White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/228181/

Version: Published Version

Article:
Deebes, M., Mahfouf, M., Omar, C. et al. (2025) A plant wide modelling framework for the multistage processes of the continuous manufacturing of pharmaceutical tablets. Journal of Pharmaceutical Innovation, 20 (4). 115. ISSN: 1872-5120

https://doi.org/10.1007/s12247-025-10017-4

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Journal of Pharmaceutical Innovation (2025) 20:115
https://doi.org/10.1007/s12247-025-10017-4

RESEARCH

A Plant Wide Modelling Framework For The Multistage Processes of The Continuous Manufacturing of Pharmaceutical Tablets

Motaz Deebes[1] · Mahdi Mahfouf[1] · Chalak Omar[2] · Syed Islam[3] · Ben Morgan[3]

Accepted: 3 June 2025
© The Author(s) 2025

Abstract

Continuous manufacturing can be seen as a promising shift in the pharmaceutical industry, offering benefits such as reduced costs and improved product quality. However, the multistage nature of continuous tablet manufacturing demands a deeper understanding of the complex interactions between process parameters, material attributes, and final product quality. This study aims to address this challenge by developing a novel, data-driven modelling framework to predict key critical quality attributes, including particle size distribution, moisture content, and tablet tensile strength across the processing stages of a pilot-scale continuous tablet manufacturing line. A sequential modelling approach was employed, integrating Random Forest and Gradient Boosting Machines to model each processing stage. These models were sequentially trained and interlinked to holistically capture process–material interactions across granulation, drying, milling, and tabletting stages. To manage error propagation between stages, Gaussian Mixture Models were incorporated for error characterisation and uncertainty reduction. The results showed that the proposed framework captured the non-linear interactions between processing parameters and the quality attributes. The incorporation of GMMs was influential in quantifying uncertainty within each process model, resulting in a final estimation of tablet tensile strength with an R^2 value of 0.90 using the integrated Random Forest model. This framework demonstrated considerable improvement in the predictive performance of the continuous manufacturing processes modelling through the integration of machine learning models and an uncertainty-aware strategy. The predictive tool is intended to support the Quality by Design (QbD) concept through systematic design space exploration and process understanding of the pharmaceutical continuous manufacturing.

Keywords Continuous manufacturing · Pharmaceutical manufacturing · Pilot plant scale · Predictive modelling · Machine learning · Gaussian mixture models · Quality by design (QbD)

Introduction

Continuous manufacturing is recognised as a transformative approach in pharmaceutical industry due to its advantages in operational cost reduction,improved efficiency, and consistent product quality [1, 2]. Traditionally, the processes of developing oral dosages such as tablets operate in a batch-oriented manner, where each unit operation such as mixing, granulation, drying, and tabletting is performed separately with manual intervention. In contrast, continuous manufacturing integrates these operations into an interconnected chain of processes in which materials are autonomously processed without intervention for feeding or discharge. In recent years,the continuous manufacturing for drug products has progressed considerably with growing regulatory support and and industrial commitment facilitated by frameworks such as Quality by Design (QbD) and ICH Q13 guidelines

✉ Motaz Deebes
mmdeebes1@sheffield.ac.uk

Mahdi Mahfouf
m.mahfouf@sheffield.ac.uk

Chalak Omar
c.omar@sheffield.ac.uk

Syed Islam
s.f.islam@sheffield.ac.uk

Ben Morgan
b.morgan@amrc.co.uk

[1] School of Electrical and Electronic Engineering, University of Sheffield, Sheffield, UK

[2] Department of Multidisciplinary Engineering Education, University of Sheffield, Sheffield, UK

[3] Advanced Manufacturing Research Centre, University of Sheffield, Sheffield, UK

Published online: 17 June 2025

which offer both scientific foundation and implementation guidance [3–5]. A number of commercial drug products manufactured using continuous processes have since received regulatory approval from agencies such as the Food and Drug Administration (FDA) and European Medicines Agency (EMA) [6–8]. While continuos manufacturing offers tangible benefits, it is still considered as complex and not fully understood [5, 9, 10]. These complexities lie in the inherent multivariate nature of pharmaceutical manufacturing, which often involves the handling of particulate materials throughout multiple processing units resulting in complex interactions between these particulates and unit operation process parameters [10, 11]. In addition, the multistage structure of such a manufacturing process is still lacking an efficient process control strategy. For example, the variations resulted by a particular processing stage may be caused by multiple sources of variations introduced by both the current and preceding processing units resulting in a challenging control action and long-time delays [1, 10, 12]. Addressing these complexities is particularly crucial in order to reduce variations in tablet manufacturing and ensuring consistent quality in the final product. Predictive process models play a significant role in addressing the challenges of continuous manufacturing of pharmaceutical tablets [5, 9, 13, 14]. They can enhance process understanding and control by predicting the processes outcomes based on their operational data.

There are several modelling approaches that deemed to be potential tools for revealing and understanding the complexity associated with continuous pharmaceutical tablet manufacturing such as mechanistic models, data-driven models, and hybrid models [5, 9, 14]. However, the behaviour of the particulate processes during continuous manufacturing involves complex multi-scale phenomena spanning from particles interactions to equipment-scale dynamics, creating challenges for comprehensive process modelling [11]. Flowsheet modelling represents a mechanistic approach where continuous manufacturing processes can be integrated and simulated using network of mathematical models [15–18]. While technically rigorous and providing valuable insights, flowsheet models often rely on constitutive equations derived from physical and chemical principles which are not always available. For instance, the description of the particle size distribution using Population Balance Models (PBM) requires detailed specification of subprocess parameters such as nucleation, growth, breakage, and aggregation in granulation processes which often need to be empirically estimated [16, 17]. Consequently, accurate estimation of these rate processes is critical for maintaining flowsheet model performance; otherwise, predictive capability may degrades due to the parameter uncertainty [9, 16]. This challenge is further compounded in plant-wide manufacturing scenarios. As the number of unit operations parameters increases, so too does the complexity of parametrisation, making plant-wide flowsheet models computationally intensive and requiring extensive calibration effort [9, 16, 18].

Data driven models, however, can capture complex interdependencies across particulate pharmaceutical processes without requiring mechanistic details [14]. These models directly utilise real process data generated either from actual operational plants or structured design of experiments to learn complex, and nonlinear relationships between process variables and quality attributes. Furthermore, they have emerged as valuable tools in pharmaceutical manufacturing, complementing the Quality by Design (QbD) paradigm emphasised by regulatory agencies [3, 5, 11]. Data driven models offer potential advantages including computational efficiency, adaptability, and reduced model maintenance requirements compared to mechanistic models [14, 19]. In addition, these models are well recognised for pharmaceutical processes optimisation and controls [20, 21]. These models, while compelling for pharmaceutical applications, face deployment challenges that impact their predictive performance. These challenges include acquiring comprehensive and representative training datasets, selecting appropriate model algorithm, and ensuring generalisability beyond training conditions [14, 19, 20]. The complexity of modelling of pharmaceutical tablet processes has thus shifted from constitutive models parameter determination to data mining and statistical validation areas.

Within this context, a solution to model the processes of continuous manufacturing of pharmaceutical tablets can be achieved with the help of machine learning techniques. Machine learning algorithms have evolved to autonomously identify complex patterns and draw inferences from data [22]. The application of machine learning in pharmaceutical tablet manufacturing has been explored in various studies [23–26]. These studies contribute with a unique insight into the complexities associated with different manufacturing processes of oral solid dosage forms. For example, the authors in [24] evaluated the capability of machine learning models such as Ridge regression and Random Forest to reveal critical quality attributes of a commercial pharmaceutical tablet product based on historical process data. Furthermore, Bekaert et al. [25] investigated the interactions between various pharmaceutical material characteristics and operational variables in a continuous direct compression line, focusing on their effects on tablet physical properties. They employed Partial Least Squares (PLS) regression to develop a predictive model that could identify optimal process configurations based on blend characteristics. A recent study made by Al Alaween et al. for the development of serial of artificial neural networks (ANNs) for modelling the continuous tabletting line. Their study focused on predicting tablet tensile strength based on observational data from the milling machine and tablet press within the manufacturing line [27].

While these studies demonstrate significant progress in utilising machine learning for the prediction and optimisation of tablet manufacturing processes, they also highlight the multifaceted challenges inherent in the field. In fact, the multivariate nature of the continuous manufacturing of pharmaceutical tablets leads to complex patterns that may be challenging to predict [11, 13]. Additionally, since the materials are continuously processed through multiple units, modelling each stage separately may be insufficient to understand the interactions associated with each process stage. For instance, some studies relied on batch processes, where the outputs do not adequately reflect the patterns of continuous manufacturing [24]. Although other studies have explored the use of continuous manufacturing lines for modelling purposes, their models often fell short in capturing the complexity of such systems. For example, the PLS algorithm used in [25] assumes linearity across all processes involved, which is not reflective of the inherently complex nature of continuous manufacturing processes. Additionally, the development of ANN-based models proposed in [27] were constrained by extremely small dataset, potentially limiting the model reliability due to overfitting and lack of generalisability. However, none of these studies have addressed the fully interconnected process stages across the manufacturing line, where interactions between process variables and material characteristics can be traced within fully interconnected data driven modelling paradigm. Additionally, the availability of diverse and comprehensive process datasets, which could be systematically utilised to develop a holistic data-driven model, has not been explored.

This research study introduces a novel plant-wide modelling framework for the continuous manufacturing of pharmaceutical tablets, based on pilot plant data, aiming to enhance process understanding and control of the interconnected multistage processes. In an earlier study, a multistage modelling framework using machine learning techniques was developed to predict granule moisture content in a continuous pharmaceutical tablet manufacturing line [23]. The study focused on integrating multiple unit operations, including twin screw granulation and fluidised bed drying, to address the complex interactions within these processes that cause variability in critical quality attributes, such as granule moisture content, which eventually affects the tablet's physical profile. In this study, machine learning models were developed for each critical unit operation within the continuous tabletting pilot plant line. These models were then sequentially integrated, with the output of each model serving as the input for the subsequent stage, culminating in the prediction of the final process outcome. The inclusion of all processing stages within this sequential modelling framework not only captured the individual complexities between the process stages but also facilitated the propagation of their effects throughout the entire continuous manufacturing line. This approach was achieved by utilising machine learning models such as Random Forest (RF) and Gradient Boosting Machines (GBM) due to their ability to capture complex, and non-linear process patterns while remaining computationally efficient [28]. Additionally,through this sequential design, the ML models were further integrated with Gaussian Mixture Models (GMMs) in order to to characterise their predictions uncertainties caused by processes variability and deviations, ultimately enabling uncertainty-aware model refinement. Moreover,the modelling development strategy is supported by efficient experimental work that strategically involves manipulating critical process parameters to produce diverse data observations. These observations, related to critical material attributes, were collected, analysed, and utilised from the pilot plant to ensure consistent modelling based on real data.

The scientific contribution of this research lies in the practical implementation of a plant-wide modelling framework designed to holistically represent the continuous manufacturing of pharmaceutical tablets. Consequently, the resulting predictive platform facilitates the estimation of intermediate critical material attributes (CMAs) across key processing stages including granulation, drying, and milling, as well as final critical quality attributes (CQA) at the tablet press stage, based on the critical process parameters (CPPs) associated with the continuous tabletting line. Thus, the primary objective of this modelling framework is to develop a data-driven tool that enables systematic exploration and understanding of the operational design space of the continuous manufacturing line, without necessitating physical experimentation. It achieves this by predicting the impact of the multivariate process interactions on the tablet's physical profile under an uncertainty-aware strategy. Additionally, this framework is aligned with regulatory compliance under the Quality by Design (QbD) paradigm, facilitating comprehensive identification and control of critical sources of variability. Within this framework, the proposed modelling approach is intended to contribute to the broader scientific efforts supporting and complementing ongoing initiatives to advance continuous processing technologies in pharmaceutical manufacturing [5, 10].

Methodology

Experimental Work

This research was conducted on a powder-to-tablet manufacturing pilot plant available at the University of Sheffield as shown in Fig. 1. The pilot plant comprises a Continuous tabletting Line (CTL) known as Consigma25 (GEA Pharma Systems, Collette™, Wommelgem, Belgium)which includes

(a) The continuous powder processing line (GEA Pharma Systems, Collette™, Wommelgem, Belgium)

(b) The continuous rotary tablet press (GEA Pharma Systems, Collette™, Wommelgem, Belgium)

Fig. 1 Continuous Tabletting Line (CTL)

a continuous powder processing line and a continuous rotary tablet press. The continuous powder processing line encompasses a series of interconnected processing stages, including a loss-in-weight feeder, twin screw granulator, segmented fluidised bed dryer, conical milling machine, and helical ribbon blender, all of which collaboratively process the powders into granules Fig. 1(a). The rotary tablet press as shown in Fig. 1(b) is directly linked to the continuous powder processing line, effectuating the transformation of the produced granules into tablets in continuous, and high speed mechanism.

The development of predictive models for the Continuous Tabletting Line requires comprehensive and representative datasets that can be effectively utilised by machine learning algorithms. However, generating such datasets presents challenges due to the complexity of the CTL's integrated operations, especially when multiple process parameters are manipulated across several unit operations to explore relationships within and between processing stages. This may necessitates an experimental design capable of capturing higher-order interactions, yet such comprehensiveness may be expensive in terms of materials, time, and operational resources. To address this, an efficient experimental design based on Taguchi Orthogonal Arrays was adopted in this study to provide representative data that can be readily utilised by machine learning algorithms [29]. The main advantage of such a technique is the systematic and balanced operation of the CTL line to provide diverse and informative data with minimal experimental burden. This strategy involved manipulating seven critical process parameters related to multiple processing units within the CTL. For each identified process parameters, three levels were deliberately selected to achieve a diverse range of effects on the resultant material attributes, particularly those concerning the physical characteristics of the tablet such as granules particle size, moisture content, and tablet tensile strength. This has resulted in a final design required only 81 experimental runs. Table 1 shows the process parameters for each processing stage within the CTL and their range levels to be maintained across the experimental runs. It is worth to note that the 81 experimental runs were conducted within a systematic research framework that aimed to balance the requirements of

Table 1 The process parameters utilised for data collection and process operation of the CTL

Unit Operation	Process Parameter	Parameter Level
Twin Screw Granulator (TSG)	L/S Ratio	0.1, 0.2, 0.3
	Screw Speed (rpm)	400, 600, 800
Fluidised Bed Dryer (FBD)	Drying Temperature (°C)	50, 60, 70
	Drying Time (s)	500, 600, 700
	Inlet Air Flow ($m^3\ h^{-1}$)	300, 350, 400
Milling Machine (MM)	Milling Mesh Size (mm)	0.99, 1.4, 1.57
Tablet Press (TP)	Compaction Force (kN)	5, 10, 15

data-driven modelling with practical constraints. For example, this selected design enabled the efficient coverage of the multivariate interactions across the CTL's unit operations through minimised experimental runs compared to more exhaustive designs, such as full factorial. This allowed for the development of predictive models across the CTL's unit operations while maintaining material and operational efficiency. The selection of these process parameters and their levels was based on initial screening tests indicating their local sensitivity to the material physical characteristics. In addition, the material attributes, including the particle size distribution of the granulation process, and both the particle size distribution and the moisture content after drying and milling processes, were assessed in this study due to their measurable impact on the final tablet physical profile. In addition, the tablet physical profile, such as the tensile strength, was also evaluated after tabelting process. Within each of these runs, a batch of 20 tablets was produced as the final process outcomes, culminating in a substantial total of 1620 tablets across all runs. This dataset formed the foundation for training the sequential machine learning models that underpin the plant-wide predictive modelling framework proposed in this study. Figure 2 is a block diagram demonstrating the methodological flow of multistage processes involved in this study. The following sections discuss a detailed overview on the systematic collection of processes data across the continuous tabletting Line and the steps involved in developing the predictive models.

