Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Development of a Hybrid Model for the Paracetamol Batch Dissolution in Ethanol Using Universal Differential Equations

Fernando Arrais R. D. Lima[a,b,*], Amyr Crissaff Silva[a], Marcellus G. F. de Moraes[c,d], Amaro G. Barreto Jr.[a], Argimiro R. Secchi[a,d], Idelfonso Nogueira[b], Maurício B. de Souza Jr.[a,d]

[a] School of Chemistry, EPQB, Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco E, 21941-914, Rio de Janeiro, RJ – Brazil
[b] Chemical Engineering Department, Norwegian University of Science and Technology, Trondheim, 793101, Norway
[c] Instituto de Química, Rio de Janeiro State University (UERJ), Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro, RJ, 20550-900, Brazil
[d] PEQ/COPPE – Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco G, G115, 21941-914, Rio de Janeiro, RJ – Brazil
* Corresponding Author: farrais@eq.ufrj.br.

ABSTRACT

Crystallization is a relevant process in the pharmaceutical industry for product purification and particle production. An efficient crystallization is characterized by crystals produced with the desired attributes. Therefore, modeling this process is a key point to achieve this goal. In this sense, the objective of this work is to propose a hybrid model to describe paracetamol dissolution in ethanol. The universal differential equations methodology is considered in the development of this model, using a neural network to predict the dissolution rate combined with the population balance equations to calculate the moments of the crystal size distribution (CSD) and the concentration. The model was developed using experimental batches. The dataset is composed of concentration measurements obtained using attenuated total reflectance-Fourier transform infrared (ATR-FTIR). The objective function of the optimization problem is to minimize the relative absolute difference between the experimental and the predicted concentration. The hybrid model efficiently predicted the concentration compared to the experimental measurements. Furthermore, the hybrid approach made predictions of the moments of the CSD similar to the population balance model proposed by Kim et al. [1], being able to successfully calculate batches not considered in the training dataset. Moreover, the performance of the hybrid model was similar to the phenomenological one based on population balance. Therefore, the universal differential equations approach is presented as an efficient methodology for modeling crystallization processes with limited information.

Keywords: Crystallization, hybrid model, pharmaceutical industry.

INTRODUCTION

Crystallization is a separation and purification process that is characterized by a phase change in which crystalline product is obtained from a solution [1]. It is a relevant unit operation and is commonly used for purification, separation and production. It is a practical method that produces pure, concentrated, appealing and convenient to handle chemical substances. Thus, crystallization processes are present in many industries such as pharmaceuticals, food and agrochemicals. Many active ingredients in pharmaceutical industries are produced in the form of crystals, such as paracetamol [2], praziquantel [3] and amoxicillin [4].

For the development of an efficient crystallization process, the product should present the desired characteristics, such as size and shape. Therefore, modeling the crystallization kinetics to calculate the crystal size distribution (CSD) is a key point to predict these characteristics and to optimize the process. Population balance

model (PBM) is the methodology usually applied for modeling crystallization processes. This approach was used to model the batch crystallization of many compounds, such as potassium dihydrogen phosphate (KDP) [5], paracetamol [2, 6, 7] and potassium sulfate [8].

Over the last years, advancements in machine learning have been fueled by greater computational power, the abundance of extensive experimental data, and the development of more efficient algorithms. Machine learning techniques have achieved significant success in analyzing and predicting complex behaviors from large datasets, as demonstrated by applications such as web searches, email filtering, and image recognition [9, 10]. In the field of crystallization, machine learning is emerging as a valuable tool for predicting and optimizing processes to meet specific product requirements [11, 12].

Even with these positive points of machine learning models, they are usually empirical models and do not consider the physical laws that describe the system. Hybrid models present as an alternative to solve these problems. Hybrid models are a combination of data-driven and first principles models [13]. These approaches combine the flexibility of data-driven with the knowledge of physical laws from first principles models [14].

Regarding crystallization processes, Lima et al. [15] combined neural networks and PBM for modeling the nucleation, crystal growth and dissolution for the potassium sulfate batch crystallization. Wu et al. [16] proposed a physics-informed recurrent neural network (PIRNN) for modeling the crystallization process of aspirin. Sitapure and Kwon [17] developed hybrid modeling approach combining time-series-transformers (TSTs) to population balance model. They used TSTs to obtain the nucleation and growth rates for the batch crystallization of dextrose. Then, the rates were used in the population balance equations to obtain the crystal size distribution (CSD). There are also hybrid model applications used to describe the batch crystallization of sugar accounting for nucleation, growth and agglomeration [18, 19].

