Civil & Mineral Engineering
UNIVERSITY OF TORONTO

Du B, Siegel JA. 2023. Estimating Indoor Pollutant Loss Using Mass Balances and Unsupervised Clustering to Recognize Decays. *Environmental Science & Technology*, **57(27)**, 10030-10038. DOI: 10.1021/acs.est.3c00756

Abstract

Low-cost air quality monitors are increasingly being deployed in various indoor environments. However, data of high temporal resolution from those sensors are often summarized into a single mean value, with information about pollutant dynamics discarded. Further, low-cost sensors often suffer from limitations such as a lack of absolute accuracy and drift over time. There is a growing interest in utilizing data science and machine learning techniques to overcome those limitations and take full advantage of low-cost sensors. In this study, we developed an unsupervised machine learning model for automatically recognizing decay periods from concentration time series data and estimating pollutant loss rates. The model uses k-means and DBSCAN clustering to extract decays and then mass balance equations to estimate loss rates. Applications on data collected from various environments suggest that the CO_2 loss rate was consistently lower than the $PM_{2.5}$ loss rate in the same environment, while both varied spatially and temporally. Further, detailed protocols were established to select optimal model hyperparameters and filter out results with high uncertainty. Overall, this model provides a novel solution to monitoring pollutant removal rates with potentially wide applications such as evaluating filtration and ventilation and characterizing indoor emission sources.

Highlights

1. Automated semantic segmentation of indoor concentration time series data.
2. Validated using control experiment and tested on CO_2 and PM in various indoor environments.
3. Detailed protocols for optimizing model performance and reducing uncertainties.
4. Coupled with mass balance equations to estimate loss rates and potentially source strengths.

Support provided by:

NSERC CRSNG

INNOVATION Canada Foundation for Innovation Fondation canadienne pour l'innovation

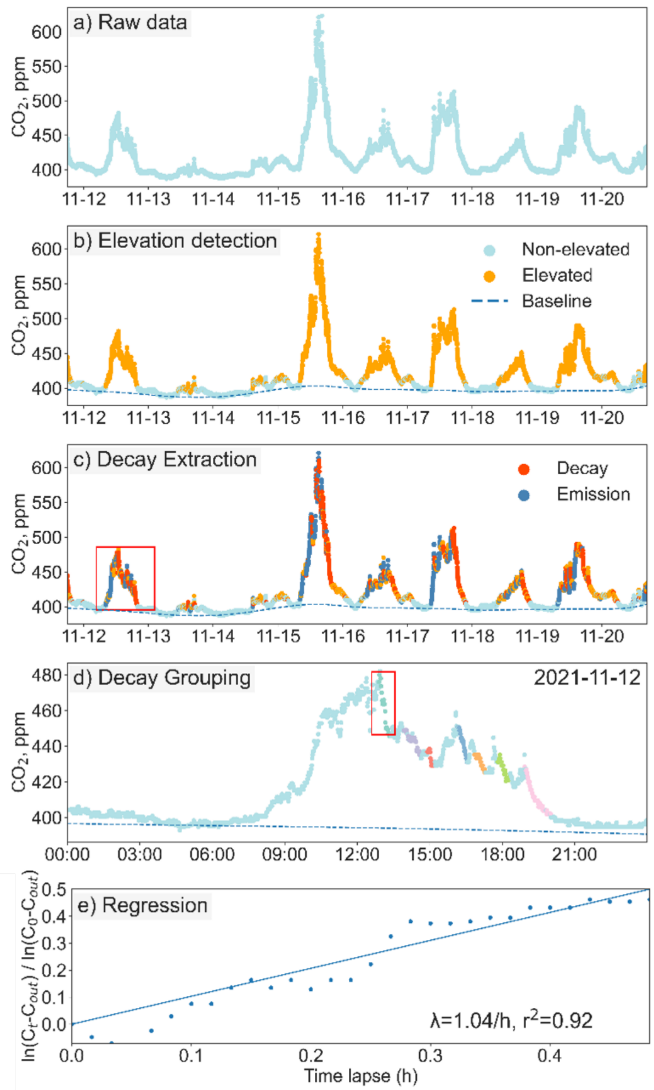

Figure 1: An illustrative example of the key steps of the model based on CO_2 data collected from an office for nine days. Data in Figure S1.d) are extracted from the red box in S1.c), and data in Figure S1.e) are further extracted from the red box in S1.d).