Reinforcement learning for ASU production planning

Ikhlass YAYA-OYÉ, Axelle ALBOT, Paul BERHAUT

paul.berhaut@airliquide.com

Agenda

Introduction

A history of successes in Reinforcement Learning

1992 **TD-Gammon**

Human grandmaster level at Backgammon, as **early as 1992**. This little-known achievement predates DeepBlue, when Gary Kasparov was beaten at chess by classical algorithms.

2013 **DQN**

Strong performance in a **large variety** of Atari games with **high-dimensional input**, **surpassing game-specific** AI algorithms on each game.

2016 **Alpha Zero**

Achieved **super human performance** at the game of Go, without prior human knowledge. **General algorithm**: results replicated in Chess and Shogi.

2018 **Open AI 5**

Super human performance in a **real time team** game where **cooperation** between multiple AI is required.
Demonstrate ability to learn in **partially observed environments.**

2022 **MuZero**

Real-life reinforcement learning application to **increase data compression rate by 4%** in video streaming applications.

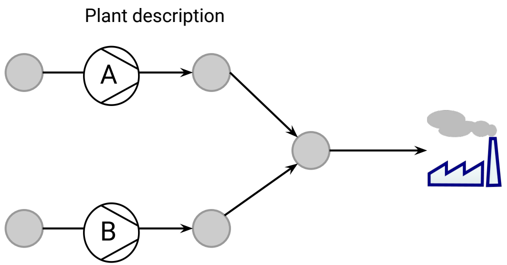
Can these techniques be applied to solve real-life Air Liquide problems ?

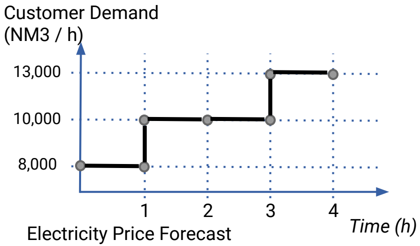
Context

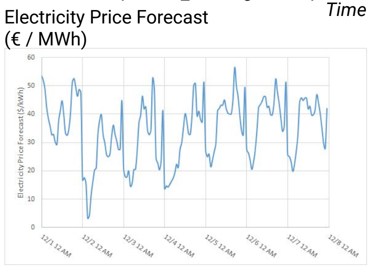
Liquefaction production scheduling

Context: Air Separation is an electro-intensive industry where electricity represents 75% of costs.

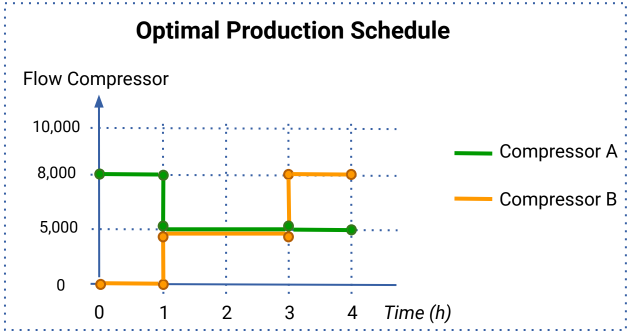
Optimization problem: Given a liquefaction system and a customer demand profile, find the optimal combination of compressor flows that minimizes total production cost







Optimization Input

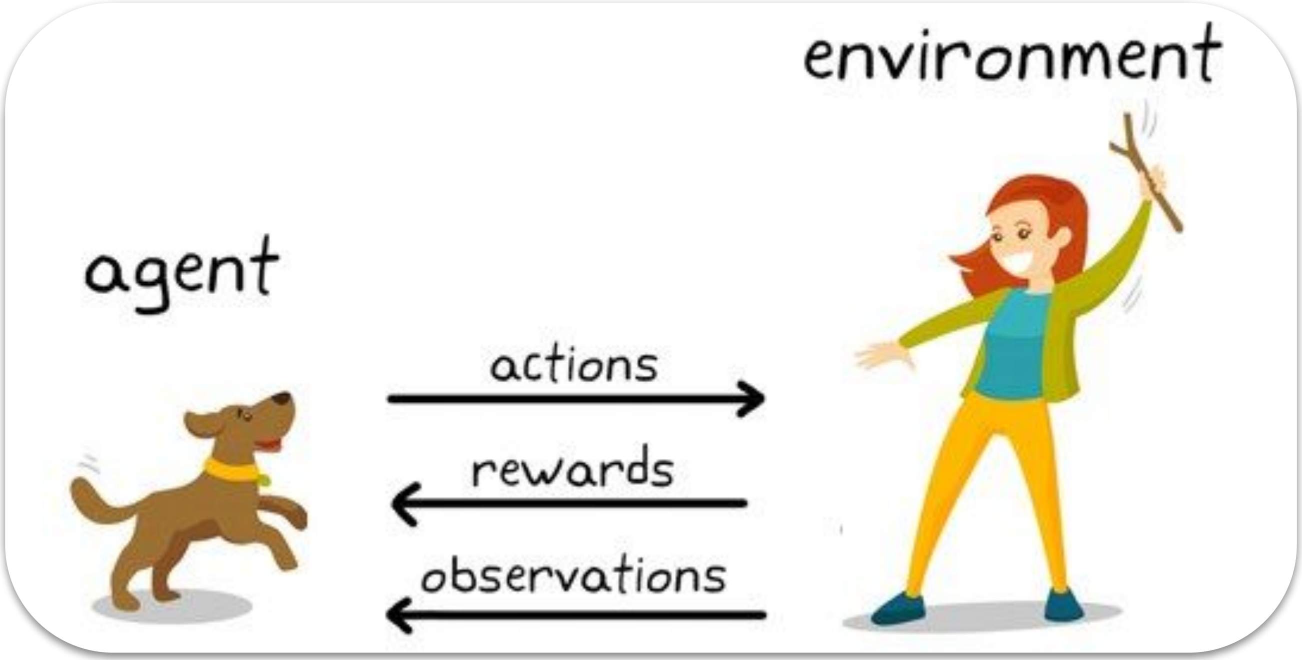


Optimization Output

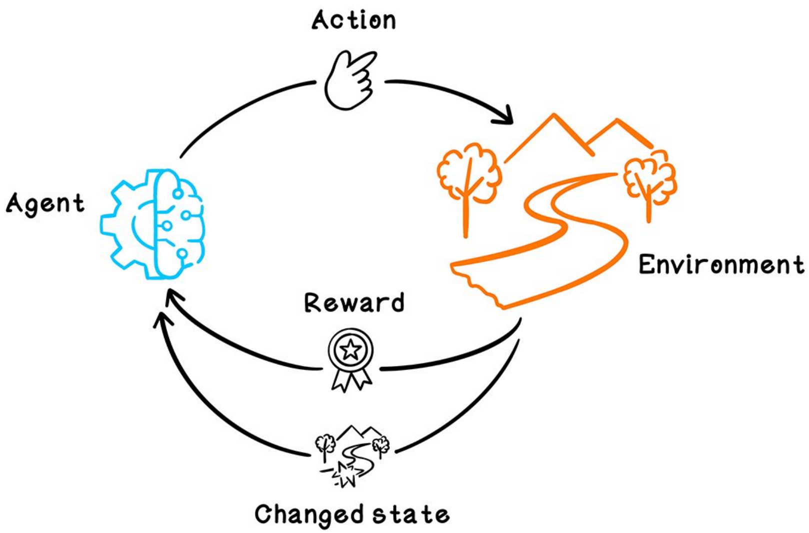
paul.berhaut@airliquide.com

Reinforcement Learning framework

What is RL ?

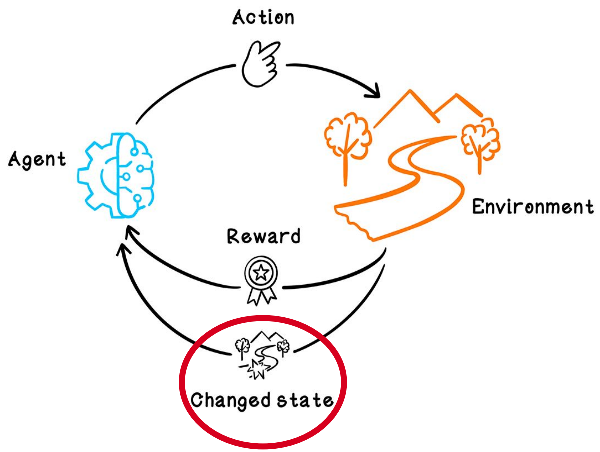


What is RL ?



