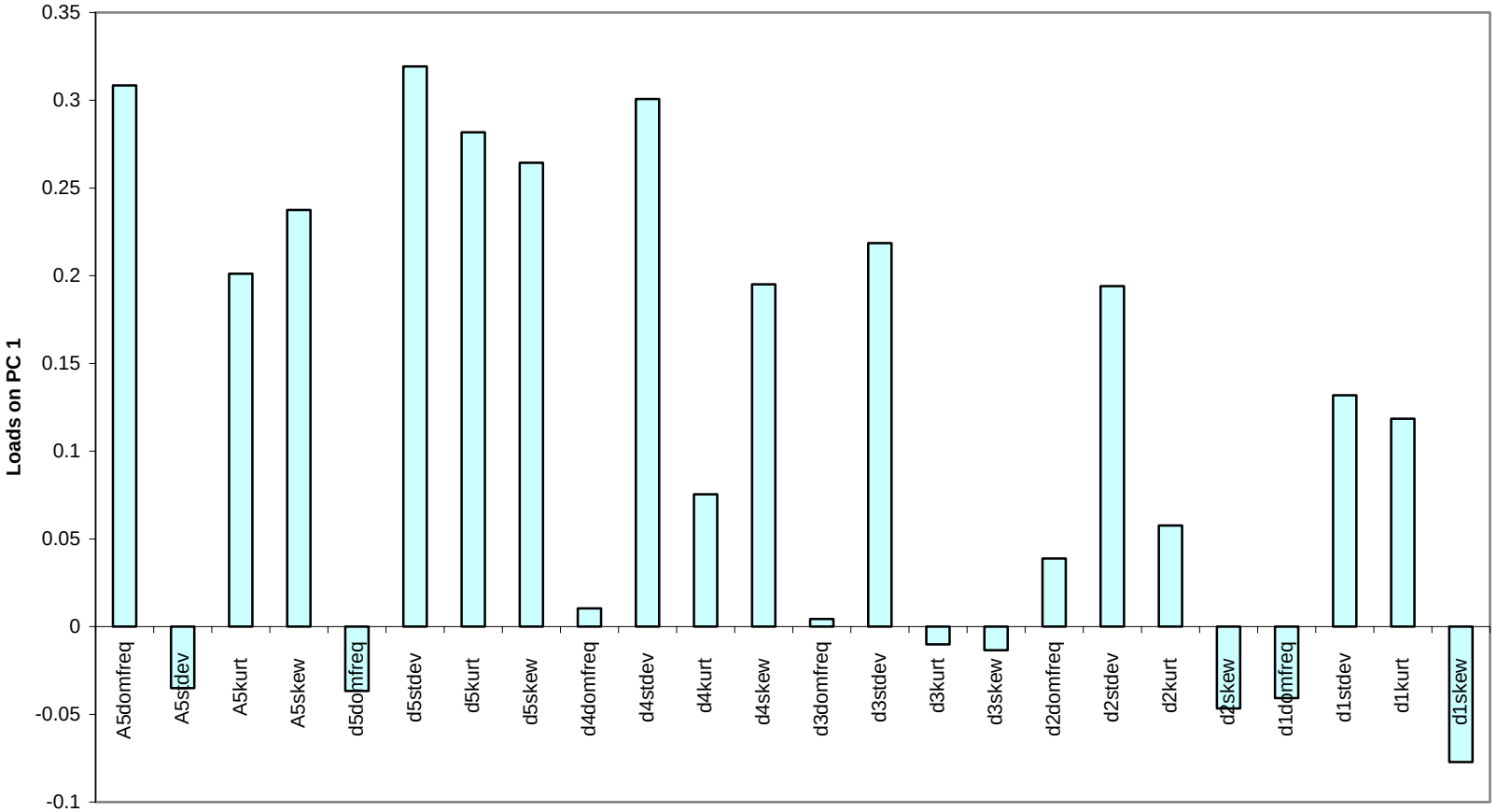

Figure 4.4: Loadings plot for the dominant principal component in a PCA model constructed to examine the effects of superficial velocity

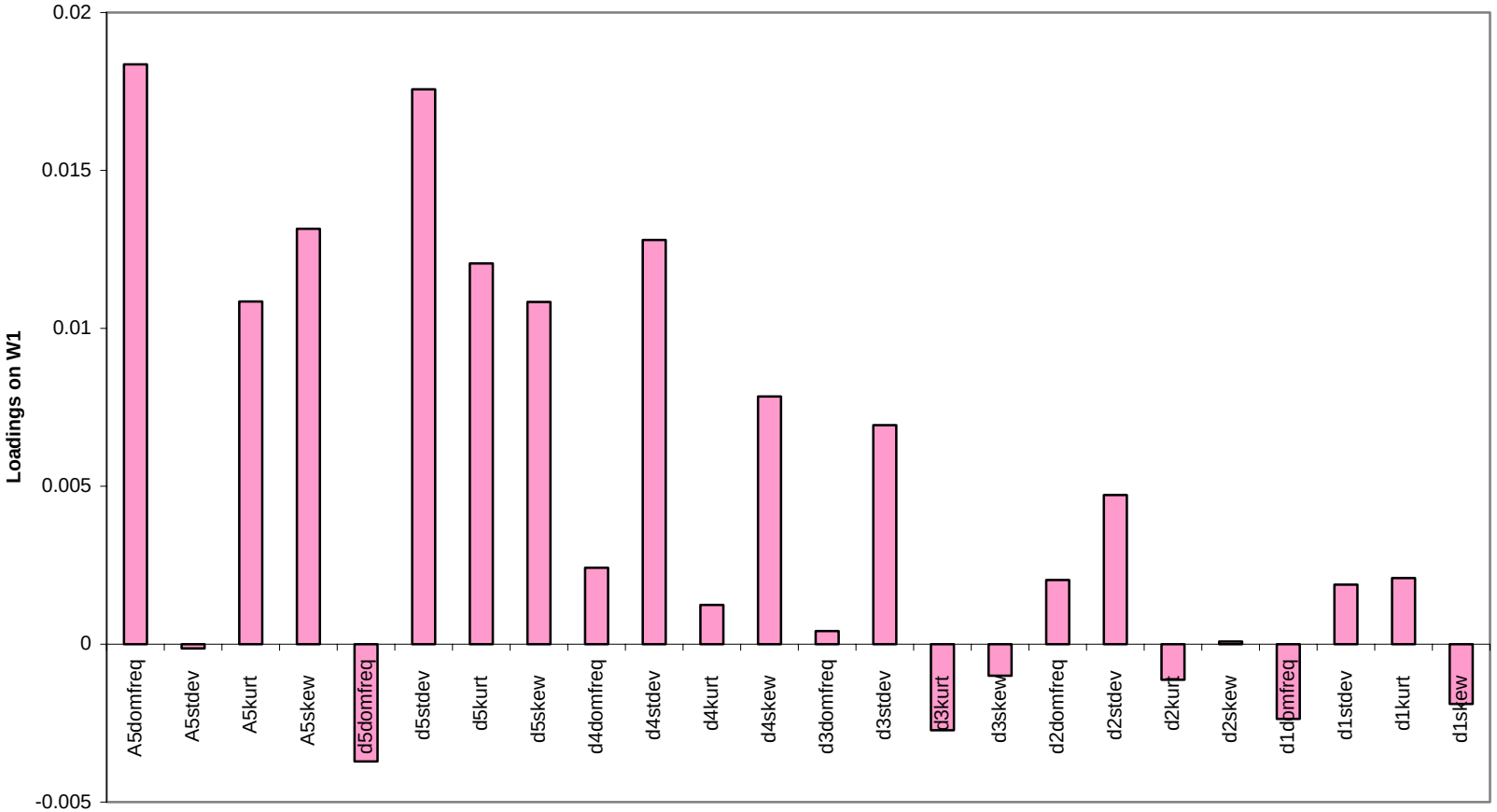

Figure 4.5: Loadings plot for the dominant weight vector in a PLS model constructed to examine the effects of superficial velocity

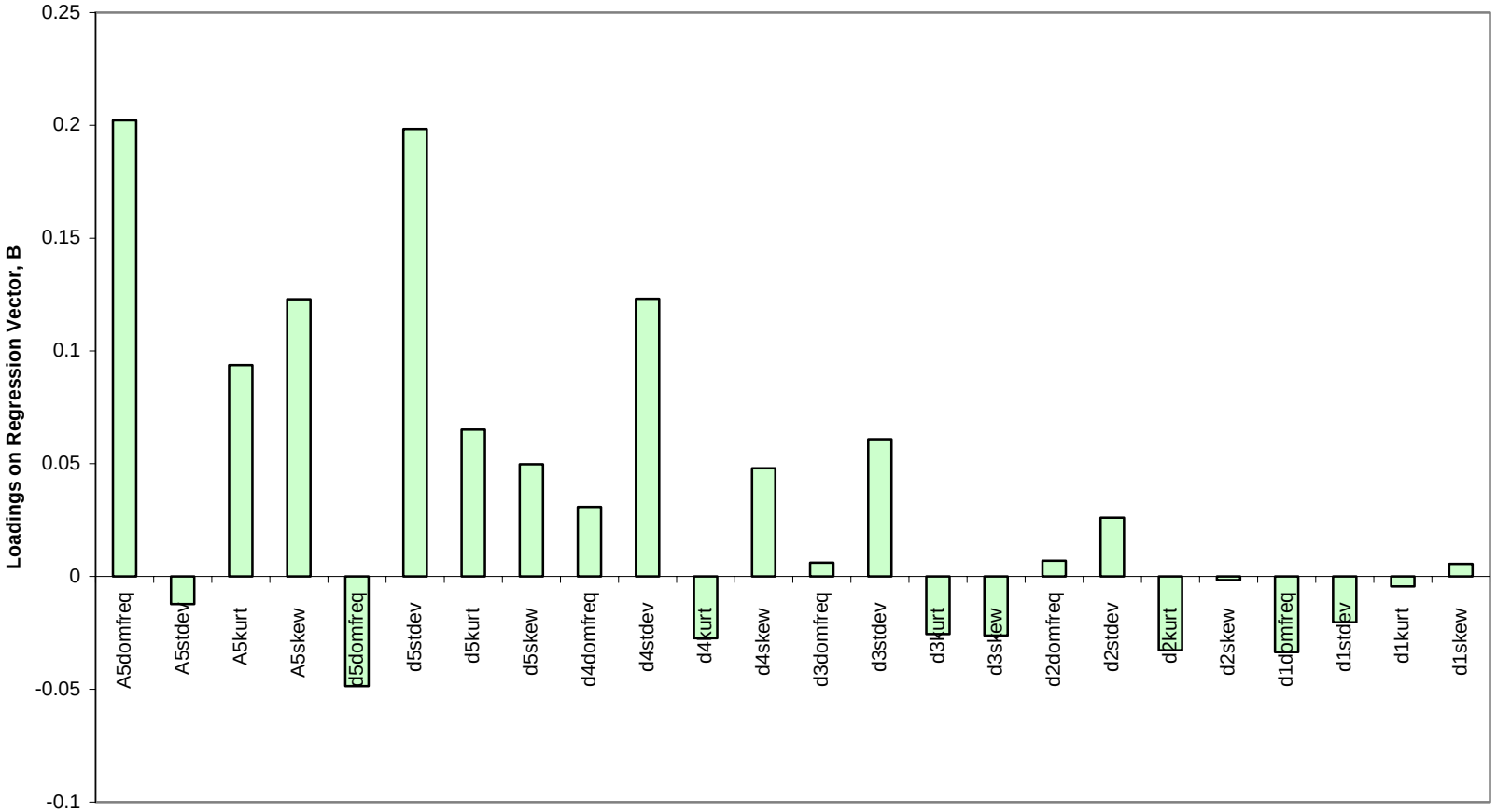

Figure 4.6: Loadings plot for the regression vector from a PLS model constructed to examine the effects of superficial velocity

components from the PCA models. However, in the case of the PLS models the decreasing correlation with the higher frequency scales is more pronounced to the extent that it appears as though there is no correlation with the microscale components. The latent variables in PLS models are essentially rotated in space to give the best description of the variance in the predicted variables. For this reason, they give a better indication of the correlations between an individual effect such as superficial velocity and the remaining variables.

Over the range examined, increasing the superficial velocity in the fluid bed dryer increases the bubbling frequency (A5 dominant frequency) and the signal amplitude of the d5, d4, and d3 detail signals. This correlation is more pronounced in the macroscale than the mesoscale. According to the PLS model results, increasing the superficial velocity (through the range tested) does not have any significant influence on the microscale interactions.

4.1.2 Effect of Sensor Position

In this study, a Piezotronics PCB-106B high-frequency piezoelectric (pressure fluctuation) sensor was used as a measurement of the system hydrodynamics. In an industrial application this measurement could be critical to the operation of the dryer. It is therefore important to determine what factors affect the signal, how they affect the signal, and how to maximize the signal quality.

The fluid bed dryer had 4 mounting positions for the pressure sensor (Chaplin, Pugsley, and Winters, 2004b). When installed, the sensor was flush with the inner wall of the bowl section of the dryer. The mounting positions were spaced 19 mm center-to-center and labeled 1 to 4 from the bottom upwards. Following an approach

similar to the investigation of superficial velocity, the effect of moving the sensor between positions was evaluated using PCA and PLS.

PCA models were calculated using data from a single superficial velocity and granulate type. A single component accounting for 22% to 28% of the variance in the data was found to be optimal. Additional retained components did not increase the error but failed to show any additional benefits for describing sensor position in the scores plots (Appendix D). The loadings plot suggests that changing the sensor position has the greatest effect on the standard deviation of the A5, d5, d4, d3, and d2 signals (Appendix D). In other words, the signal amplitude varies with sensor position at all levels except d1.

PLS models were also constructed to evaluate the influence of sensor position. PLS models retained 2 latent variables accounting for approximately 90% of the **Y**-block variance. The trends observed in the loadings plots of the dominant weight vector and the regression vector were identical to those in the loadings plot of the PCA models (Appendix D). An interesting pattern can be identified by looking at the directions of these loadings. The macroscale amplitude variables are positively correlated with sensor position and the mesoscale amplitude variables are negatively correlated with sensor position. Therefore, as you move the sensor position upwards on the dryer, the resolution of the macroscale components of the pressure signal increase whereas the higher-frequency mesoscale contributions decrease.

van der Schaaf, Schouten, and van den Bleek (1998) concluded that the amplitude of upward traveling pressure waves in a fluidized bed increased with increasing proximity to the bed surface. In their study, these upward traveling waves

were determined to be the result of bubble formation and coalescence (ie. macroscale behaviors). This agrees with the results obtained here, however the reason for the observed effect at the mesoscale is unclear.

Optimal placement of the sensor position is dependent on the particular phenomenon or conditions of interest in the dryer. Based on the results of this study, lower mounting locations may be better. They provide increased resolution of the higher-frequency mesoscale behaviors and thus a more intricate picture of the hydrodynamics. The macroscale contributions are easy to discern regardless of mounting position and microscale resolution appears relatively unaffected by sensor position. The use of multiple sensor positions at different levels in the bed might be another option for increasing the resolution of the signal at all scales.

4.1.3 Effect of Particle Size Distribution

The ability to monitor particle size distribution would provide a significant benefit for the operation of the fluidized bed drying process. Early detection and warning of agglomeration, attrition, and entrainment effects could possibly help to advert product quality problems and economic losses from rejected batches. PCA and PLS models were constructed to determine the sensitivity of the method to the particle size distribution of the granulate in the dryer.

In the dry bed study, the effects of two different granulate formations were examined. The granulate formulations were a placebo mixture of pharmaceutical binders and fillers and only differed in only one component (Chaplin, Pugsley, and Winters, 2004b). This difference resulted in different overall particle size distributions (Appendix E). The regular granulate, which was also used in the batch

drying study, had a mean particle size of 318 microns and was referred to as PSD 1. The alternative formulation, referred to as PSD 2, had a narrower and finer particle size distribution with a mean particle size of 179 microns.

The effects of particle size distribution were evaluated using the same methodology as for superficial velocity and sensor position. A PCA model containing a single principal component was found to be sufficient to identify differences in particle size distribution. The single principal component accounted for 23% to 31% of the variance in the data. Additional components did not offer significant additional information. Using scores plots, detecting differences between the granulate formulations was simple (Figure 4.7).

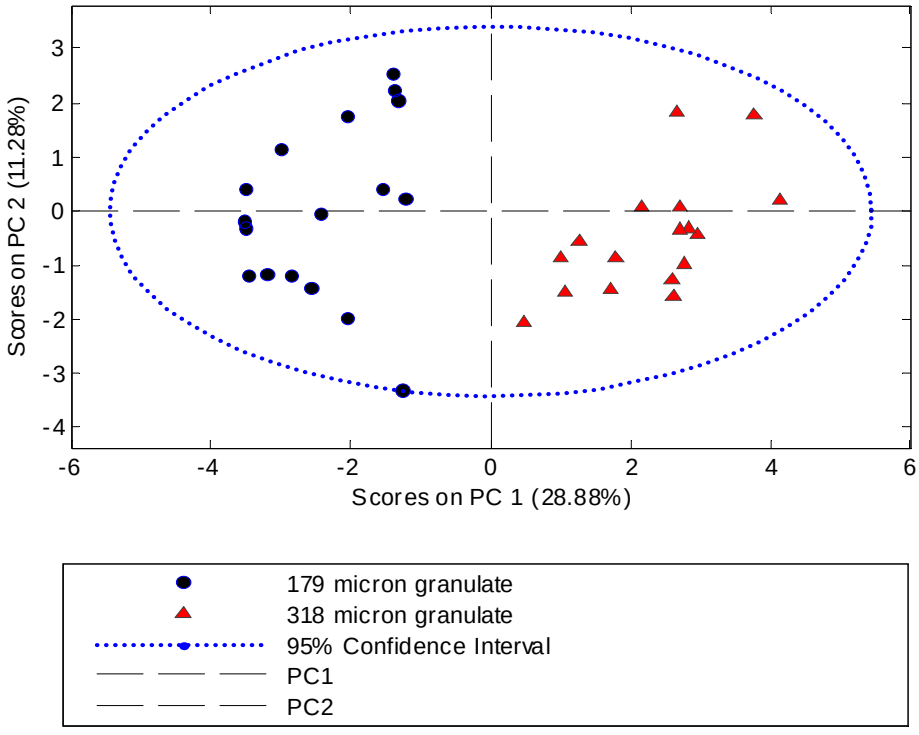

Figure 4.7: Scores plot from a PCA model constructed from data generated using the 179 micron and 318 micron granulate formulations

The PCA loadings plots showed some interesting correlations. The differences in particle size distribution were highly correlated to the *A*5 dominant frequency, and the A5, d5, d2, and d1 standard deviations. PLS models were constructed to reaffirm these correlations. PLS models with 2 retained latent variables accounted for an average of 93% of the variance in particle size distribution. The loadings plots for the dominant weight vector and the regression vector showed the same relationships as the PCA models (Appendix F).

Contributions from the high-frequency scales were explored further by trending the d1 standard deviations from the original 180-second pressure signals (Figure 4.8). The results of this plot have several implications. Firstly, the d1 standard deviation has the ability to discern between the two granulate formulations. This is an important result because it provides proof of a correlation between particle-scale phenomenon and the high-frequency components of the pressure signal. This is something previous researchers have been unable to provide. Secondly, this classification of particle size distribution is relatively insensitive to changes in both pressure sensor position and superficial velocity (within the ranges tested). This is reinforced by the previous PLS modeling results that showed that superficial velocity and sensor position did not affect the microscale or highest frequency components of the pressure signals.

The ability of the PLS models to predict particle size distribution was also evaluated. It was determined that models constructed using the data from the dry bed study could predict the mean particle size within 32 microns (95% confidence level). This could possibly be improved with enough data containing only batch-to-batch

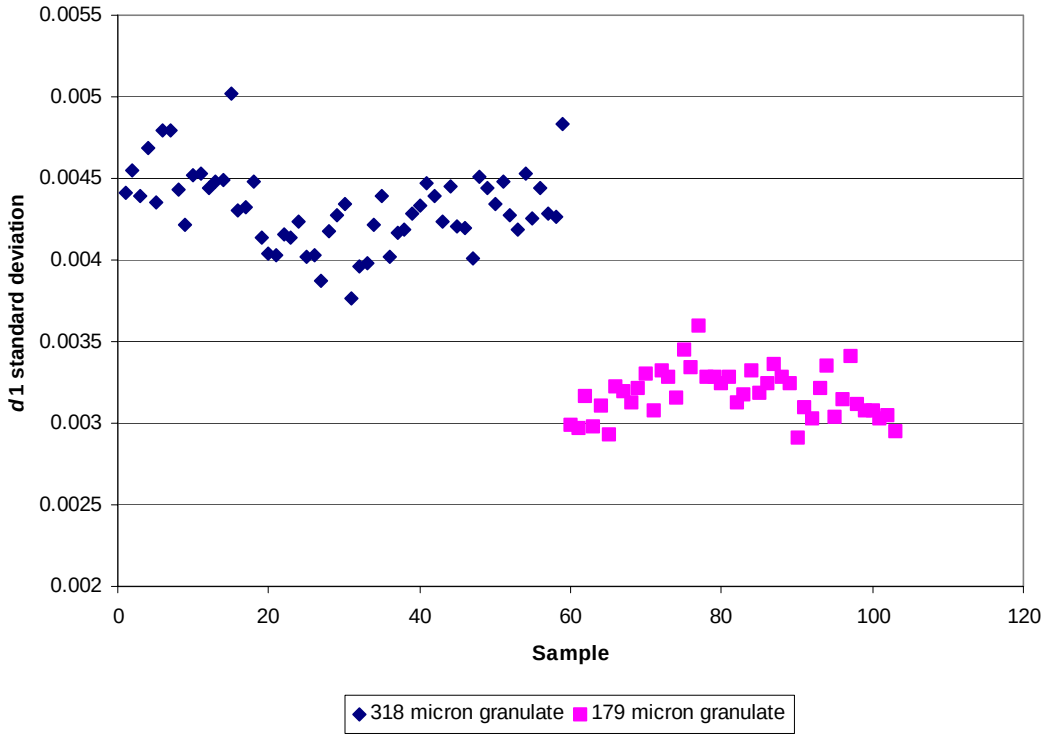

Figure 4.8: Standard deviation of the d1 level signal derived from the original 180-second pressure signals.

stochastic variations. A predictive model such as this may be useful as an online soft sensor for monitoring PSD.

Overall, it can be said that the differences in granulate particle size distribution are correlated with the bubbling frequency (A5 dominant frequency), and the signal amplitudes of the macroscale and microscale levels of the pressure signal. The high-frequency components of the pressure signal are correlated to the particle size distribution of the granulate, and thus microscale phenomenon. Furthermore, these high-frequency features are relatively insensitive to pressure sensor location and superficial velocity.

4.1.4 Classification and Control in Non-Transient Beds

While it is conceded that it is important to maintain an optimal mode of fluidization possible throughout the drying process, determination of these optimal states is difficult. Even when possible, visual observation of the bed is insufficient for detecting subtle differences in hydrodynamics. Measurements capable of quantifying hydrodynamic states, such as high-frequency pressure fluctuation measurements, have the potential for being implemented in an online process-monitoring scheme. If changes in dryer performance could be detected early, operators could intervene and correct undesired conditions before any irreversible product quality problems develop.

PCA provides a powerful multivariate approach to process monitoring applications. Using Hotelling's T^2 and the Q-statistic, PCA models are capable of detecting abnormal differences in process operation due to individual variables as well as the correlation structure between variables. Once a PCA model identifies faulty operation, contribution plots can be constructed to determine the most likely causes of the problem. This diagnostic capability is what differentiates PCA from other process monitoring approaches.

Several PCA models were constructed using data from the dry bed study to determine the suitability of the method for process monitoring. Models were built using a desired set of conditions and data from outside that set of conditions was compared with the model using Hotelling's T^2 and the Q-statistic. The sensitivity of this method to changes in superficial velocity, sensor position, and particle size distribution was evaluated.

To evaluate the use of PCA for monitoring superficial velocity, models were constructed using a data collected at the desired superficial velocity, a single granulate formulation, and single sensor position. The number of components was minimized and only 1 component was retained. Additional components tighten the constraints of a model and thus consider a more defined range of operation. In this investigation, it was desired to see how loose the constraints could be made while still retaining the ability to identify differences in operation.

Hotelling's T^2 proved to be a poor classification statistic for superficial velocity. Greater success was achieved using the *Q*-statistic (Figure 4.9). With the exception of the lowest superficial velocity, it was difficult to identify differences in velocities that were close to one another. However, when the differences in superficial velocity were greater than 0.3 m/s or where one data set contained data from the 1.95 m/s tests, classification was easily achieved. Contribution plots identified the macroscale components, particularly A5 dominant frequency and d5 standard deviation, as being the primary contributors to the faults (Figure 4.10). Based on the findings of the superficial velocity investigation, the combination of these variables would be a cue to adjust the air velocity in an online application.

PCA models were also constructed containing a range of velocities. The results were similar. Models like this could be constructed to be robust against slight changes in superficial velocity, while still providing indications of changes in fluidization regimes.

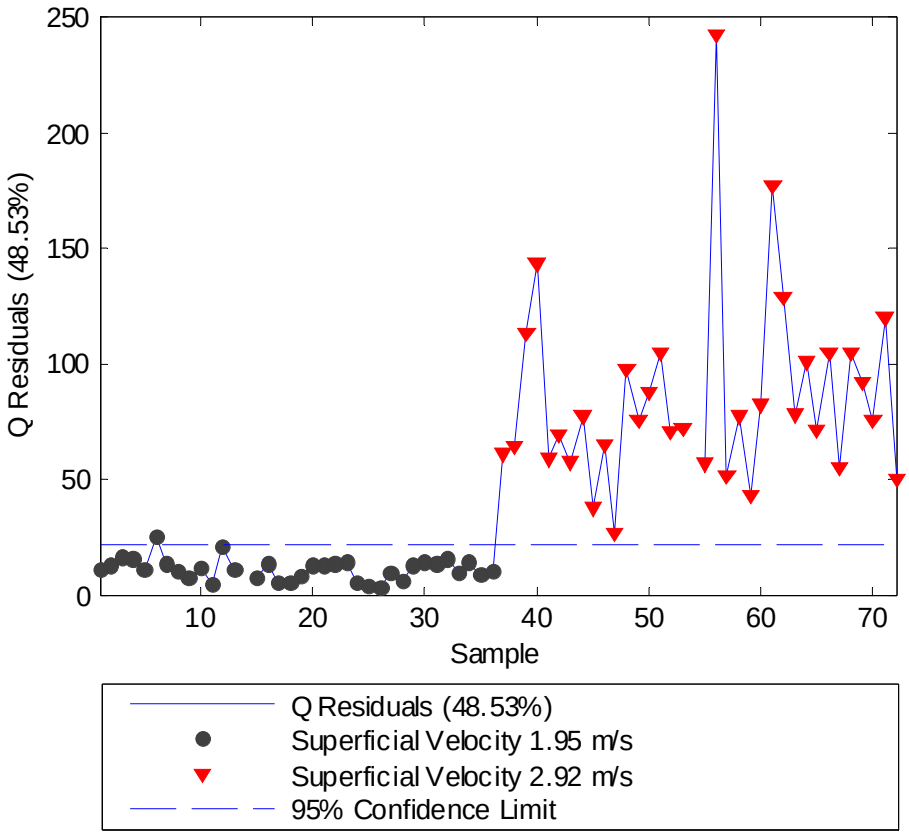

Figure 4.9: *Q*-statistic plot for experiments run with dryer superficial velocity at 1.95 m/s and 2.92 m/s

The procedure was repeated for pressure sensor positions. While it is not expected behaviour for a pressure sensor to change positions, the modeling was carried out to determine whether apparatus-scale effects were detectable using the technique. The *Q*-statistic proved once again to be the better classifier in this evaluation. There were no problems in identifying differences between the data taken at different sensor positions.

