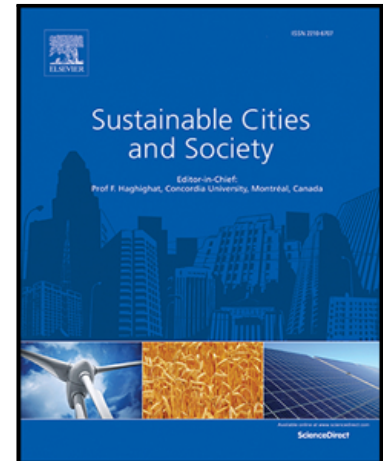


Air-quality-sensitive urban design on NO₂ Mitigation in the high-rise high-density City of London: A machine-learning land-use regression analysis

Vasilis Kazakos , Zhiwen Luo , Said Munir , Julie Fletcher , Juliet Davis , Qihong Deng

PII: S2210-6707(26)00883-8
DOI: <https://doi.org/10.1016/j.scs.2026.108008>
Reference: SCS 108008



To appear in: *Sustainable Cities and Society*

Received date: 8 March 2026
Revised date: 8 September 2026
Accepted date: 10 October 2026

Please cite this article as: Vasilis Kazakos , Zhiwen Luo , Said Munir , Julie Fletcher , Juliet Davis , Qihong Deng , Air-quality-sensitive urban design on NO₂ Mitigation in the high-rise high-density City of London: A machine-learning land-use regression analysis, *Sustainable Cities and Society* (2026), doi: <https://doi.org/10.1016/j.scs.2026.108008>

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

HIGHLIGHT

- Random Forest LUR model predicted urban NO₂ with high accuracy (Adj. R² = 0.88).
- Increased building height variability reduced local NO₂ by up to ~7 µg m⁻³.
- Targeted tree canopy expansion reduced hotspots but had complex local effects.
- Combined design and traffic interventions delivered the largest air quality gains.
- Machine learning LUR enables fine-scale appraisal of urban design scenarios.

Air-quality-sensitive urban design on NO₂ Mitigation in the high-rise high-density City of London: A machine-learning land-use regression analysis

Vasilis Kazakos¹, Zhiwen Luo^{2*}, Said Munir³, Julie Fletcher⁴, Juliet Davis², Qihong Deng^{5*}

¹School of Built Environment, University of Reading, UK

²Welsh School of Architecture, Cardiff University, Cardiff, UK

³Faculty of Environment, University of Leeds, Leeds, UK

⁴Urban Generation, London, UK

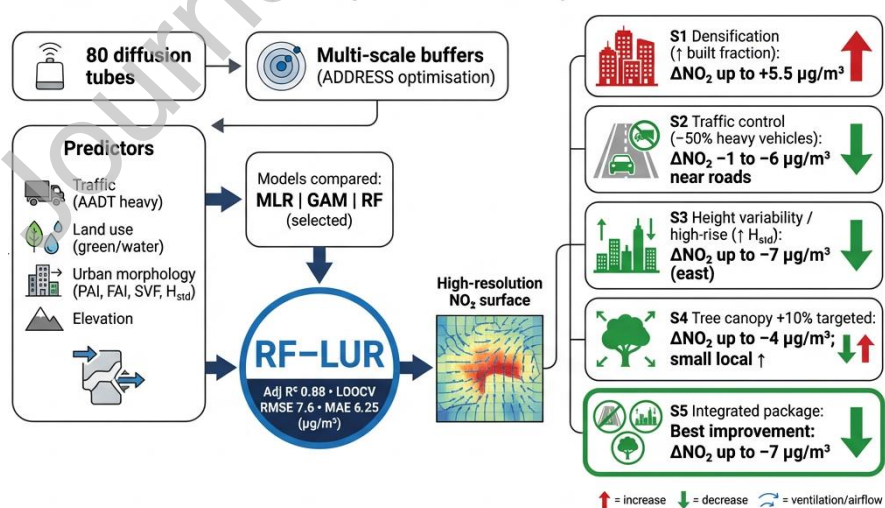
⁵School of Public Health, Zhengzhou University, China

Word count of abstract: 241

Word count of text: 5666

* Correspondence: Prof Zhiwen Luo (LuoZ18@Cardiff.ac.uk); Prof Qihong Deng (qhdeng@zzu.edu.cn)