Material Preparation

In this study, a powder blend of three pharmaceutically relevant materials was utilised. Alpha-lactose monohydrate (Pharmatose 200M) constituted 72% w/w as the bulk, microcrystalline cellulose (Pharmacel 101) accounted for 24% w/w as an excipient, and polyvinylpyrrolidone (Povidone K30) served as a binder at 4% w/w. These materials are pharmaceutically suitable for pilot-scale operations and thus enabled the development of a predictive model based on process–material complex interactions across the CTL. A total powder blend mass of 7.5 kg was mixed in a tumbler mixer (INVERSINA 20L) for 10 minutes to ensure uniformity before being transferred to the loss-in-weight (LIW) feeder hopper, which supplied a consistent mass flow to the granulation unit.

Process Operation and Data Collection

In this study, the Consimga-25 continuos tabletting line pilot plant as shown in Fig. 1 was utilised for data collection and processes operation including the Twin Screw Granulator (TSG), Fluidised Bed Dryer (FBD), Milling Machine (MM), and the Tablet Press (TP). Pre-experimental calibrations were initially conducted for the utilised process operations. Starting from the continuous powder processing line, the wet granulation process was initiated using the

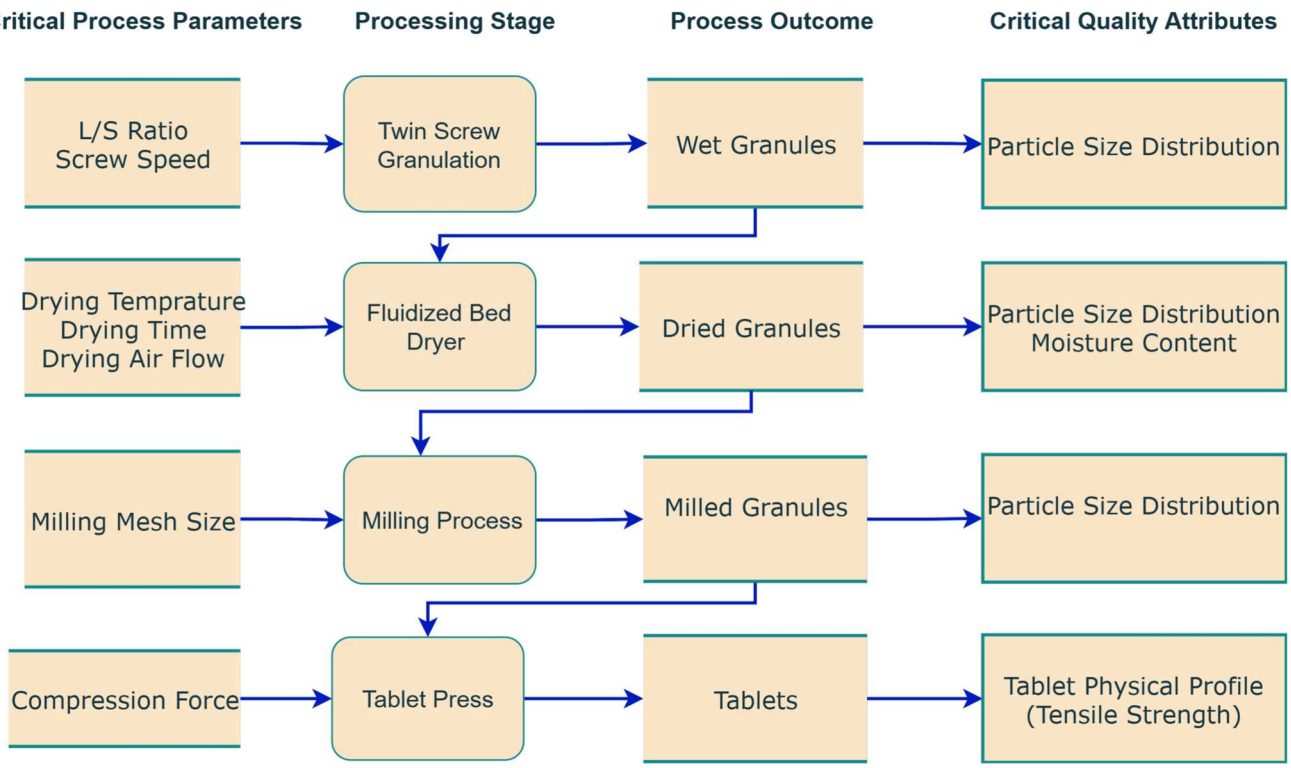


Fig. 2 The multistage processes of continuous tabletting line

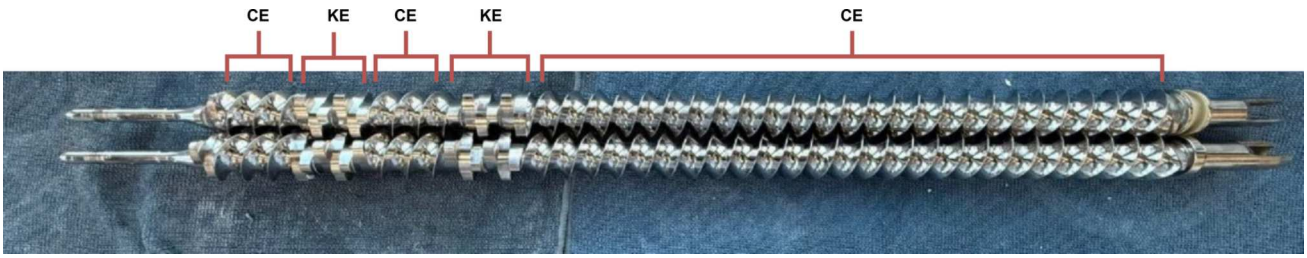


Fig. 3 The Twin Screw Granulator (TSG) configuration showing conveying elements (CE) and kneading elements (KE)

TSG. The TSG consists of two intermeshing, co-rotating screws that convey and compact the powder blend along their length during rotation [30]. Each screw had a diameter of 20 mm with a corresponding L/D ratio of 25:1. The screws were configured using a combination of conveying and kneading elements designed to facilitate consistent powder transport, liquid binder distribution, and granule formation. As shown in Fig. 3, the configuration began with three consecutive forward-conveying segments with a total length of 350 mm. These elements were positioned directly under the feed inlet in order to convey the powder blend into the downstream granulation region. The granulation region consisted of two kneading blocks, each comprising six discs staggered at 60°, and separated by a 37.5 mm forward-conveying elements. As the powder was conveyed forward, deionised water, serving as a liquid binder, was introduced via a peristaltic pump through an injection port located directly above this region. The binder was distributed along the screw length, and agglomeration was facilitated within the kneading sections, where shear and compaction forces promoted granule formation. A final conveying section was placed after the second kneading block to transport the wet mass toward discharge. It is worth mentioning that the TSG configuration remained constant across all experimental runs. In TSG, two critical process parameters were manipulated: the liquid feed rate, and the twin screw speed, while the material feed rate was fixed at 166 g/m as shown in Table 1. Moreover, the granulator was directly connected to the FBD with six cells, each cell were filled with wet granules for a duration of 180 s. The wet granules were exposed to drying air introduced at the FBD base. Here, drying occurred semi-continuously and individually per cell, enabling heat and mass transfer between the drying air and granules. Three FBD process variables were manipulated: airflow rate, drying temperature, and drying time as shown in Table 1. To maintain consistency, three cells were continuously processed per run, with the first cell disregarded to ensure stability. After each drying cycle, the cell was discharged and granules passed over inline Near-Infrared Spectroscopy (NIR) probe for moisture content measurement [31]. Thus, the average moisture content of the latter two cells was recorded. Post-drying, the granules were milled via a conical milling machine to de-agglomerate any clumped granules, achieving uniform particle size distribution. In this stage, three different mesh screen sizes were utilised as critical process parameters as shown in Table 1, while milling time and milling speed were kept constant at 100 s and 1000 rpm respectively. The resultant milled granules were then transferred to a ribbon blender where 0.6% w/w magnesium stearate was added before tablet compression. Finally, the continuous rotary tablet press was used to compress the milled granules into tablets. Here, three levels of compaction forces were manipulated as critical process parameters of the rotary tablet press as shown in Table 1. Moreover, the turret speed was maintained at 20 rpm. During each processing run, 20 tablets were randomly sampled for the analysis of their physical properties such as diameter, thickness, and breaking force using semi-automated tablet testing unit (LAB.line P4). Hence, the results for each run were averaged, and the tensile strength of the tablets was calculated. The tensile strength was determined using an equation that incorporates the tablets' physical properties, including breaking force, diameter, thickness, and a constant value for the convex cap related to the tablet press punch cap depth [32, 33], as follows:

$$\sigma = \frac{10F}{\pi D^2 \left(2.84\frac{H}{D} - \left(0.126\frac{H}{H-2H_{cap}}\right) + \left(3.15\frac{H-2H_{cap}}{D}\right) + 0.01\right)} \quad (1)$$

Where:

- F is the breaking force of the tablet,
- D is the diameter of the tablet,
- H is the tablet thickness,
- H_{cap} is the punch cap depth ($H_{cap} = 1.21$ mm).

Particle Size Distribution Analysis

In this study, particle size distribution (PSD) analysis was conducted following granulation, drying, and milling. Sample collection was performed offline and included wet, dried,

and milled granules, which were periodically extracted for characterisation using a dynamic image analyser (Camsizer X2). The procedure for collecting a representative and an accurate PSD analysis at each powder processing unit was carefully conducted. For example, after granulation, the wet granules were collected from the twin-screw granulator (TSG) outlet after steady-state conditions were reached, thereby avoiding variability associated with process transients. For each experimental run, between 40 to 50 grams of granule samples were taken using a collection tray. The collected granules were then left to dry for 48 hours at ambient conditions to allow moisture stabilisation. In addition, another round of sampling was performed following the drying and milling stages which were collected and then directly utilised for PSD measurement, as shown in Fig. 2. Prior to any measurement, all samples were split using a sample splitter, which partitioned the material into two symmetrical subsamples to ensure representativeness across the bulk granule mass [34]. These steps ensured that the samples were as representative of the bulk as possible and that sources of sampling bias such as granule inhomogeneity during transient process states and process instability were systematically minimised. To assess the size of the granules, each sample was then independently analysed by feeding them into the dynamic image analyser (Camsizer X2), where they passed through high-resolution cameras that automatically captured images of the particles. From these images, the equivalent spherical diameter of each particle was measured. This involved categorising particles into specific size classes each defined by a size interval. In this study, the PSD for each measured sample was represented using 20 size classes where each class indicates the volume fractions of particles falling within its specific size interval measured in millimetres. Furthermore, in order to aid the visual presentation of the PSDs, they were smoothed using cubic spline interpolation [34, 35]. This method was applied to generate smooth and continuous curves from the discrete size class data without altering the original measured fractions. Furthermore, similar steps were also applied to generate cumulative size distributions in order to facilitate the graphical extraction and comparison of particle size percentiles, such as D_{10}, D_{50}, and D_{90}. These parameters summaries the representation of the PSD shape [34].

Data-Driven Modelling Framework

The primary objective of this study is to develop a plant-wide modelling framework based on operational data of the state of the art continuous tabletting line (CTL). Ultimately, this work aims to establish a sequential modelling approach capable of accurately predicting the final quality attributes of pharmaceutical tablets—specifically, tensile strength—by integrating all interconnected critical process operations within the continuous manufacturing process, from powder to tablet. The sequential modelling strategy adopted in this study is designed as a multistage approach, where each unit operation within the CTL is represented by a distinct predictive model. In the CTL, critical unit operations including, twin screw granulation(TSG), fluidised bed drying(FBD), milling machine(MM), and tablet press(TP), are interconnected forming a continuous multistage process. The output of each processing stage serves as an input for the subsequent stage. This framework integrates ensemble-based machine learning (ML) algorithms to effectively represent each processing stage within the CTL. ML models were trained and interlinked to capture the relationship between these stages, thereby ensuring that variations arising from each process stage propagate to subsequent stages. This propagation of process variations through interconnected models reflects the multivariate nature of continuous manufacturing for allowing the integration of each processing stage behaviour into the final process outcome. Hence, the predictability of such system will be reinforced by modelling variability within all the subsequent stages where their effects can be traced and indicated through the sequential modelling stages.

Two ensemble-based ML algorithms were employed to model the complexities of the CTL. Ensemble modelling is a machine learning technique that combines multiple models with the same modelling objective in order to improve the predictive performance compared to a single model [22]. In this study, Random Forest (RF) and Gradient Boosting Machines (GBM) algorithms were chosen for their ability to capture intricate relationships among process parameters, material attributes, and final output characteristics based on their ensemble learning mechanism [22, 26]. Moreover, due to their non-parametric nature, these models are flexible in adapting to diverse data distributions across stages making them well-suited for multistage process modelling [36, 37]. RF is an ensemble learning method used to constructs multiple regression trees using bootstrap sampling, where each tree is trained on a randomly drawn subset of the process data with replacement [36]. Each regression tree independently partitions the feature space by recursively splitting the data that considers different features at each split which can minimise the prediction error. By combining these trees, and averaging their predictions, the RF will effectively improve the overall prediction stability with reduced variance [22]. Moreover, GBM is an ensemble learning technique that utilises the boosting algorithm [37]. GBM constructs a series of weak models (e.g. regression trees), where each model is trained on the process data. Unlike the RF, GBM builds models iteratively in a way that each trained model attempts to correct the errors of the previous models gradually improving the overall prediction. This sequential correction process is driven by a gradient-based optimisation algorithm which

utilises a differentiable loss function to minimise the squared error resulting in highly effective learning method for complex process data [28]. Both ML algorithms were utilised in this study to assess their predictive performance in determining the tensile strength of tablets. Their results were independently evaluated to compare their effectiveness in modelling the CTL. The Python programming language was primarily utilised to develop the models using the scikit-learn library [38]. The following sections describe the stepwise development of this plant-wide modelling framework, which provides a novel modelling approach for the continuous manufacturing of pharmaceutical tablets.

