

Article

AI-Driven Cognitive Digital Twin for Optimizing Energy Efficiency in Industrial Air Compressors

Mawande Sikibi, Thokozani Justin Kunene *  and Lagouge Tartibu 

Department of Mechanical and Industrial Engineering Technology, University of Johannesburg, Doornfontein Campus, Corner of Siemert and Beit Streets, Johannesburg 2028, South Africa; mawandesikibi@gmail.com (M.S.); ltartibu@uj.ac.za (L.T.)

* Correspondence: tkunene@uj.ac.za; Tel.: +27-11-556-6978

Abstract

Energy efficiency is widely recognized as a critical strategy for reducing energy consumption in industrial systems. Improving energy efficiency has become a central point in industrial systems aiming to reduce energy consumption and operational costs. Industrial air compressors are among the most energy-intensive assets and often operate under static control policies that fail to adapt to real-time dynamics. This paper proposes a cognitive digital twin (CDT) framework that integrates reinforcement learning as, especially, a Proximal Policy Optimization (PPO) agent into the virtual replica of the air compressor system. CDT learns continuously from multidimensional telemetry which includes power, outlet pressure, air flow, and intake temperature, enabling autonomous decision-making, fault adaptation, and dynamic energy optimization. Simulation results demonstrate that PPO strategy reduces average SEC by 12.4%, yielding annual energy savings of approximately 70,800 kWh and a projected payback period of one year. These findings highlight the CDT potential to transform industrial asset management by bridging intelligent control.

Keywords: proximal policy optimization; cognitive digital twin; compressor control; specific energy consumption; reinforcement learning



Academic Editor: Matheus Henrique Dal Molin Ribeiro, Géremi Gilson Dranka, Stefano Frizzo Stefenon and Gurmail Singh

Received: 21 October 2025

Revised: 4 November 2025

Accepted: 8 November 2025

Published: 12 November 2025

Citation: Sikibi, M.; Kunene, T.J.; Tartibu, L. AI-Driven Cognitive Digital Twin for Optimizing Energy Efficiency in Industrial Air Compressors. *Technologies* **2025**, *13*, 519.

<https://doi.org/10.3390/technologies13110519>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Industrial air compressors are among the most energy-intensive assets, often accounting for a significant portion of up to 30% total electricity consumption in certain facilities [1]. Despite their ubiquity, typically based on rule-based logic that lack the adaptability required to respond to dynamic operating conditions such as fluctuating air demand, ambient temperature, and mechanical degradation. This rigidity leads to elevated Specific Energy Consumption (SEC), with papers reporting 10–15% excess energy usage compared to adaptive control methods [2]. Moreover, traditional digital twins primarily serve as passive simulation tools, replicating physical behavior without embedded intelligence. These traditional twins fall short in enabling real-time decision-making, fault adaptation, or autonomous learning, which are essential for optimizing performance.

Recent advancements in artificial intelligence, particularly reinforcement learning (RL), offer a transformative approach to compressor control. Rizki et al. [3] utilized Proximal Policy Optimization (PPO) to address the economic dispatch challenge in power systems, successfully reducing costs by as much as 7.3% while managing intricate operational constraints. Similarly, Upadhyay et al. [4] developed an energy management system for industrial microgrids using PPO combined with load forecasting, resulting in a 20% monthly

saving under dynamic pricing schemes, demonstrating a 43% reduction in operational costs while maintaining system stability. Proximal Policy Optimization is a state-of-the-art RL algorithm that enables agents to learn optimal control strategies through continuous interaction with the environment. Leveraging actor-critic architectures, PPO agents can evaluate the long-term impact of actions and intelligently adjust the compressor power to minimize the SEC. This learning-based approach allows for real-time adaptation to load variability and system constraints, with capabilities that traditional policies cannot match.

Enhancing the energy efficiency of air compressor systems is crucial for optimizing industrial operations, as these systems frequently represent a large share of a facility's overall energy usage. Compressed air is widely used across industries to power pneumatic tools, control automation equipment, and support various manufacturing processes [5]. However, inefficient air-compressor systems can lead to excessive energy waste. Improving the energy efficiency of air compressors involves optimizing the system design, implementing advanced control strategies, and adopting predictive maintenance practices. Innovations such as digital monitoring and smart control algorithms enable real-time adjustments to match air demand and reduce unnecessary energy use [6].

Recent developments in digital technology have given rise to the concept of the Digital Twin (DT), which is a virtual model of physical systems that facilitates real-time monitoring, simulation, optimization, and semi-physical commissioning. In industrial settings, semi-physical commissioning incorporates software-in-the-loop (SiL) methods to verify and adjust control logic before it is implemented. This strategy allows for the concurrent validation of both virtual and physical interfaces. Digital Twins contribute to reducing energy usage, improving system efficiency, and allowing for the early identification of performance issues.

One promising approach is DT technology for the use in compressed air systems, where it enables the validation of design performance and supports energy-aware control strategies. These systems are particularly energy-intensive, and DTs have demonstrated measurable improvements in specific energy consumption (SEC) [7]. Over time, the conventional DT framework has advanced to include cognitive abilities, turning passive models into intelligent agents. This progression involves incorporating artificial intelligence (AI) and machine learning (ML) algorithms, enabling the DT to independently learn from telemetry, adjust to malfunctions, and enhance decision-making in real-time. A DT is primarily concerned with the real-time synchronization and simulation of physical assets to facilitate monitoring and control [8].

