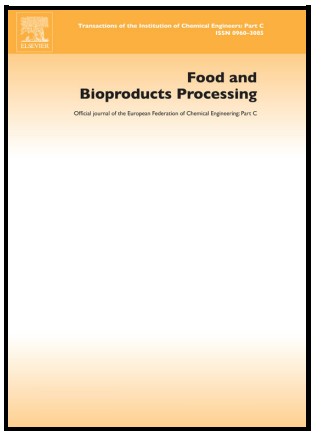



# Lactose Crystallization: Integrating Machine Learning with Process Analytical Technologies

Cha Yong Jong, Akshay Mittal, Felix Lee Jun Jie, Lee May Loo, Goh Yongkai, Qiaolin Yuan, Eunice Yeap Wan Qi, Srinivas Reddy Dubbaka, Harsha Nagesh Rao, Wong Shin Yee

PII: S0960-3085(25)00029-X

DOI: https://doi.org/10.1016/j.fbp.2025.02.008

Reference: FBP1989

To appear in: *Food and Bioproducts Processing*

Received date: 3 September 2024
Revised date: 6 February 2025
Accepted date: 25 February 2025

Please cite this article as: Cha Yong Jong, Akshay Mittal, Felix Lee Jun Jie, Lee May Loo, Goh Yongkai, Qiaolin Yuan, Eunice Yeap Wan Qi, Srinivas Reddy Dubbaka, Harsha Nagesh Rao and Wong Shin Yee, Lactose Crystallization: Integrating Machine Learning with Process Analytical Technologies, *Food and Bioproducts Processing,* (2025) doi:https://doi.org/10.1016/j.fbp.2025.02.008

# Lactose Crystallization: Integrating Machine Learning with Process Analytical Technologies

*Cha Yong Jong,* [†] *Akshay Mittal,* [†] *Felix Lee Jun Jie,* [†] *Lee May Loo,* [†] *Goh Yongkai,* [§] *Qiaolin Yuan,* [§] *Eunice Yeap Wan Qi,* [‡] *Srinivas Reddy Dubbaka,* [‡] *Harsha Nagesh Rao,* [‡] *Wong Shin Yee* [†*].

[†] Singapore Institute of Technology, 10 Dover Drive, 138683, Singapore

[§]Applied Materials South East Asia Pte. Ltd., 8 Upper Changi Road North, Applied Materials Building, 506906, Singapore.

[‡] Pfizer Asia Manufacturing Pte Ltd., Manufacturing Technology Development Centre (MTDC), 3 Biopolis Drive, 138623, Singapore.

**KEYWORDS** Crystallization, lactose recovery, machine learning, PAT, optimization

## Abstract

Lactose is recovered from whey through crystallization process, where a concentrated supersaturated solution is cooled to crystallize the lactose, leaving the impurities in the mother liquor. Designing this process requires considerations over various parameters, particularly the concentration of the feed solution and the cooling profile. To optimize the parameters, most developers depend on trial-and-error methods, a manageable task for the experienced but challenging for novices. This study presents a novel system that leverages machine learning (ML) and process analytical technologies (PAT) to streamline lactose crystallization process development, going beyond manual trial and error interpretations. The automated system initiated with Direct Chord Length (DCL) feedback control run, which provided the

foundational data for the ML model, which was then employed in subsequent AN1 and AN2 iterative runs. These iterative runs have smoother concentration and temperature curves, and it generates larger crystal with enhanced productivity and yield. The results indicate that the ML-driven approach can significantly outperform conventional methods, enabling the precise control of nucleation and growth phases to produce larger lactose crystals.

# 1. Introduction

Lactose is a disaccharide, composed of glucose and galactose. It is a primary sugar in human and mammalian milk, serving as an energy source, and is also used in infant formulas, pharmaceutical tablets, inhalers, and processed foods (Bier et al., 2023; Halfwerk et al., 2021). Industrially, lactose is typically extracted from whey, a byproduct of cheese and casein production, using crystallization.

Crystallization is a purification process, where a concentrated whey solution, typically containing around 60% dissolved lactose (Durham, 2009), is cooled to produce lactose crystals (Evdokimov et al., 2021). This process is pivotal in defining the quality of lactose products, with yield, purity, and particle size as the benchmark indices (Hourigan et al., 2013; Zeng et al., 2010). The optimization of the crystallization process hinges on critical parameters, such as the initial concentration of the whey solution and the specific cooling rate employed, both of which have been extensively studied (Raghavan et al., 2001; Shi et al., 2006). Adjustments in these parameters are known to influence the supersaturation profile, which in turn dictates the nucleation rate and crystal growth dynamics, thereby affecting the crystal size distribution and purity levels (Carpin et al., 2017; Wong & Hartel, 2014a). The meticulous control of these variables is essential for tailoring the crystallization process to yield lactose crystals that meet

the stringent specifications required for various applications, from food products to pharmaceuticals (Durham, 2009; Halfwerk et al., 2021).

The generation of large lactose crystals significantly influences the quality and yield of lactose products (Ortega-Rivas, 2009) from several perspectives. In the context of downstream filtration and centrifugation, larger crystals, typically within the 200-300 μm range, facilitate a more efficient operation. Their size allows for easier separation, resulting in faster separation rates and minimized loss of small particles into the mother liquor, thereby ensuring higher recovery and operational efficiency (Ortega-Rivas, 2009). Uniform and large crystals exhibit a lower degree of mother liquor entrapment. This reduction lessens the likelihood of impurity occlusion within the crystal lattice or on the crystal surface (Carpin et al., 2017), simplifying the washing process. Additionally, uniformly large and well-structured crystals are favoured because they facilitate easier moisture removal (Faria et al., 2003). From a product perspective, large crystals are less prone to moisture sorption and caking, unlike finer particles (< 80 μm) which, due to their increased specific surface area and impurity content, are more susceptible to humidity-induced caking (Carpin et al., 2017; Modugno et al., 2015). Therefore, advancements in crystallization technology that enable the reliable production of large, uniform lactose crystals is highly desirable.

However, in most industrial lactose crystallization processes, smaller crystals (< 100 μm) are typically formed due to secondary nucleation (Paterson, 2009; Wong et al., 2012), leading to inefficient downstream separation and lower recovery ratios (Wong & Hartel, 2014a). Current approaches to optimizing lactose crystallization are summarized in Table 1. Methods such as cooling (Wong et al., 2012), antisolvent (Bund & Pandit, 2007), and ultrasound-assisted crystallizations (Khaire & Gogate, 2021) have been employed, with most resulting in relatively small crystal sizes, such as the 63-90 μm range with ultrasound (Dhumal et al., 2008) and up

to 333 µm using anti-solvent methods (Simone et al., 2019). While mathematical modelling has provided a more systematic approach to predicting and controlling crystal size, a significant portion of the process development still relies on experimental trial and error (Halfwerk et al., 2021, 2023). These methods produce crystals that are smaller than the desired 200-300 µm size preferred for industrial applications, indicating a continued need for innovation and refinement in lactose crystallization techniques.

Over the past few decades, the food industry has increasingly embraced Artificial Intelligence (AI) to meet the growing global food demands driven by the rising world population (Addanki et al., 2022). Some of the reported applications include determining food quality (Goyache et al., 2001; Hamid et al., 2018), serving as control tools in food processing (Funes et al., 2015), classifying food types (Krupitzer & Stein, 2021; Tsakanikas et al., 2020), and making predictions through sensory imaging in crop's cultivation (Kakani et al., 2020). Typically, these applications reported a significant improvement in process output over conventional practices (Mavani et al., 2022). For crystallization process development of food materials, the use of AI is still relatively limited. Majority of the related research effort employed sophisticated computational tools and AI models to enhance the efficiency and understanding of sucrose crystallization processes (Chernyaeva et al., 2020). In the field of sucrose crystallization, advancements typically involve the development of sensors for accurate measurement of essential parameters (Meng et al., 2019). This is followed by the adoption of AI models to enhance the crystallization process. Specifically, the use of deep convolutional neural networks (DCNNs) (Zhang et al., 2020), nonlinear Model Predictive Control (MPC) (Paz Suárez et al., 2011), and/or artificial neural networks (Naik et al., 2019). The outcome of these implementations typically trends towards achieving automated quality control and higher process efficiency. These technologies represent a shift towards minimizing human

intervention in the crystallization process, demonstrating how AI and computational methods can significantly improve both the efficiency and quality of sugar production.