Plant-Wide Modelling Framework Development

The plant-wide modelling approach adapted in this study integrated ML models for processing stage into a holistic predictive framework. In the CTL, the first three unit operations such as, TSG, FBD, and MM, are primarily responsible for pharmaceutical powder processing. These units continuously process pharmaceutical powders in sequential stages, refining their physical characteristics to be readily utilised for tablet compression process. Within the Quality by Design (QbD) framework, particle size distribution (PSD) is considered a critical material attributes (CMA) [3]. The importance of the powder processing operations lies in their collective impact on the PSDs, which in turn affects subsequent processing stages. For example, variations in PSDs may arise due to the impact of critical process parameters (CPPs) across different unit operations. In addition to PSD, moisture content is also classified as a CMA [3], as it determines the remaining liquid binder content in the material post-drying, subsequently impacting granule properties and tablet performance. While these are typically classified as critical quality attributes (CQAs), they are treated here as CMAs to reflect their function as inputs to downstream processes [3]. Understanding these material attributes effects in conjunction with the CPPs of the CTL should lead to a holistic insight into tablet tensile strength prediction, which is considered as CQA [3]. Thus, in this study, a sequential modelling framework incorporating ML models was developed to predict the intermediate CMAs such as PSD and moisture content based on continuous powder processing operations. The predicted CMAs were then utilised as inputs for a subsequent ML model dedicated to the final tabletting stage, where tensile strength was ultimately predicted based on these intermediate attributes and tablet press process parameters. The framework is structured into three key predictive models as shown in Fig. 4: (1) PSD prediction based on continuous powder processing units, (2) moisture content prediction based on TSG and FBD units, and (3) Tablet Tensile Strength prediction based on the predicted PSD, and moisture content as material input features, and tablet press operational parameters. All of these models were developed through a series of steps including data pre-processing and scaling, models training, and evaluation [22]. For instance, the collected processes data mentioned in "Experimental Work" section were partitioned into input features and target outputs for each

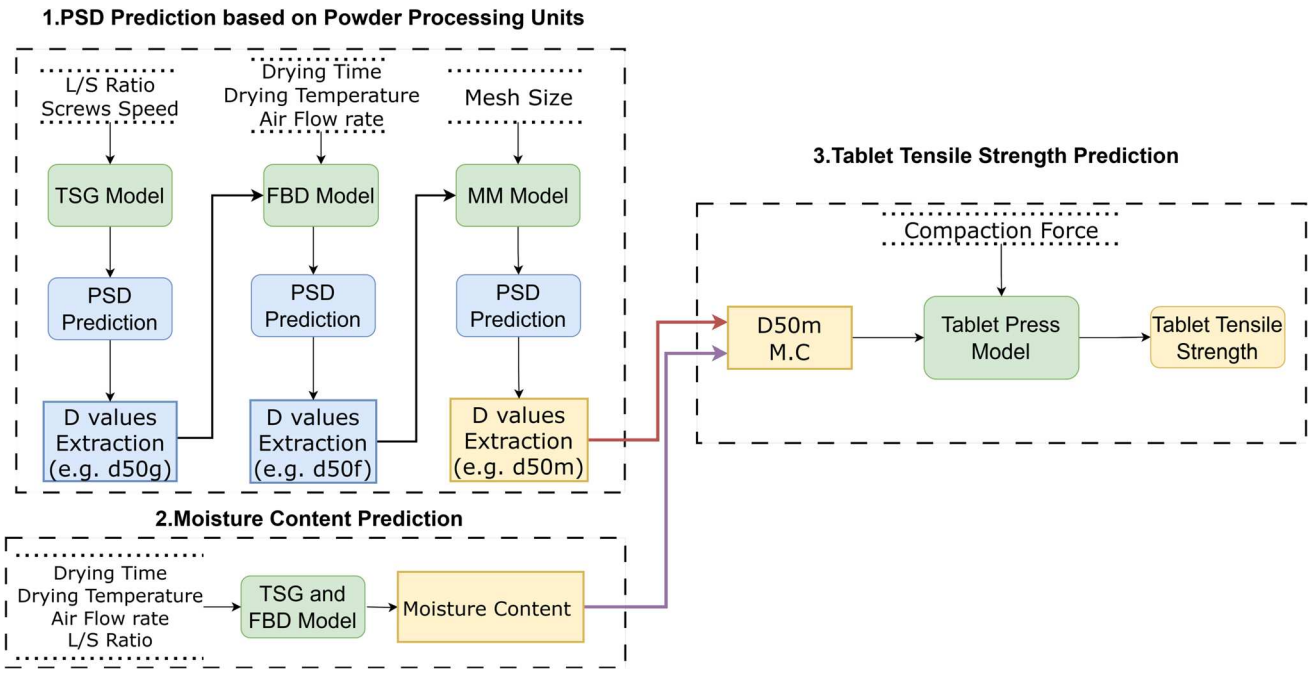


Fig. 4 Plant-wide modelling framework for the continuous tabletting line

processing stage within the CTL. The input features represents operational variables and material attributes, while target outputs refers to the process outputs to be predicted, and then utilised as input for the subsequent process model as summarised in Table 2. To ensure consistency across all stages, all input features were scaled using min-max scaling. Models training was conducted using a five-fold cross-validation approach. Finally, the performance of the trained models was assessed by calculating evaluation metrics, including the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2), to quantify and illustrate their predictive accuracy and generalisation ability within a 10% error bands. The following sub-sections discuss each modelling framework and the associated modelling process.

Particle size distributions(PSDs) Model

In this study, PSDs were sequentially measured and modelled across the three continuous powder processing units such as TSG, FBD, and MM under varying process conditions, as detailed in "Particle Size Distribution Analysis" section. The motivation for estimating the complete PSD lies in obtaining a comprehensive view of how CPPs influence powder behaviour across the full range of particle size fractions. For instance, it enables the detailed tracking of variations in fine, median, and coarse fractions in response to different processing conditions. While key percentiles such as D_{10}, D_{50}, and D_{90} are commonly used to characterise PSDs, predicting the entire distribution also allows for the derivation of additional statistical features, including spread, skewness, and distributional shapes [34]. However, modelling PSDs is complex, as they represent distributions of volume fractions across multiple size classes, each corresponding to the proportion of particles within a specific size interval. In order to solve this, a multiple-input, single-output (MISO) modelling strategy was adopted in this research where independent ML models, including RF and GBM, were developed for each size class. For example, a single ML model was trained to predict the particle size fraction corresponding to a specific size class as a function of powder processing stage parameters. Hence, the challenge of modelling the entire PSD profile was addressed in this research by treating each size class as an independent target variable. Specifically, for m size classes, m independent ML models were trained, each predicting the volume fraction of particles within its respective size class based on process parameters from TSG, FBD, and MM. Since 20 size classes were defined to represent the PSDs within each process stage, 20 individual models were developed using ensemble ML algorithms. Once the size fractions across all classes were predicted, they were reassembled to reconstruct the complete PSD profile. Each estimated fraction is then normalised by dividing it by the total sum of all predicted fractions. Furthermore, the predicted PSDs were smoothed to support visual comparison with measured PSDs, each predicted distribution was smoothed using cubic spline interpolation [34, 35]. Finally, to facilitate the extraction of key percentiles, the predicted frequency distributions were converted into cumulative distributions for the extraction of D_{10}, D_{50}, and D_{90}. Moreover,the predicted D_{50} values, together with the corresponding process parameters, were sequentially integrated as input features for the subsequent ML model. It is worth emphasising that the choice of a MISO strategy was driven by the necessity of providing an efficient and comprehensive plant-wide modelling framework with reduced computational effort and less modelling complexity. While the entire PSD distribution may indeed be modelled simultaneously using a MIMO strategy (such as ANNs), the resulting model may be computationally expensive and highly complex, owing to the network architectures and the

Table 2 Table of each unit operation used for model development as well as their related input features and process outcome

Processing Stage	Process Parameter (Input)	Material Attributes (Input)	Process Outcome (Output)
TSG	L/S Ratio Screw Speed		Wet Granule Size (D_{50})
FBD	Drying Time Drying Temperature Air Flow Rate	Wet Granule Size (D_{50})	Dried Granule Size (D_{50})
MM	Milling Mesh Size	Dried Granule Size (D_{50})	Milled Granule Size (D_{50})
TSG-FBD	L/S Ratio Drying Time Drying Temperature Air Flow Rate		Moisture Content (MC) %
TP	Compaction Force	Milled Granule Size (D_{50}) Moisture Content (MC) %	Tablet Tensile Strength (MPa)

learning algorithms that demand more time and effort to tune and initialise [22, 39]. The sequential modelling framework for continuous powder processing is illustrated in Fig. 5.

Moisture Content Model

In addition to the PSD prediction, modelling framework for predicting moisture content was developed. In this study, GBM and RF models were employed to develop an integrated predictive framework that incorporates variables from both granulation and drying stages. While the liquid-to-solid (L/S) ratio is directly related to the granulation process, its influence propagates through the drying stage. Consequently, the inclusion of L/S ratio in the drying model was essential to capture the interactions between these processes and accurately predict moisture content. In addition, the drying rates related to the FBD were also included as input features to predict moisture content. Thus, the operational parameters of these two unit operations served as the input features for predicting moisture content as depicted in Fig. 4.

Tablet Tensile Strength (T.S) Model

The development of the tablet T.S prediction model aimed to establish a data-driven framework capable of predicting T.S based on critical material attributes and process parameters within the continuos tabletting line (CTL). Since tablet mechanical properties are influenced by preceding powder processing stages, the modelling framework incorporated three key input parameters: the moisture content of the granules, the median particle size (D_{50m}) from the final powder processing stage (e.g. milling), and the tablet press

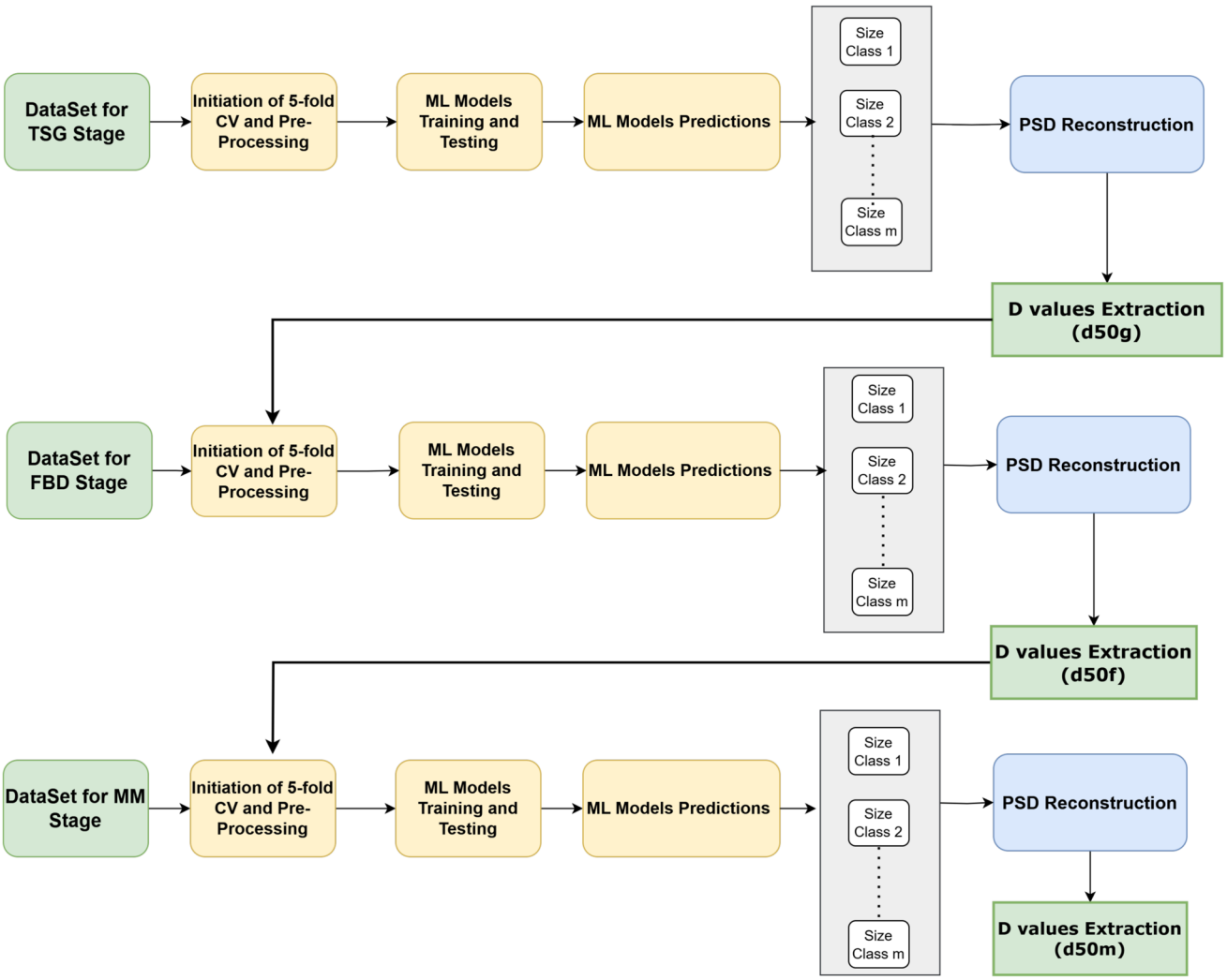


Fig. 5 PSD prediction model across the continuous powder processing units. Here, testing data(e.g. the size classes) for each stage are preserved across all folds and utilised to construct the particle size distribution in both frequency and cumulative distributions. Then, the D_{50} across all the PSDs were extracted to be incorporated as input feature with the subsequent stage process parameters. These features are then employed to predict process outputs. The resulting outputs from each stage's machine learning model are subsequently fed into the next modelling stage

compaction force. Given that the tablet press represents the final operation within the continuous manufacturing process, the inclusion of both moisture content and D_{50m} was essential, as these attributes encapsulated the cumulative effects of different processing conditions throughout the powder processing line. Thus, after achieving the predictions for intermediate material attributes, such as the median particle size (D_{50m}) and moisture content, they were incorporated within the tablet T.S model development. For example, in each fold iteration, the training subset incorporated actual measured values of moisture content and D_{50m}, while the test subset utilised predicted values of these attributes from preceding models. The predictive modelling framework for T.S is illustrated in Fig. 4

Results and Discussions

Experimental Analysis

In this section, a detailed analysis was conducted to evaluate the influence of process parameters on both the material attributes of each processing unit and the final outcome characteristics. This analysis is for elucidating the importance of utilising these parameters in data-driven model development. A comprehensive approach was utilised to examine the relationships between the process variables and material attributes related to each processing stage. By identifying and studying the relationships between these process variables and the resulting tablet characteristics, their critical importance can be confirmed and effectively integrated into the model development process.