The goal of this work is to propose a hybrid model based on universal differential equations (UDEs) approach with a neural network as a universal approximator to describe the dissolution for the paracetamol batch crystallization in ethanol. According to Xiouras et al. [9] and Lima et al. [10], the use of hybrid models for crystallization processes is not new, but there are few applications of these methodologies. Most studies proposed hybrid models for sugar crystallization. There is a lack of hybrid model applications for the pharmaceutical field and none of them considered paracetamol. Paracetamol is a pharmaceutical compound whose crystallization has been extensively explored in the literature for modeling and control studies. The hybrid model is composed of a neural network to calculate the dissolution rate combined with the population balance equations to predict the concentration and the moments of the CSD. The concentration measurements obtained by Kim et al. [2] were used to develop the hybrid model. Therefore, the hybrid approach presents flexibility to model the dissolution rate compared to the PBM.

METHODOLOGY

Population Balance Model

Kim et al. [2] proposed a PBM to describe the paracetamol batch crystallization in ethanol. The model accounts for nucleation, crystal growth and dissolution. The method of moments was applied to solve this problem.

Eq. 1 describes the i^{th} order moment, in which L is the characteristic crystal size, and $n(L)$ is the number density of crystals. The moments μ_0, μ_1, $\mu_2$and μ_3 are proportional to the total number, length, surface area, and volume of crystals, respectively.

$$\mu_i = \int_0^\infty L^i\, n(L) dL,\ i = 1,2,3,\ldots \tag{1}$$

The paracetamol dissolution PBM is described by Eq. 2, where k_v is the volume shape factor, ρ_c is the solid density of crystals (g/cm^3), and D is the dissolution rate.

$$\frac{d\mu_0}{dt} = 0 \tag{2a}$$

$$\frac{d\mu_1}{dt} = D\mu_0 \tag{2b}$$

$$\frac{d\mu_2}{dt} = 2D\mu_1 \tag{2c}$$

$$\frac{d\mu_3}{dt} = 3D\mu_2 \tag{2d}$$

$$\frac{dC}{dt} = -3k_v\rho_c\, D\mu_2 \tag{2e}$$

The dissolution rate is described by Eq. 3, where k_d is the pre-exponential rate constants for the crystal growth $((\mu m/min)(g/g)^{-\gamma_d})$, E_{a_d} is the activation energy (J/mol) and γ_d is a exponential parameters on supersaturation. R is the universal gas constant (J/(mol K)) [20].

$$D = k_d exp\left(\frac{-E_{a_d}}{RT}\right)(C_S - C)^{\gamma_d} \tag{3}$$

In Eq. 3, C (g solute/g solvent) is the concentration in the solution and C_S (g solute/g solvent) is the saturated concentration at the system temperature. The saturated concentration can be calculated by the polynomial described by Eq. 5, in which T is the temperature in Kelvin.

$$C_S = -8.707 + 9.669x10^{-2}T - 3.610x10^{-4}T^2 + 4.590x10^{-7}T^3 \quad (5)$$

To estimate the parameters of this PBM, Kim et al. [2] performed seven experimental batches, which five were used to develop the model and two to validate it. All experimental batches were developed at constant temperatures; two of them were performed at 10°C, one at 15°C, three at 20°C, and one at 30°C. The initial concentration was also changed for the batches, considering values between 0.142 and 0.202 g/g. The model was developed using concentration measurements obtained during the experimental batches with attenuated total reflectance Fourier transform infrared (ATR-FTIR).

Universal Differential Equations

Universal differential equations (UDEs) are differential equations enhanced by the inclusion of one or more universal approximators. Various functions and series have been recognized as universal approximators due to their versatility and adaptability. Neural networks excel as universal approximators in higher-dimensional settings. Their layered architecture and the use of nonlinear activation functions allow them to effectively approximate functions within multidimensional spaces [15].

In this work, a hybrid model was developed to describe the dissolution kinetics of paracetamol in ethanol based on the UDE approach. A neural network is used to calculate the dissolution rate, while the moments of the CSD and the concentration are obtained by the population balance equations. The neural network uses concentration measurements and the saturated concentration to calculate the dissolution rate. Before being used by the neural network, the inputs were scaled between zero and one. Figure 1 presents a scheme of the proposed UDE.

