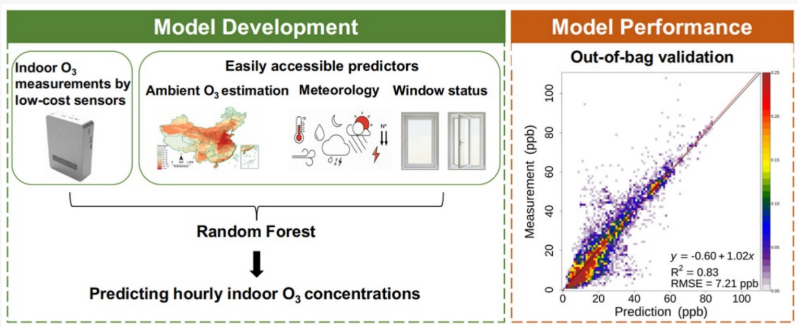
EINPRESSWIRE

AI meets air: Machine learning predicts indoor ozone exposure hour by hour

FAYETTEVILLE, GA, UNITED STATES, December 10, 2025 / EINPresswire.com/ -- Understanding how ozone behaves indoors is vital for assessing human health risks, as people spend most of their time inside. This study developed the first large-scale machine learning model capable of predicting hourly indoor ozone (O_3) concentrations using easily accessible predictors, including outdoor O_3, meteorological conditions, and window-opening behavior.



Machine learning predicts hourly indoor ozone concentrations with high accuracy.

Ozone (O_3) is a key air pollutant formed by chemical reactions between nitrogen oxides and volatile organic compounds under sunlight. In 2021, long-term O_3 exposure contributed to nearly 490,000 deaths worldwide. Although most exposure assessments rely on outdoor data, people typically spend 70%–90% of their time indoors, where ventilation, indoor sources, and building materials all affect actual O_3 levels. Traditional mechanistic models require detailed indoor parameters that are hard to obtain in large-scale studies, while linear regression models struggle with nonlinear environmental relationships. Due to these limitations, there is an urgent need to develop accurate, scalable models that can predict indoor O_3 exposure based on accessible environmental and behavioral data.

Researchers from Fudan University and the Chinese Academy of Sciences have built a machine-learning model to predict hourly indoor O_3 levels across 18 Chinese cities. The study, published (DOI: 10.1016/j.eehl.2025.100170) in Eco-Environment & Health on July 9, 2025, used random forest algorithms trained on low-cost sensor measurements combined with meteorological and ventilation data. By comparing two models—with and without window-status information—the researchers demonstrated that including ventilation behavior significantly improved prediction accuracy, marking a major step toward more realistic O_3 exposure assessment.

The team collected over 8,200 hours of indoor O_3 data using portable electrochemical sensors in 23 households. Predictor variables included outdoor O_3 levels (from high-resolution random-forest and MERRA-2 datasets), meteorological parameters (temperature, humidity, wind, solar radiation, boundary-layer height, and surface pressure), and window-opening status recorded manually by volunteers. Two random forest models were compared: one excluding and one including window status. Incorporating window behavior raised cross-validation R^2 from 0.80 to 0.83 and lowered RMSE from 7.89 to 7.21 ppb. The model accurately captured hourly O_3 fluctuations and regional differences, performing better in southern than northern China and in the cold than warm season. Predictor-importance analysis showed surface pressure, temperature, and ambient O_3 as dominant factors, with ventilation emerging as a crucial behavioral determinant. Diurnal comparisons revealed that indoor O_3 concentrations were 40% lower than outdoor levels during the day, underscoring the buffering effect of indoor environments.

“Most exposure studies still rely on outdoor O_3 data, but that’s only half the story,” said Prof. Xia Meng, senior author of the study. “Our findings show that ventilation behavior—something as simple as whether a window is open or closed—can change exposure dramatically. By integrating such behavioral data with meteorological information through machine learning, we can finally estimate indoor O_3 more precisely at large scales. This will strengthen epidemiological studies and help guide public-health interventions in urban and residential settings.”

This research introduces a practical, low-cost strategy for predicting indoor O_3 exposure in real time across large geographic areas. The model can be integrated into health-risk assessments, smart-home monitoring systems, and public-health surveillance platforms, enabling policymakers and scientists to better understand indoor–outdoor exposure differences. Future work could extend the framework to other pollutants such as fine particulate matter or nitrogen dioxide, incorporate smart sensors for automated window tracking, and expand monitoring to diverse climatic zones. Ultimately, this machine-learning approach bridges environmental modeling with daily life, promoting healthier indoor environments in rapidly urbanizing regions.

References
DOI
10.1016/j.eehl.2025.100170

Original Source URL
https://doi.org/10.1016/j.eehl.2025.100170

Funding Information
This work was funded by the National Natural Science Foundation of China (82003413 and 82030103).

Lucy Wang
BioDesign Research
email us here