Twin Screw Granulation Analysis

In granulation, the liquid-to-solid (L/S) ratio has a significant impact on the growth of granules and the distribution of particle sizes (PSD). The frequency distribution of granules at different L/S ratios is depicted in Fig. 6(a). At the lowest L/S ratio, the prevalence of granules tends to be characterised by small size, which is a result of the low wetting condition. As the L/S ratio increases, the granule distribution broadens, resulting in larger and coarser granules as in the case of L/S=0.3. Furthermore, the cumulative PSD plots in Fig. 6(b) illustrate the percentage of granules below a certain size. These cumulative distribution plots can be utilised to extract percentile values such as (D_{10}, D_{50}, D_{90}) in order to measure the quantitative impacts of varying L/S ratios [34, 40]. For instance, when the L/S ratio is low (L/S=0.1), the D_{50} of the granule distribution indicates that 50% of the granules are less than 160 micrometers, which corresponds to fine to medium granules. At a moderate L/S ratio (L/S=0.2), approximately 50% of the granules exhibit an increase in size to approximately 540 micrometres, indicating the presence of medium to large granules. At a high L/S ratio (L/S=0.3), the median particle size D_{50}, is approximately 1780 micrometres, suggesting the presence of larger and coarser granules. Thus, L/S ratios have a critical impact on the subsequent tablet manufacturing units, as the downstream processes will react differently to the different sizes of granules produced.

Fluidised Bed Dryer Analysis

The fluidised bed drying process is notably complex due to the interconnected effects of drying conditions(e.g. dry-

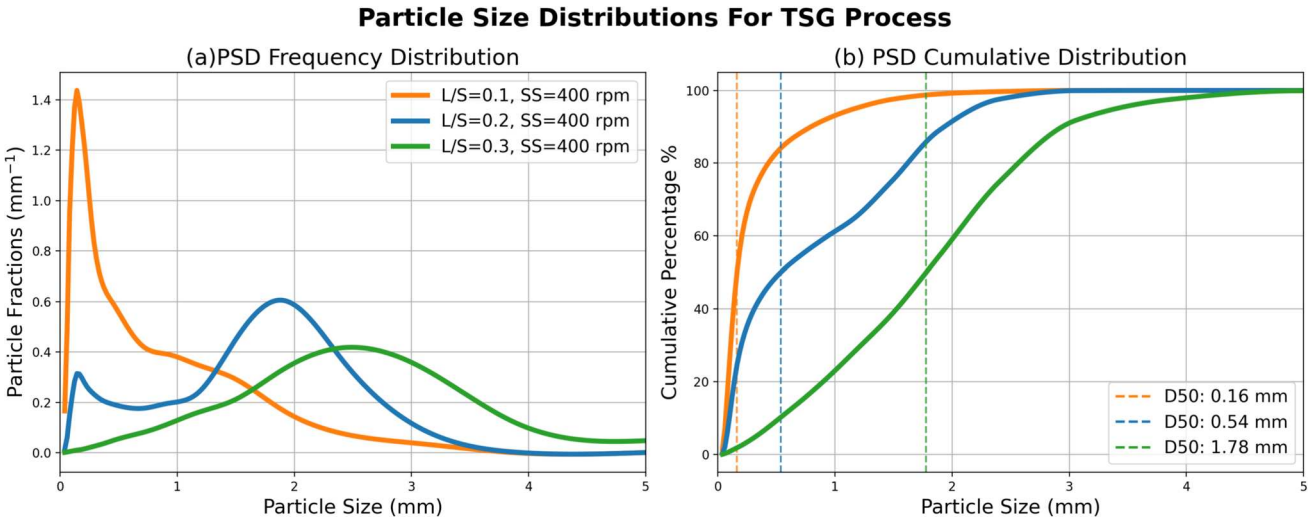


Fig. 6 Granule size distributions of twin screw granulation related to different L/S ratios

ing time, drying temperature and air flow rate), the initial L/S ratio, and the variations in granule size distribution following granulation. In particular, the L/S ratio has a pivotal influence on the moisture content of the granules, which subsequently shapes their drying characteristics and size distribution. Figure 7 demonstrates PSD samples for dried granules when the initial L/S ratio was set to 0.1, where moisture contents were fairly uniform (ranging from 4.8% to 5.5%) and the distributions were relatively narrow, concentrated around the median particle sizes (D_{50}). A slight shift toward larger granules was observed in samples with higher moisture levels. When the initial L/S ratio was further increased to 0.3, moisture contents rose more sharply (5.8% to 9.8%), accompanied by broader particle size distributions. These observations indicate that L/S ratios intensify the non-uniform nature of the drying process, underlining how tightly coupled drying conditions are with the initial moisture content of the granules.

Milling Machine Analysis

After drying, the granules were milled using a cone milling machine equipped with different mesh sizes to achieve uniform particle size distributions. Samples of these dried and milled particles were collected to examine their size reduction behaviour in relation to the mesh size and their final moisture content. Figure 8 shows the particle size distributions of milled granules related to different Moisture Content, processed through conical milling machine mesh sizes of 0.99 and 1.4 mm post-drying. At low moisture contents (e.g. 5% or less), the milled granules exhibit a relatively uniform size distribution across all mesh sizes, concentrated around smaller particle sizes. Conversely, at higher moisture content levels (e.g. 6% or above), there is a broader distribution in the particle size, with a notable shift towards larger particles. Thus, the narrower distributions at lower moisture content highlight the criticality of drying stage process in achieving a uniform particle size distribution before tablet compression.

Tablet Press Analysis

The rotary tablet press is considered the final stage within CTL, in which the granulated powder blend is processed into tablets as the final product. Tablet tensile strength is considered the final quality outcome to be analysed according to the moisture content in which it regarded as a key condition of the processed particulate material influencing the behaviour of tabletting process stage. The interactions between the compaction force and the tablet tensile strength under different ranges of moisture contents are illustrated in Fig. 9. It can be seen that tensile strength generally increases as compaction force increases, resulting in stronger tablets. For instance, at low to moderate moisture content, the increase in tablet tensile strength is gradual and exhibits less variability. This is crucial for achieving specific hardness levels to ensure tablet strength within a defined range. However, for high moisture content ranges, the produced tablets produced exhibited reduced tensile strengths and greater variability,

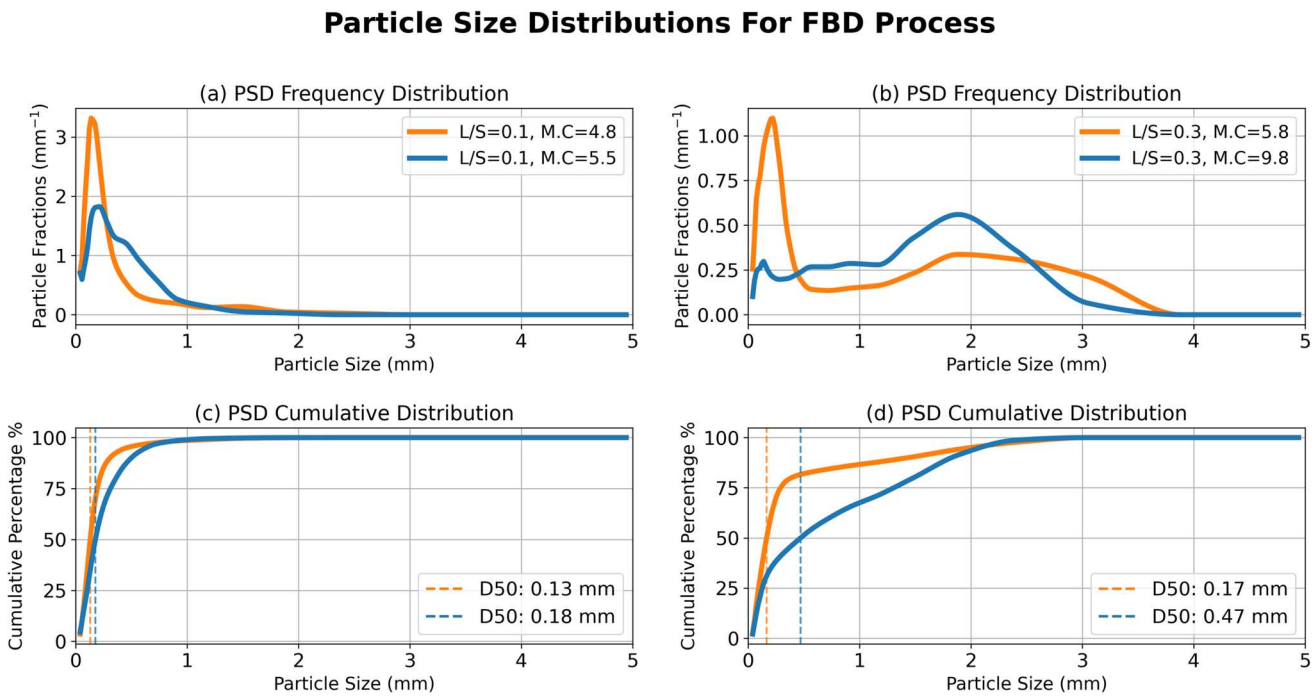


Fig. 7 Granule Size Distributions after Fluidised Bed Drying related to different Moisture Contents

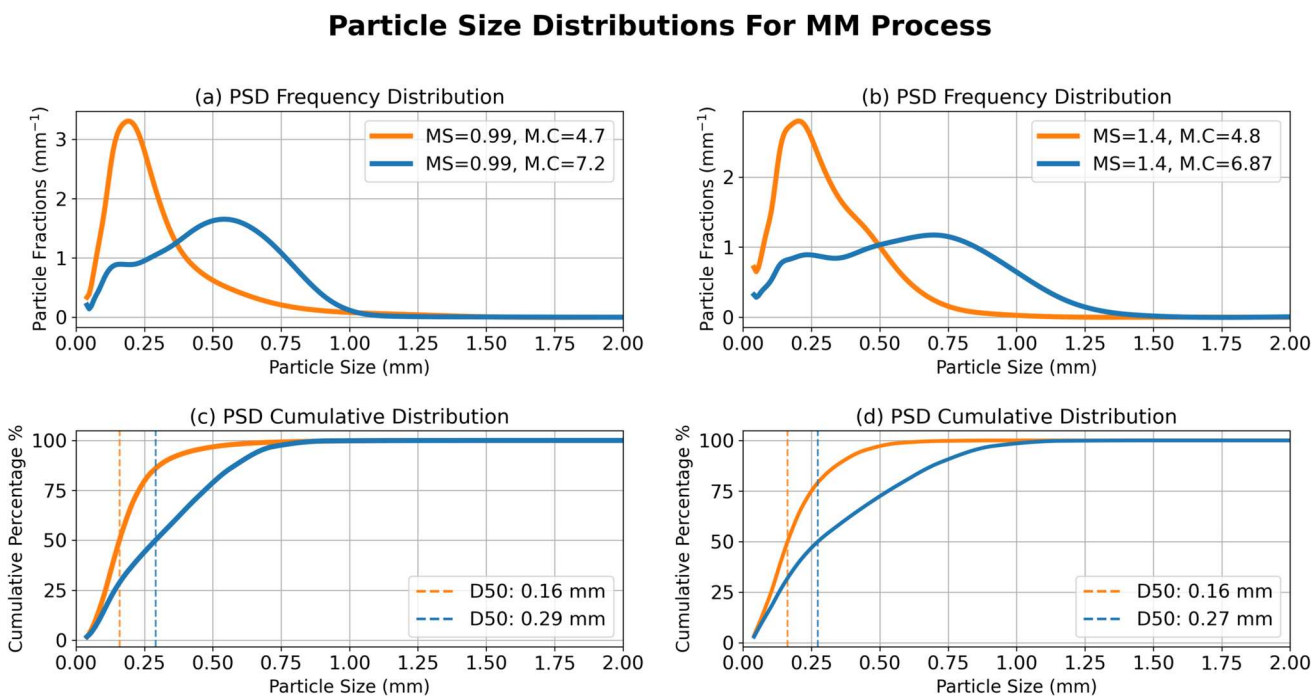


Fig. 8 Granule size distributions after milling process related to different moisture content and milling mesh sizes

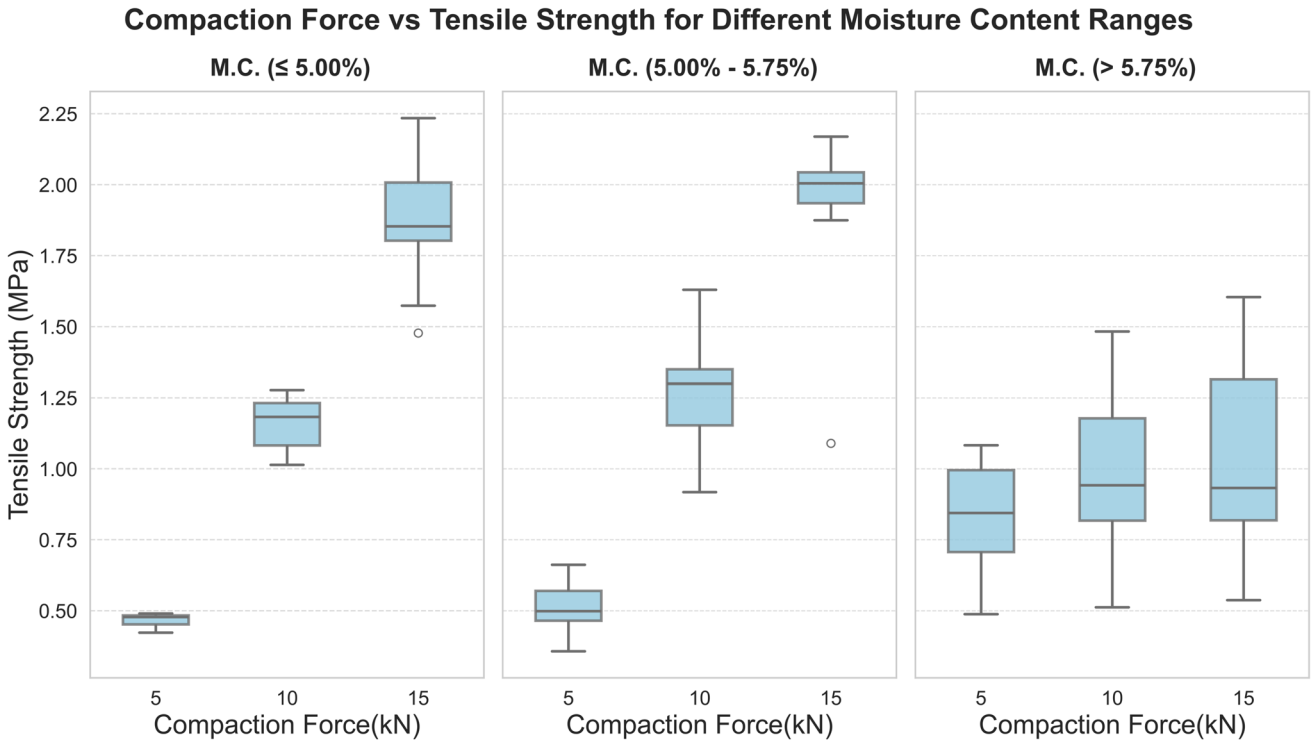


Fig. 9 Effect of compaction force on tensile strength across different moisture content ranges. Each subplot represents a distinct moisture content range. Box plots illustrate the distribution of tensile strength at varying compaction forces

despite the increase in compaction force. This non-linearity suggests that once a certain moisture level is reached, further increases in compaction force do not yield a proportional rise in tablet tensile strength. Despite the fact that the granules were dried under different conditions, those with a higher moisture content may have remained only partially dried. For example, according to the experimental design, there were instances where the drying conditions were highly aggressive, which might have led to generally acceptable moisture levels for granules with a high L/S ratio. Nevertheless, certain batches could have been less thoroughly dried, thereby contributing to the plateau effect. This is evident because the notable variability observed in tablet tensile strength at high moisture content can be attributed to multiple samples having experienced different drying conditions, which may have either increased sensitivity or reduced responsiveness to compaction force. Thus, the variability in tablet tensile strength, especially at initial higher L/S ratios, underscores the impact of drying conditions on the moisture content of granules, thereby influencing the final process outcome.