Agglomeration, attrition, and entrainment are all phenomenon that can cause problems with product quality. Therefore, being able to detect differences in particle size distribution would be of particular value. The modeling procedures were

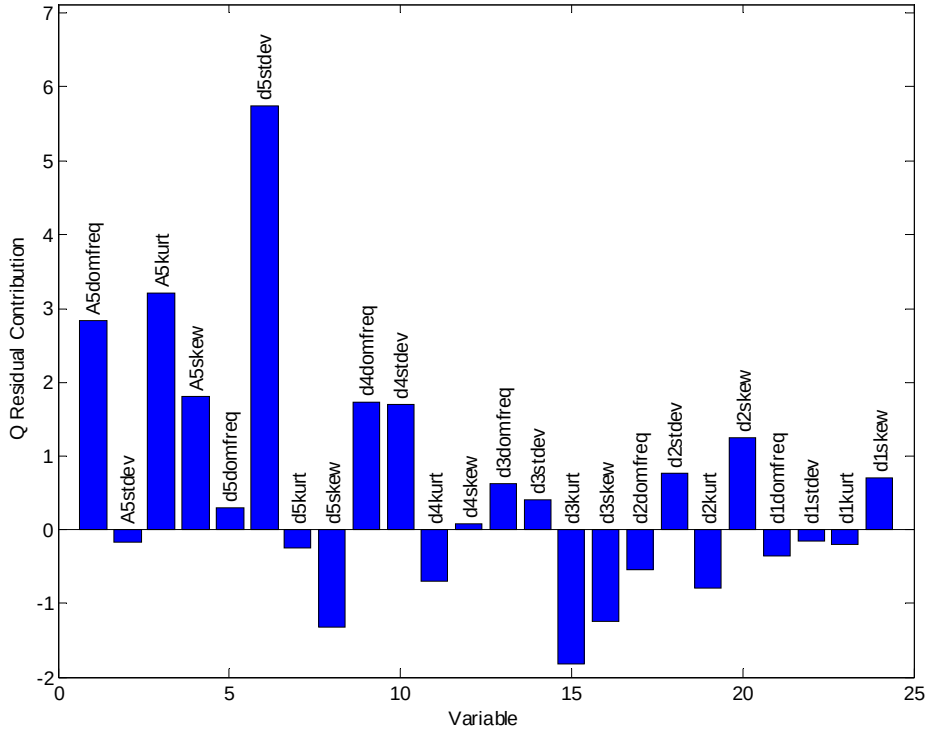

Figure 4.10: *Q*-statistic contribution plot for a point exceeding the 95% confidence limits in Figure 4.9

repeated for particle size distribution. Differences in the particle size distributions were easily detected using the *Q*-statistic. They were also detectable on Hotelling's T^2 when two principal components were retained in the model (Figure 4.11). In this instance, T_{con} plots showed that A5 dominant frequency, d5 standard deviation, d2 standard deviation, and d1 standard deviation were strong contributors for the points that exceeded the 95% confidence limits (Figure 4.12). This combination of variables would clearly identify problems in the particle size distribution.

The use of pressure fluctuations for process monitoring is not a new concept. Chaplin, Pugsley, and Winters (2004a-b) showed that the S-statistic could also detect differences between two batches of granulate with different particle size distributions.

However, in an online application the S-statistic would not be practical because it only shows that there is a difference between the current dryer operation and a chosen reference state. It fails to provide a diagnosis for that difference. For this reason, multiresolutional PCA would be the recommended technique for process monitoring.

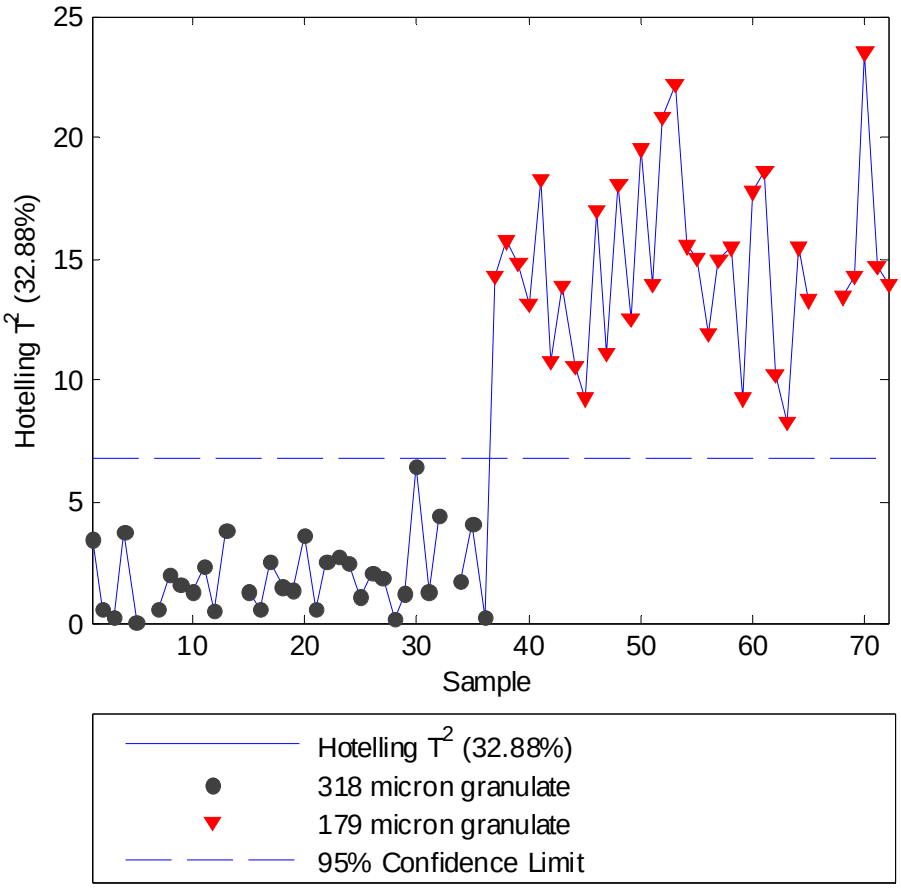

Figure 4.11: Hotelling's T^2 plot from a PCA model containing data from the 179 micron and 318 micron granulate experiments

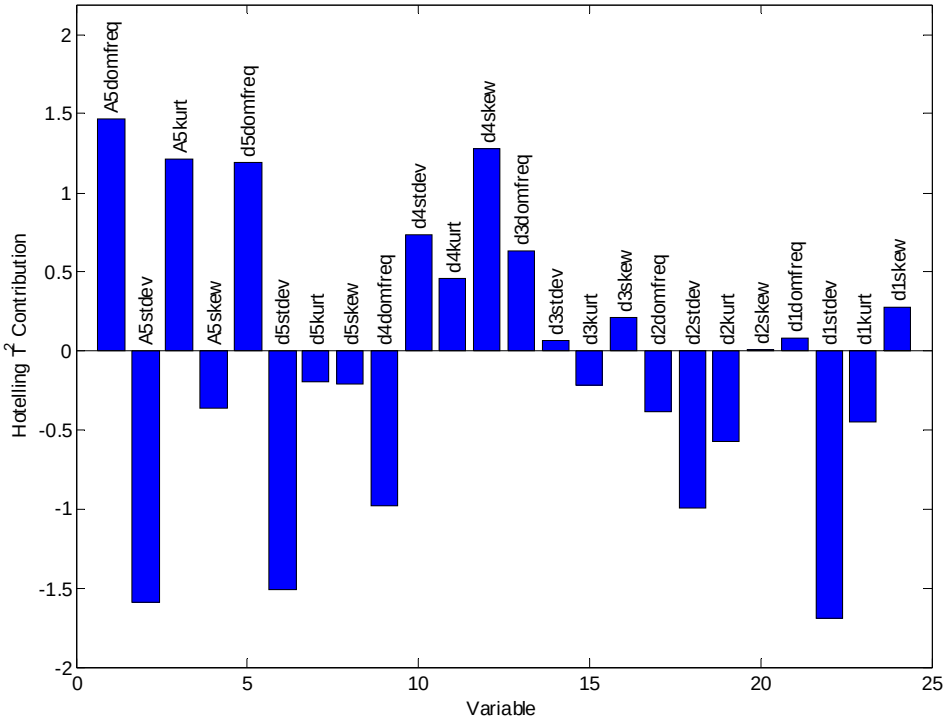

Figure 4.12: Hotelling's T^2 contribution plot for a point exceeding the 95% confidence limits in Figure 4.11

4.2 Batch Drying Study

In the batch drying study, granulate was dried from an initial 33 wt.% moisture to approximately 5 wt.% in the fluidized bed dryer. The inlet temperature, mass of granulate, and sensor position were varied between the experiments. A total of 15 complete data sets were available for the analysis. Eleven of these data sets were part of a statistically designed experimental campaign referred to as the factorial experiments, while the remaining data sets were taken from trial runs conducted before the factorial tests. The initial conditions of each experiment are tabulated in Appendix G.

As in the dry bed study, pressure fluctuations were logged at a sampling rate of 400 Hz. Additionally, periodic measurements were taken of the bed temperature, outlet air temperature, and granulate moisture content. The superficial velocity was maintained at a constant setting throughout the cycle except at the beginning where a higher velocity was required to initially fluidize the bed contents (Figure 4.13). Further details of the experimental procedures can be found in Chaplin, Pugsley, and Winters (2004).

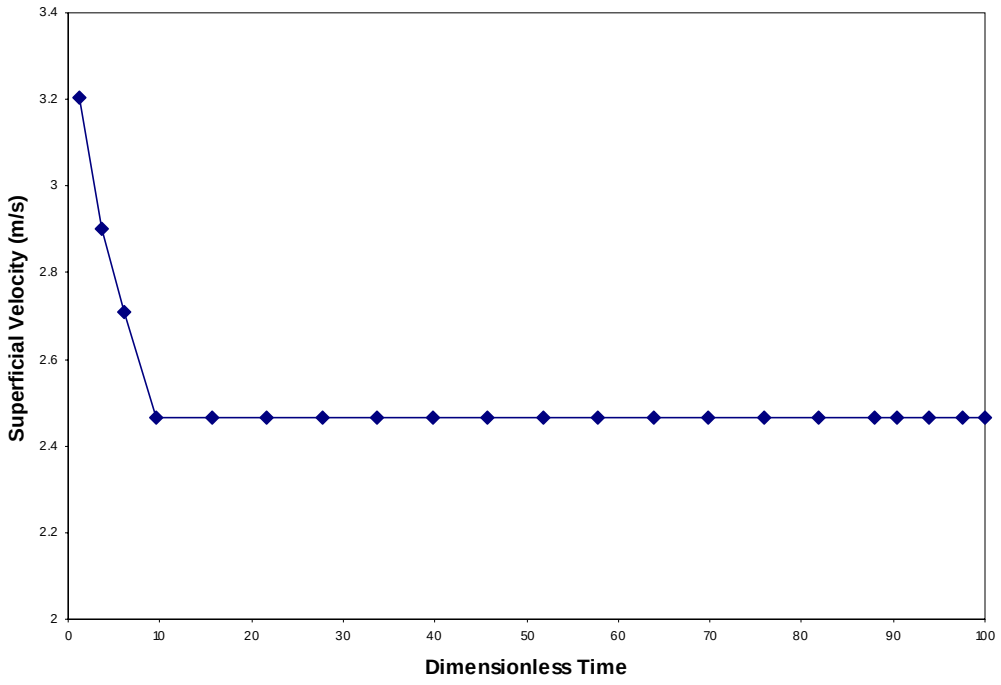

Figure 4.13: Superficial velocity as a function of dimensionless time throughout a batch drying test (Chaplin, 2005).

For reasons previously discussed, the range of moistures was divided into 100 evenly spaced points between 30.00 wt.% and 5.25 wt.% and all other variables were linearly interpolated to correspond to these values. Moisture was used as the

indicator variable because of its nearly linear trajectory and equal start and end values for all batches (Figure 4.14). Wavelet decompositions using a 7-level Daubechies 2 (db2) wavelet were performed for the 30-second intervals of the pressure signals that corresponded to each point along the moisture scale.

The wavelet approximation and detail signal features were summarized using dominant frequencies and the statistics: standard deviation, mean, median, kurtosis, and skewness. A convention was adopted to relate the wavelet approximation and

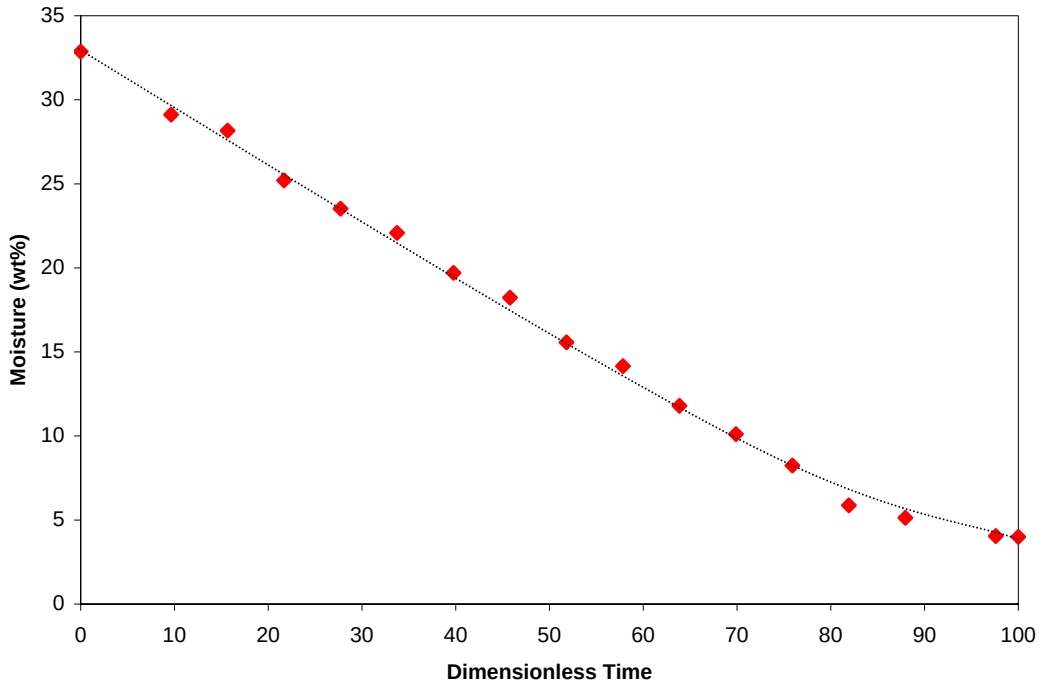

Figure 4.14: Granulate moisture as a function of time throughout a batch drying cycle (Chaplin, 2005).

detail signals to the scales of motion. In the analysis of the batch drying data, the A7, d7, and d6 are referred to as the macroscale components, the d5, d4, and d3 are

referred to as the mesoscale components, and the d2, and d1 scales are referred to as the microscale components. The statistics describing each approximation and detail signal were examined individually to determine whether any significant effects related to the modes of operation or phases in the drying process were visible.

The batch drying process can be broken down into three distinct periods or modes of operation according to the temperature and mass transfer effects governing the process. In the first of these periods, referred to as the initial equilibrium period, the temperatures of the air stream and solids reach an equilibrium state. Next, in the stage known as the constant-rate period, evaporation occurs from the saturated surface of the granulate. During this stage, the evaporation rate is dominated by heat transfer effects and the temperature of the bed remains constant. Finally, in the stage known as the falling-rate period, evaporation occurs from an unsaturated surface and the drying rate eventually becomes dependent on the rate of diffusion of moisture from within the granulate. The beginning of the falling-rate period occurs when the critical moisture content is reached and is marked by a steady increase in the temperature of the bed and the outlet air temperature. All three periods can be identified by examining the outlet air and bed temperatures (Figure 4.15). In these tests, the initial equilibrium period occurred over a longer period of time as a result of the higher initial superficial velocity used to fluidize the bed.

Several of the approximation and detail descriptive statistics appeared to have behaviors related to the drying process. The first of these was the A7 standard deviation. When this variable is plotted as a function of time, the resulting curve very closely resembles the curves seen for the bed and outlet air temperatures. A different

behavior is exhibited by the detail signal standard deviations (at all levels). These standard deviations increase at a steady-rate through time and begin to level off near the end of the cycle. The shape of these trajectories is very similar to the moisture curve and may give an indication of drying rate. Further interpretation of these effects is left to the latent variable analyses.

In the dry bed study, the A5 signal dominant frequency was interpreted as an indicator of bubbling frequency in the bed. The dominant frequencies from the approximation and detail signals were examined for the batch drying data to see if a similar relationship was evident. The d7 dominant frequency was found to be approximately 2.7 Hz throughout the entire drying cycle except where higher bed masses were used. In these cases, the dominant frequency was slightly lower near 2.5 Hz. Recalling the results from the dry bed study, this dominant frequency was within the expected range of bubbling frequencies for the superficial velocities used in the batch drying study and was therefore interpreted as being an indicator of bubbling frequency in the bed. Chaplin's analysis reported a similar dominant frequency of 2.8 Hz throughout the drying cycle except at the start where a dominant frequency of 5 Hz was observed. He attributed the initial frequency to the higher superficial velocity used to fluidize the bed (Chaplin, Pugsley, and Winters, 2004a-b).

Further investigation of the remaining detail signal dominant frequencies revealed that the d6 dominant frequency was close to 5 Hz at the start of the drying cycle and dropped to approximately 3.6 Hz at a point that roughly coincides with the end of the initial equilibrium region. It is not believed that this frequency is due to an initial higher bubbling frequency. The d6 dominant frequency more likely represents

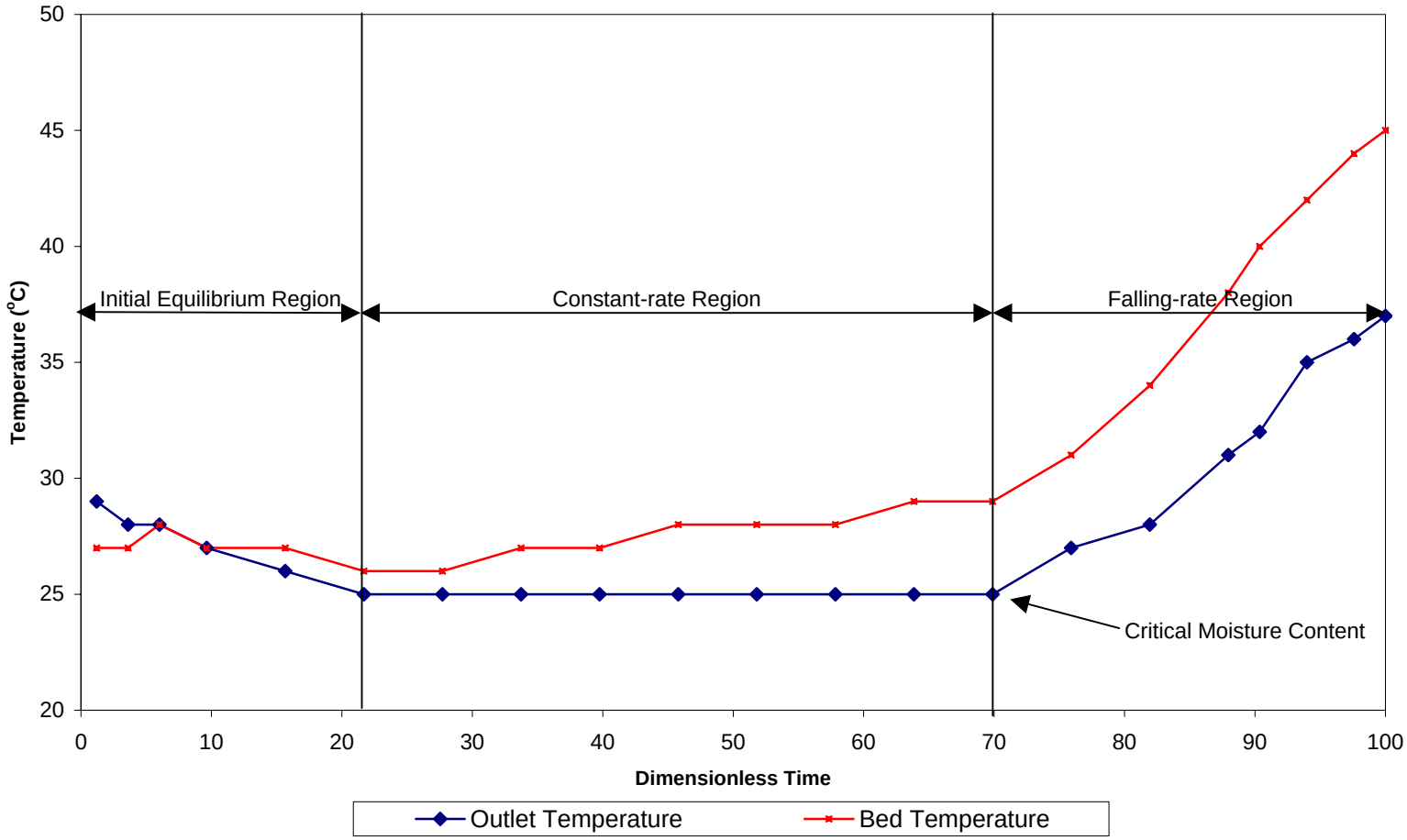

Figure 4.15: Outlet air and bed temperatures as a function of time throughout the batch drying process. The initial equilibrium, constant-rate, and falling-rate regions are identified (Chaplin, 2005).

another behavior that is amplified by the higher moisture content at the beginning of the cycle (ie. amplification of the formation or coalescence of bubbles). Traditional signal processing techniques such as FFT are subject to aliasing effects where multiple contributions compose the signal. This would explain Chaplin's observation of a higher dominant frequency at the beginning of the cycle. Further investigation using an imaging technique such as electrical capacitance tomography should be carried out to verify whether or not a shift in the bubbling frequency occurs at the beginning of the drying cycle.

There was a shift in dominant frequency observed in the d5 signal as well. However, in this case the dominant frequency increased from 8 Hz to 10 Hz at approximately 18 wt.% moisture (halfway through the drying cycle). The reason for this transition is unclear. The batch drying process involves a much more complex range of conditions than those observed in the dry bed study. In addition to the hydrodynamic characteristics of the process, the mass transfer effects appear to be visible in the pressure signal as well. Multiresolutional analysis using wavelets has the ability to detect contributions in the pressure signal that traditional signal processing techniques cannot. Data analysis techniques such as the latent variable analysis techniques used in this study can be used to correlate these signal components to other measured and observed behaviors.

MPCA, PCA, and PLS analyses were performed on the batch drying data with a number of objectives in mind. The first objective was to determine whether key states or phenomenon in the drying process were visible in model parameters.

Secondly, an attempt was made to derive a suitable method for monitoring the hydrodynamic conditions in the fluid bed dryer throughout the drying cycle. Next, models were constructed to study the effects of sensor position, inlet temperature, and bed mass. Finally, the batch drying data was used to construct a soft sensor to predict moisture content of the granulate throughout the drying process.

4.2.1 MPCA Analysis

In MPCA analyses, there are three unique ways to unfold the data matrices, two of which are meaningful. For constructing models that are to be used for classification or statistical control, the batch-wise unfolding of the data matrix is typically a good choice. When the matrix is unfolded in this way the matrices are aligned side by side at each time interval. Thus, the time trajectory is effectively removed and the principal components describe the variation in the data at each respective time step. However, if one is interested in studying the dynamic nature of a process and the correlations related to the process trajectories, time-wise unfolding of the data is recommended.

When a data set is unfolded time-wise, the data sets are aligned side by side according to batch. In this way, the scores plots essentially show the average projection of each sample or time interval about the principal components. A data set was assembled using the data from the 11 factorial experiments and then unfolded batch-wise to examine the drying process. Each of the 11 data sets consisted of 54 variables at 100 moisture intervals. This resulted in an unfolded matrix dimension of 100 x (54x11) = 100 x 594.

A MPCA model was calculated from the data and a decision was made to retain 3 components that accounted for a total 42% of the variance in the data. While further components could have been retained, they did not add significantly to the model. The scores plots for each of the 3 principal components showed deterministic behavior through the drying cycle. Each component was investigated separately.

Principal component 1 accounted for 32% of the total data variance or approximately 75% of the model variance. The points in score space had a very familiar trajectory (Figure 4.16). The shape of the curve was nearly identical to the moisture curve as seen in Figure 4.14. Inspection of the loadings plot for principal component 1 verified this relationship (Appendix H).

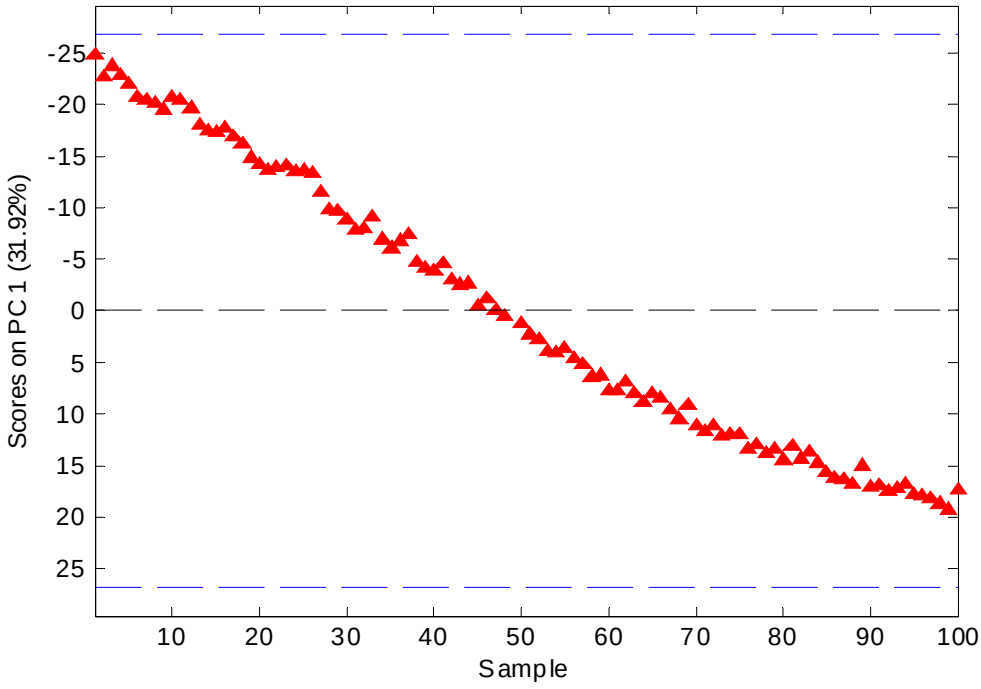

Figure 4.16: Scores plot for principal component 1 from a MPCA model constructed from the factorial data sets and unfolded time-wise.