Graphical Abstract



Abstract

Elevated nitrogen dioxide (NO₂) remains a persistent urban air-quality and public-health concern. In the UK, despite progressive emissions-control policies, exceedances have historically occurred in central London. This study quantifies neighbourhood-scale NO₂ variability in the high-rise, high-density City of London and evaluates how air-quality-sensitive urban design, alongside traffic measures, influences concentrations in a morphology-constrained street network. We developed machine-learning land-use regression (LUR) models to predict NO₂ distribution: the Random Forest (RF) model performed best (adjusted R² = 0.88; RMSE = 7.6 µg m⁻³; MAE = 6.25 µg m⁻³). The RF-LUR model was then applied to assess air-quality-sensitive urban design scenarios. Increased built fraction associated with population growth was associated with widespread NO₂ increases (up to 5.5 µg m⁻³), consistent with reduced dispersion in dense street-canyons. A 50% reduction in heavy-vehicle activity was associated with corridor-focused NO₂ decreases of 1–6 µg m⁻³. Introducing additional high-rise buildings increased height variability and was associated with NO₂ reductions up to ~7 µg m⁻³ in the eastern cluster. Targeted tree-canopy expansion was associated with hotspot reductions of up to ~4 µg m⁻³ but produced small local increases in some areas. The combined intervention package was associated with the largest overall improvement, with NO₂ reductions up to ~7 µg m⁻³. The results show that integrating urban morphology within machine-learning LUR models can provide a practical, fine-scale evidence base for spatial targeting and comparative appraisal of NO₂ mitigation options in high-density urban districts.

Keywords: Air Quality; NO₂ pollution; Urban Design; Urban Morphology; Land-Use Regression; Machine-learning.

1 Introduction

Traffic-related air pollution remains a persistent public-health concern in large cities and continues to contribute substantially to adverse health impacts across Europe (European Environment Agency, 2022), where exposure varies sharply over short distances and is shaped jointly by emissions and local dispersion conditions (Samoli et al., 2016; Fecht et al., 2016; Carey et al., 2016; Smith et al., 2017). These gradients are often strongest in dense urban cores. In the City of London, high traffic activity coincides with narrow streets, frequent junctions, and vertically complex development, creating street-canyon conditions that constrain ventilation and amplify neighbourhood-scale variability in NO_2 . Improving exposure characterisation in such environments is important both for urban epidemiology and for planning interventions that are feasible, spatially targeted, and sensitive to local urban form.

The spatial variability of NO_2 is particularly pronounced in dense urban centres, where concentrations can change substantially over distances of only a few tens of metres. In such environments, exposure patterns are influenced not only by traffic intensity but also by local street geometry, building configuration and ventilation potential. Previous studies have shown that street-canyon characteristics, including building height, aspect ratio and enclosure, can strongly influence pollutant accumulation and create exposure gradients that are not captured by regional-scale assessments (Lee et al., 2017; Jeanjean et al., 2017; Buccolieri et al., 2018). Consequently, understanding NO_2 variability in compact city centres requires approaches capable of representing both emission sources and the influence of urban form.

Recent research has increasingly highlighted the role of urban design and morphology in shaping air quality, not only through the spatial distribution of emissions but also via aerodynamic and micro-climatic processes that govern pollutant dispersion and exposure. Using dispersion modelling (ADMS-Roads), Metrangolo et al. (2024) showed that urban morphological parameters (e.g., planar area index, aspect ratio and mean building height) are strongly linked to concentration patterns, reinforcing the physical influence of built form on near-road pollution fields. At the regional scale, Augusto et al. (2025) demonstrated with WRF–CAMx that urban compaction scenarios can reduce annual NO₂ concentrations by around 17–20%, whereas an extreme dispersed morphology can increase NO₂ substantially (reported close to +98%), illustrating that morphological change can be comparable in magnitude to (or even exceed) emission-only controls depending on the scenario. Complementing these modelling studies, Ling et al. (2025) further emphasised that urban form and environmental characteristics jointly drive exposure variability and inequality, supporting the need for integrated planning responses rather than single-sector measures. Finally, a recent systematic review synthesising 79 studies concluded that vegetation effects on air quality are strongly contingent on urban form and street-scale configuration, and that combining greening with morphological considerations tends to outperform interventions considered in isolation (Yao et al., 2025).

Although computational fluid dynamics (CFD) and Large Eddy Simulation (LES) approaches can provide detailed representations of urban airflow and pollutant dispersion, their computational requirements often limit their application to individual neighbourhoods, street canyons or development sites. Such methods are therefore less suited to rapid assessment of multiple city-scale planning scenarios. Land-use regression approaches offer a complementary alternative by enabling efficient prediction of spatial air-pollution

patterns across large urban areas while requiring substantially lower computational resources. However, conventional LUR models often simplify the representation of urban morphology and are rarely applied to evaluate future planning interventions.

The objective of this study was not to reproduce detailed street-scale airflow processes at individual locations, but rather to evaluate how alternative urban-planning and transport interventions may influence the spatial distribution of NO₂ across the entire City of London. This required a modelling framework capable of integrating observations from multiple monitoring locations, urban-form indicators and future planning scenarios within a computationally efficient workflow. LUR was therefore considered particularly suitable because it enables fine-scale spatial prediction across the whole study area while maintaining direct links between NO₂ concentrations and planning-relevant variables such as traffic activity, building density, vegetation and urban morphology. To address some of the limitations of conventional LUR in complex urban environments, machine-learning techniques and detailed morphology indicators were incorporated into the modelling framework.

