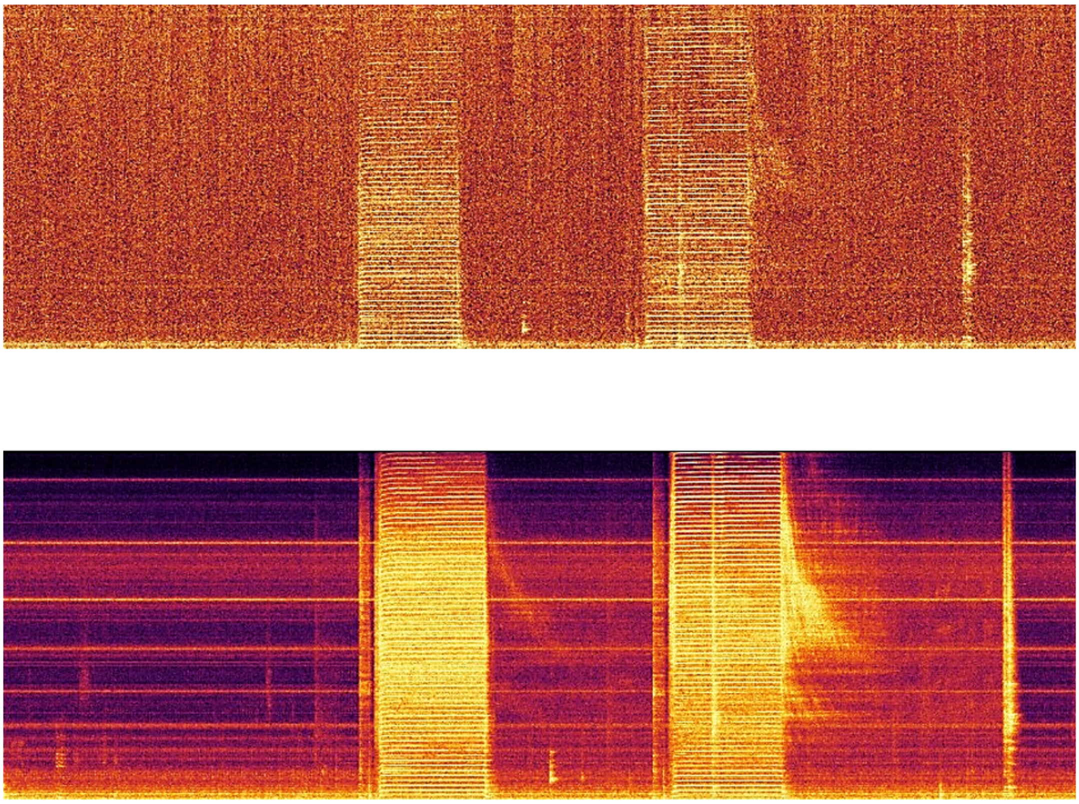
CHALMERS
UNIVERSITY OF TECHNOLOGY



ACOUSTIC EMISSION MONITORING OF BLENDING IN CONTINUOUS DIRECT COMPRESSION

Estimating residence mass through acoustic analysis

Varshaa Jagadeesan Ashok Kumar

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Master's thesis 2026

Acoustic emission monitoring of blending in continuous direct compression

Estimating residence mass through acoustic analysis

VARSHAA JAGADEESAN ASHOK KUMAR

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Division of Biomedical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2026

Acoustic emission monitoring of blending in continuous direct compression
Estimating residence mass through acoustic analysis
VARSHAA JAGADEESAN ASHOK KUMAR

Supervisors: Mats Josefson, AstraZeneca and Muhammad Ahmer, AstraZeneca
Examiner: Andreas Fhager, Department of Electrical Engineering

Master's Thesis 2026
Department of Electrical Engineering
Division of Biomedical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Spectrogram representation of acoustic measurements recorded during the blending process in a continuous direct compression (CDC) line

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2026

Acoustic emission monitoring of blending in continuous direct compression
Estimating residence mass through acoustic analysis
VARSHAA JAGADEESAN ASHOK KUMAR
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Continuous direct compression (CDC) is increasingly used in pharmaceutical manufacturing for the continuous production of oral solid dosage forms. Within a CDC line, continuous blending is an important processing step because the amount of material retained in the blender affects material transport and residence behaviour. This thesis investigates whether acoustic measurements recorded during continuous blending contain information that can be related to residence mass.

Acoustic signals covering different frequency ranges were recorded during six blender experiments performed at different rotational speeds and throughputs. Time-frequency analysis, singular value decomposition, and multivariate modelling were used to characterise the signals and assess the influence of operating conditions. Partial least squares discriminant analysis achieved 97.8% classification accuracy using frequency-band features and 100% using a higher-resolution spectral representation. Multi-output partial least squares regression further showed that both rotational speed and throughput were strongly represented in the acoustic measurements.

Residence mass was modelled using gravimetric measurements from each experimental run and evaluated using leave-one-run-out validation. High-dimensional spectral models produced strong training fits but poor prediction of omitted runs. Reducing the acoustic feature set and one PLS latent substantially improved cross-run generalisation. The best acoustic-only model achieved a training R^2 of 0.940, a LORO R^2 of 0.854, and a LORO RMSE of 266 g. Engineered process variables showed strong individual correlations with residence mass but did not improve LORO performance. The modelling procedure was also applied to separately acquired recordings using a different high-frequency measurement setup. Resampling reduced a systematic prediction offset, while a reference gravimetric measurement was required for bias correction. Overall, the results show that acoustic measurements contain information associated with both blender operating conditions and residence mass, while also highlighting the importance of model complexity, run-level validation, block selection, and measurement consistency when applying the approach to continuous pharmaceutical blending.

Keywords: Continuous direct compression, powder blending, residence mass, acoustic monitoring, partial least squares, LORO validation, pharmaceutical manufacturing.

Acknowledgements

I would like to express my heartfelt gratitude to my supervisors, Mats Josefson and Muhammad Ahmer, at AstraZeneca for their guidance, patience, and thoughtful feedback throughout this project. Their willingness to share their knowledge and perspectives has been invaluable to both the development of this work and my own learning. The Product and Process Control team has been an important part of this experience through their collaboration, valuable input, and the welcoming environment they created, making me feel truly part of the team.
My sincere thanks go to Andreas Fhager, Xuezhi Zeng, and the professors at Chalmers University of Technology for the knowledge, guidance, and strong academic foundation they have provided throughout my studies. Finally, I am grateful to my peers for the many discussions, encouragement, and shared experiences that have made my journey as a student at Chalmers both rewarding and memorable.

Varshaa Jagadeesan Ashok Kumar, Gothenburg, 08 2026

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order.

AE	Acoustic Emission
ANN	Artificial Neural Network
AUC	Area Under the Curve
CDC	Continuous Direct Compression
CSV	Comma-Separated Values
DAQ	Data Acquisition
FDA	U.S. Food and Drug Administration
FIR	Finite Impulse Response
FLAC	Free Lossless Audio Codec
GPR	Gaussian Process Regression
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
LORO	Leave-One-Run-Out
MAE	Mean Absolute Error
MRT	Mean Residence Time
PAT	Process Analytical Technology
PCA	Principal Component Analysis
PLS	Partial Least Squares
PLS-DA	Partial Least Squares Discriminant Analysis
PSD	Power Spectral Density
RMS	Root Mean Square
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
RPM	Revolutions per Minute
RTD	Residence Time Distribution
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
SVR	Support Vector Regression
TDMS	Technical Data Management Streaming
TEDS	Transducer Electronic Data Sheet
t-SNE	t-Distributed Stochastic Neighbor Embedding
VIP	Variable Importance in Projection

Nomenclature

Below lists the principal mathematical symbols used throughout this thesis.

Process and Residence Quantities

$M(t)$	Residence mass in the blender at time t.
M	Steady-state residence mass.
$\dot{m}_{\text{in}}(t)$	Inlet mass flow rate at time t.
$\dot{m}_{\text{out}}(t)$	Outlet mass flow rate at time t.
$\dot{m}$	Mass flow rate through the blender.
τ	Mean residence time.
$C(t)$	Tracer concentration measured at the blender outlet.
$E(t)$	Normalised residence time distribution.

Acoustic Signal and Spectral Quantities

$x(t)$	Time-domain acoustic signal.
$x[n]$	Discrete acoustic signal.
f	Frequency.
$\hat{S}_{xx}(f)$	Estimated power spectral density of signal x.
P_B	Mean logarithmic power spectral density within frequency band B.
N_B	Number of frequency bins contained in frequency band B.
β_0	Intercept of the logarithmic spectral regression.
β_1	Spectral slope coefficient.
ϵ	Residual term in the spectral slope regression.
$X(t_0, f)$	Short-time Fourier transform of the acoustic signal at time position t_0 and frequency f.

t_0	Time position of the analysis window in the short-time Fourier transform.
$w(t - t_0)$	Time-shifted window function used in the short-time Fourier transform.

Multivariate Analysis

$\mathbf{X}$	Data matrix used in multivariate analysis.
$\mathbf{U}$	Matrix containing left singular vectors.
$\mathbf{\Sigma}$	Diagonal matrix containing singular values.
$\mathbf{V}$	Matrix containing right singular vectors.
$\mathbf{T}$	Principal component score matrix.
$\mathbf{P}$	Principal component loading matrix.
R^2	Coefficient of determination.
Q^2	Cross-validated explained variance of the response.
r	Pearson correlation coefficient.
r_w	Weighted Pearson correlation coefficient.

Signal Alignment

t_{AE}	Time coordinate of the high-frequency acoustic recording.
t_{FLAC}	Time coordinate of the FLAC recording.
α	Slope of the linear time mapping between the AE and FLAC recordings.
β	Intercept of the linear time mapping between the AE and FLAC recordings.

Weighting and Feature Selection

m_i	Measured residence mass associated with experimental run i.
w_i	Density weight assigned to experimental run i.
d_i	Local density of residence mass observations around run i.
h	Bandwidth used in the local density calculation.
w_{ij}	Confidence weight assigned to block j of run i.

d_{ij}	Euclidean distance of block j from the feature centroid of run i.
$\tilde{d}_i$	Median block-to-centroid distance for run i.
κ	Threshold multiplier used for rejection of block-level prediction outliers based on the median absolute deviation; $\kappa = 2.0$ in this work.
$\bar{x}_w$	Weighted mean of predictor x.
$\bar{y}_w$	Weighted mean of response y.

Engineered Process Features

$\bar{\omega}$	Mean blender rotational speed within an analysis block.
T_m	Blender motor torque.
$\bar{T}_m$	Mean blender motor torque within an analysis block.
$\bar{\dot{m}}_k$	Mean mass flow rate of feeder k within an analysis block.
$\dot{m}_{\text{tot}}$	Total mean mass flow rate from the active feeders.
W_k	Net weight of feeder k.
$\dot{W}_k$	Rate of change of the net weight of feeder k.
P_{f1}	Engineered torque, speed, and throughput ratio.
P_{f2}	Engineered flow and weight-change ratio.
P_{f3}	Engineered dynamic variability ratio.
μ_x	Mean of variable x.
σ_x	Standard deviation of variable x.
$\text{CV}(x)$	Coefficient of variation of variable x, defined as σ_x/μ_x.

Prediction and Bias Correction

$\hat{m}_{\text{raw}}$	Raw predicted residence mass.
$\hat{m}_{\text{corrected}}$	Residence mass prediction after additive bias correction.
$\bar{\hat{m}}_{\text{ref}}$	Mean raw residence mass prediction for the reference recording.
m_{ref}	Gravimetrically measured residence mass of the reference recording.
b	Additive prediction bias calculated from the reference recording.
p_i	Individual block-level residence mass prediction.

$\tilde{p}$	Median of the block-level residence mass predictions.
$\mathrm{MAD}(p)$	Median absolute deviation of the block-level predictions.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Introduction

Continuous manufacturing is increasingly used in pharmaceutical production as an alternative to conventional batch processing. In a continuous system, material feeding, processing, and product formation are integrated within the same manufacturing line. This creates opportunities for flexible production, real-time monitoring, and improved process control, while also increasing the need to understand how variations and disturbances propagate between connected unit operations [1, 2].

Continuous direct compression (CDC) is one approach for the continuous production of pharmaceutical tablets. In a typical CDC line, active pharmaceutical ingredients and excipients are supplied by gravimetric feeders, mixed in a continuous blender, and transferred to downstream operations leading to tablet compression. The blending stage is important because variations introduced during feeding must be sufficiently distributed before the material reaches subsequent processing steps [4, 5].

The behaviour of a continuous powder blender depends on several interacting factors. Impeller speed, throughput, blade configuration, powder properties, and equipment design can influence material transport, residence behaviour, and mixing performance [6, 8, 5]. One quantity that describes the internal state of the blender is the amount of material present during operation. Studies of inclined linear blenders refer to this quantity as *residence mass*, while the wider continuous blending literature also uses the terms *hold-up mass* and *mass holdup* [11, 12]. In this thesis, the term *residence mass* is used throughout.

Residence mass is not determined by throughput alone. Experimental and modelling studies have shown that rotational speed, feed rate, powder properties, blade configuration, and blender geometry can influence the amount of material retained within the system [6, 8, 11, 12]. Particle tracking studies of an inclined linear blender further show that changes in rotational speed and hold-up affect powder motion and mixing behaviour [10]. These findings indicate that the amount of material present in the blender is closely connected to the internal dynamics of the powder.

Changes in powder motion provide a possible basis for indirect process monitoring. Particle impacts and interactions can generate mechanical disturbances that propagate through the equipment structure or surrounding medium and can be detected using acoustic or vibration-based measurements [18, 17]. Previous studies have shown that such signals can contain information about particle concentration, collision behaviour, and material flow [17, 20, 21].

There is also direct evidence that passive acoustic measurements can provide useful information during pharmaceutical powder processing. Allan et al. [15] monitored pharmaceutical powders during mixing and found that acoustic emission intensity changed with powder mass, impeller speed, particle density, and particle size. Selected frequency ranges were also used to follow the progression of the mixing process. Crouter and Briens [16] investigated passive acoustic emissions from a V-blender and found that the measured vibrations were influenced by fill level, particle size, equipment scale, and powder flowability. These studies show that powder quantity can affect the measured vibroacoustic response, while also demonstrating that the signal is influenced by several other process and material variables.

This dependence on multiple variables makes quantitative prediction more challenging than detecting a change in process state. A signal characteristic associated with residence mass may also change because of impeller speed, material flow, particle properties, equipment response, or the measurement configuration [15, 18, 20, 21]. A relationship identified under one operating condition may therefore not remain valid when the process conditions change. For a monitoring approach to be useful across multiple operating points, the extracted signal information must retain a meaningful relationship with residence mass despite these sources of variation.

Recent studies have shown that residence mass or the corresponding fill level can be predicted in continuous blenders using combinations of process parameters and powder properties. Jones-Salkey et al. [11], for example, developed data-driven models for steady-state fill level in an inclined linear blender using information including rotational speed, feed rate, blade configuration, and material characteristics. It appears that comparatively little work has addressed the direct estimation of residence mass from passive acoustic measurements during continuous pharmaceutical blending.

This thesis investigates whether acoustic measurements obtained during continuous powder blending contain information that can be used to estimate residence mass across different operating conditions. The recorded signals are also examined to determine how changes and transient events are represented in the time and frequency domains. The work therefore combines process understanding, acoustic signal analysis, and multivariate modelling to evaluate passive sensing as a potential approach for monitoring the internal state of a continuous blender.

1.2 Aim

The aim of this thesis is to investigate whether acoustic measurements obtained during continuous powder blending can be used to estimate the residence mass within the blender.

The study focuses particularly on identifying informative signal characteristics and determining whether the resulting predictive relationship remains useful across changes in blender operating conditions.

1.3 Objectives

To address the aim of this thesis, the work is structured around four main objectives that progress from understanding the measured signals to developing and evaluating a predictive model. Together, these objectives consider both the information contained in the acoustic measurements and the ability to use this information across different operating conditions.

The objectives of this thesis are:

1. To characterise the acoustic signals recorded during continuous powder blending and investigate how they respond to changes and disturbances in process operation.
2. To extract and evaluate relevant time-domain and frequency-domain features, and identify those associated with residence mass and blender operating conditions.
3. To develop a predictive model for estimating residence mass from the acoustic measurements.
4. To evaluate the generalisation of the developed approach across different blender speeds, throughputs, experimental runs, and an additional dataset acquired using a different acoustic measurement configuration.

1.4 Limitations

The study is based on a limited set of controlled blending experiments and therefore covers only the operating conditions represented in the available dataset. The developed relationships should not be interpreted as universal models for other formulations, blender geometries, or operating ranges without further experimental validation.

The acoustic measurements are also specific to the sensors, placement, equipment structure, and acquisition systems used in the experiments. Previous studies show that particle-generated signals depend on both the physical source and the measurement pathway, while sensor coupling and installation can affect the recorded frequency response [18, 22]. Therefore, relationships identified for one acquisition configuration cannot be assumed to transfer directly to another without additional calibration or validation.

A second set of recordings was acquired later using the same blender but a different high-frequency acoustic measurement configuration. These recordings were used to evaluate the developed model on separately acquired data.

Finally, the work focuses on signal characterisation and residence-mass estimation. Although the results may provide information relevant to future process monitoring and control, optimisation of blender operation and implementation of a closed-loop control strategy are outside the scope of this thesis.

1.5 Use of Generative AI

Generative AI tools were used in this thesis in accordance with the Chalmers regulations for the use of AI tools in thesis work [42]. The use of ChatGPT [43] was limited to supporting the writing and understanding process.
The tool was used for the following purposes:

1. Used in addition to literature to understand difficult topics and assisting in interpreting relationships between concepts.
2. Identifying opportunities to make the text more concise and improve the clarity of the written explanations.
3. Identifying and correcting grammatical and language inconsistencies in the text.

2

Theory

2.1 Continuous Direct Compression and Continuous Blending

Continuous direct compression (CDC) integrates the preparation of a powder blend and its conversion into tablets within a continuous production line. Although the exact equipment configuration can vary, a typical system consists of gravimetric feeders, a continuous blender, and a tablet press, with an additional lubrication stage included in some configurations [23, 4]. Material is therefore transferred continuously between unit operations rather than being processed as separate batches. A schematic representation of the main unit operations in a typical CDC line is shown in Figure 2.1.

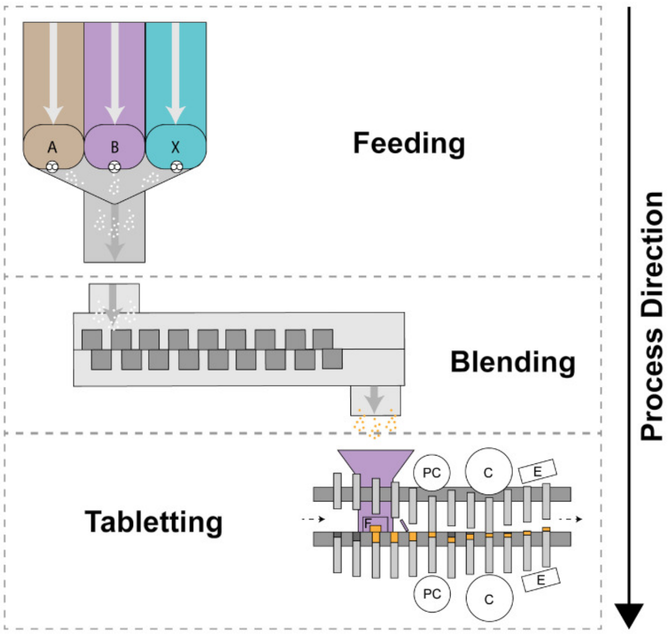


Figure 2.1: continuous direct compression, showing the feeding, blending, and tabletting stages. Reproduced from Jones-Salkey et al. [10] under the Creative Commons Attribution 4.0 licence.

The feeding stage controls the amount and proportion of each formulation component entering the process. Loss-in-weight feeders are commonly used for this purpose and continuously dispense the required materials at specified mass flow rates. Their

performance depends not only on the selected flow rate but also on powder properties and the amount of material remaining in the feeder. Studies of continuous pharmaceutical feeding have shown that variables such as flow rate, hopper level, and bulk density influence material transport and residence behaviour within screw feeders [24]. Variability introduced at this stage can subsequently propagate into the blending operation and downstream processing.
The continuous blender (refer figure 2.2 receives the individual material streams and distributes them to produce a sufficiently uniform mixture. Its behaviour is controlled by both the incoming material flow and the mechanical action of the impeller. Impeller speed and blade configuration influence the transport and mixing of the powder, while throughput determines the rate at which material passes through the equipment [4, 5]. These variables are therefore coupled rather than independent. For example, modifying the orientation or number of mixing blades can change axial transport through the blender and consequently alter the amount of material retained and the time available for mixing [5].

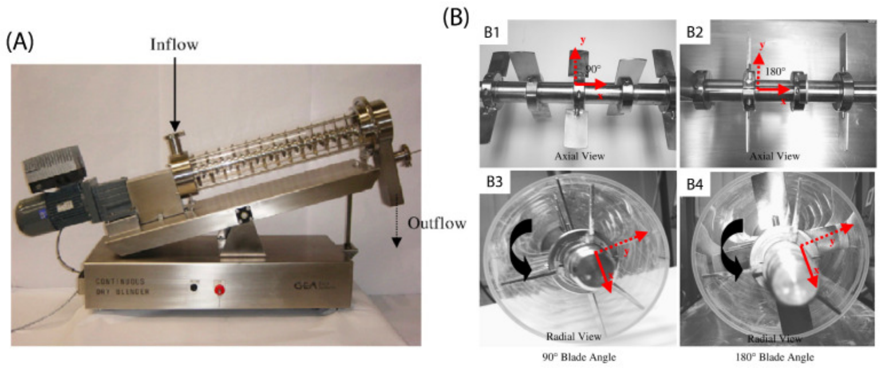


Figure 2.2: Example of a linear continuous powder blender, showing the blender profile and the agitator shaft with blade orientations. Reproduced from Jones-Salkey et al. [10], based on the blender reported by Portillo et al. [64].

The interaction between transport and mixing is important because a continuous blender must perform two functions simultaneously. It must distribute the different formulation components at the particle scale while also reducing temporal variations entering from the feeders. Van Snick et al. [4] showed that impeller configuration and rotational speed could be used to modify residence behaviour and the ability of a continuous blender to smooth steady-state feeding variability. Earlier continuous direct compression experiments also demonstrated that changing impeller speed and total feed rate can influence the uniformity and downstream properties of the manufactured tablets [23]. The operating conditions of the blender can therefore affect both its internal powder dynamics and the quality of the material leaving the unit.
Variations during continuous operation are not limited to steady fluctuations. Short interruptions, feeder refilling events, changes in material flow, or other disturbances can introduce temporary deviations from the normal process state. Karttunen et al. [13] investigated impulse and step disturbances introduced at the feeding stage of a CDC line and showed that their downstream effect depended on the residence

behaviour of the connected equipment. Increasing production rate reduced the available residence time and resulted in less smoothing of the introduced disturbances. This illustrates that the response of a continuous manufacturing line depends on both the disturbance itself and the dynamic behaviour of the unit operations through which the material travels.
For the blender, these dynamics are closely connected to the amount of powder present within the equipment and to the distribution of times for which material remains inside it. Residence mass, mean residence time, and residence time distribution therefore provide complementary descriptions of the internal process state.

2.2 Residence Mass, Residence Time and Blender Dynamics

Residence mass describes the amount of powder present within the blender at a given operating condition, whereas residence time describes how long material remains inside the equipment. The residence time distribution (RTD) further accounts for the fact that individual particles or material elements do not all spend the same amount of time inside the blender [7, 11].
For a continuous blender, the change in residence mass can be described using a transient material balance:

$$\frac{dM(t)}{dt} = \dot{m}_{\mathrm{in}}(t) - \dot{m}_{\mathrm{out}}(t), \tag{2.1}$$

where $M(t)$ is the mass of material inside the blender, and $\dot{m}_{\mathrm{in}}(t)$ and $\dot{m}_{\mathrm{out}}(t)$ are the inlet and outlet mass flow rates, respectively. During steady operation, the average inlet and outlet flow rates approach the same value and the material inventory becomes approximately constant:

$$\frac{dM}{dt} \approx 0, \qquad \dot{m}_{\mathrm{in}} \approx \dot{m}_{\mathrm{out}}. \tag{2.2}$$

Under these conditions, the mean residence time can be related to the residence mass and material throughput as

$$\overline{\tau} = \frac{M}{\dot{m}}, \tag{2.3}$$

where $\overline{\tau}$ is the mean residence time, M is the steady-state residence mass, and $\dot{m}$ is the mass flow rate through the blender [11]. This relationship shows why residence mass and residence time are related but should not be treated as the same quantity. Two operating conditions can have different residence masses while also differing in throughput, resulting in a different relationship between the amount of material retained and the time spent in the equipment.
The RTD provides a more complete description of material transport by describing the distribution of residence times rather than only their average value. It is commonly determined experimentally by introducing a tracer at the blender inlet and monitoring its concentration at the outlet. For an impulse tracer experiment, the normalized RTD can be expressed as [7]

$$E(t) = \frac{C(t)}{\int_0^\infty C(t)\,dt}, \tag{2.4}$$

where $C(t)$ is the measured tracer concentration at the outlet and $E(t)$ is the residence time distribution. The corresponding mean residence time is obtained from the first moment of the distribution:

$$\overline{\tau} = \int_0^\infty tE(t)\,dt. \tag{2.5}$$

The width and shape of the RTD contain additional information about axial transport and dispersion within the blender. Gao et al. [7] showed that changes in impeller speed, feed rate, and blade configuration alter RTD behaviour in a continuous powder mixer. In their experiments, lower blade speeds generally produced longer mean residence times, while increasing the rotational speed increased axial transport and shortened the time spent in the mixer.

Residence mass is also strongly influenced by impeller operation. Vanarase and Muzzio [6] reported a pronounced reduction in powder hold-up as impeller rotational speed increased. Similar behaviour was observed by Osorio and Muzzio [8], where both hold-up and mean residence time decreased with increasing rotational speed. More recent experiments and numerical studies have reported the same overall tendency in different continuous blender configurations [25, 26, 12]. A higher impeller speed generally promotes faster axial transport of powder through the blender, reducing the amount of material retained and shortening its residence time.

The effect of throughput is less straightforward because changes in material flow can affect residence mass and residence time differently. Bekaert et al. [25] observed that increasing throughput increased hold-up mass while reducing mean residence time in their continuous blender. This occurs because more material may be present in the equipment even though it is transported through the system at a higher rate. Jolliffe et al. [12] further showed that the response of mass holdup to throughput depends on powder properties and outlet geometry. The relationship between residence mass and throughput should therefore be considered together with the mechanical configuration and material behaviour rather than treated as a universal linear dependence.

Blade configuration provides another means of modifying powder transport. Forward-oriented blades promote axial movement, whereas arrangements containing backward-pushing elements can increase recirculation and extend the time available for mixing. Experimental studies have shown that changing blade configuration can alter residence mass, RTD, and the number of blade passes experienced by the powder [6, 8, 25]. The relative importance of these effects can also change with rotational speed.

Powder properties introduce an additional source of variation. Jolliffe et al. [12] demonstrated that density and flow behaviour affected mass holdup, particularly when combined with different outlet weir geometries. Their results showed that materials with different flow characteristics did not respond identically to changes in throughput or blender configuration. Consequently, a residence mass associated with a particular RPM and throughput cannot be assumed to remain unchanged when the formulation or equipment geometry changes.

These relationships are important for the present study because changes in residence mass are accompanied by changes in the physical state of the powder inside the blender. Particle tracking experiments in an inclined linear blender have shown that both rotational speed and hold-up influence particle trajectories and mixing behaviour [10]. Numerical simulations similarly indicate that increasing impeller speed changes the powder flow regime while reducing hold-up and mean residence time [26]. Residence mass therefore provides more than a measure of the material inventory: it is connected to the number, distribution, and motion of particles interacting with the impeller and blender structure.

This connection provides the transition to acoustic monitoring. If operating conditions modify residence mass and particle dynamics, the resulting particle-to-particle and particle-to-wall interactions may also change the mechanical and acoustic response of the system.

2.3 Acoustic Signal Generation in Powder Blending

Powder blending involves continuous movement and interaction of particles within the equipment. As particles are transported by the impeller, they collide with other particles, the impeller, and the blender walls. These interactions generate transient mechanical forces that can excite the surrounding structure and produce measurable acoustic signals. The characteristics of the generated signal depend on the nature of the impacts as well as the subsequent transmission of the resulting mechanical energy [17, 18]. The basic mechanisms of acoustic generation during particle-particle and particle-wall impacts are illustrated in Figure 2.3.

Particle-to-wall impacts are particularly important because they directly excite the blender structure. Carson et al. [17] developed a model in which particle impacts produce flexural vibrations in a vessel wall. The resulting power spectrum depends on physical properties of the particles and on their concentration. This demonstrates that the measured frequency content is not arbitrary, but can contain information related to the particulate system that generated it. Pecorari [18] similarly described the relationship between particle impact conditions and the acoustic emission generated when solid particles strike a surface.

In a powder blender, a large number of impacts occur simultaneously and the measured signal represents the combined response of these interactions. Experimental evidence from pharmaceutical blending shows that this response changes with both powder quantity and operating conditions. Allan et al. [15] found that acoustic emission intensity during pharmaceutical powder mixing increased with powder mass, impeller speed, and particle density. Particle size also influenced the response, including the distribution of energy across frequency. Crouter and Briens [16] observed that fill level affected passive vibration measurements from a V-blender because the amount of material present influenced particle motion and momentum during collisions.