This study goes beyond the implementation of AI models to monitor or optimize the crystallization process. It demonstrates the application of an automated crystallization platform (Jong et al., 2024) on lactose crystallization process development. This is an exploration of employing ML techniques to optimize a lactose crystallization process automatically, exploring its potential to replace conventional trial-and-error methodologies. Using the automated platform, the lactose crystallization process was optimized using an Adaptive Neuro-Fuzzy Inference System (ANFIS) driven algorithms (Jong et al., 2024). The platform operates under a modified feedback control framework, utilizing real-time data from PAT tools to control the crystallization process. Initial experiments, based on literature-derived cooling profiles, set the stage for understanding the basic crystallization behaviour of lactose. Subsequent iterations utilize the accrued data to intelligently adjust the process, resulting in an optimized approach for larger crystal production.

**Table 1:** Overview of various techniques in optimization of lactose crystallization. ($V_1$- Starting solution concentration, $V_2$-pH, $V_3$ - residence time, $V_4$ - seeding amount/temperature, $V_5$ - cooling rate, $V_6$ - rotation speed, $V_7$ - antisolvent concentration/addition rate, $V_8$ - sonication application frequency/duration)

| | Mode/method/Material | Objective | Optimization approach | Max size obtained | reference |
|---|---|---|---|---|---|
| 1 | Cooling (Lactose Recovery from Whey)<br><br>Laboratory scale: purified lactose<br>Plant scale: whey permeates concentrate. | Promote production of large crystals to avoid downstream filtration losses. | Mathematical modelling with appropriate crystallizer design (minimize shear rate) to control the process within specific zone in the metastable zone width to promote growth and minimize secondary nucleation | $L_{50} \approx 360\mu m$<br><br>Typical industrial process $\approx 250\mu m$ | (Wong et al., 2012) |
| 2 | Antisolvent (ethanol) assisted by ultrasound<br><br>Laboratory scale: analytical grade lactose and bovine serum albumin as protein source to mimic the constituent of whey solution | To develop an economical method to recover lactose from whey solution with low (13.5%w/v & 17.5%w/v) lactose concentration | Experiments with varying tuning parameters $V_1$, $V_2$, $V_8$ | Average diameter $12.97 \pm 5.86$ µm. | (Bund & Pandit, 2007) |

| | | | | | |
|---|---|---|---|---|---|
| 3 | Seeded cooling and anti-solvent (acetone) crystallization (laboratory scale, extra-pure alpha lactose monohydrate) | To develop more efficient design of crystallization to recover lactose from whey permeate | Experiment study on varying tuning parameters $V_1$, $V_4$, $V_5$, $V_6$, $V_7$ | Seeded cooling crystallization: Volume mean diameter (VMD) = 149µm.<br><br>Anti-solvent crystallization VMD = 333µm | (Simone et al., 2019) |
| 4 | Eutectic crystallization (laboratory scale using Delactose whey permeate) | To recover lactose from delactosed whey permeate (DLP) | Experiments and mathematical modelling | VMD = 23- 31µm. | (Halfwerk et al., 2023; Halfwerk et al. 2021) |
| 5 | Sonocrystallization | To recover lactose crystals with desirable size, shape, and surface for Dry Powder Inhalation (DPI) application | Using $L_9$-orthogonal array to design experiments to study varying parameters ($V_1$, $V_2$, $V_7$, $V_8$) | Size = 135.2 µm | (Patel & Murthy, 2011) |
| 6 | Swift cooling | Control the crystal size and morphology to produce good-quality lactose | Experiments with varying initial concentration of solute ($V_1$) | At supersaturation (S) of 6.99, the crystal size is 100 ± 1 µm | (Parimaladevi & Srinivasan, 2015) |

# 2. Materials and methods

## 2.1 Experimental setup

The experimental apparatus is shown in Figure 1. It comprises four principal components: (1) a 1L crystallization vessel (Diameter = 101mm, Height = 172 mm) with built-in electrical heating and solid-state cooling controls (METTLER TOLEDO®, OptiMax 1001 Synthesis Workstation); (2) the BLAZEMETRICS Blaze 900 probe with built-in High Dynamic Range microscopy that capture real-time images, and chord-length distributions are calculated based on the images; (3) the ReactIR probe (METTLER TOLEDO® ReactIR 15 with AgX 9.5 mm x 1.5 m fiber & Diamond probe tip) for solution concentration measurement; and (4) a desktop computer (Hewlett-Packard Z4 G4 computer with i9 64GB processor, 2TB SSD, NVIDIA Quadro RTX 4000) for data consolidation and control.

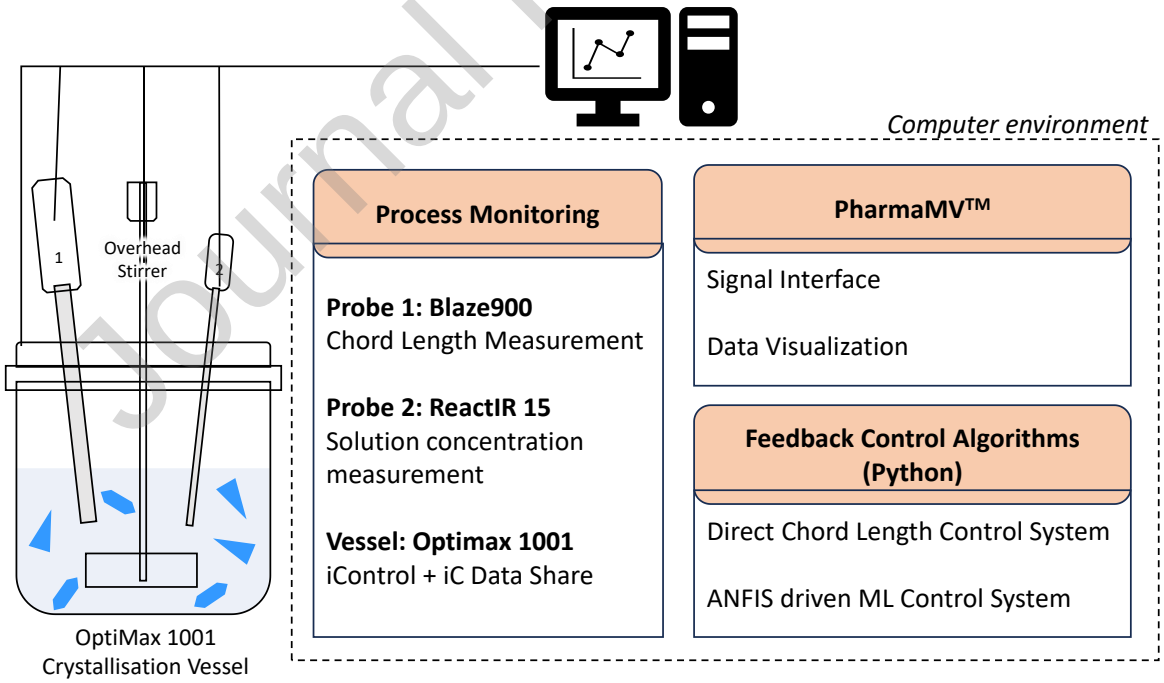



**Figure 1**: Schematic of experimental setup.

Physical connections among the hardware units are represented in Figure 1 by solid lines, signifying direct wiring. In contrast, the dotted lines illustrate the configuration of the software system. The signals procured from the probes are assimilated within the desktop computer via the PharmaMV™'s (Applied Materials, SmartFactory Rx™) platform. Utilizing these real-time data streams, the platform activates pertinent feedback control algorithms which is programmed in Python, thereafter, dispatching the requisite temperature setpoints (via PharmaMV™'s temperature profile tracking model predictive control function) to achieve the targeted process outcome. These signals are relayed back to the vessel's control interface (through PharmaMV™), thereby implementing the necessary control actions to steer the crystallization process.