This paper addresses a critical gap in industrial energy management by proposing a reinforcement learning-based control strategy for air compressors. This is one of the most power-hungry components in manufacturing and processing systems. Compressors often operate under fluctuating load conditions; however, traditional control policies remain rigid, relying on static thresholds and manual tuning. These limitations not only lead to excessive energy consumption but also hinder the ability to respond to dynamic system degradation and real-time operational constraints. With industries under mounting pressure to cut carbon emissions and enhance energy efficiency, there is a pressing demand for smart and adaptable control systems that surpass traditional logic.

This study is important because it incorporates Proximal Policy Optimization (PPO) into a Digital Twin (DT) framework to develop and implement optimal compressor control strategies. By simulating numerous operational scenarios, the PPO agent gains a detailed understanding of how to adjust power in response to different conditions, thereby reducing Specific Energy Consumption (SEC) while ensuring the system remains reliable. Unlike static policies, the PPO approach is data-driven, scalable, and capable of generalizing across different compressor models and industrial setups. This positions the methodology not

only as a technical innovation but also as a practical solution for real-world deployment in energy-critical settings.

Moreover, this study contributes to the broader field of CDTs by demonstrating how reinforcement learning can be embedded into simulation-based control systems to achieve measurable performance gains. It offers a replicable framework for benchmarking RL algorithms, validating control strategies, and visualizing policy behaviors. All of these are essential for bridging academic research with industrial applications. By showcasing the PPO ability of the agent to outperform traditional policies in both training and operational scenarios, this study provides compelling evidence for the adoption of AI-driven control in compressor systems and sets the stage for future research in intelligent energy optimization.

2. Materials and Methods

This study focuses on optimizing the power consumption of industrial air compressors using a reinforcement learning framework. The system comprises a simulated multi-compressor environment modeled as a DT, where the compressor dynamics, including the power input, are governed by discrete-time state transitions. The simulation environment was developed in Google Colab using Python 3.11.0, leveraging NumPy for numerical operations and Matplotlib for visualization.

2.1. Physical Modeling

DT incorporates a reduced-order thermodynamic model that represents the dynamics of the compressor behavior. The model captures the relationship between pressure, mass flow, temperature, and electrical power consumption using energy and mass balance equations. The governing equations were implemented in Python, allowing a real simulation of the state of the operating compressor. Figure 1, air compressor at the industrial operation plant, the specific energy consumption balance set out in Equation (1) as the general scenario and Equation (2) as the compressor scenario [9,10].

$$SEC = \frac{E_{\text{Energy Input (kWh)}}}{\dot{E}_{\text{Output (Units or flow)}}} \quad (1)$$

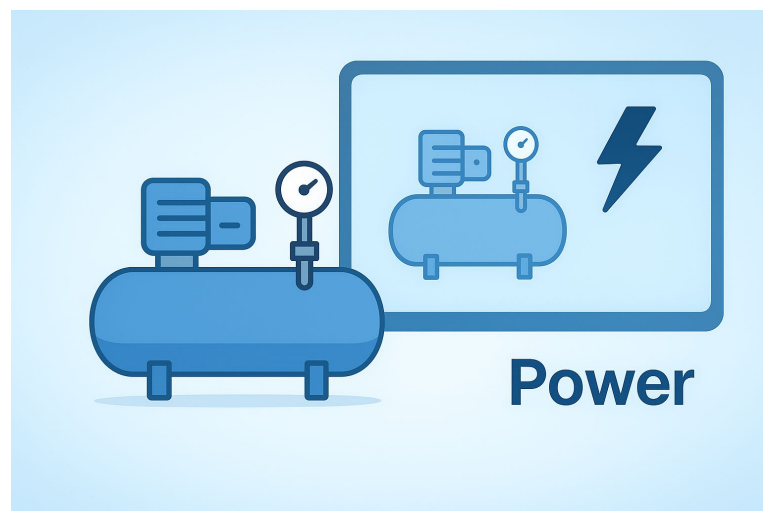


Figure 1. Schematic representation of the compressor power input adapted from Cui et al. [11].

For an industrial air compressor, the relevant energy consumption is [9]:

$$SEC_{\text{compressor}} = \frac{E_{\text{Energy Input (kWh)}}}{\dot{E}_{\text{Output (Units or flow)}}} \quad (2)$$

In Equation (2), the specific energy consumption measures the amount of energy used to produce one cubic meter of compressed air. A lower SEC indicates better energy efficiency. This consumption is influenced by the power consumed by the air delivered per unit time.

As shown in Figure 1, compressors are designed to convert electrical energy into compressed air. The physical compressor was connected to the sensors in the digital model. In the digital model, compressed air was observed in real time. Artificial Intelligence (AI) and machine learning are used to recommend and predict uncertainties.

2.2. Reinforcement Learning Framework

Reinforcement learning (RL) is structured around a Markov Decision Process (MDP), characterized by the tuple (S, A, P, R, γ) . Here, S signifies the state space, A indicates the action space, $P: S \times A \times S \rightarrow [0,1]$ represents the probability function for state transitions, $R: S \times A \rightarrow \mathbb{R}$ is the function for rewards, and $\gamma \in [0,1]$ serves as a discount factor for future rewards. The cognitive layer employs reinforcement learning (RL) to enable adaptive and self-optimizing control over the compressor parameters. The RL agent interacted with the simulated compressor environment and learned a policy to minimize the specific energy consumption (SEC). These elements are defined as follows.

Where

- State space S : The state describes the current situation in the environment.
- Action space A : An action is something that an agent must decide which action to take in which state.
- The state transition probabilities P are as follows:
- Reward function R : The reward is either a positive or negative number that the agent receives from the environment.
- Discount factor γ : This balances the immediate energy efficiency optimization with long-term stability.