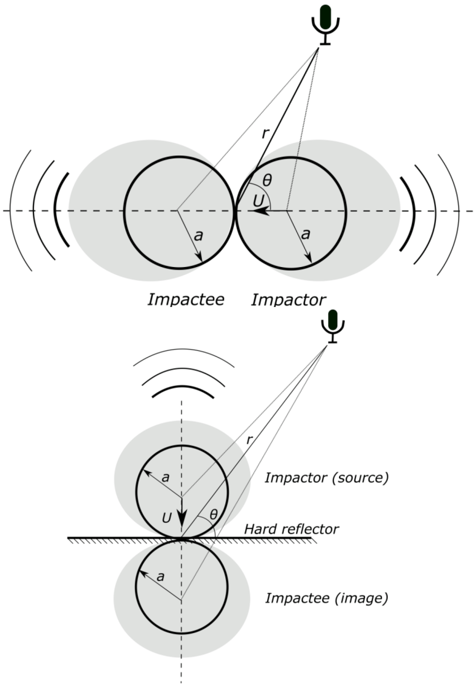


Figure 2.3: Conceptual representation of acoustic generation from particle-particle and particle-wall impacts. Mechanical energy released during collision produces acoustic waves that propagate away from the impact region. Adapted from Petrut et al. [65] under the Creative Commons Attribution 3.0 licence.

The relationship between particle motion and frequency content has also been demonstrated in controlled granular-flow experiments. Bachelet et al. [27] simultaneously measured particle motion and acoustic emissions from dry granular flows. In their experimental configuration, higher-frequency components were associated mainly with particle collisions, whereas lower-frequency components were related to more coherent motion of the granular material. This demonstrates that different forms of particle motion can contribute differently to the measured spectrum. The specific frequency ranges reported in such experiments are dependent on the particles, equipment, and measurement arrangement and should therefore not be transferred directly to a different blending system.
The recorded signal is also affected by the path between its source and the measurement point. A useful conceptual representation is

$$x(t) = s(t) * h(t) + n(t), \tag{2.6}$$

where $s(t)$ represents the signal generated by particle and equipment interactions, $h(t)$ represents the combined transmission response between the source and measurement point, $*$ denotes convolution, and $n(t)$ represents unrelated background contributions. Equation 2.6 is intended as a simplified representation rather than a complete physical model of the blender.
Transmission through granular material and the surrounding structure can alter both the amplitude and frequency content before the signal is measured. Zhai et al. [28] investigated ultrasound propagation through randomly packed granular ma-

terial and showed that packing structure and interparticle forces influenced wave velocity and dispersion, while structural disorder contributed strongly to attenuation. Although their study used actively transmitted ultrasound rather than passive signals generated during blending, it demonstrates that the state of a granular medium can influence how mechanical waves propagate through it.
These effects are relevant to residence mass estimation because changing the amount of powder within the blender can modify the number and distribution of particle interactions as well as the environment through which the generated disturbances propagate. However, acoustic response cannot be expected to depend on residence mass alone. Changes in rotational speed or throughput can modify particle velocity, collision frequency, and internal powder motion, producing additional variation in the measured signal [15, 16]. A useful acoustic indicator of residence mass must therefore be identified in the presence of these competing effects.
The measured response consequently contains information from both the powder state and the measurement pathway. How this information is captured depends on the sensing principle, sensor location, usable frequency range, and interaction between the sensor and the equipment. These aspects are considered in the following section.

2.4 Acoustic Sensing for Process Monitoring

Passive acoustic monitoring uses signals generated naturally by a process rather than introducing an external acoustic excitation. In particulate systems, these signals can arise from particle impacts, friction, equipment motion, and other mechanical interactions occurring during operation [14]. The way this activity is observed depends strongly on the sensing principle. In particular, a distinction is required between measurements obtained from the equipment structure and sound measured in the surrounding air.
Structure-borne measurements are commonly obtained using piezoelectric acoustic emission sensors or accelerometers coupled to the equipment surface. Mechanical disturbances generated by particle impacts propagate through the vessel wall before reaching the sensor. The recorded response therefore depends on both the original excitation and the structural transmission path. Pharmaceutical powder studies have successfully used this approach to follow mixing and other particulate processes [15, 16, 29]. Allan et al. [15] also showed that the frequency content recorded during powder mixing depended on the response characteristics of the acoustic transducer, demonstrating that the sensor itself forms part of the measurement system.
Airborne measurements follow a different transmission path. A microphone responds to pressure fluctuations in the surrounding air and does not require mechanical coupling to the equipment. Sound generated inside the process must therefore be transmitted through or radiated by the vessel before reaching the microphone. This approach offers the practical advantage of non-contact installation, but the measured signal can contain contributions from the blender drive, surrounding machinery, and the acoustic environment in addition to the powder-related response.
The difference between structure-borne and airborne measurements has been demonstrated experimentally. Watson et al. [29] recorded acoustic signals from both a

transducer attached to a high-shear granulator bowl and an airborne sensor. The bowl-mounted transducer showed clear changes in frequency content during granulation, whereas considerably smaller changes were detected by the airborne measurement. This illustrates that the usefulness of a sensing configuration depends on how effectively process-generated mechanical energy reaches the measurement point.

Nevertheless, airborne sound can contain useful information about particulate processes when the transmission path and signal level are suitable. Fulek et al. [30] used condenser and directional microphones positioned close to a pharmaceutical granulation vessel and demonstrated that time-frequency representations of the recordings could distinguish different process phases. Similarly, Leroy et al. [31] placed a measurement microphone close to a milling reactor and showed that changes in the recorded spectrum corresponded to changes occurring inside the vessel. These studies demonstrate that external microphones can capture process-related information without direct mechanical contact with the equipment, although the sensitivity depends on the process and measurement arrangement.

Sensor position and mechanical coupling are therefore important experimental factors. For contact sensors, changes in the interface between the transducer and the structure can modify the measured frequency response. Xu et al. [22] experimentally compared different coupling materials for piezoelectric sensors and found frequency-dependent differences in amplitude response and repeatability. This is particularly relevant when measurements obtained using different sensor installations are compared, because changes in the recorded spectrum may arise from the measurement configuration as well as from the underlying process.

The usable frequency range also depends on the sensing approach and transducer response. Powder-related information is not necessarily concentrated at a single frequency, and previous studies have extracted process information from both audible and ultrasonic regions [15, 29]. For this reason, broadband measurements can be useful when the relevant spectral regions are not known in advance. However, a wider measurement bandwidth also captures additional equipment and environmental contributions, making subsequent signal analysis necessary to separate informative patterns from unrelated components.

For residence mass monitoring, these considerations mean that the measured acoustic signal should be treated as an indirect representation of the internal powder state rather than as a direct measurement of mass. Changes in particle quantity and motion can modify the generated sound, but the recorded response is also shaped by the blender structure, transmission path, sensor characteristics, and surrounding environment. Robust estimation therefore requires signal representations that retain process-related information while reducing sensitivity to irrelevant variation.

2.5 Signal Representation and Feature Extraction

Acoustic recordings from powder processing contain information distributed across both signal amplitude and frequency. Large process changes may be visible directly in the waveform, while more subtle differences can require statistical or spectral representations. Previous pharmaceutical studies have therefore used combinations of time-domain and frequency-domain features to describe changes in particulate processes [32, 33, 19].

Time-domain features describe the amplitude distribution and temporal behaviour of a signal within a selected interval. A commonly used measure is the root mean square (RMS),

$$x_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2}, \tag{2.7}$$

where x_n is the measured signal and N is the number of samples. RMS provides a measure of the overall signal magnitude and has been used in acoustic and vibration monitoring of pharmaceutical particulate processes [32]. Other descriptors, such as standard deviation, peak amplitude, skewness, kurtosis, crest factor, and zero-crossing rate, provide complementary information about variability, asymmetry, impulsive behaviour, and signal fluctuations. In this thesis, these quantities are treated as descriptive features rather than direct measurements of specific powder properties.

Frequency-domain analysis describes how the signal is distributed across frequency. For a discrete signal $x[n]$, the discrete Fourier transform can be written as

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N}, \tag{2.8}$$

where k represents the frequency bin. The Fourier transform provides the basis for estimating the power spectral density (PSD), which represents the distribution of signal power over frequency.

For finite measurements, PSD estimates obtained from individual signal segments can contain substantial variability. Welch's method reduces this effect by dividing the signal into overlapping windowed segments, calculating a periodogram for each segment, and averaging the resulting estimates [34]. A general representation is

$$\hat{S}_{xx}(f) = \frac{1}{K}\sum_{r=1}^{K} P_r(f), \tag{2.9}$$

where $P_r(f)$ is the periodogram of segment r and K is the number of segments. Window length and overlap influence the frequency resolution, temporal localisation, and variance of the resulting spectrum. The specific values used in this work are therefore defined in the Methods chapter.

Frequency-domain representations are particularly useful in acoustic monitoring because changes in a particulate process may affect some parts of the spectrum more strongly than others. Hansuld et al. [33] analysed acoustic power within selected frequency intervals during pharmaceutical granulation and showed that different spectral regions contained useful process information. A later study found that acoustic power in particular regions was also affected by operating parameters, including impeller speed [19]. Watson et al. [29] similarly observed changes in the distribution of acoustic energy across frequency as the particulate state evolved.

Using every individual frequency bin as a separate predictor can produce a large number of highly correlated variables. One alternative is to combine neighbouring frequencies into broader bands. The mean logarithmic PSD within a band B can be represented as

$$\overline{P}_B = \frac{1}{N_B} \sum_{f_k \in B} \log_{10}\left(\hat{S}_{xx}(f_k)\right), \tag{2.10}$$

where N_B is the number of frequency bins within the selected interval. Band aggregation provides a lower-dimensional description of the spectrum and reduces the influence of isolated frequency bins. The particular band limits used in this thesis are derived from the experimental analysis rather than adopted from fixed frequency ranges reported in previous studies.

A second way of describing a frequency region is through its spectral slope. For a selected band, the relationship between logarithmic PSD and logarithmic frequency can be approximated by

$$\log_{10} \hat{S}_{xx}(f) = \beta_0 + \beta_1 \log_{10}(f) + \epsilon, \tag{2.11}$$

where β_1 represents the spectral slope. This parameter provides a compact description of how acoustic power changes across a frequency interval. Previous particulate studies support the broader observation that changes in particle state and collision behaviour can modify the shape of the measured spectrum [29].

PSD-based features describe the frequency content averaged over an analysed interval but do not retain the exact timing of short-duration changes. For non-stationary signals, the short-time Fourier transform (STFT) provides a joint representation of time and frequency. It can be expressed as

$$X(t_0, f) = \int_{-\infty}^{\infty} x(t)w(t - t_0)e^{-j2\pi f t}\, dt \tag{2.12}$$

where $w(t - \tau)$ is a window centred at time τ. The squared magnitude of the STFT can be visualised as a spectrogram, making it possible to identify when changes in spectral content occur [35].

Time-frequency representations have been used in acoustic monitoring where the signal changes during operation. Fulek et al. [30], for example, transformed microphone recordings from pharmaceutical granulation into spectrogram-based representations before classifying different process phases. In a different sensing application, Mendoza-Silva et al. [36] converted mechanical nanomotion signals into time-frequency spectrograms before applying machine-learning models. Although the physical system in that study is unrelated to powder blending, it provides a recent example of extracting discriminative information from complex mechanical signals through time-frequency representation.

For residence mass estimation, the different feature groups provide complementary descriptions of the measurements. Time-domain quantities capture overall amplitude and transient behaviour, while PSD-based features describe how acoustic energy is distributed over frequency. Band averages reduce spectral dimensionality, spectral slopes describe changes in spectral shape, and spectrograms retain the temporal development of frequency content.

2.6 Multivariate Analysis: PCA, SVD, and t-SNE

Acoustic spectra contain a large number of frequency variables, many of which are correlated because neighbouring frequency bins often respond to the same underlying changes in the measured signal. Dimensionality reduction can therefore be used to represent the dominant variation using a smaller number of variables and to examine patterns between observations.

Principal component analysis (PCA) provides a linear representation of this variation using orthogonal principal components. PCA can be obtained from the singular value decomposition (SVD) of a centred data matrix. For a centred matrix $\mathbf{X}$,

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \tag{2.13}$$

where $\mathbf{U}$ and $\mathbf{V}$ contain the left and right singular vectors, respectively, and $\boldsymbol{\Sigma}$ contains the singular values [52]. The PCA score and loading matrices can then be written as

$$\mathbf{T} = \mathbf{U}\boldsymbol{\Sigma}, \qquad \mathbf{P} = \mathbf{V}, \tag{2.14}$$

such that

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T. \tag{2.15}$$

The score matrix $\mathbf{T}$ represents the observations in the reduced principal-component space, while the loading matrix $\mathbf{P}$ describes the contribution of the original variables to each component. The components are ordered according to the amount of variance they explain, allowing a reduced number of components to be retained while preserving most of the variation in the original data.

The same decomposition can also be applied directly to a spectrogram matrix. In this case, the singular vectors provide complementary descriptions of dominant spectral and temporal patterns in the signal. SVD of a spectrogram and PCA of a feature matrix are therefore mathematically related, but they are applied to different data representations for different purposes. Spectrogram decomposition is useful for identifying dominant frequency and time behaviour within a recording, whereas PCA of spectral features provides a common lower-dimensional space for comparing observations between time periods or operating conditions.

PCA provides a linear representation of the data and may not fully describe more complex local relationships between observations. t-distributed stochastic neighbour embedding (t-SNE) is a nonlinear dimensionality reduction method developed primarily for visualisation of high-dimensional data [38]. It constructs a low-dimensional representation by defining similarities between observations and attempting to preserve their local neighbourhood relationships.

When PCA is used before t-SNE, the retained PCA score matrix $\mathbf{T}$ provides a reduced input representation for the nonlinear embedding. This reduces the number of variables supplied to t-SNE while retaining the dominant variation identified by PCA. The resulting two-dimensional maps can then be used to examine whether

observations with similar temporal or operating conditions occupy similar regions of the feature space.
t-SNE plots are interpreted qualitatively because the method primarily preserves local neighbourhood relationships. The apparent distance between widely separated groups and the relative size of clusters should therefore not be interpreted as direct quantitative measures of similarity [38].
Together, the linear PCA/SVD representation and nonlinear t-SNE visualisation provide complementary descriptions of high-dimensional acoustic data. The former identifies dominant variation and spectral or temporal structure, while the latter provides an additional view of local relationships between observations.

2.7 Predictive Modelling Approaches and Model Selection

Estimating residence mass from acoustic measurements is a multivariate regression problem in which the predictor variables can be numerous and strongly correlated. Neighbouring frequency bins describe related parts of the same spectrum, while features derived from the same recording may contain overlapping information. The available modelling approach must therefore account for this correlation while avoiding unnecessary complexity relative to the number of independent experimental conditions.
Several regression methods could be considered for this type of problem. Random Forest Regression combines predictions from multiple decision trees and can represent nonlinear relationships and interactions between predictors. Artificial Neural Networks (ANNs) provide greater flexibility and can approximate complex nonlinear mappings between input variables and a target response. Support Vector Regression (SVR) and Gaussian Process Regression (GPR) provide further nonlinear alternatives, particularly when a linear relationship between spectral measurements and the response is insufficient.
These approaches have been used in related pharmaceutical modelling problems. Jones-Salkey et al. [11] compared Random Forest Regression, symbolic regression, and an ANN for predicting steady-state fill level in an inclined linear continuous blender. The different models provided complementary forms of process information, while the ANN produced the highest predictive accuracy in their study. Beke et al. [50] also used recurrent neural networks together with residence-time-distribution models to represent the dynamic behaviour of a continuous pharmaceutical powder-blending process.
Nonlinear models may also provide advantages for high-dimensional spectral calibration. Sow et al. [51] compared PLS with Random Forest, Support Vector Regression, and Gaussian Process Regression for pharmaceutical near-infrared measurements and reported lower prediction errors for the nonlinear models in their dataset. This illustrates that no single regression approach is expected to provide the best performance for every spectral modelling problem.
Partial least squares (PLS) regression provides a different balance between dimensionality reduction, prediction, and model interpretation. PLS was developed for re-

gression problems containing correlated predictor variables and constructs a smaller set of latent components while taking the response variable into account [39]. This characteristic is particularly relevant to acoustic spectra because adjacent frequency variables and derived spectral descriptors can contain substantial redundant information.

For predictor matrix $\mathbf{X}$ and response matrix $\mathbf{Y}$, the PLS decomposition can be represented as

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}, \tag{2.16}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F}, \tag{2.17}$$

where $\mathbf{T}$ and $\mathbf{U}$ contain latent scores for the predictor and response spaces, $\mathbf{P}$ and $\mathbf{Q}$ contain their corresponding loadings, and $\mathbf{E}$ and $\mathbf{F}$ represent residual variation. The fitted model can then be expressed as

$$\hat{\mathbf{Y}} = \mathbf{X}\mathbf{B} + \mathbf{b}_0, \tag{2.18}$$

where $\mathbf{B}$ contains the regression coefficients and $\mathbf{b}_0$ represents the intercept.

PLS was selected as the primary regression approach in this thesis for three main reasons. First, the acoustic predictor variables are strongly correlated, making latent-variable modelling suitable for reducing redundant information. Second, the number of independent residence-mass experiments is limited relative to the number of possible signal features. A model based on a small number of latent components therefore provides a relatively constrained representation compared with highly flexible non-linear approaches. Third, PLS retains a useful degree of interpretability through its loadings, regression coefficients, and variable-importance measures. This is relevant because the study aims not only to predict residence mass but also to investigate which acoustic characteristics contribute to the prediction.

PLS also has direct precedent in continuous powder blending. Gao et al. [40] used projection to latent structures regression to investigate the influence of design and operating parameters on continuous-blender performance. Bekaert et al. [25] similarly applied PLS regression to establish relationships between powder properties, throughput, impeller speed, hold-up mass, mean residence time, and blending performance.

PLS can also be adapted for supervised classification through partial least squares discriminant analysis (PLS-DA). In this approach, class membership is represented numerically and modelled using the latent-variable structure of PLS [41]. In the present study, PLS-DA is used to determine whether acoustic features contain sufficient information to distinguish the different operating conditions. This analysis is exploratory and does not replace the run-level evaluation required for residence-mass prediction.

The contribution of individual predictors to a PLS model can be examined using Variable Importance in Projection (VIP) scores. A commonly used formulation for predictor j is

$$\mathrm{VIP}_j = \sqrt{p \frac{\sum_{a=1}^{A} SSY_a \left(\frac{w_{ja}}{\|\mathbf{w}_a\|} \right)^2}{\sum_{a=1}^{A} SSY_a}}, \tag{2.19}$$

where p is the number of predictors, A is the number of retained latent components, SSY_a represents the response variation associated with component a, and w_{ja} is the weight of predictor j on that component.
Larger VIP values indicate that a variable contributes more strongly to the fitted PLS representation. However, VIP should not be interpreted as evidence that a predictor has a direct causal relationship with residence mass. Its value depends on the fitted components and on correlations between predictors [44, 45]. In this work, VIP is therefore used to identify informative variables and spectral regions rather than to assign a physical mechanism to an individual feature.
The selection of PLS does not imply that the relationship between the acoustic measurements and residence mass is fundamentally linear or that PLS is expected to outperform nonlinear machine-learning methods in all cases. It was chosen as a comparatively low-complexity and interpretable approach that is compatible with strongly correlated predictors and the limited number of independent experimental runs. A larger and more diverse dataset would allow a more extensive comparison with nonlinear modelling approaches in future work. “‘

2.8 Model Generalisation and Validation

The performance of a predictive model should be evaluated not only by how well it describes the data used for fitting, but also by its ability to predict observations that were not involved in model development. This distinction is particularly important when the number of independent experiments is limited or when several observations originate from the same experimental condition.
A model can achieve a high training accuracy by fitting relationships that are specific to the calibration data. Such relationships may reflect the intended process variable, but they may also capture run-specific characteristics, measurement noise, or correlations that do not remain stable under new conditions. Generalisation therefore refers to the ability of the fitted relationship to retain predictive value when it is applied to independent observations.
This issue becomes important when a dataset contains repeated observations from the same experimental unit. Signal windows extracted from a single blending run are related because they share the same equipment settings, material condition, acquisition environment, and residence-mass reference. Treating these windows as fully independent observations can place closely related samples in both the training and additional sets. Cross-validation studies on structured data have shown that ignoring such dependencies can lead to an underestimated prediction error [46]. For a model intended to predict a new experimental run, the validation split should therefore preserve the grouping structure of the data.

Leave-one-run-out validation provides one approach for evaluating this type of generalisation. If there are R independent runs, the model is fitted using $R-1$ runs and evaluated on the remaining run. The procedure is repeated until every run has been held out once. All observations belonging to the test run remain outside the fitting process, so the resulting prediction represents transfer to an experimental condition that was not included in model calibration.
Prediction performance can be summarised using several complementary metrics. The coefficient of determination is given by

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{2.20}$$

where y_i is the measured value, $\hat{y}_i$ is the prediction, and $\bar{y}$ is the mean of the measured values. An R^2 value close to one indicates that the predictions account for a large proportion of the variation in the response. During validation, R^2 can also become negative when the prediction errors are larger than those obtained by using the mean response as a constant prediction.
The root mean square error (RMSE) is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{2.21}$$

while the mean absolute error (MAE) is

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|. \tag{2.22}$$

Both metrics retain the units of the predicted quantity and therefore provide a direct measure of the typical prediction error. RMSE gives greater weight to large errors because the residuals are squared, whereas MAE describes the average absolute deviation between measurements and predictions.
Cross-validation can also be used during model development to select features, model complexity, or other tuning parameters. However, using the same validation results both to select the best model and to report its final performance can produce an optimistic estimate of generalisation. Varma and Simon [47] showed that model tuning based on cross-validation can bias the resulting error estimate unless the complete selection procedure is repeated within each outer cross-validation fold. Cawley and Talbot [48] further demonstrated that the model-selection criterion itself can be overfitted when several alternatives are compared using a finite dataset.
Feature selection must therefore be regarded as part of model fitting rather than as an independent preprocessing step. If predictor selection is based on the response variable, it should be repeated using only the training observations within each validation fold. The same principle applies to scaling, selection of latent components, and other data-dependent modelling decisions [47, 49].
With a small number of independent experimental runs, cross-validation estimates can also be sensitive to individual observations. Differences in powder conditions,

residence mass, RPM, or throughput can make one held-out run considerably more difficult to predict than another. Performance metrics should therefore be interpreted together with the predictions for the individual runs rather than relying only on a single aggregated value.
A further assessment can be made by applying the model to data collected using a different measurement configuration. This provides information about how the model behaves when the sensing or acquisition conditions change. If a known response value from the new data is used to adjust the predictions, the subsequent results should be interpreted as model application after reference calibration rather than as an independent evaluation of predictive accuracy.

3

Methods

3.1 Process Description and Data Acquisition

3.1.1 Process Overview and Experimental Design

The data used in this study were acquired from a linear blender operating under pharmaceutical manufacturing conditions. The experiments were designed and carried out by the Product and Process Control team at AstraZeneca, and the resulting acoustic and process recordings were subsequently provided for the analysis performed in this thesis. Three gravimetric feeders supplied the components of a placebo formulation at predefined proportions to the blender. A total of seven runs were performed using combinations of different blender rotational speeds and throughput settings. Each run was operated for approximately three residence times, resulting in recording durations of approximately 5 to 13 minutes depending on the operating condition. At the end of each run, the blender was stopped and the amount of material remaining inside was measured gravimetrically. This measured quantity is referred to as the residence mass.Table 3.1 summarises the operating conditions, recording duration, and measured residence mass for each run.

Run	**Blender Speed (RPM)**	**Throughput (kg/h)**	**Duration (min)**	**Residence Mass (g)**
1	350	35	7	1380
2	450	50	5	1142.3
3	350	35	7	1383.8
4	450	20	12.5	948.8
5	250	20	12.5	1885.4
6	350	35	7	1368.8
7	250	50	5	3049.8

Table 3.1: Operating conditions and residence mass for the training runs.

Run 1 was excluded from the subsequent acoustic analysis because saturation in the high-frequency recording. The remaining six runs (Runs 2-7) formed the training dataset used for model development. Run 6 additionally contained deliberately introduced spike and pseudo-spike perturbations during blending, which were used in the analysis of transient signal behaviour.

Figure 3.1: Acoustic measurement setup configurations used in this study.

3.1.2 Acoustic Sensor Instrumentation

During each training run, two acoustic signals were recorded simultaneously using independent acquisition systems. The high-frequency airborne signal was measured using a PCB Piezotronics Model 378C01 measurement system, consisting of a Model 377C01 1/4-inch free-field prepolarized microphone and a Model 426B03 preamplifier. The signal was sampled at 299,760 Hz using the PicoScope acquisition system. The measurement system has a manufacturer-specified frequency range extending from 4 Hz to 100 kHz[53]. Throughout the subsequent analysis, this recording is referred to as the AE signal.
The second signal was recorded using a RØDE PinMic miniature omnidirectional microphone connected through a Roland QUAD-CAPTURE audio interface. This signal was sampled at 44,100 Hz and stored in FLAC format. Throughout the subsequent analysis, this recording is referred to as the FLAC signal.
Both microphones were positioned below the central section of the blender so that the two measurements represented acoustic activity from approximately the same region. The measurement configurations are shown in Figure 3.1. Further technical specifications of the sensors and acquisition systems are provided in Appendix A.1.

3.1.3 Additional Recordings

A separate set of recordings was acquired at a later stage by another team within AstraZeneca using the same linear blender. These recordings were collected separately during routine operation rather than as part of the original training experiment. They were used to examine how the developed modelling procedure behaved when applied to independently acquired data. The available operating conditions are summarised in Table 3.2.
For the additional recordings, the acquisition system was changed from the PicoScope system used during training to an NI 9223 analogue input module installed

in a cDAQ-9170 chassis. The high-frequency signal was acquired at 1,000,000 samples/s using NI FlexLogger and stored in TDMS format [55].
The RØDE PinMic was also used during the additional recordings and was sampled at 44,100 Hz through the Roland QUAD-CAPTURE interface. The measurement arrangement is shown in Figure 3.1. Further details of the sensor, signal-conditioning equipment, and acquisition hardware are provided in Appendix A.1.

Run	**Blender Speed (RPM)**	**Duration (min)**	**Residence Mass (g)**
8	300	10	Not measured
10	300	7	Not measured
11	300	6	Not measured
13	400	10	Not measured
14	400	10	Not measured
15	400	10	873

Table 3.2: Operating conditions for the additional runs.

3.2 Data Conversion and Preprocessing

The PicoScope recordings from the training experiment were exported as multiple CSV segments for each run. The segments were reconstructed sequentially into a single analysis record and stored in binary format in the original sampling rate. The high-frequency additional recordings were converted from TDMS to a compact signed 16-bit binary representation containing the scaling and timing information required to reconstruct the signal. These recordings remained at their original sampling rate of 1,000,000 Hz during conversion.
The RØDE PinMic recordings were stored in FLAC format at 44,100 Hz and were loaded directly for signal processing. Where required, the audio channels were combined into a single mono waveform.
Detailed descriptions of the PicoScope reconstruction, TDMS conversion, binary file structure, and storage requirements are provided in Appendix A.2. Resampling of the additional high-frequency recordings was performed later as part of the inference procedure rather than during the initial data conversion.

3.3 Signal Analysis

3.3.1 Spectrogram Analysis

Spectrogram analysis was performed on the AE and FLAC signals from the six training runs (Runs 2-7) to examine their spectral and temporal characteristics. The analysis was used to identify recurring process-related events, persistent spectral components, and frequency regions containing visible changes during blender operation.

For the AE signal, spectrograms were calculated using 16,384-sample segments with 75% overlap and a Blackman window. This provided a frequency resolution of approximately 18.3 Hz. For the FLAC signal, 2,048-sample segments with 50% overlap and a Hann window were used, resulting in a frequency resolution of approximately 21.5 Hz.
To examine structure across the available bandwidth, the spectrograms were displayed in separate frequency regions. For the AE signal, the ranges 0-1.5 kHz, 1.5-3 kHz, 3-10 kHz, and 10-50 kHz were examined. For the FLAC signal, the corresponding displayed ranges were 0-1.5 kHz, 1.5-5 kHz, 5-10 kHz, and 10-20 kHz. Spectral power was expressed in decibels, with the display limits set to the 5th and 95th percentiles within each frequency region to improve visual contrast.

3.3.2 Singular Value Decomposition Analysis

Singular value decomposition (SVD) was used to characterise dominant spectral and temporal patterns in the low-frequency region of the AE and FLAC signals.
For this analysis, both signals were resampled to approximately 1 kHz. Anti-alias filtering was applied during the resampling procedure so that frequency components above the supported range were suppressed before the sampling rate was reduced.
The resampled signals were subsequently low-pass filtered and analysed separately over two frequency ranges: 0-100 Hz and 0-400 Hz. A 301-tap FIR filter with a Hann window was used, with cutoff frequencies of 100 and 400 Hz respectively. The filter was applied in forward and reverse directions to avoid introducing a temporal shift.
Spectrograms of the filtered signals were calculated using 512-sample segments with 50% overlap, producing a frequency-by-time power matrix for each signal and experimental run. SVD was then applied to the squared spectral-power matrix,

$$S^2 = U\Sigma V^T, \tag{3.1}$$

where U represents the frequency patterns, Σ contains the singular values describing the relative strength of the components, and V^T describes their temporal behaviour.
The contribution of each component was calculated from its squared singular value relative to the sum of all squared singular values. Components accounting for 90% of the variance were considered for the 0-100 Hz analysis, while an 85% criterion was used for the 0-400 Hz analysis. The first three components were examined in terms of their dominant frequencies, temporal behaviour, and explained variance across the six training runs.

3.4 Signal Alignment

The AE and FLAC signals were recorded simultaneously using independent acquisition systems. The AE recording was acquired using the PicoScope system, while the FLAC recording was acquired through the Roland QUAD-CAPTURE interface and Audacity. Recording was initiated manually in the two software applications, and the recordings therefore did not share a common start time.