## 2.2 Experimental preparations

Lactose monohydrate, FlowLac® 100 (Meggle GmbH & Co. KG, Germany), was used for all experiments. To accurately measure solution concentration in terms of lactose monohydrate per gram of solution, the React-IR needs to be calibrated to convert raw spectrum into solution concentration. The IR signal was calibrated at varying concentrations (0.63 – 0.1 g lactose/g solution) and temperature (5 - 95$^{o}$C). Out of all the signals collected, 85% of the data was (randomly) used for calibration and the remaining 15% for validation. The baseline of the IR spectrum was first corrected using Savitzky-Golay second derivative. A Partial Least Square (PLS) calibration model was developed in Matlab® software using the calibration data collected. The prediction error % is 1.17% and 1.18% for both training and validation datasets, respectively. Some of the crystals were isolated from the crystallizer using vacuum filtration. The particle size distribution of isolated lactose crystals was measured with Malvern Mastersizer 3000.

In the Optimax system, a four-blade pitch-blade turbine was used for mixing. The rotational speed for the stirrer was fixed at 150 rpm. This speed was chosen to ensure adequate mixing and suspension of lactose particles, which is critical for uniform nucleation and crystal growth (McLeod et al., 2016) . Previous research also indicates that the width of the Metastable Zone Width (MSZW) can vary with changes in agitation speed, typically narrowing as the rpm increases (Wong & Hartel, 2014b). The system presented in Figure 1 works at any user-defined rotational speed. The adaptive feedback system integrated within the automated crystallization platform independently evaluates and optimizes the crystallization process at the set rpm. Should there be a need to adjust the rpm, user can restart the automated loop, to collect additional process data for the re-optimization of the cooling profile to adapt to new operational conditions.

## 2.3 The automated ANFIS driven platform.

An overview of the workflow for the automated platform (Jong et al., 2024) is shown in Figure 2. The process initiates with the preparation of a supersaturated feed solution with known concentration. Users then configure the system settings, including critical parameters such as temperature boundaries and solubility equations. Additionally, users may select the option to isolate the crystallized product between iterative runs for analysis.

For this study, the feed solution was first dissolved at a starting temperature of 91-95°C. The dissolution process was monitored by the Blaze 900 probe. If the chord count (10 – 900 μm) is less than 100, dissolution is complete. Otherwise, the system will gradually increase the temperature until the upper bound of the temperature limit. Following the dissolution, the crystallization phase will start and be controlled until the specified lower temperature boundary.

For the cooling crystallization process, the automated platform can operate within a "closed loop" framework, as depicted in the blue cycle of Figure 2. The operational sequence during a

campaign is as follows: upon the conclusion of a crystallization experiment, the platform can either enter a holding state that permits the isolation of the product for analysis or other purposes, or it can automatically trigger the dissolution phase. This dissolution re-establishes a supersaturated environment, thereby preparing the system for the subsequent iteration. This design allows for continuous automated operation without manual intervention between cycles if the user opts not to isolate the crystallized product after each iteration.

Upon finalizing the settings, the user commences the crystallization campaign by activating the "Start Campaign" function. The initial phase of the campaign employs a DCL Control System (Jong et al., 2024), which monitors and adjusts the crystallization process in real-time. All real-time PAT results are displayed on the PharmaMV™ platform. The outcome of this phase is a primary ANFIS model v1 (Figure 2), which serves as the “brain” for subsequent ANFIS-driven iteration experiments.

In order to transition to the ANFIS-driven ML control system, the platform will execute more feedback-controlled experiments. The feedback control strategy is refined through iterative model generation and training, followed by consolidating real-time PAT data to enhance the precision of the crystallization process. With each iteration, the model (beginning with version 1 (from DCL’s data) and advancing to versions 2, 3, and so on) evolves due to the accumulated data from prior runs, which then develops into an optimized temperature profile for lactose crystallization.

At the end of each crystallization experiment, the platform offers an optional step for the isolation of the product, which can be executed through filtration or other suitable separation techniques. At the end of the iterative cycle, users can review the profile and retrieve the most optimized temperature profiles from all the iterations.

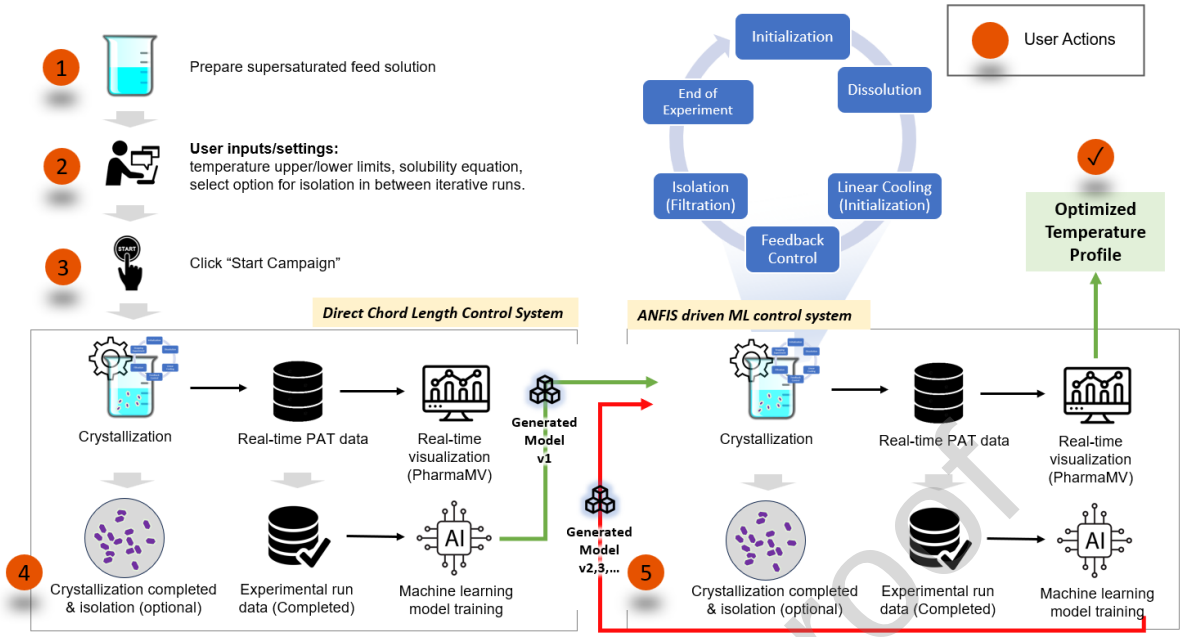



**Figure 2**: Workflow for the automated crystallization platform

# 3. Results & Discussion

## 3.1 Baseline experiments

Baseline experiments (using approximately 0.6 g lactose monohydrate/g solution) were carried out to evaluate the performance of the setup (Figure 1) against cooling profiles reported in the literature (Wong et al., 2012). Dissolution processes were carried out at 95°C and then the temperature was reduced to 77°C prior to the crystallization process. The cooling profiles utilized, as shown in Figure 3, include SLOWC, MEDIUMC, and FASTC, with cooling rates of -0.031 °C/min, -0.055 °C/min, and -0.43 °C/min, respectively. The corresponding concentration profiles and crystal size distributions are included in Figure 3 and Figure 4. In Figure 3, the supersolubility line, delineates the boundary between supersaturated metastable zone and the labile zone, marking the condition under which spontaneous nucleation is likely to be initiated (in labile zone). These supersaturation and nucleation thresholds are usually system specific, and it is influenced by specific characteristics (design, scales etc) and the hydrodynamic conditions (mixing, heat transfer etc) within the crystallizer.