Land-use regression (LUR) has traditionally been used for exposure assessment, estimating spatial variation in long-term pollutant concentrations by relating observations to land-use, traffic and other spatial predictors (Hoek et al., 2008; Beelen et al., 2014; Li et al., 2021). Conventional LUR models commonly use multiple linear regression, which provides relatively interpretable relationships but may be less able to represent the non-linearities and interactions characteristic of complex high-density urban environments. Machine-learning approaches, including Random Forest and Gradient Boosting, extend the LUR framework by capturing non-linear relationships and interactions among predictors and

have demonstrated improved predictive performance for fine-scale air-pollution mapping (Brokamp et al., 2017; Wang et al., 2018; Liu et al., 2021; Criado et al., 2023).

However, improved prediction under observed conditions does not necessarily establish that an empirical LUR model is suitable for predicting the effects of hypothetical planning interventions. Most LUR applications are designed primarily for spatial prediction or exposure assessment under conditions represented by the calibration data. Using such models for planning-scenario evaluation requires predictor values to be modified and the fitted relationships to be applied to conditions that may differ from those observed during model development. This distinction is important because both conventional and machine-learning LUR models remain empirical and correlational: they do not explicitly simulate the physical processes through which changes in traffic, building morphology or vegetation affect emissions, airflow, dispersion or atmospheric chemistry. Machine learning can improve representation of complex statistical relationships, but does not by itself convert those relationships into causal or mechanistic effects.

This creates an important methodological gap between the established use of LUR for exposure mapping and its potential use as a planning-support tool. The present study addresses this gap by evaluating whether a morphology-sensitive machine-learning LUR framework can be used for exploratory, comparative assessment of changes in planning-relevant predictors across a high-rise, high-density urban area. Rather than treating the resulting scenarios as causal forecasts of future air quality, the approach is used to examine how the fitted RF-LUR model responds to specified changes in urban form, vegetation and traffic relative to the baseline conditions. The scenario analysis therefore provides a computationally efficient means of comparing spatial responses and identifying

potentially sensitive locations, while recognising that mechanistic modelling would be required to establish the physical effects of specific interventions.

Evidence from dense urban contexts indicates that refined predictor sets, including morphological indicators, can enhance LUR performance and better reflect intra-urban variability (Sabaliauskas et al., 2015; Lee et al., 2017; Dimakopoulou et al., 2022). Recent work also highlights associations between building coverage/diversity and pollutant concentrations (Wan et al., 2022). Yet, neighbourhood-scale evidence remains limited on how design-relevant morphology features (e.g., density and height configuration) interact with traffic to shape NO₂ patterns in highly compact city centres, and how such evidence can be operationalised into planning scenarios.

Furthermore, while machine-learning-enhanced LUR models have demonstrated improved predictive performance, relatively few studies have combined detailed urban morphology indicators with scenario-based assessment of future planning interventions. Consequently, there remains a need for modelling frameworks capable of linking exposure prediction with practical evaluation of urban design, transport and greening strategies at the neighbourhood scale. Although LUR models are widely used for exposure assessment and spatial prediction, their application to urban-planning and design scenarios remains relatively limited. Most studies employ LUR to characterise existing air-quality patterns, whereas fewer have explored how empirically derived relationships can be used to evaluate hypothetical changes in traffic activity, urban form or green infrastructure. As a result, there is limited evidence regarding the usefulness and limitations of LUR-based frameworks as decision-support tools for assessing the potential air-quality implications of alternative planning strategies. Despite recent advances in machine-learning-based air-

quality modelling and increasing recognition of the importance of urban morphology (Dimakopoulou et al., 2022; Augusto et al., 2025; Ling et al., 2025), relatively few studies have combined these developments within a framework specifically designed to support urban-planning decision making. Existing studies have largely focused on improving prediction accuracy or understanding individual determinants of air pollution, whereas fewer have explored how morphology-sensitive machine-learning LUR models can be used to evaluate future planning, transport and greening interventions. This study addresses that gap by integrating detailed urban morphology indicators, Random Forest-based LUR modelling and scenario-based assessment within a single framework for a dense city-centre environment.

This study develops a neighbourhood-scale modelling and scenario-screening framework for the City of London that integrates urban morphology within an LUR approach and applies machine-learning methods to better represent non-linear and interaction effects. We address two questions: (1) how do urban design elements influence the spatial distribution and intensity of traffic-related NO₂ in a high-density, high-rise setting? and (2) which emissions- and design-oriented interventions are likely to be most effective for reducing concentrations within a morphology-constrained street network? By linking fine-scale exposure estimation to scenario-based evaluation, the study aims to support both urban health research and planning-relevant air-quality management in dense city centres.

2 Methods

The overall methodological framework adopted in this study is summarised in Figure 1. The workflow comprises data preparation, predictor-variable selection using the ADDRESS

procedure (Su et al., 2009), development and evaluation of alternative LUR models, selection of the optimal RF-LUR model, implementation of future urban-planning scenarios, and assessment of their effects on predicted NO₂ concentrations. This Scenario Assessment Process framework was designed to support rapid evaluation of potential planning and mitigation interventions within the City of London.