In a MPCA model composed from a data set with the dimensions 100 x 594, there will be 594 contributing variables to each loadings plot or principal component. Practically speaking, this is very difficult to interpret. However, a quick inspection of the plot reveals a repeating and constant pattern throughout. In order to extract meaningful information from the plot, the average of each variable was taken. This reduced the plot down to the 54 original variables describing the average contribution throughout the cycle. Further analysis of this reduced loadings plot for principal component 1 showed some other interesting, yet intuitive, relationships. For example, time was negatively correlated with moisture but positively correlated with bed temperature and moisture was negatively correlated with bed temperature.

Moisture and the standard deviations from the approximation and detail signals at all levels were found to be the strongest contributing variables to principal component 1. Kurtosis and skewness also showed significant loadings at some levels. Kurtosis is another measure of signal amplitude and skewness is a result of transient behavior. Overall, it can be said that the effect of moisture dominates the drying cycle. This agrees with the findings of Chaplin, Pugsley, and Winters (2004a-b).

Principal component 2 of the MPCA model accounted for 7.6% of the total data variance or approximately 18% of the model variance. The scores plot for principal component 2 showed points initially projecting positively, progressing to a negative projection half way through the cycle, and then returning to a positive projection (Figure 4.17). The reasons for this type of transition were not so obvious as for the first principal component.

Examination of the loadings plot showed that, unlike principal component 1, moisture was not a strong contributor to principal component 2. Rather the bed and outlet temperatures were the strongest contributors along with some of the approximation signal statistics. This shows that principal component 2 is representative of the macroscopic temperature effects throughout the drying process. To clarify, this shows a progression towards and away from the average temperature that occurs over the course of the drying cycle (ie. constant-rate period).

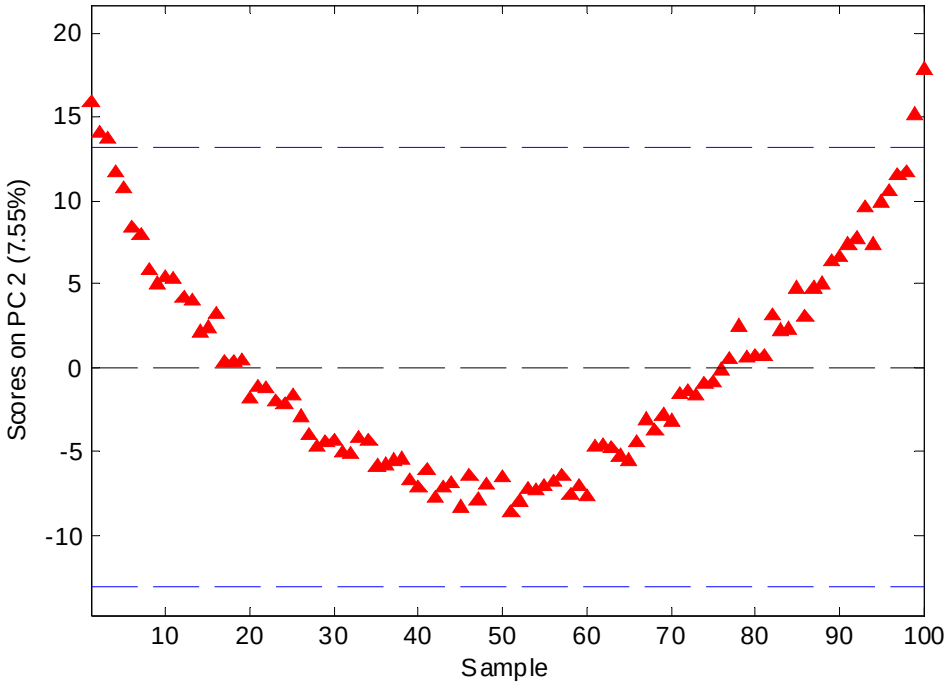

Figure 4.17: Scores plot for principal component 2 from a MPCA model constructed from the factorial data sets and unfolded time-wise.

The third principal component in the model only accounted for 2.8% of the total data variance and 7% of the MPCA model variance. However, the component was retained because the scores plot showed some very interesting behaviour (Figure 4.18). Three distinct regions corresponding to the initial equilibrium region, constant-

rate region, and the falling-rate region can be identified. The loadings plot show the significant contributions to the principal component are derived mainly from the superficial velocity, inlet air temperature, outlet gas temperature, and bed temperatures. This third component is believed to be representative of the mode of mass transfer throughout the drying process.

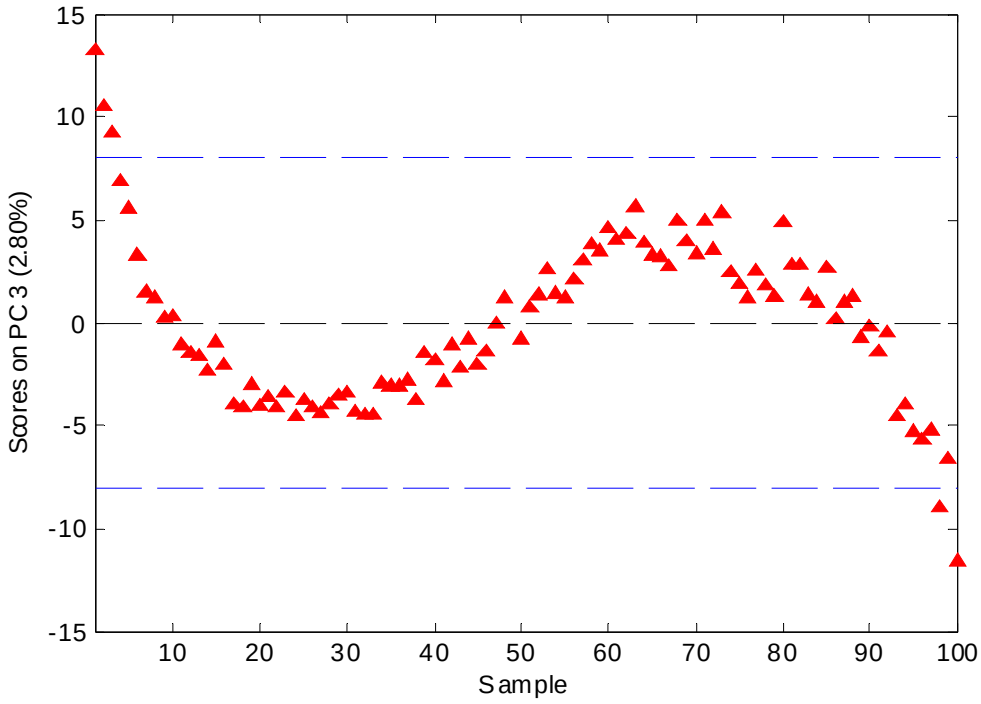

Figure 4.18: Scores plot for principal component 3 from a MPCA model constructed from the factorial datasets and unfolded time-wise.

In the MPCA model, the heat and mass transfer effects dominate the overall process behavior. Information regarding the hydrodynamic behaviors or transitions in fluidization regimes that occur throughout the drying process is difficult to distinguish next to these moisture and temperature effects. As in the dry bed study, the MPCA model was reconstructed using only the pressure signal elements to verify the structure of the model was not being dominated by the other variables. However,

when reconstructed using only the approximation and detail signals, the scores plots retained the same characteristic projection patterns observed when using the complete data set.

In order to better explore each phase of the drying cycle, the data was split into 3 groups that reflected the initial equilibrium region, the constant-rate region, and the falling-rate region. Models were constructed from each of these groups to investigate the behaviors for each period.

4.2.1.1 Initial Equilibrium Region

A MPCA model was constructed to examine the initial equilibrium region. The model was constructed using the data from the baseline tests (#1, #8, and #11) of the factorial data runs. This was intended to minimize the effects of inlet temperature, bed mass, and sensor position. The initial equilibrium region was defined as the period where the moisture in the bed was 30.00 wt.% to 24.75 wt.%. Two principal components accounting for 32.47% of the variance were retained in the model (Appendix I).

As in the overall MPCA model, the dominant principal component in this model showed a very strong correlation between moisture and the standard deviations of the detail signals. A strong loading of the d6 dominant frequency was also observed in this component. It appears as though this frequency is positively correlated with moisture, that is, as moisture decreases the d6 dominant frequency decreases. However, this correlation cannot be shown to be independent of other variables.

The second principal component showed a strong correlation between changes in velocity and the macroscale behaviour in the bed. In particular, the product temperature, A7 standard deviation, and velocity are positively correlated (Appendix I). When the initial fluidizing superficial velocity is decreased, the bed and outlet air temperatures slightly decrease. This validates the assumption that the A7 standard deviation provides information regarding the bed temperature.

4.2.1.2 Constant-Rate Region

A MPCA model was constructed to examine the constant-rate region using the baseline data used to examine the initial equilibrium region. The constant-rate region was defined as the period where the moisture in the bed was 24.50 wt.% to 13.00 wt.%. Two principal components accounting for 33.35% of the variance were retained in the model (Appendix I).

The dominant principal component in the constant-rate region once again showed a very strong correlation between moisture and the detail signal standard deviations. This model does not contribute to a further understanding of the process beyond what was previously observed.

4.2.1.3 Falling-Rate Region

A MPCA model was constructed to examine the falling-rate region using the same data as in the models for the initial equilibrium and constant-rate regions. The falling-rate region was defined as the period where the moisture in the bed was 12.75 wt.% to 5.25 wt.%. Two principal components accounting for 27.27% of the variance were retained in the model (Appendix I).

The first principal component from the model in the falling-rate region was similar to the first component in the other two models. A strong correlation between the detail signal standard deviations is seen. The next principal component was weighted nearly as heavy as the first and shows a strong correlation between the outlet temperature, bed temperature, A7 standard deviation, and time. These variables are all negatively correlated with moisture. Once again, this does not help to further understand hydrodynamic behaviours in the process but does illustrate the dominant effect of moisture throughout the drying process.

4.2.2 Classification and Control of the Drying Process

The ability to monitor the hydrodynamic state throughout the drying process would help to ensure optimal product quality and minimize economic losses. MPCA is an obvious candidate for building a statistical monitoring model of the drying process. A batch-wise unfolded data matrix consisting of the batch drying data sets would not be influenced by differences in the trajectories between batches and provide simple process monitoring plots. However, this approach proved to be cumbersome due to the large dimension of the unfolded dataset (ie. 11 x (54 x 100) = 11 x 5400). Even if only 10% of the observations were used in the models, the loadings plots would be too complicated to use for fault diagnosis.

An alternative approach to the MPCA modeling was evaluated. A PCA model was constructed for each phase of the drying process: the initial equilibrium period, the constant rate period, and the falling-rate period. The reasoning behind this approach was that, while transient, the hydrodynamic behaviors in each individual phase should be similar throughout that region. If successful, the largest drawback to

this approach would be that three sets of monitoring charts would be required over the course of the drying cycle. In theory, constructing a model at each weight % moisture would be better. However, due to the large number of models that would be required, dividing the data according to drying region was considered more practical.

The models were constructed using 10 of the 11 factorial data sets for each of the following moisture ranges: 30.00 wt.% to 24.75 wt.%, 24.50 wt.% to 13.00 wt.%, and 12.75 wt.% to 5.25 wt.%. The number of components were chosen for each model based on the usual criteria. Once the models were constructed, the factorial data set that was excluded in the formulation of the three models was projected against each model. Hotelling's T^2 and the Q-statistic plots for each model were used as the measure of performance. In the absence of data that contained known faults or operating difficulties, this was considered the best alternative for validating the models.

Hotelling's T^2 statistic and the Q-residual statistic showed the same basic characteristics. The data set projected within the 95% confidence limits where the moisture content was similar to the range used to construct each model (Figure 4.19). Where points exceeded the 95% confidence limits, contribution plots were constructed to determine which variables were responsible. In almost all cases the moisture and standard deviations were strongly weighted in the contribution plots. However, other more interesting contributing factors were noted as well. For example, points that fell outside the 95% confidence interval for the initial equilibrium model showed a strong contribution from the d6 dominant frequency. This provides some confidence that this approach is capable of detecting differences

in hydrodynamic behaviors as well as mass transfer effects throughout the drying cycle.

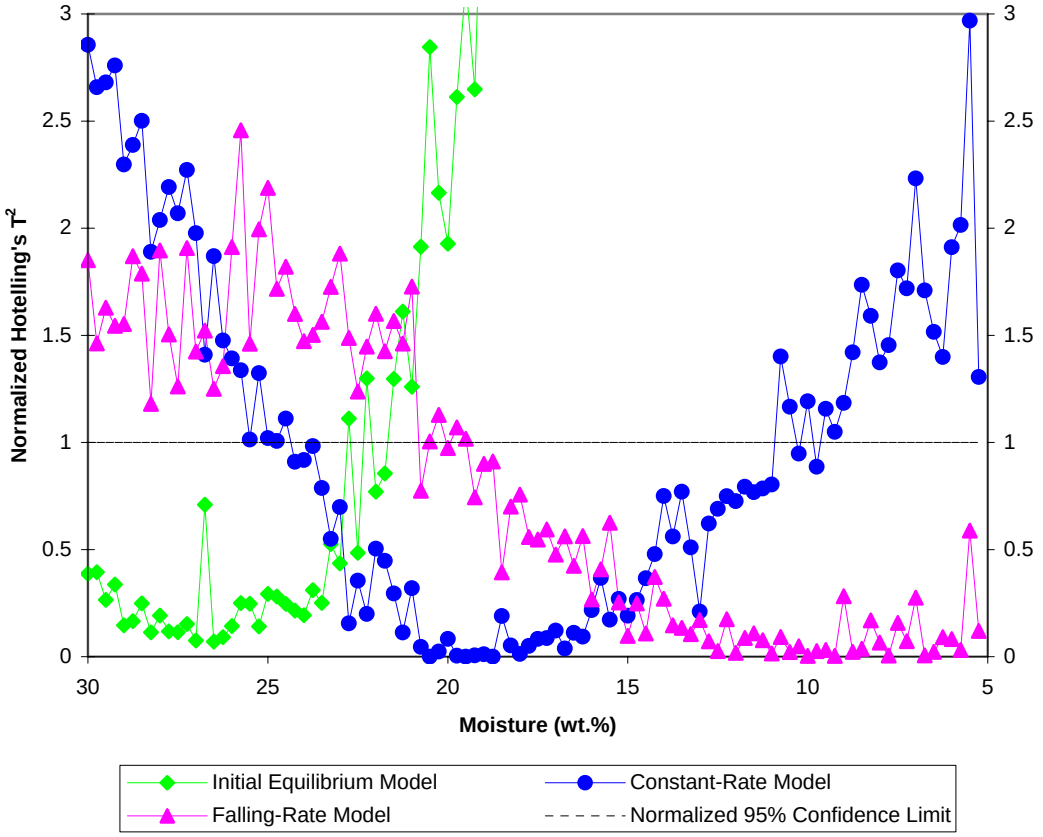

Figure 4.19: Normalized Hotelling's T^2 statistic showing new data projected on each of the three models.

The sensitivity of this approach can be improved with increased volumes of higher quality data. Typically, data used in this type of statistical process monitoring model should come from the normal mode of operation and contain only the batch-to-batch stochastic variations that occur in the process. In this study, very little data was available and there were systematic variances in the process parameters that would not normally occur in a production facility. In the batch drying study, most variables were recorded to the nearest minute. Immediately this decreases the accuracy of the

model. Where possible, measurements used in the model should be automated at a frequency that is suited to the frequency response of the instrument.

In general, this approach was found suitable to monitoring the hydrodynamic conditions in the dryer throughout the drying process. Larger volumes of higher quality data and possibly data sets with known faults are required to determine the ability of this approach to detect upset process conditions.

4.2.3 Effect of Inlet Temperature

PLS modeling is usually the best latent variable modeling approach for determining the effect of a single variable. This is particularly true where there is a strong transition in the state of the system. These transient phenomena can overshadow other more subtle effects in PCA models. The overall effect of inlet temperature on the drying process was examined by constructing a PLS model using the factorial data. Inlet temperature was taken as the *Y*-block while all other remaining variables were taken as the *X*-block. A PLS model with 4 latent variables describing 95.37% of the *Y*-block variance was selected based on cross validation and the eigenvalue curve.

The overall correlation structure of this model is best summarized with the regression vector (Appendix J). The relationships observed in the regression vector loadings plot are intuitive. The inlet temperature is negatively correlated with batch time and positively correlated with outlet temperature and bed temperature. Next to these relationships, the remaining correlations were considered insignificant.

The approximation and detail signals were not strongly correlated with the inlet temperature or cycle drying time. This was interpreted to mean that the same

hydrodynamic behaviours were occurring between batches regardless of inlet temperature. As such, the use of moisture as an indicator variable in previous MPCA and PCA modeling attempts was deemed acceptable.

4.2.4 Effect of Initial Mass

A PLS model was constructed to examine the effect of initial bed mass on the batch drying process. A similar approach was taken as with inlet temperature. Initial bed mass was used for the **Y**-block data and the remaining variables from the factorial data sets were taken as the **X**-block data. Initial mass did not have as much of an effect on the runs as inlet temperature. Three latent variables describing 42.56% of the **Y**-block variance were retained in the model.

Analysis of the regression vector revealed some interesting relationships (Appendix J). As expected, the increased bed masses were slightly correlated with increased cycle times, higher moistures, and lower outlet air temperatures. However, what is more interesting is the effect on the d7 dominant frequency, d6 dominant frequency, and mesoscale (d5 and d4) variables. The regression vector suggested that at higher initial bed masses, the d7 dominant frequencies were lower and the d6 dominant frequencies were higher. These relationships were verified in the original data sets. The lower frequencies observed at higher bed masses in the d7 dominant frequency agrees with the observations by Chaplin, Pugsley, and Winters (2004a-b). The reason for the contributions observed in the mesoscale variables is unclear.

4.2.5 Effect of Sensor Position

The effect of pressure sensor position was evaluated in the dry bed study and again here using the data collected in the factorial experiments. A 5-latent variable

PLS model was constructed that accounted for 66.01% of the **Y**-block variance. The relationships observed here were consistent with the findings in the dry bed study.

The standard deviations of the detail signals were affected to a larger degree than any of the other variables, with the exception of moisture, which once again leads to the conclusion that sensor position affects signal amplitude (Appendix J). The moisture variable showed up as a contributing factor in the correlation because it is a dominant variable in the data and highly correlated to standard deviation. As in the dry bed study, the macroscale standard deviations were positively correlated with higher sensor positions while the higher frequency standard deviations were negatively correlated with higher sensor positions. Unlike the dry bed study, there was an indication that sensor position affects the signal amplitude of the microscale contributions. However, this is not conclusive as the microscale contributions may have been due to the presence of the moisture variable rather than sensor position.

Based on the results of this PLS model, the recommended placement of the pressure sensor would be the lower mounting positions. This would help increase the resolution of the higher frequency, more difficult to distinguish, interactions. Lower frequency signal components were easily obtained at all mounting positions.

4.2.6 Prediction of Moisture Content

In the industrial batch drying of pharmaceutical granulate, the process is controlled by monitoring the outlet and bed temperatures. When a limiting value of either of these temperatures is reached, the batch is complete. The problem with this approach is that it does not consider the overall state of the bed. An ideal measure of product moisture would consider the distribution of conditions throughout the dryer

as well as allow for a more accurate indication of the drying cycle completion. With this goal in mind, ten data sets from the batch drying study were used to construct PLS models to predict moisture throughout the drying cycle.

The number of components to retain in each model was questionable and therefore two models were constructed from the same data and compared. The first model retained 3 latent variables that accounted for 94.17% of the **Y**-block variance while the other retained 10 latent variables that accounted for 95.07%. Predictions from each of the two models are shown in Figure 4.20.

The confidence intervals gave the average error in the 3 LV model as $\pm$ 1.40 % moisture. This error decreased to $\pm$ 1.31 % moisture in the 10 LV model. This is not a large enough gain in resolution to justify keeping the extra latent variables in the model. In general, the resolution of the predictions suffered from both the lack of data and non-stochastic variations in the data (ie. changes in sensor position, inlet temperature, and initial bed mass). These variations created large bias errors in some predictions where the conditions were extremes. This was the case where the lowest pressure sensor position, lowest mass, and the lowest inlet temperature appeared in the same model. As in the PCA statistic process control application, the resolution of these predictions can most likely be improved by creating the model from a larger number of batches where the only variance is due to normal stochastic errors that will be experienced through the course of regular operation.

The variables contributing the most to the moisture predictions can be determined by examining the dominant weight vector and the regression vector

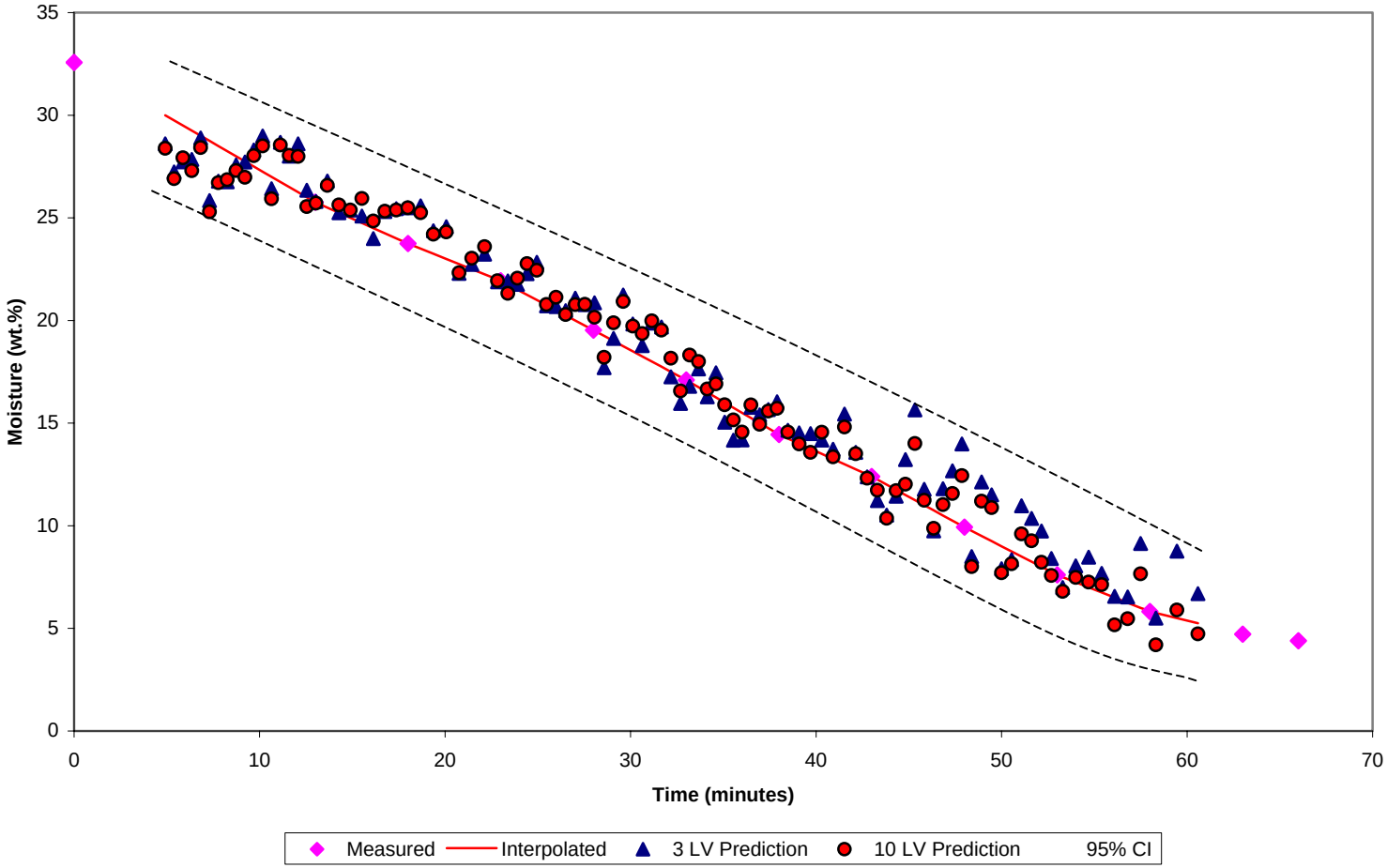

Figure 4.20: PLS model predictions of moisture throughout the drying cycle. Model built from 10 batch drying data sets. Predicted moisture is for data set not used in model.

(Appendix K). The regression vector shows some particularly interesting relationships. As previously suggested, moisture affects the standard deviation contributions at all levels. However, in this plot it would appear as though moisture affects the amplitude of the microscale components slightly more than the mesoscale components and the mesoscale components slightly more than the macroscale components. This may be an indication that the mass transfer of moisture affects the behavior of the particle scale and cluster scale interactions to a larger degree than the macroscale phenomenon. There is also evidence here that changes in the d6 dominant frequency throughout the drying cycle are related to moisture effects.

The PLS modeling approach to moisture prediction is worth further development for several reasons. Firstly, the models consider the hydrodynamic conditions within the fluid bed dryer and, as a result they provide better information regarding the state of the process. Secondly, there are no large instrumentation costs. The model is based on data from a relatively inexpensive piezoelectric sensor and pre-existing measurements. Finally, the models take into account the temperatures that are currently used for control. Therefore, the quality of those measurements is always being scrutinized. Should a thermocouple begin to fail, the model will detect and diagnose the failure (using Hotelling's T^2 and the Q-statistic).

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

PCA, MPCA, and PLS modeling techniques were applied to previously generated experimental data collected from a fluidized bed dryer containing pharmaceutical granulate. The data was collected during both steady state and transient operation of the dryer. In the experiments, the hydrodynamic conditions in the dryer were quantified using high frequency pressure fluctuation measurements. These pressure fluctuation signals were decomposed using discrete wavelet transformations in order to extract scale specific information prior to the modeling. In general, this was found to be an excellent approach to the modeling of the fluid bed dryer system. Several conclusions can be made based on the current work. These conclusions are presented separately for the dry bed study and the batch drying study.