While these experimental results identified the sensitivity of the process parameters on the quality attributes, their utility remains limited and localised to the individual process stages within the CTL. For example, the complexity of the CTL arises from the sequential nature of integrated processing stages. Each stage not only individually influences the material attributes but also propagates these influences downstream, therefore compounding the complexity as materials progress through subsequent unit operations. Hence, the plant-wide modelling approach presented in this study is intended to complement and extend these empirical insights by enabling a system-level integration of the multistage processes of the CTL through the development of the sequential ML models in order to capture such complexity between these multivariate interactions on the final product attribute such as the tablet tensile strength.

Plant-Wide Modelling Framework Results

In this study, a plant-wide modelling framework for the continuous manufacturing of pharmaceutical tablets was developed in a sequential manner using machine learning models, including Gradient Boosting Machines (GBM) and Random Forest (RF). Each processing stage was trained using relevant training data, and the predicted output for each stage—based on the testing set—was used as an input for the subsequent stage. It is important to emphasise that the modelling strategy was designed to capture variations attributable to the manipulation of critical process variables across the CTL without the inclusion of raw material properties. Specifically, the twin screw granulator (TSG) utilised powder blend with fixed material type sourced from a single manufacturer, following a standardised preparation and mixing as described in "Material Preparation" section. Nevertheless, intermediate material attributes such as granule particle size distribution (PSD) and moisture content were measured at each relevant stage and subsequently integrated as inputs to downstream models. This enabled indirect characterisation of material-process interactions within the modelling framework, while maintaining a focus on operational-level variability. The following sections discuss the results related to the development and evaluation of the plant-wide model.

Particle Size Distribution Predictions

In PSD predictions, the process parameters for the continuous powder process line including TSG,FBD and MM were used as input features, in conjunction with the D_{50} values of the process units' PSDs, for model training. Each model predicted the fraction of particles corresponding to a specific size class. Given that the PSD represents a continuous distribution, the predicted fraction values across all size classes were aggregated to reconstruct the full frequency distribution. A direct comparison between the predicted and measured PSD curves was conducted to evaluate the accuracy of the models predictions. Figure 10 presents this comparison for a representative particle sample collected at the final stage of the continuous powder processing line, specifically from the milling unit. While all predicted PSDs were utilised for evaluation, a single example is provided here to illustrate how the (GBM) model captures the overall distribution trend. The results indicate that the predicted PSD closely aligns with the measured PSD, effectively capturing the overall shape and key distribution characteristics. However, deviations were observed across multiple size classes, where the predicted fractions exhibited minor deviations relative to the actual measured values. These small deviations can be attributed to the uncertainty related to the inherent variability in the PSD data. Moreover, the non-linear relationships between process parameters and particle size fractions may further complicate the models predictive performance. This is particularly pronounced when each size class was modelled separately, potentially missing out on shared non-linear interactions across size classes. Despite this, the overlay of the predicted and actual PSD curves demonstrated that the developed models effectively captured the overall distribution shape, with deviations occurring predominantly in specific size fractions rather than across the entire range. To further assess predictive performance, cumulative distribution plots were generated based on both the predicted and measured PSD. Figure 10 illustrate these cumulative PSD curves, where the two distributions were used to extract key particle size percentiles, including D_{10}, D_{50}, and D_{90}. As the primary objective of this modelling framework was to ensure

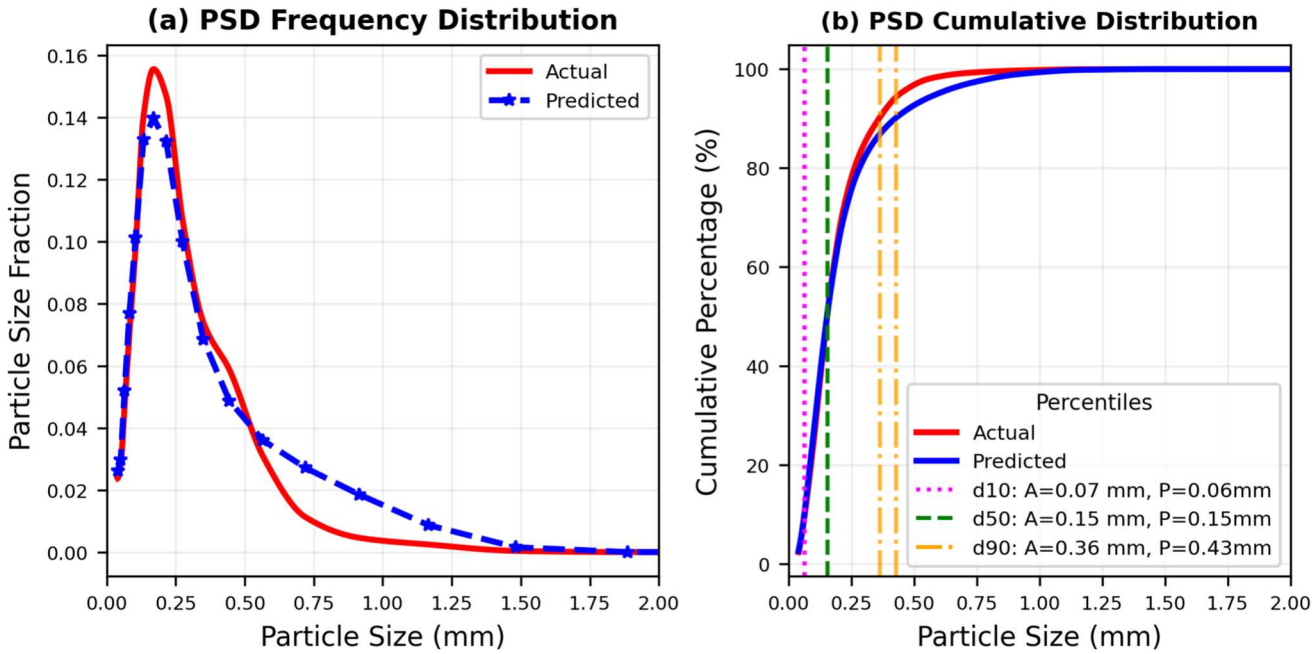


Fig. 10 Comparison of Actual and Predicted Particle Size Distribution (PSD) for the Milling Machine Model. a) The figure shows the Frequency Distribution, comparing actual (red curve) and predicted (dashed blue curve) data. b) The figure presents the cumulative distribution of actual and predicted PSDs, respectively, highlighting the extraction of the key percentiles (D10, D50, D90)

reliable predictions of the D_{50} percentile, a broader evaluation was conducted by comparing predicted D_{50} values with their corresponding measured D_{50} values across all the powder processing units within the continuous processing line.

Table 3 summarises the performance metrics for each powder processing stage related to the extracted key percentiles from the predicted PSDs. As detailed, each process stage model within the sequential framework was evaluated for its training and testing data according to the predicted key percentiles. In general, all process models showed acceptable model performances on the training data as indicated by the evaluation metrics for every key percentile values. However, as the sequential modelling progressed, a decline in the predictive performance was observed in the subsequent process models after the granulation stage. This decline was indicated by the decreasing R^2 values in unseen data for both ML algorithms. To illustrate this further, Fig. 11 presents the ML models performance at the final stage of the powder processing units, evaluating both RF and GBM models on the predicted D_{50} percentile using training and testing data. Both models achieved reasonable prediction performance on the training dataset is indicated by both the Gradient Boosting Machine (GBM) and Random Forest (RF) models by R^2 of 0.90 and a RMSE of 0.02 mm. While there are minor deviations from the actual line approximately between 0.40 to 0.50 mm, majority of the predictions were following the actual line indicating adequate fit. However, it can be observed that neither model fully captured the variability

Table 3 Evaluation Metrics for GBM and RF Models

Process Stage	Evaluation	GBM						RF					
		D10		D50		D90		D10		D50		D90	
		Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
TSG	(R^2)	0.88	0.82	0.98	0.96	0.98	0.96	0.99	0.79	0.96	0.94	0.98	0.96
	RMSE	0.08	0.09	0.1	0.13	0.14	0.18	0.02	0.11	0.13	0.16	0.11	0.16
FBD	(R^2)	0.81	0.74	0.96	0.62	0.98	0.90	0.61	0.23	0.93	0.35	0.97	0.85
	RMSE	0.01	0.01	0.02	0.08	0.06	0.17	0.01	0.01	0.03	0.11	0.09	0.21
MM	(R^2)	0.81	0.22	0.90	0.51	0.88	0.43	0.81	0.20	0.90	0.34	0.89	0.46
	RMSE	0.007	0.013	0.021	0.045	0.056	0.13	0.007	0.014	0.021	0.053	0.057	0.124

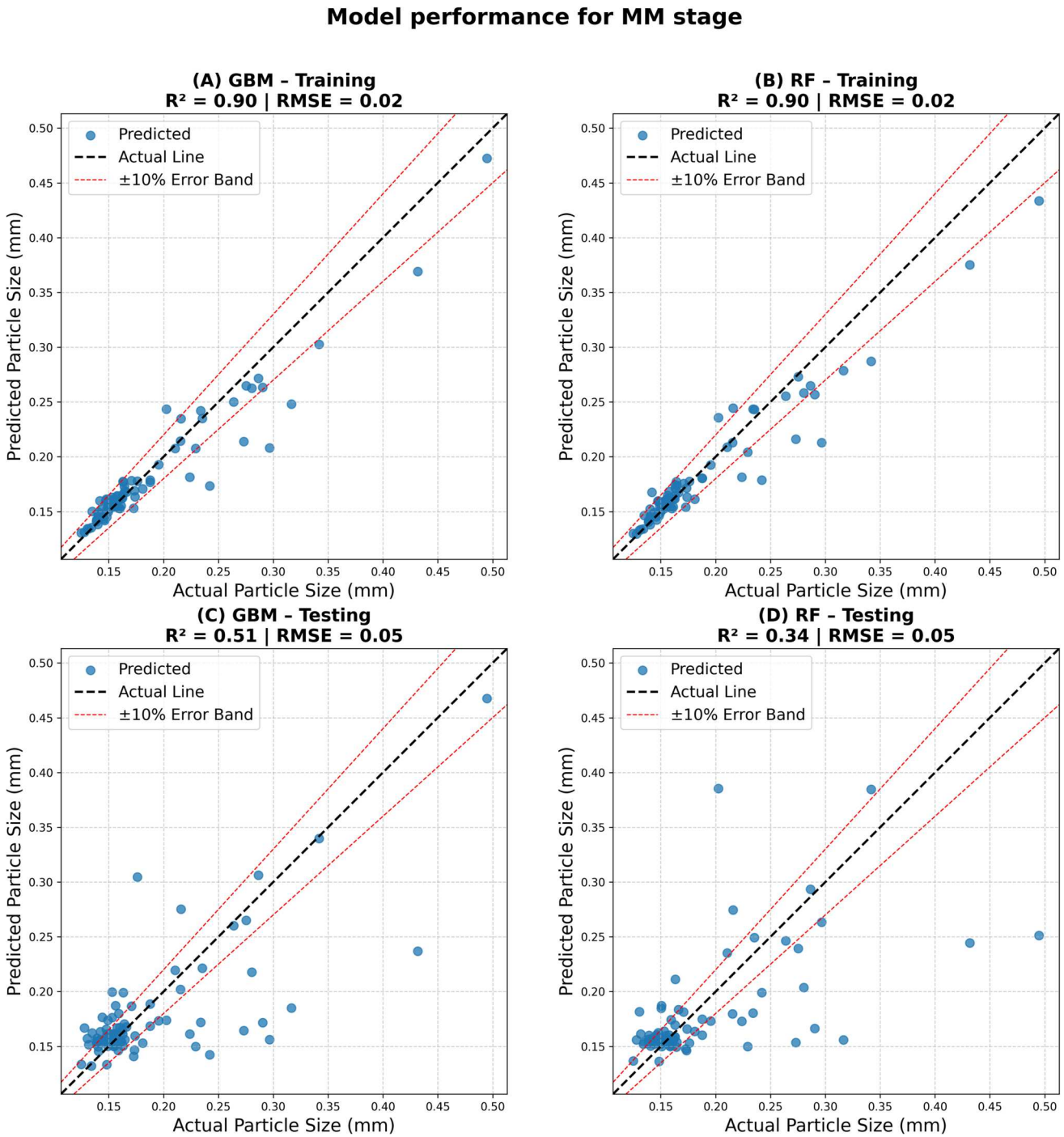


Fig. 11 The final model performance of the median particle size (D_{50m}) predictions after the extraction of key percentiles. a) Training performance of GBM, b) Training performance of RF, c) Testing performance of GBM, and d) Testing performance of RF

inherent in sequential process models, as indicated by the noticeable decrease in R^2 values within the testing dataset. For instance, both RF and GBM models achieved R^2 values of 0.51 and 0.34 respectively, indicating that a majority of the D_{50} values' variability was not fully captured within the predicted PSDs. In addition, both models showed RMSE values of approximately 0.05, indicating that on average, the predictions deviate by 0.05 mm from the actual D_{50} values'. Nevertheless, the inconsistency of the predictions falling outside the 10% error bands in the scatter plots indicates that both models struggle with uncertainty across the D_{50} range. This performance inconsistency reveals areas where the models

produce less reliable predictions. Furthermore, the uncertainty may also be attributed to the cumulative impact of prediction errors, wherein inaccuracies at one stage propagate through the sequential modelling framework, affecting subsequent models performance. While the initial results for predicting overall PSDs were promising, the performance was ultimately affected by the multi-stage nature of the processes, which reflects the inherent complexities of continuous manufacturing of pharmaceutical tablets. This decline in accuracy underscores the necessity of error mitigation strategies between interconnected process models.