Reinforcement Learning is a set of techniques applied when an **agent repeatedly interacts** with an **environment**, with the objective to maximize a **long-term reward**.

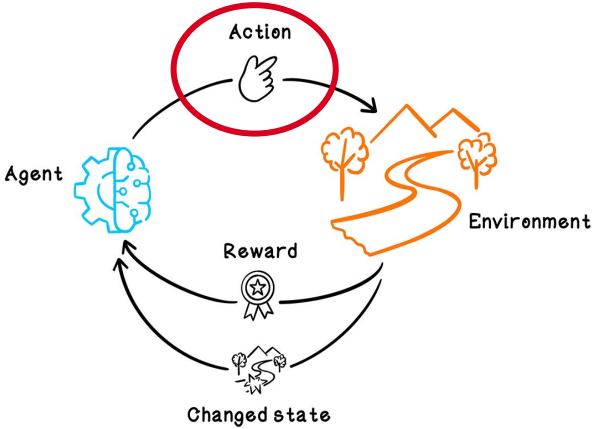
RL cycle - State



<u>A state is defined by:</u>

- The **stored quantity of liquid** at *t*
- The **forecasted electricity** price from *t* to *t + horizon*
- The **forecasted liquid demand** from *t* to *t + horizon*
- The **forecasted gas** demand at *t*
- The current **time** and **day**

RL cycle - Action

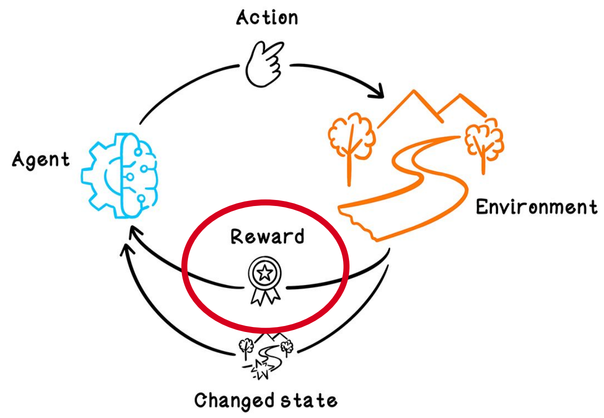


<u>An action is defined by:</u>

- Compressor flow
- LIN liquefier
- LOX liquefier

The **action chosen by the agent** defines the **ASU setpoint**.

RL cycle - Reward



- Compute the **cost of production for each element**
- When necessary, compute the **missing stock** for each element to **satisfy the demand**
- Estimate a **penalization factor** to discourage missing stock

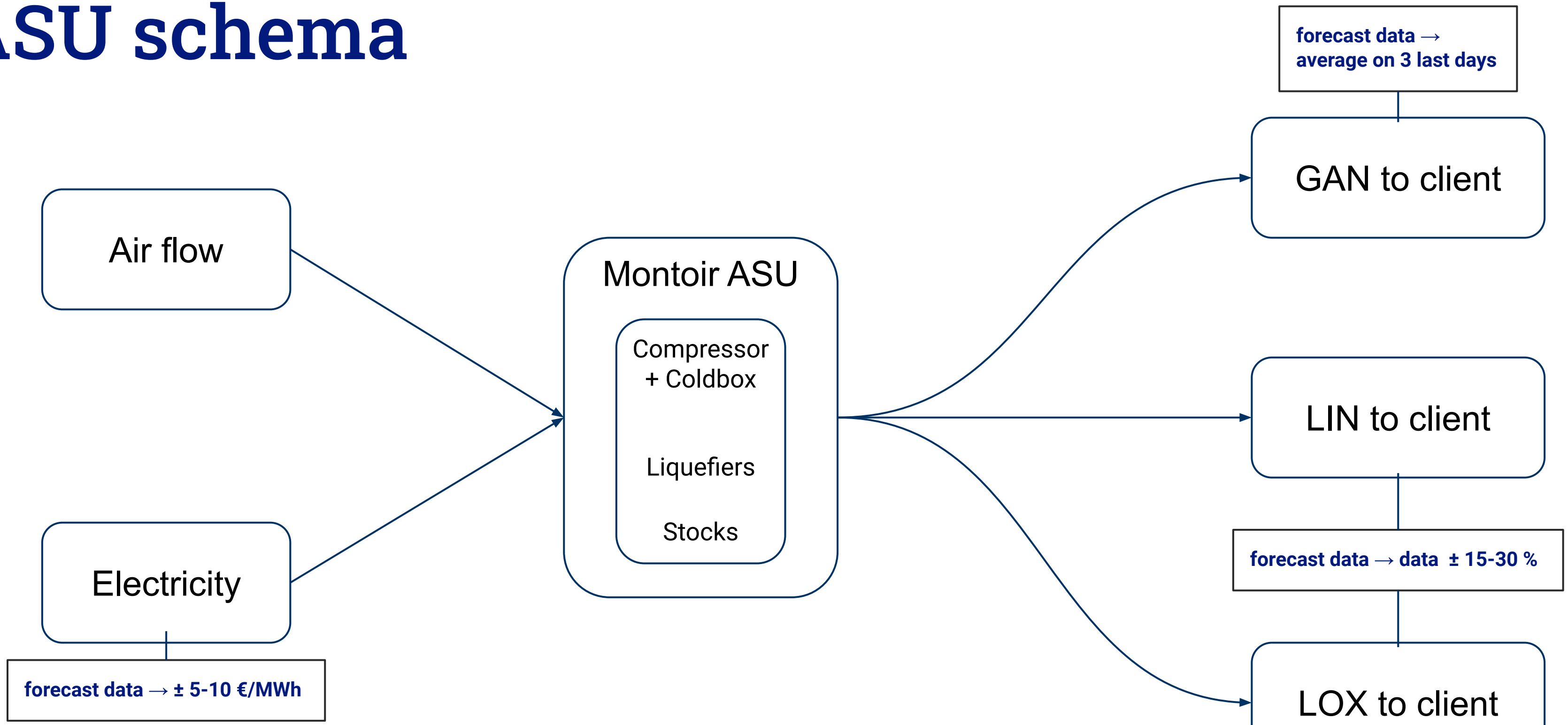
reward = - cost of electricity - penalization cost for missing stock