The cooling profiles were extracted from existing literature to our system directly, maintaining the exact cooling rates described (Figure 3). However, the concentration profiles obtained differed significantly between the two setups. In Figure 3, the dotted lines represent the concentration profiles reported in literature, while the solid lines are derived from ReactIR measurements from the system described in Figure 1. This discrepancy underscores the challenges in directly transferring optimized crystallization conditions from one system to another, highlighting that outcomes can vary significantly even when identical parameters are used. Specifically, when the FASTC profile is implemented, literature source indicate that the operation occurs within the labile zone—an area that typically results in undesirable crystal properties due to rapid nucleation and growth rates. Conversely, using the optimax system, the FASTC profile largely remained within the MSZW, suggesting a more controlled crystallization process and highlighting the potential deviations in system behavior under seemingly identical conditions. This observation illustrates the need for careful adaptation and recalibration of crystallization parameters when transitioning between different systems to ensure desired outcomes.

Compared to other two profiles, the SLOWC profile is less smooth with occasional increases in lactose concentration, despite a consistent drop in the temperature. This might be attributed to a slower cooling rate that leads to a delayed nucleation process, causing lactose to remain dissolved longer. This delay facilitates the dissolution of smaller, less stable crystals back into the solution, thereby increasing the concentration. As these smaller crystals dissolve, their solutes potentially contribute to the growth of larger, more stable crystals. However, if this redistribution of solute is not immediate, a transient increase in lactose concentration occurs until these molecules are eventually absorbed into the growth of larger crystals.

The SLOWC and MEDIUMC profiles, with its gradual temperature decrease, operates in the upper MSZW, closest to the solubility line (Figure 3). The FASTC profile, marked by its rapid

temperature fall, operates near the lower MSZW, with limited deviation into the labile zone (Figure 3). Despite the distinctly different cooling profiles (SLOWC, MEDIUMC, and FASTC, Figure 3), the crystal size distributions do not exhibit substantial differences (Figure 4). This could imply that the nucleation rate and crystal growth mechanisms are not overly sensitive to the variations in cooling rates within the range of conditions tested. However, the yield of crystals varies significantly, with the FASTC profile resulting in the lowest yield (Table 2). The rate of nucleation is expected to increase with supersaturation (Wong & Hartel, 2014a), which is typically higher in faster cooling rates like those seen in the FASTC profile. With a higher nucleation rate, there might also be secondary nucleation, that generates a large number of small crystals that dissolve more readily or are difficult to recover, thus lowering the overall yield. In contrast, the slower cooling rates in SLOWC may allow for more controlled nucleation and growth phases, giving the crystals sufficient time to reach a stable size that contributes to a higher yield.

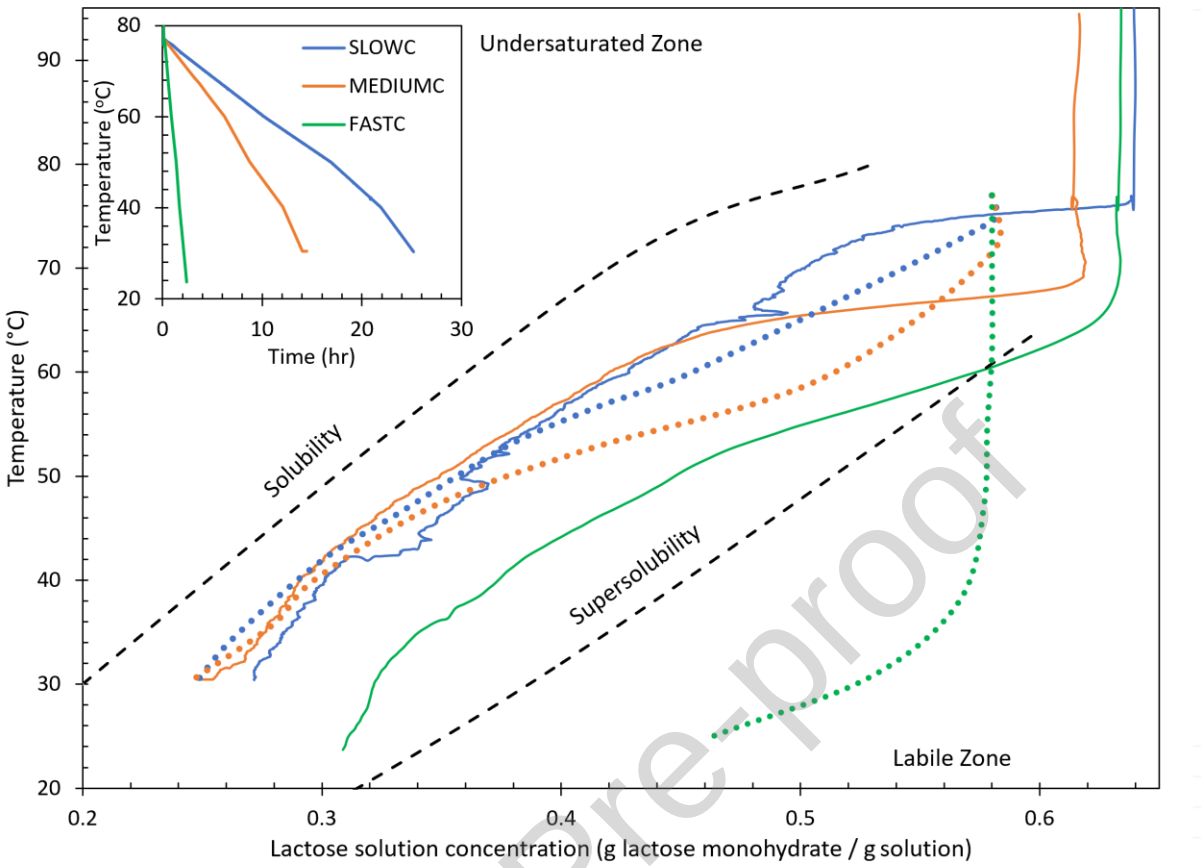



**Figure 3** Experimental concentration profile (dotted line: literature concentration profile (Wong et al., 2012), solid line: this work) Insert: baseline temperature profiles extracted from literature (Wong et al., 2012)

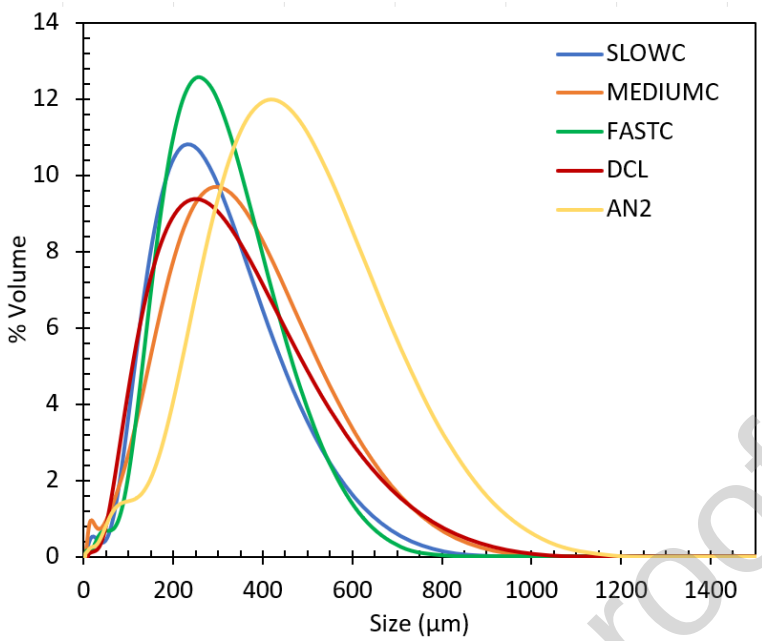



**Figure 4**: (Offline) Crystal size distributions measured using laser diffraction method

### 3.2 Automated Crystallization Platform

Moving beyond conventional practices, the automated crystallization platform employs a ML ANFIS and initiates with a DCL feedback control strategy.

