

Predicting Individual Physiological Responses to Pollution Using Transformer-Based Time-Series Models

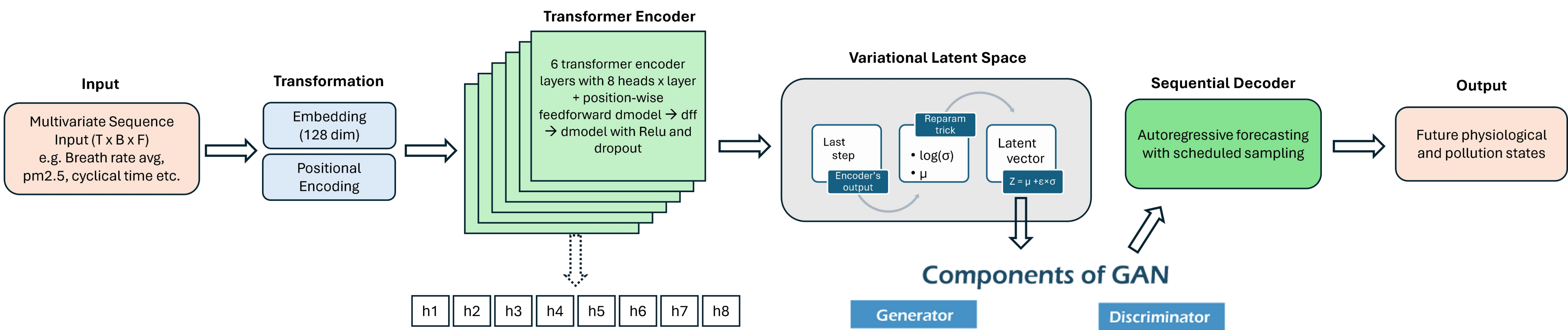
Davide Baino, supervisors:

1. Christopher Pain
2. Boyang Chen

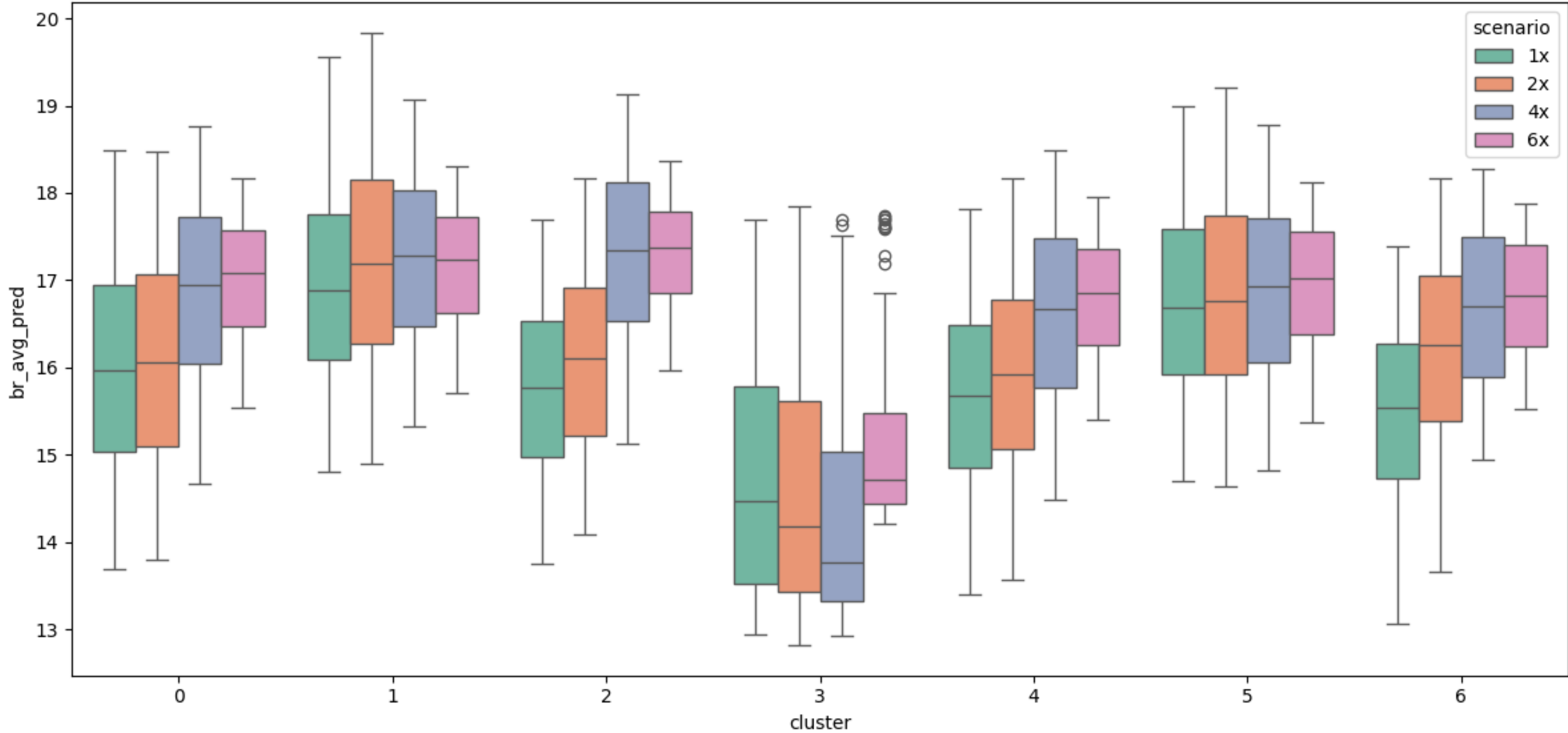
Abstract

Air pollution remains a major global health and environmental concern, contributing to an estimated seven million deaths annually because of the combined effects of outdoor and household exposure (WHO, 2025)[1]. While pollution levels are projected to decline, the ongoing impacts of climate change continue to pose serious risks. Simultaneously, advancements in wearable sensor technologies allow for the systematic collection of high-resolution physiological data over long periods of time (Roos & Slavich, 2023)[2]. This study aims to develop an identity map linking varying levels of air pollution to individual physiological responses. Such a framework will enable the prediction of health responses to pollution exposure, facilitating early warnings and personalised health recommendations. To achieve this, we propose a two-model approach: an initial general model to capture general population temporal trends, and a personalised one specialised on individual characteristics. Together, these models will enhance the precision of forecasting and contribute to more effective, data-driven health interventions when reacting to a polluted environment.

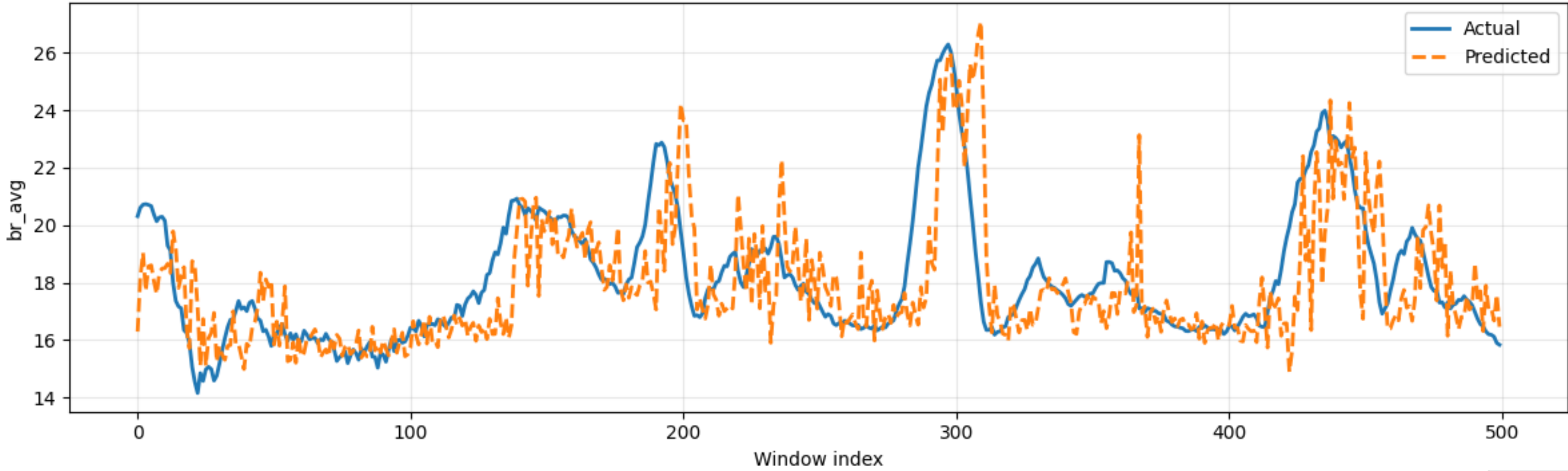
Hybrid Transformer-Gan model for physiological and time-series forecasting