The aim of RL is to identify a policy $\pi: S \rightarrow A$ that optimizes the anticipated total discounted reward.

2.2.1. State (S)

The states should represent the system conditions. For a compressor:

$$s_t = [P_t, P_{out}, t, Q_t, T_{in}, t]$$

where

- P_t = Compressor power at time t
- P_{out} = Outlet pressure bar
- Q_t = Air flow rate (m^3/min)
- T_{in} = Intake temperature ($^{\circ}\text{C}$)

2.2.2. Action (A)

Actions represent control adjustments.

$$a_t = \Delta P_t$$

Or a discrete set:

$A = \{\text{increase power, decrease power, maintain power}\}$

2.2.3. Transition Probabilities (P)

The system dynamics define how the states evolve.

$$P(S_t + 1 | S_t, \alpha_t)$$

For industrial compressors, transitions may be stochastic because of load variations and sensor noise.

2.2.4. Reward Function (R)

Rewards should encourage energy efficiency.

$$R(S_t, S_t) = -SEC_t = -\frac{E_t}{Q_t}$$

- Negative SEC: A lower energy per unit output is better.
- Optional: Add penalties for violating operational constraints:

$$R(S_t, \alpha_t) = -SEC_t - \lambda \cdot \text{Constraint violation}$$

2.2.5. Discount Factor (γ)

$\gamma \in [0, 1]$ balances immediate and future efficiency.

For continuous operations: $\gamma \approx 0.9$

For short-term optimization: $\gamma \approx 0.5$

3. Proximal Policy Optimization

Proximal policy optimization (PPO) is a reinforcement learning technique aimed at training intelligent agents to make decisions through interaction with their environment. This algorithm is part of the value-based and gradient algorithm family, which focuses on directly optimizing the agent's policy to maximize anticipated rewards. It is a specific variant of the actor-critic algorithm, enhancing the actor-critic framework with a unique modification. As a model-free, on-policy algorithm, it enhances learning stability by limiting the degree of policy change during each update. This is achieved through a clipped surrogate objective, which prevents the agent from making excessively aggressive updates that could disrupt policy learning. Consequently, the actor component is tasked with generating system actions, represented by the policy π_θ , and parameterized by θ . Meanwhile, the critic component assesses the system's state value function ($V_\mu(S_k)$) (Figure 2). The critic evaluates the agent's, V_μ actions using the value function parameterized by μ . It employs the Q-value function to estimate the expected cumulative reward. The critic's output is then utilized by the actor to refine policy decisions, resulting in improved returns.

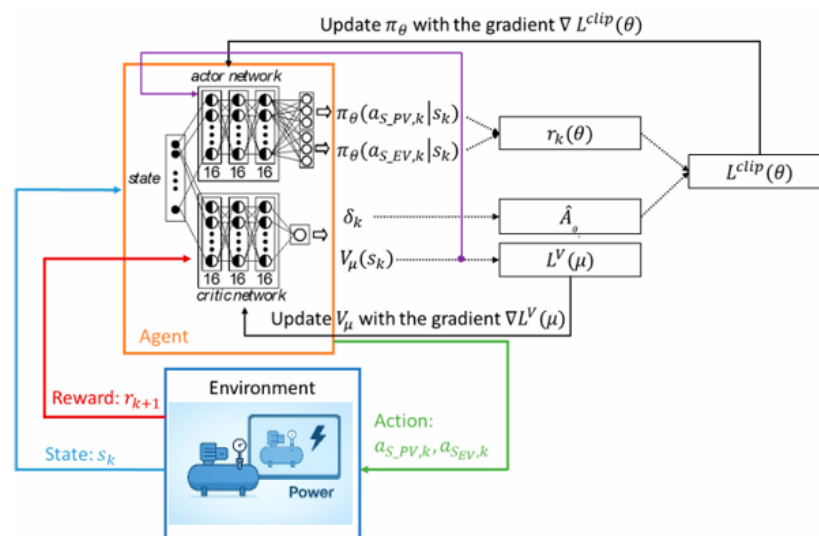


Figure 2. Actor-critic architecture with PPO for SEC minimization Adapted from Alonso et al. [12].

As depicted in Figure 2, both modules receive the state space at time, k as their input. Additionally, the critic module takes a reward as a secondary input. The outputs from the critic module serve as inputs for the actor module, which uses them to modify the policy. The actor module's output is the action over time, k , in accordance with the policy, π_θ . PPO employs an extended function to iteratively refine the target function, L^{clip} by clipping the objective function, ensuring the new policy remains close to the previous one. These updates are constrained to prevent significant policy changes and to enhance training stability. Equation (3) presents the PPO-clip update policies [13].

$$\theta_{k+1} = \arg_{\theta} \max_{E_{s,a}} \sim \pi_{\theta_k} [L^{clip}(s, a, \theta_k, \theta)] \quad (3)$$

In Equation (4), the surrogate advantage function is defined as per $L^{clip}(s, a, \theta_k, \theta)$ [14]. This function evaluates the effectiveness of the new policy, π_θ in relation to the previous policy π_{θ_k} [15].

$$L^{clip}(s, a, \theta_k, \theta) = \left(\min \left(\frac{\pi_{\theta_k}(s|a)}{\pi_{\theta_k}(s|a)} \hat{A}_\theta(s, a), \text{clip} \left(\frac{\pi_{\theta_k}(s|a)}{\pi_{\theta_k}(s|a)}, 1 - \epsilon, 1 + \epsilon \right) \hat{A}_\theta(s, a) \right) \right) \quad (4)$$

