Integrated Scheduling and Maintenance (GASUEnv)

April 24, 2025

1 Environments in the paper

1.1 Integrated scheduling and maintenance (GASUEnv)

GASUEnv simulates the operation and maintenance of compressors in an air separation unit (ASU) focused on meeting gaseous product demand. Since inventorying gaseous products is impractical, the system allows for external product purchases when demand exceeds production capacity. The ASU consists of a set C of n compressors ($n = 3$), where each compressor $c \in C$ has a maximum capacity denoted by Cap_c. The agent must decide on a daily basis whether each compressor should operate at a production level $r_c \in [0, 1]$ of its maximum capacity or undergo maintenance, based on its condition. Over an episode of length T, the agent also determines daily external purchase quantities, aiming to minimize the total operational cost, which includes both production and purchase costs.

Observation Space:
The observation state at time t is represented as a vector,

$$s(t) = (\mathbf{d}_t, \mathbf{e}_t, \mathbf{tslm}_t, \mathbf{tlcm}_t, \mathbf{cdm}_t), \quad t \in [1, T]$$

The state vector $s(t)$ captures the essential operational and maintenance-related information for the Air Separation Unit (ASU) on day t. It includes forecasts of production demands and electricity prices over a fixed horizon (S days), along with detailed maintenance indicators for each compressor. These components are defined as follows:

- **Demand Forecast** $d_t \in \mathbb{R}_+^{S\times 1}$: a vector of predicted demands from day t to $t + S - 1$, expressed in tons.
- **Electricity Price Forecast** $e_t \in \mathbb{R}_+^{S\times 1}$: a vector of corresponding day-ahead electricity prices from day t to $t + S - 1$, measured in KWh.
- **Time Since Last Maintenance** $\mathrm{tslm}_t \in \mathbb{Z}_+^{n\times 1}$: the number of days since each compressor c last underwent maintenance.

- **Time Left to Complete Maintenance** $\mathrm{tlcm}_t \in \mathbb{Z}_+^{n\times 1}$: the remaining time (in days) required to complete maintenance for each compressor c; it is non-zero only if maintenance has already started.

- **Can Do Maintenance Indicator** $\mathrm{cdm}_t \in \{0,1\}^n$: a binary vector where $\mathrm{cdm}_{ct} = 1$ indicates that compressor c is eligible for maintenance on day t.

Action Space:
The action space at each time step t is defined as:

$$a(t) = (a_{\mathrm{maintenance}}(t), a_{\mathrm{production}}(t), a_{\mathrm{purchase}}(t))$$

The action space at time t consists of operational decisions related to maintenance scheduling, compressor utilization, and external product procurement. The physical description is as below:

- **Maintenance Scheduling** $a_{\mathrm{maintenance}}(t) \in \{0,1\}^n$: a binary vector indicating whether each compressor is scheduled for maintenance at time t, with $a_{\mathrm{maintenance},c}(t) = 1$ if compressor c is under maintenance.

- **Production Rate** $a_{\mathrm{production}}(t) \in [0,1]^n$: a continuous vector representing the fraction of the maximum capacity Cap_c utilized by each compressor c.

- **External Purchase** $a_{\mathrm{purchase}}(t) \in [0, 10000]$: a scalar value indicating the quantity of external product to be purchased to meet demand when internal production is insufficient.

Reward Function:
The reward function represents the cost incurred by the agent for making decisions related to production and external purchases. At each time step t, the reward is defined as:

$$r(t) = -\left(\text{production cost}_t + \text{external purchase cost}_t\right)$$

The production cost at time t is calculated using the production rate, compressor capacity, specific energy consumption, and electricity price. The external purchase cost is incurred when demand exceeds production capacity, calculated by multiplying the purchase amount by the external price.

Cost:
The agent may incur various costs at each time step if system constraints are violated, which encourages the agent to learn an optimal policy. The potential costs are as follows:

1.1.1 Maintenance Duration Cost

This penalty is incurred if maintenance is interrupted before it is completed or prolonged, in which case tlcm_{ct} becomes negative. The penalty is defined as:

$$
\text{cost}_{ct,\text{MD}} = \begin{cases} -\rho_{\text{MD}} \cdot \exp\left(|\text{tlcm}_{ct}|\right) & \begin{array}{l} \text{if } \left(a_{\text{maintenance},c}(t) \neq 1 \text{ and } \text{tlcm}_{ct} > 0\right) \\ \textit{or } \left(a_{\text{maintenance},c}(t) = 1 \text{ and } \text{tlcm}_{ct} = 0 \text{ and } \text{tslm}_{ct} = 0\right) \\ \textit{or } \left(a_{\text{maintenance},c}(t) = 1 \text{ and } \text{tlcm}_{ct} < 0\right), \end{array} \\ 0 & \text{otherwise.} \end{cases}
$$