Starting with DCL feedback control, the platform used fundamental crystallization practices (slow cooling for crystal growth, fast heating for fines dissolution) to generate training data for the ML model. The first ML iterative run, AN1, exclusively uses data from DCL. AN2 merges data from DCL and AN1, leading to iterative enhancements in the process. Each subsequent AN iteration will adjust processing conditions by building on existing data, aiming to progressively improve the crystal size. This method not only makes the crystallization kinetics optimization process more efficient but also navigates the complex interplay between cooling rates, supersaturation levels, and crystal growth dynamics, thereby yielding larger crystals. The outputs from the platform are summarized in Figure 5, Figure 6 and Table 2.

The first iterative run, represented as DCL, produced significant fluctuations in both concentration (Figure 5) and temperature (Figure 6) profiles. The DCL algorithm reacts to real-time monitoring data based on pre-determined heating/cooling actions (Jong et al., 2024). These fluctuations are reactive in nature. As the process progressed, all deviation from the ideal state (crystal growth) will be corrected, resulting in frequent "corrections/fluctuations". After the DCL initiation, the ANFIS model was trained with DCL data, subsequent AN1 and AN2 runs exhibited a notable stabilization with smoother concentration and temperature curves, indicating an improved control over the crystallization process.

The performance enhancements are quantitatively supported by Table 2, which shows a consistent improvement in productivity, defined as yield (%) over total time. Learning from the data obtained from DCL and AN1, the AN2 iteration managed to produce comparable crystals in approximately half the time, illustrating the ML model's growing efficacy. From a crystallization kinetics perspective, the smoother concentration profiles in AN1 and AN2 runs (Figure 5) suggest an optimized navigation through the metastable zone, where conditions are ideal for growth with minimum secondary nucleation nor aggregation. Compared to the output from the manual runs (SLOWC. MEDIUMC, FASTC), the yield has improved, along with the production of larger crystals.

Real time microscopic imaging and offline imaging of isolated products are shown in Figure 7 and Figure 8, respectively. Despite the high suspension density complicating direct visualization for real-time monitoring in Figure 7, the images provide a glimpse of the crystal morphology at the end of the crystallization process. Offline microscopic images of isolated crystals in Figure 8 show that, among all crystals produced, those from AN2 are the largest. This aligns well with the particle size distribution data from the Mastersizer shown in Figure 4.

The data indicates that the application of ML to the lactose crystallization process can substantially enhance the control over crystallization kinetics, resulting in an efficient and improved production of lactose crystals. Our study demonstrates the potential of ML algorithms to supersede traditional trial-and-error methods, paving the way for more sophisticated and precise crystallization strategies in the industry.

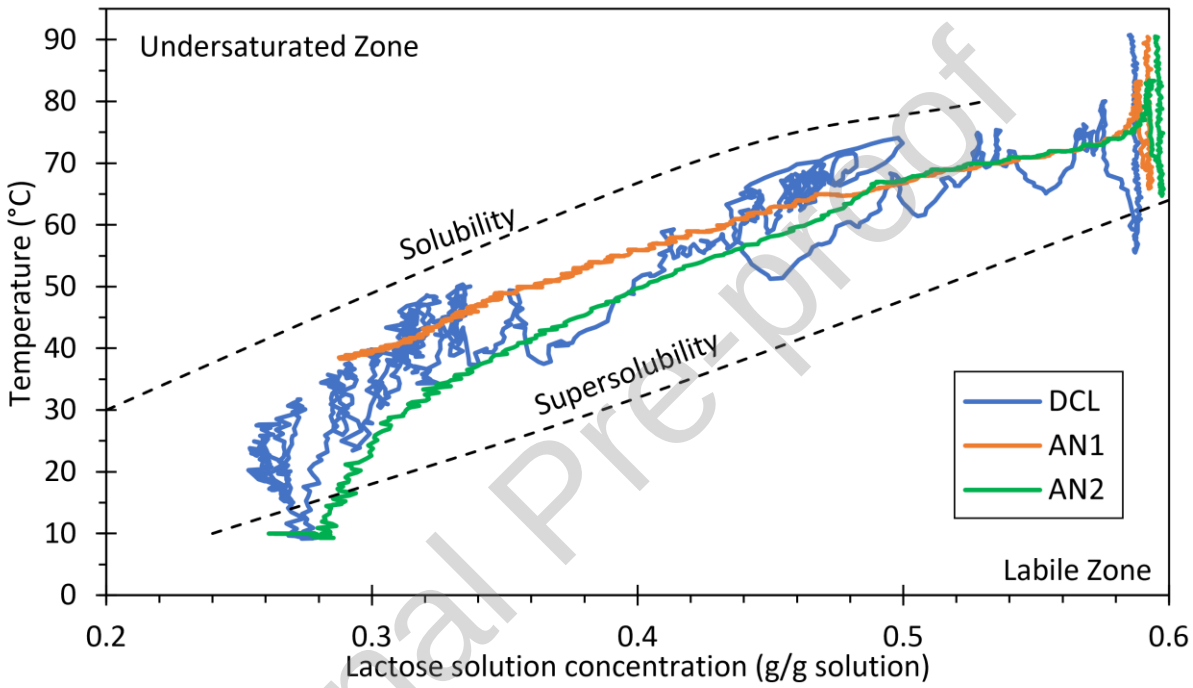



**Figure 5**: Concentration profile for crystallization operation with different iteration (from DCL to AN2)

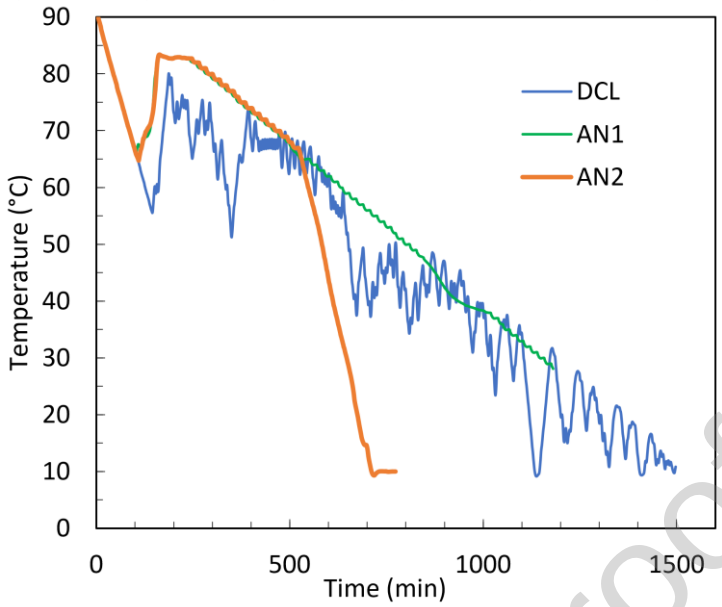
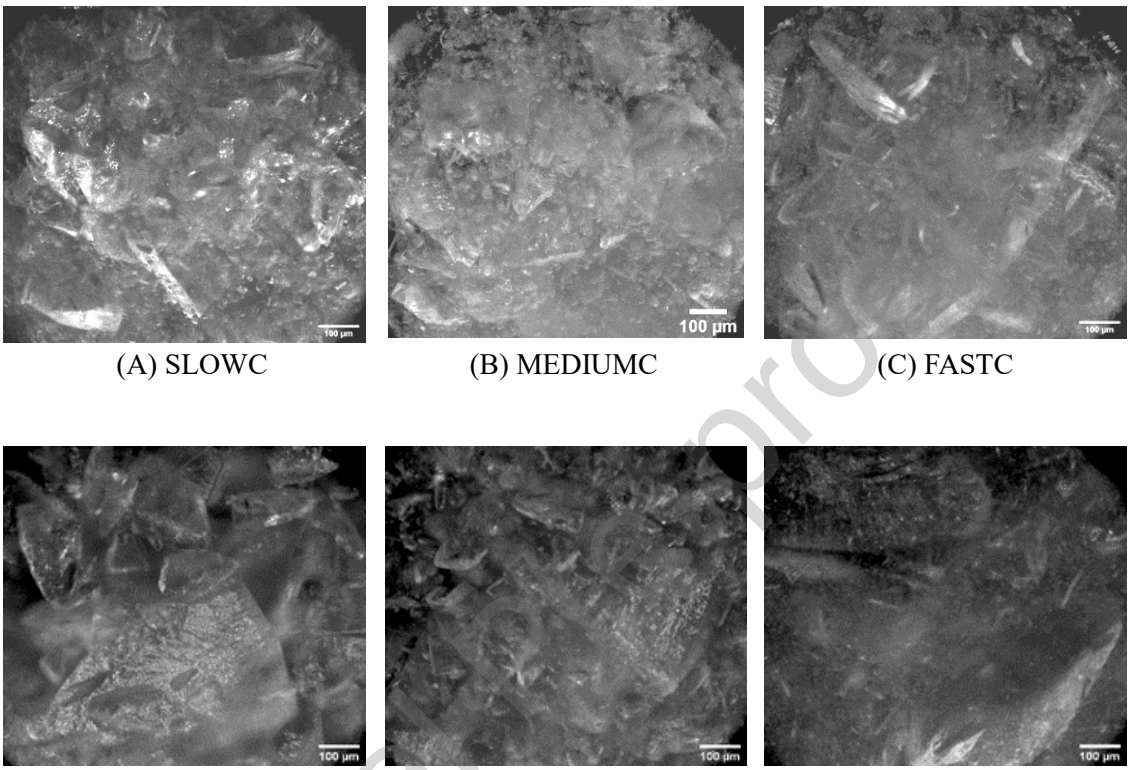