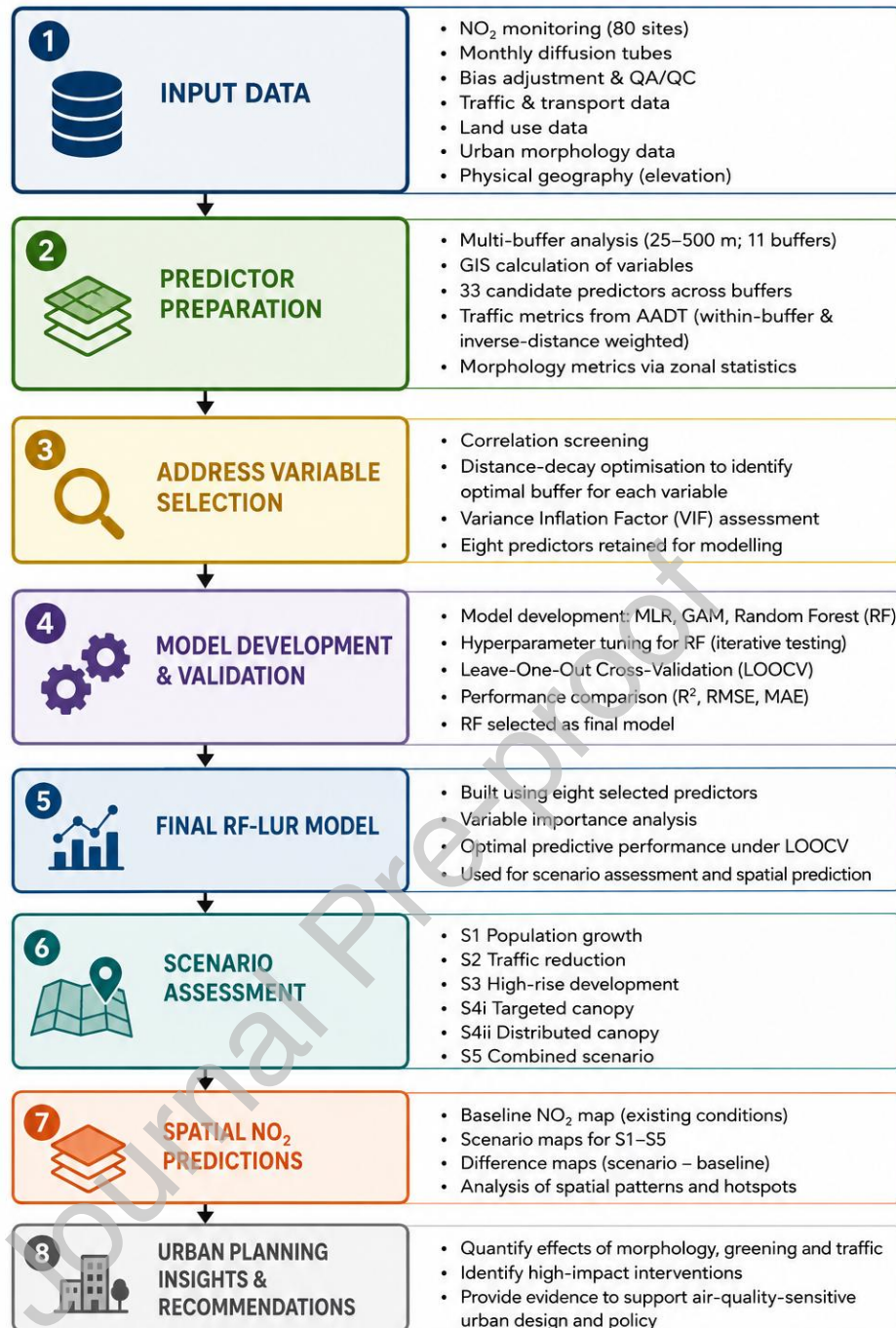


Figure 1: Scenario Assessment Process framework illustrating the workflow from data preparation and predictor selection through RF-LUR model development, scenario implementation and evaluation of NO₂ outcomes for urban-planning interventions in the City of London.

2.0 Study area

The City of London (Figure 2) is the historic and financial core of Greater London, covering 1.12 square miles ($\approx 3 \text{ km}^2$). Despite its small area, it concentrates major economic activity (e.g., the London Stock Exchange and Bank of England) and exhibits a distinctive high-density urban fabric comprising historic street patterns alongside modern high-rise development (average building height $\approx 25 \text{ m}$). The resident population is small (just over 8,000), but the daytime population increases to $\sim 600,000$ due to commuting, intensifying transport demand and associated emissions (Department for Transport, 2019). The coexistence of narrow streets and tall buildings creates frequent “street-canyon” conditions that restrict ventilation and can promote local accumulation of traffic-related pollutants such as NO_2 . These characteristics make the City of London an appropriate case study for examining how built form and traffic interact to shape fine-scale NO_2 patterns and how urban design interventions may mitigate exposure.