1.1.2 Maintenance Failure Cost

This penalty occurs if the *Time Since Last Maintenance* (tslm_{ct}) exceeds the *Mean Time to Failure* (MTTF_c) of the compressor. The penalty is defined as:

$$
\text{cost}_{ct,\text{MF}} = \begin{cases} -\rho_{\text{MF}} \cdot (\text{tslm}_{ct} - \text{MTTF}_c) & \text{if } \left(a_{\text{maintenance},c}(t) = 0 \text{ and } \text{tslm}_{ct} > \text{MTTF}_c\right), \\ -\rho_{\text{MF}} & \text{if } \left(a_{\text{maintenance},c}(t) = 0 \text{ and } \text{tslm}_{ct} = \text{MTTF}_c\right), \\ 0 & \text{otherwise.} \end{cases}
$$

1.1.3 Early Maintenance Cost

This penalty is incurred if maintenance is performed on a compressor when it is not yet eligible for maintenance (i.e., when $\text{cdm}_{ct} = 0$, indicating that the compressor has recently undergone maintenance, and $\text{tlcm}_{ct} = 0$). The penalty is proportional to the *Time Since Last Maintenance* (tslm_{ct}) of the compressor.

$$
\text{cost}_{ct,\text{EM}} = \begin{cases} -\rho_{\text{EM}} \cdot \text{tslm}_{ct} & \text{if } \left(a_{\text{maintenance},c}(t) = 1 \text{ and } \text{cdm}_{ct} = 0 \text{ and } \text{tlcm}_{ct} = 0\right), \\ 0 & \text{otherwise.} \end{cases}
$$

1.1.4 Ramp Cost

This penalty is incurred when a compressor is ramped up while under maintenance.

$$
\text{cost}_{ct,\text{RP}} = \begin{cases} -\rho_{\text{RP}} \cdot a_{\text{production},c}(t) \cdot \text{Cap}_c & \text{if } \left(a_{\text{maintenance},c}(t) = 1 \text{ and } a_{\text{production},c}(t) \neq 0\right), \\ 0 & \text{otherwise.} \end{cases}
$$

1.1.5 Demand Cost

This penalty is incurred if the total supply from production and external purchases does not meet the demand on the current day. The total supply is the sum of the production and external purchase, and if this is less than or greater than the demand, a penalty is imposed proportional to the absolute difference between demand and supply.

$$\text{cost}_{t,\text{D}} = -\rho_{\text{D}} \cdot |d_t - \text{total\_supply}_t|$$

Therefore, the total penalty at time t can be written as:

$$\text{cost}_t = \sum_{c \in C} (\text{cost}_{ct,\text{MD}} + \text{cost}_{ct,\text{MF}} + \text{cost}_{ct,\text{EM}} + \text{cost}_{ct,\text{RP}}) + \text{cost}_{t,\text{D}}$$

2 In the Appendix: Environment details

2.1 Plant model and motivation

Energy-intensive chemical processes leverage Demand Response (DR) to adjust electricity usage in response to price fluctuations, typically optimizing production on a rolling basis using forecasted demand and electricity prices. However, optimizing production scheduling alone can be detrimental, as it neglects the operational condition of essential equipment. Recent studies have addressed this by integrating condition-based maintenance into production optimization, notably for Air Separation Units (ASUs) [Xenos2016] and natural gas plants [Huang2020].

In this work, we study an ASU tasked with meeting aggregated gaseous nitrogen (GAN) and oxygen (GOX) demand using three compressors with limited capacities and maintenance needs. Downtime from maintenance may reduce production, prompting external product purchases to meet demand. We adopt a deterministic 30-day rolling forecast for demand and electricity prices, with an episode length of 31 days. Daily actions guide the system through time, balancing production, maintenance, and purchasing decisions.

2.2 Assumptions

The GASU environment is based on the following assumptions:

- We consider only two operational modes in the plant: "Work" and "Maintenance". Furthermore, in the work mode, we assume a joint production capacity for both GAN and GOX.
- Perfect forecasts are assumed to be available for both demand and electricity prices over the duration of the episode.
- Each compressor has a fixed mean time to failure ($\text{MTTF}_c \in \mathbb{Z}_+$), after which it must undergo maintenance.
- Compressors have a fixed maintenance duration, after which they can be put back into operation. This duration is denoted as the mean time to repair (MTTR_c).

- After maintenance is completed, a compressor must wait for a specified duration before the next maintenance action can be performed. This is referred to as the minimum no-repair duration (MNRD_c). This constraint is important because it allows for the possibility of scheduling early maintenance when a spike in demand is anticipated near the maximum allowable operation time of the compressor.