**Figure 6**: Temperature profile obtained from the automated crystallization platform at different iterative runs (from DCL to AN2)

| | Automated | | | Manual | | |
|---|---|---|---|---|---|---|
| | DCL | AN1 | AN2 | SLOWC | MEDIUMC | FASTC |
| Feed Concentration (g/g solution) | 0.59 | 0.59 | 0.59 | 0.64 | 0.62 | 0.63 |
| Total time (hr) | 25.62 | 24.58 | 12.88 | 25.18 | 14.45 | 2.43 |
| Yield (%) | 77.96 | 81.63 | 74.89 | 57.86 | 59.81 | 52.14 |
| $D_{10}$ (µm) | 12.85 | 11.34 | 10.06 | 11.02 | 8.81 | 11.85 |
| $D_{50}$ (µm) | 36.41 | 28.91 | 31.20 | 26.34 | 23.91 | 30.90 |
| $D_{90}$ (µm) | 87.02 | 54.41 | 62.09 | 48.22 | 43.52 | 56.54 |
| Productivity (%/min) | 0.05 | 0.06 | 0.10 | 0.04 | 0.07 | 0.36 |

**Table 2**: Process output for experiments conducted using the crystallization platform. ($D_{10}$, $D_{50}$, $D_{90}$: the $10^{th}$, $50^{th}$ and $90^{th}$ percentile size, meaning 10/50/90 % of the particles in the sample are smaller than the specific size)




(A) SLOWC (B) MEDIUMC (C) FASTC

(D) DCL (E) AN1 (F) AN2

**Figure 7**: Real-time microscopic images collected using the Blaze probe inserted directly into the crystallizer.

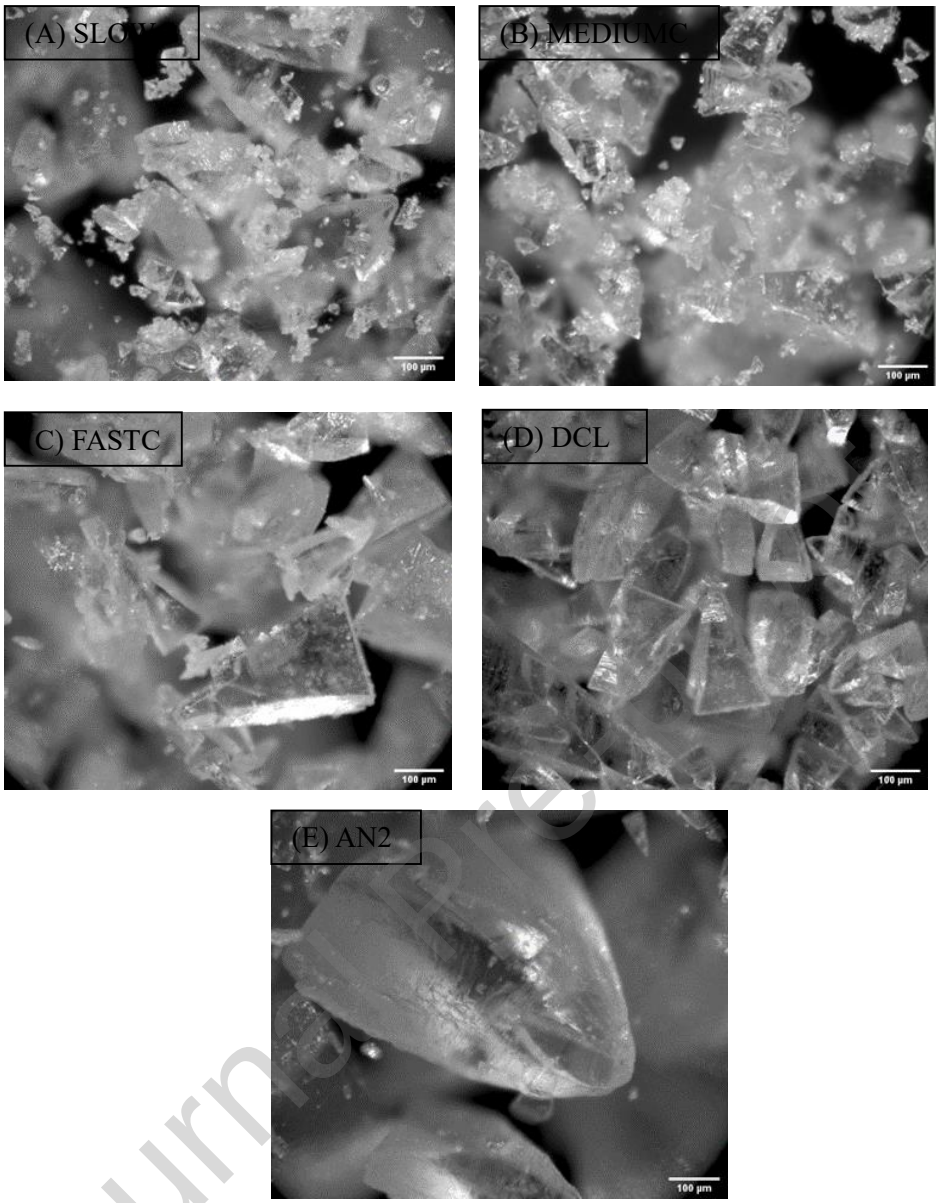



**Figure 8**: Offline microscopic images of all isolated crystal products

(Additional images are provided as supplementary information)

## 4. Conclusion

In conclusion, this study introduces a novel approach to crystallization processes by integrating a ML ANFIS within an automated platform, significantly advancing beyond traditional methodologies. The innovative use of DCL feedback as the foundational strategy, complemented by iterative enhancements through the AN series, has not only demonstrated

superior control over crystallization kinetics but also enabled the production of larger crystals. Additionally, the data generated from this process provides users with an indirect understanding of the kinetics during crystallization, offering valuable insights into the dynamics of crystal growth and behaviour under various conditions. This comprehensive approach enhances our ability to refine crystallization processes, contributing to more efficient and effective iteration techniques.