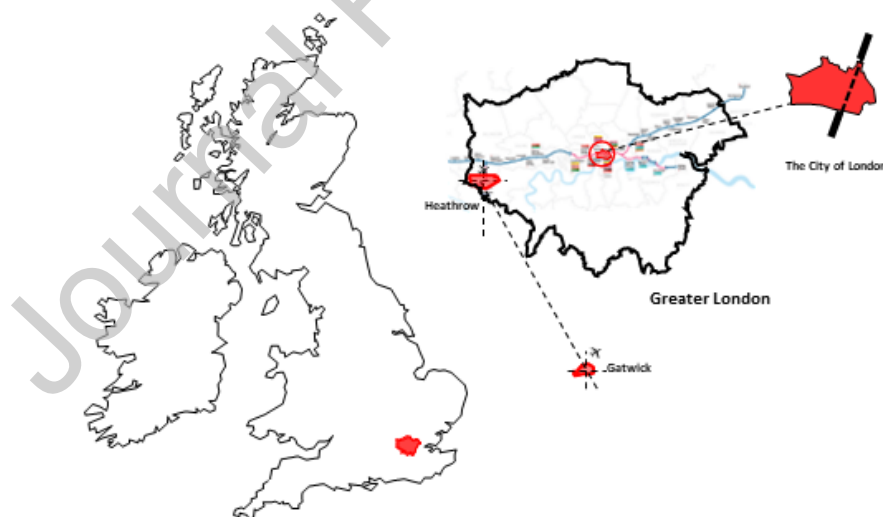


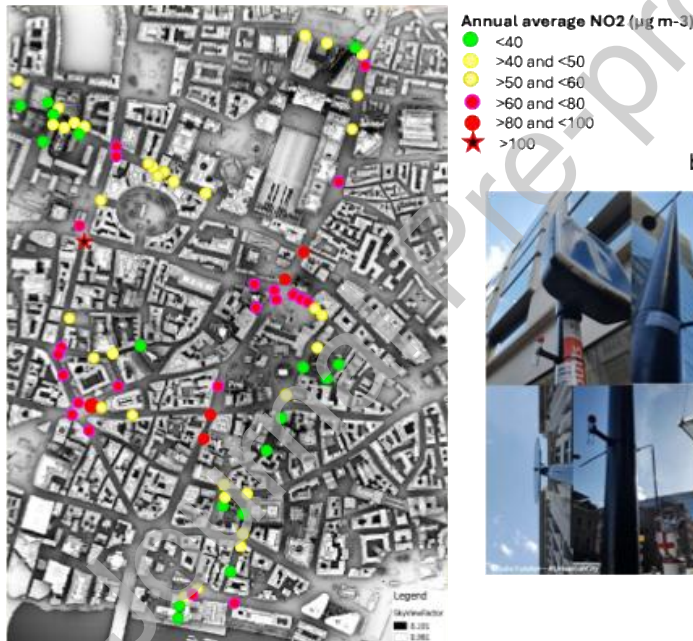
Figure 2: The City of London, the historic city core central to Greater London, is a commercial area that occupies an area just over 3 km^2 or less than 0.2% of the Greater London area. It is distinguished by its occupation patterns that are strongly linked to commercial functions; during the daytime, its population swells to nearly 600,000 but there are just over 8,000 residents.

2.1 NO₂ measurements

We used NO₂ measurements from an 80-site passive monitoring campaign in 2019 using lamppost-mounted diffusion tubes (Figure 3). Sites were selected to capture a range of built-form and street configurations expected to be relevant to ventilation and pollutant dispersion (e.g., canyon-like streets, open intersections, and areas adjacent to major roads), with the aim of sampling intra-urban variability within a compact high-rise district. Diffusion tubes were exposed continuously throughout 2019 and replaced at approximately monthly intervals in accordance with standard UK monitoring protocols. Monthly concentrations were combined to derive annual-average NO₂ concentrations and subsequently bias-adjusted using a local adjustment factor derived from co-location studies and CDF analysis (Futcher 2021; Futcher & Mills 2020). Following quality assurance procedures, 70 monitoring locations with complete and reliable annual-average NO₂ data were retained for model development, while 10 locations were excluded due to missing or unreliable measurements. The retained monitoring dataset exhibited substantial spatial variability, with annual-average NO₂ concentrations ranging from 31 to 78 $\mu\text{g m}^{-3}$ (mean = 51.8 $\mu\text{g m}^{-3}$, median = 49 $\mu\text{g m}^{-3}$, $n = 70$). Individual monitoring results are provided in Appendix Table A1. The resulting bias-adjusted annual-average NO₂ concentrations ($\mu\text{g m}^{-3}$) served as the dependent variable for model development. To assess the spatial representativeness of the retained monitoring network in terms of the predictor conditions represented across the modelling domain, distributions of three key predictors were compared between the 70 monitoring locations and the full prediction domain (1,258 locations): PAI at 500 m, heavy-vehicle AADT at 50 m, and building-height variability at 100 m. The ranges observed at the monitoring locations encompassed 83.1%, 87.3%, and 76.6% of prediction locations for PAI, heavy-vehicle AADT, and building-height

