



“Gheorghe Asachi” Technical University of Iasi, Romania



EXPLAINABLE AI FOR NATURE-BASED SOLUTIONS: THE HEAF STACKING ENSEMBLE FRAMEWORK FOR MODELING URBAN FOREST AIR POLLUTION MITIGATION

**Reyam Ammar Yasir¹, Sarmad Khaleel Ibrahim^{2*}, Hala Husham Nussrat³,
Taymoor Husham Nussrat⁴**

¹Department of Environmental Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq

²Department of Computer Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq

³Department of Environmental Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq

⁴Financial and Banking Department Baghdad College of Economic Sciences University Baghdad

Abstract

Urban vegetation is an important Nature-based Solution (NbS) to alleviate atmospheric pollution, and measuring species-specific effectiveness is a methodology bottleneck because of the non-linear interaction between the environment and the data heterogeneity. As a response to the limitations of the conventional deterministic models and single-algorithm machine learning, this paper proposes the Hybrid Ensemble AI Framework (HEAF). To address the limited standardized global inventories, we built a high-fidelity synthetic benchmark (n=15,000) built on Auto Arborist and OpenAQ architecture to test the performance of models in a stressful environment. We compared HEAF with three state-of-the-art baselines, XGBoost, Support Vector Regression (SVR), and Random Forest (RF). These findings indicate that HEAF is able to predict better with a Coefficient of Determination (R^2) of 0.9621 and 0.9531 compared to standalone XGBoost ($R^2= 0.9605$) because it stabilizes the prediction variance. More importantly, to ensure the transition between AI as a black-box and the ecological theory, the Ordinary Least Squares (OLS) validation was also performed, which statistically justified Tree Density and Canopy Diameter ($p < 0.001$) as the most significant factors in pollution reduction. These results hereby offer an urban planner with an interpretability centred, reproducible tool when it comes to maximising green infrastructure, and affirm that dense, large-canopy forests can be an important part of ensuring optimal environmental health outcomes.

Key words: air quality modeling, explainable AI (XAI), hybrid framework, nature-based solutions (NbS), stacking ensemble learning

Received: January, 2026; Revised final: May, 2026; Accepted: June, 2026

* Author to whom all correspondence should be addressed: E-mail: eng_sarmadnet@uomustansiriyah.edu.iq