Controlling Paracetamol Batch Crystallization in Ethanol by Reinforcement Learning

Fernando Arrais R. D. Lima * **Ruan R. Faria** *
Marcellus G. F. de Moraes ** **Argimiro R. Secchi** ***
Idelfonso B. R. Nogueira **** **Maurício B. de Souza Jr.** *

* Escola de Química, EPQB, Universidade Federal do Rio de Janeiro, P.O. Box 68542, Rio de Janeiro, RJ 21941-909, Brazil (e-mail: farrais@eq.ufrj.br, rrfaria@eq.ufrj.br, mbsj@eq.ufrj.br).
** Rio de Janeiro State University (UERJ), Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro, RJ 20550-900, Brazil (e-mail: marcellus@peq.coppe.ufrj.br)
*** Programa de Engenharia Química, PEQ/COPPE, Universidade Federal do Rio de Janeiro, P.O. Box 68502, Rio de Janeiro, RJ 21941-972, Brazil (e-mail: arge@peq.coppe.ufrj.br)
**** Chemical Engineering Department, Norwegian University of Science and Technology, Trondheim 793101, Norway (e-mail: idelfonso.b.d.r.nogueira@ntnu.no)

Abstract: In this work, we proposed a controller based on reinforcement learning (RL) for the unseeded batch crystallization of paracetamol in ethanol. The controller aims to achieve the crystal mean volume size and the crystal mass at five different targets by manipulating the temperature. We used the deep deterministic policy gradient (DDPG) algorithm to train the control agent. The performance of the RL controller was compared to an NMPC using a population balance model (PBM) as its internal model and tested for the five different scenarios. The controllers were also tested taking into account a 5% disturbance in the concentration measurement. Both controllers were able to reach values of the controlled variables close to the targets even accounting for disturbances. However, the RL controller was able to calculate the control action much faster than the NMPC and imposed less temperature changes, which presents as better alternative for real control applications. Therefore, the RL controller presented as a more suitable approach than an NMPC using a PBM as its internal model for controlling the paracetamol batch crystallization process.

Keywords: Crystallization, Pharmaceutical Industry, Reinforcement Learning, NMPC, Control.

1. INTRODUCTION

Crystallization is a unit operation where a solid crystalline product is typically produced from a solution (Ahn et al., 2022). For the process to be efficient, the resulting crystals must meet specific shape and size requirements that comply with product quality regulations. As a result, an effective control system is essential to ensure these standards are achieved (Nagy and Braatz, 2012; Moraes et al., 2023).

Model predictive control (MPC) is a control strategy that utilizes an internal model to manage the process, and it is commonly applied to crystallization processes in the literature (Nayhouse et al., 2015; Cao et al., 2016; Bötschi et al., 2018). One such application was developed by Sitapure and Kwon (2023), who employed a population balance model (PBM) accounting for growth and nucleation of cadmium telluride quantum dot batch crystallization within the MPC framework. They conducted simulations to regulate the mean crystal size and the desired standard deviation of the crystal size distribution (CSD) by adjusting the solute concentration.

Szilágyi et al. (2018) introduced a nonlinear model predictive control (NMPC) algorithm and evaluated its performance in simulation-based applications for controlling the crystal size distribution (CSD) in batch-cooling crystallization processes. Recently, Lima et al. (2024) developed an NMPC to control both the mean crystal size and the yield in paracetamol batch crystallization in ethanol, with temperature as the manipulated variable.

The machine learning methodology known as reinforcement learning (RL) has been standing out in the area of artificial intelligence (Faria et al., 2023, 2024b). The impressive results achieved in these areas (largely driven by advances in deep neural networks and new RL algorithms) have sparked interest within the process control community (Faria et al., 2022). RL is defined by an agent that learns autonomously within a system through numerical rewards, following a Markov decision process. The

agent's learning occurs through interactions, without the need for a process model, and can incorporate data-driven insights and simulation-based information (Sutton and Barto, 2018). As a result, RL is emerging as a promising alternative to traditional model-based control methods, such as Model Predictive Control (MPC), for managing batch processes.