→ We try to maximize future rewards, ie minimize cost of electricity

paul.berhaut@airliquide.com

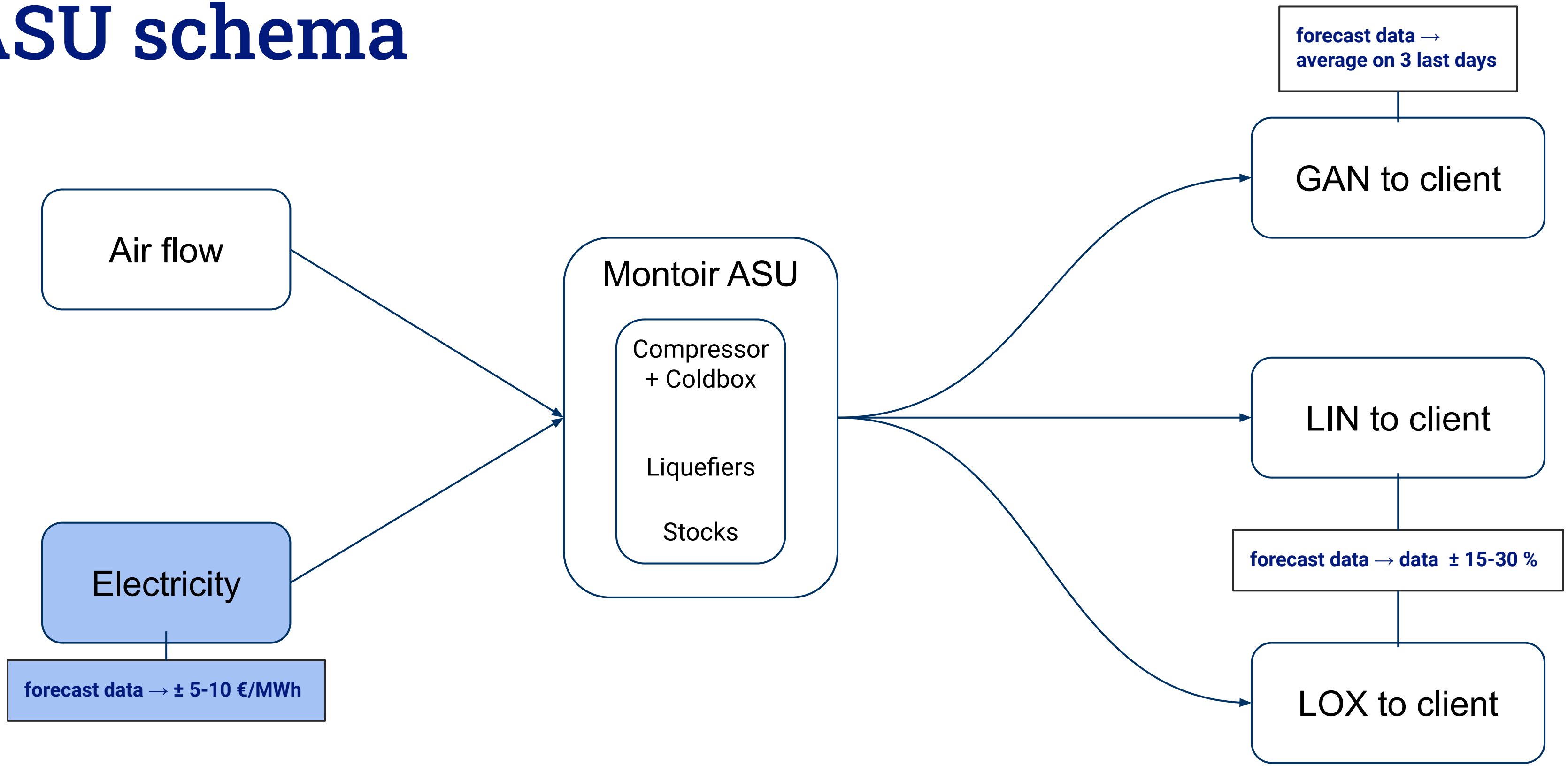
Scenarios and ASU environment

ASU schema



4 years of **hourly** historical data

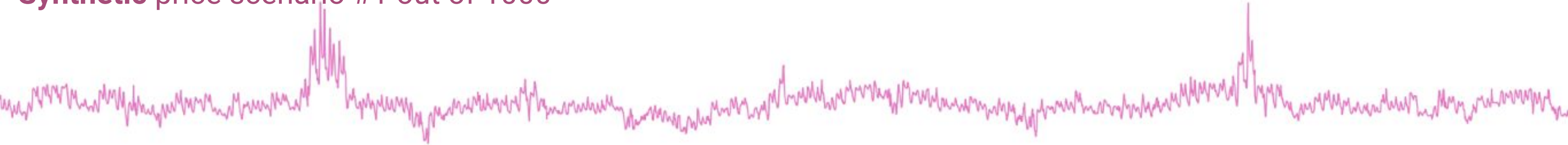
ASU schema



4 years of **hourly** historical data

Real historical electricity price data

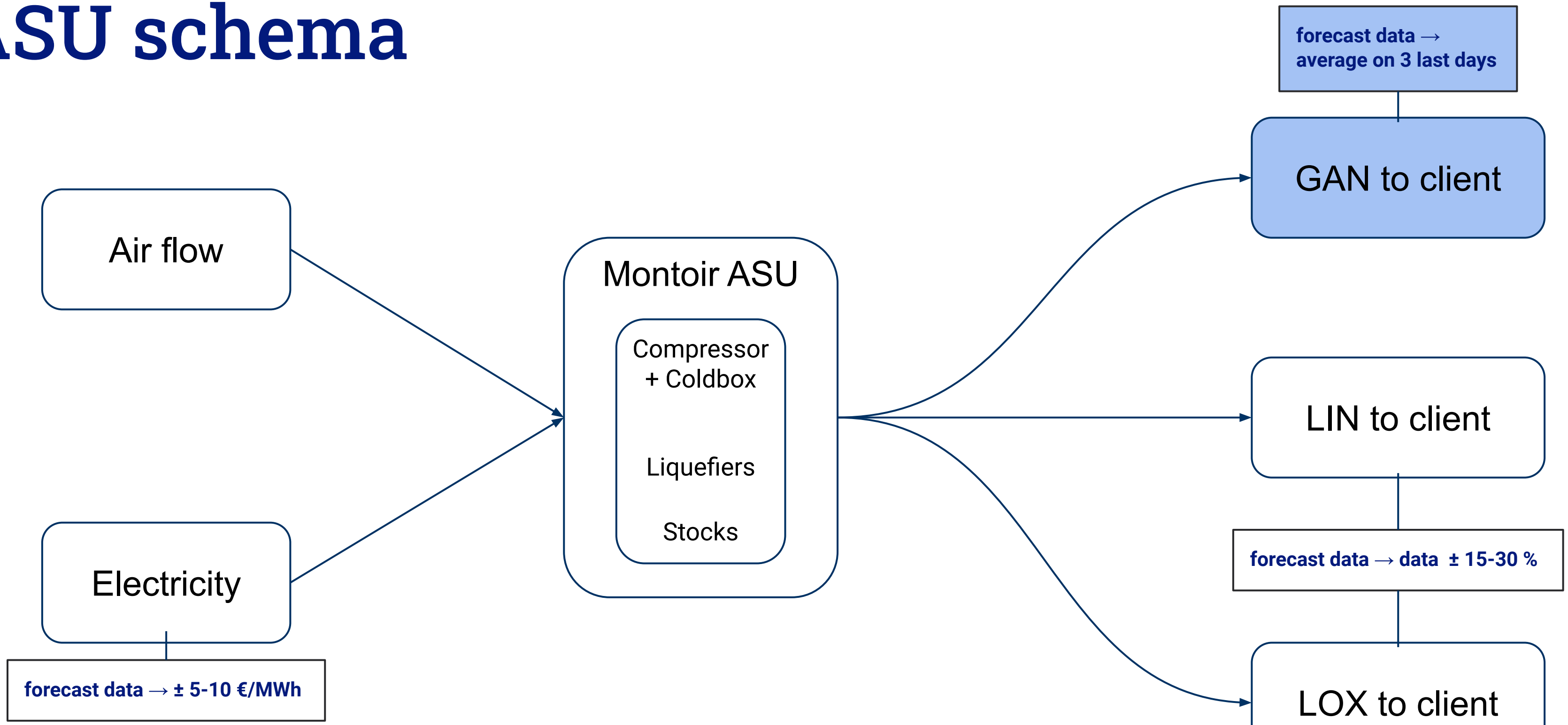
Synthetic price scenario #1 out of 1000