5.1.1 Dry Bed Study

In Chaplin's dry bed study (Chaplin, Pugsley, and Winters, 2004a-b), the variables under investigation were superficial velocity, pressure sensor position, and granulate particle size distribution. The wavelet decomposition of the pressure signals collected in these experiments proved to extract behaviors related to specific scales of motion. These behaviors could be correlated to the variables under investigation using PCA and PLS models. From the models generated specifically to

evaluate the effects of superficial velocity, sensor position, and granulate particle size distribution, the following conclusions can be made:

- Information regarding the pressure sensor position, superficial velocity, and granulate particle size distribution were found to be encoded into the pressure signals.
- PCA scores plots showed deterministic variations due to sensor position, superficial velocity, and granulate particle size distribution. It was determined that the PCA and PLS models constructed solely from the wavelet decomposed pressure signals could describe these deterministic variations.
- PCA models investigating the effects of superficial velocity showed potential for identifying differences between fluidization regimes. Data collected at the lowest superficial velocity of 1.95 m/s projected differently in scores plots than points generated at higher velocities. At 1.95 m/s, the behavior of the bed can be classified as being on the boundary between the bubbling and turbulent regimes.
- Over the range of velocities tested, increases in the superficial velocity are correlated to increases in the bubbling frequency. These increases in velocity appear to affect the macroscale and mesoscale behaviors but have little to no effect on the particle scale interactions of the microscale variables.
- Changing the pressure sensor position affects the amplitude of the pressure signal. This is seen as an increase in the signal standard deviation. Higher sensor positions are correlated to increases in the macroscale signal features and decreases in the mesoscale signal features.

- Based on loads plots from PCA and PLS models, differences in granulate particle size distribution affect both the macroscale and microscale behaviors in the fluid bed dryer.
- The two granulate formulations tested could be distinguished solely by the standard deviation of the high frequency *d*1 detail signal. This provides evidence to support a relationship between the high frequency components of the pressure fluctuation signal and the microscale behavior in the fluid bed dryer.
- The A5 dominant frequency was found to be equal to the dominant frequency of the original pressure signal and is, thus, an indicator of bubbling frequency in the fluid bed dryer. Therefore it can be said that the 5^{th} level approximation signal provides information regarding the macroscopic behavior of the bed.

In addition to the evaluation of the test variables, PCA was found to have excellent potential for the statistical process monitoring of the hydrodynamic conditions in a fluidized bed dryer containing dry pharmaceutical granulate. Using the *Q*-statistic and Hotelling's T^2 statistic, differences between granulate particle size distributions, superficial velocity, and sensor positions could be detected. Contribution plots were found to be capable of diagnosing the causes of inconsistent or 'faulty' operation.

5.1.2 Batch Drying Study

In Chaplin's batch drying study, the initial bed mass, pressure sensor position, and inlet temperatures were varied between runs. In terms of typical latent variable analyses, the actual number of runs performed was relatively low. However the

approach was capable of clearly identifying correlations between the test variables and the drying process. Models constructed specifically to examine these variables yielded the following conclusions:

- Differences in inlet air temperatures to the fluid bed dryer are correlated to the drying time, outlet air temperature, and bed temperatures. Additionally, there appears to be very little correlation between this inlet temperature and the pressure signal elements. This suggests the trajectory the process follows in multivariate space does not change at different temperatures and is more likely a function of granulate moisture content.
- Increased initial bed masses are correlated to increased overall drying cycle times, higher moistures, and lower outlet air temperatures. Additionally, increased bed masses are correlated to decreased d7 detail signal dominant frequencies, and increased d6 detail signal dominant frequencies. The explanation for the latter relationships is unclear but, due to the level of the detail signals, are most likely related to macroscale effects.
- The effect of sensor position in the batch drying study was found to be consistent with the findings of the dry bed study. Moving the pressure sensor affects the signal amplitude.

One of the main purposes of this study was to model the batch drying process by combining wavelet decompostion with latent variable analysis. However, several observations were apparent from examining the wavelet decomposed pressure signals alone:

- The low frequency d7 detail signal dominant frequency is an indicator of bubbling frequency in the drying bed. This frequency remained approximately constant throughout the drying process at 2.7 Hz except at higher initial bed masses where it was found to be slightly lower.
- The dominant frequency of the d6 detail signal was found to be approximately 5 Hz during the initial equilibrating period of the drying process before dropping to 3.6 Hz. It is speculated this shift in behavior is related to amplified effects of the formation or coalescence of bubbles in the bed due to the initially higher moisture content.
- There is a shift in the d5 signal dominant frequency from 8 Hz to 10 Hz that occurs approximately half way through the drying cycle. The reason behind this behavior is not clear.

MPCA models constructed to examine the batch drying process were successful in identifying key operating conditions and the dominant variables controlling the process in the model parameters. The following conclusions can be derived from the MPCA modeling study:

- Moisture dominants all aspects of the drying process. This relationship is evident in both the scores and loadings plots for the MPCA model of the drying process. The scores plot for the dominant principal component closely resembles the moisture curve. The loadings plot for the same principal component shows a strong correlation between the moisture content of the bed and the detail signal standard deviations at all levels.

- Principal component 2 in the MPCA model of the batch drying process is dominated by transient temperature effects throughout the cycle. The scores plot for this component can be interpreted as showing temperatures approaching and then diverging from some average temperature throughout the process. This average temperature occurs throughout the constant-rate period.
- The third principal component in the MPCA model of the drying process shows periods of operation that correspond to each of the 3 phases of the drying process: the initial equilibrium period, the constant-rate period, and the falling-rate period. These periods of operation can be identified in the scores plot.
- The hydrodynamic behaviors of the bed are overshadowed by heat and mass transfer effects in the MPCA model parameters.
- The A7 approximation signal standard deviation is correlated to the bed and outlet air temperatures of the fluid bed dryer. Therefore, the A7 approximation signal is related to the temperature effects throughout the drying process and can be considered an indicator of macroscopic behavior of the system.
- The detail signal amplitudes are highly correlated to the moisture content of the granulate throughout the batch drying process. The individual trajectories of these standard deviations is comparable to the trajectory of the moisture versus time curve.

The final objective of this study was to evaluate the applicability of PCA for the process monitoring of the fluidized bed drying process. This study was extended beyond the original scope to consider the ability of PLS to create inferential models or soft sensors capable of predicting bed moisture content throughout the drying process. Conclusions based on these studies can be summarized as:

- PCA can be used for multivariate statistical process control of the fluidized bed drying process by constructing a separate model for each of the 3 distinct periods that occur during the drying process. These models were found to be capable of detecting the differences between modes of operation and provided some evidence that they are able to detect differences in the hydrodynamic behaviors between regions. Increased volumes of higher quality data are required to properly evaluate the sensitivity of this approach.
- PLS can be used to construct inferential moisture models capable of predicting granulate moisture content throughout the drying cycle. Better resolution of this type of soft sensor may be obtained with larger volumes of higher quality data.

5.2 Recommendations for Future Work

The scope of the current study was limited to the pre-existing data generated by Chaplin. The results of this analysis illustrate the potential of the multiresolutional PCA and PLS approach for modeling the fluidized bed drying process. These methods are not limited to the drying process and can be easily extended to other fluidized bed systems as well. Multiresolutional decomposition of high frequency pressure fluctuations helps extract scale specific hydrodynamic behaviors and latent

variable analysis techniques help to correlate these behaviors to other system variables and events. In the future, more information can be extracted regarding the dynamics in fluid bed systems with data generated specifically for this type of analysis approach. However, this is a general long-term objective. Recommendations for future studies related more directly to the current work are:

- Further investigation should be carried out to determine the reason for the unusual behavior of the A5 approximation signal mean and median statistics. Experiments that monitor the behavior of a dry bed over extended periods of time would be best suited to this task. This effect was amplified at lower superficial velocities.
- Combining PCA with wavelet decomposed pressure signals showed excellent potential for characterizing differences between fluidization regimes. This should be investigated further by conducting experiments over a wider range of superficial velocities. In particular, data should be collected at superficial velocities lower than 1.9 m/s and compared to behavior at superficial velocities above 2.1 m/s.
- Lower pressure sensor positions should be favored over higher mounting positions. Macroscale behaviors are easily distinguished at all positions whereas higher frequency behaviors can benefit from increased resolution. In future experimental work, only a single mounting position should be used.
- An experimental campaign should be conducted to generate data specifically for evaluating the sensitivity of PCA models for monitoring the hydrodynamic conditions in the fluid bed dryer. The experiments intended to generate data

for the models should be run under the similar conditions with only batch-to-batch stochastic variations that would be seen in a typical industrial process. Validation runs should be conducted in which faulty operation is induced or increasing variances are introduced into the operating conditions.

- An investigation into the cause behind the shifting dominant frequencies in the d6 and d5 detail signals should be conducted. It is suspected that the initially higher dominant frequency in the d6 detail signal is not the result of an increased bubbling frequency at higher moisture contents. This study may best be accomplished using an imaging technique such as ECT or alternatively by using a fiber optic voidage probe in conjunction with the pressure fluctuation measurements.
- The ability of PLS to construct inferential models or soft sensors of bed moisture throughout the drying process should be explored further. Experiments should be conducted to generate data specifically for this task in order to accurately determine what level of accuracy can be achieved.
- Where possible, measurements of the fluidized bed dryer that are used in PCA or PLS modeling of the system should be automated and logged directly using a data acquisition system capable of acquiring data at a rate equal to the frequency response of the instruments.
- Experiments should be carried out using a pressure transducer with a higher sampling rate than 400 Hz to determine whether hydrodynamic behaviors can be detected at higher frequencies. This work would be best supplemented with experiments designed to investigate microscale phenomenon.

6. LITERATURE CITED

Alsberg, B.K., A.M. Woodward, and D.B. Kell, 1997. An introduction to wavelet transforms for chemometricians: a time-frequency approach. *Chemometrics and Intelligent Laboratory Systems*, **37**: 215-239.

Baffi, G., E. Martin, and J. Morris, 2002. Prediction intervals for non-linear projection to latent structures regression models. *Chemometrics and Intelligent Laboratory Systems*, **61**: 151-165.

Bakshi, B.R., 1998. Multiscale PCA with Application to Multivariate Process Monitoring. *American Institute of Chemical Engineering Journal*, **44**: 1596-1610.

Bjork, A. and L.G. Danielsson, 2002. Spectra of wavelet scale coefficients from process acoustic measurements as input for PLS modeling of pulp quality. *Journal of Chemometrics*, **16**: 521-528.

Burris, C.S., R.A. Gopinath, and H. Guo, 1998. Introduction to Wavelets and Wavelet Transforms A Primer. Prentice-Hall, New Jersey.

Chaplin, G., 2005. Monitoring Fluidized Bed Dryer Hydrodynamics using Pressure Fluctuations and Electrical Capacitance Tomography. Ph.D. thesis. University of Saskatchewan, Canada.

Chaplin, G., T. Pugsley, and C. Winters, 2004a. Application of Chaos Analysis to Fluidized Bed Drying of Pharmaceutical Granulate. In Fluidization XI: Present and Future for Fluidization Engineering. U. Arena, R. Chirone, M. Miccio, P. Salatino (Eds.), Engineering Foundation: New York, NY. 419-426.

Chaplin, G., T. Pugsley, and C. Winters, 2004b. Application of chaos analysis to pressure fluctuation data from a fluidized bed dryer containing pharmaceutical granule. *Powder Technology*, **142**: 110-220.

Chaplin, G., T. Pugsley, and C. Winters, 2005. The S-statistic as an early warning of entrainment in a fluidized bed dryer containing pharmaceutical granulate. *Powder Technology*, **149**: 148-156.

Chen, G. and T.J. McAvoy, 1998. Predictive on-line monitoring of continuous processes. *Journal of Process Control*, **8**: 409-420.

Chong, Y.O., D.P. O'Dea, E.T. White, P.L. Lee, and L.J. Leung, 1987. Control of quality of fluidization in a tall bed using the variance of pressure fluctuations. *Powder Technology*, **53**: 237-246.

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, **41**: 909 – 995.

Daubechies, I., 1992. Ten lectures on wavelets. Society for Industry and Applied Mathematics. Capital Press, Vermont.

de Jong, S., 1993. SIMPLS: an alternative approach to partial least squares regression. *Chemometrics Intell. Lab. Syst.*, **18**: 251-263.

Denham, M.C., 1997. Prediction intervals in partial least squares. *Journal of Chemometrics*, **11**: 39-52.

Duchesne, C. and J.F. MacGregor, 2001. Jackknife and bootstrap methods in the identification of dynamic models. *Journal of Process Control*, **11**: 553-564.

Eastment, H.T. and W.J. Krazanowski, 1982. Cross Validatory Choice of the Number of Components from a Principal Components Analysis. *Technometrics*, **24**: 73-77.

Ellis, N., L.A. Briens, J.R. Grace, H.T. Bi, and C.J. Lim, 2003. Characterization of dynamic behaviour in gas-solid turbulent fluidized bed using chaos and wavelet analyses. *Chemical Engineering Journal*, **96**: 105-116.

Faber, K. and B.R. Kowalski, 1997. Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares. *Journal of Chemometrics*, **11**: 181-238.

Faber, N. M., 2000. Comparison of two recently proposed expressions for partial least squares prediction error. *Chemometrics and Intelligent Laboratory Systems*, **52**: 123-134.

Gabor, D., 1946. Theory of Communication. *Journal of IEE (London)*. **93**: 429-457.

Geladi, P. and B.R. Kowalski, 1986. Partial Least Squares Regression: A Tutorial. *Analytical Chimica Acta*, **185**: 1-17.

Guo, Q., G. Yue, T. Suda, and J. Sato, 2003. Flow characteristics in a bubbling fluidized bed at elevated temperature. *Chemical Engineering and Processing*, **42**: 439-447.

Guo, Q., G. Yue, and J. Werther, 2002. Dynamics of Pressure Fluctuation in a Bubbling Fluidized Bed at High Temperature. *Industrial and Engineering Chemistry Research*, **41**: 3482-3488.

Haar, A., 1910. Theorie der orthogonalen Funktionen-Systeme. *Mathematische Analen*, **69**: 221-371.

Höskuldsson, A., 1988. PLS Regression Methods. *J. Chemometrics*, **2**: 211-228.

Jackson, J.E. and G.S. Mudholkar, 1979. Control Procedures for Residuals Associated With Principal Component Analysis. *Technometrics*, **21**: 341-349.

Jackson, J.E. ,1980. Principal components and factor analysis: I. principal components. *Journal of Quality Technology*, **12**: 201–213.

Johnsson, F., R.C. Ziejerveld, J.C. Schouten, C.M. van den Bleek, and B. Leckner, 2000. Characterization of fluidization regimes by time-series analysis of pressure fluctuations. *International Journal of Multiphase Flow*, **26**: 663-715.

Kosanovich, K.A. and M.J. Piovoso, 1997. PCA of Wavelet Transformed Data for Monitoring. *Intelligent Data Analysis*, **1**: 85-99.

Kourti, T. and J.F. MacGregor, 1996. Multivariate SPC Methods for Process and Product Monitoring. *Journal of Quality Technology*, **28**: 409-428.

Kourti, T., 2003. Multivariate dynamic data modeling for analysis and statistical control of batch processes, start-ups and grade transitions. *J. Chemometrics*, **17**: 93-109.

Kresta, J., J.F. MacGregor, and T.E. Marlin, 1991. Multivariate Statistical Monitoring of Process Operating Performance. *Can. J. Chem. Engng*., **69**: 35-47.

Kresta, J.V., T.E. Marlin, and J.F. MacGregor, 1994. Development of inferential process models using PLS. *Computers and Chemical Engineering*, **18**: 597-611.

Li, J., 2000. Compromise and resolution – Exploring the multi-scale nature of gas-solid fluidization. *Powder Technology*, **111**: 50-59.

Lu, N., F. Wang, and F. Gao, 2003. Combination method of Principal Component and Wavelet Analysis for Multivariate Process Monitoring and Fault Diagnosis. *Industrial and Engineering Chemical Research*, **42**: 4198-4207.

Lingren, F., P. Geladi, and S. Wold, 1993. The Kernel Algorithm for PLS. *Journal of Chemometrics*, **7**: 45-59.

Lu, X. and H. Li, 1999. Wavelet analysis of pressure fluctuation signals in a bubbling fluidized bed. *Chemical Engineering Journal*, **75**: 113-119.

Mallet, S.G., 1989. A Theory for Multiresolution Signal Decomposition: the Wavelet Representation. *IEEE Transactions on Pattern and Machine Intelligence*, **11**: 674-693.

Nomikos, P. and J.F. MacGregor, 1994. Monitoring Batch Processes Using Multiway Principal Component Analysis. *AIChE Journal*, **40**: 1361-1375.

Nomikos, P. and J.F. MacGregor, 1995a. Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics*, **37**: 41–59.

Nomikos, P. and J. F. MacGregor, 1995b. Multi-way partial least squares in monitoring batch processes. *Chemometrics and Intelligent Laboratory Systems*. **30**: 97-108.

Pearson, K.,1901. On lines and planes of closest fit to systems of points in space. *Phil. Mag., Ser. B.*, **2**: 559–572.

Ren, J. and J. Li, 1998. Wavelet Analysis of Dynamic Behaviour in Fluidized Beds. In Fluidization IX Proceedings of the Ninth Engineering Foundation Conference on Fluidization. Durango, Colorado. May 17-22, 1998. United Engineering Foundation, New York.

Ren, J., Q. Mao, J. Li, and W. Lin, 2001. Wavelet analysis of dynamic behavior in fluidized beds. *Chemical Engineering Science*, **56**: 981-988.

Teppola, P., S.P. Mujunen, P. Minkkinen, T. Puijola, and P. Pursiheimo, 1998. Principal component analysis, contribution plots, and feature weights in the monitoring of sequential process data from a paper machine's wet end. *Chemometrics and Intelligent Laboratory Systems*, **44**: 307-317.

Teppola, P. and P. Minkkinen, 2000. Wavelet-PLS regression models for both exploratory data analysis and process monitoring. *Journal of Chemometrics*, **14**: 383-399.

Trygg, J. and S. Wold, 1998. PLS regression on wavelet compressed NIR spectra. *Chemometrics and Intelligent Laboratory Systems*, **42**: 209-220.

Trygg, J., N. Kettaneh-Wold, and L. Wallbacks, 2001. 2D wavelet analysis and compression of on-line industrial process data. *Journal of Chemometrics*, **15**: 299-319.

van den Bleek, C.M., M.-O. Coppens, and J.C. Schouten, 2002. Application of chaos analysis to multiphase reactors. *Chemical Engineering Science*, **57**: 4763-4778.

van der Schaaf, J., J.C. Schouten, and C.M. van den Bleek, 1998. Origin, propagation, and attenuation of pressure waves in gas-solid fluidized beds. *Powder Technology*, **95**: 220-233.

van Ommen, J.R., J.C. Schouten, and C.M. van den Bleek, 1999. Monitoring fluidization dynamics for detection of changes in fluidized bed composition and operating conditions. *ASME Journal of Fluids Engineering*, **121**: 887-894.

van Ommen, J.R., J.C. Schouten, M.L.M. vander Stappen, C.M. van den Bleek, 1999. Response characteristics of probe-transducer systems for pressure measurements in gas-solid fluidized beds: how to prevent pitfalls in dynamic pressure measurements. *Powder Technology*, **106**: 199-218.

van Ommen, J.R., J.C. Schouten, M.-O. Coppens, and C.M. van den Bleek, 2000. Early Warning of Agglomeration in Fluidized Beds by Attractor Comparison. *AIChE Journal*, **46**: 2183-2197.

Werther, J., 1999. Measurement techniques in fluidized beds. *Powder Technology*, **102**: 15-36.

Westerhuis, J.A., T. Kourti, and J.F. MacGregor, 1999. Comparing Alternative Approaches for Multivariate Statistical Analysis of Batch Process Data. *J. Chemometrics*, **13**: 397-413.

Westerhuis, J.A., S.P. Gurden, and A.K. Smilde, 2000a. Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, **51**: 95-114.

Westerhuis, J.A., S.P. Gurden, and A.K. Smilde, 2000b. Standardized Q-stastistic for improved sensitivity in the monitoring of residuals in MSPC. *Journal of Chemometrics*, **14**: 335-349.

Wise, B.M., N.B. Gallagher, S. Watts Butler, D.D. White Jr. and G.G. Barna, 1999. A Comparison of Principal Component Analysis, Multiway Principal Component Analysis, Trilinear Decomposition, and Parallel Factor Analysis for Fault Detection in a Semiconductor Etch Process. *J. Chemometrics*, **13**: 379-396.

Wise, B.M., N.B. Gallagher, R. Bro, and J.M. Shaver, PLS_Toolbox 3.5 for use with MATLAB Users Guide. Eigenvector Research, Inc., Washington, 2004.

Wold, S., 1978. Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics*, **20**: 397-405.

Wold, H., 1982. Soft modeling: the basic design and some extensions. In K.-G. Joreskog, H. Wold (Eds.), *Systems Under Indirect Observation, vol. 2*, North-Holland, pp. 1-53.

Wold, S., K. Esbensen and P. Geladi ,1987. Principal component analysis. *Chem. Intell. Lab. Sys*, **2**: 37–52.

Zhao, G.-B. and Y.-R. Yang, 2003. Multiscale Resolution of Fluidized-Bed Pressure Fluctuations. *AIChE Journal*, **49**: 869-882.

APPENDIX A

WAVELET RECONSTRUCTION ERRORS

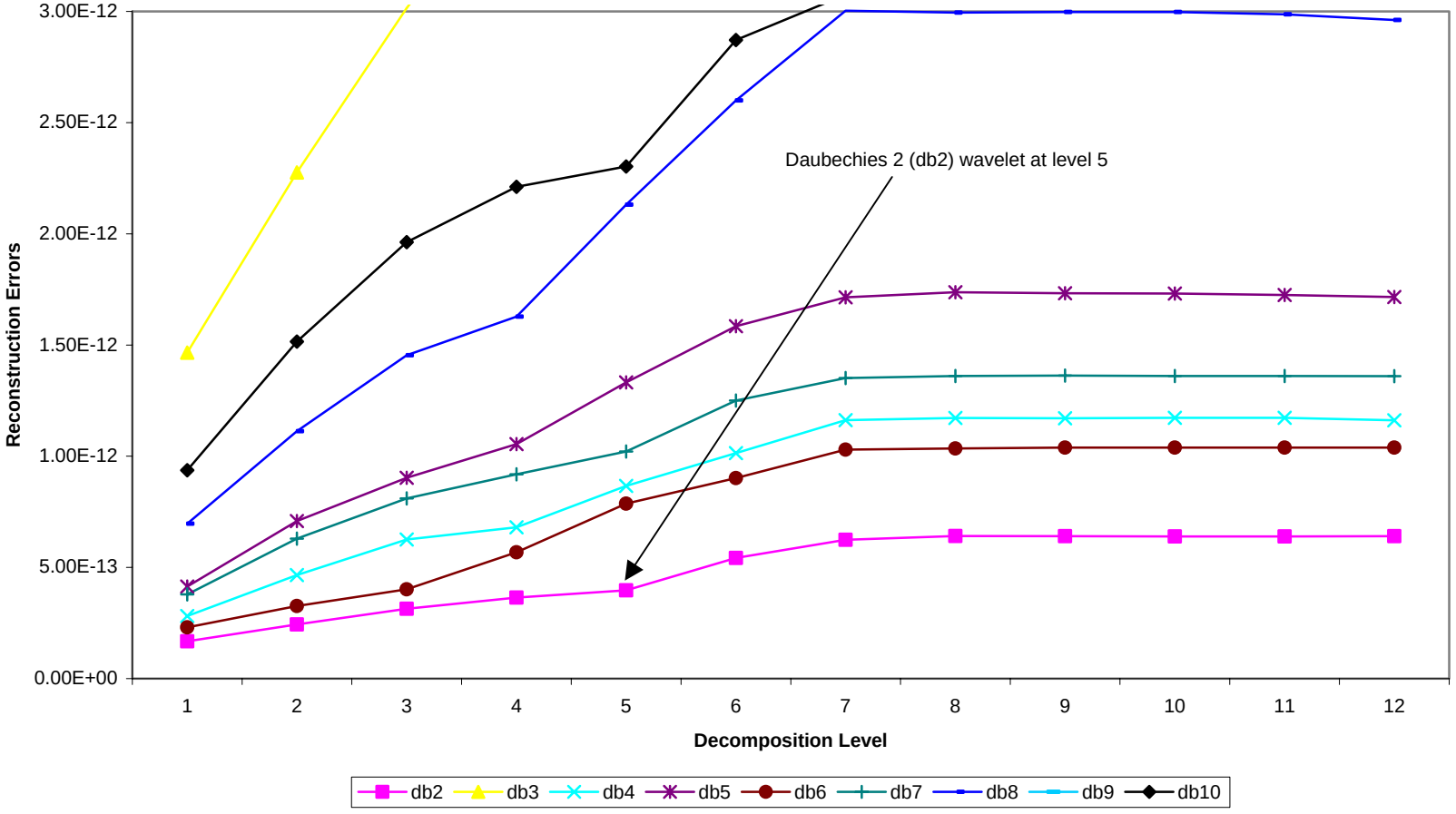

Figure A.1: Reconstruction error for wavelet decomposed pressure signals from the dry bed study