2.3 Action Correction Functions

To ensure feasibility and consistency with compressor state constraints, a sanitization step is applied to the raw agent actions before execution. For each compressor $c \in C$, the action is adjusted as follows:

- If $\mathrm{tslm}_{ct} \geq \mathrm{MTTF}_c$ and $a_{\mathrm{maintenance},c}(t) \neq 1$, then:

$$\Rightarrow \quad a_{\mathrm{maintenance},c}(t) \leftarrow 1 \quad \text{and} \quad a_{\mathrm{production},c}(t) \qquad \leftarrow 0$$

- If $a_{\mathrm{maintenance},c}(t) = 1$ and $a_{\mathrm{production},c}(t) > 0$, then:

$$\Rightarrow \quad a_{\mathrm{production},c}(t) \leftarrow 0$$

- If $\mathrm{cdm}_{ct} = 0$ and $a_{\mathrm{maintenance},c}(t) = 1$, then:

$$\Rightarrow \quad a_{\mathrm{maintenance},c}(t) \leftarrow 0$$

- If $\mathrm{tlcm}_{ct} > 0$ and $a_{\mathrm{maintenance},c}(t) \neq 1$, then:

$$\Rightarrow \quad a_{\mathrm{maintenance},c}(t) \leftarrow 1 \quad \text{and} \quad a_{\mathrm{production},c}(t) \qquad \leftarrow 0$$

2.4 Environment Transition Function

The environment transition function in **GASUEnv** defines how the environment state evolves in response to the agent's actions at each discrete time step $t \in T$. Let $s(t) = (\mathbf{d}_t, \mathbf{e}_t, \mathbf{tslm}_t, \mathbf{tlcm}_t, \mathbf{cdm}_t)$ be the observation vector at time t. The transition to $s(t+1)$ is governed by the following procedures:

2.4.1 Information State Update

This update incorporates changes in demand and electricity price signals. The updated states are retrieved from the simulated perfect-forecast arrays of demand (`D`) and electricity prices (`E`) for the next S days as follows:

$$\mathbf{d}_{t+1} \leftarrow \mathtt{D}[t, t+S], \quad e_{t+1} \leftarrow \mathtt{E}[t, t+S] \qquad \forall t \in T$$

2.4.2 Compressor Physical Condition Transition

The following updates track the evolution of maintenance status and compressor readiness for each compressor $c \in C$, based on operational decisions. The initial physical state at the start of the simulation is given by $(tlcm_{c,0}, tslm_{c,0}, cdm_{c,0})$, and future states are derived accordingly.

$$\mathrm{tslm}_{c,t+1} = \begin{cases} 0, & \text{if } a_{\text{maintenance},c}(t) = 1 \\ \mathrm{tslm}_{ct} + 1, & \text{otherwise} \end{cases}$$

$$\mathrm{tlcm}_{c,t+1} = \begin{cases} \mathrm{MTTR}_c - 1, & \text{if } a_{\text{maintenance},c}(t) = 1 \wedge \mathrm{cdm}_{ct} = 1 \\ \mathrm{tlcm}_{ct} - 1, & \text{if } a_{\text{maintenance},c}(t) = 1 \wedge \mathrm{cdm}_{ct} = 0 \\ \mathrm{tlcm}_{ct}, & \text{otherwise} \end{cases}$$

$$\mathrm{cdm}_{c,t+1} = \begin{cases} 1, & \text{if } \mathrm{tslm}_{c,t+1} \geq \mathrm{MNRD}_c \\ 0, & \text{otherwise} \end{cases}$$

2.5 Representativeness and Generalizability

The model can be used for independent production scheduling by relaxing the maintenance constraints, allowing for simpler operation-focused planning. Additionally, the model is highly extensible; incorporating additional compressors requires only minor modifications to the compressor configuration in the input JSON file, making the framework scalable to more complex plant setups.

The environment's capability can be readily enhanced to account for uncertainty in demand and electricity price forecasts. This extension would enable the development and benchmarking of robust or stochastic reinforcement learning policies, and constitutes a key direction for future work.

2.6 Limitations

The current environment could be made more realistic and robust by incorporating condition-based maintenance strategies, where the timing of maintenance is informed by compressor health indicators rather than fixed schedules.

Moreover, the current assumption of external gas purchases during supply shortfalls could be replaced by a more detailed process-based mechanism such as the activation of an internal vaporizer, which vaporizes stored cryogenic liquids (liquid oxygen or nitrogen) to meet demand. This would improve the physical fidelity of the environment and allow for more nuanced control strategies.

2.7 MILP counterpart

Define all the variables, parameters, and the optimization model to derive the optimal action. To be done later.