variability, respectively. The comparison indicates that the monitoring network captured a substantial proportion, but not the full range, of these predictor conditions across the study domain. Locations characterised by lower PAI and building-height variability, and by very high heavy-vehicle traffic intensity, were less well represented in the monitoring dataset. The use of annual-average concentrations reduces the influence of short-term seasonal variability, including differences between leaf-on and leaf-off periods, and is consistent with the study objective of modelling long-term spatial patterns of NO₂ across the City of London.

a) NO₂ monitoring sites and concentration on sky view factor map



b) Examples of monitoring sites



Figure 3. NO₂ diffusion tube locations in the City of London covering a range of urban sites; a) NO₂ monitoring sites and concentration overlaying a sky view factor (SVF) map; b) Examples of monitoring sites

The monitoring locations were selected to provide broad spatial coverage across the City of London and to capture a wide range of traffic, land-use and urban-morphology conditions. The network included sites adjacent to major roads, street canyons, urban squares and lower-traffic environments, thereby ensuring that the dataset represented

the diversity of exposure conditions present within the study area and provided sufficient variability in observed NO₂ concentrations for model development.

2.2 Land-use regression modelling

Here, the LUR air pollution modelling method is used to investigate the quantitative association between intraurban air pollution concentration and urban planning/design factors. The air pollutant concentrations depend on the emission intensity of pollution sources, which determines how fast the pollutants can be produced and released into the ambient air, and the dynamic potential of pollution dispersion, which describes how fast the concentrated air pollutants can disperse.

The dynamic potential of pollution dispersion is analysed by using two categories of potential independent variables: 1) physical geography, and 2) urban/building morphology. Most of these variables were chosen based on previous LUR models for NO₂ (Molter and Lindley, 2021; Ma et al., 2020; Munir et al., 2020; Miskel et al., 2015). We also included predictor variables representing the distinct features of City of London's urban form that would help explain the relationship between unique morphology and urban air pollution levels. A total of 33 predictor variables were selected for model development (Table 1). The study analysed the predictors using 11 different buffer sizes (i.e., 25m, 50m, 100m, 150m, 200m, 250m, 300m, 350m, 400m, 450m, and 500m), leading to a total of 363 candidate independent variables for LUR development.

City of London is a highly urbanised area where there is a low variety of traditional land-use parameters. Consequently, the land-use variables considered in this study focused on the principal non-built surface types present within the study area. Vegetation (green space areas excluding trees and tree canopy cover) and water coverage were selected

because, although they occupy a relatively small proportion of the City of London compared with the built environment and transport infrastructure, they can influence pollutant dispersion, deposition and local microclimatic conditions. Therefore, the traditional land-use parameters considered in this study are the vegetation (i.e., green space areas excluding trees and tree coverage) and water coverage. In this category, data from Open Street (OS) MasterMap (Edina Digimap) were used for green space and water areas, and the latest tree canopy cover vector map provided by the London City Hall. For each of those parameters, the area (m^2), calculated by ArcGIS Pro, was used as an input for the LUR model.

The morphological parameters considered in this study are the plan area index (PAI), the mean building height (H_{mean}), the maximum building height (H_{max}), the standard deviation of building heights (H_{std}), the aspect ratio (AR), the frontal area index (FAI), the roughness length (RL), and the sky view factor (SVF). Raster digital surface model (DSM) map of City of London derived from Edina Digimap was used as an input in QGIS 3.18.3 to calculate the morphology parameters. The QGIS plugin, Urban Multi-scale Environmental Predictor (UMEP) (Lindberg et al., 2018), was then used to calculate the above parameters and produce the related vector maps. Subsequently, ArcGIS Pro zonal statistics were used to calculate the parameters' values for each buffer zone, which were included in the LUR model. The predictor variables related to traffic and transport included the annual average daily traffic (AADT) (GLA 2019), the bus stops, the Railway-Underground stations, the major roads and the major intersections. All information was extracted from maps derived from OS MasterMap, and the calculations were conducted in ArcGIS Pro.

For each monitoring location, circular buffers ranging from 25 m to 500 m were generated. Area-based variables (e.g., green space, water and tree canopy) were calculated as the area contained within each buffer. Transport-related variables were represented either as counts within buffers (e.g., bus stops, rail/underground stations and major intersections) or as the distance to the nearest feature. Traffic-related variables were derived from AADT information at multiple spatial scales, together with inverse-distance weighted traffic metrics. For the latter, the traffic value associated with the relevant road feature was weighted by the reciprocal of the distance between the receptor and the road feature ($V \times 1/d$), where d is the distance in metres. This represents a reciprocal distance-decay function; no additional fitted distance-decay parameter was applied.