The application of reinforcement learning (RL) to control crystallization processes is a relatively recent development in the literature. The first instance of RL being used for crystallization control was by Zhang et al. (2020), who developed control policies to determine the optimal field shapes and orientations for rapidly eliminating grain boundaries and restoring circular crystal morphologies from anisotropic forms. Following this, Manee et al. (2021) applied RL to control the crystallization of NaCl in a water-ethanol system, adjusting antisolvent addition and temperature to regulate the mean crystal size. They demonstrated the effective implementation of this control strategy in an experimental setup.

Anandan et al. (2022) developed an inverse reinforcement learning approach to control paracetamol batch crystallization. The machine learning model was trained using crystallization processes controlled by PID and MPC, and then applied in simulations of paracetamol crystallization. Manee et al. (2022) employed convolutional neural networks (CNNs) as soft sensors combined with reinforcement learning to control sodium chloride crystallization in water, using ethanol as an antisolvent. They manipulated both temperature and antisolvent addition to control crystal size in an experimental setup. Meng et al. (2023) applied reinforcement learning in simulations to control crystal size in the crystallization of KDP in water and aspirin in ethanol-water mixtures. In the KDP system, temperature was the key manipulated variable, while both temperature and antisolvent addition were adjusted in the aspirin crystallization system.

In this work, we adopted RL for controlling the unseeded batch crystallization of paracetamol in ethanol. Paracetamol is a medicine and its crystallization has been of interest in the literature for control applications (Griffin et al., 2017; Grover et al., 2020; Anandan et al., 2022; Lima et al., 2024). We used the population balance model (PBM) developed by Kim et al. (2023) to train the agent. The controller aims to maintain the crystal mean volume size and the crystal mass at the targets by manipulating the temperature for five scenarios. Different from previous works, the current controller uses the temperature, the current values of the controlled variables and their deviations from the setpoints as inputs. The performance of the RL based controller was compared to an NMPC and they were also tested considering disturbances in the concentration measurement.

2. DEEP REINFORCEMENT LEARNING BASED CONTROL

This work employs reinforcement learning to control the unseeded batch crystallization of paracetamol in ethanol. To simulate this process, we used the PBM previously developed by Kim et al. (2023). More details about the

PBM can be found in Kim et al. (2023) and Lima et al. (2024).

The RL problem is formulated as an MDP (Markov decision process), as shown in Fig. 1. The agent learns the control task based on the experience obtained when interacting with the simulated process environment (PBM), guided by rewards. Specifically, the Markov state is defined by the current value of the control action, controlled variable and its deviation in relation to the setpoint. The agent executes the control action (the increments of the manipulated variable $\Delta u(k) = \Delta T(k)$) based only on the current state (s_k), which causes the environment to transition to a new state $(s_{k+1} = p(s_{k+1}|s_k, a_k))$ approximated with PBM. The reward (r_k) used to guide the control agent's learning considers the absolute setpoint error and control action deviation with relative importance weighted by $Q = [Q_1, \cdots, Q_{N_y}]$ and $R = [R_1, \cdots, R_{N_u}]$, with N_y and N_u representing the size of these respective vectors.
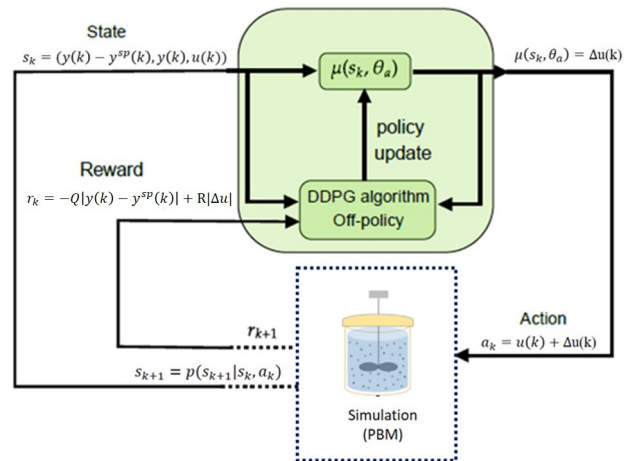

Fig. 1. MDP scheme for controlling the paracetamol batch crystallization in ethanol.

To produce the optimal policy, the agent executes this process repeatedly (episodes). A self-parameterized policy $(\theta_a \in R^k)$ is assumed, in which the objective is to maximize the cumulative sum of rewards $(R(\tau))$ obtained regardless of the sampled trajectory $\tau = s_1, a_1, \cdots, s_T, a_T$, initial state $(p(s_1))$ and state transition s_{k+1}, as shown in (1), with γ representing the discount factor.

$$\theta_a{}^* = \underset{\theta_a}{argmax} \int p(s_1) \prod_{k=1}^{T} p(s_{k+1}|s_k, a_k) \pi_{\theta_a}(a_k|s_k) R(\tau)$$

$$R(\tau) = \sum_{k=1}^{T} \gamma^{i-1} r(s_k, a_k, s_{k+1})$$

(1)

In this study, the deep deterministic policy gradient (DDPG) (Lillicrap, 2015) algorithm was considered for training the control agent. Fig. 2 details the structure of the critic and actor employed, in which the critic network estimates the value of rewards by solving (2), then the actor uses the gradient from (1) to improve the policy by ascending gradient, according to the theorem of deterministic gradient, as shown in (3). More details about
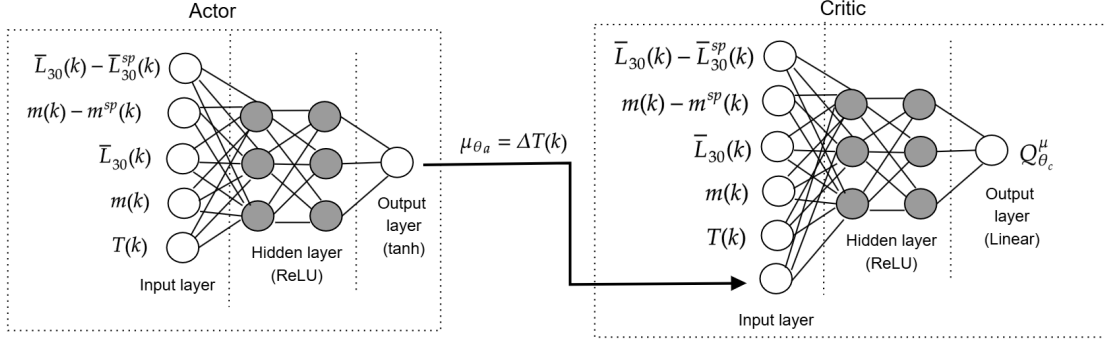
Fig. 2. Structures of actor and critical networks for training the control agent applied to the crystallization process.

the implementation of the algorithm can be seen in Faria et al. (2024a).

$$\boldsymbol{\theta_{c,k+1}} = \underset{\boldsymbol{\theta_c}}{argmin}[Q^\mu(s_k, a_k, \boldsymbol{\theta_{c,k}}) - Q^\mu(s_k, a_k)] \quad (2)$$

$$\boldsymbol{\theta_{a},k+1} = \boldsymbol{\theta_{a},k} + E[\nabla_{\theta_k}\mu_{\boldsymbol{\theta_a}}(s_k)\nabla_a Q^\mu(s_k, a_k, \boldsymbol{\theta_c})]|_{a=\mu_{\boldsymbol{\theta_a}}} \quad (3)$$

Once trained, the resulting actor network implements the control actions in the actual process. Therefore, this controller does not depend on an internal model to compute control actions as it evaluates an approximate function by a neural network, potentially resulting in an inherently adaptive and computationally efficient controller.

The training of the control agent was based on the DDPG algorithm, in which the definition of the structure of the actor and critical networks, parameters of the ADAM optimizer and parameters of the algorithm itself are summarized in Table 1. Specifically, N, D, K were chosen to balance computational and learning efficiency of the algorithm. The DDPG algorithm was selected because many works successfully applied this approach for regulatory control (Faria et al., 2022, 2023). The number of episodes was selected based on a sensitivity analysis. The policy used for testing was the one for the last episode. The structure of the actor and critic networks was selected based on the hyperparameter analysis carried out in Faria et al. (2022). Based on Faria et al. (2022), the critic and actor networks have different optimal configurations. The actor presents more layers and a lower learning rate compared to the critic network, since it approximates the deterministic gradient. Therefore, a more conservative actor network than the critical network was used in order to improve the convergence of the algorithm and, consequently, improve the training of the reinforcement learning agent.

The nominal process conditions (initial condition) and desired values were varied during each episode, randomly selecting, considering a uniform distribution, from 5 different previously defined cases, meaning that 50,000 PBM-based simulations were used to fill the buffer, with the following 20,000 updates for the algorithm to learn. Additionally, the parameters referring to the relative importance of states and inputs for the objective function were $Q = [1, 0.01]$ and $R = [0.01]$. The same tuning was considered for

Table 1. The DDPG algorithm parameters.

Parameters	Value
DDPG	
Discount factor (γ)	0.99
Batch size (K)	250
Buffer (D)	50000
Episodes (N)	70000
Time constant (κ)	0.005
Actor Network	
Activation Function	ReLU, Tanh
Layers	6
Neurons	(5, 250, 250, 120, 80, 1)
Critic Network	
Activation Function	ReLU, Linear
Layers	4
Neurons	(6, 250, 150, 1)
DNN training algorithm	
Optimizer	ADAM
Actor learning rate (α_a)	0.0035
Critic learning rate (α_c)	0.0035
Decay learning rate (ϑ)	0.9

NMPC and the limit of control increments for NMPC and DDPG were $-1 \leq \Delta T(k) \leq 1$.

The performance of the RL controller was compared to an NMPC previously developed by Lima et al. (2024). The controllers were applied to maintain the crystal mean volume size and the crystal mass in the targets by manipulating the temperature. They were tested for five cases as presented in Table 2. The conditions in Table 2 were considered to train the RL controller. Before starting the process control, the crystallization was maintained at the initial temperature for 60 min in all cases to promote the initial growth. After that, the control is started and the process is controlled for 100 min. A sampling time of one minute was adopted. The performance of the controllers was also tested considering a 5% disturbance in the concentration measurement at 110 min. All simulations were performed with a computer having the following specifications: Intel Core i7-12700, CPU 2.10 GHz, and 32 GB of RAM.

3. RESULTS AND DISCUSSION

The NMPC used in this work is the same developed by Lima et al. (2024) with prediction horizon equals to 10 and

Case	$\bar{L}_{30}^{sp}$ [μm]	m^{sp} [g]	S_0	T_0 [°C]
1	225	9.0	1.35	30.0
2	200	9.0	1.35	30.0
3	175	7.0	1.30	33.4
4	160	8.0	1.25	40.0
5	170	7.5	1.30	35.0

a control horizon of 5. Table 3 presents the performance of the RL controller and the NMPC for the five cases considered. Comparing the values of the controlled variables in the end of each batch, both approaches presented a close performance. The controllers were able to achieve values of the controlled variables close to the targets for all batches. Specially for Case 1, the NMPC presented an offset for both controlled variables, which is expected as only one manipulated variable is used to control two variables. On the other hand, the RL controller was able to reach values closer to the targets. Comparing all cases, Case 1 is the most difficult to control because it aims to produce the largest crystals with the highest mass, considering a high initial supersaturation and unseeded crystallization. Even with all these difficulties, the RL controller was able to perform better than the NMPC. Moreover, the RL controller imposed less temperature changes than the NMPC in most cases, as presented in Table 3.

Table 3. $\bar{L}_{30}$ and m values at the end of the controlled batches with the NMPC and the RL, and the total control effort.

Case	System	$\bar{L}_{30}$ (μm)	m (g)	$\sum \Delta u(k)^2$
	Target	225.0	9.00	-
1	NMPC	213.2	9.85	9.94
	RL	221.6	9.27	11.45
	Target	200.0	9.00	-
2	NMPC	197.9	9.15	11.73
	RL	202.2	9.12	5.29
	Target	175.0	7.00	-
3	NMPC	175.0	6.99	36.95
	RL	174.9	7.00	28.38
	Target	160.0	8.00	-
4	NMPC	160.0	7.99	163.92
	RL	159.1	8.03	22.61
	Target	170.0	7.50	-
5	NMPC	170.0	7.50	48.07
	RL	172.4	7.49	30.65

Fig. ?? presents the performance of the controllers from Case 1 of Table 2. Both controllers were able to achieve the setpoints and maintain the control variables in the desired values. The temperature changes imposed by the manipulated variables were similar, imposing a decrease in this value to promote crystal growth and nucleation, followed by an increase to reach the equilibrium. The RL controller presented an overshooting that was not observed for the NMPC.

Table 4 presents the mean time that each controller took to calculate the optimal control action for all simulations. For the five cases, the RL controller calculated the optimal control action much faster than the NMPC. Both controllers are able to calculate the control action fast enough for the sampling time of one minute. However, obtaining this value faster is a relevant advantage. In the crystalliza-

tion system used by Kim et al. (2023), the temperature change imposed by the controller is not done immediately. There is a dead time after imposing this temperature change. Therefore, calculating the control action faster is an advantage for a quicker change in the system.

Table 4. Mean time for the RL based controller and the NMPC to calculate the optimal temperature.

Case	NMPC time (s)	RL time (s)
1	5.0×10^{-2}	3.1×10^{-4}
2	1.6×10^{-1}	2.6×10^{-4}
3	1.0×10^{-1}	2.8×10^{-4}
4	1.5×10^{-1}	2.3×10^{-4}
5	2.2×10^{-1}	3.2×10^{-4}

The crystallization system used by Kim et al. (2023) to perform the crystallization batches includes attenuated total reflectance-Fourier transform infrared (ATR-FTIR). This equipment was used to get the absorbance and this value was used to obtain the concentration with a calibration curve. Situations like a change in the position of the ATR-FTIR probe can lead to disturbances in the concentration measurement. In order to check the performance of the controllers for this issues, we considered a perturbation of 5% in the concentration measurement at 110 min (i.e., $C(110) = C(110) + 0.05 \times C(110)$). Table 5 presents the performance of the controllers for the simulations account for disturbances in the concentration measurement. Compared to the values in Table 3, the performance of the controllers was similar accounting for disturbance. The RL controller imposed less temperature changes, as previously observed.

Table 5. $\bar{L}_{30}$ and m values at the end of the controlled batches with the NMPC and the RL, and the total control effort for the simulations with disturbance in the concentration measurement.

Case	System	$\bar{L}_{30}$ (μm)	m (g)	$\sum \Delta u(k)^2$
	Target	225.0	9.00	-
1	NMPC	213.3	9.85	13.20
	RL	220.8	9.32	12.55
	Target	200.0	9.00	-
2	NMPC	197.8	9.15	17.90
	RL	200.5	8.98	11.93
	Target	175.0	7.00	-
3	NMPC	175.0	6.99	38.80
	RL	177.1	7.28	28.94
	Target	160.0	8.00	-
4	NMPC	160.0	7.99	166.00
	RL	164.6	8.13	56.38
	Target	170.0	7.50	-
5	NMPC	170.0	7.50	49.97
	RL	173.3	7.62	31.63

Fig. 4 presents the performance of the controllers for Case 2 from Table 2 considering the disturbance in the concentration measurement. At 110 min, a change in the manipulated variable behavior can be observed. However, the performance of the controllers is still efficient to keep the controlled variables at the targets.

Table 6 presents the mean time demanded by each controller to obtain the control action for the five simulations with disturbance. The computational cost was similar to
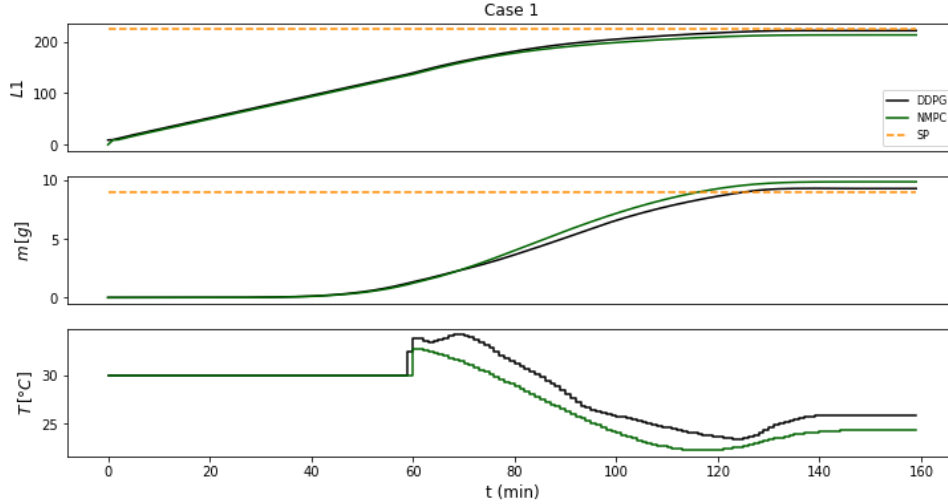
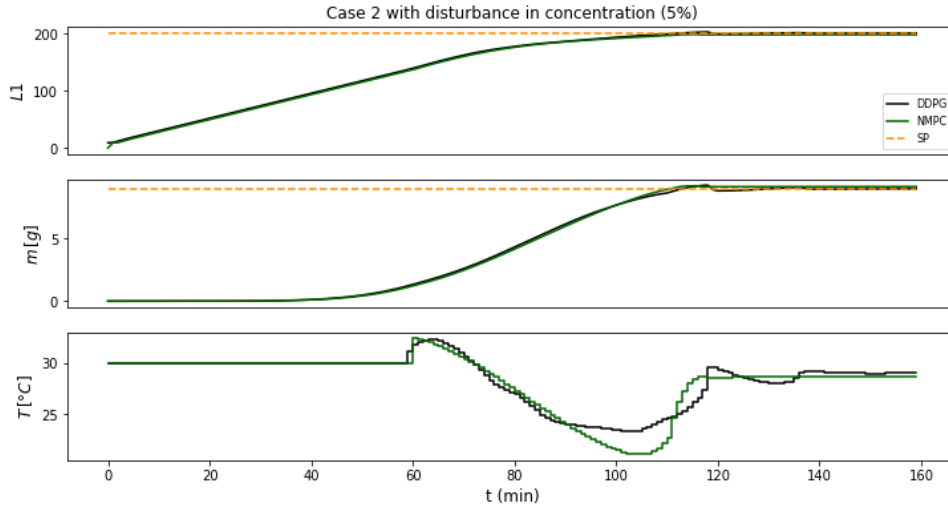Fig. 3. Performance of the NMPC and the RL for Case 1.

Fig. 4. Performance of the NMPC and the RL for Case 2 with disturbance in the concentration measurement.

what was previously presented in Table 4. Therefore, the NMPC and the RL controller presented efficient performance to control the unseeded batch crystallization of paracetamol in ethanol. However, the RL controller obtained the control action faster and the fewer temperature changes imposed, being a more suitable option to control this process.

Table 6. Mean time for the RL based controller and the NMPC to calculate the optimal temperature for the simulations with disturbance.

Case	NMPC time (s)	RL time (s)
1	3.8×10^{-2}	3.1×10^{-4}
2	1.9×10^{-1}	4.7×10^{-4}
3	9.2×10^{-2}	2.2×10^{-4}
4	1.5×10^{-1}	2.9×10^{-4}
5	2.9×10^{-1}	3.6×10^{-4}

4. CONCLUSION

In this work, we proposed a controller based on RL and showed its potential for controlling batch crystallization processes. The RL controller was used to control the crystal mean volume size and the crystal mass by manipulating the temperature and its performance was compared to an NMPC using a PBM as its internal model. Five cases were simulated, accounting for different setpoints and initial conditions. Both controllers were able to efficiently achieve the targets, even accounting for disturbance in the concentration measurement. However, the RL was able to calculate the optimal control action and imposed less temperature changes to the system compared to the NMPC. Therefore, the RL controller proved to be a good alternative to the NMPC and for real crystallization applications.

REFERENCES

Ahn, B., Bosetti, L., and Mazzotti, M. (2022). Secondary nucleation by interparticle energies. ii. kinetics. *Crystal Growth & Design*, 22(1), 74–86. doi: https://doi.org/10.1021/acs.cgd.1c00928.

Anandan, P.D., Rielly, C.D., and Benyahia, B. (2022). Optimal control policies of a crystallization process using inverse reinforcement learning. In L. Montastruc and S. Negny (eds.), *32nd European Symposium on Computer Aided Process Engineering*, volume 51 of *Computer Aided Chemical Engineering*, 1093–1098. Elsevier. doi:https://doi.org/10.1016/B978-0-323-95879-0.50183-1.

Bötschi, S., Rajagopalan, A.K., Morari, M., and Mazzotti, M. (2018). Feedback control for the size and shape evolution of needle-like crystals in suspension. i. concepts and simulation studies. *Crystal Growth & Design*, 18(8), 4470–4483. doi: https://doi.org/10.1021/acs.cgd.8b00473.

Cao, Y., Kang, J., Nagy, Z.K., and Laird, C.D. (2016). Parallel solution of robust nonlinear model predictive control problems in batch crystallization. *Processes*, 4(3), 20. doi:https://doi.org/10.3390/pr4030020.

Faria, R.R., Capron, B., Secchi, A., and Souza Jr., M. (2024a). A data-driven tracking control framework using physics-informed neural networks and deep reinforcement learning for dynamical systems. *Engineering Applications of Artificial Intelligence*, 127, 107256. doi: https://doi.org/10.1016/j.engappai.2023.107256.

Faria, R.R., Capron, B.D.O., de Souza Jr, M.B., and Secchi, A.R. (2023). One-layer real-time optimization using reinforcement learning: A review with guidelines. *Processes*, 11(1), 123. doi: https://doi.org/10.3390/pr11010123.

Faria, R.R., Capron, B.D.O., Secchi, A.R., and de Souza, M.B.J. (2024b). Gas-lift optimization using physics-informed deep reinforcement learning. *Industrial & Engineering Chemistry Research*, 63(32), 14199–14210. doi:10.1021/acs.iecr.3c04615.

Faria, R.R., Capron, B.D.O., Secchi, A.R., and de Souza Jr, M.B. (2022). Where reinforcement learning meets process control: Review and guidelines. *Processes*, 10(11), 2311. doi:https://doi.org/10.3390/pr10112311.

Griffin, D.J., Kawajiri, Y., Rousseau, R.W., and Grover, M.A. (2017). Using mc plots for control of paracetamol crystallization. *Chemical Engineering Science*, 164, 344–360. doi:https://doi.org/10.1016/j.ces.2017.01.065.

Grover, M.A., Griffin, D.J., Tang, X., Kim, Y., and Rousseau, R.W. (2020). Optimal feedback control of batch self-assembly processes using dynamic programming. *Journal of Process Control*, 88, 32–42. doi: https://doi.org/10.1016/j.jprocont.2020.01.013.

Kim, Y., Kawajiri, Y., Rousseau, R.W., and Grover, M.A. (2023). 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(6), 2866–2881. doi: https://doi.org/10.1021/acs.iecr.2c03465.

Lillicrap, T. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*. doi:https://doi.org/10.48550/arXiv.1509.02971.

Lima, F.A.R.D., de Moraes, M.G.F., Grover, M.A., Barreto Junior, A.G., Secchi, A.R., and de Souza, M.B.J. (2024). Neural network inverse model controllers for paracetamol unseeded batch cooling crystallization. *Industrial & Engineering Chemistry Research*. doi: https://doi.org/10.1021/acs.iecr.4c02060.

Manee, V., Baratti, R., and Romagnoli, J.A. (2021). Optimal strategies to control particle size and variance in antisolvent crystallization operations using deep rl. *Chemical Engineering Transactions*, 86, 943–948. doi: https://doi.org/10.3303/CET2186158.

Manee, V., Baratti, R., and Romagnoli, J.A. (2022). Learning to navigate a crystallization model with deep reinforcement learning. *Chemical Engineering Research and Design*, 178, 111–123. doi: https://doi.org/10.1016/j.cherd.2021.12.005.

Meng, Q., Anandan, P.D., Rielly, C.D., and Benyahia, B. (2023). Multi-agent reinforcement learning and rl-based adaptive pid control of crystallization processes. In A.C. Kokossis, M.C. Georgiadis, and E. Pistikopoulos (eds.), *33rd European Symposium on Computer Aided Process Engineering*, volume 52 of *Computer Aided Chemical Engineering*, 1667–1672. Elsevier. doi: https://doi.org/10.1016/B978-0-443-15274-0.50265-1.

Moraes, M.G.F., Lima, F.A.R.D., Lage, P.L.d.C., de Souza, M.B.J., Barreto, A.G.J., and Secchi, A.R. (2023). 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(24), 9515–9532. doi: https://doi.org/10.1021/acs.iecr.3c00739.

Nagy, Z.K. and Braatz, R.D. (2012). Advances and New Directions in Crystallization Control. *Annual Review of Chemical and Biomolecular Engineering*, 3(1), 55–75. doi:10.1146/annurev-chembioeng-062011-081043.

Nayhouse, M., Tran, A., Kwon, J.S.I., Crose, M., Orkoulas, G., and Christofides, P.D. (2015). Modeling and control of ibuprofen crystal growth and size distribution. *Chemical Engineering Science*, 134, 414–422. doi: https://doi.org/10.1016/j.ces.2015.05.033.

Sitapure, N. and Kwon, J.S.I. (2023). Model predictive control of cadmium telluride (cdte) quantum dot (qd) crystallization. In *2023 American Control Conference (ACC)*, 3251–3256. doi: https://doi.org/10.23919/ACC55779.2023.10156018.

Sutton, R.S. and Barto, A.G. (2018). *Reinforcement learning: An introduction*. MIT press.

Szilágyi, B., Agachi, P., and Nagy, Z.K. (2018). Chord length distribution based modeling and adaptive model predictive control of batch crystallization processes using high fidelity full population balance models. *Industrial & Engineering Chemistry Research*, 57(9), 3320–3332. doi:https://doi.org/10.1021/acs.iecr.7b03964.

Zhang, J., Yang, J., Zhang, Y., and Bevan, M.A. (2020). Controlling colloidal crystals via morphing energy landscapes and reinforcement learning. *Science advances*, 6(48), eabd6716. doi: https://doi.org/10.1126/sciadv.abd6716.