The model was developed using the concentration measurements obtained by Kim et al. [2] during the batch experiments. Similarly to Kim et al. [2], two experimental batches were used to test the model. However, four batches were used to train and one to validate the model to avoid overfitting. The optimization problem consisted of designing the neural network that minimizes Eq. 6, where C_i is the experimental concentration, $\hat{C}_i$ is the predicted concentration and N is the number of measurements. As this model accounts only for undersaturated conditions, the constraint described in Eq. 7 was considered for the optimization problem.

$$F_{obj} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{C}_i - C_i}{C_i}\right| \quad (6)$$

$$\hat{C}_i / C_S(T_i) \leq 1 \quad (7)$$

The hybrid model was developed in Python using the deep learning library, Pytorch [21] and the neural differential equations library TorchDyn [22]. The optimizer Adam was used to train the hybrid model. Early stopping was also considered in the training to avoid overfitting with a criterion of stopping the training whenever the model exceeded 70 epochs without lowering validation error.

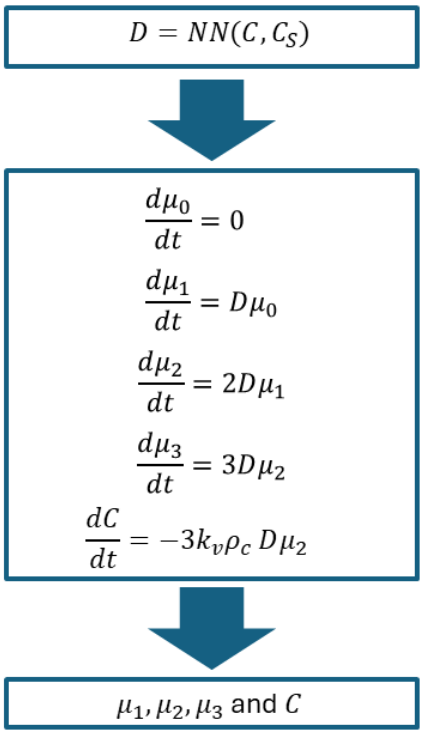

Figure 1. Hybrid model scheme.

RESULTS AND DISCUSSION

The hybrid model was composed of a neural network with one hidden layer containing 30 neurons and using hyperbolic tangent as the activation function. An adaptive learning rate was considered with an initial value equal to 10^{-5}. The learning rate is reduced to 10% of the initial value after 500 epochs.

Table 1 presents the relative mean absolute error (RMAE) between the predicted and measured concentrations for the PBM and the hybrid model. While the performance of the PBM was better for the test dataset, the hybrid model was able to make predictions closer to the experimental values for the training and validation datasets. However, both approaches presented close values of RMAE in all cases.Fig. 2 presents the performance of the hybrid model and the PBM to predict the solution concentration for the validation batch. The PBM and the hybrid model showed close performances. Their concentration predictions presented a similar behavior and achieved the same equilibrium concentration at the end of the batch. Compared to the experimental values, the UDE model made predictions closer to the concentration

measurements than the PBM.

Table 1: Relative mean absolute error (RMAE) between the concentration measurements and the predicted values by the PBM and the hybrid model.

Model	Training	Test	Validation
PBM	6.50×10^{-3}	5.18×10^{-3}	4.31×10^{-3}
UDE	3.50×10^{-3}	8.29×10^{-3}	3.18×10^{-3}

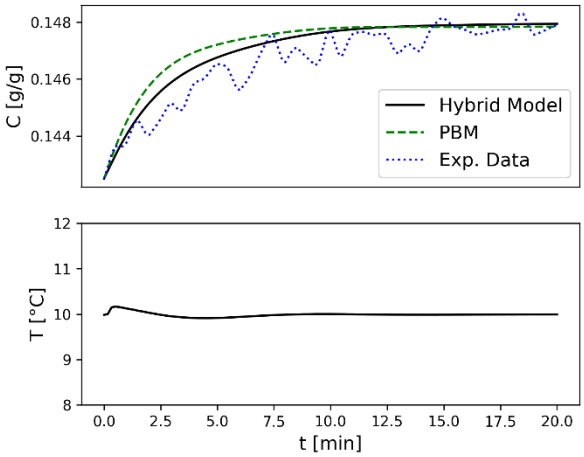

Figure 2. Performance of the PBM and hybrid model to predict the concentration for the validation batch.