Moisture Content Prediction Results

In addition to PSD prediction, the predictive performance for moisture content was also evaluated based on process conditions in twin-screw granulation (TSG) and fluidised bed drying (FBD). The GBM and RF models exhibited acceptable generalisation performance in predicting moisture content. Both models effectively captured the underlying complexities associated with process parameters and resulting moisture content, as reflected in Fig. 12. The predictive performance of moisture content, as modelled by GBM and RF, was directly influenced by the input features of TSG and FBD, which exhibit nonlinear relationships but are less complex than PSD prediction. For instance, both models achieved (R^2) value of 0.94 and RMSE of 0.25 and 0.26 on unseen data. The predictive performance of both models followed a similar pattern, with GBM slightly outperforming RF. While these results indicate that moisture content prediction is less complex compared to PSD prediction, it is still affected by uncertainty. For example,the scatter plots in Fig. 12 present few instances where the predicted moisture content values slightly exceeded the 10% error margin. While such deviations do not necessarily impact the overall performance, they may result from model uncertainty and be influenced by processes variations.

Tablet Tensile Strength Prediction Results

In the final stage of the plant-wide modelling framework, the predictive performance of final tensile strength (T.S), incorporating predictions from both PSD and moisture content(MC) models, was assessed. Both RF and GBM models were provided with predicted material attributes (e.g. D_{50m} and MC) alongside compaction force from the tablet press. Figure 13 presents the predictive performance results for both algorithms on the training and testing datasets for the tablet tensile strength. It can be seen that both models exhibited strong predictive capability on the training data, achieving R^2 values of 0.94 and 0.96 for the GBM and RF models and RMSE OF 0.12 and 0.11, respectively. The predicted tablet tensile strength values were closely aligned with the actual measurements, with the majority of data points falling within the $\pm 10\%$ error bands. Minor deviations from the ideal prediction line were observed particularly at lower tensile

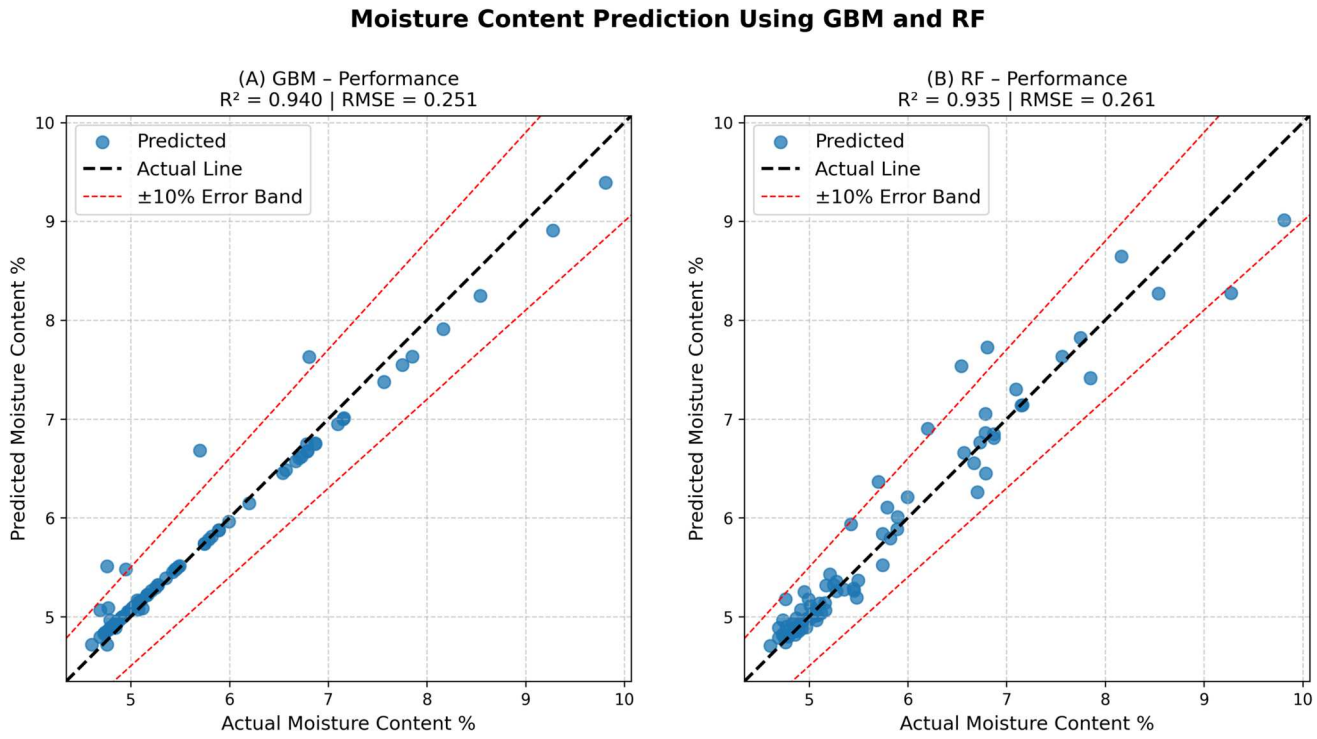
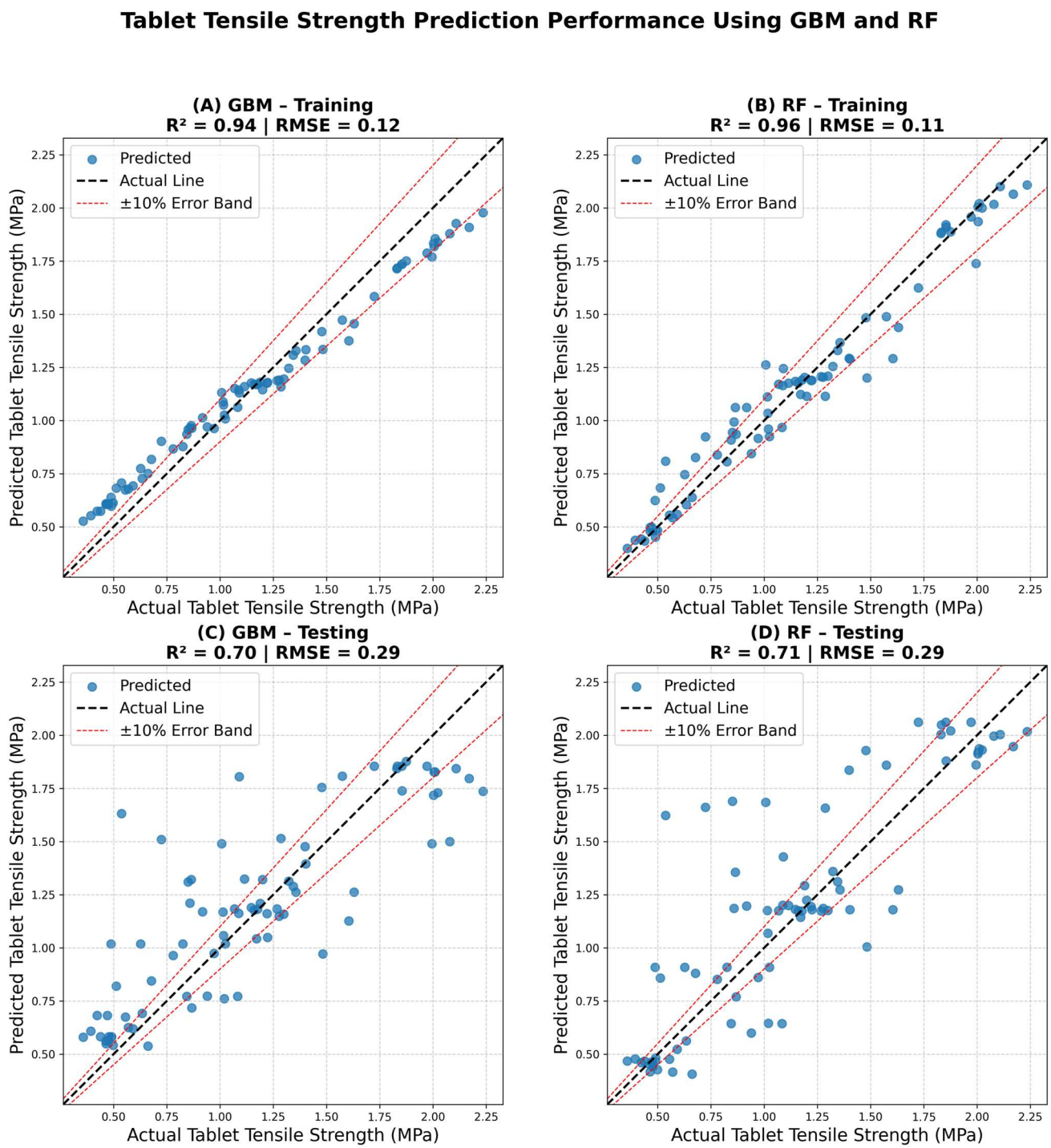


Fig. 12 The models performance of Moisture Content Predictions based on A) GBM model performance and B) RF model performance



.

Fig. 13 The models performance of Tablet Tensile Strength Predictions based on a) Training performance of GBM, b) Training performance of RF, c) Testing performance of GBM, and d) Testing performance of RF

strength values. The models demonstrated low to moderate performance on the testing sets across all the folds, with R^2 values of 0.70 and 0.71 for GBM and RF, respectively. While these values seem moderate, they are still affected by uncertainty in the predicted data. For example, GBM captures 70% of the variance in the actual data, but the majority of its predictions deviate from the actual line by RMSE of 0.29 MPa. These deviations can also be observed in RF, where its predictions deviate from the actual line by RMSE of 0.29 MPa. The limitations in predictive accuracy suggest that, while ML approaches are powerful, they were unable to fully capture the multivariate dependencies between D_{50m}, MC, and T.S, particularly in the presence of accumulated systematic errors. The observed model inaccuracies may stem from process-material interactions or random noise, and the models may struggle to distinguish between these sources of variation. Thus, these findings underscore the necessity of uncertainty-aware modelling techniques, as conventional ML models alone may be insufficient for maintaining predictive reliability across all stages of pharmaceutical tablet manufacturing.

While the predictive performance observed in Figs. 11 and 13 may be attributed to the multivariate and non-linear nature of the process data, the ML models may have been constrained in their capacity to fully capture the underlying relationships between input features and target outputs across the multistage processes of the CTL. Additionally, random noise due to the process variability may have further impaired training effectiveness. To mitigate these effects, model development followed systematic procedures to preserve predictive performance [22]. For instance, to ensure measurement integrity, pre-experimental calibrations were conducted on the unit operations to promote operational consistency across process conditions. Samples were collected and analysed in duplicate, with averaged readings used to minimise potential measurement bias, as detailed in "Experimental Work" section. In terms of data preparation and model validation, input features were scaled using min-max scaling to ensure uniform feature ranges, and a 5 folds cross validation strategy was applied across the entire process datasets for consistent training and evaluation. In addition, both ML algorithms utilised in this study were configured to capture the underlying complexity of process interactions while avoiding overfitting to processes noise. This was achieved through systematic tuning of model parameters such as the number of estimators and tree depth, with the learning rate additionally adjusted for the GBM model to ensure overall stability in model development. As shown in Table 3, the models achieved moderate to high training R^2 values across the continuous powder processing stages of the CTL. This reflects an appropriate compromise between model complexity and generalisability within the constraints of plant-wide sequential modelling. However, as the sequence of models progresses towards the final processing stage model, a decline in performance is observed. This degradation in predictive performance observed in testing folds within the sequential stages remains a limitation, underscoring the inherent complexity and variability of continuous manufacturing. The systematic model development steps contributed to achieving stable performance across the sequential models. While not optimal, they nonetheless show moderate predictive capability of the ML models for such a type of complex systems.

Uncertainties Reduction for Process Model Robustness

In order to improve the accuracy of the processes models predictions, the integration of Gaussian Mixture Models (GMMs) was undertaken across the utilised machine learning (ML) models within each processing stage. The motivation behind this approach arises from the recognition that prediction errors often contain complex patterns and hidden interactions that the primary machine learning models may not fully capture due to the inherent variability and non-linearity of the processes involved. For instance, since the sequential modelling approach was employed, in which the predicted outputs from one stage were used as inputs for the subsequent stage, this sequential dependency implied that any uncertainties in the initial stage models propagated through the entire modelling chain, potentially degrading the overall predictive performance of the plant-wide model. To address these challenges, in this study, GMMs were employed to characterise the structure of each ML model residuals (i.e., the errors between predicted and actual values) with the aim of capturing systematic patterns through probabilistic inference from the conditional error distributions [27, 39, 41]. This approach provides systematic method for refining and enhancing the predictive accuracy of primary machine learning models by adjusting their errors. Here, the GMM represents the error data distribution as a weighted sum of pre-defined gaussian components, each with its own mean, covariance, and mixing coefficient (weight) [41]. The probability density function of the error data, represented as GMM, is as follows [39]:

$$p(\mathbf{x}_e) = \sum_{j=1}^{J} \pi_j \mathcal{N}(\mathbf{x}_e|\mu_j, \Sigma_j) \tag{2}$$

Where:

- $\mathbf{x}_e$: The error data consisting of relevant input features and residual errors vector.
- J : The total number of Gaussian components used to approximate the distribution in the GMM.

- π_j: The mixing coefficient for the j-th Gaussian component representing the probability that a data error was generated from component, satisfying:

$$\sum_{j=1}^{J} \pi_j = 1 \quad \pi_j > 0,$$

- $\mathcal{N}(\mathbf{x}_e|\mu_j, \Sigma_j)$: The multivariate probability function given by:

$$\mathcal{N}(\mathbf{x}_e|\mu_j, \Sigma_j) = \frac{e^{\left[-\frac{1}{2}(\mathbf{x}_e-\mu_j)^\top \Sigma_j^{-1}(\mathbf{x}_e-\mu_j)\right]}}{(2\pi)^{d/2}|\Sigma_j|^{1/2}} \tag{3}$$

Where:

  - μ_j: the **mean** of the j-th Gaussian component,
  - Σ_j: the **covariance** of the j-th Gaussian component,
  - d: the dimensionality of $\mathbf{x}_e$.