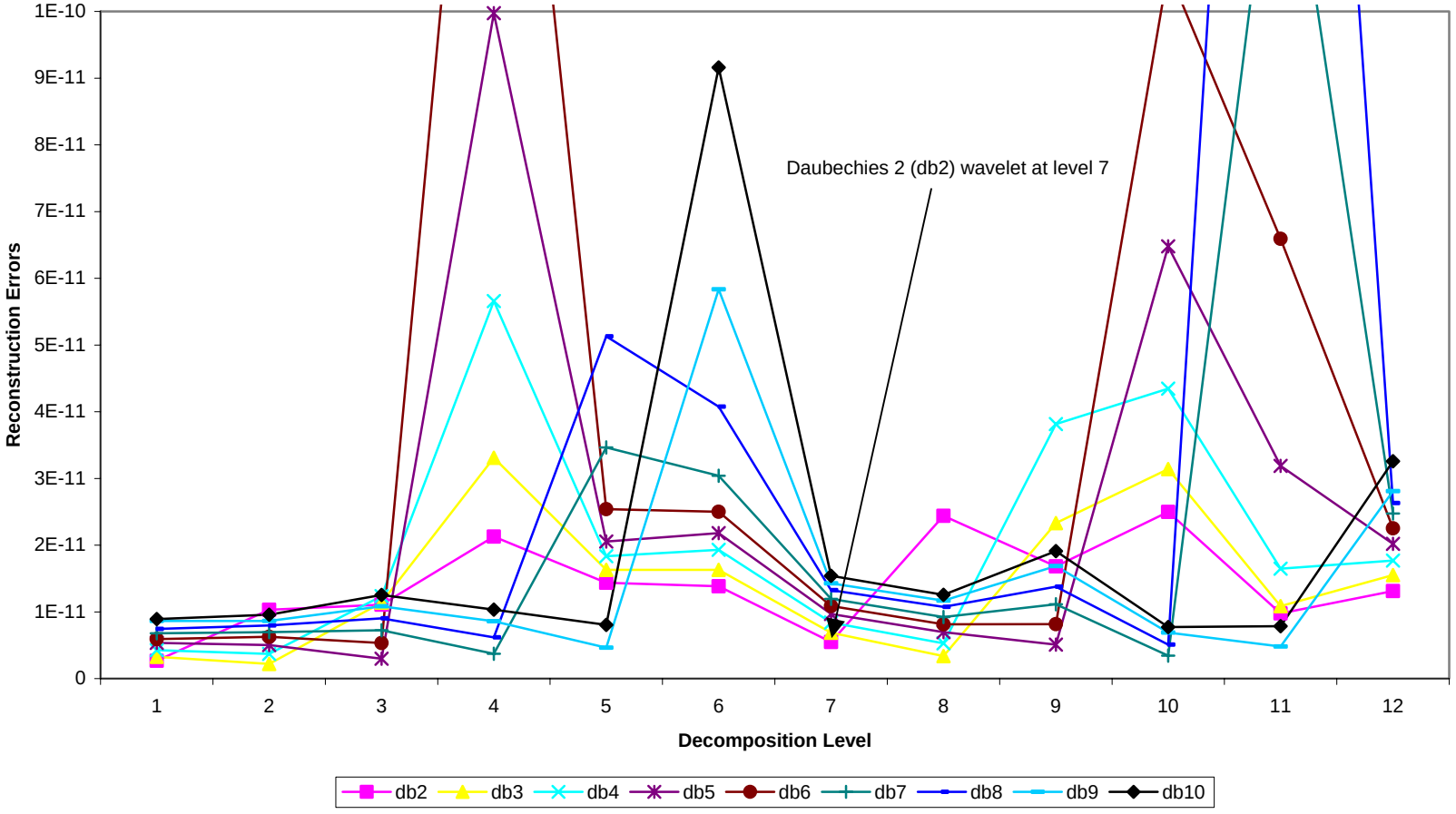

Figure A.2: Reconstruction error for wavelet decomposed pressure signals from the batch drying study

APPENDIX B

DRY BED STUDY EXPERIMENTAL CONDITIONS

Table B.1: Dry bed study experimental parameters

Test Date	Pressure Signal File #	Granulate Batch #	Test ID #	Superficial Velocity (m/s)	Pressure Sensor Location	Initial Bed Mass (g)	Granulate Mean Particle Size (microns)
August 19, 2002	190802a	1198-24	1	1.95	1	2006	318
August 19, 2002	190802b	1198-24	2	1.95	1	2006	318
August 19, 2002	190802c	1198-24	3	1.95	1	2006	318
August 19, 2002	190802d	1198-24	4	2.19	1	2006	318
August 19, 2002	190802e	1198-24	5	2.19	1	2006	318
August 19, 2002	190802f	1198-24	6	2.19	1	2006	318
August 19, 2002	1908902g	1198-24	7	2.44	1	2006	318
August 19, 2002	190802h	1198-24	8	2.44	1	2006	318
August 19, 2002	190802i	1198-24	9	2.44	1	2006	318
August 19, 2002	190802j	1198-24	10	2.68	1	2006	318
August 19, 2002	190802k	1198-24	11	2.68	1	2006	318
August 19, 2002	190802l	1198-24	12	2.68	1	2006	318
August 19, 2002	190802m	1198-24	13	2.92	1	2006	318
August 19, 2002	190802n	1198-24	14	2.92	1	2006	318
August 19, 2002	190802o	1198-24	15	2.92	1	2006	318
August 21, 2002	210802a1	1198-24	16	1.95	2	2001	318
August 21, 2002	210802a2	1198-24	17	1.95	2	2001	318
August 21, 2002	210802a3	1198-24	18	1.95	2	2001	318
August 21, 2002	210802a4	1198-24	19	2.19	2	2001	318
August 21, 2002	210802a5	1198-24	20	2.19	2	2001	318
August 21, 2002	210802a6	1198-24	21	2.19	2	2001	318

Table B.1: Dry bed experimental parameters (continued)

Test Date	Pressure Signal File #	Granulate Batch #	Test ID #	Superficial Velocity (m/s)	Pressure Sensor Location	Initial Bed Mass (g)	Granulate Mean Particle Size (microns)
August 21, 2002	210802a7	1198-24	22	2.44	2	2001	318
August 21, 2002	210802a8	1198-24	23	2.44	2	2001	318
August 21, 2002	210802a9	1198-24	24	2.44	2	2001	318
August 21, 2002	210802a10	1198-24	25	2.68	2	2001	318
August 21, 2002	210802a11	1198-24	26	2.68	2	2001	318
August 21, 2002	210802a12	1198-24	27	2.68	2	2001	318
August 21, 2002	210802a13	1198-24	28	2.92	2	2001	318
August 21, 2002	210802a14	1198-24	29	2.92	2	2001	318
August 21, 2002	210802a15	1198-24	30	2.92	2	2001	318
August 21, 2002	210802b1	1198-24	31	1.95	3	2010	318
August 21, 2002	210802b2	1198-24	32	1.95	3	2010	318
August 21, 2002	210802b3	1198-24	33	1.95	3	2010	318
August 21, 2002	210802b4	1198-24	34	2.19	3	2010	318
August 21, 2002	210802b5	1198-24	35	2.19	3	2010	318
August 21, 2002	210802b6	1198-24	36	2.19	3	2010	318
August 21, 2002	210802b7	1198-24	37	2.44	3	2010	318
August 21, 2002	210802b8	1198-24	38	2.44	3	2010	318
August 21, 2002	210802b9	1198-24	39	2.44	3	2010	318
August 21, 2002	210802b10	1198-24	40	2.68	3	2010	318
August 21, 2002	210802b11	1198-24	41	2.68	3	2010	318
August 21, 2002	210802b12	1198-24	42	2.68	3	2010	318

Table B.1: Dry bed experimental parameters (continued)

Test Date	Pressure Signal File #	Granulate Batch #	Test ID #	Superficial Velocity (m/s)	Pressure Sensor Location	Initial Bed Mass (g)	Granulate Mean Particle Size (microns)
August 21, 2002	210802b13	1198-24	43	2.92	3	2010	318
August 21, 2002	210802b14	1198-24	44	2.92	3	2010	318
August 21, 2002	210802b15	1198-24	45	2.92	3	2010	318
August 22, 2002	220802r1	1198-55,56	46	1.95	1	2006	179
August 22, 2002	220802r2	1198-55,56	47	1.95	1	2006	179
August 22, 2002	220802r3	1198-55,56	48	1.95	1	2006	179
August 22, 2002	220802r4	1198-55,56	49	2.19	1	2006	179
August 22, 2002	220802r5	1198-55,56	50	2.19	1	2006	179
August 22, 2002	220802r6	1198-55,56	51	2.19	1	2006	179
August 22, 2002	220802r7	1198-55,56	52	2.44	1	2006	179
August 22, 2002	220802r8	1198-55,56	53	2.44	1	2006	179
August 22, 2002	220802r9	1198-55,56	54	2.44	1	2006	179
August 22, 2002	220802r10	1198-55,56	55	2.68	1	2006	179
August 22, 2002	220802r11	1198-55,56	56	2.68	1	2006	179
August 22, 2002	220802r12	1198-55,56	57	2.68	1	2006	179
August 22, 2002	220802r13	1198-55,56	58	2.92	1	2006	179
August 22, 2002	220802r14	1198-55,56	59	2.92	1	2006	179
August 22, 2002	220802r15	1198-55,56	60	2.92	1	2006	179
August 23, 2002	230802r1	1198-24	61	1.95	4	2000	318
August 23, 2002	230802r2	1198-24	62	1.95	4	2000	318
August 23, 2002	230802r3	1198-24	63	1.95	4	2000	318

Table B.1: Dry bed experimental parameters (continued)

Test Date	Pressure Signal File #	Granulate Batch #	Test ID #	Superficial Velocity (m/s)	Pressure Sensor Location	Initial Bed Mass (g)	Granulate Mean Particle Size (microns)
August 23, 2002	230802r4	1198-24	64	2.19	4	2000	318
August 23, 2002	230802r5	1198-24	65	2.19	4	2000	318
August 23, 2002	230802r6	1198-24	66	2.19	4	2000	318
August 23, 2002	230802r7	1198-24	67	2.44	4	2000	318
August 23, 2002	230802r8	1198-24	68	2.44	4	2000	318
August 23, 2002	230802r9	1198-24	69	2.44	4	2000	318
August 23, 2002	230802r10	1198-24	70	2.68	4	2000	318
August 23, 2002	230802r11	1198-24	71	2.68	4	2000	318
August 23, 2002	230802r12	1198-24	72	2.68	4	2000	318
August 23, 2002	230802r13	1198-24	73	2.92	4	2000	318
August 23, 2002	230802r14	1198-24	74	2.92	4	2000	318
August 23, 2002	230802r15	1198-24	75	2.92	4	2000	318
August 27, 2002	270802r1	1198-55,56	76	1.95	2	2002	179
August 27, 2002	270802r2	1198-55,56	77	1.95	2	2002	179
August 27, 2002	270802r3	1198-55,56	78	1.95	2	2002	179
August 27, 2002	270802r4	1198-55,56	79	2.19	2	2002	179
August 27, 2002	270802r5	1198-55,56	80	2.19	2	2002	179
August 27, 2002	270802r6	1198-55,56	81	2.19	2	2002	179
August 27, 2002	270802r7	1198-55,56	82	2.44	2	2002	179
August 27, 2002	270802r8	1198-55,56	83	2.44	2	2002	179
August 27, 2002	270802r9	1198-55,56	84	2.44	2	2002	179

Table B.1: Dry bed experimental parameters (continued)

Test Date	Pressure Signal File #	Granulate Batch #	Test ID #	Superficial Velocity (m/s)	Pressure Sensor Location	Initial Bed Mass (g)	Granulate Mean Particle Size (microns)
August 27, 2002	270802r10	1198-55,56	85	2.68	2	2002	179
August 27, 2002	270802r11	1198-55,56	86	2.68	2	2002	179
August 27, 2002	270802r12	1198-55,56	87	2.68	2	2002	179
August 27, 2002	270802r13	1198-55,56	88	2.92	2	2002	179
August 27, 2002	270802r14	1198-55,56	89	2.92	2	2002	179
August 27, 2002	270802r15	1198-55,56	90	2.92	2	2002	179
August 27, 2002	270802br1	1198-55,56	91	1.95	3	2003	179
August 27, 2002	270802br2	1198-55,56	92	1.95	3	2003	179
August 27, 2002	270802br3	1198-55,56	93	1.95	3	2003	179
August 27, 2002	270802br4	1198-55,56	94	2.19	3	2003	179
August 27, 2002	270802br5	1198-55,56	95	2.19	3	2003	179
August 27, 2002	270802br6	1198-55,56	96	2.19	3	2003	179
August 27, 2002	270802br7	1198-55,56	97	2.44	3	2003	179
August 27, 2002	270802br8	1198-55,56	98	2.44	3	2003	179
August 27, 2002	270802br9	1198-55,56	99	2.44	3	2003	179
August 27, 2002	270802br10	1198-55,56	100	2.68	3	2003	179
August 27, 2002	270802br11	1198-55,56	101	2.68	3	2003	179
August 27, 2002	270802br12	1198-55,56	102	2.68	3	2003	179
August 27, 2002	270802br13	1198-55,56	103	2.92	3	2003	179
August 27, 2002	270802br14	1198-55,56	104	2.92	3	2003	179
August 27, 2002	270802br15	1198-55,56	105	2.92	3	2003	179

APPENDIX C

SCORES PLOTS FROM OVERALL PCA MODELS

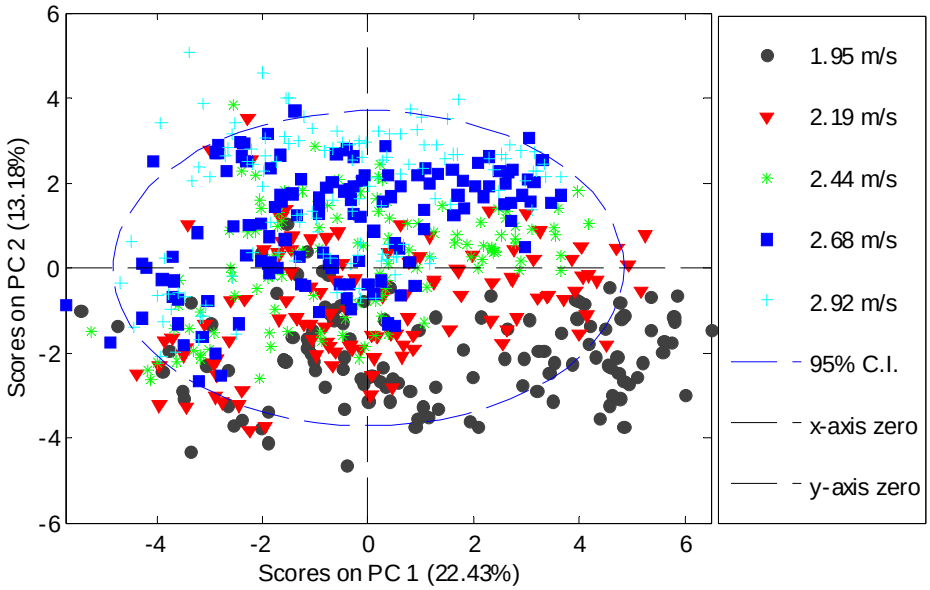

Figure C.1: Scores plot from an overall PCA model containing all 630 data points. Superficial velocities of each point are specified.

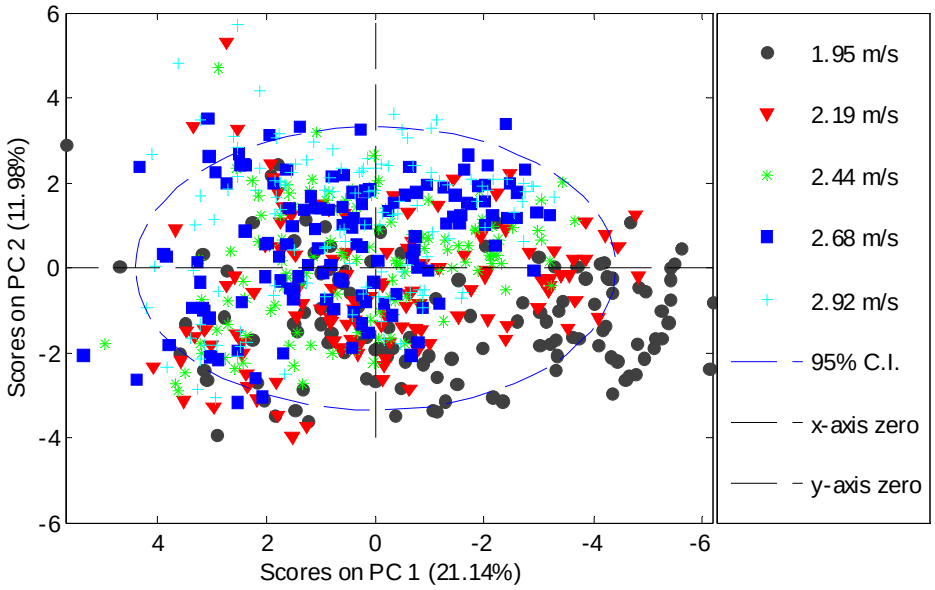

Figure C.2: Scores plot from an overall PCA model containing all 630 points. Reduced data set including only pressure fluctuation data was used to construct this model. Superficial velocities of each point are specified.

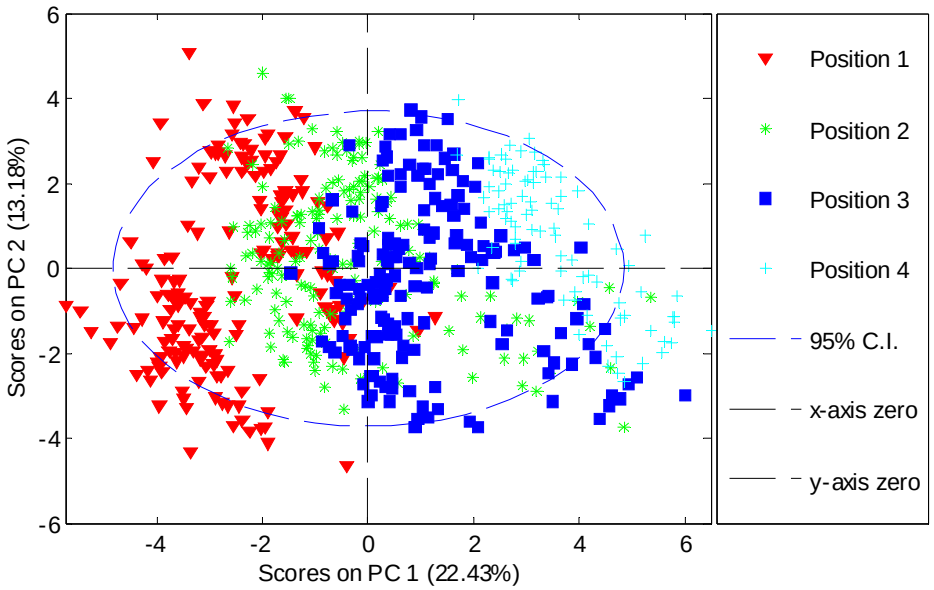

Figure C.3: Scores plot from an overall PCA model containing all 630 points. Pressure sensor position of each point is specified.

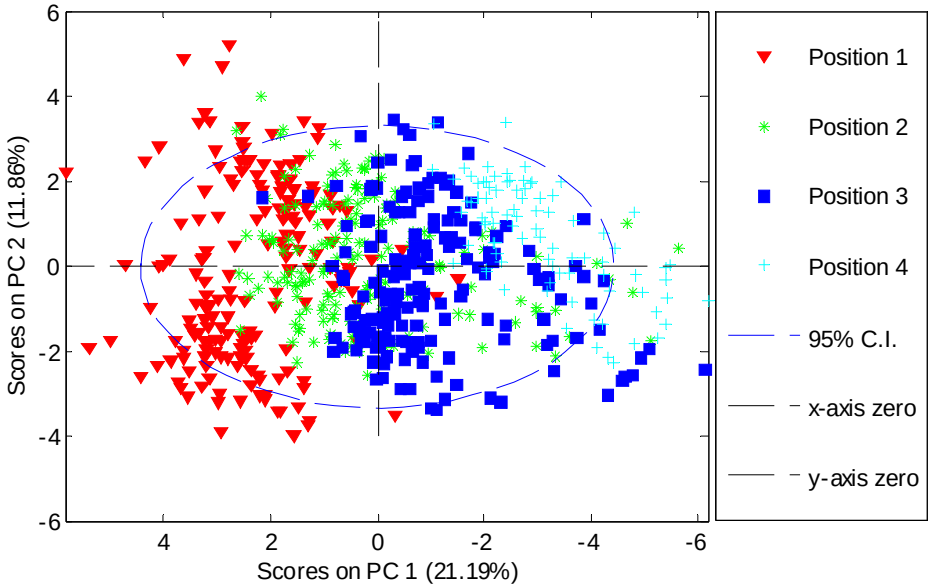

Figure C.4: Scores plot from an overall PCA model containing all 630 data points. Reduced data set containing only pressure fluctuation data was used to construct this model. Pressure sensor position of each point is specified.

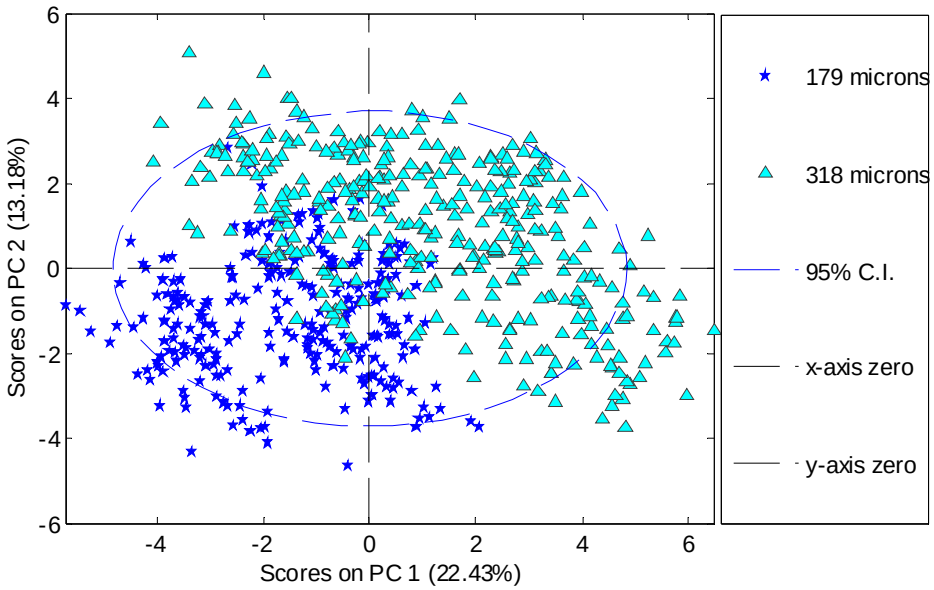

Figure C.5: Scores plot from an overall PCA model containing all 630 data points. Mean particle size for the granulate used in each test is specified.

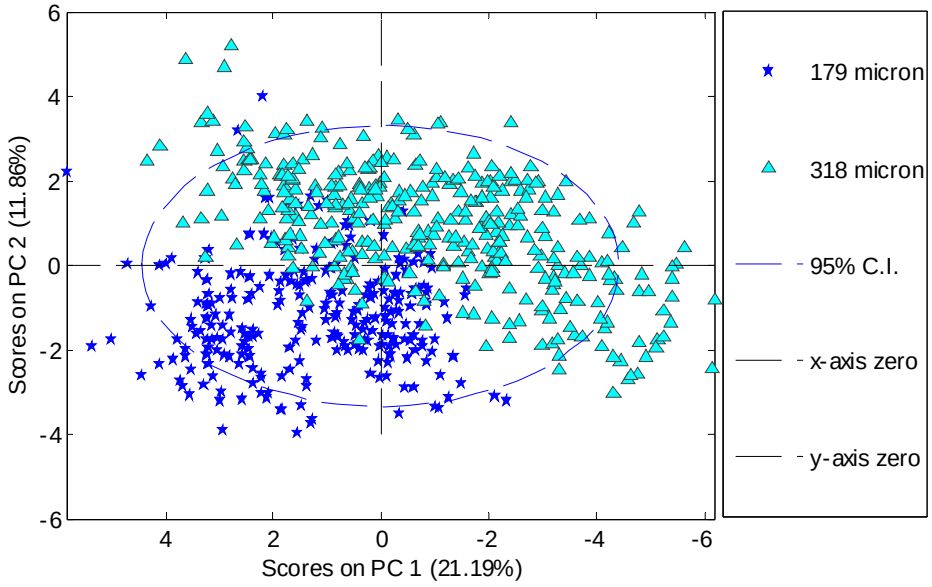

Figure C.6: Scores plot from an overall PCA model containing all 630 points. Reduced data set containing only pressure fluctuation data was used to construct this model. Mean particle size for the granulate used in each test is specified.

APPENDIX D

SCORES AND LOADINGS PLOTS FROM DRY BED EVALUATION OF PRESSURE SENSOR EFFECT

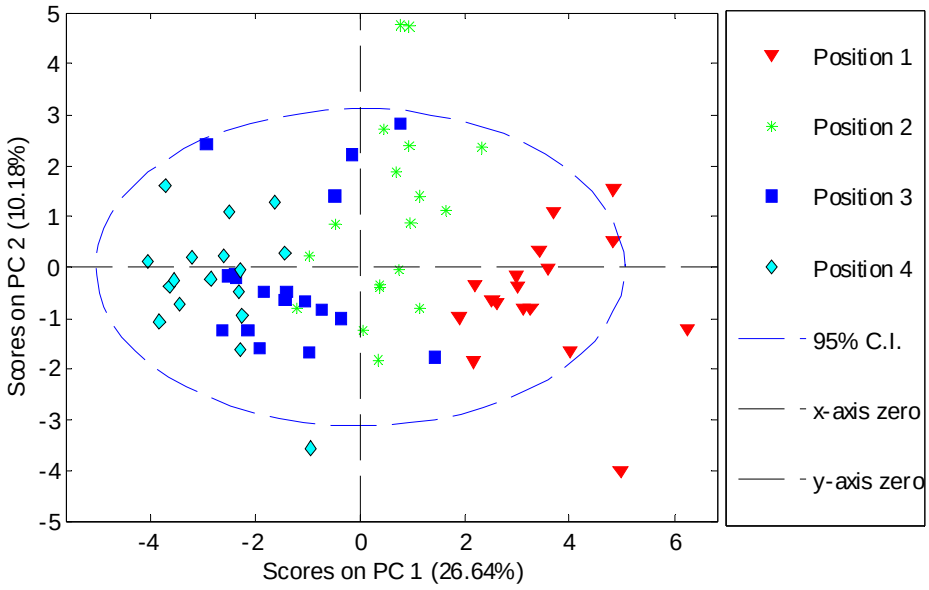

Figure D.1: Scores plot from a PCA model examining the effect of sensor position. Model calculated at a superficial velocity of 2.44 m/s using regular granulate. No deterministic variation is seen along PC 2.

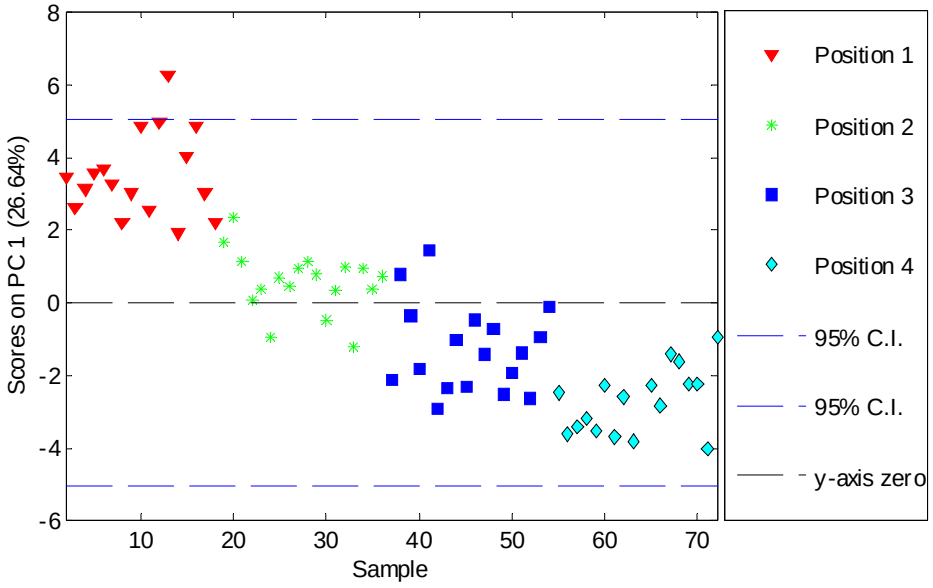

Figure D.2: Scores plot from a PCA model examining the effect of sensor position with only 1 retained component. Model calculated at a superficial velocity of 2.44 m/s using regular granulate.