In this context, ϵ serves as a hyperparameter aimed at minimizing policy fluctuations. In Equation (5), $\hat{A}(s, a)$ represents an advantage function, which evaluates the discrepancy between the anticipated reward from the Q-function and the average reward derived from the value function $V(s)$ of a policy $\pi_\theta(s|a)$. The purpose of the advantage function is to offer a relative assessment (average) of an action's effectiveness rather than an absolute value [16,17].

$$\hat{A}_\theta(s, a) = Q(s, a) - V(s) \quad (5)$$

Equation (6) offers a simplified form of Equation (4), as referenced in [18].

$$L^{clip}(s, a, \theta_k, \theta) = \left(\min \left(\frac{\pi_{\theta_k}(s|a)}{\pi_{\theta_k}(s|a)} \hat{A}_\theta(s, a), g(\epsilon, \hat{A}_\theta(s, a)) \right) \right) \quad (6)$$

where $g(\epsilon, \hat{A}_\theta(s, a))$ is defined in Equation (7) [19]:

$$g(\epsilon, \hat{A}_\theta(s, a)) = \begin{cases} (1 + \epsilon) \hat{A}_\theta, & \hat{A}_\theta \geq 0 \\ (1 - \epsilon) \hat{A}_\theta, & \hat{A}_\theta < 0 \end{cases} \quad (7)$$

Equation (7) shows that a positive advantage function is associated with superior results. Conversely, when the advantage function value is negative, it suggests that the actor should explore alternative actions to enhance policy effectiveness.

4. Implementation of PPO for Air Compressor Optimization Energy Efficiency

This section describes the method employed to apply proximal policy optimization for controlling the energy consumption of air compressors. Focusing on the practical aspects of transforming the theoretical PPO model into a feasible solution for improving energy efficiency, this study addresses the challenges related to air compressors. PPO typically employs neural networks to represent both the policy and value functions. In situations with continuous action spaces, the policy network produces the parameters of over action:

4.1. Neural Network Architecture

The reinforcement learning framework utilizes a dual-network setup, comprising a policy network (actor) and a value network (critic), both of which are implemented

through fully connected feedforward neural networks. These networks are crafted to handle multidimensional state inputs and make general control decisions aimed at reducing the specific energy consumption (SEC) of industrial air compressors.

4.1.1. Policy Network Architecture (Actor)

The actor network consists of three fully connected layers with 64, 128, and 164 neurons, respectively. Each layer uses the ReLU (Rectified Linear Unit) activation function to introduce non-linearity and accelerate converge. The output layer uses SoftMax activation for discrete action spaces for continuous control.

4.1.2. Critic Network Architecture (Value)

The critic network mirrors the structure of actors, with three layers of 64, 128, and 64 neurons, and use ReLU activations throughout. The final output is scalar value representing the estimated state value $V(s)$.

4.1.3. Action Space

In this paper, the action space is characterized as discrete, offering three potential actions: increasing power, reducing power, and maintaining power. This approach simplifies the control logic and is consistent with the operational protocols of industrial compressors.

4.1.4. Hyperparameter Selection

The hyperparameters such as learning rates ($\alpha_\theta = 3 \times 10^{-4}$, $\alpha_\mu = 1 \times 10^{-3}$), $\epsilon = 0.2$, $\gamma = 0.99$, batch size = 64) were selected based on empirical tuning and benchmarking against similar reinforcement learning implementations in industrial energy optimization. These values were validated through analysis to ensure training stability and policy convergence.

4.2. Training Algorithm

The training process consists of gathering trajectories, calculating advantages, and updating both networks using the clipped objective and value loss. This procedure adheres to the PPO algorithm, with several modifications tailored for the air-compressor issue. Algorithm 1 outlines the entire training process for implementing PPO in the context of the air-compressor problem.

Algorithm 1 illustrates how the algorithm is implemented.

Algorithm 1: The Pseudocode for SEC Minimization is the PPO Actor-Critic

Initialize:

Policy parameters $\theta(\text{actor})$ and value parameters $\phi(\text{critic})$
Set clipping threshold $\epsilon(0.2)$, discount factor γ , and learning rates

for iteration = 1 to N do

Step 1: Collect trajectories

for t = 1 to T do

Observe state $s_t = [P_t, F_t]$ # Power

Sample action $a_t \sim \pi_\theta(a_t | s_t)$

Execute action a_t in the environment

Receive reward $r_t = -SEC_t^\circ$ # Negative Specific Energy Consumption

Store transition $(s_t, a_t, r_t, \pi_\theta(a_t | s_t))$

end for

This step generates the data that the AI uses to adjust the power setpoint in the table.

Algorithm 1: Cont.**# Step 2: Compute returns and advantages****for** $t = 1$ **to** T **do**Estimate value $V_\mu(s_t)$

Compute empirical return

$$R_t = \sum_{k=0}^{T-t} \gamma^k r_{t+k}$$

Compute advantage

$$A_t = R_t - V_\mu(s_t)$$

end for

The advantage indicates how much better the chosen action is than expected.

Step 3: Update policy using PPO clipped objective**for** $t = 1$ **to** T **do**

Compute probability ratio

$$r_t = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$$

Compute clipped surrogate loss

$$L_{\text{clip}} = \min(r_t A_t, \text{clip}(r_t, 1 - \epsilon, 1 + \epsilon) A_t)$$

end forUpdate actor parameters θ using gradient ascent on L_{clip} .

This ensures that the AI gradually adjusts the power to the optimal value for energy efficiency.