In addition to the initial offset, a progressive difference between the two time axes was observed over the duration of each run. Consequently, applying only a constant time shift was insufficient to maintain alignment throughout the recording. A time mapping was therefore required to account for both the initial start time difference and the observed difference in time scale.
Alignment between the signals was possible because both acoustic channels recorded the same blender operation simultaneously. Process events such as feeder activity and other short duration mechanical disturbances produced changes that were present in both recordings. Although the two sensors differed in bandwidth and signal response, the timing of these common events remained correlated and could therefore be used to establish temporal correspondence between the recordings.
The alignment procedure used recurring process events that produced measurable changes in both acoustic signals as temporal reference points. Preliminary inspection of the recordings showed that the 100 to 400 Hz range contained strong temporal features that were present in both signals. This frequency range was therefore selected for event detection and cross sensor matching.
The lower frequency region was preferred for alignment because the two measurement systems shared substantially more spectral content in this range. At higher frequencies, differences in sensor bandwidth and response resulted in increasingly different signal characteristics, reducing the similarity required for reliable matching.
The alignment was performed in three stages: coarse peak detection, coarse peak matching, and cross correlation refinement. The resulting anchor pairs were then used to determine a linear mapping between the two time axes.

3.4.1 Coarse Peak Detection

For alignment, both acoustic signals were resampled to approximately 1 kHz, which retained the 100-400 Hz frequency range selected for event detection while reducing the computational cost of the spectrogram calculation.
Spectrograms were computed using a segment length of 512 samples with 50% overlap. At each time frame, the spectral power between 100 and 400 Hz was summed to produce a one-dimensional band-energy envelope for each signal. The envelopes were normalised before peak detection so that event matching depended primarily on their temporal patterns rather than differences in absolute signal magnitude between the two measurement systems.
The strongest peaks in the resulting envelopes were retained as candidate alignment events. The number of retained peaks was selected separately for each run according to the approximate number of recurring events visible in the recordings.

3.4.2 Coarse Peak Matching

The detected peaks from the two band-energy envelopes were used to establish an initial set of corresponding events. For each AE peak, the nearest unused FLAC peak within a maximum temporal separation of 60 s was identified as a candidate match.

The relatively wide search interval allowed for the unknown initial time difference between the two independently started recordings. Each FLAC peak was used only once. The resulting event pairs provided approximate anchor locations for the subsequent cross-correlation refinement.

3.4.3 Cross-Correlation Refinement

The coarse event pairs were refined using local cross-correlation. For this step, both signals were resampled to approximately 4 kHz and band-pass filtered between 100 and 400 Hz using a fourth-order Butterworth filter applied in zero-phase form. This retained the common low-frequency activity used for alignment while reducing differences associated with the wider sensor bandwidths.

The absolute value of the filtered signal was calculated and smoothed using a 50 ms moving-average window to obtain an amplitude envelope. Using the envelope reduced the influence of waveform polarity and detailed differences in sensor response while preserving the timing of changes in acoustic activity.

For each coarse event pair, a template extending approximately ± 2 s around the AE event was extracted. A wider search region extending approximately ± 8 s around the corresponding FLAC event was used to refine its location.

Because the resampled AE and FLAC signals had slightly different sampling rates, the AE template was interpolated to the FLAC sampling rate before comparison. The two segments were normalised to zero mean and unit variance, and cross-correlation was calculated between them. The lag corresponding to the maximum correlation was used to determine the refined temporal correspondence between the two recordings.

3.4.4 Robust Linear Drift Model

The refined event pairs were used to determine the relationship between the AE and FLAC recording time axes. For each experimental run, an initial linear mapping was fitted as

$$t_{\mathrm{FLAC}} = \alpha t_{\mathrm{AE}} + \beta, \tag{3.2}$$

where β represents the relative time offset and α accounts for a difference in the time scales of the two recordings.

Residuals between the matched FLAC event times and those predicted by the initial mapping were calculated. Event pairs with absolute residuals greater than 1.5 s were treated as mismatches and excluded. The linear mapping was then refitted using the remaining event pairs.

The resulting slope and intercept were determined separately for each experimental run. Alignment quality was evaluated using the root-mean-square residual of the retained event pairs. The final mapping was subsequently used whenever corresponding intervals were extracted from the AE and FLAC recordings.

3.5 Unsupervised Multivariate Analysis

3.5.1 Window Selection

For the unsupervised multivariate analysis, time windows were selected manually using the spectrogram analysis described in Section 3.3.Windows containing filling events, process pauses, environmental disturbances, or deliberately introduced spike events were excluded so that the selected intervals represented continuous blender operation without prominent transient disturbances.

For the per-run analysis, 4 s windows were selected from three temporal regions of each run: early, middle, and late. For the cross-run analysis, only windows from the late region were retained. This reduced the influence of the initial filling and start-up period and allowed acoustic measurements from comparable later portions of the six runs to be examined together.

3.5.2 Feature Extraction

For each selected window, the power spectral density (PSD) was calculated independently for the AE and FLAC signals using Welch's method. For the AE signal, a segment length of 16,384 samples with 75% overlap, corresponding to 12,288 samples, was used. The resulting spectrum extended from 0 to approximately 150 kHz. For the FLAC signal, a segment length of 2,048 samples with 75% overlap, corresponding to 1,536 samples, was used, covering frequencies up to 22,050 Hz.

The PSD values were transformed using $\log_{10}$. The AE and FLAC spectra were then independently normalised to the interval $[0, 1]$ using min max scaling before being concatenated into a single feature vector for each window. The two acoustic recordings were aligned before feature extraction. The AE time boundaries corresponding to each FLAC window were determined using the inverse of the time mapping obtained in Section 3.4.4. This ensured that the spectral features from both signals represented the same portion of the blending process.

3.5.3 Dimensionality Reduction

The combined feature matrix (windows $\times$ frequency bins) was standardised to zero mean and unit variance before dimensionality reduction. Since neighbouring frequency bins and spectral regions contained correlated information, PCA was applied to identify the linear directions of maximum variance in the spectral feature space. The number of components retained was determined using a 90% cumulative explained variance threshold.

For the individual-run analyses, this threshold retained between 11 and 13 principal components across Runs 3 to 7. For the pooled cross-run analysis using late windows from Runs 2 to 7, 4 principal components were retained.

PCA scores were used to examine the distribution of the selected windows, while the corresponding loadings were inspected to identify the frequency regions contributing most strongly to the observed variation.

t-SNE was subsequently applied to the retained PCA score matrix, $\mathbf{T}$, to produce two-dimensional visualisations of the local neighbourhood structure. The perplexity parameter was adjusted according to the number of observations, typically between 3 and 15 depending on the analysis, and 2,000 iterations were used.

3.5.4 Analysis Configurations

Two complementary analyses were performed:

- **Per-run analysis** - For Runs 3 to 7, early, mid, and late windows were projected jointly into PCA and t-SNE space. Run 2 was not included in this analysis because adequate uninterrupted intervals were not available across the required temporal regions. Points were coloured by temporal position to assess whether the acoustic signature changed over the run duration.
- **Cross-run analysis (late windows only)** - Late windows from all six training runs, Runs 2 to 7, were pooled into a common feature matrix and projected jointly. Points were coloured by run identity, RPM, and throughput to assess whether the acoustic signatures differed by operating condition.

3.6 Supervised Classification and Regression

Supervised analysis was performed to determine how strongly the acoustic measurements reflected the operating conditions of the blender. Two complementary approaches were used. PLS-DA was applied to distinguish the discrete combinations of blender speed and throughput, while PLS regression was used to predict RPM and throughput as continuous variables. This analysis was used to assess the process information contained in the acoustic measurements before developing the residence mass model. It was not used as the final evaluation of residence mass prediction.

3.6.1 Feature Extraction

For each run, late windows were extracted using 1-second windows with 50% overlap. Corresponding AE and FLAC segments were obtained using the time mapping described in Section 3.4.4. Two spectral representations were compared:

- **Frequency bands -** The power spectral density was integrated across 20 logarithmically spaced bands ranging from 100 Hz to the Nyquist frequency of each sensor. This produced a compact representation capturing energy distribution across broad frequency regions.
- **Full spectrum -** The Welch PSD was interpolated onto 100 uniformly spaced frequency bins from 100 Hz to the Nyquist frequency of each sensor. This representation retained greater spectral detail and was used to examine whether information contained within narrower frequency regions was lost through band aggregation.

In both cases, eight time-domain statistical features (RMS, peak amplitude, crest factor, kurtosis, skewness, zero crossings, mean, and standard deviation) were computed per sensor and concatenated with the spectral features, yielding a combined AE and FLAC feature vector per window.

3.6.2 PLS-DA Classification

Each window was assigned a class label based on its combined operating condition (e.g., “RPM450_TP50”). The feature matrix was standardised to zero mean and unit variance. The target variable was one-hot encoded across all unique operating conditions.

The number of PLS components was selected using two criteria evaluated jointly: stratified 5-fold cross-validation accuracy across a grid of 1 to 12 components, and the Q^2 metric (cross-validated explained variance of the response). The model with the highest Q^2 was preferred when Q^2 exceeded 0.5 and the $R^2 - Q^2$ gap remained below 0.2; otherwise, maximum CV accuracy was used.

A permutation test containing 200 repetitions was used to evaluate whether the observed classification performance exceeded that obtained after randomly rearranging the class labels. For each permutation, the model was refitted using the shuffled labels. The proportion of permutations that achieved an accuracy equal to or greater than the original model was used to calculate the permutation p value. VIP scores were subsequently calculated to identify the variables contributing most strongly to the classification model. Because the cross validation folds were constructed from individual acoustic windows, windows originating from the same experimental run could occur in both the training and validation folds. The resulting classification metrics were therefore interpreted as a measure of operating condition separability within the available dataset rather than as evidence of generalisation to a completely unseen run.

3.6.3 PLS Regression

To assess whether the operating conditions could be estimated as continuous variables rather than discrete classes, a multi-output PLS regression model was fitted jointly to RPM and throughput. The same full-spectrum acoustic representation was used, consisting of 100 interpolated frequency variables and eight statistical features from each acoustic signal. Both the predictor matrix and the two-response target matrix were standardised before model fitting.

Models with different numbers of latent components were evaluated using five-fold cross-validation. Component selection was based on the mean cross-validated Q^2 across the RPM and throughput responses. Prediction performance was subsequently evaluated separately for the two responses using R^2, RMSE, and MAE. Response-specific VIP scores were also calculated to examine whether RPM and throughput were associated with different regions of the acoustic feature space.

The cross-validation folds were constructed at the acoustic-window level, so windows originating from the same experimental run could occur in both the training and validation folds. The resulting metrics were therefore used to assess the operating-condition information contained in the acoustic measurements rather than prediction performance for a completely unseen experimental run.

3.7 Process Data Integration and Signal Masking

In addition to the acoustic recordings, process data were available from the blender and feeder instrumentation. The recorded variables included blender motor speed and torque, together with mass flow rate, net weight, and volumetric mode status for the gravimetric feeders. The process measurements were provided as timestamped CSV files. These signals were used both to identify periods suitable for acoustic analysis and, where applicable, to derive additional process related features during later modelling stages.

3.7.1 Signal Masking Strategy

The gravimetric feeders periodically entered volumetric mode during feeder refilling events. During these periods, the feeder operates without weight-based feedback control. Additionally, periods of zero mass flow indicate that the blending line is paused or transitioning between operating states. In both cases, the acoustic signature does not correspond to steady-state blending conditions and must be excluded.

To ensure that only acoustically representative blending data were used for feature extraction and modelling, a binary validity mask was computed for each run. A time region was classified as **valid** if the following conditions were simultaneously satisfied:

1. Volumetric mode was **inactive** (feeder operating in gravimetric control, no refilling occurring)
2. Mass flow rate was greater than zero (process actively running)

Intervals where volumetric mode was active or mass flow was zero were classified as invalid and removed from the subsequent acoustic analysis. The purpose of this mask was to restrict the dataset to periods representing active operation under gravimetric feeder control. Selection of the later portion of each run for residence mass modelling was performed separately as described in Section 3.8.

3.7.2 Implementation

The volumetric mode and mass flow signals were first merged onto a common process time axis. Forward fill interpolation was used because the process variables were recorded as discrete time series, and each recorded state was considered valid until a new measurement became available. The combined process timeline was then evaluated using the two validity conditions defined in Section 3.7.1. Consecutive samples satisfying both conditions were grouped into contiguous valid time intervals.

The process data and acoustic recordings were stored using different time references. The valid windows, originally expressed in the process data timestamp domain, were converted to the AE and FLAC time references using the datetime-to-AE offset and the warp function Section 3.4.4) respectively. This conversion allowed the same process validity intervals to be applied consistently to both acoustic recordings.An example of the resulting process-based masking procedure is shown in Figure 3.2, illustrating the intervals retained and excluded from subsequent acoustic analysis.

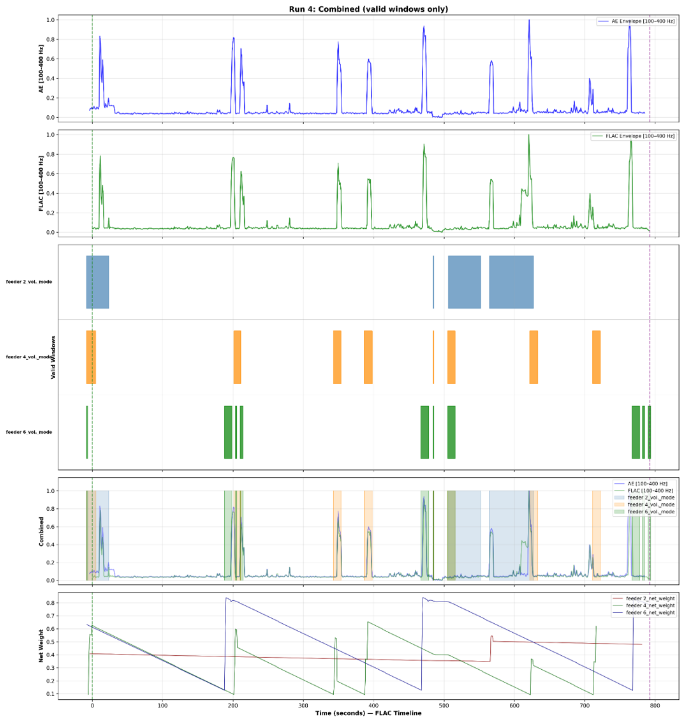


Figure 3.2: process based masking procedure used to identify valid operating intervals before acoustic feature extraction. Regions not satisfying the process validity conditions were excluded from subsequent analysis.

3.7.3 Integration with Acoustic Features

The validity mask was applied before subsequent feature extraction and residence mass modelling. Acoustic windows located within invalid process intervals were excluded from the analysis. Windows crossing the boundary between valid and invalid regions were either excluded or restricted to the corresponding valid interval according to the processing step in which they were used. This procedure reduced the contribution of feeder refill periods, process pauses, and other intervals that did not represent the operating condition selected for model development.

The process signals themselves (motor speed, motor torque, mass flow rate) were also utilised as features in later modelling stages, where their mean values within each analysis block provided complementary information to the acoustic features.

3.8 Acoustic and Process Feature Integration

Following the process-based masking described in Section 3.7, residual transient events were additionally excluded. This reduced the influence of short-duration acoustic disturbances that remained after the process-based masking.

For the feature analysis described in this section, only the final third of each run was used. This reduced the influence of the initial filling and start-up period and focused the analysis on the later part of each experiment, closer in time to the gravimetric residence mass measurement obtained at the end of the run. Since residence mass was measured only once for each experiment, the same run-level residence mass value was associated with all observations extracted from that run.

3.8.1 Block Structure

Within the retained regions, the acoustic recordings were divided into 1 s windows with 50% overlap. Windows were generated only within continuous stretches of valid data and were not allowed to cross an excluded interval. Each window produced one observation in the feature matrix.

This window-based representation was used for the acoustic and process analysis described in this section. The block-based representation used for residence mass prediction is described separately in Section 3.9.2.

3.8.2 Acoustic Features

The acoustic representation followed the full-spectrum feature structure described in Section 3.6.1. For each 1 s window, each acoustic signal contributed 100 spectral variables and eight time-domain statistical variables.

For the analysis in this section, the power spectral density was calculated using Welch's method with 8,192-sample segments and a Hann window. The PSD was interpolated onto 100 uniformly spaced frequencies from 100 Hz to the Nyquist frequency of the corresponding signal and transformed using $\log_{10}$.

The corresponding AE and FLAC intervals were obtained using the time mapping described in Section 3.4.4, ensuring that the two sets of acoustic features represented the same period of blender operation.

3.8.3 Process Features

Process measurements were assigned to each acoustic window using the corresponding window-centre time. Because the process variables were recorded as discrete time series, the most recent available measurement was retained until a new value became available.

Mass flow rate and net weight from feeders 2, 4, and 6 were included, contributing six process variables to each observation.

3.8.4 Combined Feature Vector

The acoustic and process variables were concatenated to form one feature vector for each valid window. Each observation contained 100 AE spectral variables, eight AE statistical variables, 100 FLAC spectral variables, eight FLAC statistical variables, and six process variables, giving a total of 222 predictors.

3.8.5 Train, Validation, and Test Split

The valid data within the last one-third of each run were split chronologically into training (60%), validation (20%), and test (20%) partitions based on the cumulative duration of the retained valid intervals. The chronological split preserved the

temporal ordering of the observations within each run rather than randomly mixing windows from different portions of the same recording. Windows from the six runs were then pooled within their corresponding partitions to form the training, validation, and test matrices.
Because observations from the same experimental run were present across these partitions, this split was used for the intermediate acoustic and process modelling rather than as the final measure of generalisation between independent runs. Cross-run performance of the residence mass model was evaluated separately using the leave-one-run-out procedure described in Section 3.10.5.
For the PLS model containing acoustic and process variables, models with different numbers of latent components were fitted using the training partition and evaluated on the chronological validation partition. The number of components producing the highest average validation R^2 across RPM and throughput was retained. Training R^2 and cross-validated Q^2 calculated from the training data were additionally examined as measures of model quality. The test partition was not used during component selection.

3.9 Feature Engineering for Residence Mass Prediction

The initial spectral representations contained a large number of predictors relative to the number of independent residence mass measurements. Feature engineering was therefore used to reduce the spectral dimensionality before development of the residence mass model.
Initial PLS models were evaluated using higher dimensional spectral representations to determine whether the relationships identified from the acoustic signals remained useful across experimental runs. The resulting model behaviour motivated the use of a reduced representation based on broader frequency regions. The performance of the different spectral representations is presented in the Results chapter.
The residence mass modelling procedure was initially developed using acoustic features only. After the acoustic feature representation, block structure, sampling strategy, and weighting procedure had been established, the engineered process features described in Section 3.11 were added to evaluate whether process information provided complementary information for residence mass prediction.

3.9.1 Dimensionality Reduction Through VIP Analysis

VIP scores obtained from the initial PLS analyses were examined to identify frequency regions that contributed strongly to the fitted models. The purpose of this step was not to select individual frequency bins directly, but to determine whether groups of neighbouring frequencies contained related information.
Adjacent spectral variables were expected to be correlated because they represented neighbouring portions of the same acoustic spectrum. Retaining each frequency value as an independent predictor would therefore increase model dimensionality while introducing substantial redundancy.

The VIP distributions were considered together with comparisons of candidate frequency partitions to define broader spectral regions. Several candidate band configurations were evaluated during model development. Their performance was compared using the cross-run evaluation procedure described in Section 3.10.5, and the reduced band configuration used in the subsequent modelling stages was selected based on cross-run performance rather than training fit alone.
The detailed VIP rankings and comparison of the different spectral representations are presented in the Results chapter.

3.9.2 Band Definition and Feature Computation

The reduced acoustic representation consisted of three frequency bands for each acoustic signal.
For the AE signal, the bands were defined as

$$5{,}000 \text{ to } 20{,}000 \text{ Hz}, \qquad 20{,}000 \text{ to } 50{,}000 \text{ Hz}, \qquad 50{,}000 \text{ to } 100{,}000 \text{ Hz}.$$

For the FLAC signal, the corresponding bands were

$$1{,}000 \text{ to } 5{,}000 \text{ Hz}, \qquad 5{,}000 \text{ to } 15{,}000 \text{ Hz}, \qquad 15{,}000 \text{ to } 22{,}050 \text{ Hz}.$$

The band boundaries were selected empirically from the VIP analysis and the comparison of candidate band configurations. They were used as broad spectral regions for feature extraction rather than being assigned to individual particle interactions or specific physical mechanisms.
For each analysis block, the acoustic signals were divided into 1 s windows with 50% overlap. The power spectral density was calculated for each window using Welch's method with 4,096-sample segments and a Hann window. The resulting PSD estimates were averaged across the windows contained within each block.
Two descriptors were calculated for each frequency band. The first was the mean logarithmic PSD,

$$\bar{P}_{\text{band}} = \frac{1}{|F_{\text{band}}|} \sum_{f \in F_{\text{band}}} \log_{10}\left(S(f) + \epsilon\right), \tag{3.3}$$

where F_{band} represents the frequency bins contained within the selected band, $S(f)$ is the PSD estimate, and $\epsilon = 10^{-12}$ was included to avoid numerical problems for values approaching zero.
The second descriptor was the spectral slope. Within each band, a linear regression was fitted between logarithmic PSD and logarithmic frequency,

$$\log_{10} S(f) = \beta_0 + \beta_{\text{band}} \log_{10} f + \epsilon_f, \qquad f \in F_{\text{band}}, \tag{3.4}$$

where β_{band} is the fitted spectral slope within the selected frequency range.
In addition to the twelve band-based descriptors, the logarithm of the RMS amplitude was calculated for each acoustic signal. The complete acoustic representation therefore contained 14 candidate features per block,

$$6 \text{ bands} \times 2 \text{ descriptors} + 2 \text{ RMS features} = 14 \text{ features}. \tag{3.5}$$

3.9.3 Time Block Segmentation and Late-Region Selection

Rather than computing a single feature vector for the complete retained portion of each run, the acoustic data were divided into fixed-duration blocks. Each block produced one feature vector containing the 14 acoustic features defined in Section 3.9.2. This provided several acoustic observations from each experimental run while retaining the run-level residence mass as the response variable.
Only the final third of each recording was considered for residence mass modelling. The beginning of this region was defined as

$$t_{\text{boundary}} = t_{\text{total}} \left(1 - f_{\text{late}}\right), \tag{3.6}$$

where t_{total} is the total recording duration and $f_{\text{late}} = 0.333$ is the fraction of the recording retained for this stage of the analysis.
The same late-region selection described in Section 3.8 was used for residence-mass modelling. Within this region, the acoustic data were divided into fixed-duration blocks.
The process validity mask and transient exclusion procedure described in Sections 3.7 and 3.8 were applied within the retained region. Blocks were generated only from valid intervals and were not allowed to extend across excluded regions.
Within each block, the acoustic signals were further divided into 1 s windows with 50% overlap. The PSD was calculated for each window and averaged across the windows contained within the block. This provided one spectral representation per block while reducing short-term variation between individual windows.
Corresponding intervals from the AE and FLAC signals were obtained using the time mapping described in Section 3.4.4, ensuring that the extracted features represented the same period of blender operation.
Several block durations were evaluated during model development to examine the balance between the number of available blocks and the stability of the resulting feature estimates. The comparison of block durations and the final selected duration are presented in the Results chapter.

3.10 Block Sampling and Model Training Strategy

The residence mass model was developed using six training runs, with one gravimetrically measured residence mass available for each run. The measured values ranged from 948.8 g to 3049.8 g. Each run, however, produced several valid acoustic blocks, and the number of available blocks differed according to recording duration and the amount of data remaining after masking and transient removal.
The residence mass values were also unevenly distributed across the calibration range. Five runs were located between approximately 948 and 1886 g, while one run represented the upper part of the range at approximately 3050 g. This high residence mass measurement was a valid experimental condition and therefore needed to retain sufficient influence during model fitting rather than being treated as an atypical observation simply because it was sparsely represented.

The model training procedure was therefore designed to address three aspects of the available data: unequal numbers of blocks between runs, variation among blocks belonging to the same run, and the uneven distribution of residence mass across the calibration range.

3.10.1 Block Duration Optimisation

Block duration determined the amount of acoustic data represented by each feature vector. Shorter blocks produced more observations from each run but averaged over fewer 1 s analysis windows, whereas longer blocks provided greater averaging within each feature vector but reduced the number of available blocks.
Four block durations of 10, 15, 20, and 30 s were evaluated using the same frequency-band configuration and model structure. The initial comparison was performed using convergence-based sampling so that the effect of block duration could be examined under the same sampling procedure.
A duration of 15 s was carried forward to the subsequent sampling comparison because it retained approximately 4 to 12 valid blocks per run while still averaging several 1 s spectral estimates within each block. The influence of block duration on cross-run prediction performance is presented in Section 4.5.2

3.10.2 Sampling Methods

Four block selection strategies were evaluated: convergence based sampling, stratified sampling, random sampling, and a combined approach. The purpose of comparing these strategies was to determine how representative blocks could be selected while giving each experimental run a comparable contribution to model fitting.
convergence based sampling- A provisional PLS model was fitted using the currently accepted blocks. Predictions were calculated for the individual blocks, and blocks with residuals greater than $k = 2.0$ times the median absolute residual were identified for removal. The procedure was repeated until the accepted block set converged, defined as no further change in the retained blocks, or until a maximum of 30% of the available blocks had been removed.
Stratified sampling - The available blocks within each run were divided into equally spaced temporal groups. Blocks were selected across these groups so that the retained observations covered different parts of the analysed period rather than being concentrated within one temporal region.
Random sampling - Blocks were selected uniformly from the available observations within each run without considering their temporal position or their agreement with a provisional model. This provided a sampling strategy without model based rejection.
Combined sampling - convergence based sampling was first applied to identify atypical blocks. Stratified sampling was subsequently applied to the remaining blocks to maintain temporal coverage.
The stochastic sampling procedures were repeated 100 times. For each repetition, the selected blocks were used to construct the training dataset and fit the model. The sample producing the highest training R^2 was retained according to the implemented

model development procedure. The effect of the different sampling strategies and repeated sampling on model performance is presented in the Results chapter.

3.10.3 Density and Confidence Weighting

Two weighting procedures were applied to address different sources of imbalance in the training data. Density weighting operated between experimental runs, while confidence weighting operated between blocks belonging to the same run.

Density weighting was introduced because the residence mass measurements were concentrated within one part of the calibration range, with only one run representing the upper mass region. Without correction, the group of closely spaced residence mass values would contribute more strongly to the regression objective than the sparsely represented upper part of the range.

For run i, the local density was calculated as

$$d_i = \sum_{j=1}^{n} \mathbf{1}\left(|m_i - m_j| \leq h\right), \tag{3.7}$$

and the corresponding density weight was

$$w_i^{(d)} = \frac{1}{d_i}, \tag{3.8}$$

where m_i is the measured residence mass of run i, n is the number of runs included in the current training set, $\mathbf{1}(\cdot)$ is the indicator function, and h is a bandwidth equal to 30% of the residence mass range.

The resulting weights were normalised so that their sum was equal to the number of runs. Runs located in densely populated parts of the residence mass range therefore received lower relative weights, while sparsely represented conditions retained greater influence during model fitting.

Confidence weighting was used to account for variation among blocks from the same experimental run. For each block, the Euclidean distance between its feature vector and the feature centroid of the corresponding run was calculated. The confidence weight for block j in run i was defined as

$$w_{ij}^{(c)} = \frac{1}{1 + d_{ij}/\widetilde{d}_i}, \tag{3.9}$$

where d_{ij} is the distance of block j from the run centroid and $\widetilde{d}_i$ is the median distance of all blocks belonging to that run.

Blocks with feature values closer to the typical behaviour of the run therefore received greater influence, while blocks further from the run centroid contributed less without being removed entirely. The confidence weights were used when combining the selected blocks into a representative run level feature vector.

The density weights were subsequently applied during model fitting by scaling the run level feature vectors and corresponding residence mass values by the square root of their respective weights.

3.10.4 Feature Selection and PLS Fitting

The number of candidate features remained large relative to the six independent residence mass measurements. Feature selection was therefore performed before fitting the final PLS model.
Candidate variables were ranked according to the magnitude of their weighted Pearson correlation with residence mass. For predictor x and response y, the weighted correlation was calculated as

$$r_w(x,y) = \frac{\sum_i w_i(x_i - \bar{x}_w)(y_i - \bar{y}_w)}{\sqrt{\sum_i w_i(x_i - \bar{x}_w)^2}\sqrt{\sum_i w_i(y_i - \bar{y}_w)^2}}, \tag{3.10}$$

where the weighted means are

$$\bar{x}_w = \frac{\sum_i w_i x_i}{\sum_i w_i}, \qquad \bar{y}_w = \frac{\sum_i w_i y_i}{\sum_i w_i}. \tag{3.11}$$

The correlation based ranking provided a simple way to reduce the predictor set without introducing a more complex multivariable selection procedure. This was particularly important given the small number of independent residence mass measurements.
For the acoustic only model, the three highest ranked acoustic variables were retained. When the three process features described in Section 3.11 were added to the candidate pool, four variables were retained from the combined set.
The selected predictors were used to fit a PLS regression model with one latent component. Restricting the model to one component limited model complexity relative to the small number of independent training runs.

3.10.5 Leave-One-Run-Out Validation

Model generalisation was evaluated using leave one run out (LORO) cross validation. In each fold, one complete experimental run was withheld and the model was developed using the remaining five runs.
The complete training procedure was repeated within each fold. This included block sampling, calculation of density and confidence weights, construction of representative run level features, feature ranking, scaling, and PLS fitting. Density weights were recalculated using only the residence mass values available within the current training fold.
The withheld run was not involved in feature ranking or model fitting. Its prediction was obtained from the available valid blocks using the model developed from the remaining five runs. The measured residence mass of the withheld run was used only after prediction to calculate the prediction error.
Repeating this procedure for all six runs produced one prediction for each withheld experimental condition. These predictions were combined to calculate LORO R^2, RMSE, and MAE at the run level.
Feature selection was repeated independently within every fold. The features selected in different folds were therefore allowed to vary, so that the reported LORO

R^2, RMSE, and MAE assessed the complete modelling procedure rather than the performance of a feature set selected beforehand using all six residence mass measurements.