## Abbreviation List

| Abbreviation | Definition |
|---|---|
| ML | Machine Learning |
| PAT | Process Analytical Technologies |
| ANFIS | Adaptive Neuro-Fuzzy Inference System |
| PLS | Partial Least Square |
| MSZW | Metastable Zone Width |
| MPC | Model Predictive Control |
| DCNNs | Deep Convolutional Neural Networks |
| SLOWC | Slow Cooling |
| MEDC | Medium Cooling |
| FASTC | Fast Cooling |
| DCL | Direct Chord Length |
| AN1 | ANFIS-Driven Iterative Run 1 |
| AN2 | ANFIS-Driven Iterative Run 2 |

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Chat GPT 4.0 to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

# Author Information

**Corresponding Author**

*E-mail: ShinYee.Wong@singaporetech.edu.sg (S.Y. Wong)

**Credit author statement**

The authors affirm their contributions to the paper as follows: Cha Yong Jong, Akshay Mittal, and Shin Yee Wong participated in conceptualizing the study. Cha Yong Jong, and Akshay Mittal conducted crystallization experiments, and coded the machine learning platform. Lee May Loo participated in the writing and organization of the article. Yong Kai Goh and Qiaolin Yuan managed the reactor temperature controller, spectral calibration, and established communication to integrate PATs, reactor, and PharmaMV$^{TM}$ software. Eunice Wan Qi Yeap, Srinivas Reddy Dubbaka, and Harsha Nagesh Rao provided industrial insights regarding demand and methodology. All authors participated in the formal analysis and discussion of results. The finalization of the paper involved collaborative contributions from all authors, who have provided approval for the final version of the manuscript.

**Credit author statement**

**Cha Yong Jong:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - original draft, and Writing - review & editing **Akshay Mittal:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation; Visualization **Felix Lee Jun Jie:** Methodology and Investigation **Lee May Loo:** Writing - original draft, and Writing - review & editing **Yong Kai Goh:** Data curation, Software **Qiaolin Yuan:** Data curation,

Software **Eunice Wan Qi Yeap:** Supervision, Validation **Srinivas Reddy Dubbaka:** Supervision, Validation **Harsha Nagesh Rao:** Supervision, Validation **Shin Yee Wong:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, and Writing - review & editing

**Funding Sources**

This work was supported by the Pfizer- Pharma Innovation Programme Singapore (PIPS) Grant.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data availability**

Data will be made available on request.



# References

Addanki, M., Patra, P., & Kandra, P. (2022). Recent advances and applications of artificial intelligence and related technologies in the food industry. *Applied Food Research*, *2*(2), 100126. https://doi.org/10.1016/j.afres.2022.100126

Bier, R., Eder, C., Schiele, S. A., & Briesen, H. (2023). Selective anomer crystallization from aqueous solution: Monitoring lactose recovery under mutarotation limitation via ATR-FTIR and theoretical rate analysis. *Journal of Dairy Science*. https://doi.org/10.3168/jds.2023-23487

Bund, R. K., & Pandit, A. B. (2007). Sonocrystallization: Effect on lactose recovery and crystal habit. *Ultrasonics Sonochemistry*, *14*(2), 143–152. https://doi.org/10.1016/j.ultsonch.2006.06.003

Carpin, M., Bertelsen, H., Dalberg, A., Bech, J. K., Risbo, J., Schuck, P., & Jeantet, R. (2017). How does particle size influence caking in lactose powder? *Journal of Food Engineering*, *209*, 61–67. https://doi.org/10.1016/J.JFOODENG.2017.04.006

Chernyaeva, S. N., Korobova, L. A., Tolstova, I. S., Safonova, Yu. A., Medvedkova, I. E., & Matytsina, I. A. (2020). Optimization of Functioning of Crystallization Compartment in Sugar Production. *Proceedings of the Russian Conference on Digital Economy and Knowledge Management (RuDEcK 2020)*. https://doi.org/10.2991/aebmr.k.200730.026

Dhumal, R. S., Biradar, S. V., Paradkar, A. R., & York, P. (2008). Ultrasound assisted engineering of lactose crystals. *Pharmaceutical Research*, *25*(12), 2835–2844. https://doi.org/10.1007/s11095-008-9653-9

Durham, R. J. (2009). Modern approaches to lactose production. *Dairy-Derived Ingredients: Food and Nutraceutical Uses*, 103–144. https://doi.org/10.1533/9781845697198.1.103

Evdokimov, I. A., Gridin, A. S., Volodin, D. N., Kulikova, I. K., & Slozhenkina, M. I. (2021). Investigation of Crystallization Process of Lactose in Milk Serum Permeate. *IOP Conference Series: Earth and Environmental Science*, *852*(1). https://doi.org/10.1088/1755-1315/852/1/012031

Faria, N., Pons, M. N., Feyo de Azevedo, S., Rocha, F. A., & Vivier, H. (2003). Quantification of the morphology of sucrose crystals by image analysis. *Powder Technology*, *133*(1–3), 54–67. https://doi.org/10.1016/S0032-5910(03)00078-0

Funes, E., Allouche, Y., Beltrán, G., & Jiménez, A. (2015). A Review: Artificial Neural Networks as Tool for Control Food Industry Process. *Journal of Sensor Technology*, *05*(01), 28–43. https://doi.org/10.4236/jst.2015.51004

Goyache, F., Bahamonde, A., Alonso, J., Lopez, S., del Coz, J. J., Quevedo, J. R., Ranilla, J., Luaces, O., Alvarez, I., Royo, L. J., & Diez, J. (2001). The usefulness of artificial intelligence techniques to assess subjective quality of products in the food industry. *Trends in Food Science & Technology*, *12*(10), 370–381. https://doi.org/10.1016/S0924-2244(02)00010-9

Halfwerk, R., Yntema, D., Van Spronsen, J., & Van der Padt, A. (2021a). A sub-zero crystallization process for the recovery of lactose. *Journal of Food Engineering*, *308*. https://doi.org/10.1016/j.jfoodeng.2021.110677

Halfwerk, R., Yntema, D., Van Spronsen, J., & Van der Padt, A. (2021b). A sub-zero crystallization process for the recovery of lactose. *Journal of Food Engineering*, *308*, 110677. https://doi.org/10.1016/j.jfoodeng.2021.110677

Halfwerk, R., Yntema, D., Van Spronsen, J., & Van der Padt, A. (2023). Recovery of lactose from simulated delactosed whey permeate by a low-temperature crystallization process. *Journal of Dairy Science*, *106*(9), 5958–5969. https://doi.org/10.3168/jds.2023-23249

Hamid, A., Sudrajat, A., Kawangit, R. M., Don, A. G., Huda, M., Jalal, B., Akbar, W., Onn, A., & Maseleno, A. (2018). Determining basic food quality using SAW. *International Journal of Engineering and Technology(UAE)*, *7*(4), 3548–3555. https://doi.org/10.14419/ijet.v7i4.18835

Hourigan, J. A., Lifran, E. V., Vu, L. T. T., Listiohadi, Y., & Sleigh, R. W. (2013). Lactose: Chemistry, Processing, and Utilization. In *Advances in Dairy Ingredients* (pp. 21–41). Wiley. https://doi.org/10.1002/9781118448205.ch2

Jong, C. Y., Mittal, A., Tristan, G., Noller, V., Chan, H. L., Goh, Y., Yeap, E. W. Q., Dubbaka, S. R., Nagesh, H. R., & Wong, S. Y. (2024). ANFIS-Driven Machine Learning Automated Platform for Cooling Crystallization Process Development. *Organic Process Research & Development*. https://doi.org/10.1021/acs.oprd.3c00505

Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, *2*, 100033. https://doi.org/10.1016/j.jafr.2020.100033

Khaire, R. A., & Gogate, P. R. (2021). Understanding the role of different operating modes and ultrasonic reactor configurations for improved sonocrystallization of lactose. *Chemical*

*Engineering and Processing - Process Intensification*, *159*. https://doi.org/10.1016/j.cep.2020.108212