Synthetic price scenario #2 out of 1000

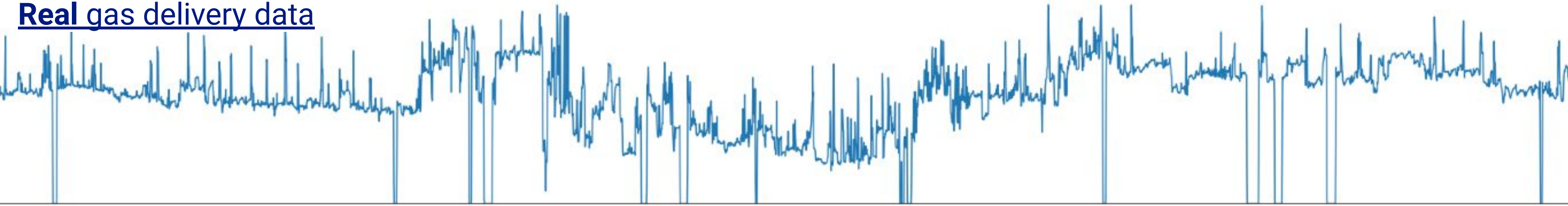
Synthetic price scenario #3 out of 1000

ASU schema

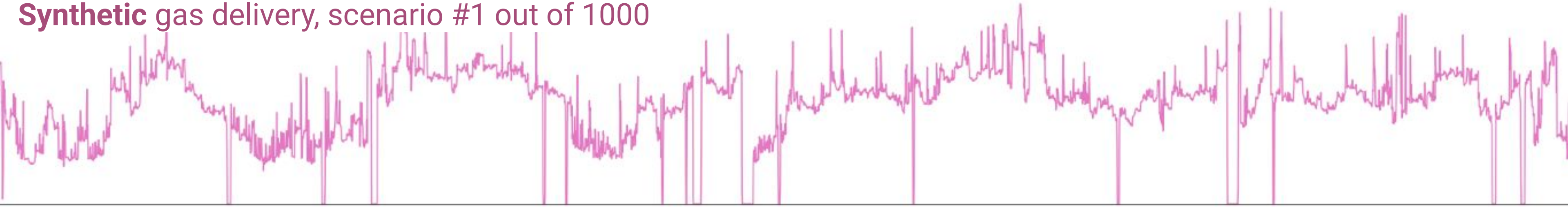


4 years of **hourly** historical data

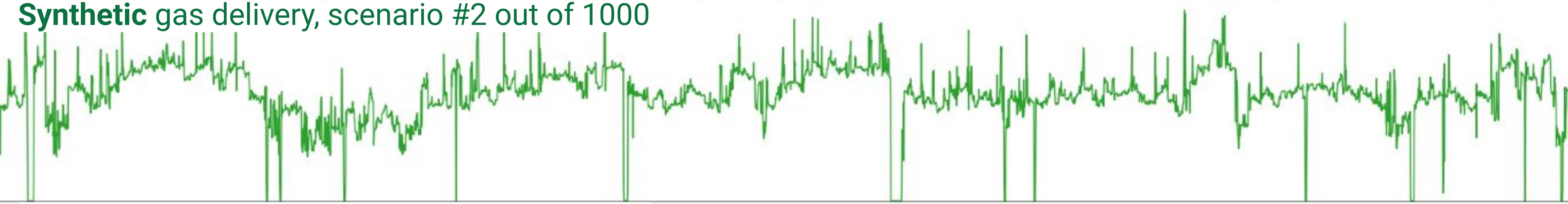
Real gas delivery data



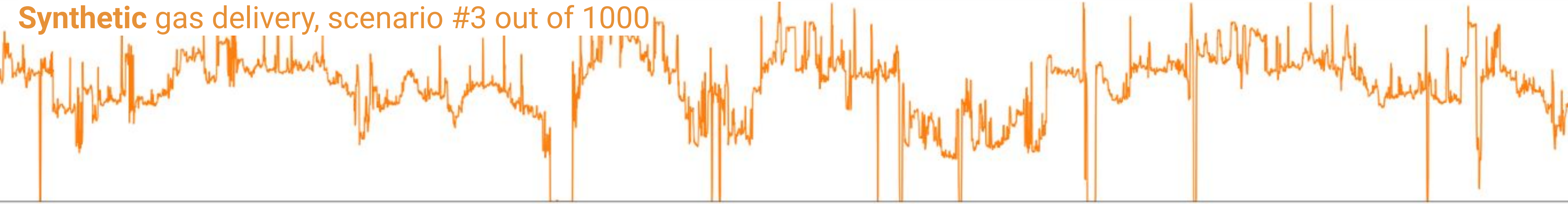
Synthetic gas delivery, scenario #1 out of 1000



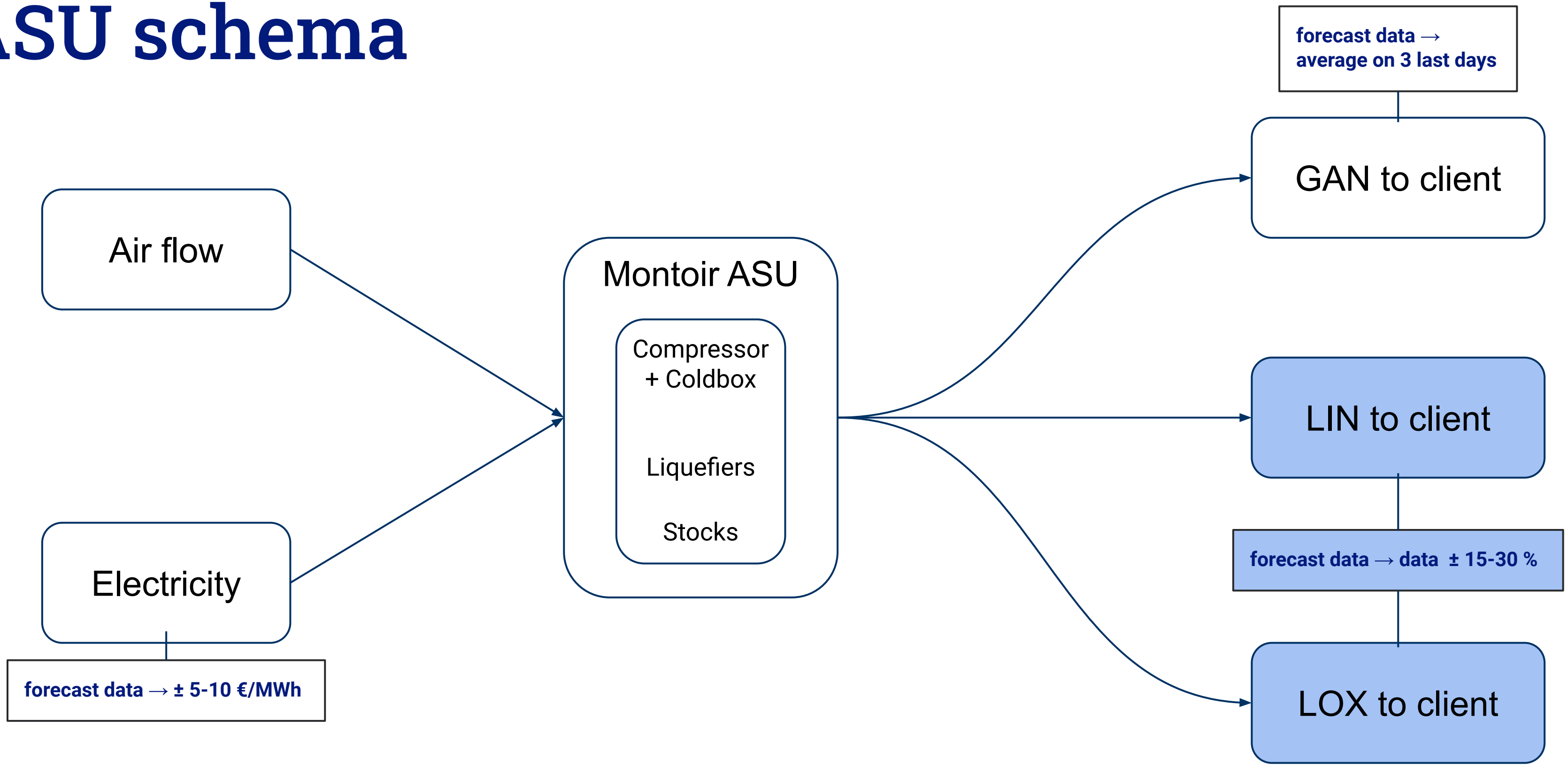
Synthetic gas delivery, scenario #2 out of 1000