3.11 Process Variable Feature Engineering

The process variables used in the earlier acoustic and process analysis described in section 3.11 were included individually as feeder mass flow and net-weight measurements. For residence mass prediction, these raw variables were not carried forward directly. Instead, three composite features were constructed from blender speed, motor torque, feeder mass flow, and feeder net weight to provide a compact representation of the process state.
The residence mass modelling procedure was first developed using acoustic features only. After the acoustic representation, block structure, sampling strategy, and weighting procedure had been established, the three engineered process features defined below were added to evaluate whether process information provided complementary information for residence mass prediction. The features were treated as empirical process descriptors rather than direct measurements of residence mass.

3.11.1 Composite Feature Definitions

The process signals were obtained from the CSV files described in Section 3.7 were recorded at approximately 1 Hz. The variables used to construct the composite features were blender speed, motor torque, feeder mass flow rate, and feeder net weight.
For each 15 s block, the process measurements corresponding to the same time interval as the acoustic features were extracted. The mean blender speed, $\bar{\omega}$, and mean torque, $\bar{T}_m$, were calculated from the available samples within the block. The total throughput was computed as mean feeder mass flow, calculated as

$$\dot{m}_{\text{tot}} = \sum_k \bar{\dot{m}}_k, \tag{3.12}$$

where k represents feeders 2, 4, and 6.
For each feeder, the change in net weight over the block was estimated from the slope of a linear fit to the recorded net weight,

$$\dot{W}_k = \frac{dW_k}{dt}. \tag{3.13}$$

These block level quantities were then used to construct the following three process features.
Process feature 1: torque, speed, and throughput ratio.
The first feature combined the mean torque, blender speed, and total feeder mass flow:

$$P_{f1} = \frac{\bar{T}_m/\bar{\omega}}{\dot{m}_{\text{tot}}}. \tag{3.14}$$

This feature was introduced as an empirical representation of the relationship between the mechanical loading of the blender, its rotational speed, and the rate at which material was supplied. It was not interpreted as a direct measure of mechanical energy.

Process feature 2: flow and weight change ratio.

The second feature combined feeder mass flow and the change in net weight relative to blender speed:

$$P_{f2} = \frac{\sum_k \bar{\dot{m}}_k + \sum_k \dot{W}_k}{\bar{\omega}}. \tag{3.15}$$

This feature was used as an empirical descriptor combining material delivery, changes in feeder weight, and blender speed. It was included to determine whether this combination provided information complementary to the acoustic features.

Process feature 3: dynamic variability ratio.

The third feature compared the relative variability of motor torque and mass flow within each block:

$$P_{f3} = \frac{\mathrm{CV}(T_m)}{\mathrm{CV}(\dot{m})}, \tag{3.16}$$

where the coefficient of variation was calculated as

$$\mathrm{CV}(x) = \frac{\sigma_x}{\mu_x}. \tag{3.17}$$

The ratio therefore describes how strongly the relative torque variation changed compared with the relative variation in material flow during the same block. It was included to examine whether short term process dynamics provided additional information beyond the average operating conditions.

3.11.2 Integration with Acoustic Features

The three engineered process features were appended to the 14 acoustic features defined in Section 3.9.2, producing a combined candidate pool of 17 features per block.

Feature selection was then performed using the weighted correlation procedure described in Section 3.10.4. For the combined acoustic and process model, the four highest ranked features were retained within each training fold before fitting the one component PLS model.

The process features were included as candidate variables rather than being forced into the model. This allowed the feature selection procedure to determine whether they provided information complementary to the acoustic measurements. The selected features and the effect of including process variables on cross run prediction performance are presented in the Results chapter.

3.12 Inference Pipeline and Bias Correction

The trained model was subsequently applied to the additional recordings described in Section 3.1.3. Although the same linear blender was used, the high-frequency signal was acquired using a different acquisition configuration. During training, the signal was acquired using the PicoScope system at 299,760 Hz, whereas the additional recordings were acquired using the NI cDAQ system at 1,000,000 Hz. Additional preprocessing was therefore required before applying the feature extraction procedure developed using the training data.

The converted high-frequency recordings were reconstructed from the binary representation described in Appendix A.2.2. The signal was then resampled from 1,000,000 Hz to approximately 300 kHz before acoustic feature extraction. Anti-alias filtering was applied as part of the rate conversion so that the resulting sampling rate was comparable to that of the high-frequency training recordings. The FLAC recordings were retained at their original sampling rate of 44,100 Hz.

3.12.1 Threshold Adjustments for Inference

The additional recordings differed from the training recordings in their acoustic baseline and in the amount of data retained after process masking. Two parameters in the data selection procedure were therefore adjusted during inference.

During training, residual transient events were excluded using a normalised FLAC envelope threshold of 0.2. For inference, this threshold was increased to 0.7. The higher threshold reduced excessive rejection of usable data from the additional recordings caused by differences in the acoustic baseline while continuing to remove the strongest transient events.

The definition of the retained late region was also modified. During model development, the final third was defined relative to the total recording duration. For inference, the final third was instead calculated from the cumulative duration of the valid data remaining after process masking and transient exclusion. This ensured that extended invalid intervals did not reduce the amount of usable data available from the later part of the additional recording.

All subsequent blocks used for prediction were generated from these valid, peak free intervals using the same 15 s block structure and acoustic feature extraction procedure used during model development.

3.12.2 Batch Bias Correction

Application of the trained model to the additional recordings showed a systematic prediction offset between the training and additional measurement configurations. A reference based additive correction was therefore included in the inference procedure. Run 15 was used as the reference because its residence mass had been measured gravimetrically and was available as a known value of 873 g. The bias was calculated using the model predictions from the valid late-region blocks of this reference recording. If no valid late-region blocks were available, all valid blocks were used instead.

The bias was defined as

$$b = \overline{\hat{m}}_{\text{ref}} - m_{\text{ref}}, \tag{3.18}$$

where $\overline{\hat{m}}_{\text{ref}}$ is the mean raw prediction over the reference blocks used for calibration and m_{ref} is the measured residence mass of Run 15.
The same additive correction was subsequently applied to all block predictions,

$$\hat{m}_{\text{corrected}} = \hat{m}_{\text{raw}} - b. \tag{3.19}$$

A positive value of b therefore indicates overprediction of the reference measurement, while a negative value indicates underprediction.
For the run-level inference results, the corrected prediction was averaged over all valid blocks in each recording. This reporting set could contain more blocks than the late-region subset used to calculate the reference bias. Consequently, the reported mean prediction for Run 15 was not constrained to equal exactly 873 g even though Run 15 was used for calibration.
Because the measured residence mass of Run 15 was used to determine the correction, the predictions for the additional recordings represent model application after a one point reference calibration.

3.12.3 Robust Smoothing

The inference pipeline produced a sequence of block level residence mass predictions for each additional recording. To reduce the influence of isolated extreme predictions when examining their temporal behaviour, a robust smoothing procedure was applied to predictions obtained from valid blocks.
Outlying block predictions were identified using the median absolute deviation criterion,

$$|p_i - \tilde{p}| > \kappa \, \text{MAD}(p), \tag{3.20}$$

where p_i is an individual block prediction, $\tilde{p}$ is the median of the block predictions, and $k = 2.0$. The median absolute deviation was calculated as

$$\text{MAD}(p) = \text{median}\left(|p_i - \tilde{p}|\right). \tag{3.21}$$

Predictions satisfying the MAD criterion were excluded from the smoothing step. A cubic smoothing spline was then fitted through the remaining block predictions to provide a continuous representation of the temporal prediction trend.
The smoothed sequence was used for visualisation of the variation in model predictions over the recording. The spread of the individual block predictions was reported as an empirical measure of within run prediction variability and was not interpreted as a formal confidence interval for residence mass.

4

Results

4.1 Acoustic Signal Characterisation

4.1.1 Spectrogram Observations

Visual inspection of the spectrograms from Runs 2 to 7 revealed several recurring acoustic patterns in both the AE and FLAC recordings. The most prominent time localised features were associated with known process events, while persistent tonal components were visible over longer portions of the recordings.

Each run contained approximately 7 to 10 filling related events, which appeared as bright vertical structures extending across a broad frequency range. Corresponding events were visible in both acoustic recordings, although their relative amplitudes differed between the two signals. Their simultaneous occurrence showed that both measurement systems responded to common events in the blending process. This shared temporal information later provided suitable landmarks for signal alignment.

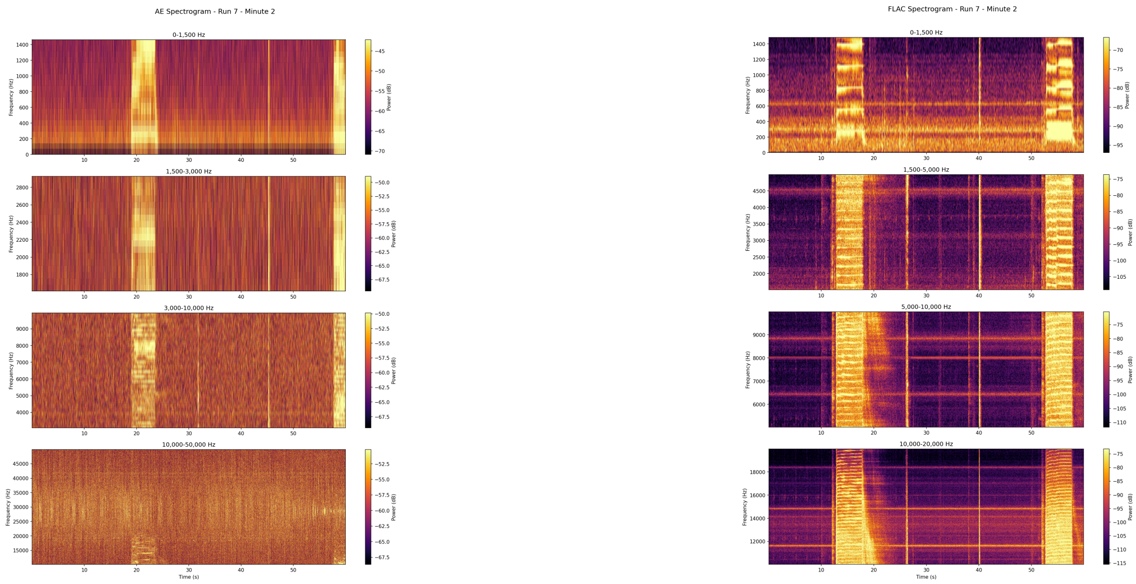
A stronger and more sustained broadband event was observed at approximately 10 min intervals in some recordings. This event corresponded to the feeder header vibration described in Section 3.5.1 and was clearly visible in both acoustic signals.

Periods in which the blending line was paused showed a substantial reduction in the structured spectral content present during normal operation. This indicated that a considerable part of the recurring acoustic activity was associated with operation of the blending process rather than continuously present environmental noise.

Short duration broadband events were also occasionally observed without a corresponding documented process event. These were treated as unidentified external disturbances and were excluded from later analysis where required.

Persistent horizontal components were present throughout the recordings. In the AE signal, prominent regions were observed at approximately 0 to 100 Hz, 200 to 400 Hz, 600 to 800 Hz, 2.4 to 2.6 kHz, 4 to 5 kHz, 15 to 20 kHz, 25 to 30 kHz, and around 40 kHz. The FLAC recording contained a greater density of tonal components below 5 kHz, together with distinct components at approximately 5.5, 6.5, 8, 9, 11, 14, 15, 16, and 19 kHz.

These persistent components show that the recorded spectra contained repeatable tonal structure during blender operation. Their exact physical origin could not be determined from the acoustic measurements alone and may include contributions from rotational excitation, structural response, and the acoustic transfer characteristics of the measurement path. A representative one-minute interval from the AE and FLAC recordings is shown in Figure 4.1. The figure illustrates both the time-localised broadband events observed in the two signals and the persistent frequency



(a) AE recording

(b) FLAC recording

Figure 4.1: Representative spectrograms of the same one-minute interval from the AE and FLAC recordings. The display ranges were adjusted separately for the two signals to make their temporal and spectral structure visible. Colour intensity should therefore be interpreted within each panel rather than compared directly between the two recordings.

components present during blender operation.

Overall, the spectrograms showed that the two recordings contained both common transient events and persistent spectral structure while responding differently across their respective frequency ranges. The common transient events provided the basis for temporal alignment of the two signals.

4.1.2 SVD Analysis

The initial SVD analysis of the 0 to 100 Hz region showed that the AE signal was strongly dominated by the first component across all six runs, accounting for more than 99% of the variance in most cases. The dominant frequency patterns contained peaks at approximately 1.95, 3.90, and 5.85 Hz. These frequencies remained at approximately the same positions across runs performed at different blender speeds and therefore did not follow the corresponding fundamental rotational frequencies. Their physical origin could not be determined from the available measurements.

The FLAC signal showed greater variation in its low-frequency SVD structure. A particularly clear response was observed in Run 6 around $t = 138$ s, corresponding to the deliberately introduced disturbance. As shown in Figure 4.2, the FLAC temporal component exhibited a pronounced response around this time, indicating that the low-frequency signal was sensitive to the disturbance. However, this single event was not considered sufficient to establish a quantitative detection limit.

The 0 to 100 Hz analysis therefore provided an initial indication of low-frequency temporal behaviour, but relatively little common spectral structure was observed

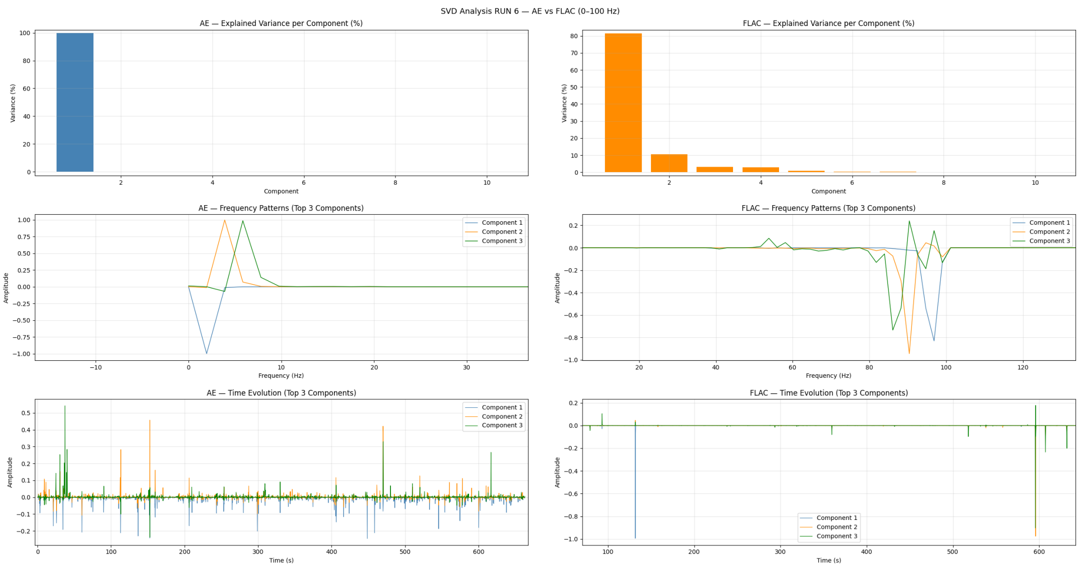


Figure 4.2: SVD results for Run 6 in the 0 to 100 Hz range, showing the AE signal on the left and the FLAC signal on the right. The AE representation was strongly dominated by the first component, while the FLAC temporal component showed a pronounced response during the period containing the deliberately introduced disturbance at $t = 138$ s

between the two acoustic measurements. The broader 0 to 400 Hz analysis presented in the following section revealed clearer recurring features and was therefore examined in greater detail.

Extending the analysed frequency range to 400 Hz revealed additional spectral structure that was not apparent in the 0-100 Hz analysis. The disturbance-related temporal response observed in the narrower analysis remained visible around $t = 138$ s in the 0-400 Hz representation. In addition, a recurring spectral component between approximately 271 and 282 Hz was observed across several runs, as illustrated for Run 6 in Figure 4.3.

For the AE signal, Runs 4 and 7 remained strongly dominated by the first SVD component, whereas Runs 2, 3, 5, and 6 distributed a larger proportion of the variance across the first three components, as shown in Table 4.1.

Run	RPM	Throughput (kg/h)	Comp. 1 (%)	Comp. 2 (%)	Comp. 3 (%)	Components for 85%
2	450	50	53.79	26.62	7.95	3
3	350	35	72.88	8.32	7.84	3
4	450	20	99.90	0.04	0.03	1
5	250	20	58.47	23.58	16.46	3
6	350	35	51.84	25.95	14.82	3
7	250	50	86.54	8.04	3.01	1

Table 4.1: SVD variance distribution for the AE signal in the 0 to 400 Hz range.

Strong frequency components between approximately 271 and 282 Hz coincides numerically with different high-order multiples of the rotational frequency at the three blender speeds. However, the component remained within a similar absolute frequency range as the blender speed changed and therefore did not behave as a single speed-synchronous rotational order.
The repeated occurrence of this narrow frequency region suggests a reproducible equipment-related spectral feature. It may reflect structural response excited during blender operation, potentially including contributions from rotational excitation, but its physical origin could not be established from the available measurements. The present experiments were also not sufficient to determine whether changes in this component were directly associated with the mixing behaviour of the powder.
Run 5 differed from the other runs, with the second and third components showing prominent frequencies around 169 to 172 Hz. This indicates that the distribution of low-frequency acoustic energy was not identical across all operating conditions, although the origin of these components could not be determined from the available measurements.
The temporal components in the 0 to 400 Hz analysis showed more sustained patterns than those observed below 100 Hz. In Run 4, pronounced events occurred at approximately 198, 462, and 747 s, corresponding to a spacing of approximately 264 s. Run 7 showed events at approximately 13, 120, and 224 s, with a spacing of approximately 107 s.

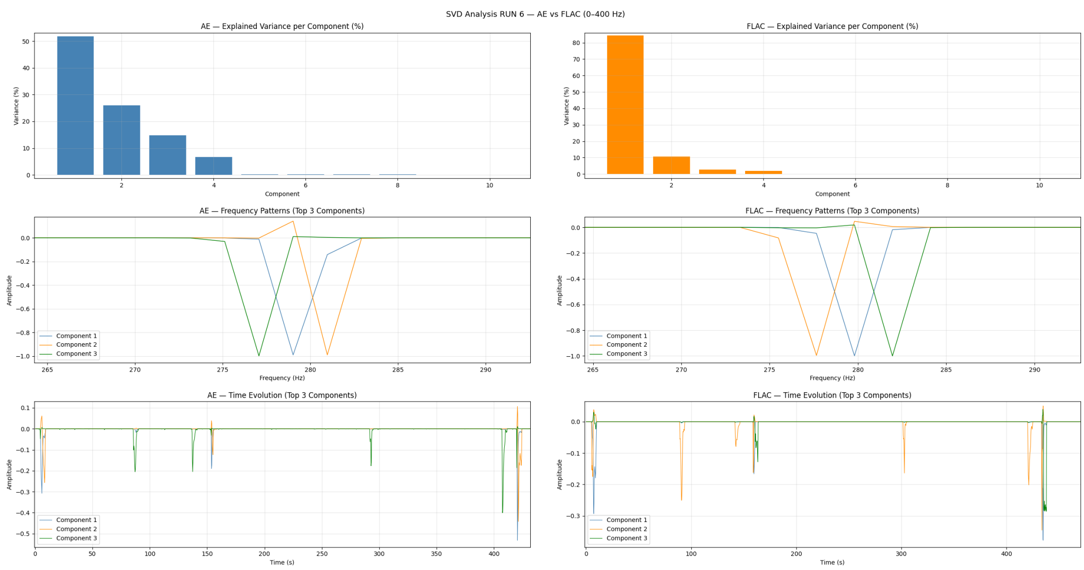


Figure 4.3: SVD results for Run 6 in the 0 to 400 Hz range, showing the AE signal on the left and the FLAC signal on the right. The broader frequency range revealed recurring spectral structure around 271 to 282 Hz across the measurements.

For the FLAC signal, between two and four components were sufficient to account for 85% of the variance in the 0 to 400 Hz analysis. A similar spectral component around 273 to 282 Hz was also observed in the FLAC recording, showing that this frequency region was present in both acoustic measurements.
Run 7 additionally contained a component around 329 Hz that was not prominent

Run	RPM	Throughput (kg/h)	Comp. 1 (%)	Comp. 2 (%)	Comp. 3 (%)	Components for 85%
2	450	50	46.17	23.27	21.35	3
3	350	35	62.09	30.34	4.99	2
4	450	20	38.16	21.10	20.98	4
5	250	20	40.06	25.88	23.61	3
6	350	35	84.46	10.59	2.74	2
7	250	50	74.98	7.64	7.22	3

Table 4.2: SVD variance distribution for the FLAC signal in the 0 to 400 Hz range.

in the other run and sensor combinations. The FLAC temporal components for Runs 4 and 5 also contained recurring events with spacing of the same order as the corresponding residence times. The similarity in time scale is noteworthy, but the SVD analysis alone was not sufficient to establish a direct residence cycle mechanism. Taken together, the SVD results show that the low frequency recordings contained both recurring spectral components and time dependent process events. Extending the analysis from 100 to 400 Hz revealed considerably more common structure between the two acoustic recordings, particularly around 275 to 282 Hz. This common low frequency information also supported the use of the 100 to 400 Hz region for the signal alignment procedure.

4.2 Signal Alignment Outcome

The alignment procedure was applied to all six training runs. Table 4.3 summarises the fitted linear time mapping parameters, the number of matched anchor pairs, and the corresponding alignment residuals.

Run	Slope (α)	Intercept (β, s)	Matched Pairs	Inliers	Outliers	RMS Residual (s)
2	1.0343	−7.86	9	9	0	0.548
3	1.0314	−7.69	7	4	3	0.045
4	1.0280	−3.73	10	10	0	0.549
5	1.0305	+0.57	10	10	0	0.632
6	1.0306	+0.62	11	11	0	0.230
7	1.0312	−8.66	9	8	1	0.488

Table 4.3: Linear time mapping parameters defined in Equation 3.2 and alignment results for the six training runs.

The fitted slope, α, in the linear time mapping defined in Equation 3.2 showed a consistent deviation from unity across all runs, ranging from 1.0280 to 1.0343. This corresponds to an approximately 2.8 to 3.4% difference between the reconstructed AE time scale and the FLAC time scale. A constant time shift alone would therefore

not have maintained correspondence between the two recordings over the complete run duration.
The similarity of the fitted slope across the six runs indicates that the relative time scale difference was systematic. The consistent deviation of the fitted slope from unity indicates a systematic difference between the reconstructed time scales of the two recordings. The available data were not sufficient to determine whether this difference originated from the acquisition clocks, recording software, reconstructed AE time base, or another part of the acquisition procedure.
The fitted intercept, β, varied considerably between runs, which was consistent with the manual initiation of the recordings. Runs 2, 3, and 7 showed offsets between approximately -7 and -9 s, while Runs 5 and 6 had offsets close to zero. Run 4 showed an intermediate offset of approximately -3.7 s. Under the time mapping defined in Equation 3.2, negative values of β indicate that the AE recording began earlier than the FLAC recording, whereas values close to zero indicate that the two recordings were initiated at approximately the same time.
Alignment quality was evaluated using the RMS residual of the retained anchor pairs relative to the fitted linear mapping. The RMS residual ranged from 0.045 s for Run 3 to 0.632 s for Run 5. Run 3 retained four of the seven initial matched pairs after outlier rejection, and its particularly low residual should therefore be considered together with the smaller number of anchors used for the final fit.
For all six runs, the RMS residual remained below approximately 0.7 s. This was small relative to the 15 s blocks subsequently used for residence mass modelling and supported the combination of corresponding AE and FLAC intervals during feature extraction. An example of the resulting alignment is shown in Figure 4.4, where corresponding events in the AE and FLAC recordings occur at similar positions after application of the fitted time mapping. The dotted lines are the peaks identified and used to align.

4.3 Multivariate Structure in Acoustic Data

4.3.1 Per-Run Temporal Structure

PCA and t-SNE applied separately to Runs 3 to 7 showed temporal structure in the acoustic feature space. Windows selected from the early, mid, and late portions of these recordings generally occupied different regions of the reduced-dimensional representations, indicating that the combined acoustic spectrum changed over the duration of the runs. Figure 4.5 shows the early, middle, and late intervals selected from the recording are shown in Figure 4.5(a), followed by the corresponding dimensionality-reduction results in panels (b)-(d).
Early windows showed greater dispersion in several runs, whereas late windows were generally more tightly grouped. This indicates that the acoustic characteristics became less variable during the later portion of the recordings. The pattern was consistent with the transition from the initial filling and start-up period to later operation.
Since residence mass was measured only once at the end of each experiment, the clustering of the late windows was not interpreted as direct evidence that residence

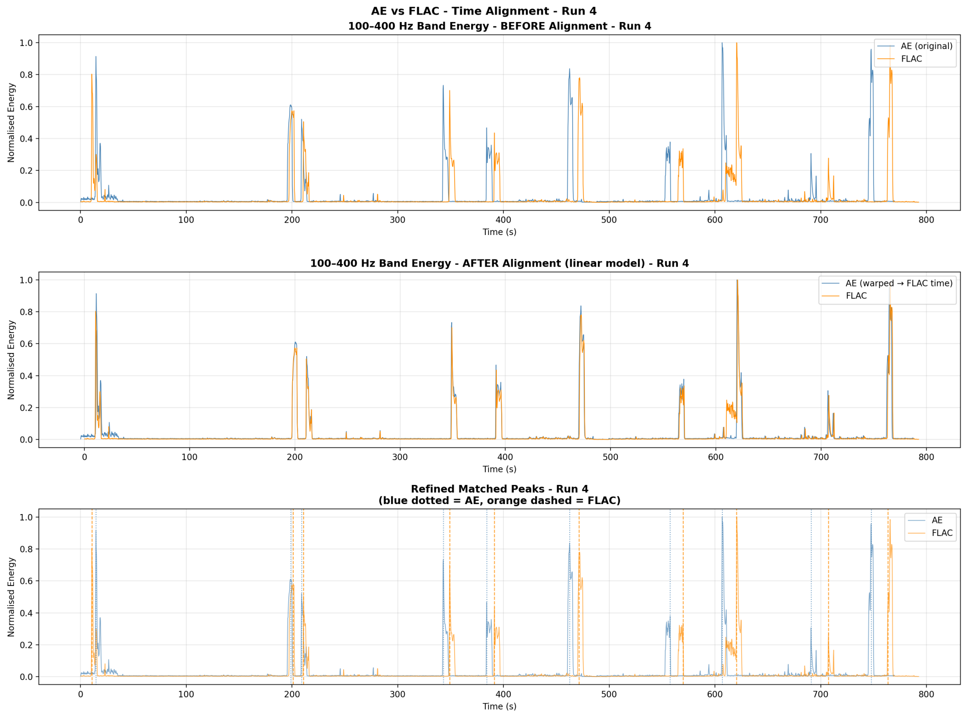
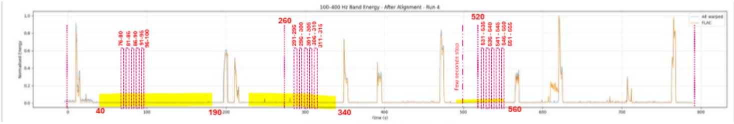


Figure 4.4: Representative alignment between the AE and FLAC recordings based on matched acoustic events.

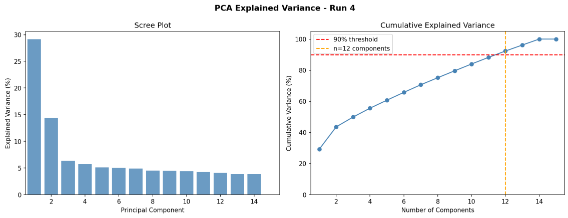
mass had reached equilibrium. Instead, the reduced variability showed that the acoustic measurements became more consistent during the portion of the run closest in time to the gravimetric residence mass measurement.
This temporal behaviour supported the use of late windows for the subsequent cross run comparison, where the objective was to reduce the influence of the initial operating period and compare acoustic measurements from similar later portions of the experiments.