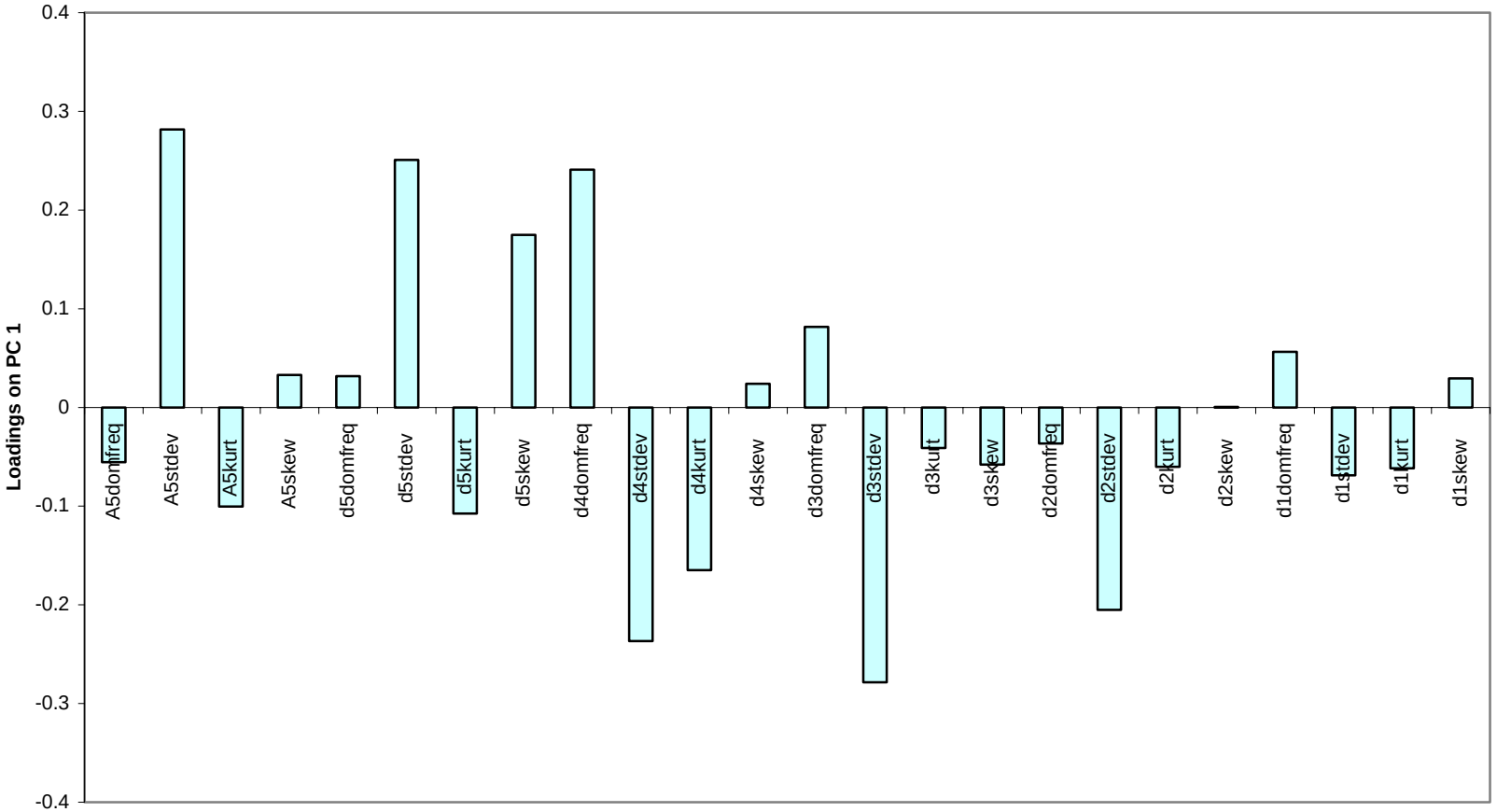

Figure D.3: Loadings plot for a 1 component PCA model evaluating the effects of sensor position. Models were calculated at a single superficial velocity and granulate type.

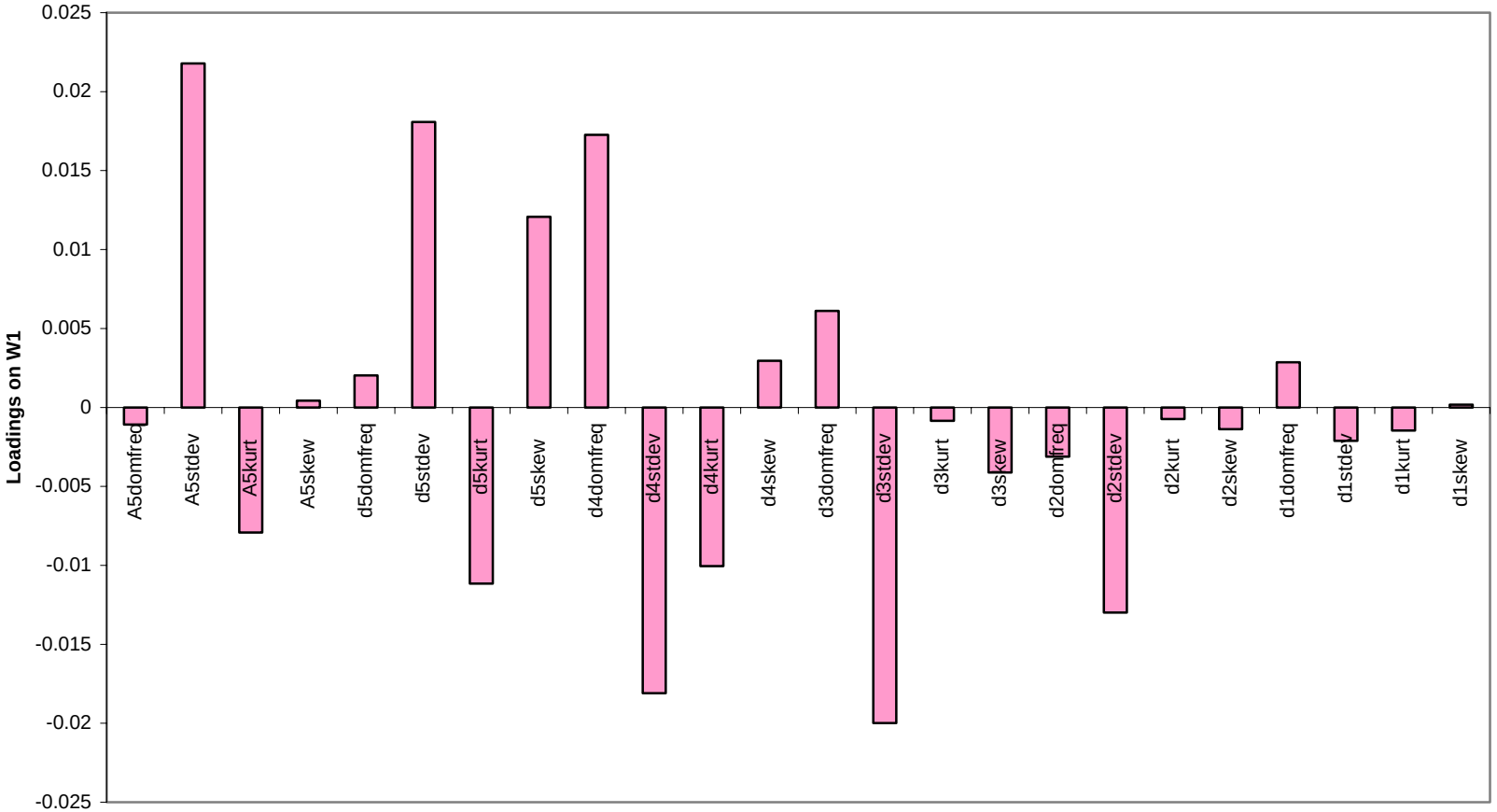

Figure D.4: Loadings plot showing the dominant latent variable in a PLS model of pressure sensor position.

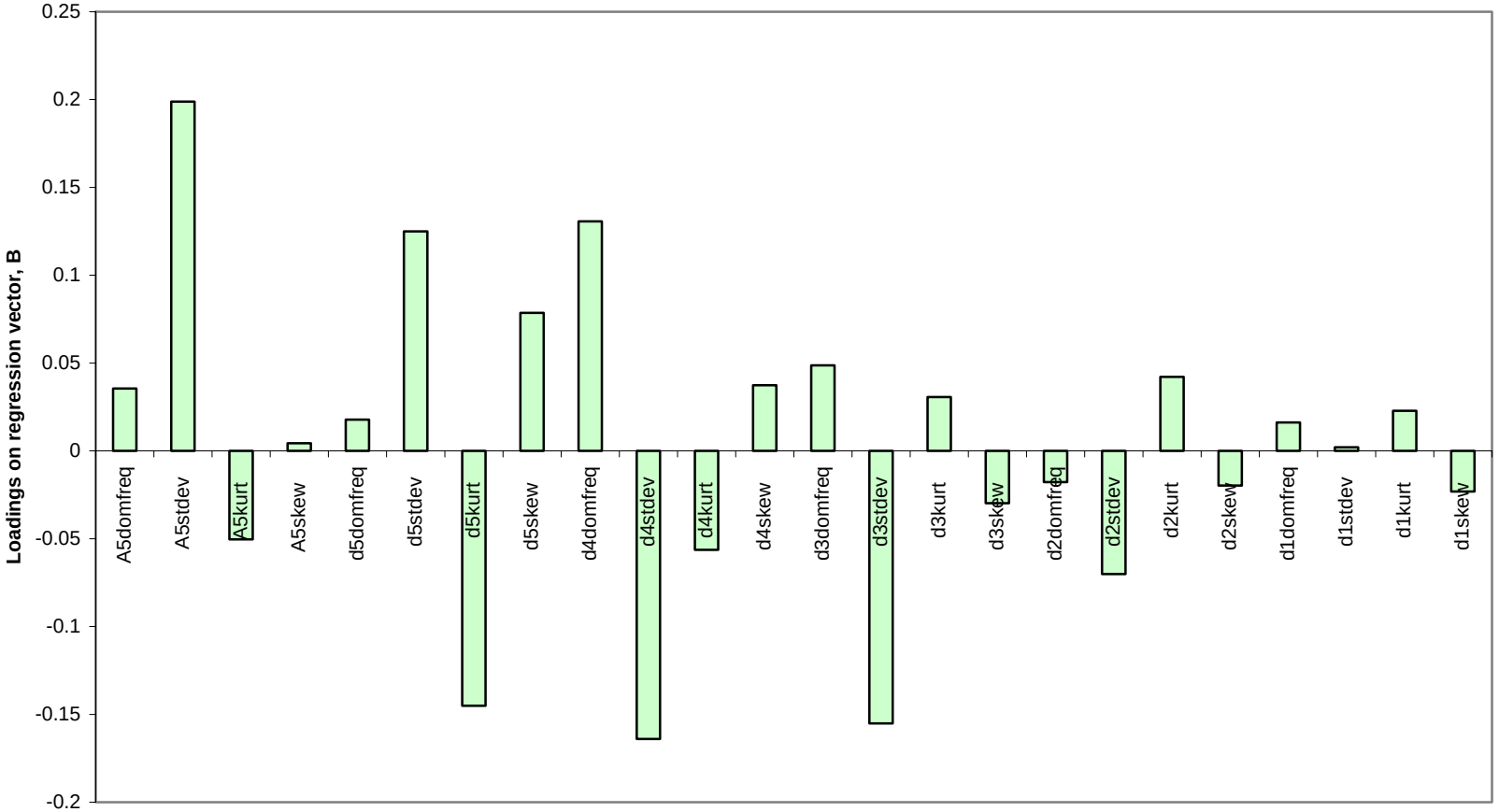

Figure D.5: Loadings plot showing the regression vector for a PLS model of sensor position.

APPENDIX E

DRY BED STUDY GRANULATE PARTICLE SIZE DISTRIBUTIONS

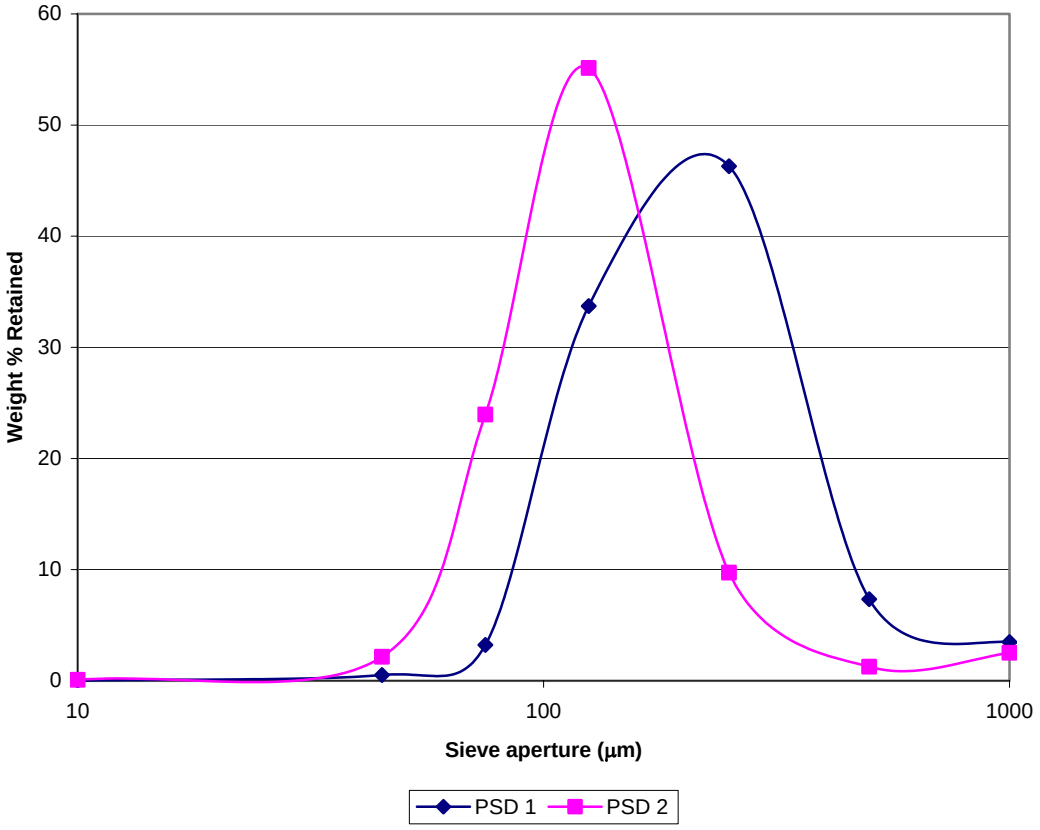

Figure E.1: Dry bed study granulate particle size distributions (Chaplin, Pugsley, and Winters, 2004b)

Table E.1: Dry bed study granulate particle size distribution (Chaplin, Pugsley, and Winters, 2004b)

Sieve Size	Aperture (µm)	PSD 1 Ind. Wt.% Retained	PSD 2 Ind. Wt.% Retained
10 Mesh	2000	5.41	5.14
18 Mesh	1000	3.49	2.52
35 Mesh	500	7.34	1.26
60 Mesh	250	46.29	9.73
120 Mesh	125	33.71	55.14
200 Mesh	75	3.23	23.96
325 Mesh	45	0.52	2.16
pan	10	0.01	0.09
Total	N/A	**100.00**	**100.00**

APPENDIX F

PLS WEIGHT AND REGRESSION VECTORS FROM PARTICLE SIZE DISTRIBUTION MODEL

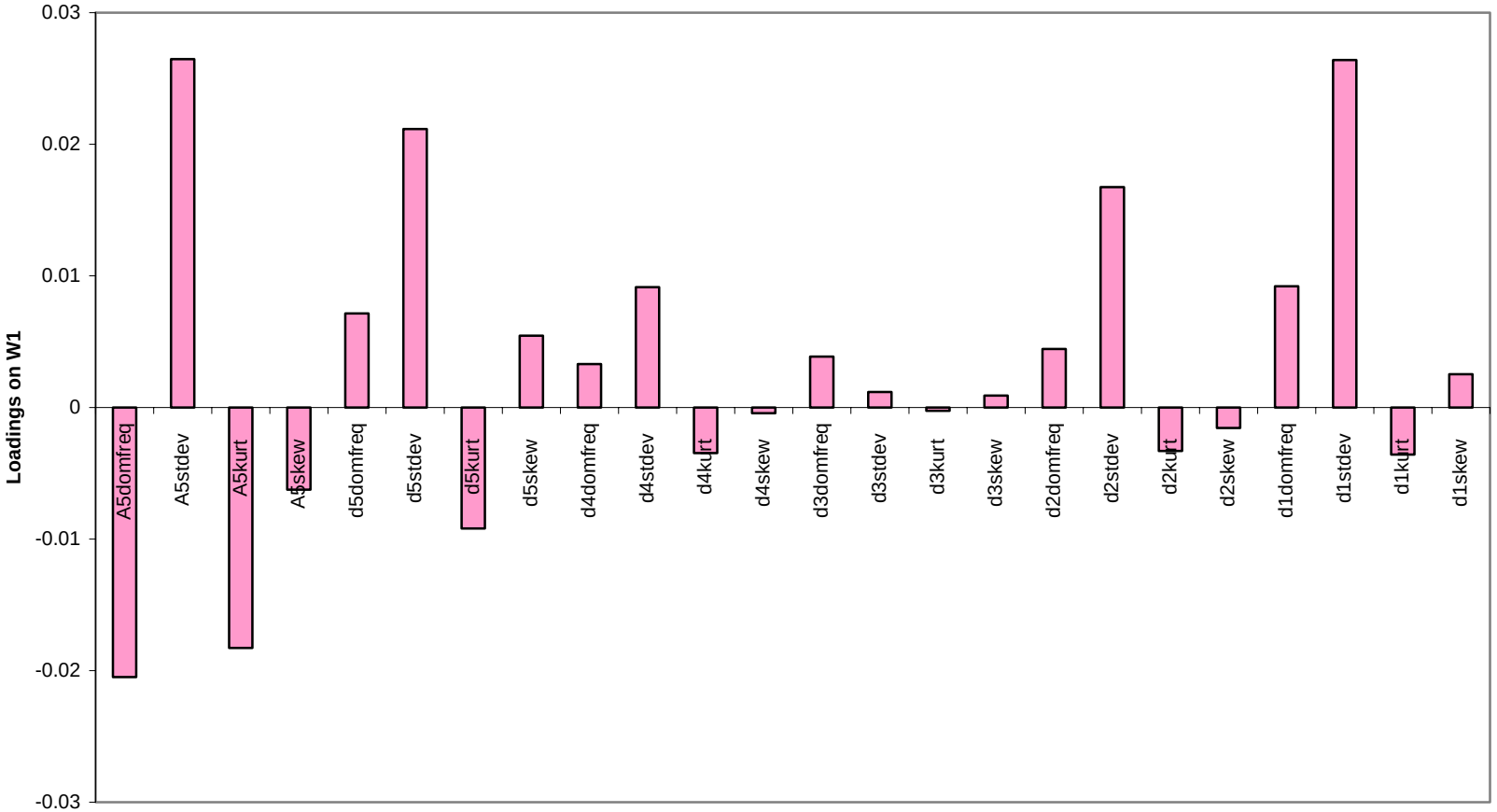

Figure F.1: Dominant weight vector from PLS models constructed to examine the effect of particle size distribution.

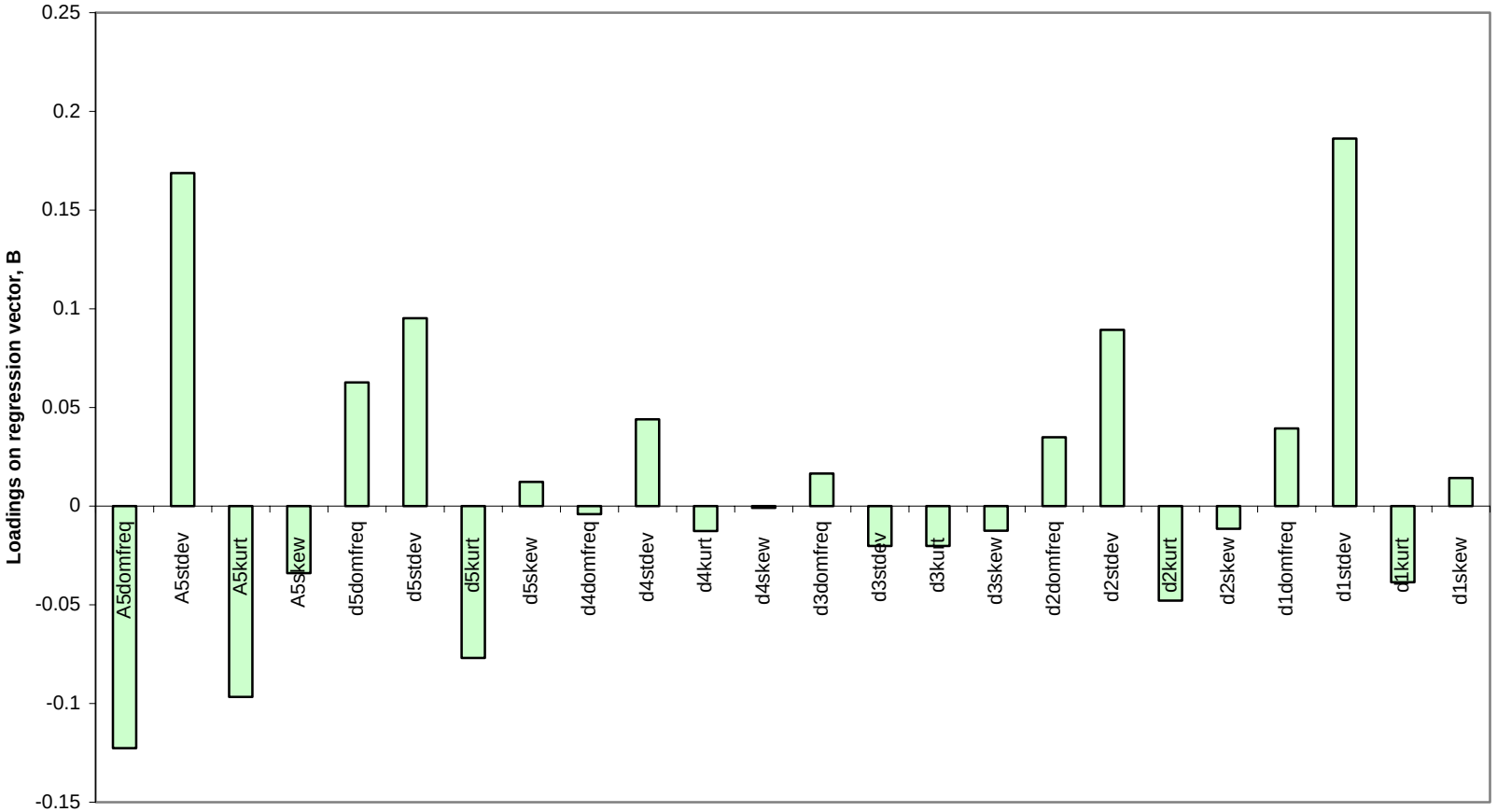

Figure F.2: Regression vector loadings from a PLS model constructed to examine the effects of particle size distribution.

APPENDIX G

BATCH DRYING STUDY EXPERIMENTAL CONDITIONS

Table G.1: Batch drying study experimental parameters

Experiment Date	Test ID #	Pressure Signal File #	Initial Bed Mass (kg)	Inlet Temperature (^{o}C)	Sensor Position	Superficial Velocity (m/s)
August 2, 2002	1	290802b	3.00	65	2	2.46
September 3, 2002	2	030902a	2.75	55	1	2.46
September 3, 2002	3	030902b	3.25	75	1	2.46
September 3, 2002	4	030902c	2.75	75	3	2.46
September 10, 2002	5	100902a	3.25	55	1	2.46
September 10, 2002	6	100902b	2.75	55	3	2.46
September 10, 2002	7	100902c	3.25	75	3	2.46
September 16, 2002	8	160902a	3.00	65	2	2.46
September 16, 2002	9	160902b	3.25	55	3	2.46
September 16, 2002	10	160902c	2.75	75	1	2.46
September 23, 2002	11	230902a	3.00	65	2	2.46
July 17, 2002	12	170702b	2.50	65	-	2.34
August 1, 2002	13	010802b	2.56	65	-	2.60 – 2.08
August 2, 2002	14	020802a	2.56	70	-	2.60 – 2.08
August 2, 2002	15	020802b	3.00	65	-	2.60 – 2.08

APPENDIX H

LOADING PLOTS FOR MPCA MODEL OF THE BATCH DRYING PROCESS

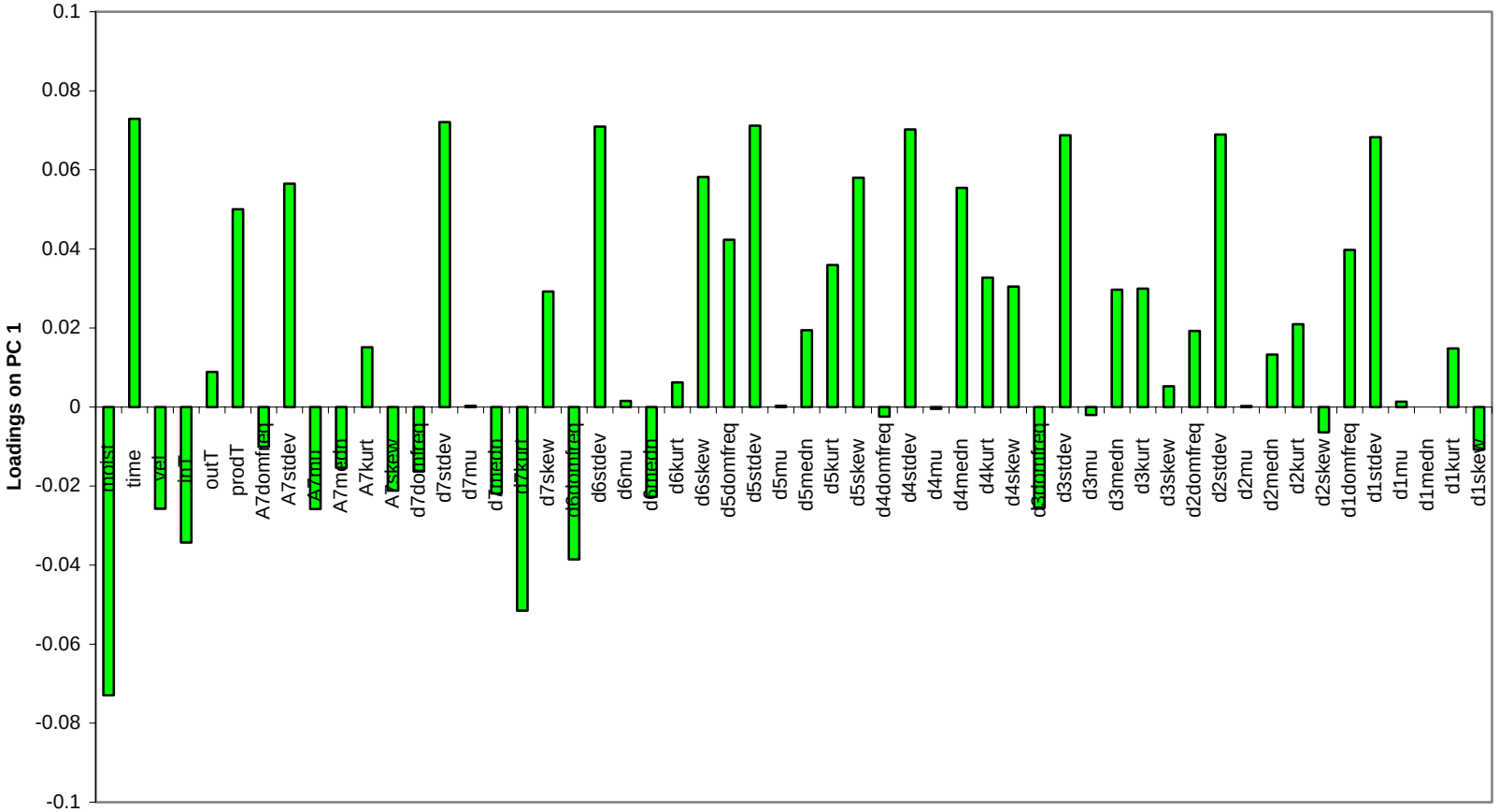

Figure H.1: Loadings plot for PC 1 from a MPCA model of the batch drying process.