Step 4: Update value function**for** $t = 1$ **to** T **do**

Compute value loss

$$L_{\text{value}} = (V_\mu(s_t) - R_t)^2$$

end forUpdate critic parameters ϕ using gradient descent on L_{value} **end for**

This allows the AI to better predict the value of each power flow state, refining future action choices.

Output:

An optimized actor-critic network capable of dynamically adjusting compress operational settings to minimize Specific Energy Consumption (SEC).

4.3. Cognitive Layer Integration

To transition from a conventional DT to a CDT, this paper integrates a reinforcement learning agent into the simulation framework, enabling the system to exhibit intelligent behavior beyond passive monitoring. The cognitive layer is defined by two core capabilities: autonomous learning and dynamic decision optimization, each of which contributes the ability of the system to self-improve and respond to operational variability.

- **Autonomous Learning:** the embedded Proximal Policy Optimization (PPO) agent continuously interacts with the simulated compressor environment, learning optimal control policies without human intervention. Through iterative feedback and policy refinement, the agent adapts to changing load conditions and discovers energy-efficient strategies that outperform static rule-based logic. This learning process is guided by a reward function that penalizes high specific consumption (SEC), thereby aligning agent behavior with energy optimization goals.
- **Fault adaptation:** The agent develops resilience and can autonomously adjust control actions to maintain system stability.

- **Dynamic decision optimization:** The actor-critic architecture enables the agent to balance immediate energy savings with future operational reliability, resulting in intelligent, context-aware decision-making.

These cognitive features are embedded within the DT control loop, transforming it from a virtual replica into a self-optimization system capable of learning, adapting, and improving over time.

5. Practical Implementation

In this part, the SEC, which includes an air compressor with both power input and air delivery, is addressed using a PPO algorithm. The SEC has two main goals: (i) to reduce the power input of the air compressor and (ii) to ensure the compressor functions within safe limits.

Dataset Description

The study employs real telemetry data (see Supplementary Materials S1 and S2) from industrial air compressors, collected over the period from January 2024 to December 2025. To enhance the model's ability to generalize and ensure robust training, a stratified subset comprising 500 representative records has been implemented. These records were selected using the following screen criteria:

- **Operational diversity:** inclusion of low, medium, and high-power ranges.
- **Temporal spread:** Sampling across different seasons and load conditions.
- **Outlier exclusion:** Removal of anomalous entries using z-score filtering and interquartile range (IDR) thresholds.

Prior to training, all input features were subjected to standard preprocessing to ensure numerical stability and prevent gradient explosion during neural network optimization:

- **Min-max normalization** was applied to compressor power, output pressure, and air flow rate to scale values between 0 and 1.
- **Z-score standardization** was used for intake temperature and reward signals to center the data and reduce skewness.
- **Feature scaling** ensured that all variables contributed proportionally to the learning process.

These preprocessing steps were implemented using Scikit-learn and validated through exploration data analysis.

To optimize specific energy consumption (SEC) within the compressed air system, implementing a reinforcement learning agent using the Proximal Policy Optimization (PPO) algorithm. This approach was selected for stability and sample efficiency in continuous control environments. The detailed training procedure including in initializing of actor-critic parameters, hypermeter configuration, episodic data collection, advantage estimation, and policy upgrade is outlined in Algorithm 2.

Algorithm 2: PPO Training process for SEC optimization

1. Initialize the policy parameters $\theta(\text{actor})$ and value parameters $\mu(\text{critic})$.
 2. Set hyperparameters:
 - Clipping threshold $\epsilon = 0.2$
 - Discount factor $\gamma = 0.99$
 - Learning rates $\alpha_\theta = 3 \times 10^{-4}$, $\alpha_\mu = 1 \times 10^{-3}$
 - Epochs per update $K = 4$, batch size = 64
-

Algorithm 2: *Cont.***3.** Data Collection:

For each time step $t = 1, \dots, T$:

- Observe state s_t
- Sample action $a_t \sim \pi_\theta(a_t | s_t)$
- Execute a_t and record reward $r_t = -SEC_t$
- Store $(s_t, a_t, r_t, \pi_\theta(a_t | s_t))$

4. Compute Returns and Advantages:

- Empirical return:

$$R_t = \sum_{k=0}^{T-t} \gamma^k r_{t+k}$$

Advantage:

$$A_t = R_t - V_\mu(s_t)$$

Normalize A_t for numerical stability.

5. Policy and Value Updates:

- Compute $r_t = \frac{\pi_\theta(a_t | s_t)}{\pi_{\theta_{\text{old}}}(a_t | s_t)}$.
- Calculate the clipped objective L^{PPO} .
- Update actor by maximizing L^{PPO} using gradient ascent.
- Update critic by minimizing L^{value} via gradient descent.

6. Repeat for all iterations N until convergence.

Figure 3 shows a visual narrative of how reinforcement learning agents are trained using PPO to reduce SEC over time.