Fig. 3 presents the performance of the hybrid model and the PBM to predict the solution concentration for one of the test batches. Once again, both methodologies were able to make close concentration predictions. The RMAE value, in Table 1, for the hybrid model for the test dataset was higher than the PBM because of the offset that the UDE model left in the equilibrium concentration compared to the experimental data.

The values presented in Table 1 and the performances showed in Figs 2 and 3 suggest that the hybrid model presented a better performance for the training and validation dataset. On the other hand, the PBM was more efficient for the test dataset. These facts suggest that the UDE model is able to make efficient predictions, but the PBM was more effective for conditions not considered in the training dataset. Therefore, the performance of the PBM was better for extrapolation. This can be expected as the PBM is based on physics laws.

Fig. 4 presents a simulation of one experimental batch using the hybrid model and the PBM. In this simulation, μ_1, μ_2, μ_3 and the solution concentration were calculated. Regarding the concentration predictions, both approaches were able to achieve the same equilibrium concentration at the end of the batch. However, PBM presented a faster response compared to the UDE trajectories. This behaviour was also observed in Fig. 2, showing that the PBM is overestimating this characteristic when compared to the experimental data. The same thing happens for the predictions of the moments of the CSD. Both models were able to achieve similar equilibrium values at the end of the batch. The same differences in the transient response are also present in the moments.

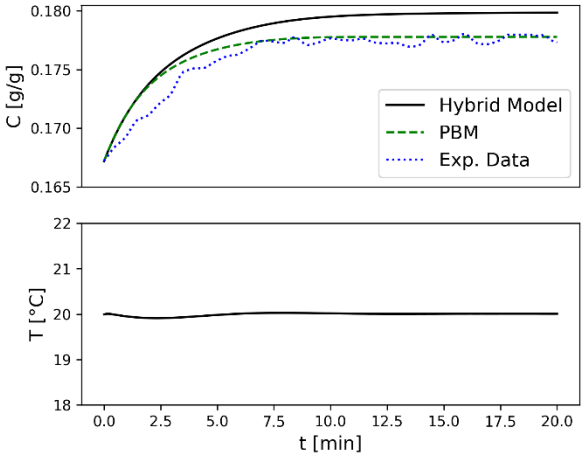

Figure 3. Performance of the PBM and hybrid model to predict the concentration for one of the test batches.

The differences presented in Fig. 4 between the predictions of the hybrid model and the PBM are consequence of the challenge when dealing with dissolution. The dissolution is a quick phenomenon. Therefore, when the crystals are added in the undersaturated solution during the batch, they are quickly dissolved until the equilibrium is reached. This fact makes it difficult to find an efficient model for dissolution. The errors calculated in Table 1 and the profiles obtained in Figs. 2 and 3 suggest that the hybrid model was able to make predictions closer to the real measurements when the crystals are dissolving compared to the PBM.

The proposed hybrid approach based on UDE presented a good performance for modeling the dissolution of paracetamol in ethanol. The PBM presented the best performance for extrapolations. On the other hand, the UDE model was able to make better predictions than the PBM when the crystals were dissolving. Even with these advantages and disadvantages, replacing the dissolution rate by a neural network was an effective alternative for modeling this phenomenon, as suggested by Lima et al. [15].

CONCLUSION

In this work, a hybrid model based on UDE was developed for describing the paracetamol dissolution in ethanol. There is a lack of studies proposing hybrid models for crystallization or dissolution of pharmaceutical compounds and none of them considered paracetamol

