Student Project
Aggregation of experts for day-ahead time-series forecasting

Air Liquide R&D
Data Science Team

2025-2026

5 months

About Air Liquide R&D

Air Liquide is the world leader in gases, technologies and services for industry and health. Present in 80 countries with approximately 65,000 employees, the Group serves more than 3.5 million customers and patients.
Air Liquide places diversity at the heart of its activities and is committed in particular to promoting professional equality and the employment of workers with disabilities.

The Innovation Campus Paris (78 - Les Loges en Josas) **is Air Liquide's largest R&D center**.
It develops innovative solutions for all of the Group's activities. More than 250 researchers work in 48 laboratories equipped with state-of-the-art equipment and experimental platforms, allowing for exploration of numerous research areas.

Project Description

In time series forecasting, combining predictions from multiple models (or "experts") is used to improve accuracy and robustness. The Python library opera provides algorithms from the online learning literature that aggregate expert forecasts based on their historical performance, often coming with theoretical worst-case guarantees.

However, these standard methods are built for a sequential predict -> observe -> update loop. This project aims to investigate the performance of these established algorithms against a simpler, intuitive baseline—an iterative linear regression on a sliding window of past performance. Furthermore, it seeks to adapt these methods for a "day-ahead" forecasting context, common in the energy sector, where this sequential loop is broken.

The project aims to answer the following questions:

- What are the theoretical and practical trade-offs between established online aggregation algorithms (e.g., from opera) and a sliding-window linear regression? This includes comparing theoretical guarantees (e.g., long-term error bounds) versus the implicit "forgetting" mechanism of a sliding window.
- How can online aggregation methods be adapted to a day-ahead forecasting context, where predictions for an entire future period (e.g., the next 24 hours) are made at once, and the ground truth is received in a single block later?
- What is the empirical performance of these different strategies on real-world time series datasets? Do the practical results align with theoretical expectations?

paul.berhaut@airliquide.com / claire.roman@airliquide.com

Student Project
Aggregation of experts for day-ahead time-series forecasting

The project will involve a literature review of expert aggregation methods, followed by the implementation and benchmarking of the different strategies. The evaluation will compare the methods on predictive accuracy, stability of the expert weights, and computational cost.

Deliverables

- A synthetic report on the state-of-the-art of expert aggregation for time series and the theoretical comparison of the studied methods.
- Practical guidelines for practitioners to choose appropriate aggregation strategies, particularly for a day-ahead context.
- A repository containing all the code, experiments, and results.

Bibliography

- Gaillard, P., & Goude, Y. (2022). opera-python: Online Prediction by Expert Aggregation in Python. https://github.com/Dralliag/opera-python
- Cesa-Bianchi, N., & Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge University Press.

✉ paul.berhaut@airliquide.com / claire.roman@airliquide.com