In this research, the error data were constructed using the residuals from the primary machine learning models along with a subset of relevant input features corresponding to each process stage model. Hence, a feature selection strategy was employed to identify the most relevant input variables for error modelling. The selection process was based on the correlation between input features and residual errors, ensuring that only the most correlated variables were incorporated into the GMM error modelling framework. This was necessary in order to ensures that the GMM will not be affected by high data dimensionality, which could introduce unnecessary complexity and reduce model performance.

Once the error dataset was constructed, it served as training data for developing the GMM in order to identify hidden distributions each represented by a Gaussian component. Because the optimal number of Gaussian components J was unknown, a maximum value $J_{\max}$ was specified, and candidate GMMs with $J = 1, 2, \ldots, J_{\max}$ were considered. Each GMM model was initialised using a K-means clustering algorithm to estimate suitable starting values for its parameters $\Theta = \{\pi_j, \mu_j, \Sigma_j\}$. With these initial parameters, the Expectation-Maximisation (EM) algorithm was then employed to refine the estimates by maximising the log-likelihood [41]. EM is a two-step iterative algorithm utilised for parameter estimation in probabilistic models [39, 42]. In the E-step, the probability that each data point $\mathbf{x}_{e,n}$ belongs to component j was computed:

$$\gamma_{nj} = \frac{\pi_j \mathcal{N}(\mathbf{x}_{e,n}|\mu_j, \Sigma_j)}{\sum_{k=1}^{J} \pi_k \mathcal{N}(\mathbf{x}_{e,n}|\mu_k, \Sigma_k)} \tag{4}$$

In the M-step, the GMM parameters were updated to maximise the log-likelihood based on the computed probabilities from E-step:

$$\mu_j = \frac{\sum_{n=1}^{N} \gamma_{nj}\, \mathbf{x}_{e,n}}{\sum_{n=1}^{N} \gamma_{nj}},$$

$$\Sigma_j = \frac{\sum_{n=1}^{N} \gamma_{nj}\, (\mathbf{x}_{e,n} - \mu_j)(\mathbf{x}_{e,n} - \mu_j)^\top}{\sum_{n=1}^{N} \gamma_{nj}},$$

$$\pi_j = \frac{1}{N} \sum_{n=1}^{N} \gamma_{nj}.$$

The EM algorithm was iterated until the estimated parameters had converged to an optimal solution. Consequently, the Bayesian Information Criterion (BIC) was computed for each GMM candidate [41]. This criterion assists in selecting the optimal GMM by finding a balance between the model's goodness of fit and its complexity. As a result, the GMM model with the lowest BIC was selected to be used for refining the primary ML model predictions. This chosen GMM was utilised to reduce the prediction uncertainty through error correction. For any new input data point, the probabilities of this point belonging to each Gaussian component in the GMM were calculated. These probabilities measure how likely that the error pattern resulted from a particular process model input conditions matches the patterns characterised by each component of the GMM. Using these probabilities as weights, a conditional error was computed by taking a weighted average of the component means. By adding this conditional error term to the primary model's predictions, a refined output was obtained that accounts for the structured error patterns identified by the GMM. Figure 14 illustrates this integration strategy for the sequential modelling frameworks for Powder Processing units.

Integration of GMMs with Machine Learning Models Results

Within a sequential framework, Gaussian Mixture Models (GMMs) were integrated with the machine learning algorithms employed in this research study, leading to notable improvements in modelling performance and a reduction in prediction errors. GMMs were utilised to capture the error between actual and predicted data at each processing stage. Through this approach, residuals were clustered into distinct distributions, allowing for the quantification of uncertainty under varying conditions. To evaluate error propagation across sequential models, an analysis was conducted to track prediction errors at each stage using error propagation plots. This assessment was based on Root Mean Square Error (RMSE) values for each model representing a distinct

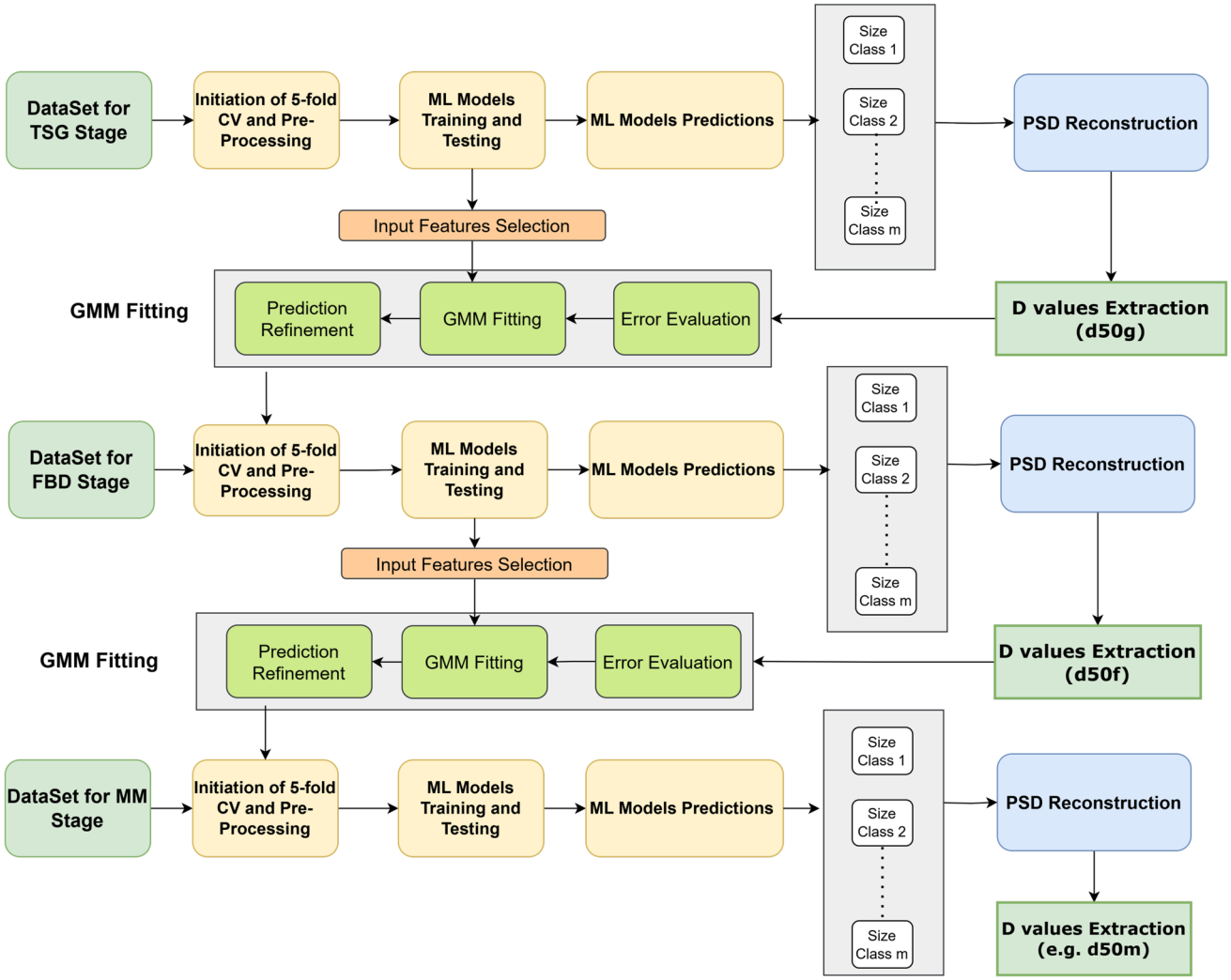


Fig. 14 PSD sequential modelling approach using GMMs for error modelling. Here, the D_{50} across all the predicted PSDs were extracted to be incorporated for error evaluations with actual data. The errors of these predictions are then correlated with relative input features and then modelled using GMMs. The GMMs outputs are subsequently employed to refine the predictions from all involved machine learning models. The resulting refined outputs from each stage's GMM are subsequently fed into the next modelling stage

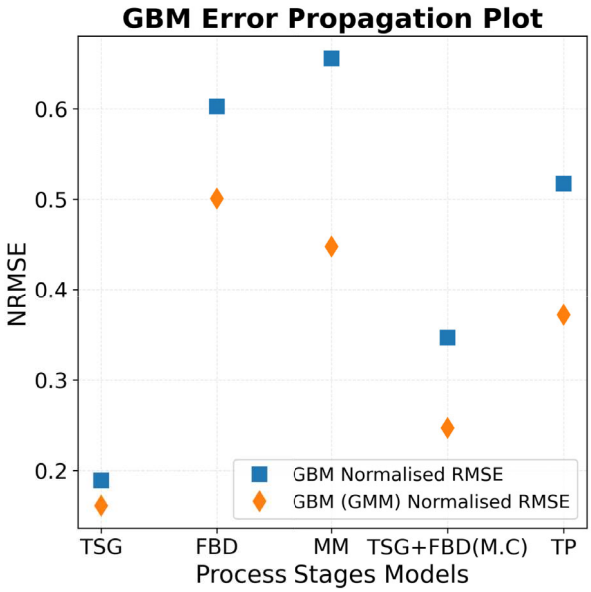
processing stage. RMSE quantifies the average magnitude of the deviation between predicted and actual values. However, since RMSE values are scale-dependent and each stage operates on a different unit scale, normalisation was necessary. This normalisation process facilitated a more interpretable measure of error that remained independent of the actual value magnitudes. Thus, the Normalised Root Mean Square Error (NRMSE) was computed by [43]:

$$\text{NRMSE} = \frac{\text{RMSE}}{\text{Standard deviation of actual values}}$$

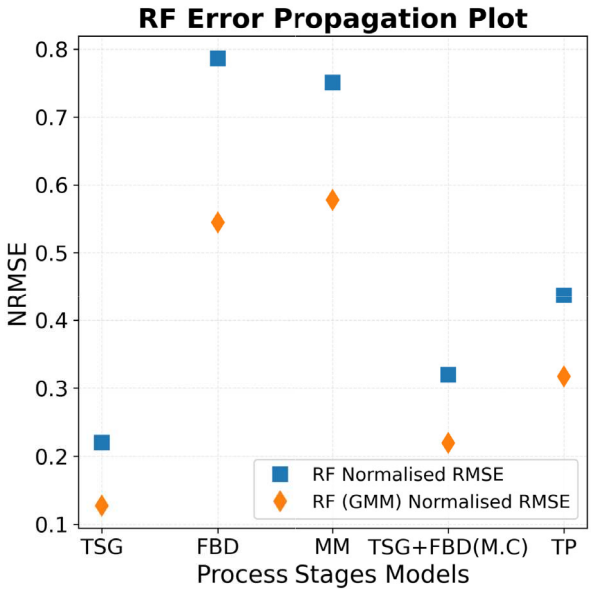
NRMSE provides a standardised approach to assessing error accumulation, indicating how relative error evolves across different stages. This method scales RMSE relative to inherent data variability, offering insights into the extent to which model prediction errors exceed natural variation in the actual data. For instance, NRMSE value close to 0 suggests adequate predictive performance, whereas NRMSE values exceeding 1 indicate poor performance, as prediction errors exceed expected variability. Hence, by comparing NRMSE values before and after the application of GMMs, it was possible to assess whether the GMM effectively captured the underlying hidden pattern and reduced the prediction error.

The total computational time required for executing the plant wide modelling was recorded for both ML models each incorporating the GMM. For GBM-GMM model, the complete process was executed in 27.7 seconds, whereas RF-GMM model achieved the same in 60.5 seconds. These

timings reflect the overall computational efficiency from model initialisation to completion. In addition, the sequential predictive performance in terms of error propagation across the process models was analysed. Figure 15 presents the error propagation plots of GBM and RF algorithms, both individually and when integrated with GMMs. When GBM and RF were employed independently to model each processing stage sequentially, error propagation was observed to commence relatively at the first processing stage Twin Screw Granulation (TSG) and continued to escalate through to the Tablet Press (TP) stage as indicated by the increasing NRMSE magnitudes. This pattern suggests that certain unit operations introduce more variability in intermediate material properties, which becomes more challenging for the models to capture. This propagation effect was particularly pronounced in the TP stage, where prediction errors impacted the accurate predictions of the tablet tensile strength. In contrast, substantial improvements were observed when GMMs were integrated with both algorithms. The GMM versions of each model tend to exhibit lower NRMSE values than the corresponding individual implementations, indicating improved predictive accuracy. For example, Fig. 15(a) illustrates the error propagation trends associated with the GBM-GMM integrated model, where a decline in NRMSE was evident from the third stage onwards. Similarly, the RF-GMM based model has showed a outstanding improvement in the overall prediction as shown in Fig. 15(b), where the TP stage exhibited an overall error reduction of 27%. Furthermore, the model performance plots for tablet tensile strength prediction, presented in Fig. 16, illustrate the impact of GMM integration on both GBM and RF models. For GBM, integrating GMM resulted in an increase in R^2 to 0.86, along with an RMSE of 0.20 MPa, indicating that the majority of predictions fall within the 10% error band. Similarly, the RF model exhibited notable improvements when augmented with GMM, achieving an R^2 value of 0.90 and an RMSE of 0.17 MPa, signifying closer alignment of predicted values with actual data while maintaining consistency within the 10% error bands. It is worth emphasising that while GBM and RF serve as the primary predictive models capturing complex, non-linear relationships among critical material-process parameters, and final quality attribute, GMMs were specifically employed to model residual errors produced by these primary models. The integration of GMM complements the ML predictions by explicitly addressing prediction uncertainties through the probabilistic characterisation of residual errors. The fundamental rationale underpinning the effectiveness of GMM lies in its capacity to represent complex error distributions via multiple Gaussian components, systematically accounting for structured errors and reducing predictive uncertainty across the CTL data. Hence, despite limitations in predictive performance noted in Figs. 11 and 13, the integrated framework effectively leverages the predictive strengths of ML models and the probabilistic error-handling capability of GMM. For instance, the RF-GMM based model



(a) Error Propagation Plot for GBM Models.



(b) Error Propagation Plot for the RF models.