Synthetic gas delivery, scenario #3 out of 1000

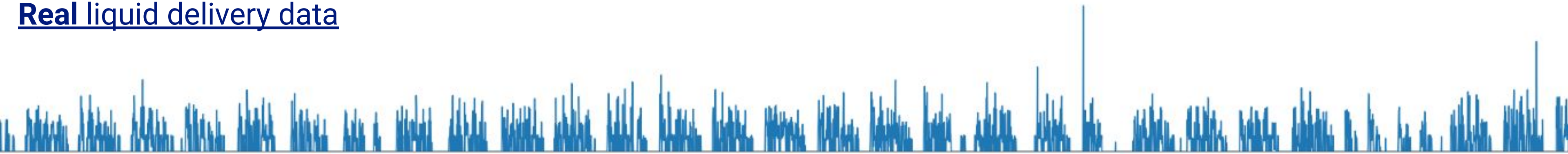


ASU schema



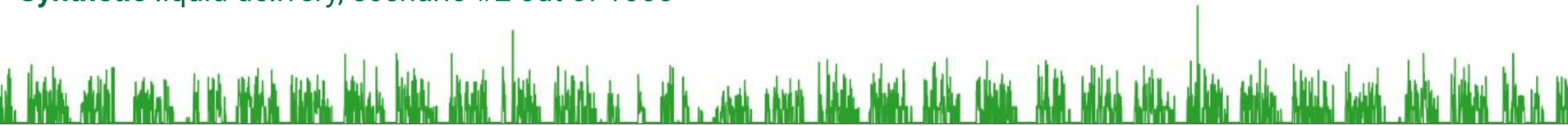
4 years of **hourly** historical data

Real liquid delivery data

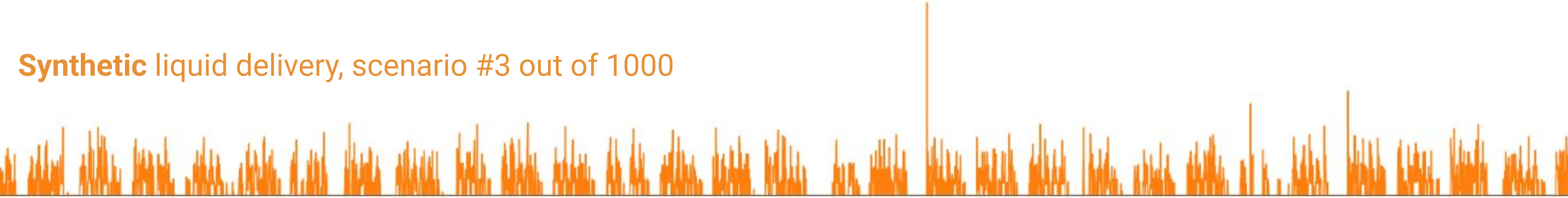


Synthetic liquid delivery, scenario #1 out of 1000

Synthetic liquid delivery, scenario #2 out of 1000



Synthetic liquid delivery, scenario #3 out of 1000



paul.berhaut@airliquide.com

Comparative results

paul.berhaut@airliquide.com

Agent training in the ASU environment

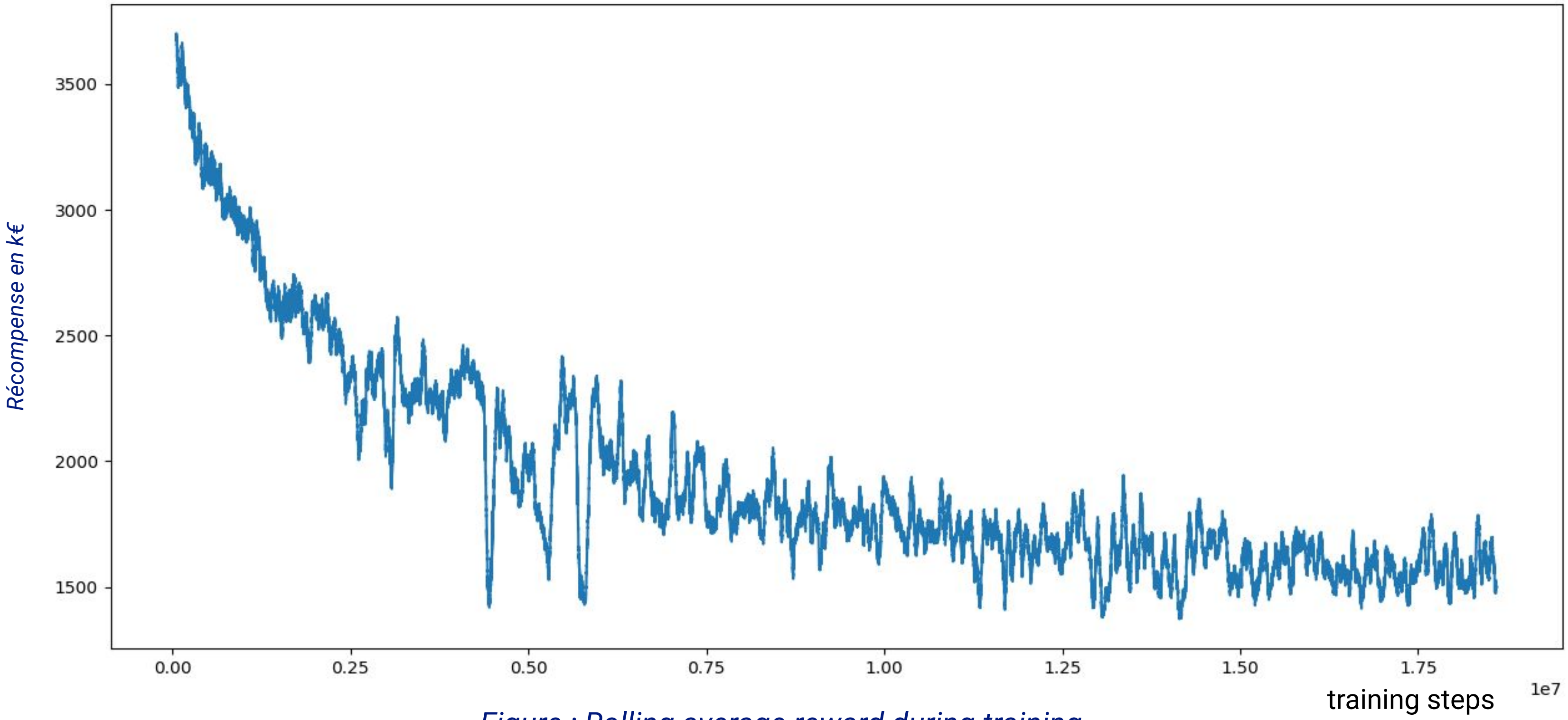


Figure : Rolling average reward during training

paul.berhaut@airliquide.com

Results on test environments

Environment type:

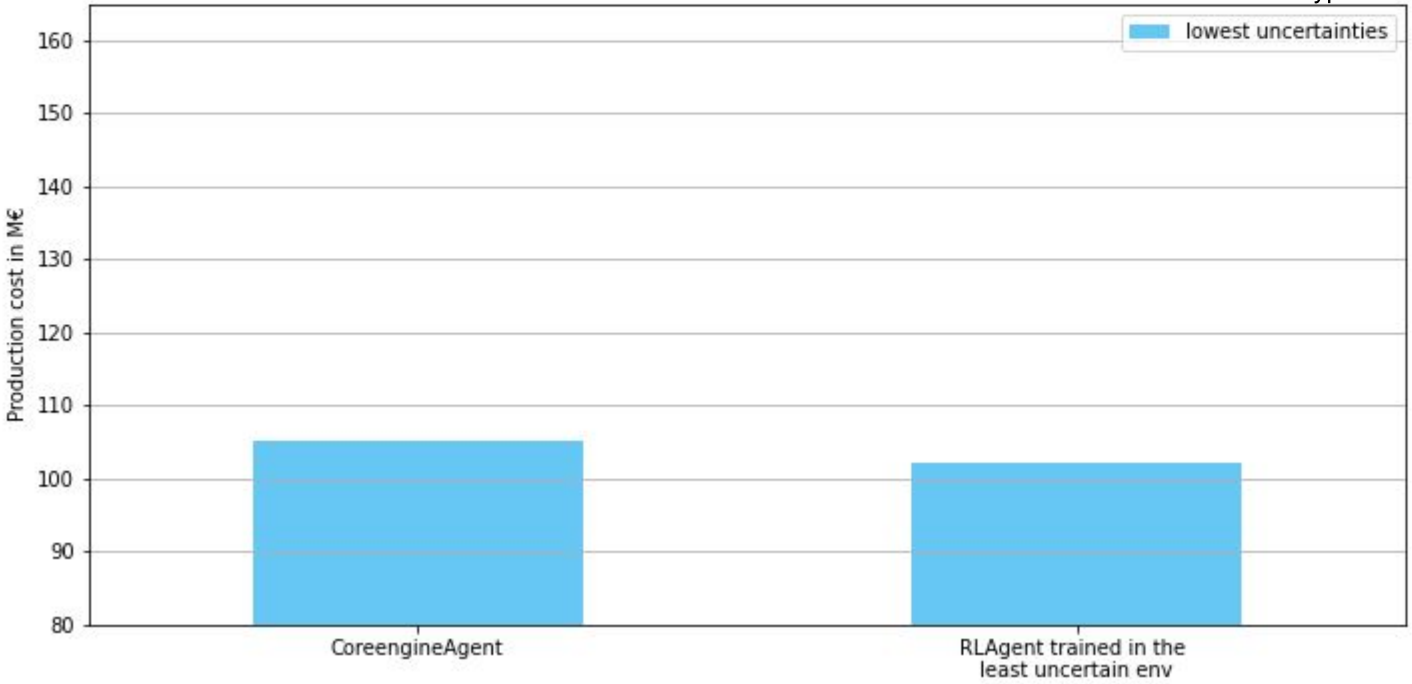


Figure : Metric obtained on environments per agent

Results on test environments

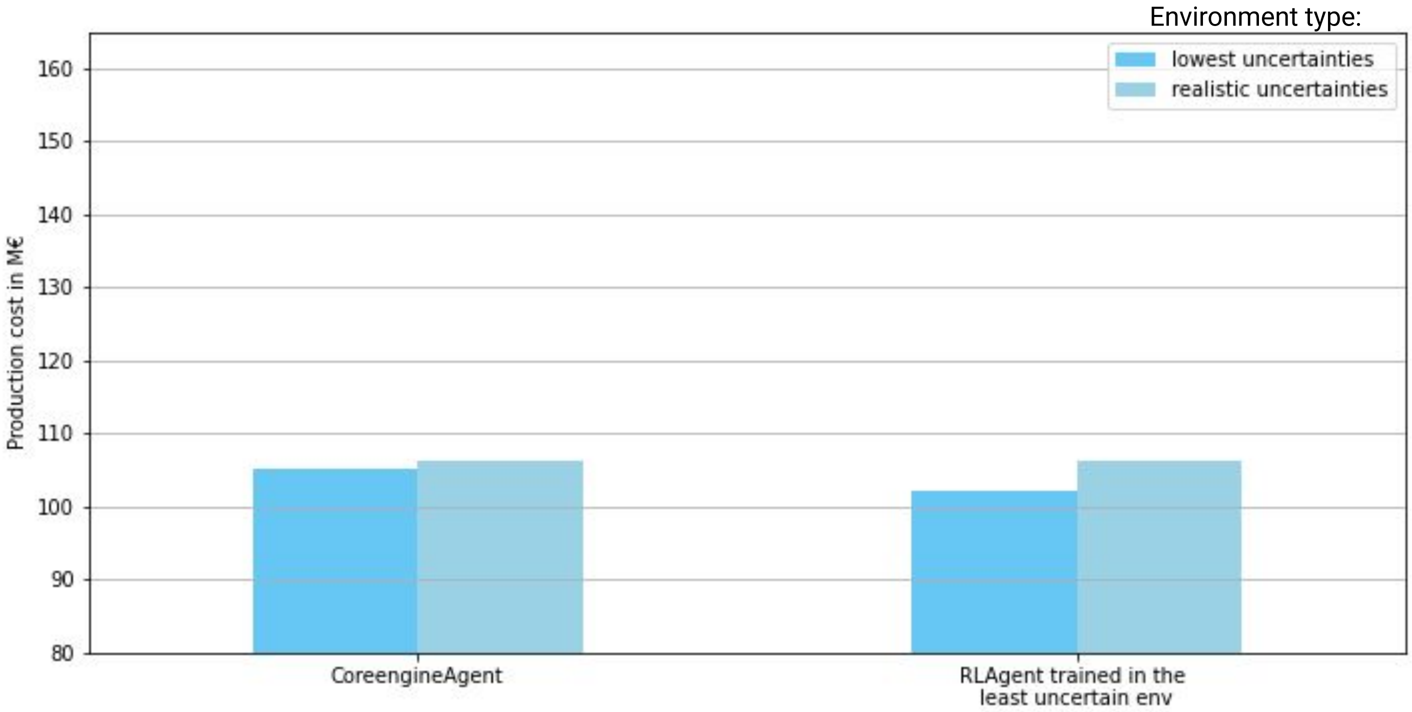


Figure : Metric obtained on environments per agent

Results on test environments

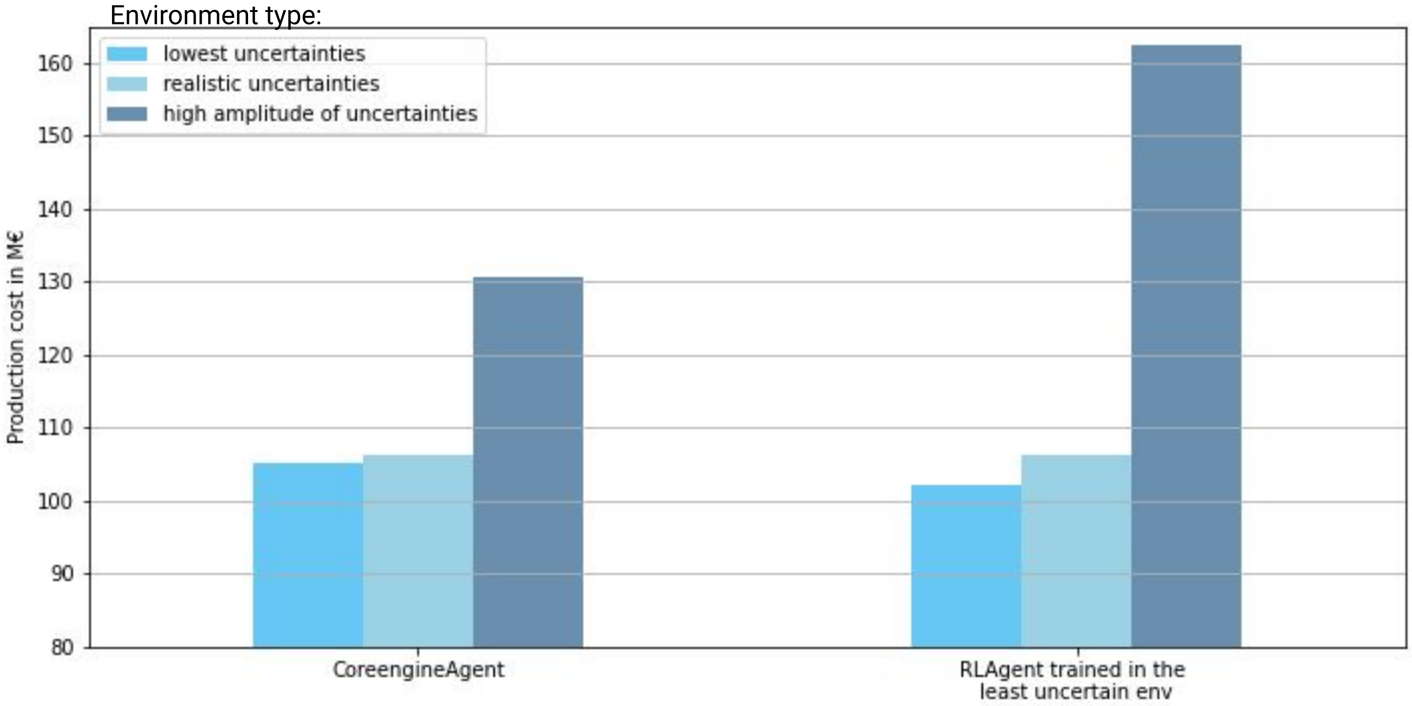


Figure : Metric obtained on environments per agent