Finally, elevation was the only geography-related parameter used in the LUR model. First, a DSM and a DTM raster map were derived from Edina Digimap, and then a DEM map was created for the City of London using the raster calculation tool in ArcGIS Pro.

Prior to model development, candidate predictor variables were screened using the A Distance Decay REgression Selection Strategy (ADDRESS; Su et al., 2009). For each candidate predictor, correlations with measured NO_2 concentrations were evaluated across the range of buffer distances, and the predictor with the strongest association at its optimal buffer distance was selected first. From the second step onwards, candidate predictors were evaluated against the residuals from the preceding regression model, and the variable with the strongest residual correlation at its optimal buffer distance was added sequentially. Following Su et al. (2009), this forward-selection procedure continued while the newly added predictor contributed significantly to model performance and was

terminated when the contribution of an additional predictor was not statistically significant ($p > 0.10$). The final predictor set retained for model development comprised tree canopy area (250 m), heavy-vehicle AADT (50 m), distance to rail stations, Plan Area Index (PAI_500 m), building-height variability (100 m), inverse-distance weighted heavy-vehicle AADT, elevation and water area (100 m). Collectively, these variables represent traffic activity, transport accessibility, urban morphology, land cover and physical geography. Elevation was retained because it met the data-driven ADDRESS selection criteria rather than because a strong direct topographic effect on NO_2 was assumed. Given the relatively limited topographic variation within the City of London, elevation may capture modest differences in local terrain and dispersion conditions, but may also act as a spatial proxy for unmeasured or incompletely represented characteristics that co-vary with elevation. Its inclusion should therefore be interpreted primarily in terms of its predictive contribution within the empirical model rather than as evidence of a direct causal effect of elevation on NO_2 concentrations

Table 1: Candidate predictor variables considered for the initial LUR model

Domain	Predictor variable	Metric / unit	Spatial derivation (GIS)	Expected direction of effect on NO_2
Traffic & transport	AADT (total)	vehicles day ⁻¹	Buffer summary (multiple radii)	Positive
	AADT (heavy vehicles)	vehicles day ⁻¹	Buffer summary	Positive
	AADT (HDV)	vehicles day ⁻¹	Buffer summary	Positive
	AADT (LDV)	vehicles day ⁻¹	Buffer summary	Positive
	AADT (car & taxi)	vehicles day ⁻¹	Buffer summary	Positive
	AADT (bus)	vehicles day ⁻¹	Buffer summary	Positive
	Traffic intensity (inverse-distance)	vehicles	Distance-based / inverse-distance	Positive

	weighted AADT)*	$\text{day}^{-1} \cdot \text{m}^{-1}$	weighting to nearest major road segment	
	Major roads	m (road length) and/or m (distance)	Buffer road length; distance to nearest major road	Buffer: Positive; Distance: Negative
	Bus stops	count (within buffer) / m (distance)	Buffer count; distance to nearest stop	Buffer: Positive; Distance: Negative
	Major intersections	count / m	Buffer count; distance to nearest intersection	Buffer: Positive; Distance: Negative
	Rail/Underground stations	count / m	Buffer count; distance to nearest station	Buffer: Positive; Distance: Negative (often proxying centrality/traffic)
Land-use / surface cover	Green space	m^2 (area) / m (distance)	Buffer area; distance to nearest polygon	Typically Negative (deposition/less traffic), but context-dependent
	Water	m^2 / m	Buffer area; distance to nearest waterbody	Often Negative / mixed (proxy for openness)
	Tree canopy	m^2 / m	Buffer area; distance to nearest canopy	Mixed (deposition $\downarrow$ vs reduced ventilation $\uparrow$)
Urban / building morphology	Frontal Area Index (FAI)	dimensionless	Buffer summary	Positive (reduced ventilation)
	Plan Area Index (PAI)	dimensionless	Buffer summary	Positive (higher building coverage)
	Mean building height H_{mean}	m	Buffer summary	Mixed (depends on canyon/roughness regime)
	Maximum building height H_{max}	m	Buffer summary	Mixed
	Height variability H_{std}	m	Buffer summary	Often Negative (enhanced mixing), context-dependent
	Roughness length z_0	m	Buffer summary	Mixed (roughness can increase mixing aloft but also indicates dense fabric)
	Sky View Factor (SVF)	0–1 (dimensionless)	Buffer summary	Negative (more open sky $\rightarrow$ better ventilation)
Physical geography	Elevation	m a.s.l.	Point value or buffer mean	Mixed / potentially reflects local terrain/dispersion conditions or

				spatially correlated characteristics
--	--	--	--	--------------------------------------