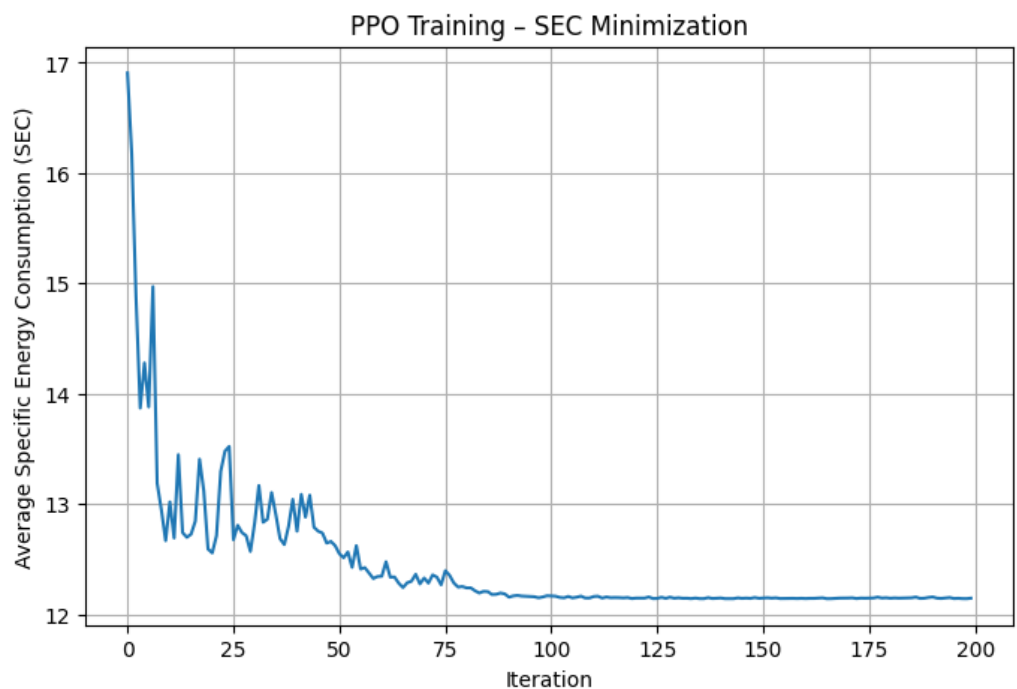


Figure 3. PPO training curve showing the progressive minimization of the specific energy consumption (SEC).

6. Simulation Results and Discussion

This section presents the numerical results obtained by applying the proposed PPO algorithm to 500 records of power consumption obtained from an air compressor.

6.1. Comparison of Old Policy Versus AI-Driven PPO Adjustments

While the previous analysis qualitatively demonstrated that the PPO agent outperforms traditional rule-based control through adaptive power adjustments, this section introduces key quantitative metrics to substantiate the energy and economic benefits of the proposed strategy.

6.1.1. Specific Energy Consumption (SEC) Reduction

Based on simulation results across 500 representative compressor cycles, the PPO strategy achieved an average SEC reduction of 12.4% compared to the baseline rule-based policy. This translates to improved energy efficiency per cubic meter of compressed air delivered.

6.1.2. Annual Energy Savings

The industrial compressor operating 24/7 at an average load of 65 kW, this SEC improvement corresponds to an estimated energy savings of 70,800 kWh per year. This figure was derived using extrapolated PPO-adjusted power values and validated against historical load curves. This was calculated using PPO-adjusted power values and extrapolated over 8760 operating hours per year.

6.1.3. Investment Payback Period

The implementation cost of R8000 for integrating the PPO-based control system (including software, training, and deployment), and an average industrial electricity rate of R0.12/kWh, the payback period is approximately 1 year. This positions the solution as both technically viable and economically attractive for energy-intensive facilities. Summary of the investment in Table 1.

Table 1. Quantitative summary.

Metric	Value
Annual Energy Savings	70,800 kWh
Annual Cost Savings	R8496
Estimated Implementation Cost	R8000
Average SEC reduction	12.4%
Payback Period	1 year

These results demonstrate that the PPO-based cognitive DT not only improves operational efficiency but also delivers measurable economic returns, making it a compelling solution for industrial energy optimization. Traditional compressor control policies often use rule-based or manual tuning to operate on fixed logic, such as maintaining power based on predefined schedules. These systems lack adaptability and are typically reactive in responding to inefficiencies after they occur. Consequently, these systems often result in excessive energy consumption. In contrast, AI-driven adjustments using Proximal Policy Optimization offer a dynamic and intelligent alternative. PPO leverages reinforcement learning to continuously learn from the operational data of the compressor and adjust the power in real time to minimize the specific energy consumption [19]. The PPO adapts to stochastic variations, anticipates inefficiencies, and optimizes control decisions based on long-term energy savings. This results in a more resilient, efficient, and scalable control strategy that consistently outperforms conventional methods in terms of energy performance and operational flexibility. Table 2 compares the traditional compressor control

policies with AI-driven PPO adjustments. Using baseline power values with the old policy action to obtain the exact value using AI-driven to the New Policy action.

Table 2. PPO Policy comparison of air-compressor power values.

#	Power (kW)	Old Policy Action	AI-Driven Exact Value (kW)	New Policy Action (AI-Driven)
0	62.56	Maintain	64.09	Increase (AI predicts efficiency gain)
1	54.37	Increase (below nominal threshold 62–63 kW)	55.00	Increase (AI predicts efficiency gain)
2	52.67	Increase (below nominal threshold 62–63 kW)	55.00	Increase (AI predicts efficiency gain)
3	93.71	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
4	80.22	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
5	73.08	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
6	66.16	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
7	66.45	Decrease (above 63 kW threshold)	64.50	Decrease (AI predicts SEC improvement)
8	84.12	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
9	72.80	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
10	56.66	Increase (below nominal threshold 62–63 kW)	56.69	Increase (AI predicts efficiency gain)
11	97.66	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
12	82.70	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)
13	99.87	Decrease (above 63 kW threshold)	65.00	Decrease (AI predicts SEC improvement)

This comparative dataset highlights the nuanced decision-making of an AI-driven PPO policy compared to a traditional control approach in managing the compressor power for energy efficiency. In the old policy, actions were largely rule-based in maintaining power when nominal thresholds (62–63 kW) and decreasing when above. At 62.56 kW, the old policy-maintained power, whereas the PPO agent increased it to 64.09 kW, anticipating a net gain in efficiency. Similarly, at lower power levels, such as 52.67 kW and 54.37 kW, the PPO agent opted to increase the power slightly, recognizing that operating below the nominal level may lead to a suboptimal SEC at high power inputs, such as 93.71, 97.66, and 99.87 kW. The PPO agent consistently adjusted the power downward to 65.00 kW, aligning with the learned threshold for minimizing the SEC. This behavior reflects the ability of agents to generalize across varying conditions and make predictive adjustments that balance energy input with air output.