4.3.2 Cross-Run Separation by Operating Condition

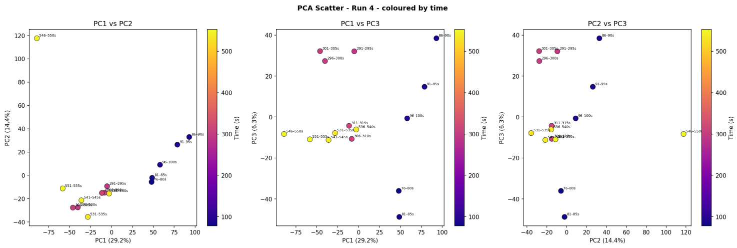
When late windows from Runs 2 to 7 were pooled into a common feature matrix, systematic differences were observed between the experimental conditions.
In the PCA score plots (shown in figure 4.6), observations from runs operated at similar blender speeds showed grouping along the principal component directions. Throughput related grouping was also visible, although the extent of separation varied between conditions. This indicates that part of the dominant variation in the combined AE and FLAC spectra was associated with the operating conditions of the blender.
The t-SNE representation showed similar local grouping of observations according to operating condition. Runs with different combinations of RPM and throughput occupied distinguishable regions of the two dimensional representation. Since t-SNE primarily preserves local neighbourhood relationships, the relative distances between widely separated groups were not interpreted quantitatively.
The similar grouping observed in the PCA and t-SNE representations showed that



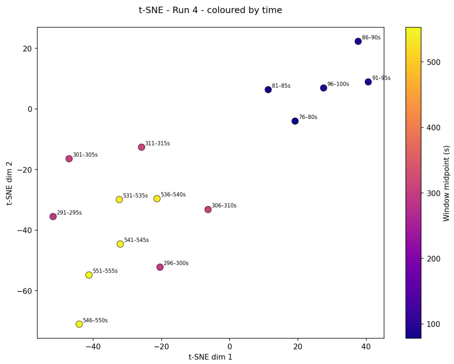
(a) Selected time intervals



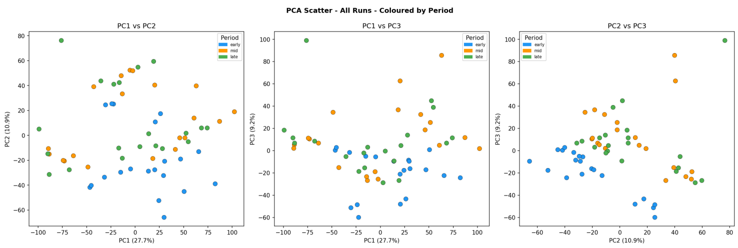
(b) PCA explained variance



(c) PCA score representation

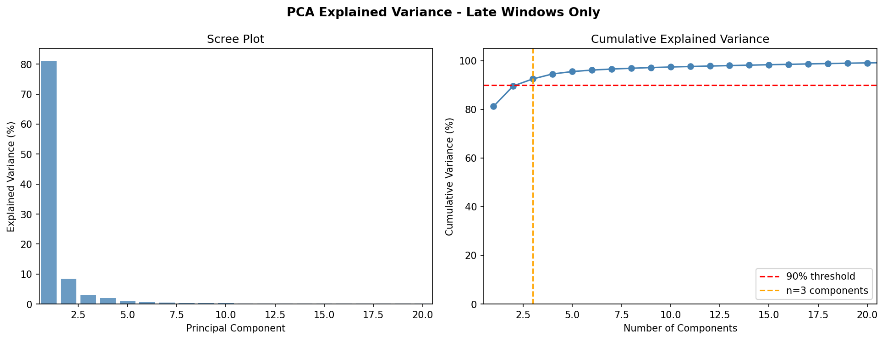


(d) t-SNE projection

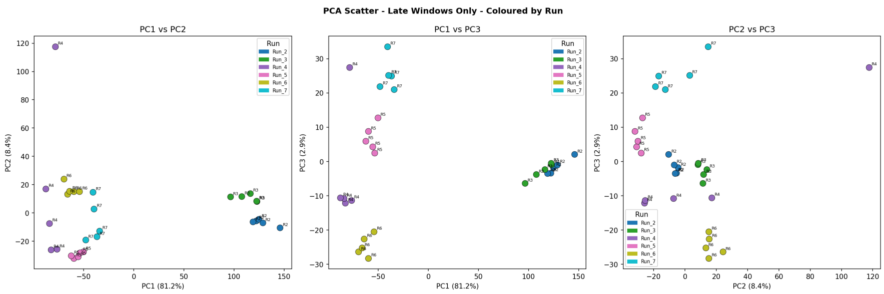


(d) Cross run pca score

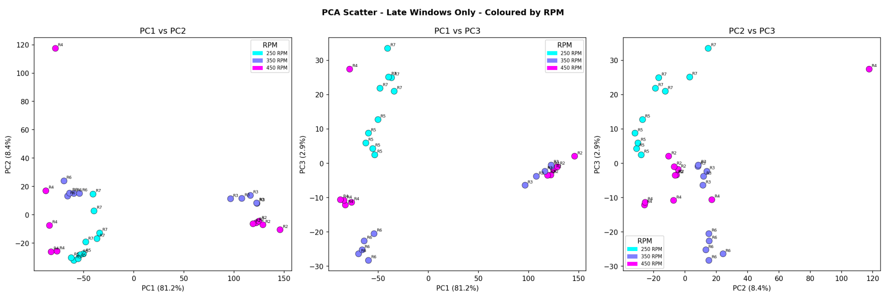
Figure 4.5: Dimensionality reduction analysis for Run 4. (a) Time intervals selected from the early, middle, and late regions of the recording. (b) PCA explained variance used to determine the number of retained principal components, with 12 components retained according to the predefined 90% cumulative explained-variance criterion used before t-SNE (c) PCA score representation of the selected observations. (d) t-SNE projection obtained from the retained PCA score matrix. (e) cross runs distribution of observations from the early, middle, and late regions.



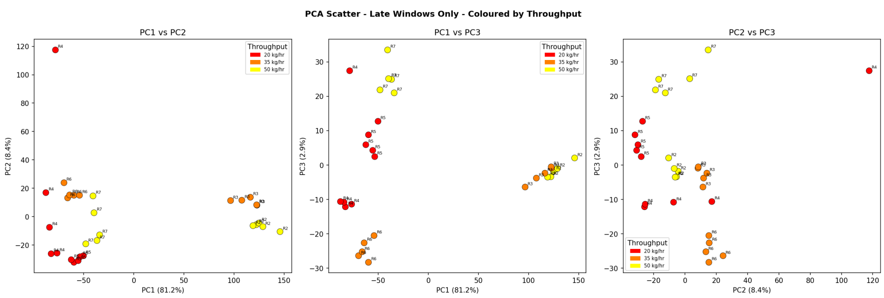
(a) PCA explained variance



(b) PCA scores coloured by run



(c) PCA scores coloured by blender speed



(d) PCA scores coloured by throughput

Figure 4.6: Cross-run PCA of the selected late windows from Runs 2-7. (a) Explained variance used to determine the retained PCA dimensionality, with four principal components required to explain at least 90% of the variance. The PCA score representation is subsequently shown with the observations coloured by (b) run, (c) blender speed, and (d) throughput. The different colourings allow the distribution of the late-window observations to be compared with the corresponding operating conditions.

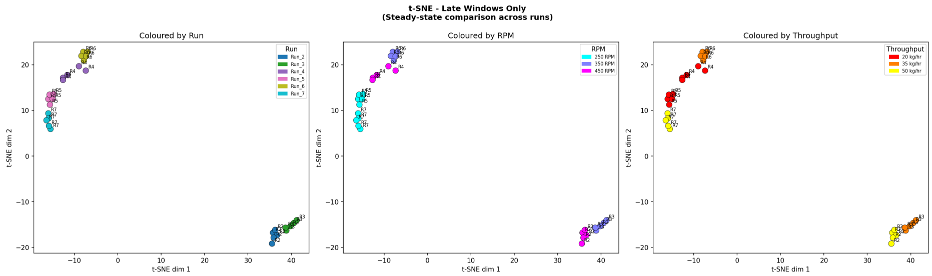


Figure 4.7: t-SNE projection of the selected late-window observations from Runs 2-7. The projection was obtained from the PCA score matrix after retaining four principal components.

the acoustic measurements contained information associated with changes in blender speed and throughput before supervised modelling was applied.

This result was also important for the subsequent residence mass analysis. Since the experimental runs differed in both operating conditions and measured residence mass, acoustic variation associated with RPM and throughput could contribute to apparent relationships between acoustic features and residence mass. Supervised classification and regression were therefore used to quantify how strongly these operating variables were represented in the acoustic measurements.

4.4 Classification and Regression of Operating Conditions

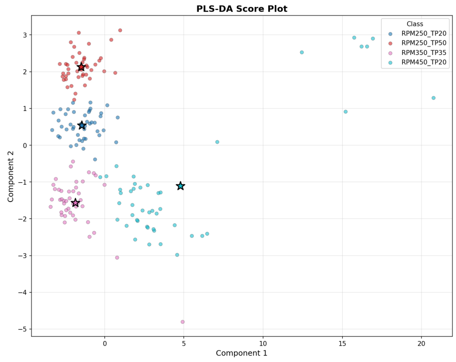
The supervised analyses in this section were used to examine how strongly blender operating conditions were represented in the acoustic measurements. Since acoustic windows from the same six experimental runs could occur in both model development and evaluation subsets, the reported classification and regression results were interpreted as measures of operating-condition information within the available dataset rather than as independent cross-run prediction performance. Generalisation to a completely omitted experimental run was evaluated separately for residence mass using leave-one-run-out validation in Section 4.5.

Variable Importance in Projection (VIP) scores were used to identify variables that contributed strongly to the fitted PLS models. These scores describe variable importance within the fitted model and were not interpreted as evidence of a unique physical origin for the corresponding acoustic frequencies.

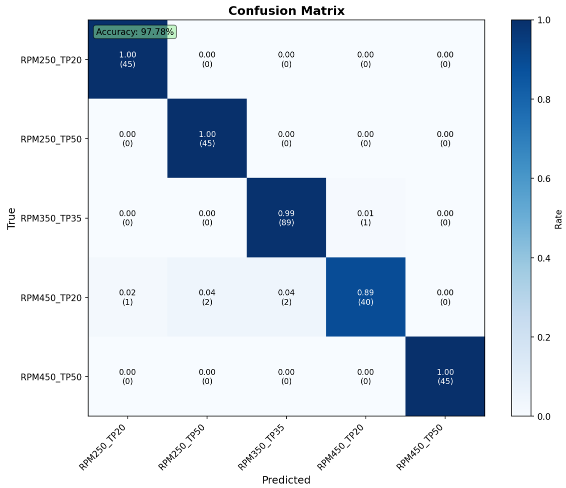
4.4.1 PLS-DA: Frequency Bands

The PLS-DA model based on 20 logarithmically spaced frequency bands retained 12 latent components, with a cross-validated Q^2 of 0.811. The corresponding R^2 was 0.903. The model achieved an overall classification accuracy of 97.8%, with a macro-averaged F1 score of 0.976 and a ROC AUC of 0.999.

The largest confusion occurred for the 450 RPM and 20 kg/h condition, for which the recall was 88.9% and five windows were misclassified as other operating condi-



(a) PLS-DA scores



(b) Confusion matrix

Figure 4.8: PLS-DA classification results using 20 logarithmically spaced frequency bands. (a) PLS-DA score representation of the operating conditions, where points represent individual acoustic windows and stars indicate the centroids of the corresponding operating-condition classes. (b) Confusion matrix showing the classification of the acoustic windows. The model achieved an overall classification accuracy of 97.8%.

tions. All remaining conditions achieved recall values of 98.9% or higher, while the 250 RPM and 20 kg/h and 450 RPM and 50 kg/h conditions were classified without error.

VIP analysis identified high frequency AE bands as the most discriminative variables. The 24,088 to 34,721 Hz band had the highest VIP score of 2.19, followed by 34,721 to 50,047 Hz with a VIP of 2.10. The higher frequency bands between 72,138 and 103,981 Hz and between 103,981 and 149,879 Hz also contributed strongly, with VIP scores of 1.66 and 1.53, respectively. The corresponding VIP ranking is shown in Figure 4.9

The AE band between 50,047 and 72,138 Hz and the FLAC band between 294 and 385 Hz both had VIP scores of 1.39. The latter overlaps with the recurring 275 to 282 Hz component identified in the SVD analysis in Section 4.1.2. Its contribution to the classification model indicates that variation within this low-frequency region

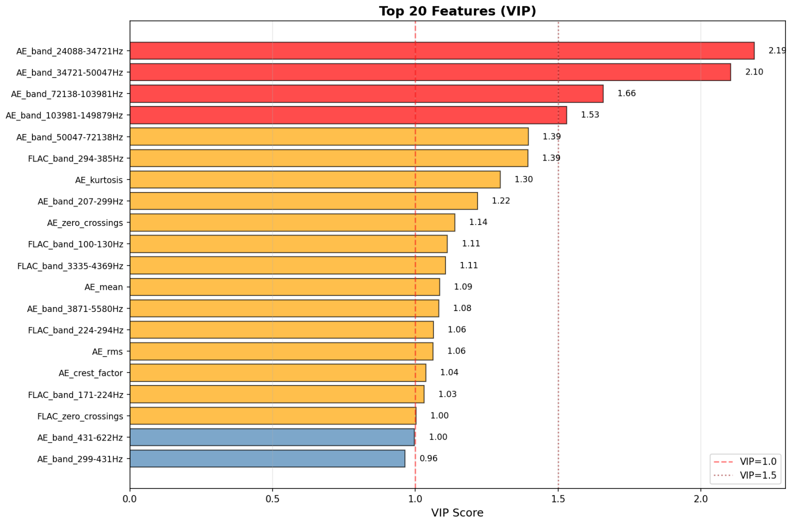


Figure 4.9: VIP scores for the frequency-band PLS-DA model. High-frequency AE bands contributed strongly to the classification, together with selected FLAC frequency regions and time-domain acoustic variables.

was associated with differences between operating conditions.

Among the time domain variables, AE kurtosis, AE zero crossings, and AE mean amplitude appeared among the 20 highest ranked variables, with VIP scores of 1.30, 1.14, and 1.09, respectively. The classification model therefore used a combination of spectral and time domain information rather than relying on a single frequency region.

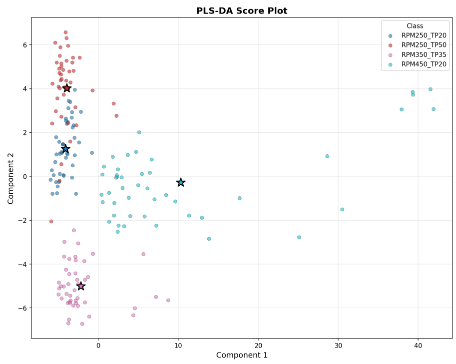
The high classification accuracy demonstrates strong separation of the operating conditions represented in the available dataset.

4.4.2 PLS-DA: Full Spectrum

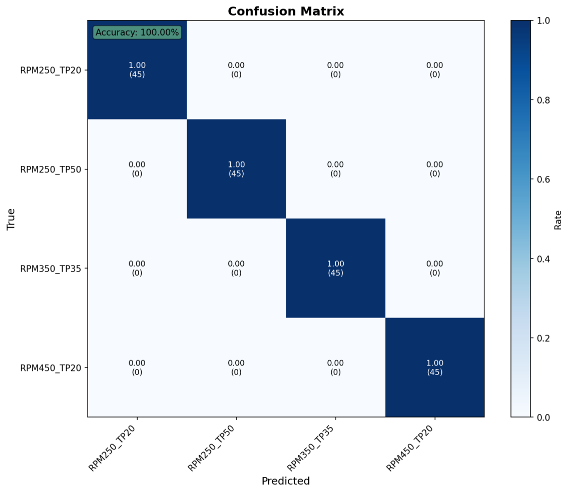
Increasing the spectral representation to 100 uniformly interpolated frequency bins per acoustic signal resulted in a PLS-DA model containing 12 latent components. The selected model achieved an R^2 of 0.950 and a cross-validated Q^2 of 0.899. Classification accuracy was 100%, with precision, recall, and F1 scores equal to 1.0 for every operating condition. The ROC AUC was also 1.0. The PLS-DA score representation and confusion matrix for the full-spectrum model are shown in Figure 4.10, where complete classification of the available acoustic windows was obtained.

VIP analysis showed that the FLAC frequency region around 12.5 to 12.75 kHz was the highest ranked part of the spectrum, with adjacent frequency bins having VIP scores of 2.44 and 2.42. AE frequencies at approximately 28,854 and 27,341 Hz followed with VIP scores of 2.25 and 2.08, respectively. The FLAC frequency around 8,305 Hz had a VIP of 1.96, while AE frequencies around 25,828 and 30,366 Hz had VIP scores of 1.93 and 1.83. The corresponding full-spectrum VIP scores are shown in Figure 4.11.

AE zero crossings and AE mean amplitude were again present among the higher ranked variables, with VIP scores of 1.83 and 1.66, respectively.



(a) PLS-DA scores



(b) Confusion matrix

Figure 4.10: PLS-DA classification results using the full spectral representation. (a) PLS-DA score representation of the operating conditions. (b) Confusion matrix showing complete classification of the acoustic windows within the available dataset.

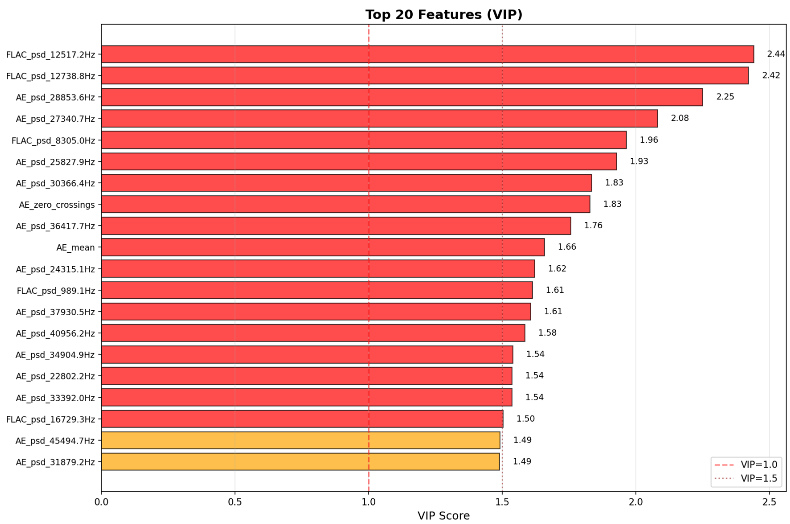


Figure 4.11: VIP scores for the full-spectrum PLS-DA model. The highest-ranked variables included FLAC frequencies around 12.5 to 12.75 kHz and AE frequencies around 27 to 30 kHz, together with selected time-domain acoustic variables.

The increase from 97.8% to 100% classification accuracy indicates that the higher-resolution representation retained some information that was reduced by the broader frequency bands. The difference in overall performance was nevertheless small, showing that the lower-dimensional band representation already retained most of the information required to separate the operating conditions within the available dataset.

4.4.3 PLS Regression: Acoustic Features Only

To determine whether blender speed and throughput could be estimated as continuous responses from the acoustic measurements, a multi-output PLS regression model was fitted jointly to RPM and throughput. Six latent components were retained based on the mean cross-validated Q^2 across the two responses, which reached 0.511. The variation in training R^2 and cross-validated Q^2 with model complexity is shown in Figure 4.12, from which six latent components were retained.

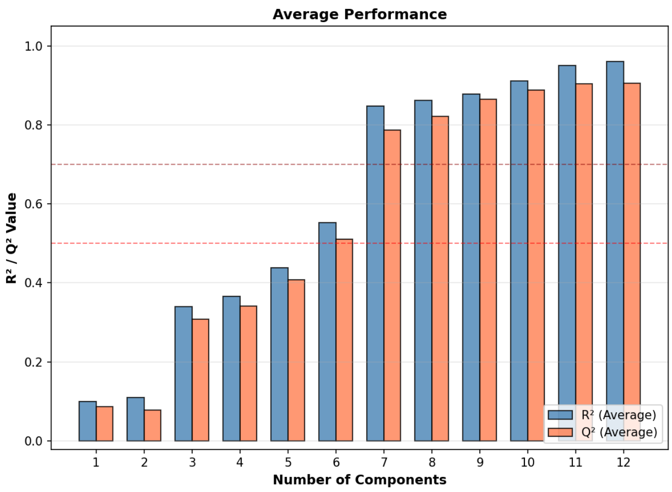


Figure 4.12: Cross-validated R^2 and Q^2 evaluation for the acoustic-only multi-output PLS regression model. Six latent components were retained for prediction of RPM and throughput.

The observations in the resulting PLS latent-variable space are shown in Figure 4.13 and performance is summarised in Table 4.4.

Output	R^2	RMSE	MAE	RMSE (% range)	MAE (% range)	Range
RPM	0.936	20.95	16.25	10.5%	8.1%	250 to 450
Throughput	0.931	3.27 kg/h	2.52 kg/h	10.9%	8.4%	20 to 50 kg/h

Table 4.4: PLS regression performance for RPM and throughput using acoustic features.

The correlation between measured and predicted values was 0.97 for RPM and 0.96 for throughput. A moderate negative correlation of approximately −0.46 was present between RPM and throughput across the experimental design. Although the two responses were modelled jointly, predictions and performance metrics were evaluated separately for RPM and throughput. The prediction of both responses indicates that the acoustic feature space contained information associated with each operating variable beyond the correlation between them.

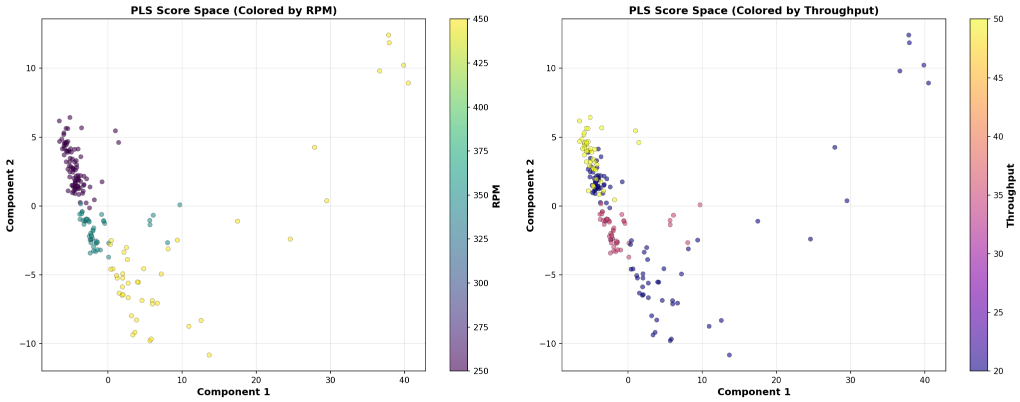


Figure 4.13: PLS score-space representation for the acoustic-only regression model. The projection shows the distribution of the acoustic observations in the latent-variable space used for prediction of RPM and throughput.

The response-specific VIP distributions for RPM and throughput are shown in Figure 4.14. The VIP distributions differed between the two responses. For RPM, the highest-ranked variables included several AE frequencies. Frequencies at approximately 28,854, 112,059, 109,034, and 90,880 Hz had VIP scores of 1.69, 1.50, 1.44, and 1.44, respectively. FLAC frequencies around 1,211 and 324 Hz also contributed, with VIP scores of 1.41 and 1.39.

Throughput showed a different importance pattern. FLAC frequencies around 12,517 and 12,739 Hz had the highest VIP scores of 3.28 and 3.24, respectively. The FLAC frequency around 8,305 Hz had a VIP score of 2.54, while the AE frequency at approximately 28,854 Hz and the FLAC frequency around 989 Hz had VIP scores of 2.30 and 2.01. In the overall VIP ranking, the 12,517 and 12,739 Hz FLAC variables remained highest, followed by the AE variable at approximately 28,854 Hz and the FLAC variable at approximately 8,305 Hz.

The different VIP profiles show that RPM and throughput were associated with different regions of the acoustic feature space. RPM placed greater importance on several AE frequency variables, whereas throughput was more strongly associated with frequencies in the FLAC recording.

Together with the classification results, this confirms that both operating variables are represented in the acoustic feature space.

4.4.4 PLS Regression: Acoustic and Process Features

The process measurements were subsequently added to the acoustic feature set to examine whether they provided complementary information for prediction of RPM and throughput. At this stage, the individual feeder mass flow rates and net weight signals were used directly as model inputs. The retained observations from the final

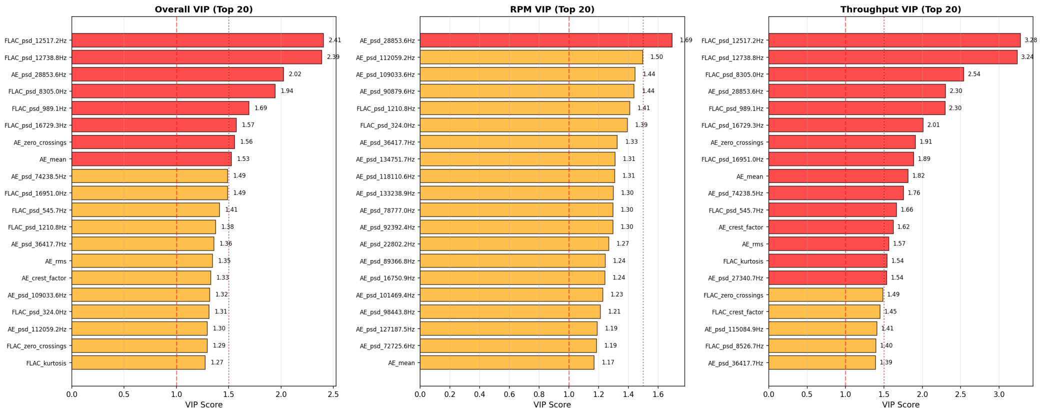


Figure 4.14: VIP distributions for RPM and throughput in the acoustic-only PLS regression model. The two responses showed different importance patterns across the AE and FLAC feature space.

Output	Metric	Training	Validation	Test
RPM	R^2	0.976	0.903	0.845
RPM	RMSE (% range)	6.5%	13.1%	16.7%
Throughput	R^2	0.991	0.963	0.927
Throughput	RMSE (% range)	3.2%	6.6%	9.3%
Average	R^2	0.984	0.933	0.886

Table 4.5: PLS regression performance for RPM and throughput using acoustic and process variables.

third of each run were divided chronologically into training, validation, and test partitions as described in Section 3.8.5. An example of the chronological partitioning applied to the retained observations is shown in Figure 4.15.

Models containing different numbers of latent components were fitted using the training partition and evaluated on the chronological validation partition. Ten latent components produced the highest average validation R^2 across RPM and throughput and were therefore retained for the final model. At this model complexity, the average cross-validated Q^2 calculated from the training data was approximately 0.976. The corresponding training R^2 and cross-validated Q^2 values across model complexity are shown in Figure 4.16.

The corresponding PLS score-space representations for the training, validation, and test partitions are shown in Figure 4.17. The prediction performance of the retained ten-component model is summarised in Table 4.5.

The inclusion of process variables substantially changed the VIP ranking (refer to figure 4.18). The overall VIP distribution was dominated by feeder mass flow measurements. Feeder 4 mass flow had the highest score at approximately 3.2, followed by feeder 6 mass flow at approximately 2.8 and feeder 2 mass flow at approximately 2.5. Feeder 2 net weight followed with a VIP score of approximately 1.9. The highest-ranked acoustic variable was the FLAC frequency at 12,739 Hz,

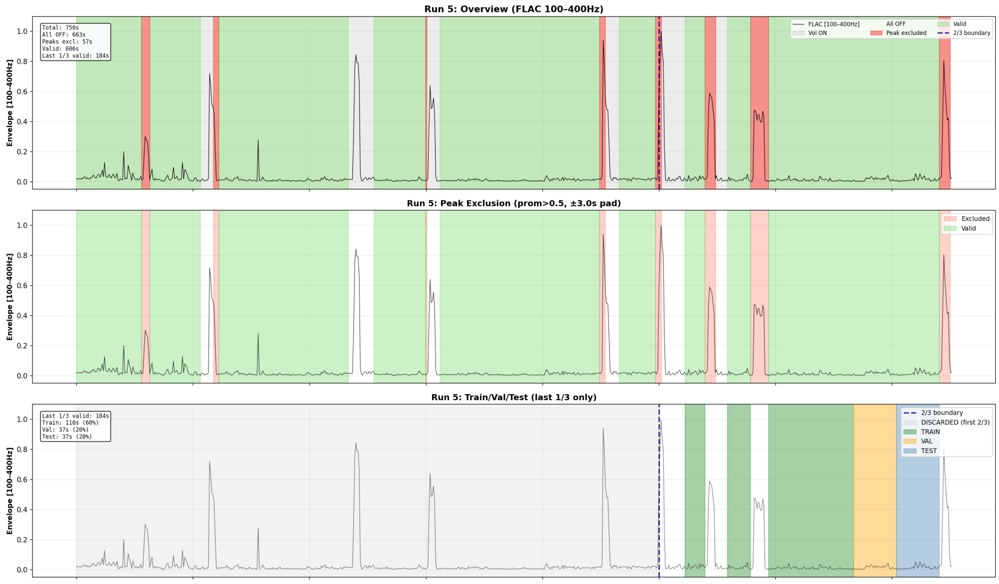


Figure 4.15: Example of the chronological partitioning used for the combined acoustic and process model. The retained observations from the selected part of the run were divided into training, validation, and test regions in temporal order.