Fig. 15 The plots show error propagation through process stages with GBM and RF alone (blue squares) and when GMMs is integrated (yellow diamonds)

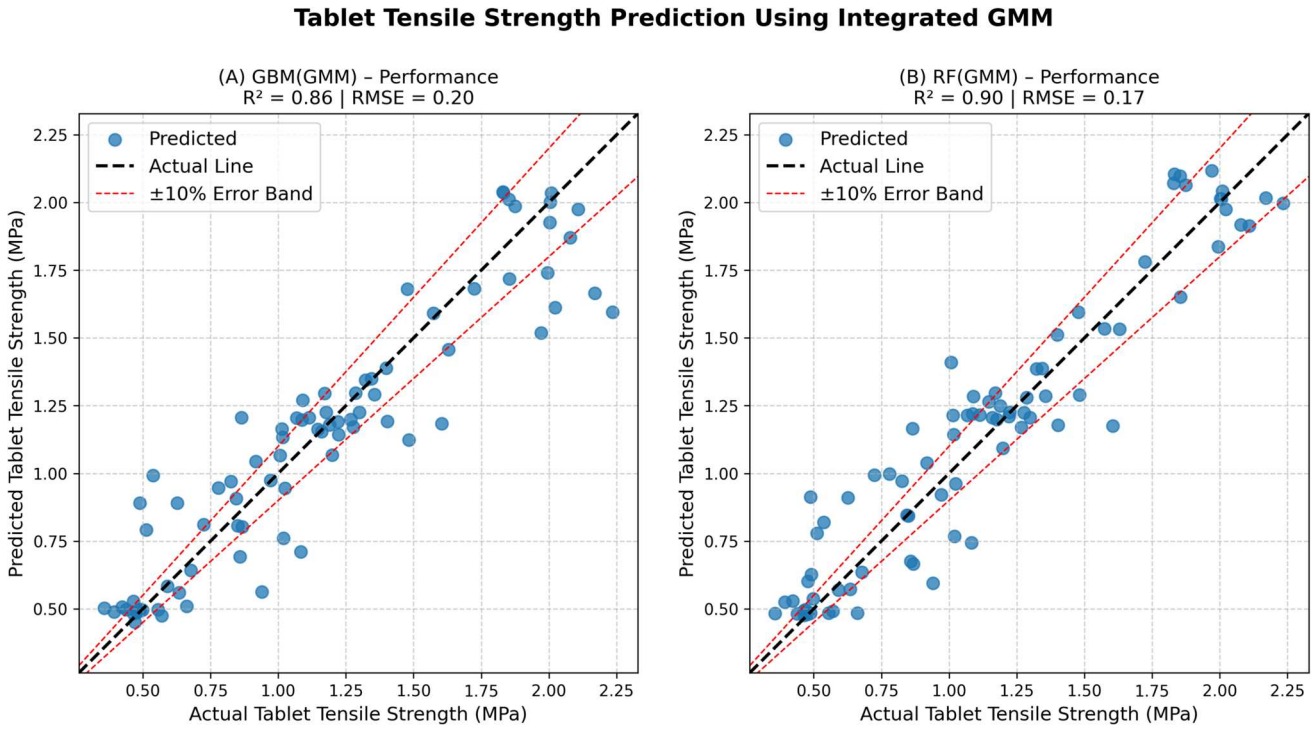


Fig. 16 The performance of GBM and RF models on unseen accumulated process stages data after the integration of error modelling approach using GMMs

performance in Fig. 16 demonstrates reduced prediction dispersion and adheres more closely to the actual tablet tensile strength line, compared to the RF model predictions illustrated in Fig. 13. This is because the GMM refines the model predictions at each process stage to improve the final prediction of the final tensile strength. Similarly, the GBM-GMM model followed a similar performance in mitigating the uncertainty presented by the primary model, aiming to improve the overall predictions. Nonetheless, deviations beyond the error bands within both integrated models persist. This may be attributed to the uncertainties resulting from process variability and complex process interactions, which could not be fully captured or explained by either model. Furthermore, limitations inherent in the ML model approaches constrain their ability to mitigate such uncertainties without further refinement, expanded training datasets, or more advanced modelling strategies. The overall enhancement in predictive accuracy underscores the effectiveness of GMMs in refining model predictions and mitigating error propagation across sequential stages, implying that GMM integration reduces predictive uncertainty across the entire manufacturing chain. Incorporating GMM to model errors between process stages reveals hidden patterns, thus leading to lower overall error propagation and more consistent predictions of end-product quality.

Conclusion

Advancements in modelling the continuous manufacturing of pharmaceutical tablets have been presented in this research. A comprehensive study was conducted on a pilot plant manufacturing line, to show how process parameters systematically influence each processing stage and the final product, highlighting the complexity of multivariate systems. This resulted in a representative dataset that facilitated the development of a plant-wide modelling framework in which all the utilised processing stages of the continuous tabletting line were sequentially integrated to form a plant-wide model from powder to tablet using machine learning techniques. Due to the complex behaviour of multivariate processing systems and their inherent variability, this sequential framework integrated GMMs to capture the variations between processing stages, enhancing model predictive performance. As a result, remarkable improvements in predictive accuracy were achieved, with R^2 value of 0.90 using the Random Forest model. The findings provide a powerful tool for implementing Quality by Design (QbD) in the pharmaceutical industry, revealing intricate process patterns and dependencies. This modelling framework for the continuous manufacturing of pharmaceutical tablets underscores the necessity of incorporating all critical processing

stages that sequentially transform powder into tablets. The inclusion of these interconnected stages enables the tracking and modelling of process variations that may arise between process stages. Consequently, this comprehensive approach facilitates the development of predictive tool that functions as an experimentation platform enabling systematic exploration and interpretation of the operational design space without the need for physical operation. While the developed plant-wide modelling framework has demonstrated promising results, certain limitations should be acknowledged. For example, the developed modelling framework focused on operational-level variability under fixed raw material, yet the framework could be extended to incorporate raw material attributes for broader applicability. The pilot-plant dataset, though comprehensive, may not encompass the full range of possible process variations. Although the integration of GMMs has improved the primary ML models predictive capabilities and reduced error propagation across stages, some degree of residual uncertainty remains. Hence, future research will focus on the utilisation of more advanced modelling algorithms particularly those that offer enhanced learning capabilities, inherently address uncertainty, and facilitate interpretability such as Fuzzy Inference Systems. These advancements are expected to contribute towards the development of holistic models capable of identifying complex patterns and providing explanations of the multivariate process interactions within the continuous tablet manufacturing. In conclusion, this study lays the groundwork for more promising predictive modelling in the continuous manufacturing of pharmaceutical tablets, offering a path toward optimised and controlled pharmaceutical manufacturing processes in line with a 'right-first-time' concept.

Acknowledgements All authors express their gratitude to DFE Pharmaceuticals for supplying the alpha-lactose monohydrate powder (Pharmatose 200M) and the microcrystalline cellulose powder (Pharmacel 101) used in this study. The generous provision of these essential pharmaceutical materials by DFE Pharmaceuticals was crucial for conducting the experimental work and collecting the data that underpinned this research. All authors also wish to acknowledge the insightful comments provided by all reviewers.

Author Contributions CRediT Author Statement: M.D.: Conceptualization, Methodology, Investigation, Formal analysis, Visualisation, Data curation, Writing-Original Draft Preparation. M.M.: Supervision, Conceptualization, Validation, Resources, Writing - Review & Editing. C. O.: Resources, Validation, Formal analysis, Writing - Review & Editing. S.I: Resources, Validation, Writing -Review & Editing. B. M: Resources, Validation, Writing - Review & Editing.

Funding The authors did not receive support from any organization for the submitted work.

Data Availability Data will be made available from the corresponding author upon request.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, Moore CM, Yu LX, Woodcock J. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10:191–9.
2. Domokos A, Nagy B, Szilágyi B, Marosi G, Nagy ZK. Integrated continuous pharmaceutical technologies—a review. Organic Process Res Develop. 2021;25(4):721–39.
3. Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, Woodcock J. Understanding pharmaceutical quality by design. AAPS J. 2014;16:771–83.
4. ICH. Continuous manufacturing of drug substances and drug products q13. 2021.
5. Destro F, Barolo M. A review on the modernization of pharmaceutical development and manufacturing – trends, perspectives, and the role of mathematical modeling. Int J Pharm. 2022;620:5.
6. Wahlich J. Review: Continuous manufacturing of small molecule solid oral dosage forms. Pharmaceutics. 2021;13(8).
7. FDA Center for Drug Evaluation and Research. Office of pharmaceutical quality: 2020 annual report. 2021. Available at https://www.fda.gov/media/145830/download. Accessed 29 Apr 2025.
8. BioPharm International. Barriers to the development of continuous manufacturing and ich q13. PDA–FDA Joint Regulatory Conference, 18 Sept. 2023. 2023. Available at: https://www.biopharminternational.com/view/barriers-to-the-development-of-continuous-manufacturing-and-ich-q13-pda-fda-joint-regulatory-conference-2023-. Accessed 29 Apr 2025.
9. Rogers AJ, Hashemi A, Ierapetritou MG. Modeling of particulate processes for the continuous manufacture of solid-based pharmaceutical dosage forms. Processes. 2013;1(2):67–127.
10. Vanhoorne V, Vervaet C. Recent progress in continuous manufacturing of oral solid dosage forms. Int J Pharm. 2020;579:4.
11. Rogers A, Ierapetritou M. Challenges and opportunities in modeling pharmaceutical manufacturing processes. Comput Chem Eng. 2015;81(10):32–9.
12. Khinast J, Bresciani M. Continuous manufacturing: Definitions and engineering principles. Contin Manufac Pharmaceut. 2017;1–31.
13. Shi G, Lin L, Liu Y, Chen G, Luo Y, Wu Y, Li H. Pharmaceutical application of multivariate modelling techniques: a review on the manufacturing of tablets. RSC Adv. 2021;11:8323–45.

14. Solle D, Hitzmann B, Herwig C, Pereira Remelhe M, Ulonska S, Wuerth L, Prata A, Steckenreiter T. Between the poles of data-driven and mechanistic modeling for process operation. Chemie Ingenieur Technik. 2017;89(5):542–61.
15. Boukouvala F, Niotis V, Ramachandran R, Muzzio FJ, Ierapetritou MG. An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. Comput Chem Eng. 2012;42(1):30–47. European Symposium of Computer Aided Process Engineering - 21.
16. Boukouvala F, Chaudhury A, Sen M, Zhou R, Mioduszewski L, Ierapetritou MG, Ramachandran R. Computer-aided flowsheet simulation of a pharmaceutical tablet manufacturing process incorporating wet granulation. J Pharm Innov. 2013;8:11–27.
17. Wang Z, Pan Z, He D, Shi J, Sun S, Hou Y. Simulation modeling of a pharmaceutical tablet manufacturing process via wet granulation. Complexity. 2019;2019(1):3659309.
18. Metta N, Ghijs M, Schäfer E, Kumar A, Cappuyns P, Van Assche I, Singh R, Ramachandran R, De Beer T, Ierapetritou M, Nopens I. Dynamic flowsheet model development and sensitivity analysis of a continuous pharmaceutical tablet manufacturing process using the wet granulation route. Processes. 2019;7(4).
19. Dong Y, Yang T, Xing Y, Du J, Meng Q. Data-driven modeling methods and techniques for pharmaceutical processes. Processes. 2023;11(7).
20. Boukouvala F, Muzzio FJ, Ierapetritou MG. Dynamic data-driven modeling of pharmaceutical processes. Indust Eng Chem Res. 2011;50(6):6743–54.
21. Wafa'H A, Mahfouf M, Salman AD. When swarm meets fuzzy logic: Batch optimisation for the production of pharmaceuticals. Powder Technol. 2021;379:174–83.
22. Kuhn M, Johnson K. Applied Predictive Modeling. Springer, 2013.
23. Deebes M, Mahfouf M, Omar C. A new data-driven modelling framework for moisture content prediction in continuous pharmaceutical tablet manufacturing. In: International Conference on Industrial Engineering and Applications. Springer; 2024. p. 107–20.
24. Mäki-Lohiluoma E, Säkkinen N, Palomäki M, Winberg O, Ta HX, Heikkinen T, Kiljunen E, Kauppinen A. Use of machine learning in prediction of granule particle size distribution and tablet tensile strength in commercial pharmaceutical manufacturing. Int J Pharm. 2021;609:121146.
25. Bekaert B, Van Snick B, Pandelaere K, Dhondt J, Di Pretoro G, De Beer T, Vervaet C, Vanhoorne V. Continuous direct compression: Development of an empirical predictive model and challenges regarding pat implementation. Int J Pharmaceut: X. 2022;4:100110.
26. Lou H, Lian B, Hageman MJ. Applications of machine learning in solid oral dosage form development. J Pharm Sci. 2021;110(9):3150–65.
27. AlAlaween WH, Mahfouf M, Omar C, Al-Asady RB, Monaco D, Salman AD. Serial artificial neural networks characterized by gaussian mixture for the modelling of the consigma25 continuous manufacturing line. Powder Technol. 2024;434:119296.
28. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Frontiers in Neurorobotics. 2013;7.
29. Taguchi G, Chowdhury S, Wu Y. Introduction to orthogonal arrays. Taguchi's Quality Engineering Handbook. 2004;584–96.
30. Seem TC, Rowson NA, Ingram A, Huang Z, Yu S, de Matas M, Gabbott I, Reynolds GK. Twin screw granulation - a literature review. Powder Technology. 2015;276(5):89–102.
31. Luypaert J, Massart D, Vander Heyden Y. Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta. 2007;72(3):865–83.
32. Pitt K, Newton J, Stanley P. Tensile fracture of doubly-convex cylindrical discs under diametral loading. Journal of Materials Science. 1988;23.
33. Monaco D, Reynolds GK, Tajarobi P, Litster JD, Salman AD. Modelling the effect of l/s ratio and granule moisture content on the compaction properties in continuous manufacturing. Int J Pharm. 2023;633:2.
34. Merkus HG. Particle Size Measurements: Fundamentals, Practice, Quality, vol. 17. Springer; 2009.
35. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction, vol. 2. Springer; 2009.
36. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
37. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. 2001;29(5):1189–232.
38. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
39. Bishop CM, Nasrabadi NM. Pattern recognition and machine learning, vol. 4. Springer; 2006.
40. Litster J. Design and Processing of Particulate Products. Cambridge University Press; 2016.
41. Yang YY, Mahfouf M, Panoutsos G. Probabilistic characterisation of model error using gaussian mixture model–with application to charpy impact energy prediction for alloy steel. Control Engineering Practice. 2012;20(1):82–92. Special Section: IFAC Conference on Analysis and Design of Hybrid Systems (ADHS'09) in Zaragoza, Spain, 16th-18th September, 2009.
42. McLachlan GJ, Krishnan T. The EM algorithm and extensions. John Wiley & Sons; 2007.
43. Zainuddin Z, Lim EA. A comparative study of missing value estimation methods: Which method performs better?. In: 2008 International Conference on Electronic Design. 2008. p. 1–5.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.