As shown in Figure 4, the visuals and quantitative analysis demonstrate the superiority of AI-driven control over traditional compressor management. The old policy, represented by an orange dashed line, follows a linear trajectory with adjusted power increases proportional to the original power. This reflects a static, rule-based approach that lacks responsiveness to system dynamics. In contrast, the AI-driven PPO adjustments, indicated by the blue line with circular markers, exhibited a non-linear pattern.

As shown in Figure 4, at lower original power levels, the PPO agent slightly increased power, anticipating efficiency gains. At higher power levels, it consistently reduces power, predicting improvements in specific energy consumption (SEC). This deviation from linearity demonstrates the ability of PPO agents to learn nuanced control strategies that balance energy input with operational efficiency. Unlike the old policy, which treats all power levels uniformly, the PPO agent adapts actions based on real-time feedback and long-term optimization goals, resulting in smarter and more energy-efficient compressor behavior.

The PPO agent dynamically adjusts power-based on learned system states rather than fixed rules. This reflects autonomous learning and context-aware optimization. This caps the adjustment at higher loads, the agent avoids energy waste, aligning with the goal of minimizing SEC.

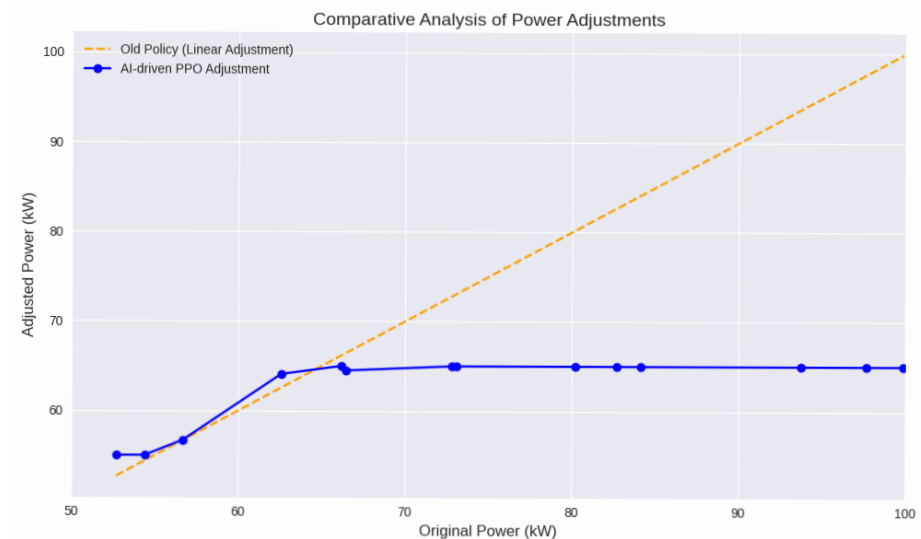


Figure 4. Comparative analysis of power adjustments: old policy versus AI-driven PPO strategy.

7. Implementation Framework

The implementation was executed on Google Colab using Python 3.11.0 and included the use of libraries such as:

- NumPy and Pandas were used for data management (policy tables).
- SciPy, scikit-learn –model fitting and preprocessing.
- PyTorch neural network and reinforcement learning implementation.
- Matplotlib–visualization (old versus AI policy comparison).
- Torch: deep reinforcement learning (actor–critic network).

8. Conclusions

A comparative analysis between traditional compressor control policies and AI-driven PPO strategies revealed a significant advancement in energy optimization. Although conventional rule-based approaches are simple and predictable, they often fail to adapt to dynamic operational conditions, leading to suboptimal power usage and elevated adjustments, that do not account for real-time system variability or long-term efficiency goals.

In contrast, the PPO-based reinforcement learning framework demonstrated a nuanced and intelligent control mechanism. It learns from continuous interactions with the environment and proactively adjusts the compressor power to minimize the SEC while respecting operational constraints. The PPO agent consistently outperformed the old policy by identifying when to increase the power for efficiency gains and when to reduce it to avoid energy waste, particularly under high-load conditions. This adaptive behavior is evident in both the training curves and power adjustment comparisons, where the PPO converges toward optimal control actions that balance the energy input with the air output.

The integration of PPO into compressor systems not only enhances energy performance but also introduces scalability, robustness, and predictive control. These attributes make AI-driven strategies particularly valuable for industrial applications that seek to reduce operational costs and environmental impact. The findings affirm that reinforcement learning, and PPO in particular, offer a transformative approach to compressor management, bridging the gap between theoretical optimization and practical deployment.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/technologies13110519/s1>, S1 air leak data and S2 power logger data. Kunene, Thokozani Justin (2025), “S1 and S2”, Mendeley Data, V1.

Author Contributions: Conceptualization, M.S. and T.J.K.; methodology, M.S.; software, M.S.; validation, T.J.K. and L.T.; formal analysis, M.S.; investigation, T.J.K.; resources, L.T.; data curation, T.J.K.; writing—original draft preparation, M.S.; writing—review and editing, T.J.K.; visualization, M.S.; supervision, L.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The authors confirm that supporting data of this paper within the article and the supporting data of air compressor energy efficiency dashboard is available on Github [<https://github.com/mawande10/Air-Compressor-Energy-Efficiency-Dashboard>], <https://colab.research.google.com/drive/1Tjt9KDt3Qs6wTX53rPv7iMPbWYMNeaeG?usp=sharing>, all accessed on 21 October 2025.