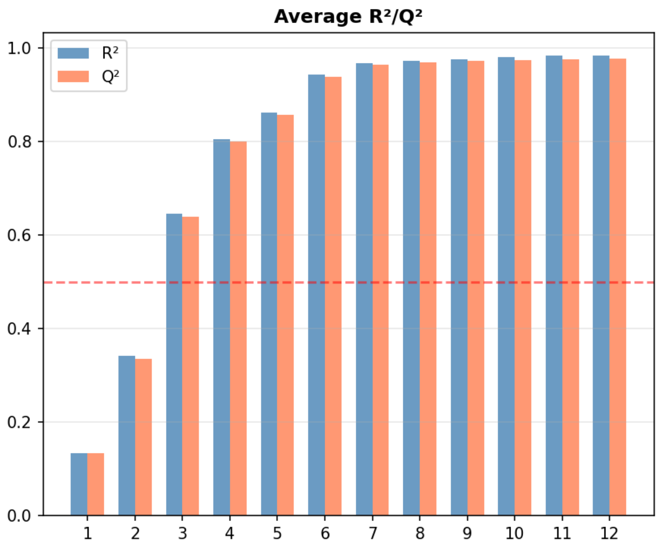
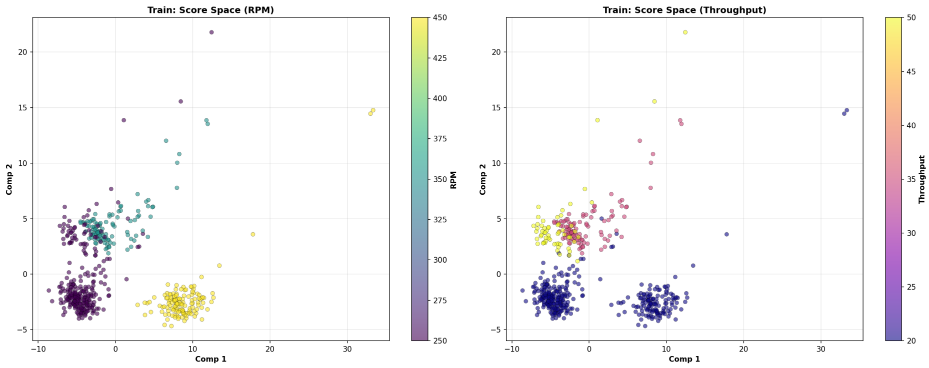
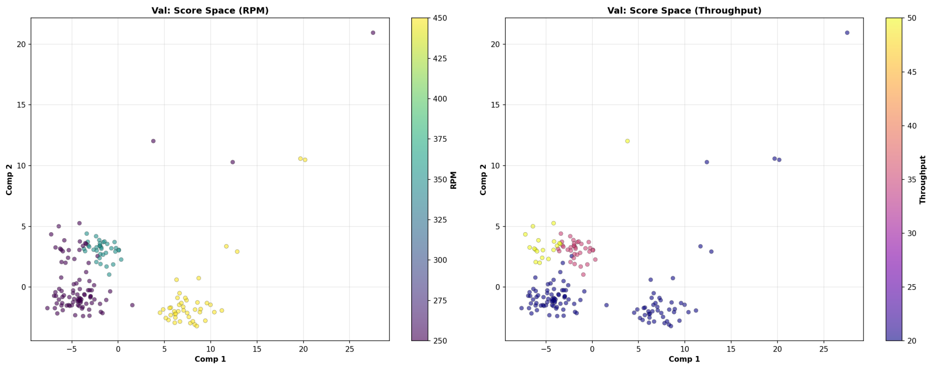


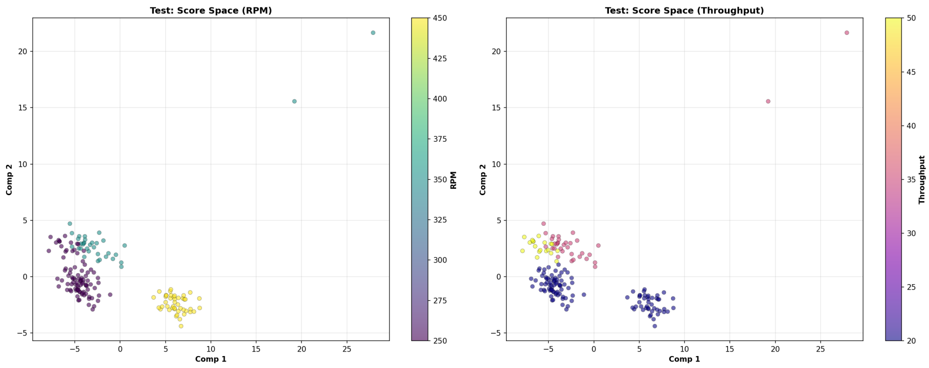
Figure 4.16: Training R^2 and five-fold cross-validated Q^2 for the PLS regression model containing acoustic and process variables as a function of the number of latent components.



(a) Training observations



(b) Validation observations



(c) Test observations

Figure 4.17: PLS score-space representations for the model containing acoustic and process variables, shown separately for the chronological training, validation, and test partitions. The first two latent components are shown to illustrate the distribution of observations within the model score space.

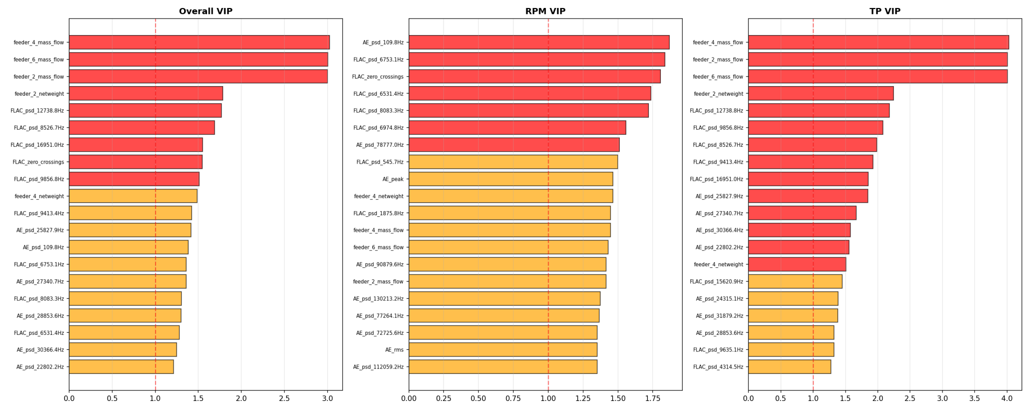


Figure 4.18: VIP distributions for the PLS regression model containing acoustic and process variables. Feeder measurements contributed strongly to throughput prediction, while acoustic variables remained more important for RPM prediction.

with a VIP score of approximately 1.7. This variable had ranked among the most important variables in the acoustic model but moved below the feeder measurements when the process variables were included.

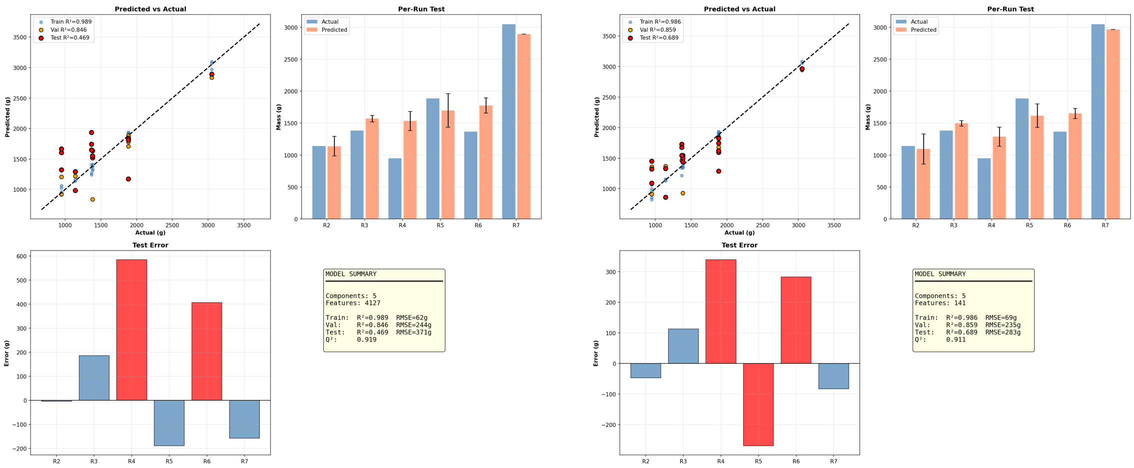
The response-specific VIP distributions showed a clear difference between RPM and throughput. For throughput, the four highest-ranked variables were feeder 4 mass flow, feeder 2 mass flow, feeder 6 mass flow, and feeder 2 net weight, with VIP scores of approximately 4.2, 3.5, 3.2, and 2.5, respectively. The strong contribution of the mass flow measurements was expected because these variables directly describe the material feed rates contributing to the throughput condition.

For RPM, the ranking remained dominated by acoustic variables. The AE frequency variable at approximately 109.8 kHz had the highest VIP score at approximately 1.8, followed by the FLAC frequency at 6,753 Hz with a VIP score of approximately 1.5, FLAC zero crossings at approximately 1.4, and the FLAC frequency at 6,531 Hz at approximately 1.3. Process variables, including feeder 4 net weight and the mass flow rates of feeders 4 and 6, appeared further down the ranking with VIP scores below 1.0.

These results show that the acoustic and process measurements contributed differently to the two responses. Direct feeder measurements provided particularly strong information for throughput, whereas acoustic variables retained greater importance for RPM prediction.

The combined model achieved a higher average training R^2 than the acoustic model, increasing from 0.934 to 0.984. At their respective retained model complexities, the average cross-validated Q^2 was 0.511 for the six-component acoustic model and approximately 0.976 for the ten-component model containing acoustic and process variables. The latter component count was selected from the chronological validation performance rather than from the Q^2 value.

The combined model nevertheless showed a reduction in average R^2 from 0.984 for the training partition to 0.886 for the test partition, corresponding to a decrease of 0.098. The additional process predictors therefore improved the description of the training observations but did not provide an equivalent improvement on the later chronological test data.



(a) Full spectrum (b) 50-bin representation

Figure 4.19: Residence mass predictions obtained during model fitting using (a) the full-spectrum representation containing 4127 predictors and (b) the reduced 50-bin representation containing 141 predictors. Both representations achieved similarly high training performance despite the substantial difference in predictor dimensionality.

The reduction from training to later prediction performance further motivated the evaluation of residence mass using complete experimental runs as the independent validation units. This run-level evaluation is presented in Section 4.5.

4.5 Residence Mass Prediction

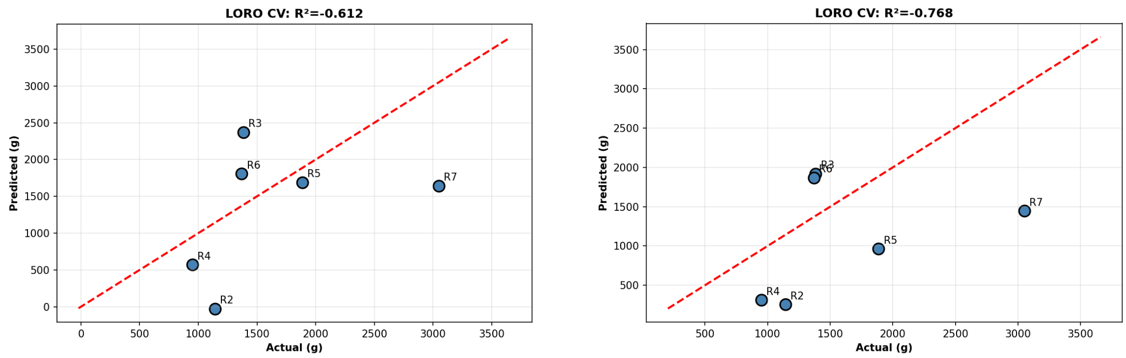
The initial residence mass models showed that good performance within the available experimental runs did not necessarily translate to prediction of an omitted run. The full-spectrum representation contained 4127 predictors, whereas reducing the spectrum to 50 frequency bins resulted in 141 predictors. Despite this substantial reduction in dimensionality, both representations produced similarly high training R^2 values.

Their chronological validation performance was also similar, with $R^2 = 0.846$ for the full-spectrum model and $R^2 = 0.859$ for the 50-bin model. A larger difference appeared on the chronological test partition, where the full-spectrum model achieved $R^2 = 0.469$ compared with $R^2 = 0.689$ for the 50-bin representation. The corresponding training behaviour is shown in Figure 4.19.

When complete runs were withheld, however, both representations showed poor generalisation. As shown in Figure 4.20, the full-spectrum model produced a LORO R^2 of -0.77, while the 50-bin representation produced a LORO R^2 of -0.612.

These results showed that reducing the number of spectral predictors from 4127 to 141 was not sufficient to obtain reliable prediction of an omitted experimental run. The subsequent analysis therefore moved to a more compact representation based on broad frequency regions.

The negative LORO R^2 values indicate that prediction of the omitted runs was less



(a) Full spectrum (b) 50-bin representation

Figure 4.20: Leave-one-run-out residence mass predictions using (a) the full-spectrum representation and (b) the 50-bin representation. The full-spectrum model produced a LORO R^2 of -0.77, while the 50-bin model produced a LORO R^2 of -0.612. Although reducing the predictor dimensionality improved the result, both models showed poor prediction of an omitted experimental run.

accurate than using the mean residence mass of the corresponding training runs. The contrast between the strong within-run results and the poor LORO performance was particularly important because the chronological training, validation, and test partitions still contained observations originating from the same experimental runs, whereas LORO withheld one complete experimental run.

Although reducing the predictor space from 4127 to 141 variables improved both the chronological test result and the LORO result, the remaining predictor space was still large relative to the six independently measured residence mass conditions. The subsequent analysis therefore focused on broader frequency-band representations and examined the effects of band definition, block duration, block sampling, weighting, and process information on cross-run prediction performance.

4.5.1 Band Configuration Comparison

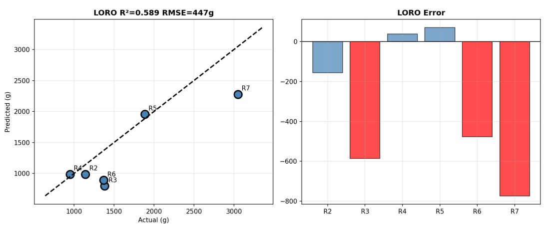
Following the poor LORO performance of the higher-dimensional spectral representations, the acoustic spectra were represented using broader frequency bands. Several candidate band configurations were compared using 30 s blocks and convergence-based sampling while keeping the remaining model-development procedure unchanged. Table 4.6 summarises the resulting LORO performance and the features retained for each configuration. The broader band configurations produced positive LORO R^2 values and repeatedly retained similar AE frequency regions together with a high-frequency FLAC feature.

Performance deteriorated as the spectral representation was divided into a larger number of narrower bands. The contrast between the broader and finer spectral representations is illustrated in Figure 4.21, where the broader configuration retained positive LORO performance while the finer representation produced substantially poorer cross-run prediction.

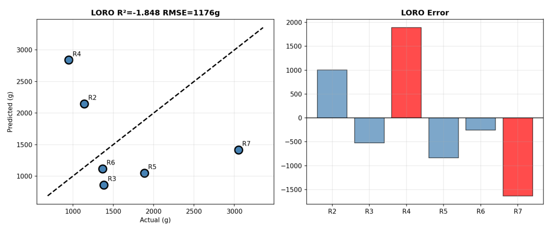
For the available dataset, increasing the spectral resolution therefore did not improve prediction of an omitted run. Based on this comparison, the reduced broad-band

Configuration	Total Features	LORO R^2	Retained Features
2 AE + 4 FLAC	14	0.59	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 15 to 22 kHz
4 AE + 4 FLAC	18	0.59	Same as above
4 AE + 5 FLAC	20	0.59	Same as above
5 AE + 4 FLAC	20	0.37	Same as above
7 AE + 6 FLAC	28	−0.98	AE 8 to 20 kHz slope AE 40 to 70 kHz slope FLAC 16 to 22 kHz
8 AE + 5 FLAC	28	−1.85	AE 5 to 12 kHz slope AE 20 to 35 kHz slope FLAC 18 to 22 kHz

Table 4.6: Effect of frequency band configuration on LORO residence mass prediction. The retained features were obtained using the same feature selection procedure for each configuration.



(a) 2 AE + 4 FLAC



(b) 8 AE + 5 FLAC

Figure 4.21: Representative leave-one-run-out residence mass predictions for two frequency-band configurations. (a) The broader 2 AE + 4 FLAC representation achieved a positive LORO R^2 of 0.59. (b) Increasing the spectral resolution to 8 AE + 5 FLAC resulted in a LORO R^2 of −1.85. The comparison illustrates the deterioration in cross-run prediction observed when the spectrum was divided into a larger number of narrower frequency bands.

representation described in Section 3.9.2 was retained for the subsequent modelling stages.

4.5.2 Effect of Block Duration, Sampling, and Weighting

The initial block duration comparison suggested that prediction performance improved when longer blocks were used with convergence based sampling. With this sampling procedure, 10 and 15 s blocks produced negative LORO R^2 values, whereas 20 and 30 s blocks produced positive values. The 30 s configuration achieved a LORO R^2 of 0.59. The effect of block duration under convergence-based sampling is illustrated in Figure 4.22, which compares the retained blocks and LORO predictions for representative block durations.

Duration (s)	**Total Blocks**	**Training R^2**	**LORO R^2**	**LORO RMSE (g)**
10	59	0.953	−1.28	1052
15	40	0.969	−1.37	1074
20	29	0.981	0.44	520
30	19	0.958	0.59	447

Table 4.7: Comparison of block durations using convergence based sampling and the selected frequency band configuration.

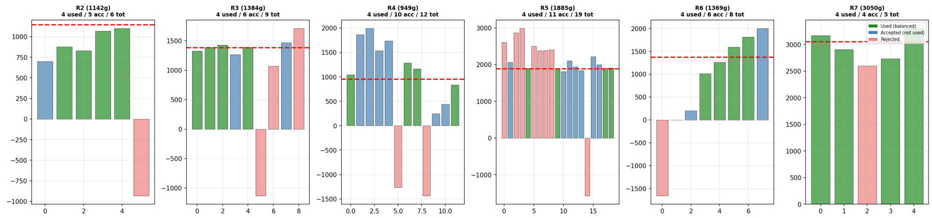
This interpretation changed when the sampling procedure was modified. Using 15 s blocks together with random sampling, density weighting, and confidence weighting increased the LORO R^2 to 0.854. The improvement relative to the 30 s convergence result showed that block duration could not be considered independently from the strategy used to select blocks from each run.

The four sampling methods were therefore compared using the same 15 s block duration and reduced frequency band representation. The results are shown in Table 4.8.

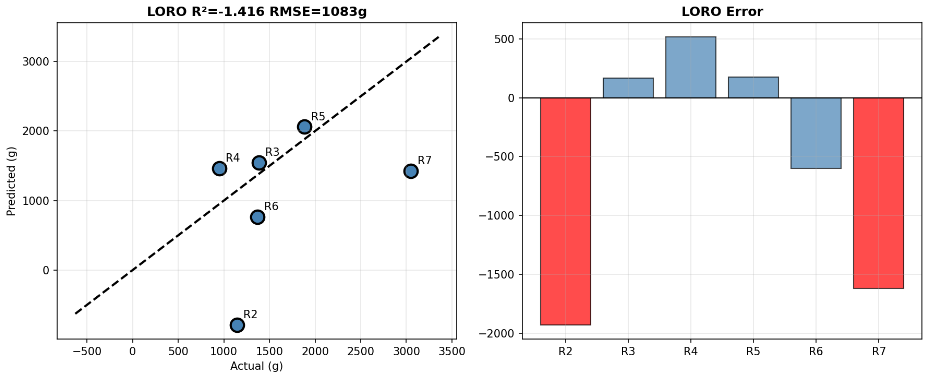
Convergence based sampling produced the highest training R^2 of 0.969 but the lowest LORO performance, with an R^2 of -1.37. Random sampling produced a lower training R^2 of 0.940 but the highest LORO R^2 of 0.854. The difference between the training and LORO results shows that training fit alone was not a reliable criterion for selecting the sampling strategy.

The behaviour of the convergence procedure can be related to its convergence based sampling step. Blocks that differed strongly from the provisional fitted relationship were progressively removed, causing the retained observations to agree more closely with the current training model. With only a small number of independent experimental runs, this procedure reduced the variability represented during model fitting without improving prediction of an omitted run.

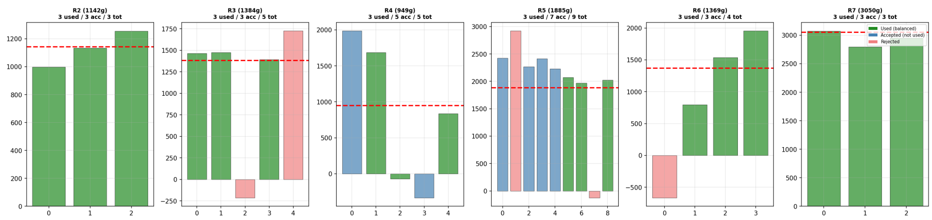
Random sampling did not use the fitted model to decide which blocks should be retained. It therefore preserved a broader representation of the valid block to block variability within each experimental run.



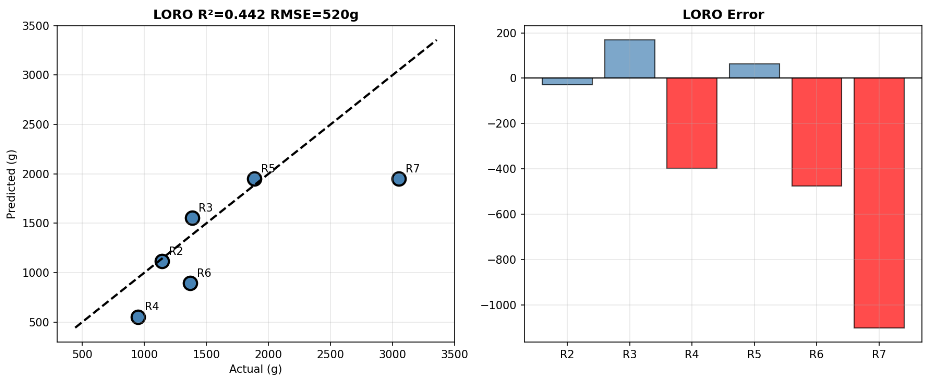
(a) Retained blocks, 10 s



(b) LORO prediction, 10 s



(c) Retained blocks, 20 s



(d) LORO prediction, 20 s

Figure 4.22: Effect of block duration on block selection and leave-one-run-out residence mass prediction under convergence based sampling. The 10 s representation produced poorer cross-run prediction, whereas the 20 s representation achieved a positive LORO result.

Method	Train R^2	LORO R^2	RMSE (g)	Selected Features
Random	0.940	0.854	266	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 15 to 22 kHz mean
Stratified	0.898	0.658	407	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 5 to 15 kHz mean
Combined	0.926	0.511	487	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 15 to 22 kHz slope
Convergence	0.969	-1.37	1074	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 15 to 22 kHz slope

Table 4.8: Comparison of block sampling methods for the acoustic only residence mass model using 15 s blocks and three selected features.

Density weighting was also important because the residence mass values were unevenly distributed. Run 7 represented the only experimental condition in the upper part of the measured residence mass range. The weighting procedure increased the influence of this sparsely represented condition rather than allowing the cluster of lower residence mass values to dominate the regression objective.

4.5.3 Run-Level Cross-Validation Performance

The individual LORO predictions obtained using the best performing acoustic only configuration are shown in Table 4.9 and figure 4.23. Each prediction was produced by a model fitted without using the corresponding experimental run during model development.

Run	Actual (g)	Predicted (g)	Error (g)	Error (%)	RPM	TP (kg/h)
2	1142	1431	+289	+25.3	444	49.2
3	1384	1251	−133	−9.6	347	34.5
4	949	1221	+272	+28.7	441	19.5
5	1885	2263	+378	+20.1	247	19.8
6	1369	1071	−298	−21.8	344	33.8
7	3050	3191	+141	+4.6	248	48.9

Table 4.9: Per run leave one run out predictions for the acoustic only residence mass model using random sampling with density and confidence weighting.

Run 7 had the smallest relative prediction error at 4.6%, despite representing the only residence mass measurement in the upper part of the calibration range. This result was consistent with the intended effect of density weighting, which increased

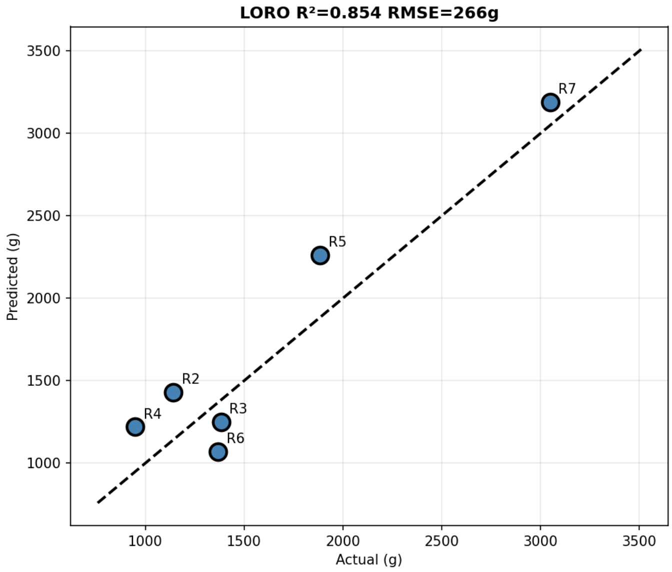


Figure 4.23: Leave-one-run-out residence mass predictions for the best performing acoustic-only model using 15 s blocks, random sampling, and density and confidence weighting. Each point represents a prediction for one experimental run obtained from a model developed without that run. The model achieved a LORO R^2 of 0.854 and an RMSE of 266 g.

the contribution of sparsely represented residence mass conditions during model fitting.

The largest relative errors occurred for Runs 2 and 4, with errors of 25.3% and 28.7%, respectively. Both runs were operated at approximately 440 RPM, whereas the remaining runs were operated between approximately 247 and 347 RPM. This common operating condition suggests that some of the remaining prediction error may have been associated with blender speed.

This observation is consistent with the operating-condition analysis in Section 4.4, which showed that RPM was strongly represented in the acoustic measurements. However, the available runs were not sufficient to isolate blender speed as the cause of the larger errors. The result motivated the subsequent evaluation of whether additional process information improved residence mass prediction across operating conditions.

4.5.4 Selected Acoustic Features

The acoustic-only modelling identified a compact group of variables that showed strong relationships with residence mass. The overall correlation ranking of the candidate acoustic features is shown in Figure 4.24. The strongest individual correlation was obtained for the FLAC spectral slope between 15 and 22.05 kHz, with

$$r = +0.923,$$

followed by the mean FLAC spectral power in the same frequency range,

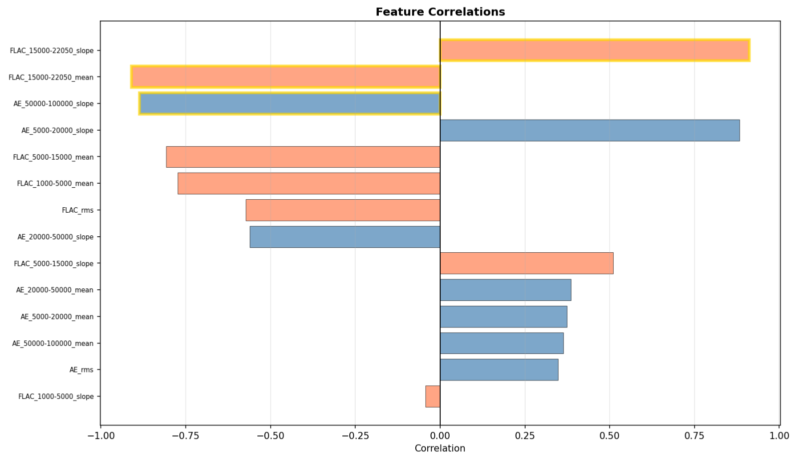


Figure 4.24: Correlation of the candidate acoustic features with residence mass. The strongest individual relationships were observed for the FLAC spectral slope and mean spectral power between 15 and 22.05 kHz, together with the AE spectral slope between 5 and 20 kHz.

$$r = -0.911,$$

and the AE spectral slope between 5 and 20 kHz,

$$r = +0.882.$$

The correlation ranking describes the relationship between each acoustic variable and residence mass individually. Within the PLS model, the predictors are considered together through the latent-variable representation, and their relative contributions can therefore differ from the individual correlation ranking. The corresponding PLS weights for the best-performing acoustic-only model are shown in Figure 4.25. The three variables retained by the best-performing acoustic-only model were the AE spectral slope between 5 and 20 kHz, the AE spectral slope between 50 and 100 kHz, and the mean FLAC spectral power between 15 and 22.05 kHz. The AE slope between 50 and 100 kHz showed an overall negative correlation with residence mass of

$$r = -0.878.$$

The two AE variables describe changes in spectral shape across different frequency regions, while the FLAC variable represents the average spectral magnitude within its selected frequency range. The final feature set therefore combined information from both spectral shape and spectral magnitude and included contributions from both acoustic measurements.

These relationships show that residence mass related information was distributed across several complementary acoustic descriptors rather than being represented by a single dominant feature. The observed correlations and PLS weights describe empirical relationships within the available experiments and do not by themselves establish a unique physical mechanism for the spectral changes.

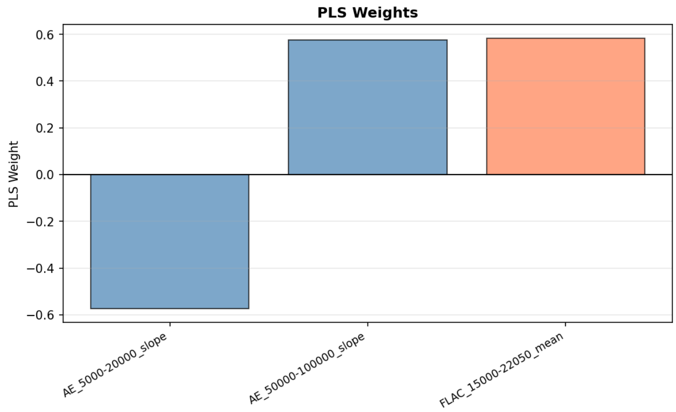


Figure 4.25: PLS weights for the best-performing acoustic-only residence mass model. The fitted model placed strong contributions on the AE spectral slopes between 5 and 20 kHz and 50 and 100 kHz, together with the mean FLAC spectral power between 15 and 22.05 kHz.

4.5.5 Effect of Process Features and Sampling Sensitivity

The three engineered process variables described in Section 3.11 were subsequently added to the candidate feature pool. Their inclusion substantially changed the correlation ranking. The dynamic ratio, P_{f3}, showed the strongest individual correlation with residence mass,

$$r = -0.934,$$

while P_{f2} showed a positive correlation of

$$r = +0.857.$$

The first process feature, P_{f1}, showed a weaker relationship,

$$r = -0.320.$$

Table 4.10 shows the highest ranked variables from the combined acoustic and process feature set.

For random and stratified sampling, the four selected predictors were the AE slope from 5 to 20 kHz, the AE slope from 50 to 100 kHz, P_{f2}, and P_{f3}. The two process features therefore replaced the FLAC mean power feature selected in the acoustic only model. The corresponding PLS weights for the combined acoustic and process model are shown in Figure 4.26.