paul.berhaut@airliquide.com

Results on test environments

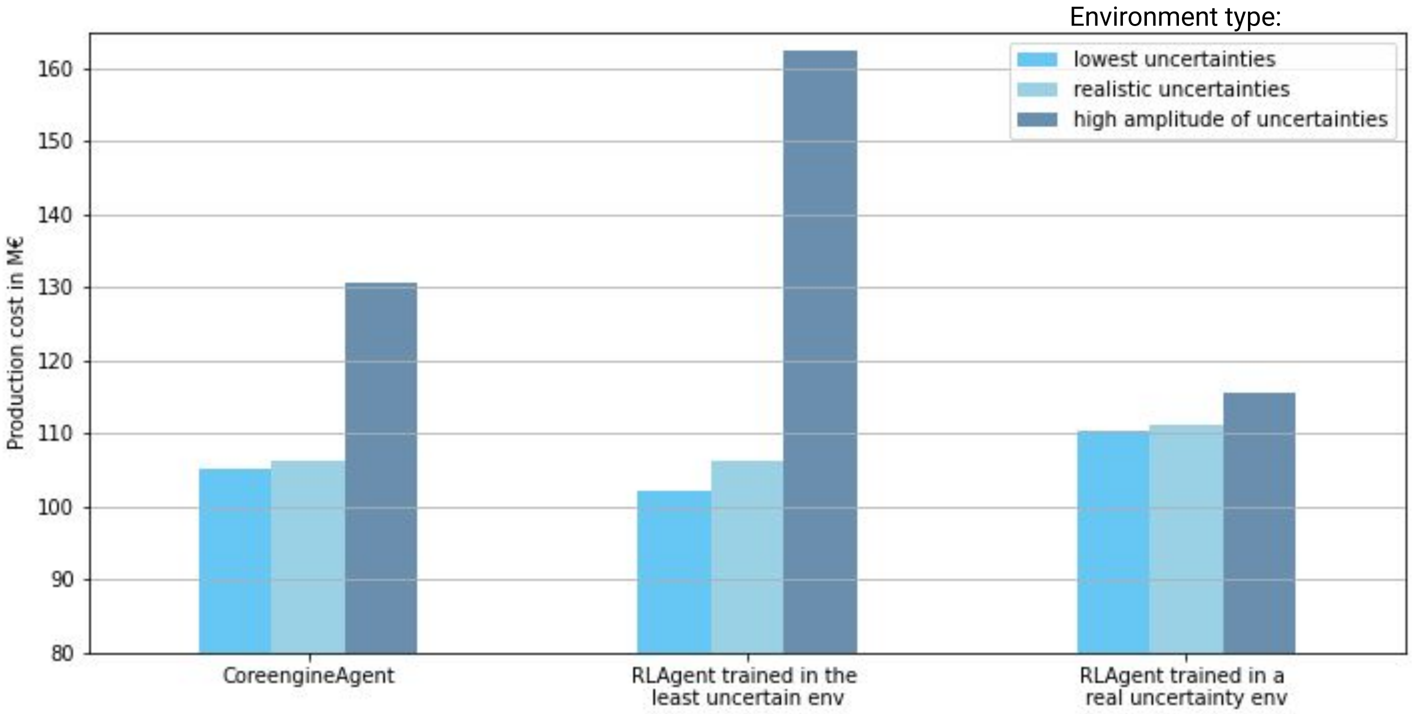
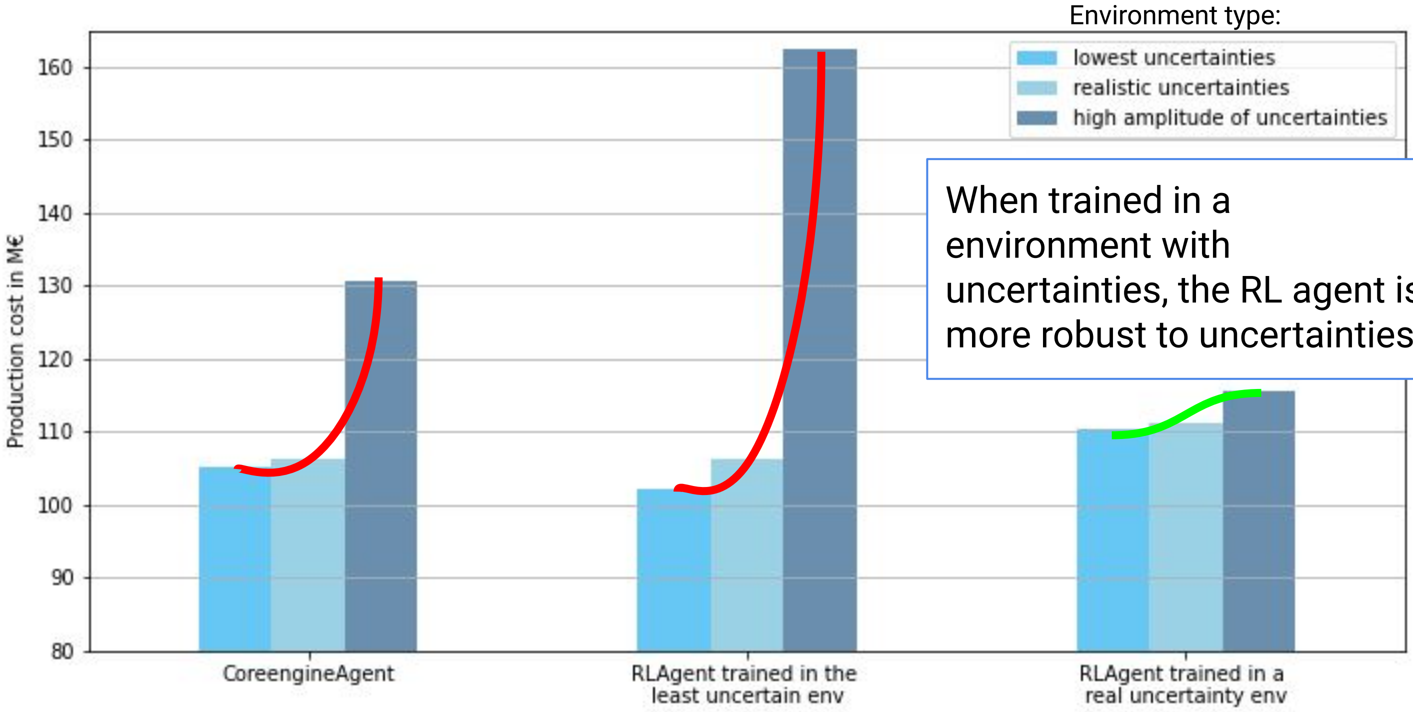


Figure : Metric obtained on environments per agent

Results on test environments



When trained in a environment with uncertainties, the RL agent is more robust to uncertainties