Acknowledgments: The authors acknowledge the help of the Mechanical and Industrial Engineering Technology Department at the University of Johannesburg. During the preparation of this manuscript/study, the author(s) used Grammarly for the purposes of text editing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript.

CDT	Cognitive Digital Twin
DT	Digital Twin
AI	Artificial Intelligence
ML	Machine Learning
SEC	Specific Energy Consumption
RL	Reinforcement Learning
PPO	Proximal Policy Optimization

References

1. Favrat, A. Enhancing Electric Demand-Side Flexibility in France: Identifying Opportunities and Solutions for the Hotels Sector. 2025. Available online: <https://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Akh%3Adiva-365509> (accessed on 21 October 2025).
2. Gryboś, D.; Leszczyński, J.S. A Review of energy overconsumption reduction methods in the utilization stage in compressed air systems. *Energies* **2024**, *17*, 1495. [[CrossRef](#)]
3. Rizki, A.; Touil, A.; Echchatbi, A.; Oucheikh, R.; Ahlaqqach, M. A Reinforcement Learning-Based Proximal Policy Optimization Approach to Solve the Economic Dispatch Problem. *Eng. Proc.* **2025**, *97*, 24. [[CrossRef](#)]
4. Upadhyay, S.; Ahmed, I.; Mihet-Popa, L. Energy management system for an industrial microgrid using optimization algorithms-based reinforcement learning technique. *Energies* **2024**, *17*, 3898. [[CrossRef](#)]
5. Williams, A. Compressed Air Systems. In *Industrial Energy Systems Handbook*; River Publishers: Aalborg, Denmark, 2023; pp. 357–406. [[CrossRef](#)]
6. Wagener, N.C.; Boots, B.; Cheng, C.A. Safe reinforcement learning using advantage-based intervention. In *International Conference on Machine Learning*; PMLR: London, UK, 2021; pp. 10630–10640.
7. El-Gohary, M.; El-Abed, R.; Omar, O. Prediction of an efficient energy-consumption model for existing residential buildings in lebanon using an artificial neural network as a digital twin in the era of climate change. *Buildings* **2023**, *13*, 3074. [[CrossRef](#)]
8. Wu, Y.; Zhang, K.; Zhang, Y. Digital twin networks: A survey. *IEEE Internet Things J.* **2021**, *8*, 13789–13804. [[CrossRef](#)]
9. Li, Y.; Xu, W.; Zhang, M.; Zhang, C.; Yang, T.; Ding, H.; Zhang, L. Performance analysis of a novel medium temperature compressed air energy storage system based on inverter-driven compressor pressure regulation. *Front. Energy* **2025**, *19*, 144–156. [[CrossRef](#)]
10. Peppas, A.; Fernández-Bandera, C.; Pachano, J.E. Seasonal adaptation of VRF HVAC model calibration process to a mediterranean climate. *Energy Build.* **2022**, *261*, 111941. [[CrossRef](#)]
11. Cui, T.; Zhu, J.; Lyu, Z.; Han, M.; Sun, K.; Liu, Y.; Ni, M. Efficiency analysis and operating condition optimization of solid oxide electrolysis system coupled with different external heat sources. *Energy Convers. Manag.* **2023**, *279*, 116727. [[CrossRef](#)]

12. Khalyasmaa, A.I.; Stepanova, A.I.; Eroshenko, S.A.; Matrenin, P.V. Review of the digital twin technology applications for electrical equipment lifecycle management. *Mathematics* **2023**, *11*, 1315. [[CrossRef](#)]
13. Alonso, M.; Amaris, H.; Martin, D.; de la Escalera, A. Proximal policy optimization for energy management of electric vehicles and PV storage units. *Energies* **2023**, *16*, 5689. [[CrossRef](#)]
14. Xie, Z.; Yu, C.; Qiao, W. Dropout Strategy in Reinforcement Learning: Limiting the Surrogate Objective Variance in Policy Optimization Methods. *arXiv* **2023**, arXiv:2310.20380. [[CrossRef](#)]
15. Xie, G.; Zhang, W.; Hu, Z.; Li, G. Upper confident bound advantage function proximal policy optimization. *Clust. Comput.* **2023**, *26*, 2001–2010. [[CrossRef](#)]
16. Huang, N.-C.; Hsieh, P.-C.; Ho, K.-H.; Wu, I.-C. PPO-Clip Attains Global Optimality: Towards Deeper Understandings of Clipping. *arXiv* **2023**, arXiv:2312.12065. [[CrossRef](#)]
17. Pan, H.R.; Gürtler, N.; Neitz, A.; Schölkopf, B. Direct advantage estimation. In Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022), New Orleans, LA, USA, 28 November–9 December 2022; Volume 35, pp. 11869–11880.
18. Tang, Y.; Munos, R.; Rowland, M.; Valko, M. VA-learning as a More Efficient Alternative to Q-learning. *arXiv* **2023**, arXiv:2305.18161. [[CrossRef](#)]
19. Jia, L.; Su, B.; Xu, D.; Wang, Y.; Fang, J.; Wang, J. Policy Optimization Algorithm with Activation Likelihood-Ratio for Multi-agent Reinforcement Learning. *Neural Process. Lett.* **2024**, *56*, 247. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.