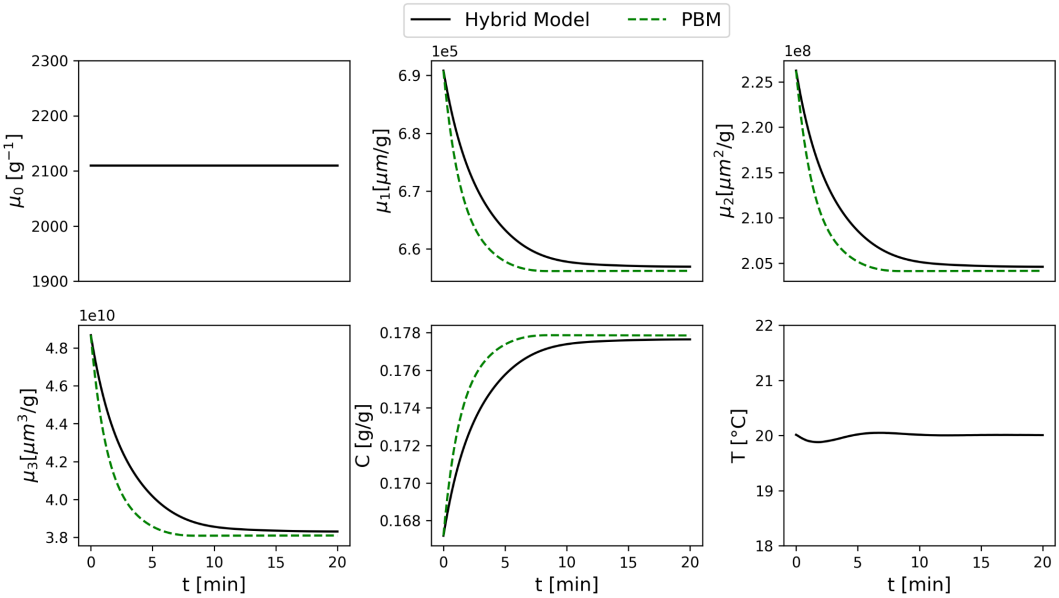

Figure 4: Performance of the hybrid model and the PBM to predict the moments of the CSD and the concentration of the solution considering the same conditions of one experimental batch.

[9, 10]. The dataset used in this work consisted of seven experimental batches developed by Kim et al. [2], where four of them were used to train the model, one for validation and two to test it. The developed hybrid model was able to make predictions of the solution concentration close to the experimental measurements, presenting low RMAE values.

The performance of the hybrid model was also compared to the PBM developed by Kim et al. [2]. The values calculated by the hybrid model of solution concentration and moments of CSD were close to the one obtained with the PBM. The RMAE between the concentration measurements and the predicted values by the PBM was similar to the RMAE calculated with the hybrid model. The hybrid methodology based on UDE presents as an efficient alternative for modeling crystallization kinetics. As a future work, the nucleation and crystal growth will be considered for developing the hybrid model of paracetamol crystallization in ethanol.

ACKNOWLEDGEMENTS

Coordination for the Improvement of Higher Education Personnel (CAPES, Finance Code 001), the National Council for Scientific and Technological Development (CNPq, Grants 303587/2020-2, 311153/2021-6, and 313337/2021-7), and the Research Support Foundation of the State of Rio de Janeiro (FAPERJ, Grants E 26/111.671/2013 and E-26/201.148/2022).

REFERENCES

1. Ahn B., Bosetti L., Mazzotti M. Secondary Nucleation by Interparticle Energies. II. Kinetics. *Crystal Growth & Design* 22:74–86 (2022) https://doi.org/10.1021/acs.cgd.1c00928
2. Kim Y., Kawajiri Y., Rousseau R.W., Grover M.A. Modeling of nucleation, growth, and dissolution of paracetamol in ethanol solution for unseeded batch cooling crystallization with temperature-cycling strategy. *Industrial & Engineering Chemistry Research* 62:2866–2881 (2023). https://doi.org/10.1021/acs.iecr.2c03465
3. Moraes M. G. F., Barreto Jr. A. G., Secchi A. R., Souza Jr. M. B., Lage P. L. C., Myerson A. S. Polymorphism of Praziquantel: Role of Cooling Crystallization in Access to Solid Forms and Discovery of New Polymorphs. *Crystal Growth & Design* 23(2):1247-1258 (2023) https://doi.org/10.1021/acs.cgd.2c01381
4. McDonald M. A., Bommarius A. S., Rousseau R. W., Grover M. A. Continuous reactive crystallization of β-lactam antibiotics catalyzed by penicillin G acylase. Part I: Model development. *Comput. Chem. Eng.* 123:331–343 (2019). http://dx.doi.org/10.1016/j.compchemeng.2018.12.029
5. Moraes M.G., Grover M.A., Souza Jr. M.B., Lage P.L., Secchi A.R. Optimal control of crystal size and shape in batch crystallization using a bivariate population balance modeling. *IFAC-PapersOnLine* 54 (3): 653–660 (2021). http://dx.doi.org/10.1016/j.ifacol.2021.08.316