Despite their high individual correlations with residence mass, inclusion of the process features did not improve LORO performance. All residence mass PLS models in this comparison used one latent component, Table 4.11 summarises the results for the combined model.

The best combined acoustic and process configuration in Table 4.11 used random sampling without a fixed seed and achieved a LORO R^2 of 0.760 with a LORO RMSE of 341 g. In comparison, the best acoustic-only model achieved a LORO R^2 of 0.854 and a LORO RMSE of 266 g. The addition of the process features therefore

Rank	Feature	Correlation
1	P_{f3} dynamic variability ratio	−0.934
2	FLAC 15 to 22.05 kHz slope	+0.923
3	FLAC 15 to 22.05 kHz mean	−0.911
4	AE 5 to 20 kHz slope	+0.882
5	AE 50 to 100 kHz slope	−0.878
6	P_{f2} flow and net-weight-change ratio	+0.857
7	FLAC 5 to 15 kHz mean	−0.827
8	FLAC 1 to 5 kHz mean	−0.794
9	FLAC 0.1 to 1 kHz slope	−0.750
10	FLAC RMS	−0.582
	...	
19	P_{f1} torque, speed, and throughput ratio	−0.320

Table 4.10: Correlation ranking of the acoustic and engineered process features used for residence mass modelling.

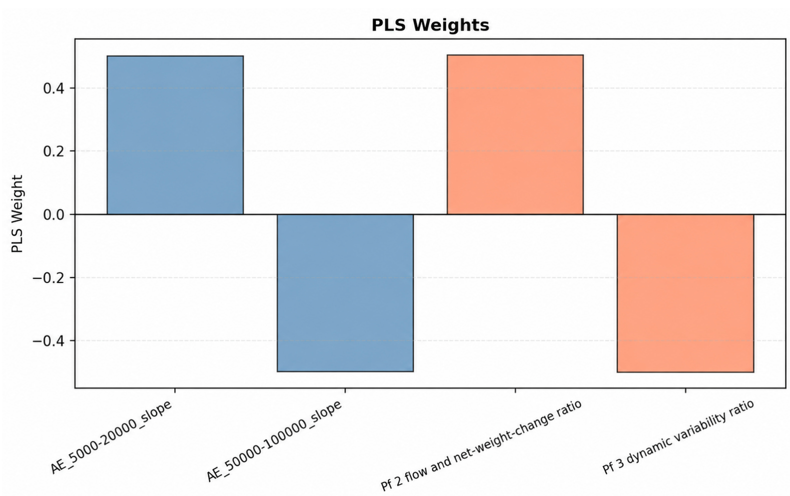


Figure 4.26: PLS weights for the residence mass model containing acoustic and engineered process variables. The fitted model combined the AE spectral slopes from 5 to 20 kHz and 50 to 100 kHz with the engineered process variables P_{f2} and P_{f3}. The weights describe the contribution of the retained predictors within the PLS latent-variable representation.

Seed	Method	Train R^2	LORO R^2	LORO RMSE (g)	Bias (g)
Fixed	Random	0.993	0.214	618	−424
Fixed	Stratified	0.993	0.430	526	−512
Fixed	Combined	0.975	−0.360	812	−344
None	Random	0.991	0.760	341	−367
None	Stratified	0.991	0.550	468	−481
None	Combined	0.976	−0.360	813	−240
None	convergence	0.992	0.147	643	+1679

Table 4.11: Sampling method comparison for the residence mass model including acoustic and engineered process features. All models used one PLS latent component.

reduced the LORO R^2 by approximately 0.09 and increased the LORO RMSE by 75 g.

This result shows that strong correlation within the available six runs did not necessarily correspond to better prediction of an omitted condition. The engineered process features distinguished differences among the available experimental conditions, but the relationships represented by these variables were less stable under LORO evaluation than the three-feature acoustic representation.

The combined model was also sensitive to the blocks selected during sampling. For random sampling, use of a fixed seed produced a LORO R^2 of 0.214, whereas the best result from 100 stochastic repetitions reached 0.760. The variation in the blocks retained under different random seeds is illustrated in Figure 4.27. This difference shows that model performance depended substantially on the particular blocks selected from the relatively small number available in each run. The resulting difference in run-level prediction performance is shown in Figure 4.28.

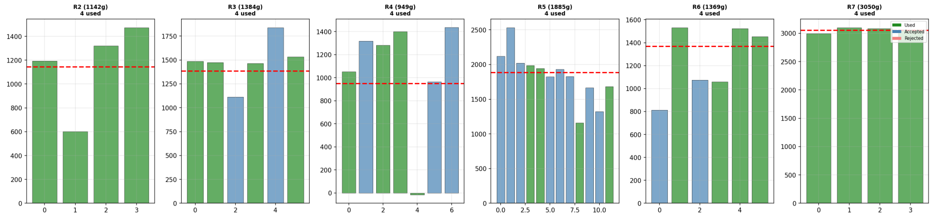
The best of 100 procedure increased the probability of finding a favourable block combination according to the implemented model selection criterion. Consequently, the maximum LORO performance obtained from the stochastic procedure should be interpreted together with this sampling sensitivity rather than as a completely deterministic estimate of cross run performance.

4.5.6 Final Model Comparison

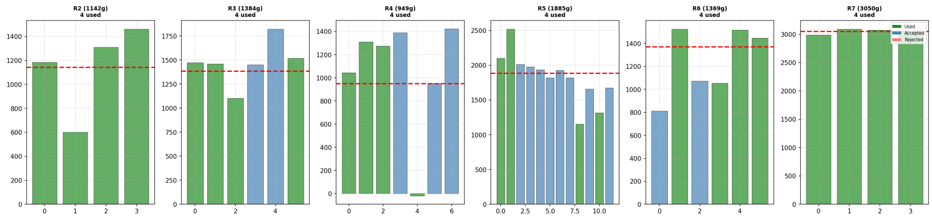
The comparison of spectral representations, block durations, sampling procedures, weighting strategies, and feature sets identified the acoustic-only model with random sampling as the best-performing configuration under LORO validation. The selected configuration is summarised in Table 4.11.

The acoustic-only model achieved a training R^2 of 0.940, a LORO R^2 of 0.854, and a LORO RMSE of 266 g. It therefore provided the strongest cross-run performance among the investigated configurations and was retained as the reference model for assessing the effect of adding engineered process features.

Adding the engineered process variables did not improve cross-run prediction. The best combined acoustic and process model achieved a LORO R^2 of 0.760 and a

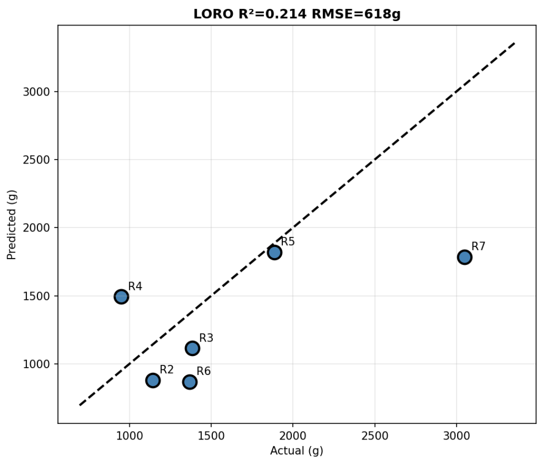


(a) Fixed-seed random sampling

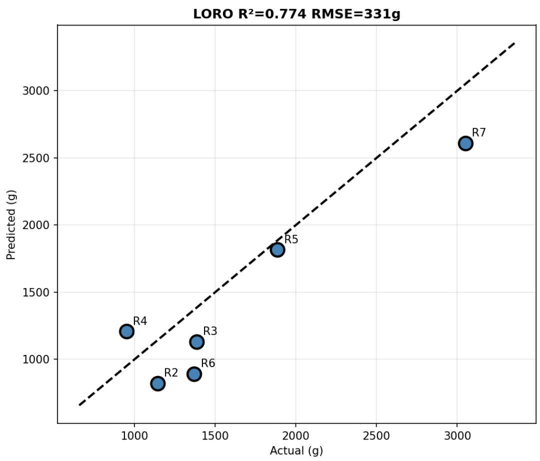


(b) Selected stochastic random sampling

Figure 4.27: Effect of the random seed on block selection for the combined acoustic and process residence mass model using random sampling. Different seeds resulted in different combinations of retained blocks from the available runs.



(a) Fixed-seed random sampling



(b) Selected stochastic random sampling

Figure 4.28: Leave-one-run-out residence mass predictions for the combined acoustic and process model using random sampling. The fixed-seed configuration achieved a LORO R^2 of 0.214, whereas the selected stochastic configuration achieved a LORO R^2 of 0.760.

Parameter	Value
Features	AE 5 to 20 kHz slope AE 50 to 100 kHz slope FLAC 15 to 22.05 kHz mean
Block duration	15 s
Sampling method	Random, best of 100 draws
Weighting	Density + confidence
PLS components	1
Training R^2	0.940
LORO R^2	0.854
LORO RMSE	266 g

Table 4.12: Best performing acoustic only residence mass model under LORO validation.

LORO RMSE of 341 g. The combined model was nevertheless carried forward to the additional recordings to examine whether the acoustic and process representation could be applied under the separately acquired measurement conditions.

4.6 Application to Additional Recordings

The final inference procedure was applied to the separately acquired recordings using the combined acoustic and process model. The same blender and high-frequency PCB microphone were used as in the training experiments, but the acquisition hardware and sampling rate differed. The additional recordings were therefore used to examine how the developed modelling procedure behaved when applied to data acquired through a different acquisition configuration.

4.6.1 Application and Resampling

As described in Section A.2.2, the additional high-frequency recordings were resampled to approximately 300 kHz before application of the trained model. The effect of this preprocessing step was also examined by comparing the prediction behaviour with that obtained from the original 1 MHz recordings.

Processing the additional recordings at their original sampling rate produced a larger systematic prediction offset relative to the training data. After resampling to approximately 300 kHz, the magnitude of the bias required for the subsequent reference correction decreased. This indicates that consistency of the spectral representation between the training and additional recordings was important for model application.

However, the reduction in prediction offset cannot be attributed to sampling rate alone, since the additional recordings also differed in sampling rate, signal conditioning, and acquisition hardware.

4.6.2 Bias Correction and Predictions

Run 15 was used as the reference recording because its residence mass had been measured gravimetrically as 873 g. This measurement was used to calculate the additive reference correction described in Section 3.12.2.
The reference correction was examined for models obtained using random, stratified, combined, and convergence based block sampling. The calculated reference bias depended on the sampling procedure and on the fitted model. Table 4.13 summarises the bias values obtained from the bias-corrected executions.

Method	**Seed**	**Batch Bias (g)**
Random	Fixed	−424
Stratified	Fixed	−512
Combined	Fixed	−344
Random	None	−470
Stratified	None	−446
Combined	None	−278
convergence	None	+1679

Table 4.13: Reference bias calculated from Run 15 for the acoustic and process models obtained using different block sampling procedures.

For the random, stratified, and combined sampling procedures, the reference prediction was lower than the measured residence mass of Run 15, resulting in negative bias values. The convergence based model behaved differently and produced a large positive bias. This indicates that the prediction offset obtained for the additional measurement configuration depended on the fitted model and the observations used during model development.
The no fixed seed models also showed some variation between repeated executions because different combinations of training blocks could be selected. For example, the random sampling model gave a LORO R^2 of 0.760 in one execution and 0.774 in another. This behaviour is consistent with the sampling sensitivity described in Section 4.5.5. A direct comparison between these two executions would therefore combine the effect of a different fitted model with the effect of bias correction.
To isolate the effect of the additive correction, the before and after comparison was therefore made using random sampling with a fixed seed. This ensured that the same training blocks and the same fitted PLS model were used in both cases. The model performance was consequently unchanged, with a training R^2 of 0.993, a LORO R^2 of 0.214, and a LORO RMSE of 618 g.
For this model, the reference bias calculated from Run 15 was approximately −424 g. Using the bias definition

$$b = \overline{\hat{m}}_{\mathrm{ref,raw}} - m_{\mathrm{ref}}, \tag{4.1}$$

and the measured reference mass of 873 g, the mean raw prediction over the selected Run 15 reference region was approximately

Run	Measured Mass (g)	Predicted (g)	Block Std. (g)	RPM	TP (kg/h)
15	873	956	856	399	29.6
14		248	926	400	29.7
13		2841	999	400	29.7
11		3242	389	299	29.7
10		1025	3662	298	29.0
8		3371	799	298	29.3

Table 4.14: Bias corrected residence mass predictions for the additional recordings using the selected acoustic and process model with random sampling and no fixed seed. Run 15 had a measured residence mass of 873 g and was used as the reference for the additive correction. Block standard deviation describes the variation among valid block predictions within each recording.

$$\overline{\hat{m}}_{\text{ref,raw}} = 873 - 424 = 449 \text{ g}. \tag{4.2}$$

The effect of applying the correction is shown in Figure 4.29. The correction shifts the overall prediction level to the measured reference value while leaving the relative variation between the block-level predictions unchanged.

The comparison shows that the reference correction changes the overall prediction level but does not change the relative variation between individual block predictions. It therefore corrects a systematic offset in the reference region but does not reduce temporal variability in the model output.

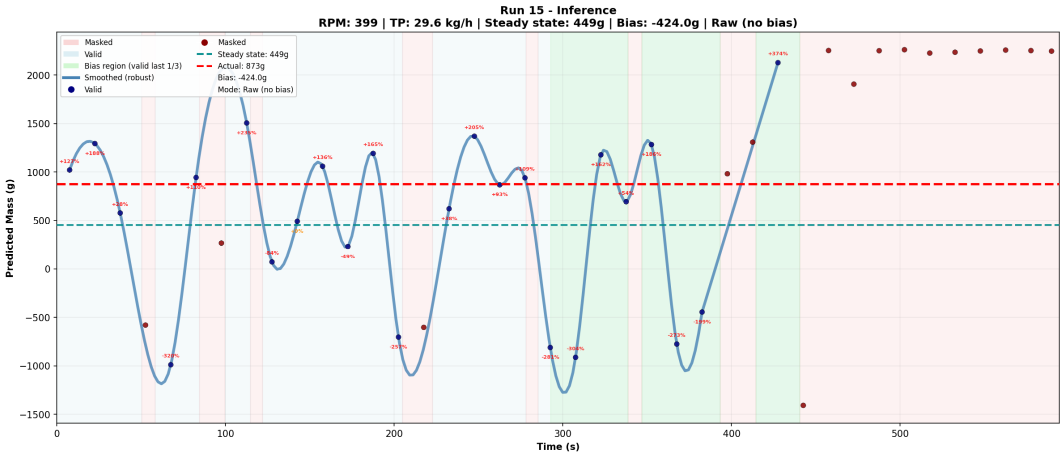
For application to the additional recordings, the selected random sampling model without a fixed seed was used with the reference correction applied. In the final execution, this model gave a LORO R^2 of 0.774 and a LORO RMSE of 331 g. The calculated Run 15 reference bias was -470.7 g. The resulting run level predictions are summarised in Table 4.14.

The corrected reference region of Run 15 was centred on the measured value of 873 g, whereas the run level mean reported in Table 4.14 was 956 g. These values represent different subsets of the recording. The reference correction was determined from the selected late reference region, while the reported run level value was calculated from the broader set of valid blocks. The two values are therefore not expected to be identical.

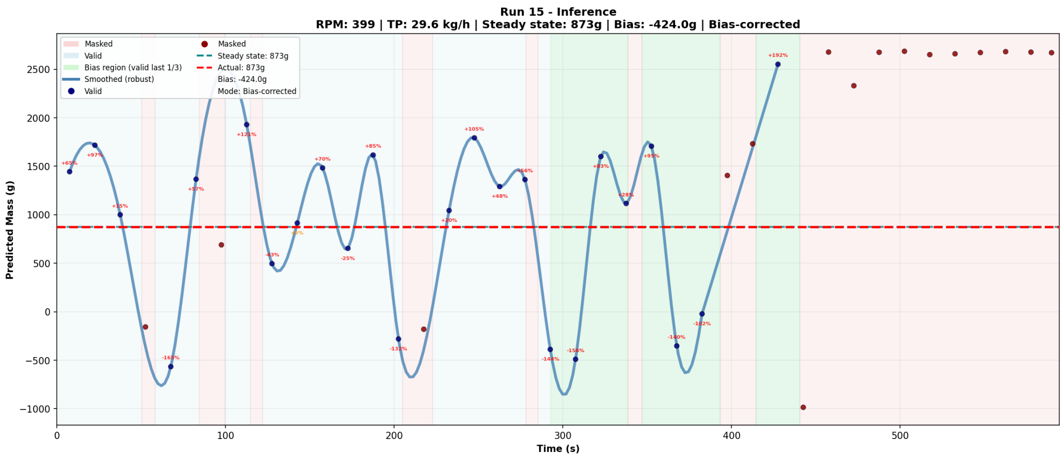
Only Run 15 had a measured residence mass among the additional recordings, and this measurement was used directly to calculate the reference correction. Quantitative prediction errors could therefore not be calculated for Runs 8 to 14. The predictions from these recordings were instead used to examine the behaviour and temporal stability of the inference procedure, as discussed in Section 4.6.3.

4.6.3 Temporal Prediction Behaviour

The block level prediction sequences were also examined over time to assess how the model output behaved within individual recordings. For this analysis, the selected



(a) Without bias correction



(b) With bias correction

Figure 4.29: Residence mass prediction for Run 15 using the same random sampling model with a fixed seed. (a) Without bias correction, the mean prediction over the selected reference region was approximately 449 g. (b) Applying the calculated reference bias of -424 g shifted the reference region to the measured residence mass of 873 g. The temporal pattern of the prediction sequence remained unchanged because the correction was additive.

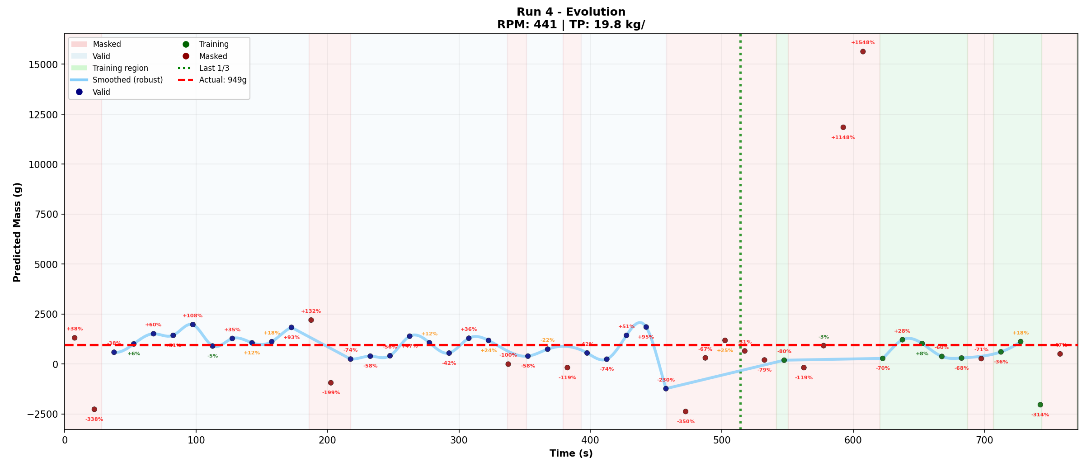


Figure 4.30: Representative temporal residence mass prediction for a training run using the selected random sampling model. The figure shows the block level predictions together with the valid process regions and the smoothed prediction trend.

random sampling model without a fixed seed was used with the reference bias correction applied. All evolution plots presented in this section were generated from the same fitted model.

A representative prediction sequence from the training data is shown in Figure 4.30. The prediction evolution was considered together with the process validity information used during model development. Interpretation was focused primarily on the valid regions, since these corresponded to the process conditions retained for feature extraction and model fitting.

The smoothed prediction curve was used only to provide a clearer visual representation of the temporal trend. The individual block predictions were retained for calculation of the run level prediction statistics.

The same prediction procedure was then examined for the additional recordings. A representative example is shown in Figure 4.31. Compared with the training data, several of the additional recordings showed greater variation between individual block predictions. However, the recording duration and the amount of valid data differed between runs. The block standard deviations reported in Table 4.14 should therefore be interpreted as descriptive measures of prediction spread within each recording rather than as directly comparable measures of stability across all runs.

Differences in temporal behaviour among the additional recordings cannot be attributed uniquely to blender speed or throughput. The recordings differed in the amount of valid data available and may also have been affected by transient process behaviour, differences in the acoustic response, sensor and acquisition characteristics, and variation in the process variables supplied to the model. The available measurements do not allow the individual contribution of these factors to be separated.

The temporal results also illustrate an important distinction between reference bias correction and prediction variability. As shown in Section 4.6.2, the additive correction changes the overall prediction level while preserving the relative temporal

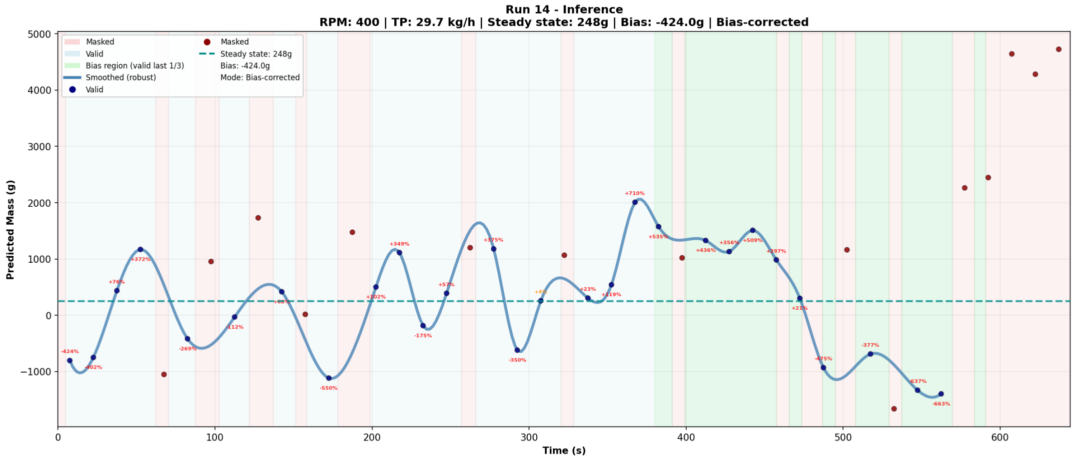


Figure 4.31: Representative temporal residence mass prediction from an additional recording using the selected random sampling model with the Run 15 reference bias correction applied. The block level predictions and smoothed trend illustrate the temporal behaviour of the model output under the additional measurement conditions.

variation in the model output. Bias correction can therefore reduce a systematic prediction offset but does not remove variation between individual block predictions.

Some recordings showed gradual changes in the predicted residence mass over time. These trends may reflect changes in the process state or in the acoustic and process variables used by the model. However, residence mass was not measured continuously during either the training or additional recordings. Changes in the predicted value over time therefore cannot be interpreted as direct measurements of changes in residence mass.

Overall, the evolution plots were useful for examining how the model behaved within individual recordings and for identifying periods of relatively consistent or variable prediction behaviour. Since continuous reference measurements of residence mass were not available, the temporal analysis was used to assess the behaviour of the inference procedure rather than its time resolved prediction accuracy.

5

Conclusions

The results of this study show that acoustic measurements contain information related to the operating state of the continuous blender and can also be used to develop a predictive relationship with residence mass. At the same time, the results highlight the difference between describing relationships within the available experiments and developing a model that remains predictive for an omitted experimental condition.

5.1 Acoustic Response and Influence of Operating Conditions

The spectrogram and SVD analyses showed that both acoustic measurements contained repeatable information associated with blender operation. Filling related events, process interruptions, and other disturbances were visible in both recordings despite their different frequency ranges. The occurrence of common events also provided the basis for alignment of the independently recorded time axes.

Previous studies have similarly shown that passive acoustic and vibration measurements can reflect changes occurring during powder blending [15, 58]. However, the physical interpretation of individual frequency components remains uncertain because the measured response can depend on particle behaviour, equipment structure, signal transmission, and sensor characteristics. The recurring spectral components observed in this work should therefore be regarded as repeatable signal characteristics rather than being assigned directly to a specific physical mechanism.

The multivariate analyses showed a strong influence of operating condition on the acoustic feature space. PLS-DA achieved 97.8% classification accuracy using the frequency band representation and 100% using the higher resolution spectral representation. Multi-output PLS regression also showed that RPM and throughput could be estimated from the acoustic measurements within the available dataset.

This dependence on operating condition is important for residence mass modelling. Previous continuous blending studies have shown that throughput and impeller speed influence blender fill behaviour and mass holdup [25, 11, 12]. Since these variables also affected the acoustic measurements in the present work, part of the acoustic variation associated with residence mass may also reflect the operating condition under which that mass occurred.

5.2 Residence Mass Prediction and Model Generalisation

Residence mass prediction was more challenging than prediction of the operating conditions. The initial high dimensional spectral representation produced a training R^2 close to 0.99 but a negative LORO R^2, showing that a strong training fit did not necessarily translate into prediction of an omitted run.

Reducing the spectral representation to broader frequency regions substantially improved LORO performance. Configurations containing fewer spectral regions repeatedly retained similar acoustic predictors, whereas increasingly fine spectral partitioning caused cross run performance to deteriorate. Within the available dataset, the reduced representation therefore provided a better balance between spectral information and model complexity.

The final acoustic-only model retained three predictors: the AE spectral slopes between 5 and 20 kHz and between 50 and 100 kHz, together with the mean FLAC spectral power between 15 and 22.05 kHz. Using 15 s blocks, random sampling, density and confidence weighting, and one PLS latent component, the model achieved a training R^2 of 0.940, a LORO R^2 of 0.854, and a LORO RMSE of 266 g.

The selected variables showed strong empirical relationships with residence mass, but the experiments do not establish their exact physical origin. Their main value in the present work is therefore as acoustic descriptors that remained comparatively stable when complete runs were omitted from model development.

5.3 Influence of Sampling, Weighting, and Process Features

The block sampling procedure had a strong effect on model performance. Residual based rejection produced a high training fit but poor LORO performance. By progressively removing blocks that disagreed with the provisional fitted relationship, the procedure may have reduced natural variation that was useful for prediction of an omitted condition.

Random sampling did not use agreement with the fitted model as a criterion for retaining observations and therefore preserved a broader representation of the variation present during valid blender operation. This produced better LORO performance despite a lower training fit. Density weighting addressed a separate issue by increasing the influence of the sparsely represented upper part of the residence mass range rather than allowing the more densely sampled lower range to dominate the regression.

The engineered process variables showed strong individual relationships with residence mass. The strongest, P_{f3}, had a correlation of $r = -0.934$. Nevertheless, adding the process features reduced prediction performance. The best combined acoustic and process model achieved a LORO R^2 of 0.760 and a LORO RMSE of 341 g, compared with 0.854 and 266 g for the acoustic-only model.

The combined model was also more sensitive to the particular blocks selected during stochastic sampling. These results show that neither training fit nor individual feature correlation was sufficient for model selection. Evaluation using complete omitted runs provided a more useful comparison of the investigated configurations.

5.4 Application to Additional Recordings

Application to the additional recordings introduced a change in the high-frequency acquisition chain. The blender and PCB microphone remained the same, while the acquisition system and sampling rate changed from the PicoScope system at 299,760 Hz to the NI cDAQ system at 1 MHz.

Resampling the additional high-frequency recordings to approximately the training sampling rate reduced the systematic prediction offset. This indicates that consistency of the signal representation was important for application of the spectral model. However, the observed difference cannot be attributed to sampling rate alone, since the acquisition hardware and other properties of the measurement chain may also have contributed.

Run 15 was the only additional recording with a gravimetrically measured residence mass. Its measured value of 873 g was therefore used to estimate an additive bias correction. The corrected results should consequently be interpreted as model application after a one point reference calibration rather than as independent validation of prediction accuracy.

The remaining recordings were evaluated from their prediction behaviour and block level stability. Differences in within run variability were observed, but the absence of independent gravimetric residence mass measurements prevents quantitative assessment of their prediction accuracy. The additional recordings therefore demonstrated application of the procedure to separately acquired data while also showing the importance of maintaining a consistent sensing and acquisition configuration.

5.5 Limitations

The main limitation of this study is the small number of independently measured residence mass conditions. Although many acoustic blocks were available, all blocks from one run shared the same gravimetrically measured residence mass. Model development and LORO evaluation were therefore based on six independent residence mass values. The measured range was also uneven, with only one run representing the upper part of the range.

Residence mass was measured only once at the end of each training run. It was therefore not possible to determine directly how residence mass changed during the recording or to verify short term changes in the model prediction against a continuous reference.

RPM, throughput, and residence mass also changed between the available experimental conditions. Since RPM and throughput were strongly represented in the acoustic signals, their individual influence could not be fully separated from the effect associated with residence mass.

A further limitation is the sensitivity of the stochastic block sampling procedure. The best-of-100 approach was useful for investigating the influence of block selection, but different selections produced noticeable differences in LORO performance. The highest result from this procedure should therefore not be treated as a deterministic estimate of performance on future data. Finally, the changed measurement chain used for the additional recordings required resampling and reference bias correction, showing that direct transfer between different measurement configurations cannot yet be assumed.

5.6 Future Work and Real-Time Implementation

Future experiments should first increase the number and distribution of independent residence mass measurements. More uniform coverage of the residence mass range, repeated operating conditions, and different combinations of RPM and throughput that produce similar residence masses would help distinguish residence mass effects from operating condition effects. Where possible, repeated gravimetric measurements during longer experiments would also provide a stronger reference for evaluating temporal prediction behaviour.

Further model development should use a fixed sensing and acquisition configuration, including sensor type, mounting arrangement, sampling rate, signal conditioning, and preprocessing. A fixed and reproducible block selection procedure should also replace the stochastic best-of-100 approach before independent validation.