Figure : Metric obtained on environments per agent

Results on test environments

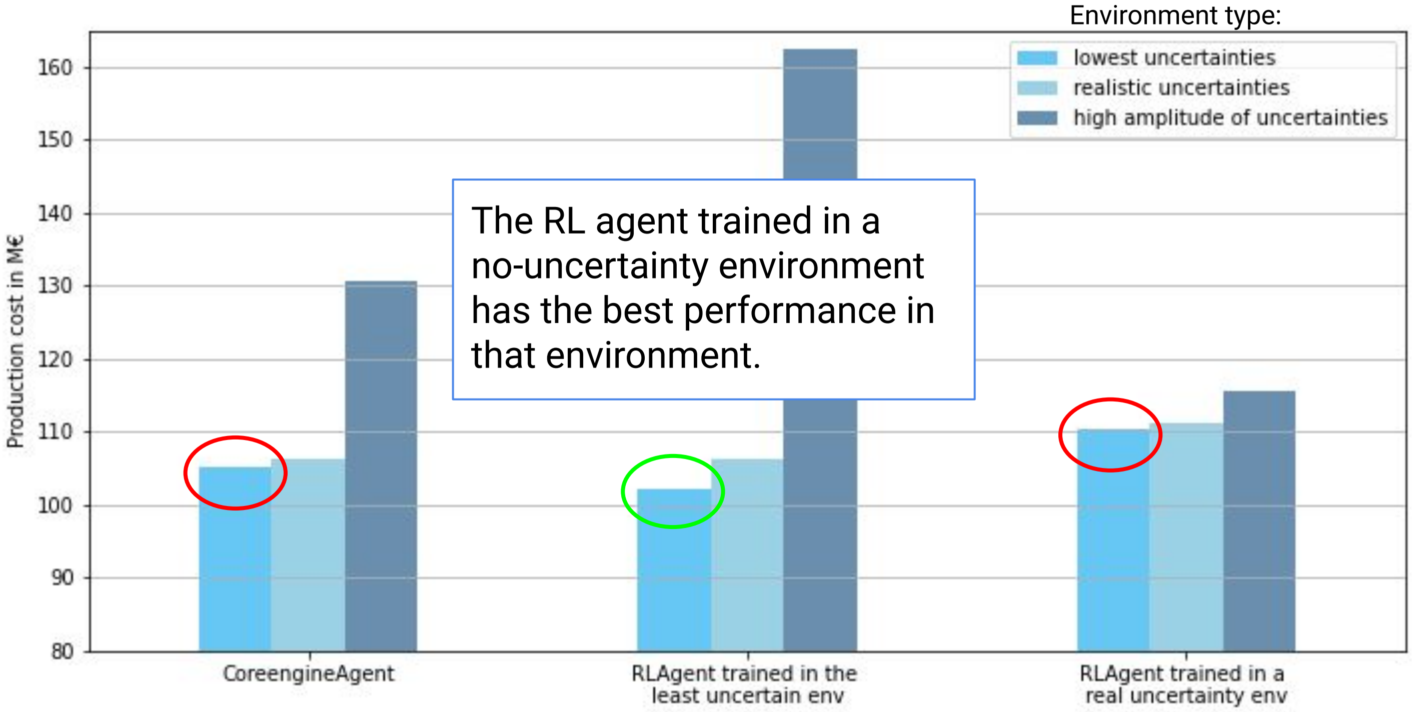
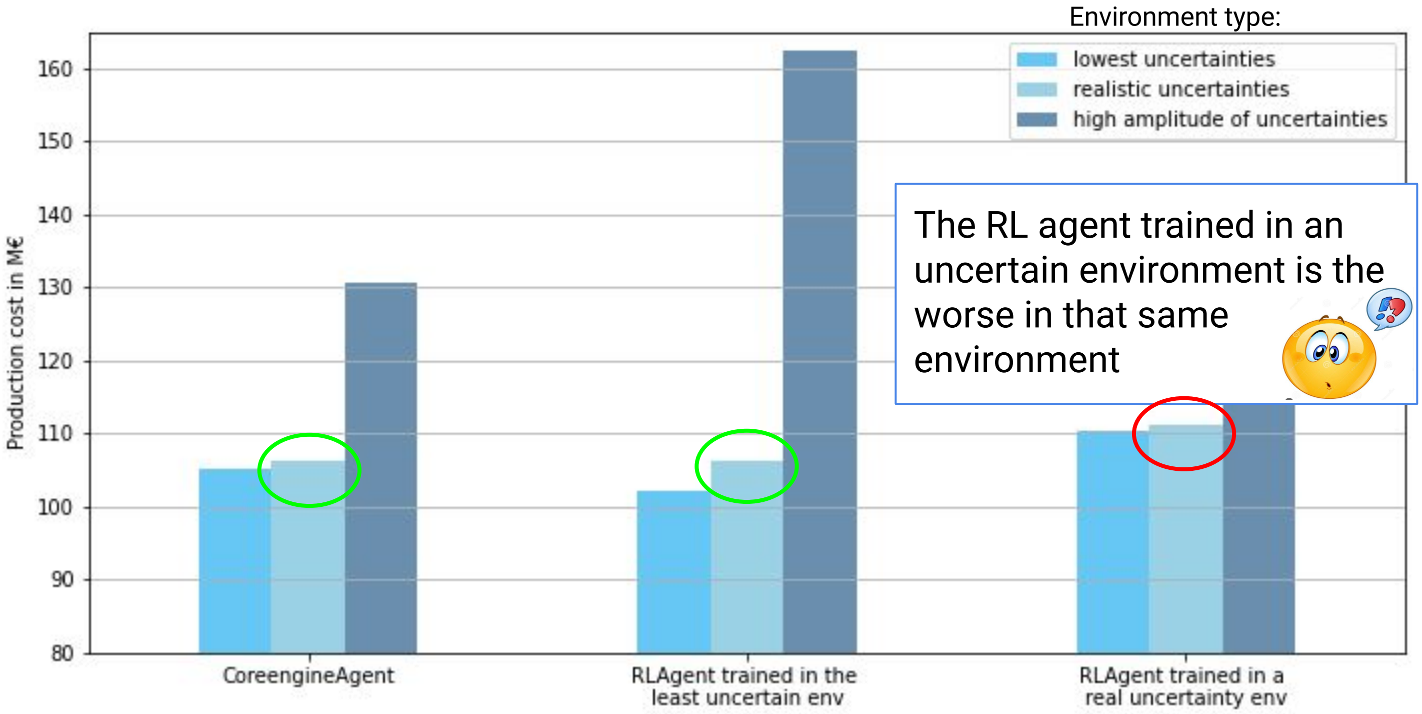


Figure : Metric obtained on environments per agent

Results on test environments



The RL agent trained in an uncertain environment is the worse in that same environment

Figure : Metric obtained on environments per agent

Testing agent with a different metric

Previous metric → **penalizing price = 500 € / MWh**

- Too punitive metric ?
- Penalty quite far from the reality

New metric → **penalizing price = 1.25 * current electricity price**

paul.berhaut@airliquide.com

Results on test environments (second metric)

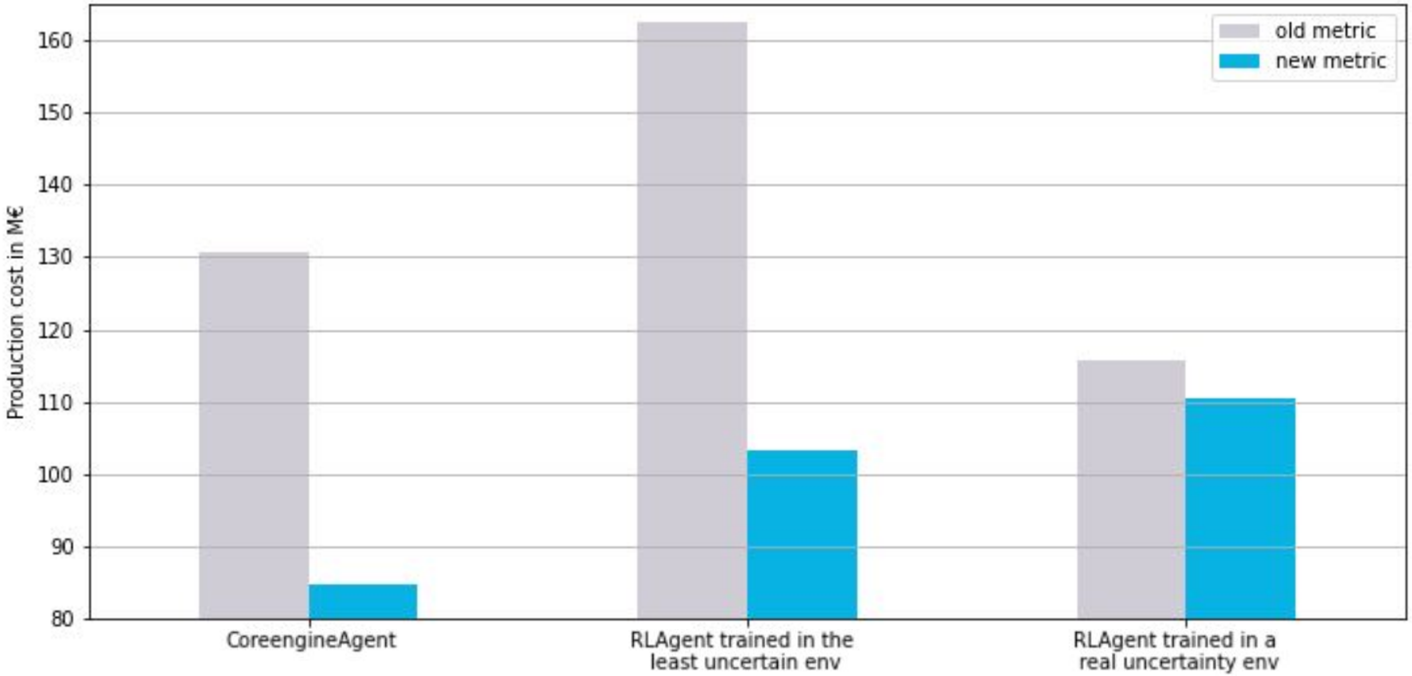


Figure : Both metrics obtained on environment with high uncertainties per agent

Thank you for your attention !

paul.berhaut@airliquide.com