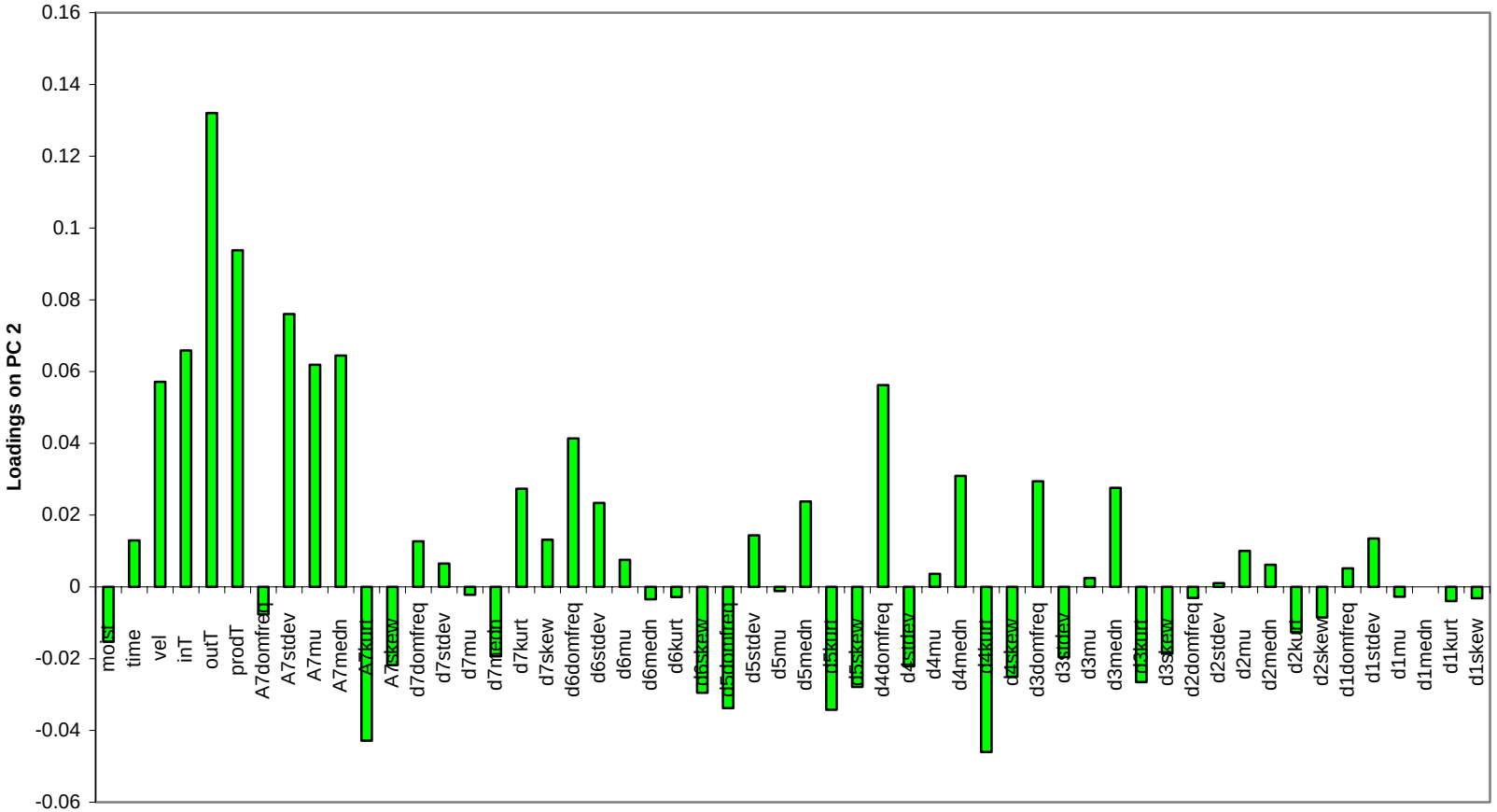

Figure H.2: Loadings plot for PC 2 from a MPCA model of the batch drying process.

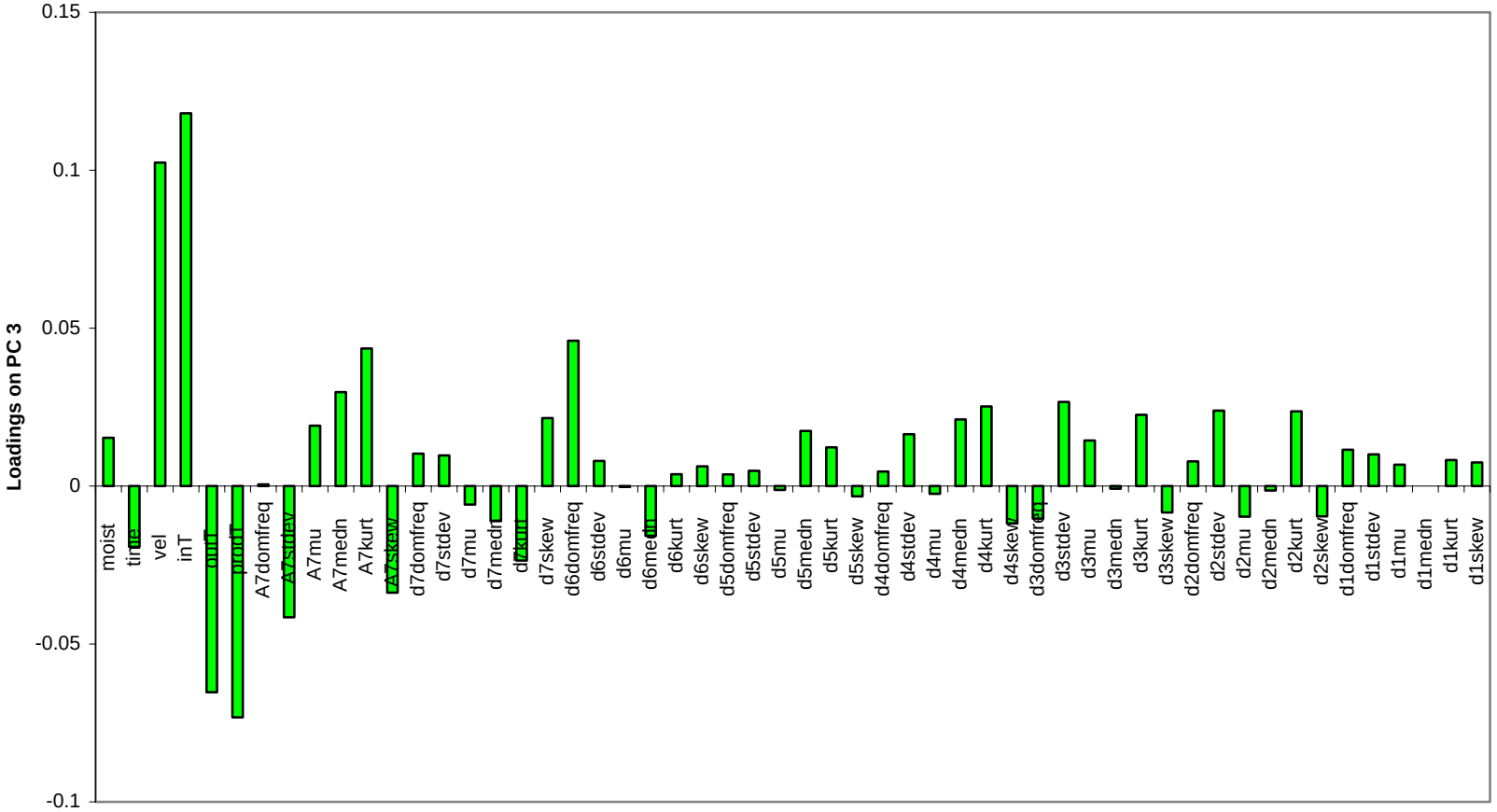

Figure H.3: Loadings plot for PC 3 from a MPCA model of the batch drying process.

APPENDIX I

LOADING PLOTS FROM MPCA MODELS OF DRYING PROCESS

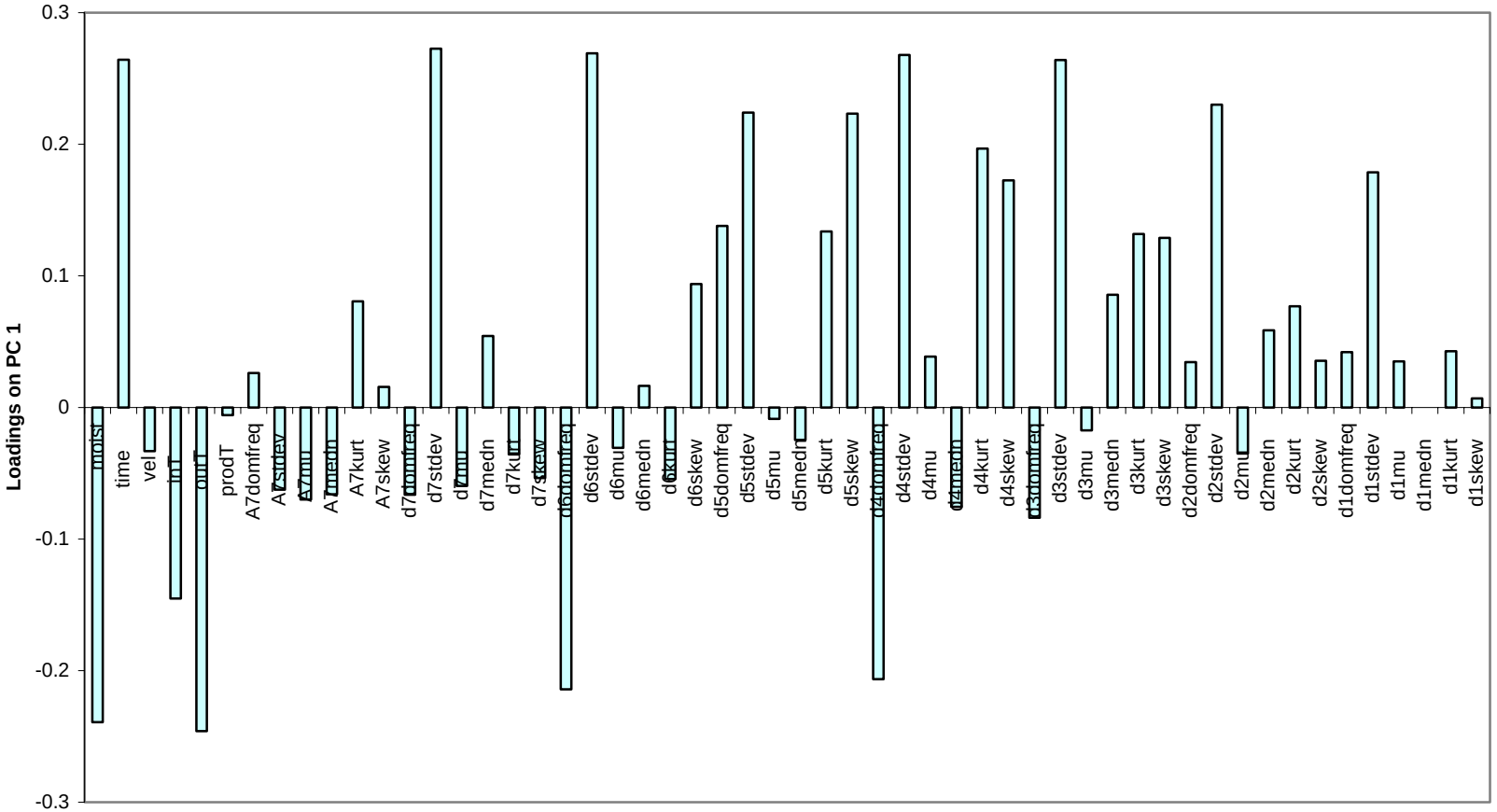

Figure I.1: Loadings plot for PC 1 from a MPCA model of the initial equilibrium region of the batch drying process.

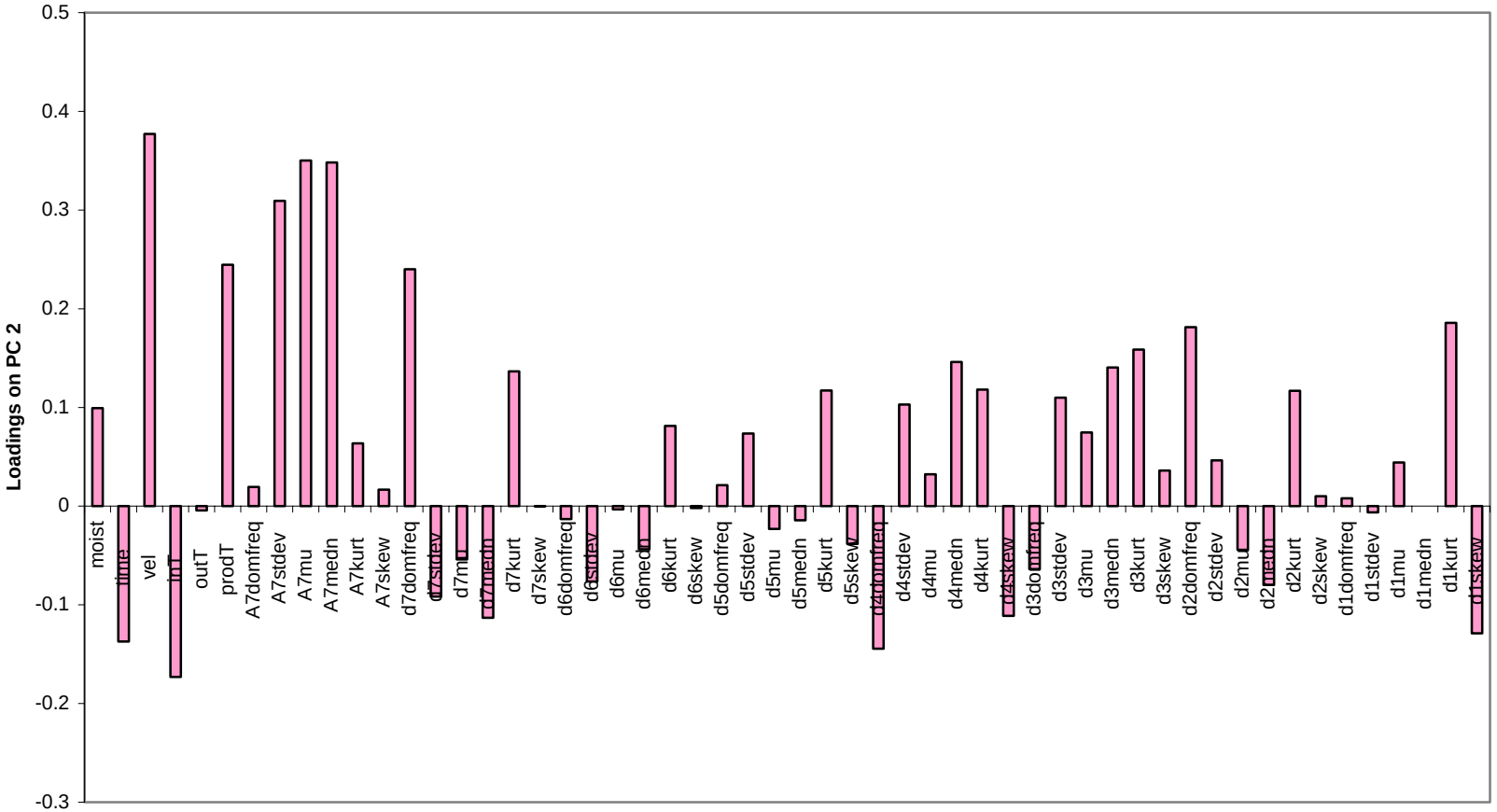

Figure I.2: Loadings plot for PC 2 from a MPCA model of the initial equilibrium region of the batch drying process.

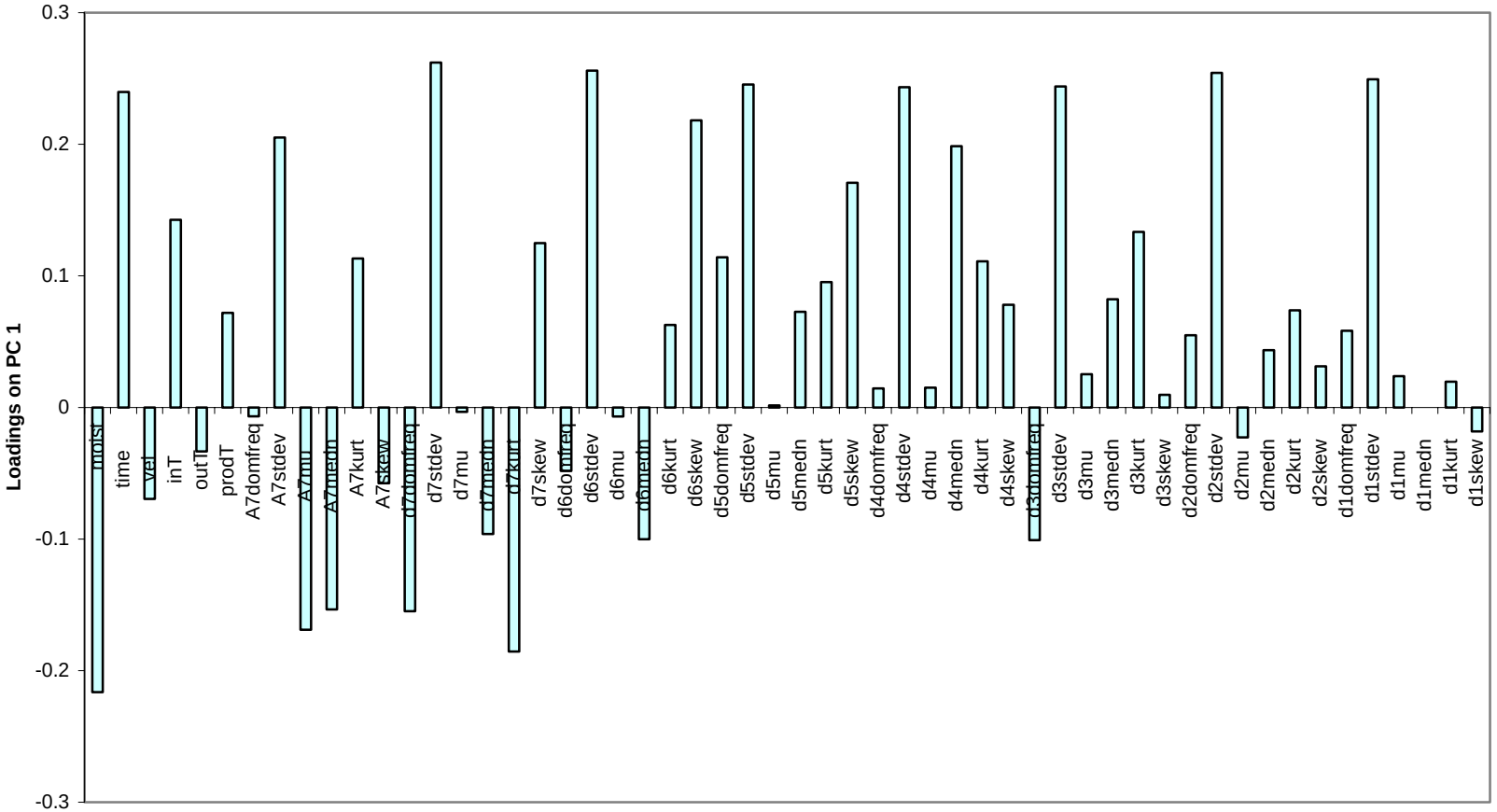

Figure I.3: Loadings plot for PC 1 from a MPCA model of the constant-rate region of the batch drying process.

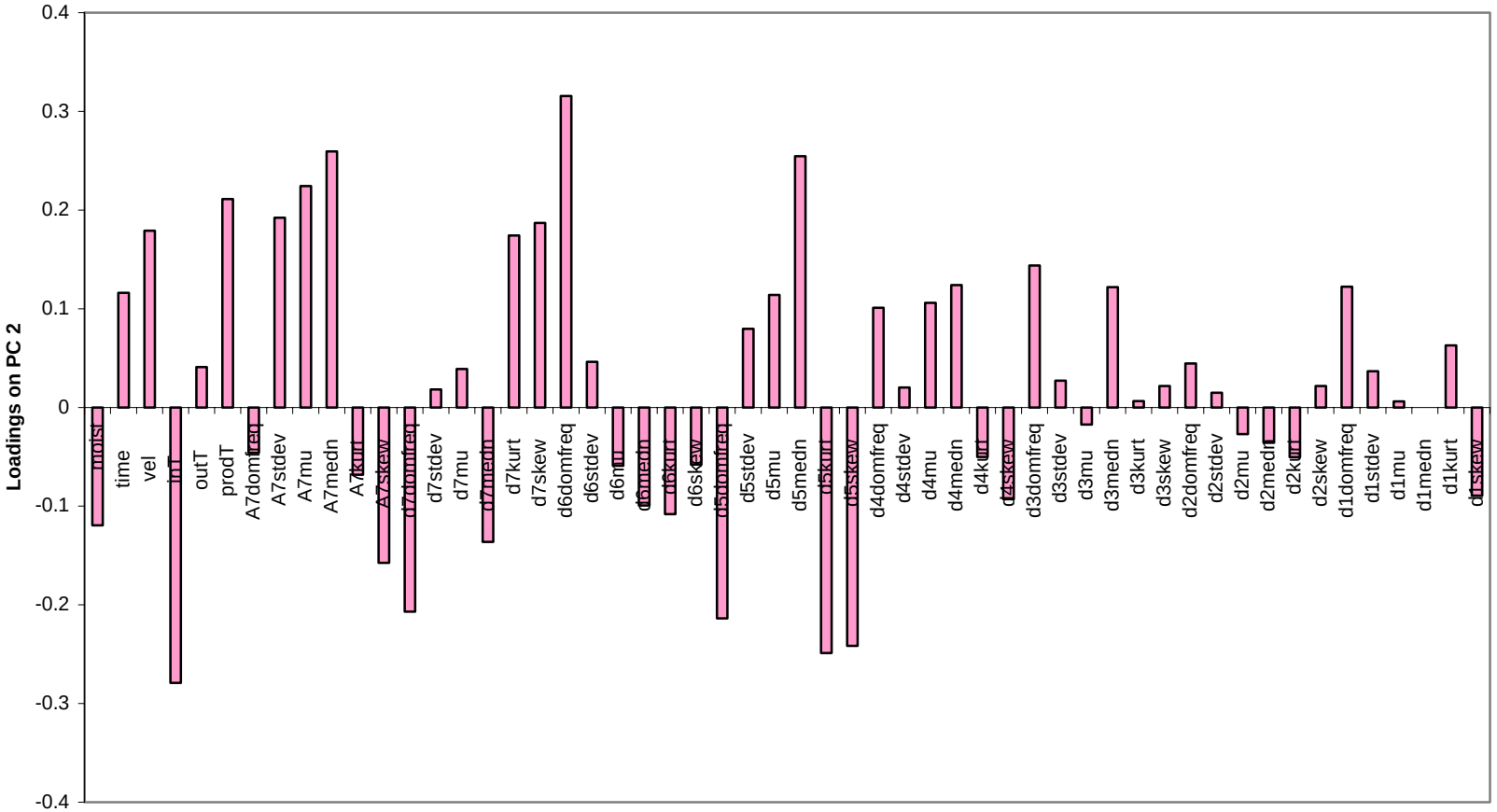

Figure I.4: Loadings plot for PC 2 from a MPCA model of the constant-rate region of the batch drying process.