Following this validation, the current offline procedure could be adapted for real time operation. Acoustic and process signals could be acquired using a common time reference, process validity conditions evaluated automatically, and the selected features calculated over consecutive or rolling blocks. Each valid block could then be supplied to the PLS model to provide an updated residence mass estimate.

For initial practical use, the prediction would be more appropriate as a monitoring soft sensor than as a direct input to automatic process control. The estimate could also be accompanied by information describing prediction stability and whether the current operating condition lies within the range represented during model development. After sufficient independent validation, such a system could contribute to broader real time process monitoring approaches in continuous pharmaceutical manufacturing, consistent with the process understanding and monitoring principles described in ICH Q13 and the FDA PAT framework [63, 62].

A validated residence mass estimate could complement existing feeder, blender, and analytical measurements by providing information about the amount of material retained in the blender without interrupting production for gravimetric measurement. The present study does not demonstrate improvements in product quality or patient outcomes, but it provides a basis for evaluating acoustic residence mass monitoring as a future real time PAT tool.

Bibliography

[1] C. L. Burcham, A. J. Florence, and M. D. Johnson, "Continuous Manufacturing in Pharmaceutical Process Development and Manufacturing," *Annual Review of Chemical and Biomolecular Engineering*, vol. 9, pp. 253-281, 2018. doi: 10.1146/annurev-chembioeng-060817-084355.

[2] S. L. Conway, J. G. Rosas, P. Overton, N. Tugby, P. Cryan, F. Witulski, S. Hurley, L. Wareham, A. Tantuccio, M. Ramasamy, A. Lalloo, M. Gibbs, and R. F. Meyer, "Implementation of a fully integrated continuous manufacturing line for direct compression and coating at a commercial pharmaceutical facility - Part 1: Operational considerations and control strategy," *International Journal of Pharmaceutics*, vol. 642, Art. no. 122820, 2023. doi: 10.1016/j.ijpharm.2023.122820.

[3] L. Pernenkil and C. L. Cooney, "A review on the continuous blending of powders," *Chemical Engineering Science*, vol. 61, no. 2, pp. 720-742, 2006. doi: 10.1016/j.ces.2005.06.016.

[4] B. Van Snick, J. Holman, V. Vanhoorne, A. Kumar, T. De Beer, J. P. Remon, and C. Vervaet, "Development of a continuous direct compression platform for low-dose drug products," *International Journal of Pharmaceutics*, vol. 529, no. 1-2, pp. 329-346, 2017. doi: 10.1016/j.ijpharm.2017.07.003.

[5] J. Palmer, G. Reynolds, F. Tahir, I. K. Yadav, E. Meehan, J. Holman, and G. Bajwa, "Mapping key process parameters to the performance of a continuous dry powder blender in a continuous direct compression system," *Powder Technology*, vol. 362, pp. 659-670, 2020. doi: 10.1016/j.powtec.2019.12.028.

[6] A. U. Vanarase and F. J. Muzzio, "Effect of operating conditions and design parameters in a continuous powder mixer," *Powder Technology*, vol. 208, no. 1, pp. 26-36, 2011. doi: 10.1016/j.powtec.2010.11.038.

[7] Y. Gao, A. Vanarase, F. Muzzio, and M. Ierapetritou, "Characterizing continuous powder mixing using residence time distribution," *Chemical Engineering Science*, vol. 66, no. 3, pp. 417-425, 2011. doi: 10.1016/j.ces.2010.10.045.

[8] J. G. Osorio and F. J. Muzzio, "Effects of processing parameters and blade patterns on continuous pharmaceutical powder mixing," *Chemical Engineering and Processing: Process Intensification*, vol. 109, pp. 59-67, 2016. doi: 10.1016/j.cep.2016.07.012.

[9] P. M. Portillo, A. U. Vanarase, A. Ingram, J. P. K. Seville, M. G. Ierapetritou, and F. J. Muzzio, "Investigation of the effect of impeller rotation rate, powder flow rate, and cohesion on powder flow behavior in a continuous blender using positron emission particle tracking," *Chemical Engineering Science*, vol. 65, pp. 5658-5668, 2010. doi: 10.1016/j.ces.2010.06.036.

[10] O. Jones-Salkey, A. L. Nicusan, C. R. K. Windows-Yule, A. Ingram, D. Werner, S. Clifford, and G. K. Reynolds, "Application of Positron Emission Particle Tracking (PEPT) for the evaluation of powder behaviour in an incline linear blender for Continuous Direct Compression (CDC)," *International Journal of Pharmaceutics*, vol. 645, Art. no. 123361, 2023. doi: 10.1016/j.ijpharm.2023.123361.

[11] O. Jones-Salkey, C. R. K. Windows-Yule, A. Ingram, L. Stahler, A. L. Nicusan, S. Clifford, L. Martin de Juan, and G. K. Reynolds, "Using AI/ML to predict blending performance and process sensitivity for Continuous Direct Compression (CDC)," *International Journal of Pharmaceutics*, vol. 651, Art. no. 123796, 2024. doi: 10.1016/j.ijpharm.2024.123796.

[12] H. G. Jolliffe, M. A. Velazco-Roa, L. Martin de Juan, M. Prostredny, C. Mendez Torrecillas, G. Reynolds, D. McElhone, and J. Robertson, "Characterisation of a continuous blender: impact of physical properties on mass holdup behaviour," *Powder Technology*, vol. 449, Art. no. 120440, 2025. doi: 10.1016/j.powtec.2024.120440.

[13] A.-P. Karttunen, J. Poms, S. Sacher, A. Sparén, C. Ruiz Samblás, M. Fransson, L. Martin de Juan, J. Remmelgas, H. Wikström, W. K. Hsiao, S. Folestad, O. Korhonen, S. Abrahmsén-Alami, and P. Tajarobi, "Robustness of a continuous direct compression line against disturbances in feeding," *International Journal of Pharmaceutics*, vol. 574, Art. no. 118882, 2020. doi: 10.1016/j.ijpharm.2019.118882.

[14] J. W. R. Boyd and J. Varley, "The uses of passive measurement of acoustic emissions from chemical engineering processes," *Chemical Engineering Science*, vol. 56, pp. 1749-1767, 2001. doi: 10.1016/S0009-2509(00)00540-6.

[15] P. Allan, L. J. Bellamy, A. Nordon, and D. Littlejohn, "Non-invasive monitoring of the mixing of pharmaceutical powders by broadband acoustic emission," *Analyst*, vol. 135, no. 3, pp. 518-524, 2010. doi: 10.1039/B922446G.

[16] A. Crouter and L. Briens, "Passive acoustic emissions from particulates in a V-blender," *Drug Development and Industrial Pharmacy*, vol. 41, no. 11, pp. 1809-1818, 2015. doi: 10.3109/03639045.2015.1009913.

[17] G. Carson, A. J. Mulholland, A. Nordon, M. Tramontana, A. Gachagan, and G. Hayward, "Estimating particle concentration using passive ultrasonic measurement of impact vibrations," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 2, pp. 345-352, 2009. doi: 10.1109/TUFFC.2009.1043.

[18] C. Pecorari, "Characterizing Particle Flow by Acoustic Emission," *Journal of Nondestructive Evaluation*, vol. 32, pp. 104-111, 2013. doi: 10.1007/s10921-012-0163-7.

[19] E. M. Hansuld, L. Briens, A. Sayani, and J. A. B. McCann, "The effect of process parameters on audible acoustic emissions from high-shear granulation," *Drug Development and Industrial Pharmacy*, vol. 39, no. 2, pp. 331-341, 2013. doi: 10.3109/03639045.2012.681055.

[20] J. G. Whiting, A. Springer, and F. Sciammarella, "Real-time acoustic emission monitoring of powder mass flow rate for directed energy deposition," *Additive Manufacturing*, vol. 23, 2018. doi: 10.1016/j.addma.2018.08.015.

[21] G. Zheng, Y. Yan, Y. Hu, W. Zhang, L. Yang, and L. Li, “Mass-Flow-Rate Measurement of Pneumatically Conveyed Particles Through Acoustic Emission Detection and Electrostatic Sensing,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-13, 2021. doi: 10.1109/TIM.2020.3039619.

[22] R. Xu, R. E. Beltrán-Gutiérrez, M. Käding, A. Lange, S. Marx, and J. Ostermann, “Frequency dependent amplitude response of different couplant materials for mounting piezoelectric sensors,” *NDT & E International*, vol. 141, Art. no. 102993, 2024. doi: 10.1016/j.ndteint.2023.102993.

[23] M. A. Järvinen, J. Paaso, M. Paavola, K. Leiviskä, M. Juuti, F. Muzzio, and K. Järvinen, “Continuous direct tablet compression: Effects of impeller rotation rate, total feed rate and drug content on the tablet properties and drug release,” *Drug Development and Industrial Pharmacy*, vol. 39, no. 11, pp. 1802-1808, 2013. doi: 10.3109/03639045.2012.738681.

[24] B. Van Snick, A. Kumar, M. Verstraeten, K. Pandelaere, J. Dhondt, G. Di Pretoro, T. De Beer, C. Vervaet, and V. Vanhoorne, “Impact of material properties and process variables on the residence time distribution in twin screw feeding equipment,” *International Journal of Pharmaceutics*, vol. 556, pp. 200-216, 2019. doi: 10.1016/j.ijpharm.2018.11.076.

[25] B. Bekaert, W. Grymonpré, A. Novikova, C. Vervaet, and V. Vanhoorne, “Impact of blend properties and process variables on the blending performance,” *International Journal of Pharmaceutics*, vol. 613, Art. no. 121421, 2022. doi: 10.1016/j.ijpharm.2021.121421.

[26] C. Zheng, L. Li, B. J. Nitert, N. Govender, T. Chamberlain, L. Zhang, and C.-Y. Wu, “Investigation of granular dynamics in a continuous blender using the GPU-enhanced discrete element method,” *Powder Technology*, vol. 412, Art. no. 117968, 2022. doi: 10.1016/j.powtec.2022.117968.

[27] V. Bachelet, A. Mangeney, R. Toussaint, J. de Rosny, M. I. Arran, M. Farin, and C. Hibert, “Acoustic Emissions of Nearly Steady and Uniform Granular Flows: A Proxy for Flow Dynamics and Velocity Fluctuations,” *Journal of Geophysical Research: Earth Surface*, vol. 128, Art. no. e2022JF006990, 2023. doi: 10.1029/2022JF006990.

[28] C. Zhai, E. B. Herbold, and R. C. Hurley, “The influence of packing structure and interparticle forces on ultrasound transmission in granular media,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 28, pp. 16234-16242, 2020. doi: 10.1073/pnas.2004356117.

[29] N. J. Watson, M. J. W. Povey, G. K. Reynolds, B. H. Xu, and Y. Ding, “Acoustic emission monitoring from a lab scale high shear granulator: A novel approach,” *International Journal of Pharmaceutics*, vol. 465, no. 1-2, pp. 262-274, 2014. doi: 10.1016/j.ijpharm.2014.01.032.

[30] R. Fulek, S. Ramm, C. Kiera, M. Pein-Hackelbusch, and U. Odefey, “A Machine Learning Approach to Qualitatively Evaluate Different Granulation Phases by Acoustic Emissions,” *Pharmaceutics*, vol. 15, no. 8, Art. no. 2153, 2023. doi: 10.3390/pharmaceutics15082153.

[31] C. Leroy, S. Mittelette, G. Félix, N. Fabregue, J. Špačková, P. Gaveau, T.-X. Métro, and D. Laurencin, “Operando acoustic analysis: a valuable method

for investigating reaction mechanisms in mechanochemistry," *Chemical Science*, vol. 13, no. 21, pp. 6328-6334, 2022. doi: 10.1039/D2SC01496C.

[32] L. Briens, D. Daniher, and A. Tallevi, "Monitoring high-shear granulation using sound and vibration measurements," *International Journal of Pharmaceutics*, vol. 331, no. 1, pp. 54-60, 2007. doi: 10.1016/j.ijpharm.2006.09.012.

[33] E. M. Hansuld, L. Briens, J. A. B. McCann, and A. Sayani, "Audible acoustics in high-shear wet granulation: Application of frequency filtering," *International Journal of Pharmaceutics*, vol. 378, no. 1-2, pp. 37-44, 2009. doi: 10.1016/j.ijpharm.2009.05.042.

[34] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70-73, 1967. doi: 10.1109/TAU.1967.1161901.

[35] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1558-1564, 1977.

[36] S. Mendoza-Silva, F. Alijani, L. V. Naarden, R. Broer, L. Smeets, T. Riepe, I. Roslon, and A. Japaridze, "Single-Cell Nanomotion and Machine Learning for Parallel Bacterial Identification and Antibiotic Screening," *ACS Sensors*, vol. 11, no. 3, pp. 2668-2677, 2026. doi: 10.1021/acssensors.5c04649.

[37] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417-441, 1933.

[38] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579-2605, 2008.

[39] S. Wold, A. Ruhe, H. Wold, and W. J. Dunn, "The Collinearity Problem in Linear Regression: The Partial Least Squares (PLS) Approach to Generalized Inverses," *SIAM Journal on Scientific and Statistical Computing*, vol. 5, no. 3, pp. 735-743, 1984. doi: 10.1137/0905052.

[40] Y. Gao, F. Boukouvala, W. Engisch, W. Meng, F. J. Muzzio, and M. G. Ierapetritou, "Improving Continuous Powder Blending Performance Using Projection to Latent Structures Regression," *Journal of Pharmaceutical Innovation*, vol. 8, pp. 99-110, 2013. doi: 10.1007/s12247-013-9152-3.

[41] M. Barker and W. Rayens, "Partial least squares for discrimination," *Journal of Chemometrics*, vol. 17, no. 3, pp. 166-173, 2003. doi: 10.1002/cem.785.

[42] Chalmers University of Technology, "Regulations for the use of AI tools in thesis work," Dec. 2023. [Online]. Available: `https://www.chalmers.se/en/education/your-studies/masters-and-bachelors-thesis/regulations-for-the-use-of-ai-tools/`

[43] OpenAI, "ChatGPT," generative AI language model used for language assistance and idea refinement, 2025. [Online]. Available: `https://chat.openai.com`

[44] I.-G. Chong and C.-H. Jun, "Performance of some variable selection methods when multicollinearity is present," *Chemometrics and Intelligent Laboratory Systems*, vol. 78, no. 1-2, pp. 103-112, 2005. doi: 10.1016/j.chemolab.2004.12.011.

[45] B. Mahieu, E. M. Qannari, and B. Jaillais, "Extension and significance testing of Variable Importance in Projection (VIP) indices in Partial Least Squares regression and Principal Components Analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 242, Art. no. 104986, 2023. doi: 10.1016/j.chemolab.2023.104986.

[46] D. R. Roberts, V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, F. Hartig, and C. F. Dormann, "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure," *Ecography*, vol. 40, no. 8, pp. 913-929, 2017. doi: 10.1111/ecog.02881.

[47] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BMC Bioinformatics*, vol. 7, Art. no. 91, 2006. doi: 10.1186/1471-2105-7-91.

[48] G. C. Cawley and N. L. C. Talbot, "On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation," *Journal of Machine Learning Research*, vol. 11, pp. 2079-2107, 2010.

[49] D. Krstajic, L. J. Buturovic, D. E. Leahy, and S. Thomas, "Cross-validation pitfalls when selecting and assessing regression and classification models," *Journal of Cheminformatics*, vol. 6, Art. no. 10, 2014. doi: 10.1186/1758-2946-6-10.

[50] A. K. Beke, M. Gyürkés, Z. K. Nagy, G. Marosi, and A. Farkas, "Digital Twin of Low Dosage Continuous Powder Blending: Artificial Neural Networks and Residence Time Distribution Models," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 169, pp. 64–77, 2021. doi: 10.1016/j.ejpb.2021.09.006.

[51] A. Sow, I. Traore, T. Diallo, M. Traore, and A. Ba, "Comparison of Gaussian Process Regression, Partial Least Squares, Random Forest and Support Vector Machines for a Near Infrared Calibration of Paracetamol Samples," *Results in Chemistry*, vol. 4, Art. no. 100508, 2022. doi: 10.1016/j.rechem.2022.100508.

[52] G. H. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, vol. 2, pp. 205–224, 1965. doi: 10.1137/0702016.

[53] PCB Piezotronics, "Model 378C01: 1/4-inch free-field prepolarized microphone and preamplifier system," PCB Piezotronics, product documentation.

[54] RØDE Microphones, "PinMic Discreet Pin-through Lavalier Microphone," product documentation.

[55] National Instruments, "NI-9223 C Series Voltage Input Module Specifications," National Instruments, product documentation.

[56] National Instruments, "cDAQ-9170: 1-Slot, USB Type-C CompactDAQ Chassis," NI, product documentation. Available: `https://www.ni.com/en/shop/hardware/compactdaq-chassis/model-cdaq-9170`

[57] Roland Corporation, *QUAD-CAPTURE USB 2.0 Audio Capture Owner's Manual*, 2011. Available: `https://cdn.roland.com/assets/media/pdf/QUAD-CAPTURE_e05_W.pdf`

[58] K. Wilson and L. Briens, "Investigation of passive acoustic emissions during powder mixing in a V-blender," *Powder Technology*, vol. 408, Art. no. 117754, 2022. doi: 10.1016/j.powtec.2022.117754.

[59] B. Nagy, A. Farkas, M. Gyürkés, S. Komaromy-Hiller, B. Démuth, B. Szabó, D. Nusser, E. Borbás, G. Marosi, and Z. K. Nagy, "In-line Raman spectroscopic monitoring and feedback control of a continuous twin-screw pharmaceutical powder blending and tableting process," *International Journal of Pharmaceutics*, vol. 530, no. 1–2, pp. 21–29, 2017. doi: 10.1016/j.ijpharm.2017.07.041.

[60] D. L. Galata, L. A. Mészáros, M. Ficzere, P. Vass, B. Nagy, E. Szabó, A. Domokos, A. Farkas, I. Csontos, G. Marosi, and Z. K. Nagy, "Continuous blending monitored and feedback controlled by machine vision-based PAT tool," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 196, Art. no. 113902, 2021. doi: 10.1016/j.jpba.2021.113902.

[61] Y.-S. Huang, S. Medina-González, B. J. Straiton, J. Keller, Q. Marashdeh, M. Gonzalez, Z. K. Nagy, and G. V. Reklaitis, "Real-Time Monitoring of Powder Mass Flowrates for Plant-Wide Control of a Continuous Direct Compaction Tablet Manufacturing Process," *Journal of Pharmaceutical Sciences*, vol. 111, no. 1, pp. 69–81, 2022. doi: 10.1016/j.xphs.2021.06.005.

[62] U.S. Food and Drug Administration, "PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance," *Guidance for Industry*, 2004.

[63] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, "Q13: Continuous Manufacturing of Drug Substances and Drug Products," *ICH Harmonised Guideline*, Step 4, 16 November 2022.

[64] P. M. Portillo, M. G. Ierapetritou, and F. J. Muzzio, "Effects of rotation rate, mixing angle, and cohesion in two continuous powder mixers—A statistical approach," *Powder Technology*, vol. 194, no. 3, pp. 217–227, 2009. doi: 10.1016/j.powtec.2009.04.010.

[65] T. Petrut, T. Geay, C. Gervaise, P. Belleudy, and S. Zanker, "Passive acoustic measurement of bedload grain size distribution using self-generated noise," *Hydrology and Earth System Sciences*, vol. 22, pp. 767–787, 2018. doi: 10.5194/hess-22-767-2018.

A

Appendix 1 - Experimental Instrumentation and Data Preparation

This appendix provides additional information on the acoustic measurement systems and data preparation procedures used in this study. The main Methods chapter describes the measurement configurations and processing steps required for the subsequent analysis, while further technical specifications and details of the file conversion procedures are provided here.

A.1 Acoustic Measurement Systems

The same high-frequency PCB microphone was used for both the training experiment and the additional recordings, while the acquisition systems and sampling rates differed. During training, the high-frequency microphone was recorded using the PicoScope system at 299,760 samples/s. For the additional recordings, the same microphone signal was acquired using an NI cDAQ system at 1,000,000 samples/s. The RØDE PinMic was used for the audible-frequency measurement in both recording sets. The main characteristics of the sensors and acquisition systems are summarised in Table A.1.

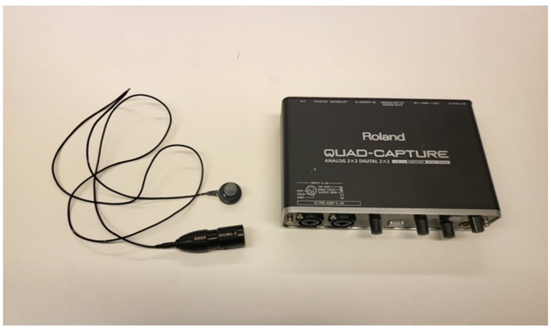
A.1.1 Training Measurement System

The high-frequency airborne signal used during model development was recorded using a PCB Piezotronics Model 378C01 measurement system. The system consisted of a Model 377C01 1/4-inch free-field prepolarized microphone and a Model 426B03 preamplifier with TEDS. The complete 378C01 system has a manufacturer-specified nominal sensitivity of 2.0 mV/Pa and a free-field frequency range of 4 Hz to 100 kHz at $+2/-3$ dB [53]. The corresponding ± 2 dB frequency range is 5 Hz to 80 kHz. The signal was sampled at 299,760 samples/s using PicoScope acquisition software.
The audible-frequency signal was recorded using a RØDE PinMic miniature microphone. The PinMic is a permanently polarised omnidirectional microphone with a manufacturer-specified frequency range of 60 Hz to 18 kHz and a nominal sensitivity of 21 mV/Pa [54]. The signal was acquired at 44,100 samples/s through a Roland QUAD-CAPTURE audio interface and stored in FLAC format.
Both microphones were positioned below the central section of the blender, as shown

(a)

(b)

(c)



(d)

Figure A.1: Instrumentation used for the training acoustic measurements: (a) PCB Piezotronics 378C01 microphone system : (b) 1/4" ICP PREAMPLIFIER Model 426B03 and (c) 1/4" FREE FIELD PREPOLARIZED MICROPHONE Model 377C01 (d) RØDE PinMic microphone, and Roland QUAD-CAPTURE audio interface. [53, 54, 57].

Measurement	Sensor and acquisition system	Sampling rate	Relevant specification
Training high frequency	PCB Piezotronics 378C01, consisting of 377C01 microphone and 426B03 preamplifier	299,760 samples/s	1/4-inch free-field, prepolarized microphone system; 4 Hz to 100 kHz frequency range ($+2/-3$ dB); nominal sensitivity 2.0 mV/Pa; TEDS compliant [53]
Training and additional audible frequency	RØDE PinMic with Roland QUAD-CAPTURE	44,100 samples/s	Omnidirectional microphone; 60 Hz to 18 kHz frequency range; nominal sensitivity 21 mV/Pa [54]
Additional high frequency	Same PCB Piezotronics 378C01 system, NI 9223 in cDAQ-9170 chassis	1,000,000 samples/s	16-bit, ±10 V, four-channel simultaneous analogue input module [55]

Table A.1: Summary of the acoustic sensors and acquisition systems used in the study.

in Figure 3.1 in the Methods chapter. The two measurements therefore represented acoustic activity from approximately the same region, although their different bandwidths and measurement systems resulted in different representations of the recorded acoustic activity.

A.1.2 Measurement System for the Additional Recordings

The same PCB Piezotronics high-frequency microphone used during model development was also used for the additional recordings. The microphone was a Model 377C01 1/4-inch free-field prepolarized microphone, forming part of the Model 378C01 measurement system described in Section 3.1.1.

The main difference in the additional recordings was the acquisition system. The high-frequency microphone signal was digitised using an NI 9223 analogue input module installed in a National Instruments cDAQ-9170 chassis. The NI 9223 provides four simultaneously sampled differential analogue inputs, 16-bit resolution, an input range of ±10 V, and a maximum sampling rate of 1 MS/s per channel [55]. The signal used in this work was acquired at 1,000,000 samples/s using NI FlexLogger and stored in TDMS format.

The RØDE PinMic was again used for the audible-frequency measurement and was recorded at 44,100 samples/s through the Roland QUAD-CAPTURE interface.

The RØDE PinMic was again used for the airborne audible-frequency measurement and was recorded at 44,100 samples/s through the Roland QUAD-CAPTURE interface.

Figure A.2: Instrumentation used for acquisition of the additional high-frequency recordings: NI cDAQ-9170 chassis. [56].

A.2 Data Formats and Conversion

The different acquisition systems produced recordings in different file formats. The high-frequency training recordings were exported from PicoScope as segmented CSV files, the additional high-frequency recordings were stored in TDMS format, and the RØDE PinMic recordings were stored as FLAC files.
The conversion procedures described below were used to provide efficient access to the long recordings during repeated signal processing. The sampling rates of the high-frequency signals were retained during this initial conversion. Resampling of the additional recordings was performed later during model application, as described in Section 3.12.

A.2.1 PicoScope CSV Reconstruction

The PicoScope recordings from the training experiment were exported as multiple CSV files for each run. Each file contained a local time axis and the corresponding measured voltage. Since the time coordinate restarted for each segment, the individual files were reconstructed into a single analysis record before subsequent signal processing.
Each CSV segment was read in its original order and the local time coordinate was shifted relative to the beginning of that segment. The segments were then arranged sequentially using their sample interval and number of samples. A separation corresponding to one sampling interval was used between consecutive segments, and the corresponding voltage samples were concatenated in the same order.
The reconstructed recording was stored as interleaved time and voltage values in 64-bit floating-point format,

$$[t_0, v_0, t_1, v_1, \dots, t_N, v_N]. \tag{A.1}$$

One binary file was produced for each run at the original sampling rate of 299,760 samples/s. The binary representation reduced the computational overhead associated with repeatedly parsing the segmented text files during signal analysis.

A.2.2 TDMS Conversion

The high-frequency additional recordings were originally stored in National Instruments TDMS format. In addition to the recorded voltage samples, the files contained waveform metadata required to reconstruct the signal, including the sampling interval and timing information.
For subsequent processing, the voltage samples were converted to a compact signed 16-bit integer representation. The minimum and maximum voltage values of the recording were used to define a linear mapping across the integer range from $-32{,}767$ to $+32{,}767$. The voltage represented by one integer increment was

$$\Delta V = \frac{V_{\max} - V_{\min}}{65{,}534}, \tag{A.2}$$

where $V_{\min}$ and $V_{\max}$ are the minimum and maximum voltage values used during conversion.
The use of a finite 16-bit integer range introduced quantisation and the conversion was therefore not mathematically lossless. The voltage resolution of the reconstructed signal was determined by Equation A.2.
Each converted file began with a 40-byte header consisting of five 64-bit floating-point values. The fields stored in the header are shown in Table A.2.

Field	**Description**
t_0	Start time of the stored signal segment (s)
Δt	Sampling interval (s)
N	Number of samples
$V_{\min}$	Minimum voltage used during scaling (V)
scale	Voltage represented by one integer increment (V/-count)

Table A.2: Header structure of the converted high-frequency recordings.

Following the header, the quantised voltage values were stored as contiguous signed 16-bit integer samples. The stored header parameters allowed the signal amplitude and time axis to be reconstructed when the binary file was loaded.
The converted recordings retained their original sampling rate of 1,000,000 samples/s. Resampling to approximately 300 kHz was performed only when the trained model was subsequently applied to the additional recordings.

A.2.3 FLAC Data Preparation

The RØDE PinMic recordings from both sets of measurements were stored in FLAC format at 44,100 samples/s. Since FLAC uses lossless audio compression, the recordings were loaded directly for subsequent signal processing without conversion to an intermediate storage format.
Where recordings contained two audio channels, the channels were combined into a single mono waveform before analysis. The resulting signal was retained at its original sampling rate of 44,100 samples/s.

A.2.4 Summary of Data Formats

Table A.3 summarises the data formats used during the subsequent analysis.

Signal	**Recording set**	**Original format**	**Analysis format**	**Sampling rate**
High-frequency airborne signal	Training	Segmented PicoScope CSV	Interleaved float64 time and voltage	299,760 Hz
High-frequency airborne signal - Additional	Additional	TDMS	40-byte header and int16 samples	1,000,000 Hz
RØDE PinMic	Both	FLAC	FLAC loaded directly	44,100 Hz

Table A.3: Signal formats used after data preparation.

A.3 Storage Requirements

The storage requirements of the high-frequency recordings were substantial because of their sampling rates and recording durations. The values presented below describe the converted analysis formats rather than the sizes of the original CSV and TDMS source files.

For the reconstructed PicoScope signal, every sample was represented by one 64-bit time value and one 64-bit voltage value. The required data size was therefore

$$S_{\text{Pico}} = 16N, \tag{A.3}$$

where S_{Pico} is given in bytes and N is the number of samples. At 299,760 samples/s, this corresponds to approximately 4.80 MB/s, or approximately 288 MB for each minute of recording.

For the converted TDMS recordings, every voltage sample was represented by a 16-bit integer in addition to the fixed 40-byte header. The corresponding file size was approximately

$$S_{\text{TDMS}} = 40 + 2N, \tag{A.4}$$

in bytes. At 1,000,000 samples/s, the signal data therefore required approximately 2.0 MB/s, corresponding to approximately 120 MB for each minute of recording.

The approximate storage requirements of the converted formats are summarised in Table A.4.

Based on the recording durations in the training dataset, the reconstructed PicoScope binary representation corresponds to approximately 1.44 GB for a 5-minute recording, 2.01 GB for a 7-minute recording, and 3.60 GB for a 12.5-minute recording. These values are calculated from the stored numerical representation and should not be interpreted as measured sizes of the original CSV files.

Recording	Stored representation	Approx. data rate	Approx. size/min
PicoScope training signal	Two float64 values per sample	4.80 MB/s	288 MB
Additional high-frequency signal	40-byte header and int16 samples	2.00 MB/s	120 MB
RØDE PinMic	FLAC compression	Variable	Variable

Table A.4: Approximate storage requirements of the signal formats used for analysis.

DEPARTMENT OF Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY