

<https://doi.org/10.1038/s44334-024-00010-z>

# Energy-efficient and quality-focused control of conveyor belt dryers in petrochemical production

Check for updates

Muhammad Waseem<sup>1</sup>, Kshitij Bhatta<sup>1</sup>, Chen Li<sup>1</sup>, Nabeel Haider<sup>2</sup> & Qing Chang<sup>1</sup> ✉

The petrochemical industry's dryers, which operate at high temperatures, consume substantial energy and require precise temperature control to maintain product quality. To address these challenges, conveyor belt dryers are used for continuous drying, but optimizing their operation is complex. Traditional control methods often rely on operator experience or fixed settings, which may not be optimal. This paper introduces a novel approach using a Graph Neural Network and Multiagent Deep Deterministic Policy Gradient algorithm. This method models dryers' components—chambers, fans, and belts—as cooperating agents, optimizing temperature, fan speed, and belt speed based on quality predictions. The algorithm, trained with industrial data, improves decision-making by integrating quality predictions into its reward function. Extensive experiments show a 5% improvement in product quality and a 3% reduction in energy consumption per time step, translating to significant energy savings. This approach enhances both efficiency and product quality control in conveyor belt dryers.

The petrochemical industry is not only crucial for national production and daily life but is also characterized by high pollution and carbon emissions. Within petrochemical plants, the drying process stands out as a major consumer of energy, constituting 10–20% of the total energy consumption<sup>1</sup>. It goes beyond being a mere production stage; instead, it plays a vital role in determining both product quality and energy utilization. Numerous dryers have been discussed in the literature based on specific applications and products to be dried, including spray dryers, rotary dryers, and belt dryers<sup>2</sup>. While each type has its advantages and disadvantages, a conveyor belt dryer emerges as a versatile and widely adopted solution, particularly in petrochemical drying, where the product is typically in granular form, requiring continuous and uniform processing<sup>3</sup>. Effectively addressing the challenge of implementing reliable process control in petrochemical drying is crucial for minimizing energy losses and ensuring product quality.

In conveyor belt dryers, attaining the desired product quality with optimal energy consumption depends on maintaining a delicate balance among various contributing factors such as temperature, load, and speed, among others. Managing these parameters is a challenging task, compounded by uncontrollable climatic conditions that can substantially influence the impact of these controllable factors. Moreover, while the energy consumption profile can be estimated analytically<sup>4</sup>, predicting product quality is not as straightforward due to the intricate non-linear relationships among these contributing factors.

Common control methods in petrochemical drying still rely on manual empirical operation, which is subjective and demands specialized operators. Alternative automatic control approaches include feedback control, computer simulation control, and model predictive control (MPC). Feedback control, exemplified by proportional, integral, and derivative (PID) tuning, adjusts the drying process based on real-time product moisture detection to minimize deviations<sup>5</sup>. However, it may not align well with independent events of the inlet product, drying conditions, and outlet product. Computer simulation control utilizes AI algorithms to model and solve non-linear processes<sup>6</sup>, yet obtaining effective coefficients is challenging due to the randomness of in-field drying. Boltzmann coefficient generators<sup>7</sup> have been developed to generate real-time coefficients based on field measurements, but this significantly increases control costs and poses challenges in accurately measuring data online, especially in the presence of large-scale dynamic disturbances. MPC establishes an analytical model of the drying system, incorporates disturbance information through online optimization, and calculates optimal future inputs based on the current state<sup>8,9</sup>. It addresses time delay effects and is suitable for controlling drying with high inertia, non-linearity, and multiple disturbances during the dehydration process. However, these approaches fall short of achieving complex control objectives, particularly when influenced by factors such as material properties, climatic conditions, and the drying process itself.

<sup>1</sup>Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA, 22904, USA. <sup>2</sup>DYNEA Pakistan Petrochemicals, Gadoon Amazai Topi, 23454 Swabi, Pakistan. ✉e-mail: [qc9nq@virginia.edu](mailto:qc9nq@virginia.edu)

Machine learning (ML), particularly reinforcement learning (RL) and deep reinforcement learning (DRL) has emerged as a transformative approach for addressing complex control problem<sup>10,11</sup>. Their capacity to learn from data, adapt to changing conditions, and optimize outcomes makes them well-suited for the intricate challenges posed by conveyor belt dryers. However, employing these directly for control purposes is not straightforward. The complexities lie in balancing both the product quality and the energy consumption profile, demanding a tailored approach for algorithmic modification to meet the specific control requirements. Simply applying these algorithms without customization might overlook the nuanced balance necessary for optimizing both quality and energy efficiency. Therefore, this study integrates a graph convolution network (GCN) model with the multiagent deep deterministic policy gradient (MADDPG) algorithm. The GCN model is first trained with data derived from industrial drying lines and then validated with real data from a petrochemical dryer. After validating the model, it becomes an integral part of the overall control framework, and the MADDPG algorithm is employed to control the drying process. The trained GCN model is incorporated into the MADDPG algorithm's reward function to estimate the product quality and energy consumption after each time step. The primary contributions of this paper encompass:

- Establish a comprehensive model that utilizes a GCN framework, drawing upon real dryer data. The model integrates a detailed description of a conveyor belt dryer, encompassing its dynamic modeling, to enable accurate predictions of product quality in response to fluctuating input conditions.
- Formulate the drying process as a multiagent control problem, enabling real-time adjustment of process parameters in response to changes in climate conditions. This approach ensures energy-efficient production while adhering to quality standards.
- Utilize the MADDPG algorithm to enhance energy efficiency in the conveyor belt dryer, treating fans, chambers, and belts as distinct agents, to improve the efficiency of the control.

The rest of the paper is organized as follows: “Results” presents the results, “Discussion” offers a comprehensive discussion, and “Methods” outlines the methods for evaluating quality, energy consumption, and problem formulation.

## Results

In this section, we conduct simulation experiments to validate the effectiveness of the GCN modeling and control framework. The experiments involve comparing the production output (dried product) by employing the control policy learned from the proposed method and the output without control. The algorithm is evaluated based on the production quality (samples within the desired product weight range  $X$ ) and the associated energy consumption of the chambers and conveyor belts. We aim to demonstrate three key findings: (1) The GCN model accurately predicts the sample product weight (2) The proposed MADDPG-based control framework results in higher yield, and (3) reduces the total energy consumption.

### Experiment parameter setting

The simulation experiments are conducted using the Python platform. We create and analyze 50 distinct conveyor belt dryers, which encompass various types such as food drying, grain drying, and chemical drying, among others. Through these experiments, we can verify the effectiveness and resilience of the proposed method across the 50 diverse drying systems. Each dryer is constructed by randomly selecting from the following sets:

$$\text{No. of fans, } f \in \{2, 3, 4, 5, 6\}$$

$$\text{No. of conveyor belts, } b \in \{1, 2, 3, 4, 5, 6\}$$

$$\text{No. of chambers, } C \in \{3, 4, \dots, 18\}$$

$$\text{Fan speed, } v_{f_i} \in [700, 1400], i = 1, 2, \dots, n$$

$$\text{Conveyor belt speed, } v_{b_i} \in [10, 20], i = 1, 2, \dots, n$$

$$\text{Chamber temperature, } t_{C_i} \in [60, 110], i = 1, 2, \dots, n$$

$$\text{Ambient temperature, } t_{amb} \in [15, 50]$$

$$\text{Average humidity, } h \in [15, 80]$$

To illustrate the application of the proposed method, a comprehensive numerical case study involving a conveyor belt dryer is provided. This dryer consists of 9 chambers, 3 fans, and 3 conveyor belts, similar to the configuration shown in Fig. 5. The specific parameters for this setup are presented in Tables 1–4.

### Industrial data

In this research, we employ data derived from a petrochemical plant<sup>12</sup> engaged in the processing of molding compound powder (MCP). It's important to note that the data exclusively pertains to the drying process associated with MCP. Currently, within the industry, a conventional approach is utilized to optimize the drying process. This approach relies on operators' experience, wherein dryer parameters are initially set and subsequently adhered to. While this methodology generally aligns with quality standards, it falls short when it comes to guaranteeing product quality under sudden fluctuations in uncontrollable parameters, such as ambient temperature and humidity. Furthermore, there hasn't been any prior exploration into issues related to energy consumption or production optimization.

For analysis and control framework training, we have gathered a dataset encompassing 1000 hours of operational data from the aforementioned dryer system. Each data sample is collected following a time interval of 2 hours. This dataset encompasses information regarding the speeds of three conveyor belts, three fans, and the temperatures of nine drying chambers. The conveyor belts maintain consistent speeds of 40, 40, and 50 RPM, while the temperatures in the drying chambers are carefully controlled within the range of 80–102 °C. Notably, the temperature tends to be higher in the initial chambers and gradually decreases toward the end of the dryer system. Similarly, the fans can operate within a speed range of 800–1200 RPM. These variables are adjustable, falling under the category of controllable parameters, and in typical weather conditions, they ensure the desired product quality. However, it's important to acknowledge that sudden changes in climatic conditions, i.e., ambient temperature and humidity, can lead to variations in product quality. While it may be feasible to manage them under strict conditions, in the situation at hand, these variables remain uncontrolled and can fluctuate due to changing weather conditions. Therefore, addressing these variations through adjustments in controllable parameters, such as conveyor belt speeds, chamber temperatures, and fan speeds, becomes imperative.

Based on the available data for the dryer system, there are 15 controllable parameters, as outlined in the adjacency matrix in Table 7, along with two uncontrollable parameters: humidity and ambient temperature. For all 17 parameters, the output, which is the product weight, is observed. At each time point, the product weights are recorded based on the specific settings of these parameters. While humidity and ambient temperature cannot be controlled, the 15 controllable parameters are adjusted to keep the product within the desired weight range of 120–130 grams. Any deviation outside this range results in a product that does not meet quality standards and is classified as scrap. The primary control objective is to maintain product quality by adhering to the specified weight standards while minimizing energy consumption.

### GCN- MADDPG training

Utilizing the specified system parameters, the proposed method involves two key training phases: GCN training and MADDPG training. First, the

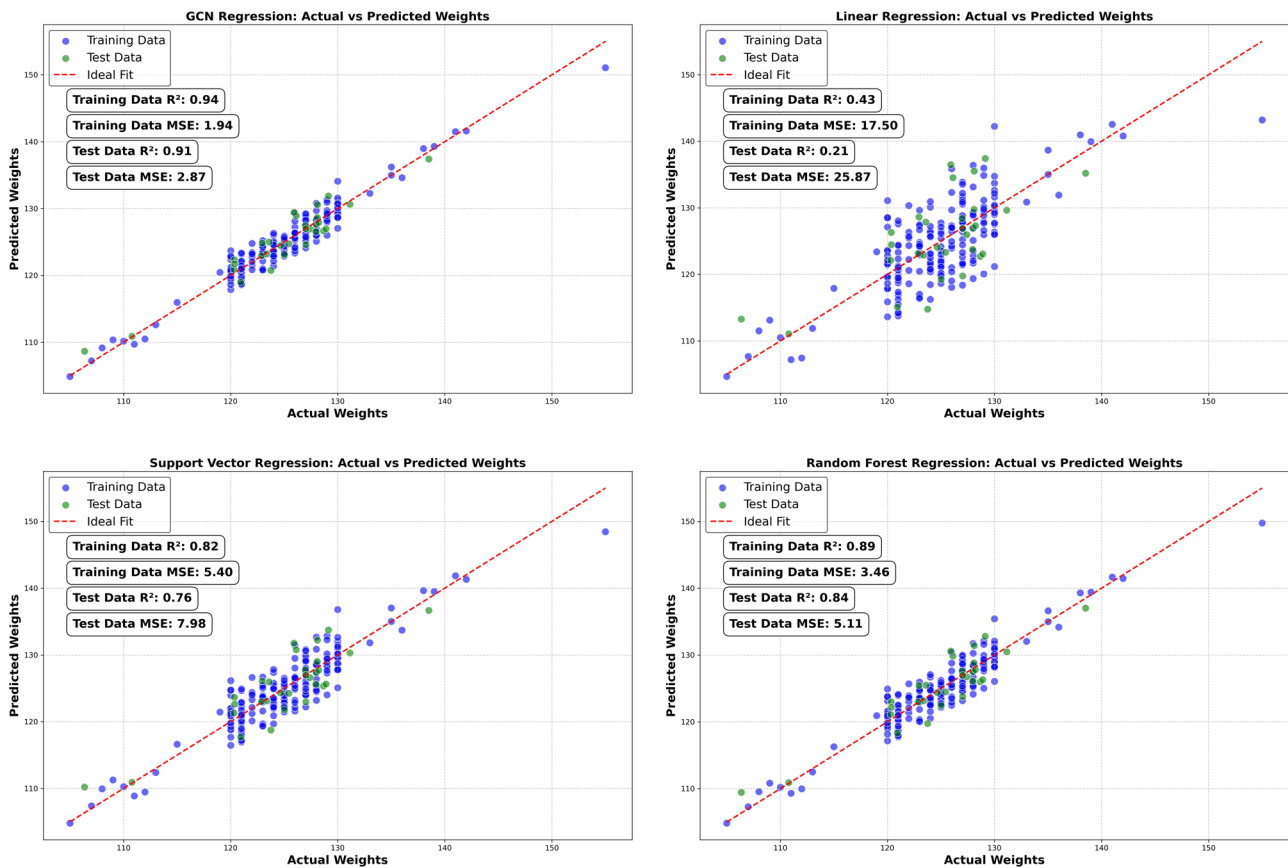


Fig. 1 | Comparison of prediction models for performance and accuracy. a GCN prediction model, b Linear regression model, c Support vector regression model, d Random forest regression model.

Table 1 | Parameters of the fans

Parameters	$f_1$	$f_2$	$f_3$
Initial speed (RPM)	1000	1000	1000

Table 2 | Parameters of the Conveyor belts

Parameters	$b_1$	$b_2$	$b_3$
Speed (m/sec)	1.25	1.25	1.32
Length (m)	7	7	7
Width (m)	2	2	2
Drive roller diameter (m)	0.3	0.3	0.3

Table 3 | Parameters of the drying chambers

Parameters	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$
Initial temperature (°C)	102	101	100	100	99	95	90	85	80

GCN model is trained offline using raw labeled data from industrial dryers to estimate product quality, specifically the sampled product weight.

To validate the effectiveness of GCN model in prediction, a comparative study is conducted. As shown in Fig. 1, the training dataset comprises 200 production samples, each taken over a two-hour period. During this time, the dryer operates under a specific set of input parameters, and at the end of two hours, a product sample is tested to measure its weight. The predicted product weight is then compared to the actual sample weight.

Table 4 | Other parameters

Parameter	Value
$C_{est}$	0.05 kWh
Timestep, $t$	2 hr
$\eta$	70%
$h_1$ (m)	0.076
$\gamma$	0.95
$\alpha_a$	0.0005
$\alpha_c$	0.001
$X$	[120, 130]

To evaluate the performance of the proposed GCN regression model, it is compared with several other regression models: linear regression (LR), support vector regression (SVR), and random forest regression (RFR). The GCN regression model achieved the best results, with an R<sup>2</sup> score of 0.91 and a mean squared error (MSE) of 2.87. This indicates that the GCN model explains a large proportion of the variance in the data and provides highly accurate predictions, suggesting its effectiveness in capturing complex relationships within the data. In contrast, LR performs poorly with an R<sup>2</sup> score of 0.21 and an MSE of 25.87, which can be attributed to its limited ability to handle the non-linear dependencies in the data. SVR exhibits a moderate performance with an R<sup>2</sup> score of 0.76 and an MSE of 7.98, indicating that while it handles non-linearity better than LR, it still falls short compared to the GCN model. RFR also performs well with an R<sup>2</sup> score of 0.84 and an MSE of 5.11, demonstrating its capability to capture non-linear relationships but not reaching the accuracy level of the GCN model. Overall,

the GCN model outperforms the others, highlighting its superior ability to model the data accurately.

Once the GCN model is trained, it is incorporated into the reward function of the MADDPG algorithm. This integration allows the GCN model to guide the MADDPG algorithm in optimizing the control of the dryer system to achieve the desired product quality.

The MADDPG algorithm is then trained in an online manner to continuously adjust control inputs in the dryer, considering the estimated product quality. Control decisions are activated at each time step  $t$  throughout the training process. The neural network architecture includes two fully connected hidden layers, each containing 64 hidden units. The reward function conforms to Equation (19), and the necessary training parameters are detailed in Table 4.

Upon completing the training process, a machine control policy is derived based on the updated network parameter  $\theta$ . Figure 2 shows the convergence curve for the proposed GCN-MADDPG algorithm. It is observed that achieving a stable policy takes around 25 episodes, where each episode corresponds to 24 hours of drying time.

### Yield comparison

As our primary focus extends beyond energy consumption to encompass system yield, Fig. 3 provides a comparison between the actual product quality output and that achieved under trained policy. This comparison involves the examination of sample weights obtained from both actual data and the simulation model trained through the proposed method. As previously discussed, disruptions to the dryer setting occur when there are

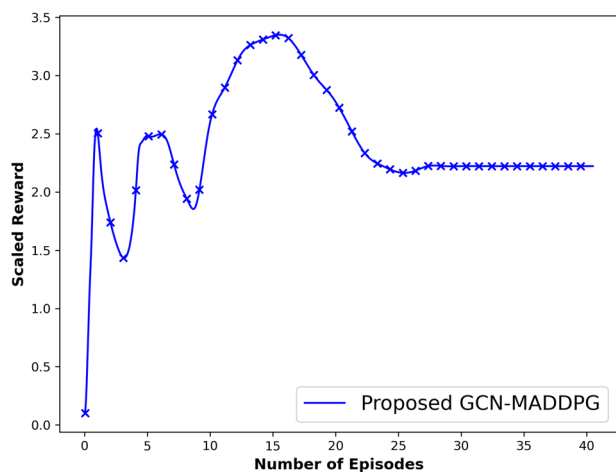


Fig. 2 | Training data, reward versus time.

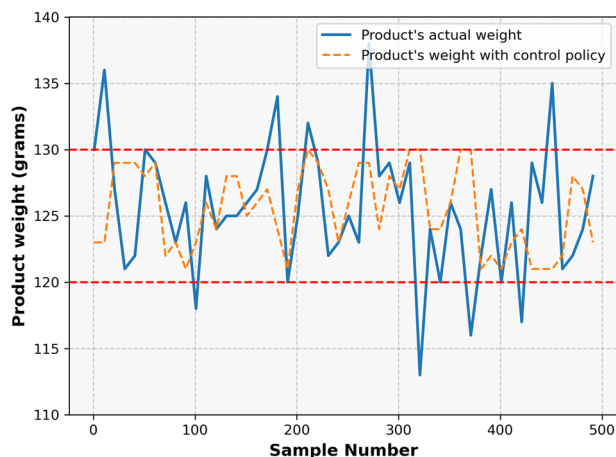


Fig. 3 | Quality comparison between actual and controlled parameters.

unforeseen changes in uncontrollable parameters. In such instances, the product quality, quantified as the sample weight recorded after each time step  $t$ , may deviate from the desired range  $X$ , resulting in subpar outcomes or scrap. However, our proposed algorithm demonstrates adaptability by adjusting controllable parameters, i.e., chamber temperatures, fan speeds, and conveyor belt speeds, to counteract these unexpected variations. This adjustment mechanism is designed to maintain product quality within the desired range, ultimately enhancing the system yield.

Table 5 displays the outcomes over 500 timesteps, representing 500 samples, where 25 samples were rejected due to deviating from the desired quality range 'X' under actual operational conditions. Conversely, by utilizing the trained policy, no samples were rejected. This result emphasizes the system's adeptness in dynamically adjusting controllable parameters, ensuring consistent product quality within the desired range and effectively preventing any scrap production.

### Energy comparison

Utilizing Eq. (14) as our basis, we calculate the energy consumption attributed to temperature fluctuations. This equation indicates an average energy consumption of 47.5 kWh for the chambers, based on real-world data obtained from the industrial dryer. However, under the trained policy, the average energy consumption notably reduces to approximately 45 kWh. This reduction signifies the achievement of a more optimized and efficient performance under controlled conditions.

Equation (15) offers a stable estimation of conveyor belt energy consumption, particularly in scenarios featuring constant speed and uniform load variations. However, energy consumption may exhibit variance under controlled settings, contingent upon the control policy. The pivotal role of Eqs. (15), (16), and (7) in determining this variance is emphasized. Table 6 presents a detailed breakdown of calculations, including mass computed via Eq. (7) and load determined by the product of mass and gravitational acceleration ( $g = 10 \text{ m/s}^2$ ). Utilizing data from Tables 2 and 4, Eqs. (15) and (16) facilitate the computation of  $P_{mech}$  and  $P_{input}$ . Furthermore,  $P_{input}$  is converted to kilowatts ( $1 \text{ kW} = 1000 \text{ Nm/s}$ ) to measure the belt energy consumption over 1 hour, adjusted for a 2-hour time step. This table serves as a valuable reference, elucidating the energy dynamics within the system.

Table 6 not only provides insights into energy consumption under actual (industrial) conditions but also presents data reflecting the average load and speeds observed when operating under the trained policy. These observations serve as the basis for calculating energy consumption. Specifically, the average speeds recorded for belts 1, 2, and 3 under the controlled policy are  $\sim 1.15 \text{ m/s}$ ,  $1.15 \text{ m/s}$ , and  $1.17 \text{ m/s}$ , respectively.

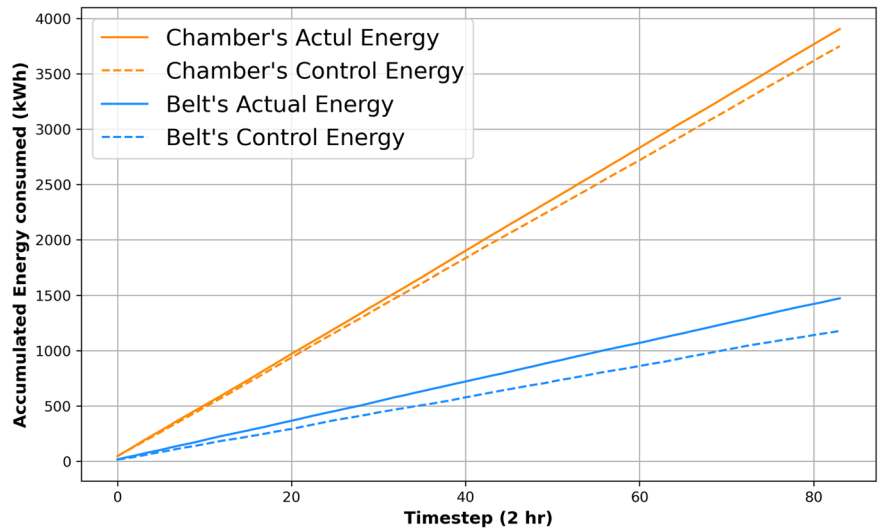
In Fig. 4, we offer a comprehensive comparison of the accumulated energy consumption between chambers and belts under real-world (industrial) conditions and controlled settings. This comparison relies on average data collected per time step over a week (168 hours). Notably, there's an average 3% reduction in energy consumption per time step observed for both chambers and belts. Over the week-long comparison period, this translates to an average energy savings of 154 kWh for controlled chambers and 295 kWh for controlled belts, respectively. It's important to note that the cumulative nature of energy consumption over time can result in a broad y-axis range, making it challenging to distinguish minor differences between two values.

It is important to acknowledge certain limitations in our work. We did not consider the impact of motors' random disruptions in conveyor belt dryers, which can potentially affect both product quality and energy consumption. Additionally, we overlooked energy losses, particularly those arising from electric motors and within drying chambers. In future research, we aim to address these factors and explore operations under conditions of random disturbance to further enhance the efficiency and robustness of conveyor belt drying systems.

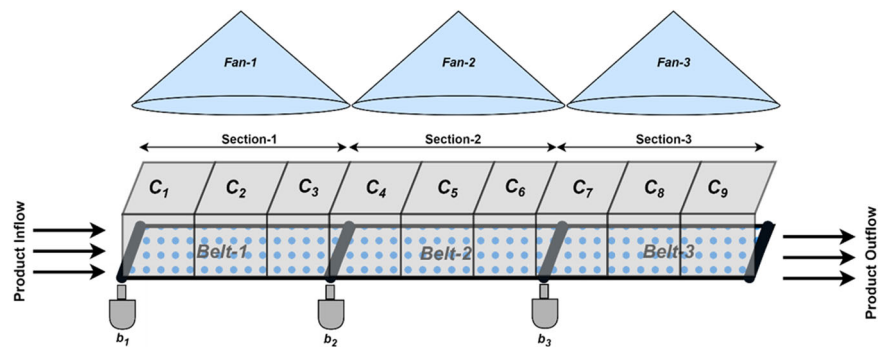
### Discussion

This research investigates conveyor belt dryers, focusing specifically on their applications within the petrochemical industry. It entails gathering

**Fig. 4** | Comparing energy consumption under an actual and controlled setting.



**Fig. 5** | Structure of a conveyor belt dryer.



and analyzing real-world data derived from industrial conveyor belt dryers. Given the intricate interplay among various drying variables—such as chamber temperatures, conveyor belt speeds, fan speeds, and product quality—the study adopts a machine-learning approach, centering around a GCN modified for regression tasks. The GCN model undergoes training on industrial data to predict product quality based on input parameters. It's imperative to note that product quality isn't solely dependent on controllable parameters. Uncontrollable factors like ambient temperature and humidity also exert a significant influence. Abrupt shifts in these uncontrollable variables can push product quality beyond acceptable limits. To mitigate this, a MADDPG algorithm is employed for real-time control amid such uncertainties. The MADDPG algorithm, well-suited for managing extensive state-action spaces and continuous action values, treats each drying chamber, conveyor belt, and fan as an autonomous agent. Moreover, energy consumption emerges as a significant concern within the drying process. The heat sources in the chambers and the power driving the conveyor belts constitute major energy consumers in conveyor belt dryers. Hence, the energy associated with these aspects is monitored and analyzed. The MADDPG algorithm plays a pivotal role in controlling both product quality and energy consumption within the conveyor belt drying system. To evaluate the effectiveness of the control algorithm, a case study is conducted, configuring a conveyor belt dryer identical to the industrial setup. The results are then compared, encompassing production quality and energy consumption, between the actual industrial scenario and the controlled setting. Encouragingly, the controlled algorithm yields notable reductions in total energy consumption and improvements in product quality.

## Methods

### System description

An industrial conveyor belt dryer is constructed using a series of drying chambers, with each drying chamber serving as the fundamental module that forms the entire system. To optimize performance, these drying chambers are grouped to create drying sections. Within a drying section, all chambers share a common conveyor belt, which evenly distributes the product to be dried as it enters the section. Naturally, the product undergoes redistribution as it exits one drying section and enters the subsequent one.

For illustration purposes, a conveyor dryer comprising three belt conveyors  $b_i$ , each run by an electric motor is given in Fig. 5. The dryer is divided into 9 chambers, represented as  $C_i$  and each belt  $b_i$  is associated with three chambers (or one section) and one fan  $F_i$ . The product arrives constantly at an arrival rate  $q_t$  and each belt  $b_i$  moving at speed  $v_{b_i}^t$  feeds to the next one. Each chamber  $C_i$  has an individual temperature setting  $t_i$ . The fans can operate at a variable speed  $v_{F_i}^t$ .

In this system, the following notations are used.

1.  $C_i$ , where  $i = 1, 2, \dots, n$  represents the  $i^{\text{th}}$  Chamber;
2.  $b_i$ , where  $i = 1, 2, \dots, n$  represents a motor controlling the  $i^{\text{th}}$  belt;
3.  $F_i$ , where  $i = 1, 2, \dots, n$  denotes the  $i^{\text{th}}$  fan;
4.  $v_{b_i}^t$  represents the speed of the conveyor belt  $b_i$  at time  $t$ ;
5.  $v_{F_i}^t$  represents the speed of the fan  $F_i$  at time  $t$ ;
6.  $t_{C_i}^t$  is the temperature of the chamber  $C_i$  at time  $t$
7.  $e_{b_i}^t$  is the energy consumption rate of the belt  $b_i$  at time  $t$ ;
8.  $H_t$  is the average humidity in the dryer at time  $t$ ;
9.  $t_{amb}$  is the ambient temperature;

10.  $q_i^t$  is the product arrival rate of the belt  $b_i$  at time  $t$ ;
11.  $L_i$  is the length of the conveyor belt  $b_i$ ;
12.  $w$  is the width of each belt  $b_i$  and remains constant for all the belts;
13.  $h_i^t$  is the thickness of the product layer on the belt  $b_i$  at time  $t$ ;
14.  $c$  is the junction point on each conveyor belt which defines the starting and ending of the conveyor's length. It joins the length of a conveyor into a loop. The arrival rate  $q_t$  and belt speed  $v_{b_i}^t$  can change only at point  $c$ ;
15.  $Th_{b_i}^t$  is the throughput of the conveyor belt  $b_i$  at time  $t$ ;
16.  $Q_i^t$  denotes the amount of product processed by the belt  $b_i$  at time  $t$ ;
17.  $d_i^t$  represents the product departing rate of the conveyor belt  $b_i$  at time  $t$ ;
18.  $q_{ev}^i$  represents the amount of solvent evaporated in section  $i$ ;
19.  $c_{et}$  is the cost of energy associated with unit degree change in the chamber's temperature;
20.  $P_{input}$  represents the electric input power;
21.  $P_{mech}$  represents the mechanical input power;
22.  $\eta$  is the efficiency of an electric motor;
23.  $o_i^t$  represents the observation of agent  $i$  at time  $t$ ;
24.  $X$  denotes an acceptable range for the sample weight i.e.,  $X = [\min, \max]$ ;

The product arrives at an arrival rate  $q_t$  to the first conveyor belt  $b_1$ . As the product arrives continuously and concurrently with the conveyor's movement, it accumulates as a layer with a thickness of  $h$  on the conveyor belt. The processed amount  $Q_i(x)$  by belt  $b_i$  at position  $x$  is the product of speed, width, and thickness:

$$Q_i(x) = v_{b_i} h_i(x) w \tag{1}$$

Let  $q$  be the product arrival rate,  $t$  be the time,  $w$  be the width of the conveyor,  $v$  be the speed of the conveyor belt, and  $h(x)$  be the initial thickness of the product layer at position  $x$  along the conveyor. The integral of the processed amount over the conveyor length should equal the total product input:

$$\int_0^L Q(x) dx = qt \tag{2}$$

$$\int_0^L (vh(x)w) dx = qt \tag{3}$$

With constant speed and width,

$$vw \int_0^L h(x) dx = qt \tag{4}$$

$$\int_0^L h(x) dx = (qt)/(vw) \tag{5}$$

where  $L$  is the length of the conveyor.

Equation (5) implies that the integral of thickness over the conveyor length is constant for a given product arrival rate, time, speed, and conveyor width.

Similarly, let  $q_1^t$  and  $q_2^t$  be the product arrival rate to conveyors  $b_1$  and  $b_2$  at time  $t$  respectively;  $\rho$  be the density of the product, the conservation of mass equation for a small-time interval  $\Delta t$  can be written as

$$q_1(t)\Delta t = q_2(t)\Delta t \tag{6}$$

where  $q_i = A_i v_i \rho_i$  and  $A$  is the cross-section area ( $A = wh$ ). Based on the conservation of mass equation, the input i.e., product arrival rate  $q_i^t$  should be equal to the product departing rate  $d_i^t$ . Since all the conveyors are interlinked and the product departing rate of conveyor belt  $b_i$  is the product

arrival rate of the conveyor belt  $b_{i+1}$  i.e.,  $q_{i+1}^t = d_i^t$ , therefore

$$\begin{aligned} q_1^t &= q_2^t \\ A_1 v_1^t \rho_1 &= A_2 v_2^t \rho_2 \\ (wh_1^t) v_1^t \rho_1 &= (wh_2^t) v_2^t \rho_2 \end{aligned} \tag{7}$$

Since the product is being heated along the conveyor, the density of the same product on the conveyor  $b_i$  will always be greater than that of the conveyor belt  $b_{i+1}$  because of the continuous evaporation. Let the amount of evaporation in section  $i$  be represented by a constant  $q_{ev}^i$ , we can write that  $\rho_i = \rho_{i+1} + q_{ev}^i$ . Furthermore, our throughput represents the volume of the product rather than the mass, so we neglect the density variations while estimating the throughput.

$$h_1^t v_1^t = h_2^t v_2^t \tag{8}$$

$$\frac{h_1^t}{h_2^t} = \frac{v_2^t}{v_1^t} \tag{9}$$

Thus, we can see that the speed and layer thickness are inversely related, and the variation in one is balanced by the other. Hence, the overall quantity remains the same i.e.,

$$Th_{b_1}^t = Th_{b_2}^t = \dots = Th_{b_n}^t = Q_i^t \tag{10}$$

At each time step  $t$ , the throughput (processed product) of  $i^{th}$  conveyor belt  $Q_i^t$  can be simply calculated based on Eq. (1). Similarly, once  $Q_i^t$  is found, we can use  $q_i = wh_i v_i$  to find the thickness of the product layer  $h_i$  on belt  $b_i$ . In the dryer system modeling, production throughput is not the sole critical parameter; the quality of the output is equally significant. Typically, quality is measured by assessing the dryness level of the product. To evaluate this, the weight of the product is measured at the end of each time step  $t$ . If the product weight falls within a predefined range  $X$ , the product batch (where one batch corresponds to the quantity processed in a single time step,  $t$ ) is considered acceptable. Otherwise, it necessitates reprocessing through the dryer. It is worth noting that this paper does not delve into the details of the reprocessing, and any product batch failing to meet the quality standard is treated as scrap.

Furthermore, it is important to note that although the quantity of product, denoted as  $Q_i^t$  remains constant, the load itself changes as the product progresses through the dryer. This change in load arises from the fact that, as the product traverses the dryer, its mass decreases due to evaporation, subsequently resulting in variations in density. To accurately estimate this density variation, we leverage industrial data and chart the density fluctuations across the drying chambers, as exemplified in Fig. 6. This visual representation of density changes serves as a crucial reference point for estimating the amount of mass that undergoes evaporation. This information, in turn, aids in determining the load on the succeeding conveyor belt, which is used to estimate the belt's energy consumption in "GCN Modeling".

Since the load depends on the product mass, i.e.,  $load = mg$ , Eqs. (7–10) are used to estimate the mass i.e.,  $q_i = wh_i v_i \rho_i$ , where  $\rho = m/V$ . This will be further used in "GCN Modeling" to estimate the energy.

Building upon the insights gleaned from the density transitions depicted in Fig. 6, we can discern that the average density difference between two successive sections amounts to approximately  $350 \frac{kg}{m^3}$ . In practical terms, this signifies that, on average, each section undergoes evaporation resulting in the removal of ~350 kg of solvents. Subsequently, for section  $i + 1$ , we can derive the density, denoted as  $\rho_{i+1}$ , by subtracting this value from the density of the previous section,  $\rho_i$ . This relationship is expressed as  $\rho_{i+1} = \rho_i - 350$ . Additionally, it is prudent to assume that the capacity of the conveyor motors exceeds the maximum load by a substantial margin, specifically at least 30%. This precautionary measure ensures that the conveyor system can

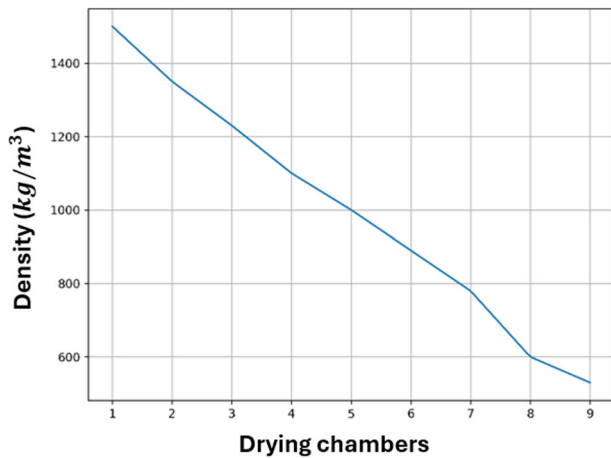


Fig. 6 | Density variation along the drying chambers.

Table 5 | Production comparison under an actual and controlled setting

No. of batches	Actual setting	Controlled setting
Total batches	500	500
Rejected batches	25	0

accommodate the variability in loads and operate efficiently under different conditions.

### GCN modeling

In the existing literature, rigorous modeling techniques based on the fundamental principles of heat and mass transfer are employed to estimate product quality<sup>4,13</sup>. However, these models often fall short of establishing a comprehensive model by incorporating environmental conditions and process parameters such as temperature, humidity, fan speed, conveyor belt speed, and product quality due to complexities among these factors. Therefore, in this paper, we adopt a machine-learning approach to establish a sophisticated model to describe the connections between these input parameters and sample product quality. Specifically, we employ a GCN model<sup>14</sup>, to predict the weight at specific sets of input parameters.

The choice of GCN to model the system is driven by the data’s inherent graph-like structure, the interactions between components in the system, and the suitability of this approach for facilitating the control of the system. In this framework, system components such as fans, chambers, and belts are represented as distinct nodes within a graph, which can later be treated as agents for control purposes. The relationships and interdependencies among these agents are identified and captured in an adjacency matrix, as shown in Table 7. This table outlines the interconnections between the 15 nodes:  $B_i$  represents the three belts,  $C_i$  nine chambers, and  $F_i$  the three fans. A value of 1 in the matrix indicates a relationship between two nodes, while a 0 denotes no connection.

The adjacency matrix is then used to construct a graphical representation of the system. Each graph corresponds to a specific combination of input parameters, such as fan speeds, chamber temperatures, and conveyor belt speeds. Additionally, each graph is paired with a corresponding label indicating sample weight. These graph-label pairs form the input dataset for training the GCN model.

The intricate interconnections and dependencies between these nodes pose a challenge for traditional regression models, which are typically designed for simpler, linear relationships and do not account for relationships between components in the data. They may struggle with tasks that involve dependencies or interactions. GCNs, however, offer significant advantages when dealing with graph-structured data. Unlike conventional

methods that often assume independent and identically distributed inputs, GCNs are adept at capturing and utilizing the relationships between interconnected nodes. This capability enables GCNs to model complex dependencies and interactions within the data more effectively, resulting in superior performance in scenarios where the data naturally forms a graph-like structure, as is the case here.

The GCN model is designed to learn these complex relationships from empirical data. By training on real-world industry data, the model is enabled to capture patterns and dependencies that may not be apparent through traditional analytical approaches. Once trained on authentic industrial data, the GCN model becomes a valuable tool for estimating product weight based on input parameters. It can be applied to new, unseen graphs, allowing for continuous and reliable predictions. This versatility makes it well-suited for testing and validation in real-world scenarios. The core mathematical operation in each GCN layer can be succinctly described as:

$$H^{(l+1)} = \sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (11)$$

where  $\sigma(\cdot)$  is the activation function,  $\hat{A}$  denotes the adjacency matrix,  $\hat{D}$  is the degree matrix,  $H^{(l)}$  represents the node features, and  $W^{(l)}$  is the learnable weight matrix.

Equation (11) encapsulates the propagation of information between product instances through successive layers. After training the model successfully, the sample weight  $w$  can be predicted as follows;

$$w_{pred} = GCN(G) \quad (12)$$

where  $G$  is the individual graph representing a combination of different input parameters, and  $GCN$  is the trained model.

While GCNs are traditionally used for classification tasks, they can be adapted for regression by modifying their final layer and choosing an appropriate loss function. In our approach, the final layer of the GCN outputs continuous values, allowing it to predict real-numbered variables, such as sample product weight, in our case. The primary adaptation involves using the MSE loss function, which quantifies the discrepancy between the predicted and actual sample product weight. The GCN model tries to minimize the MSE loss function given in Eq. (13)

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

where  $L$  represents the MSE loss;  $n$  is the number of graphs, i.e., each combination of input parameters;  $y$  denotes the true expected weights;  $\hat{y}$  signifies the predicted weights. By optimizing this loss function, the GCN model is fine-tuned to make accurate continuous predictions, making it well-suited for modeling this system.

### Energy consumption evaluation

Managing energy consumption in the context of a conveyor belt dryer presents a substantial challenge. The primary contributor to energy consumption during drying is the energy needed to provide the latent heat required for evaporation<sup>15</sup>. Accurately estimating and calculating this energy demand necessitates a comprehensive mathematical model that accounts for all relevant factors. However, the detailed mathematical modeling of this energy aspect goes beyond the scope of our current paper. Given that we have adopted an ML approach to control the drying process, we introduce a constant, denoted as  $c_{et}$ , which represents the cost of energy associated with a unit-degree change in chamber temperature for each time step. This energy cost value plays a pivotal role in the reward function of our control algorithm, as will be elucidated later in our discussion. The overall energy consumed in maintaining a specific temperature within the dryer can be approximated

**Table 6 | Comparison of belts' energy consumption under an actual and controlled setting**

Parameters	Actual setting			Controlled setting		
	Belt, $b_1$	Belt, $b_2$	Belt, $b_3$	Belt, $b_1$	Belt, $b_2$	Belt, $b_3$
Mass (kg)	890	640	470	900	600	425
Load (N)	8900	6400	4700	9000	6000	4250
$P_{mech}$ (Nm/s)	11125	8000	6204	10350	6900	4972.5
$P_{input}$ (Nm/s)	15892.8	11428.5	8862.8	14785.7	9857.14	7103.5
Belt energy consumption, $e_{b_i}^t$ (kWh)	7.94	5.71	4.43	7.35	4.9	3.55

**Table 7 | Adjacency matrix**

Nodes	$B_1$	$B_2$	$B_3$	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$F_1$	$F_2$	$F_3$
$B_1$	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0
$B_2$	1	1	1	0	0	0	1	1	1	0	0	0	0	1	0
$B_3$	1	1	1	0	0	0	0	0	0	1	1	1	0	0	1
$C_1$	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0
$C_2$	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0
$C_3$	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0
$C_4$	0	1	0	1	1	1	1	1	1	1	1	1	0	1	0
$C_5$	0	1	0	1	1	1	1	1	1	1	1	1	0	1	0
$C_6$	0	1	0	1	1	1	1	1	1	1	1	1	0	1	0
$C_7$	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1
$C_8$	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1
$C_9$	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1
$F_1$	1	0	0	1	1	1	0	0	0	0	0	0	1	1	1
$F_2$	0	1	0	0	0	0	1	1	1	0	0	0	1	1	1
$F_3$	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1

using Eq. (14).

$$chambers' \text{ energy consumption} = \sum_{i=1}^n t_{C_i}^t c_{et} \tag{14}$$

where  $n$  is the number of chambers,  $t_{C_i}^t$  represents the temperature of the chamber  $C_i$  at time  $t$ , and  $c_{et}$  is the energy constant associated with unit degree change in the chamber's temperature.

In addition to the energy consumption associated with heat transfer, conveyor belts represent another significant source of energy usage<sup>16</sup>. Extensive research has been dedicated to understanding and mitigating energy consumption related to heat transfer<sup>17</sup>. However, there is a notable research gap when it comes to the energy demands imposed by the robust electric motors responsible for driving conveyor loads. These motors play a pivotal role in the energy consumption of a conveyor belt dryer since their operation is closely linked to the load on the conveyor. The energy consumption of these motors is contingent on their speed and the product's weight being conveyed.

Most electric motors are designed to operate efficiently within a load range of 50%–100% of their rated capacity. Typically, peak efficiency occurs at around 75% of the rated load. For instance, a 10-horsepower (hp) motor has a suitable load range of 5–10 hp, with its highest efficiency point at 7.5 hp. Below ~50% load, a motor's efficiency tends to drop significantly<sup>18</sup>. However, it's important to note that the specific range of optimal efficiency can vary among individual motors and tends to be broader for larger motors.

To estimate the energy consumption per unit time (typically measured in watts or kilowatts) for an electric motor in a conveyor belt dryer system

that features variable speed and varying product weight, we can apply Eq. (15).

$$\text{Energy Consumption Rate, } e_{b_i}^t = \frac{P_{input}}{\Delta t} \tag{15}$$

where  $P_{input}$  is the electric input power and  $\Delta t$  is the time interval over which the estimated energy consumption is to be found. It is usually measured in seconds.

The electric input power  $P_{input}$  depends on the mechanical power  $P_{mech}$  and motor efficiency  $\eta$  i.e.,

$$P_{input} = \frac{P_{mech}}{\eta} \tag{16}$$

$$\text{Similarly, } P_{mech} = Load \left( N \frac{2\pi}{60} \right)$$

where  $Load = mg$  represents the weight of the product on the conveyor being conveyed. Equations (7–10) from “Methods” are employed to calculate this mass of the product on the conveyor.  $N$  is the speed, and it depends on the circumference of the drive roller, which can be calculated based on its diameter i.e.,  $circumference = \pi D$ . The speed should be changed to radians per second by multiplying with  $\frac{2\pi}{60}$  if it is in RPM.

**Control problem formulation**

In a conveyor belt dryer system, the control strategy is not trivial, as it poses a greater challenge due to several control inputs<sup>19</sup>. In this work, the control

inputs include fans' speeds, belts' speeds, and chambers' temperatures, which determine the quality and quantity of the product dried, and the associated energy consumption. Due to the lack of closed-form representation for the system, the applications of traditional control methods become challenging<sup>20</sup>. Furthermore, the complexity of the problem, which involves a significant number of states, actions, and unknown transition probabilities, poses difficulties in utilizing operations research techniques. As a result, the problem is formulated as a decentralized partially observable Markov decision process (Dec-POMDP) and tackled using multi-agent reinforcement learning (MARL)<sup>21</sup>. The objective of the control problem is to determine an optimal policy, mapping states to actions, to maximize the reward.

The Dec-POMDP framework is particularly suitable for decision-making and coordination in multi-agent environments. It can be represented by a tuple  $(n, S, \{U_i\}, \{O_i\}, P, r, \gamma)$ <sup>22</sup>. Within this framework, there are  $n$  agents denoted by  $i \in \{1, 2, \dots, n\}$  that collaborate to achieve a common task. At each decision point, agent  $i$  selects an action  $u_i \in U$  forming a joint action  $\mathbf{u} \in \mathbf{U} \equiv U^i, \forall i \in \{1, 2, \dots, n\}$ . These actions influence the true state of the environment  $s \in S$  according to a transition probability function  $P(s' | s, \mathbf{u}) : S \times U \times S \rightarrow S$ . Consequently, a global reward function  $r(s, \mathbf{u}) : S \times U \rightarrow \mathbb{R}$  is generated. Each agent possesses access only to partial observations, denoted as  $o^i \in O^i$  and selects actions based on a policy  $\pi(\mu^i | \sigma^i)$  dependent on its local observation-action history  $\sigma^i \in \Sigma \equiv (O^i \times U)$ . The primary objective of each agent is to maximize the discounted sum of rewards over an episode, using a discount factor  $\gamma$ .

**Dec-POMDP framework for MARL**

Within the framework of MARL, control decisions typically involve the generation of a probability distribution encompassing the action space. At each discrete time step  $t$ , an action is stochastically selected for each agent. This action selection results in a transition to a new state and the accrual of a corresponding reward. The dynamics of this transition to the new state are dictated by a transition probability, a feature inherent to the system. Agents, through a learning process, adapt their action selections based on their environmental observations, collectively known as the policy. The effectiveness of this policy is evaluated by considering the cumulative sum of rewards obtained over an episode, taking into account a discount factor. The primary aim of MARL is to compute a policy that maximizes this cumulative sum of rewards. Notably, in this system, the transition probabilities remain unknown. Consequently, we employ a model-free MARL algorithm, a strategy that obviates the need for making assumptions about transition probabilities. Instead, it concentrates solely on defining the observations, actions, and rewards in the learning process.

*Observations:* Observations refer to information that each agent has about the current state of the system. These "raw observations" only reflect local information from individual agents and may result in sub-optimal performance in coordinated control problems. The observation of each

agent  $i$  is represented as,

$$o_i^t = \begin{cases} v_{b_1}^t, v_{f_1}^t, t_{C_1}^t, t_{C_2}^t, t_{C_3}^t, H_t, t_{amb}; & \text{if } i \in [f_1, b_1, C_1, C_2, C_3] \\ v_{b_2}^t, v_{f_2}^t, t_{C_4}^t, t_{C_5}^t, t_{C_6}^t, H_t, t_{amb}; & \text{if } i \in [f_2, b_2, C_4, C_5, C_6] \\ v_{b_3}^t, v_{f_3}^t, t_{C_7}^t, t_{C_8}^t, t_{C_9}^t, H_t, t_{amb}; & \text{if } i \in [f_3, b_3, C_7, C_8, C_9] \end{cases} \quad (17)$$

where  $v_{b_i}^t$  is the speed of the conveyor belt  $b_i$  at time  $t$ ;  $v_{f_i}^t$  is the speed of the fan  $f_i$  at time  $t$ ;  $t_{C_i}^t$  is the temperature of the chamber  $C_i$  at time  $t$ ;  $H_t$  is the humidity at time  $t$ ; and  $t_{amb}$  is the ambient temperature at time  $t$ .

Action: An agent's control actions encompass the adjustment of fan or conveyor belt speeds and the regulation of chamber temperatures, depending on its type. These control actions have a direct effect on the quantity and quality of the dried product, as well as the energy consumption. The control action is defined as

$$A = \{v_{b_i}^t, v_{f_i}^t, t_{C_j}^t\} \quad (18)$$

where  $v_{b_j}^t$  represents the action of the conveyor belt  $b_j$ ;  $v_{f_i}^t$  represents the action of the fan  $f_i$ ; and  $t_{C_j}^t$  represents the action of the chamber  $C_j$ .

Reward: A global reward describes how good the combined action of all agents is. Thus, a proper reward setting is required to ensure that the resulting policy does what is expected. A good reward setting requires domain knowledge of the system and is designed in a way that increases the overall discounted sum of this reward pushing the system to work as per expectations. A variety of possible reward settings were explored for the problem and the following was chosen to be the best.

$$R = Q^t - \sum_{j=1}^k \sum_{i=1}^n (t_{C_j}^t c_{et} + e_{b_i}^t) + \theta^t \quad (19)$$

where  $Q^t$  is the amount of product dried at time  $t$ ,  $t_{C_j}^t$  is the temperature of chamber  $j$  at time  $t$ ,  $c_{et}$  is the energy consumption rate, and  $e_{b_i}^t$  represents the energy consumption of conveyor belts, which can be calculated based on Eq. (15) at time  $t$ ; and  $\theta^t$  is a constant defined as follows.

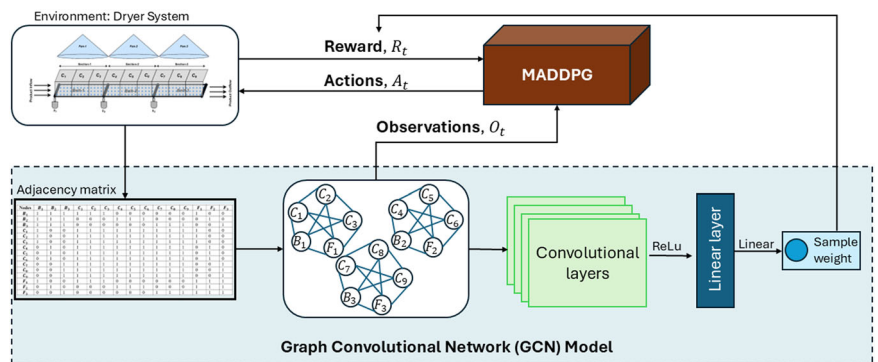
$$\theta^t = \begin{cases} 1; & \text{if } w_{pred} \in X \\ -1; & \text{if } w_{pred} \notin X \end{cases} \quad (20)$$

where  $X$  is the desired range for the predicted sample weight  $w_{pred}$ .

**MARL implementation using MADDPG algorithm**

Contrary to single-agent reinforcement learning algorithms, MARL algorithms can deal with large action spaces. Furthermore, using Actor-Critic MARL algorithms allows for medium to large state spaces unlike very often used value-based algorithms like SARSA, Q-learning, and Deep Q-learning.

**Fig. 7 |** Process flow for the proposed GCN-MADDPG algorithm.



However, in a multi-agent environment with continuous actions, coordination among the agents is often more subtle and requires fine-grained control, which MADDPG<sup>23</sup> can facilitate.

MADDPG, an extension of the Deep Deterministic Policy Gradients (DDPG) algorithm<sup>24</sup>, is tailored for coordinating the continuous actions of multiple agents in multi-agent environments. Unlike DDPG, which relies solely on self-observation and action data, MADDPG equips critics with additional information, incorporating the observations and actions of all agents. The training process follows a centralized training and decentralized execution approach, where agents independently act based on their policies during execution but collaborate during training. MADDPG employs actor-critic networks, utilizing policy gradients, to optimize agent policies. These networks take into account local observations, actions of other agents, and environmental states during training, enabling cooperative adaptation. However, during execution, agents rely solely on their local observations. Each agent in MADDPG employs two primary neural networks, actor and critic, along with target networks for stability. Critic network updates minimize a loss function, while policy networks are updated to maximize expected rewards. Soft updates ensure target network stability. This comprehensive approach empowers MADDPG to effectively address complex multi-agent scenarios, promoting cooperation and adaptability in dynamic environments.

Figure 7 represents a broad overview of our proposed methodology where the GCN and MADDPG are integrated. The environment, i.e., the dryer system, is employed to extract the relationship between different agents and develop an adjacency matrix. The adjacency matrix is used to graph the relationship between the agents at each time step. This graph is fed as input to the trained GCN model to predict the product weight, which will be used in the reward estimation for MADDPG. At the same time, the MADDPG also receives the graph-structured data and the predicted weight. It also receives energy consumption from the environment, and based on the total reward, it updates the control actions for the interacting agents, and this loop continues until the model is trained.

### Data availability

Data is a property of the DYNEA Pakistan petrochemical and cannot be shared openly.

### Code availability

The code may be provided upon request.

Received: 24 June 2024; Accepted: 30 September 2024;

Published online: 04 December 2024

### References

- Wang, S., Yu, Y., Jiang, T. & Nie, J. Analysis on carbon emissions efficiency differences and optimization evolution of China's industrial system: an input-output analysis. *Plos One* **17**, e0258147 (2022).
- Singh, P. & M. K. Gaur. Review on development, recent advancement and applications of various types of solar dryers. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 1–21 (2020).
- Mujumdar, A. S. and A. S. Menon. Drying of solids: principles, classification, and selection of dryers. In: *Handbook of industrial drying*. CRC Press. 1–39 (2020).
- Defraeye, T. Advanced computational modelling for drying processes—a review. *Appl. Energy* **131**, 323–344 (2014).
- EL-Mesery, H. S., EL-Seesy, A. I., Hu, Z. & Li, Y. Recent developments in solar drying technology of food and agricultural products: a review. *Renew. Sustain. Energy Rev.* **157**, 112070 (2022).
- Khan, M. I. H., Sablani, S. S., Joardder, M. & Karim, M. Application of machine learning-based approach in food drying: Opportunities and challenges. *Dry. Technol.* **40**, 1051–1067 (2022).
- Dibak, M., Klein, L., Krämer, A. & Noé, F. Temperature steerable flows and Boltzmann generators. *Phys. Rev. Res.* **4**, L042005 (2022).
- Arroyo, J., Manna, C., Spiessens, F. & Helsen, L. Reinforced model predictive control (RL-MPC) for building energy management. *Appl. Energy* **309**, 118346 (2022).
- Li, H. & S. Chen. A neural-network-based model predictive control scheme for grain dryers. *Dry. Technol.* **38**, 1–13 (2019).
- Zou, J., Chang, Q., Arinez, J. & Xiao, G. Analysis and evaluation of control action end-state impact on manufacturing system future production. *IEEE Access* **6**, 60187–60197 (2018).
- Huang, J., Chang, Q. & Arinez, J. Deep reinforcement learning based preventive maintenance policy for serial production lines. *Expert Syst. Appl.* **160**, 113701 (2020).
- Waseem, M., Ghani, U., Habib, T., Noor, S. & Khan, T. Productivity enhancement at molding compound manufacturing plant by applying time and motion analysis. *Mehran Univ. Res. J. Eng. Technol.* **40**(4), 761–774 (2021).
- Calín-Sánchez, Á. et al. Comparison of traditional and novel drying techniques and its effect on quality of fruits, vegetables and aromatic herbs. *Foods* **9**, 1261 (2020).
- Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv* <https://arxiv.org/abs/1609.02907> (2016).
- Menon, A., Stojceska, V. & Tassou, S. A. A systematic review on the recent advances of the energy efficiency improvements in non-conventional food drying technologies. *Trends Food Sci. Technol.* **100**, 67–76 (2020).
- He, D., Pang, Y. & Lodewijks, G. Green operations of belt conveyors by means of speed control. *Appl. Energy* **188**, 330–341 (2017).
- Zalpour, R., Singh, M., Kaur, P. & Singh, S. Refractance window drying—a revisit on energy consumption and quality of dried bio-origin products. *Food Eng. Rev.* **14**, 257–270 (2022).
- Ranges, L. *Determining electric motor load and efficiency*. US Department of Energy. Diakses dari <https://www.energy.gov/sites/prod>. (1997).
- Waseem, M., Ghani, U., Habib, T. & Noor, S. Productivity enhancement with material handling system design and human factors analysis—a case study. *Mehran Univ. Res. J. Eng. Technol.* **40**, 556–569 (2021).
- Waseem, M., C. Li, and Q. Chang. Dynamic modeling and analysis of multi-product flexible production line. *Int. J. Comput. Integr. Manuf.* **37**, 1–15 (2023)
- Bhatta, K., J. Huang, & Q. Chang. Dynamic robot assignment for flexible serial production systems. *IEEE Robot. Autom. Lett.* <https://doi.org/10.1109/LRA.2022.3182822>. (2022).
- Dibangoye, J. and O. Buffet. Learning to act in decentralized partially observable MDPs. In: *International Conference on Machine Learning*. PMLR (2018).
- Lowe, R. et al. Multi-agent actor-critic for mixed cooperative-competitive environments. In: *Advances in Neural Information Processing Systems* 30 (NIPS 2017).
- Lillicrap, T. P. et al. Continuous control with deep reinforcement learning. *arXiv* <https://arxiv.org/abs/1509.02971> (2015).

### Acknowledgements

We acknowledge the cooperation of DYNEA Pakistan Petrochemicals for providing the required data.

### Author contributions

M.W.: paper writing, modeling, and problem formulation; K.B.: coding and figures; C.L.: literature review, result analysis; N.H.: data collection; Q.C.: funding acquisition and supervision.

### Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Qing Chang.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024