6. Nagy Z. K., Fujiwara M., Woo X. Y., Braatz R. D. Determination of the Kinetic Parameters for the Crystallization of Paracetamol from Water Using Metastable Zone Width Experiments. *Ind. Eng. Chem. Res.* 47: 1245–1252 (2008). https://doi.org/10.1021/ie060637c
7. Li H., Kawajiri Y., Grover M. A., Rousseau R. W. Modeling of Nucleation and Growth Kinetics for Unseeded Batch Cooling Crystallization. *Ind. Eng. Chem. Res.* 56: 4060–4073 (2017). https://doi.org/10.1021/acs.iecr.6b04914
8. Moraes M.G.F., Lima F.A.R.D., Lage P.L.C., Souza Jr. M.B., Barreto Jr.,A.G., Secchi, A.R. Modeling and predictive control of cooling crystallization of potassium sulfate by dynamic image analysis: Exploring phenomenological and machine learning approaches. *Industrial & Engineering Chemistry Research* 62:9515–9532 (2023). https://doi.org/10.1021/acs.iecr.3c00739
9. Xiouras C., Cameli F., Quilló G. L., Kavousanakis M. E., Vlachos D. G., Stefanidis G. D. Applications of Artificial Intelligence and Machine Learning Algorithms to Crystallization. *Chemical Reviews* 122 (15): 13006-13042 (2022). https://doi.org/10.1021/acs.chemrev.2c00141
10. Lima F. A. R. D., Moraes M. G. F., Barreto Jr. A. G., Secchi A. R., Grover M. A., Souza Jr. M. B. Applications of machine learning for modeling and advanced control of crystallization processes: Developments and perspectives. *Digital Chemical Engineering* 14:100208, 2025. https://doi.org/10.1016/j.dche.2024.100208
11. Szilagyi B., Nagy, Z.K. Real-time feasible model-based crystal size and shape control of crystallization processes. *Computer Aided Chemical Engineering* 46: 1273–1278 (2019). http://dx.doi.org/10.1016/B978-0-12-818634-3.50213-7.
12. Zheng Y., Wang X., Wu Z. Machine learning modeling and predictive control of the batch crystallization process. *Ind. Eng. Chem. Res.* 61 (16): 5578–5592 (2022). http://dx.doi.org/10.1021/acs.iecr.2c00026
13. Sharma N., Liu Y.A. A hybrid science-guided machine learning approach for modeling chemical processes: A review. *AIChE J*. 68 (5): e17609 (2022). http://dx.doi.org/10.1002/aic.17609
14. Zendehboudi S., Rezaei N. Lohi, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* 228: 2539–2566 (2018). http://dx.doi.org/10.1016/j.apenergy.2018.06.051
15. Lima F.A.R.D., Rebello C.M., Costa E.A., Santana V.V., Moares M.G., Barreto Jr. A.G., Secchi A.R., Souza Jr. M.B., Nogueira I.B. Improved modeling of crystallization processes by universal differential equations. *Chem. Eng. Res. Des.* 200: 538-549 (2023). http://dx.doi.org/10.1016/j.cherd.2023.11.032
16. Wu G., Yion W.T.G., Dang K.L.N.Q., Wu, Z. Physics-informed machine learning for MPC: Application to a batch crystallization process. *Chem. Eng. Res. Des.* 192: 556–569 (2023). http://dx.doi.org/10.1016/j.cherd.2023.02.048
17. Sitapure N., Kwon J.S.-I. Introducing hybrid modeling with time-series-transformers: A comparative study of series and parallel approach in batch crystallization. *Ind. Eng. Chem. Res.* 62 (49): 21278–21291 (2023). http://dx.doi.org/10.1021/acs.iecr.3c02624
18. Georgieva P., Meireles M., Feyo de Azevedo S. Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena. *Chem. Eng. Sci.* 58 (16): 3699–3713 (2003). http://dx.doi.org/10.1016/S0009-2509(03)00260-4
19. Oliveira C., Georgieva P., Rocha F., Feyo de Azevedo S. Artificial neural networks for modeling in reaction process systems. *Neural Comput. Appl.* 18: 15–24 (2009). http://dx.doi.org/10.1007/s00521-008-0200-8
20. Lima F.A.R.D., Moraes M.G.F., Grover, M.A., Barreto Jr. A.G., Secchi A.R., Souza Jr., M.B. Neural network inverse model controllers for paracetamol unseeded batch cooling crystallization. *Ind. Eng. Chem. Res.* 63 (45): 19613–19627 (2024). http://dx.doi.org/10.1021/acs.iecr.4c02060
21. Paszke A, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *ArXiv* (2019). https://doi.org/10.48550/arXiv.1912.01703
22. Poli M., Massaroli S., Yamashita A., Asama H., Park J. TorchDyn: A Neural Differential Equations Library. *ArXiv* (2020). https://doi.org/10.48550/arXiv.2009.09346