br_avg Forecast Across Pollution Scenarios (10-90% range)



Hourly Forecast — br_avg



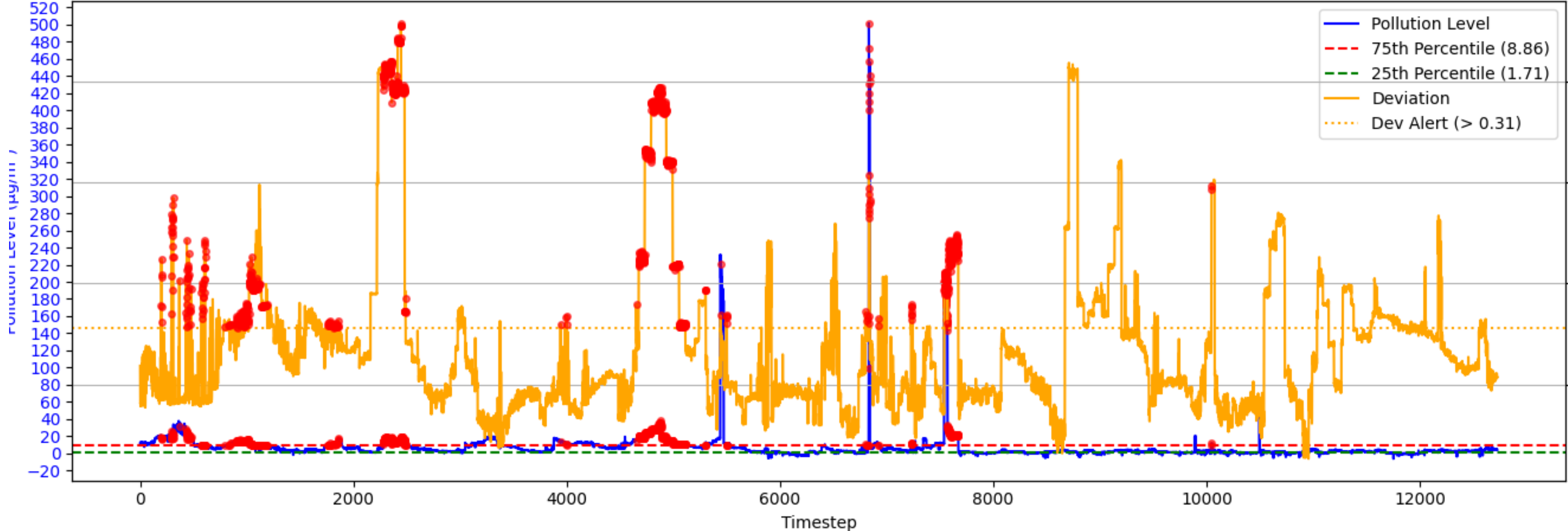
A threshold base system – why is it useful?

Here I will add some comment on clustering – why did we do it? Why is it different than just deciding on what we want?

Something else to say about clustering goes in here

Individual prediction to pollution levels
Why is it important? Why is it useful?

Pollution & Risk Over Time



References

... include here important references (if any)... ..