Krupitzer, C., & Stein, A. (2021). Food Informatics—Review of the Current State-of-the-Art, Revised Definition, and Classification into the Research Landscape. *Foods*, *10*(11), 2889. https://doi.org/10.3390/foods10112889

Mavani, N. R., Ali, J. M., Othman, S., Hussain, M. A., Hashim, H., & Rahman, N. A. (2022). Application of Artificial Intelligence in Food Industry—a Guideline. In *Food Engineering Reviews* (Vol. 14, Issue 1, pp. 134–175). Springer. https://doi.org/10.1007/s12393-021-09290-z

McLeod, J. S., Paterson, A. H. J., Bronlund, J. E., & Jones, J. R. (2016). The effect of agitation on the nucleation of α-lactose monohydrate. *International Dairy Journal*, *61*, 114–119. https://doi.org/https://doi.org/10.1016/j.idairyj.2016.04.007

Meng, Y., Yu, S., Zhang, J., Qin, J., Dong, Z., Lu, G., & Pang, H. (2019). Hybrid modeling based on mechanistic and data-driven approaches for cane sugar crystallization. *Journal of Food Engineering*, *257*, 44–55. https://doi.org/10.1016/J.JFOODENG.2019.03.026

Modugno, C., Paterson, A. H. J., & McLeod, J. (2015). Lactose Caking: Influence of the Particle Size Distribution and the Water Content. *Procedia Engineering*, *102*, 114–122. https://doi.org/10.1016/J.PROENG.2015.01.112

Naik, R. R., Gandhi, N. S., Thakur, M., & Nanda, V. (2019). Analysis of crystallization phenomenon in Indian honey using molecular dynamics simulations and artificial neural network. *Food Chemistry*, *300*, 125182. https://doi.org/10.1016/j.foodchem.2019.125182

Ortega-Rivas, E. (2009). Bulk Properties of Food Particulate Materials: An Appraisal of their Characterisation and Relevance in Processing. *Food and Bioprocess Technology*, *2*(1), 28–44. https://doi.org/10.1007/s11947-008-0107-5

Parimaladevi, P., & Srinivasan, K. (2015). Achievement of favorable uniform crystal size distribution of alpha-lactose monohydrate (α-LM) through swift cooling process. *Journal of Food Engineering*, *151*, 1–6. https://doi.org/10.1016/j.jfoodeng.2014.11.019

Patel, S. R., & Murthy, Z. V. P. (2011). Effect of process parameters on crystal size and morphology of lactose in ultrasound-assisted crystallization. *Crystal Research and Technology*, *46*(3), 243–248. https://doi.org/10.1002/crat.201000694

Paterson, A. H. J. (2009). Production and uses of lactose. In *Advanced Dairy Chemistry* (Vol. 3, pp. 105–120). Springer New York. https://doi.org/10.1007/978-0-387-84865-5_4

Paz Suárez, L. A., Georgieva, P., & Feyo de Azevedo, S. (2011). Nonlinear MPC for fed-batch multiple stages sugar crystallization. *Chemical Engineering Research and Design*, *89*(6), 753–767. https://doi.org/10.1016/J.CHERD.2010.10.010

Raghavan, S. L., Ristic, R. I., Sheen, D. B., & Sherwood, J. N. (2001). The bulk crystallization of α-lactose monohydrate from aqueous solution. *Journal of Pharmaceutical Sciences*, *90*(7), 823–832. https://doi.org/10.1002/jps.1036

Simone, E., Tyler, A. I. I., Kuah, D., Bao, X., Ries, M. E., & Baker, D. (2019). Optimal design of crystallization processes for the recovery of a slow-nucleating sugar with a complex chemical equilibrium in aqueous solution: The case of lactose. *Organic Process Research and Development*, *23*(2), 220–233. https://doi.org/10.1021/acs.oprd.8b00323

Tsakanikas, P., Karnavas, A., Panagou, E. Z., & Nychas, G.-J. (2020). A machine learning workflow for raw food spectroscopic classification in a future industry. *Scientific Reports*, *10*(1), 11212. https://doi.org/10.1038/s41598-020-68156-2

Wong, S. Y., Bund, R. K., Connelly, R. K., & Hartel, R. W. (2012). Designing a lactose crystallization process based on dynamic metastable limit. *Journal of Food Engineering*, *111*(4), 642–654. https://doi.org/10.1016/j.jfoodeng.2012.03.003

Wong, S. Y., & Hartel, R. W. (2014a). Crystallization in Lactose Refining-A Review. *Journal of Food Science*, *79*(3). https://doi.org/10.1111/1750-3841.12349

Wong, S. Y., & Hartel, R. W. (2014b). Crystallization in Lactose Refining—A Review. *Journal of Food Science*, *79*(3), R257–R272. https://doi.org/https://doi.org/10.1111/1750-3841.12349

Yuping, S., Baomin, L., & Richard, H. (2006). *Crystal refining technologies by controlled crystallization*. https://patents.google.com/patent/US20060128953A1/en

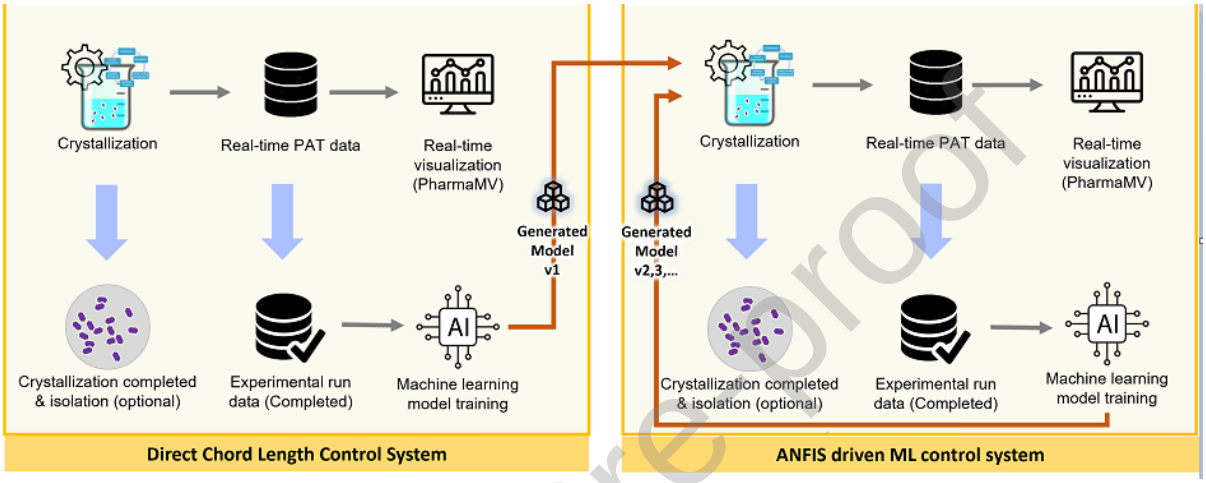
Zeng, X. M., Martin, G. P., Marriott, C., & Pritchard, J. (2010). The Influence of Crystallization Conditions on the Morphology of Lactose Intended for Use as a Carrier for Dry Powder Aerosols. *Journal of Pharmacy and Pharmacology*, *52*(6), 633–643. https://doi.org/10.1211/0022357001774462

Zhang, J., Meng, Y., Wu, J., Qin, J., Hui wang, Yao, T., & Yu, S. (2020). Monitoring sugar crystallization with deep neural networks. *Journal of Food Engineering*, *280*. https://doi.org/10.1016/j.jfoodeng.2020.109965

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

## Graphical abstract




## Highlights

- Automated crystallization platform integrating ML and PAT for lactose recovery.
- Utilized Adaptive Neuro-Fuzzy Inference System (ANFIS) for dynamic process adjustment.
- Achieved quantitative enhancements in lactose crystal size and yield productivity.
- AI-driven control approach outperforms conventional crystallization methods