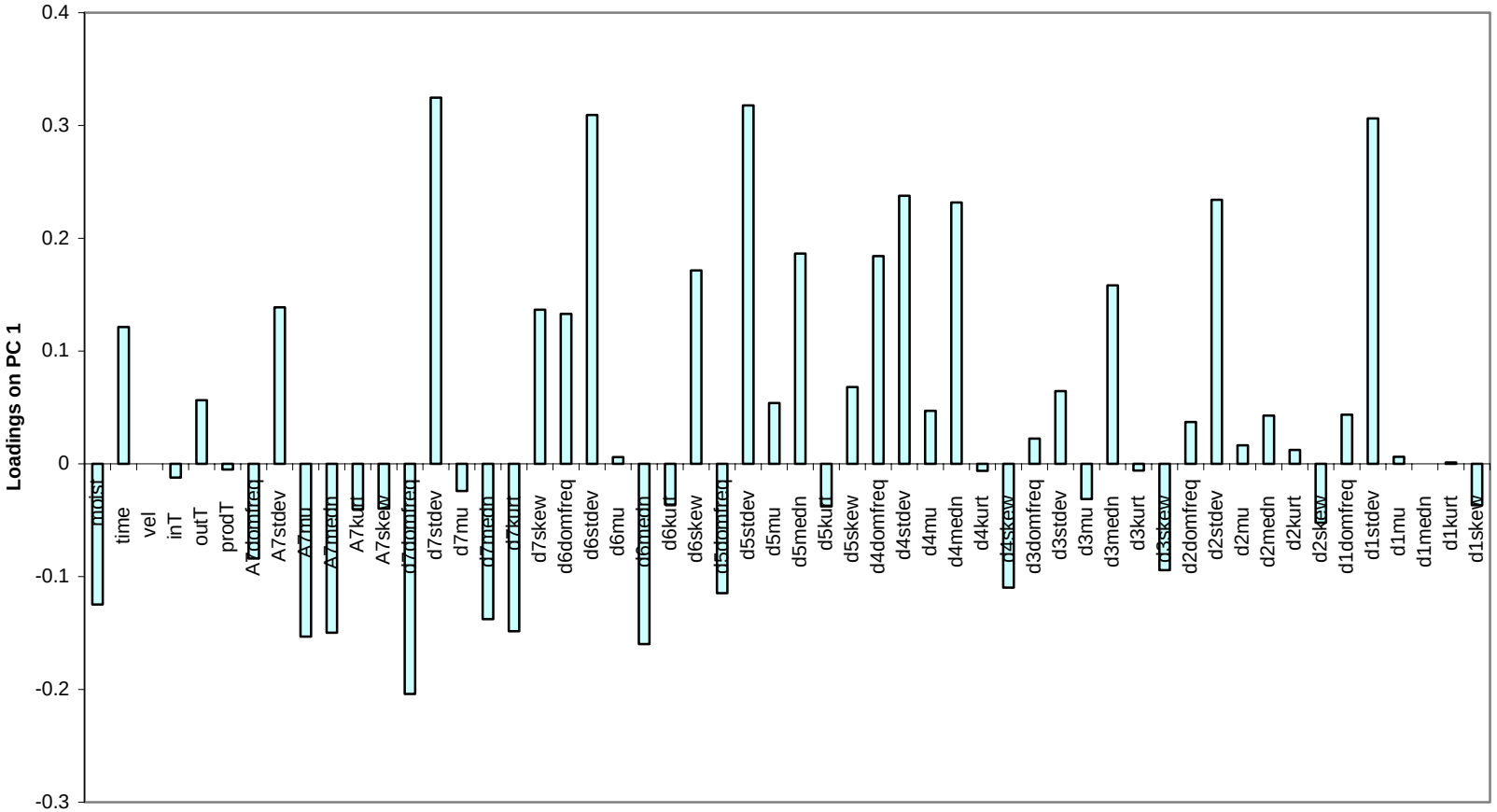

Figure I.5: Loadings plot for PC 1 from a MPCA model of the falling-rate region of the batch drying process.

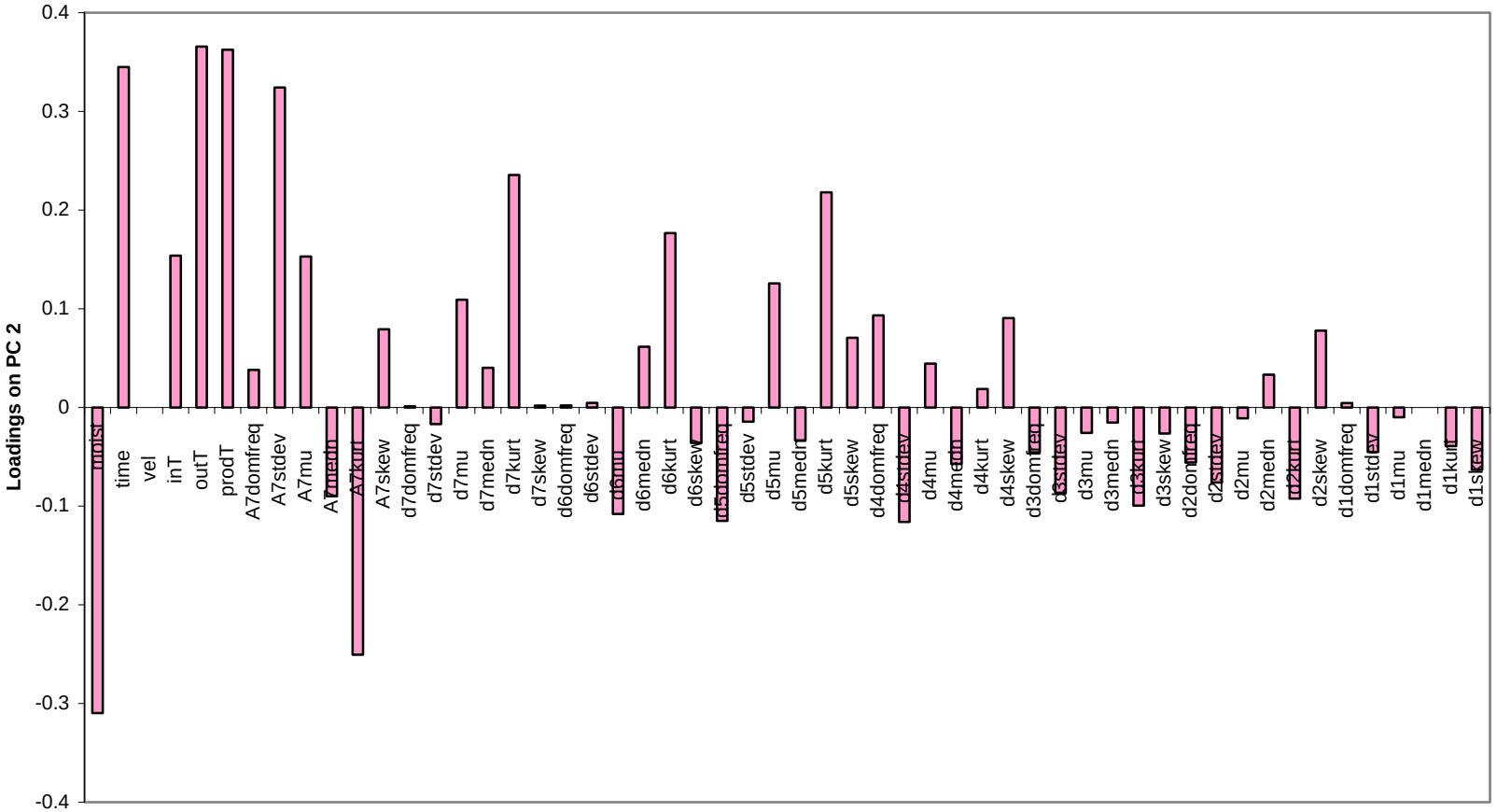

Figure I.6: Loadings plot for PC 2 from a MPCA model of the falling-rate region for the batch drying process.

APPENDIX J

REGRESSION VECTORS FROM PLS MODELS OF FACTORIAL TEST VARIABLES

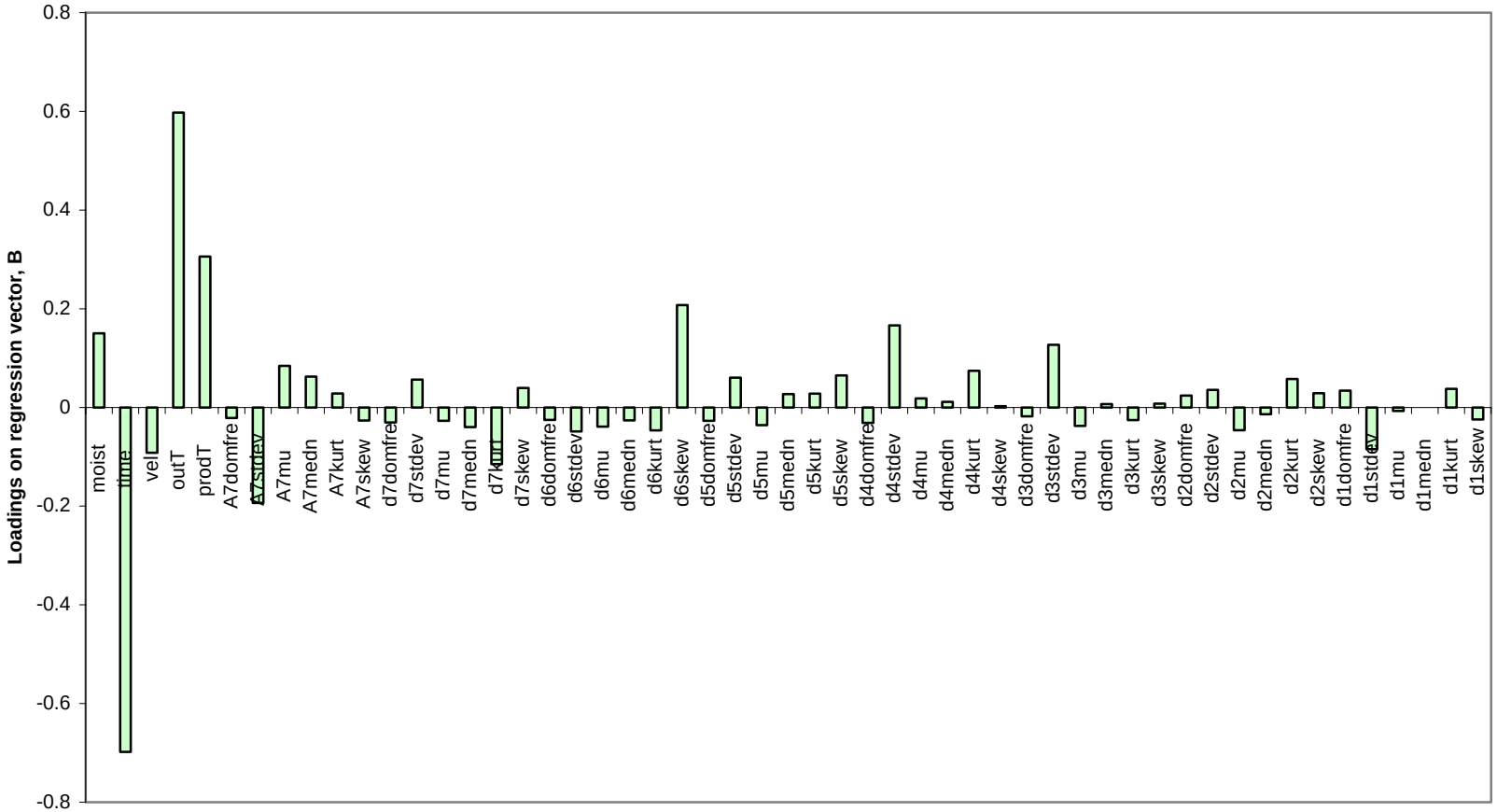

Figure J.1: Loadings plot for a regression vector from a PLS model of inlet temperature to the fluid bed dryer during the batch drying process.

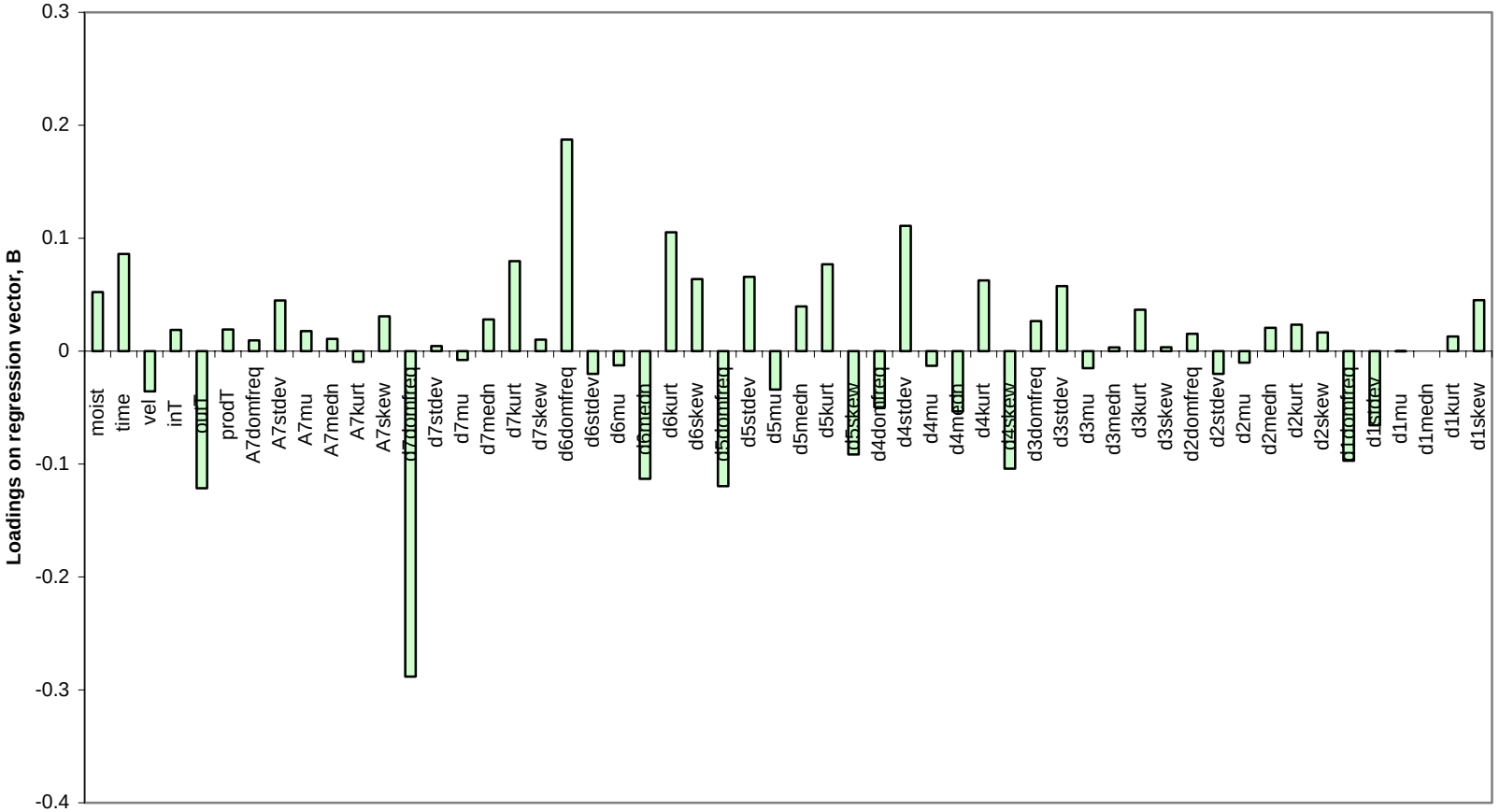

Figure J.2: Loadings plot for a regression vector from a PLS model of initial batch mass in the fluid bed dryer.

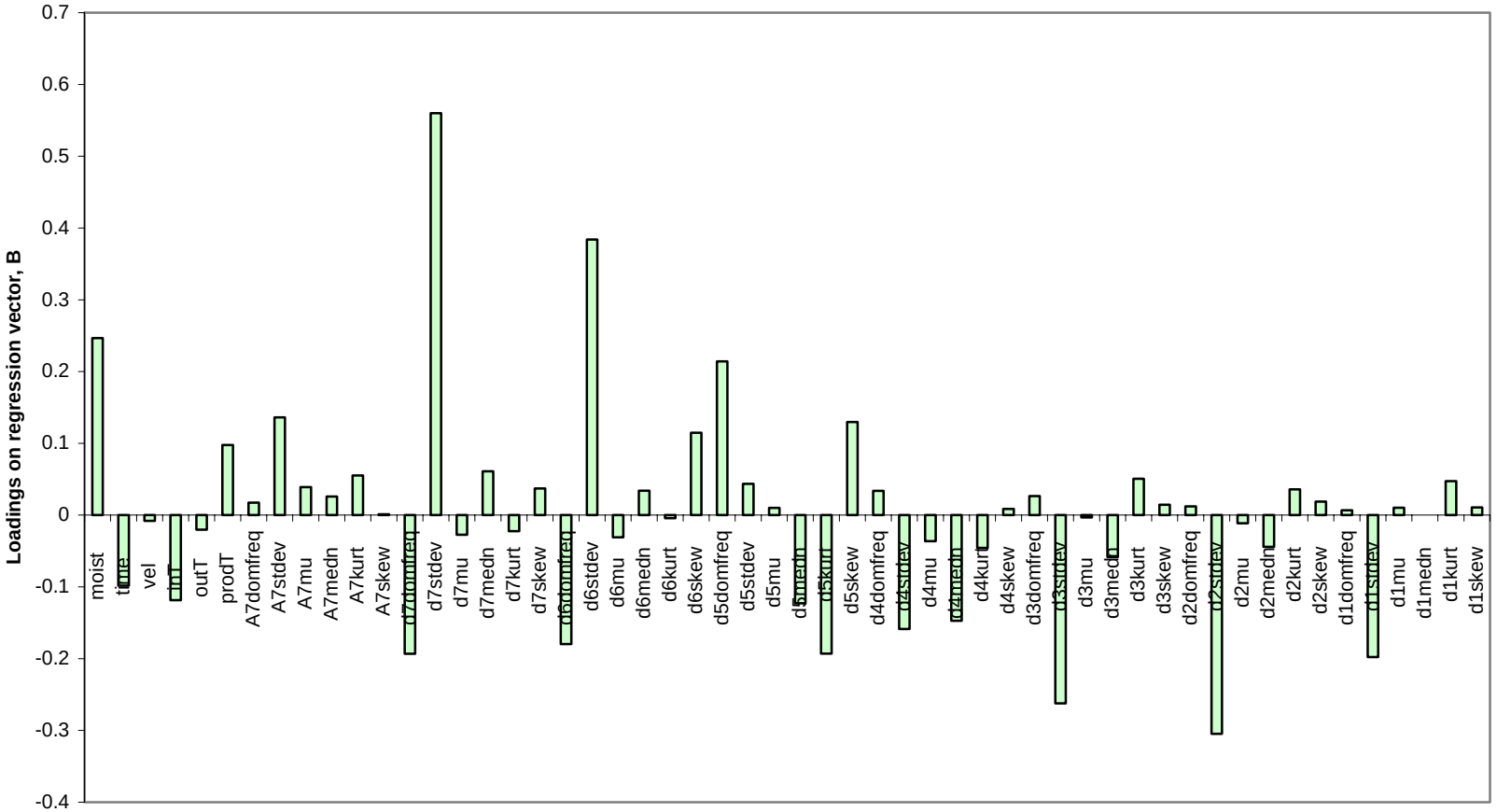

Figure J.3: Loadings plot for a regression vector from a PLS model of pressure sensor position in the fluid bed dryer.

APPENDIX K

MOISTURE MODEL WEIGHT AND REGRESSION VECTOR LOADING PLOTS

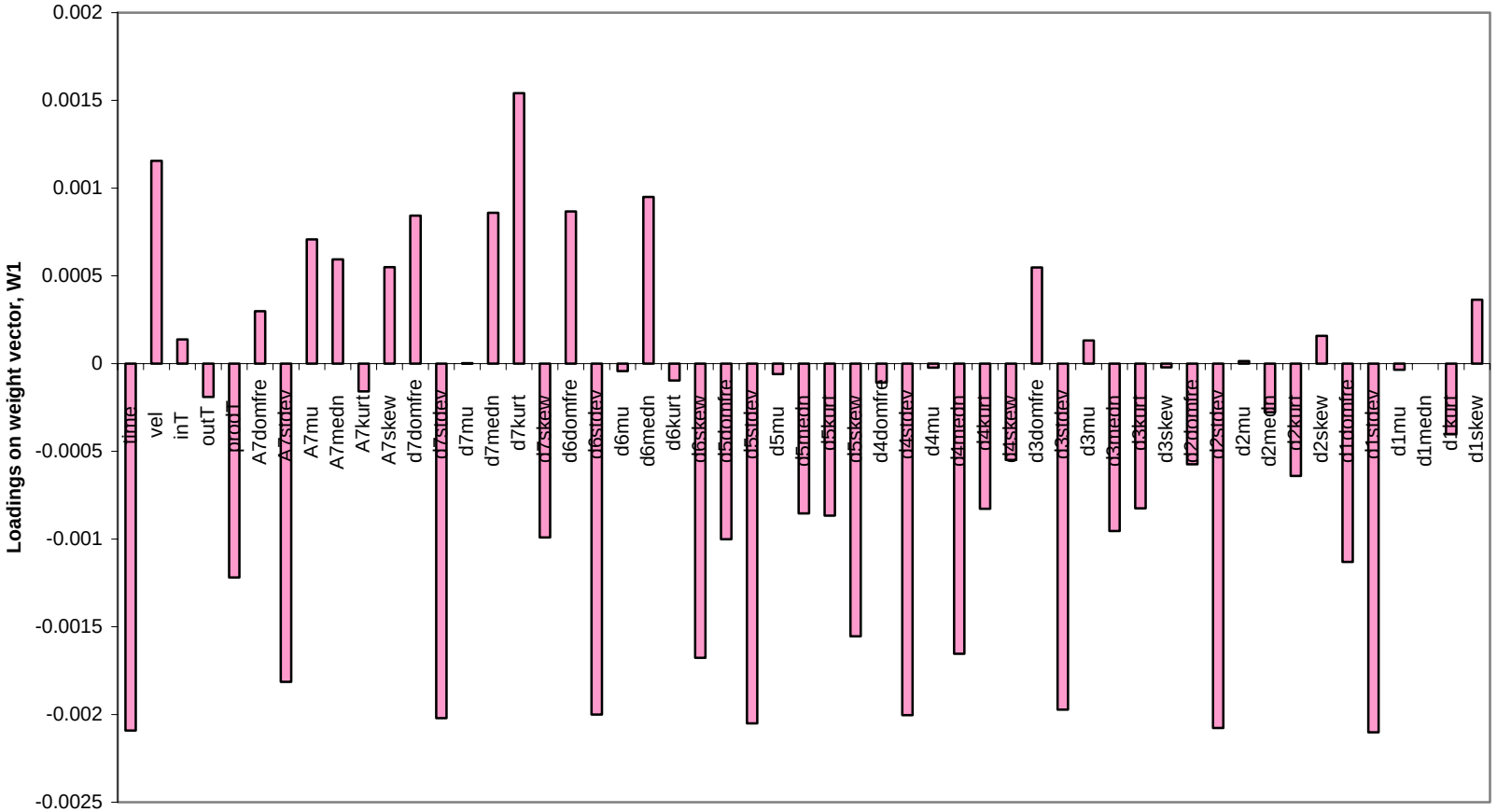

Figure K.1: Loadings plot for the dominant weight vector in a PLS model constructed to predict moisture throughout the batch drying process.

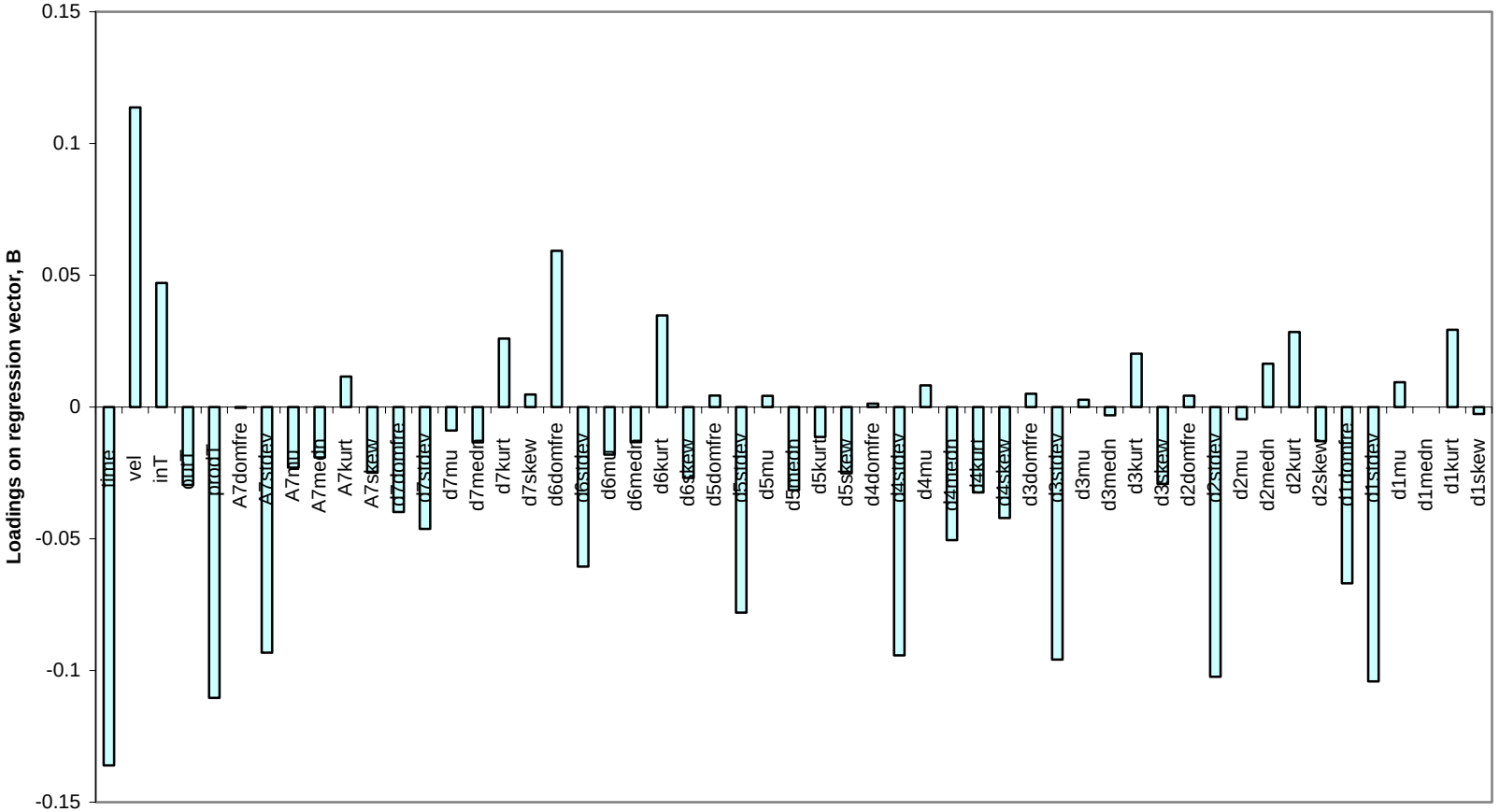

Figure K.2: Loadings plot for a regression vector from a PLS model constructed to predict moisture throughout the batch drying process.