2.3 Machine-learning approaches

In our study, we evaluated several statistical machine-learning approaches to enhance the predictive power of LUR models, focusing on Random Forest (RF), Generalized Additive Models (GAM), and Multiple Linear Regression (MLR). This comparison aimed to identify the model that offers superior performance in capturing the complex relationship between urban environmental factors and air pollution levels. By analysing a diverse range of urban morphology and pollution data, we assessed each model's accuracy, interpretability, and computational efficiency. Model fit was assessed using adjusted R^2 calculated from the fitted models, while predictive performance was evaluated using Leave-One-Out Cross-Validation (LOOCV) and summarised using the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The RF model was implemented in R using the RandomForest package. Model hyperparameters were selected through iterative testing of alternative parameter combinations and comparison of predictive performance. The final model configuration was selected on the basis of minimising prediction error while maintaining model stability and avoiding overfitting.

Following quality assurance procedures, 70 monitoring locations were retained for model development and validation. Model performance was evaluated using LOOCV, whereby each monitoring location was excluded in turn and predicted using a model trained on the remaining 69 locations.

The modelling framework was evaluated using adjusted R^2 , RMSE and MAE derived from LOOCV. The modelling framework was evaluated using adjusted R^2 , RMSE and MAE

derived from LOOCV. In addition, a global Moran's I test was applied to the Random Forest residuals to assess residual spatial autocorrelation. The test indicated no statistically significant spatial autocorrelation (Moran's I = -0.019, $p = 0.525$), suggesting that the final model adequately captured the dominant spatial structure in the monitoring data.

2.4 Urban design intervention scenarios

To translate model results into planning-relevant insights, we simulated five intervention scenarios affecting (i) development intensity, (ii) traffic activity, (iii) vertical form, and (iv) vegetation. Each scenario modified the relevant predictors in GIS, recalculated the affected LUR inputs at the same buffer distances, and predicted new NO₂ surfaces using the best-performing model. Scenario outputs were summarised as absolute changes (ΔNO_2 , $\mu\text{g m}^{-3}$) relative to the baseline.

The scenario analysis was designed as a set of exploratory sensitivity tests examining how the fitted 2019 RF-LUR model responds to hypothetical changes in selected urban-form, vegetation and traffic-related predictors relative to the 2019 baseline. The scenarios were not intended to provide forecasts of absolute future NO₂ concentrations. Unless explicitly modified within a scenario, predictor relationships and broader background air-quality conditions were retained as represented in the 2019 calibration dataset. Accordingly, the scenario outputs should be interpreted as relative modelled responses to the specified predictor changes under the fitted 2019 RF-LUR framework, rather than projections of future air quality under expected 2036 fleet, emissions or background conditions.

Scenario outcomes were expressed as point-estimate differences in predicted NO₂ relative to the baseline model, formal prediction intervals were not calculated, and the magnitude

and spatial coherence of the predicted changes were therefore used to support their qualitative interpretation.

2.4.1 Scenario 1: Population growth (horizontal densification)

To define an exploratory horizontal-densification scenario, projected workday population growth to 2036 (City Plan 2036) was used to estimate a corresponding change in building footprint area (City of London, 2021). Utilising UK Census 2011 data, we projected workday population density for 2018 as our baseline and predicted changes up to 2036. The composition of the City of London workday population was broken down as follows: 68% from the Greater London Area (GLA), 30% from outside GLA, and 2% local residents. We leveraged official borough-level resident projections from the Office for National Statistics (ONS) along with origin-destination data to estimate future workday population increases, assuming a constant influx from areas outside GLA due to lack of specific data. To quantify the impact of this population growth on the urban landscape, we applied a non-linear equation from Ward and Grimmond (2017) that links population to the building area:

$$y=10.45\ln(x)-24.49 \quad (1)$$

where x is population and y is building area.

The equation was adopted from Ward and Grimmond (2017), who developed the relationship using observed population density and building-cover data across Greater London boroughs. The relationship was selected because it was developed specifically for Greater London and was used by those authors to construct urban-development scenarios in which projected population growth was translated into changes in building cover, providing a locally relevant empirical basis for representing a plausible direction of

densification associated with population growth. However, the original relationship was derived at borough scale, whereas the present study applied it at Output Area level to spatially distribute the scenario within the City of London. This finer-scale application assumes that the borough-scale population–building-cover relationship remains informative at local scales and was not independently validated at Output Area level. The relationship was therefore used as a scenario-development tool rather than as a deterministic predictor of future building footprints and Scenario 1 should be interpreted as an exploratory representation of horizontal densification.

The relationship was calibrated against observed building cover (linear calibration; $R^2 = 0.69$; Figure 4). Percentage increases in building area were computed at Output Area level and allocated to building polygons (≈ 25 m resolution). Because the baseline model identified building density (PAI at 500 m) as influential, PAI was recalculated after allocation. This scenario assumes additional floor-area demand is met by expanding building footprints into available non-tree green space (vertical intensification is treated separately in Scenario 3).

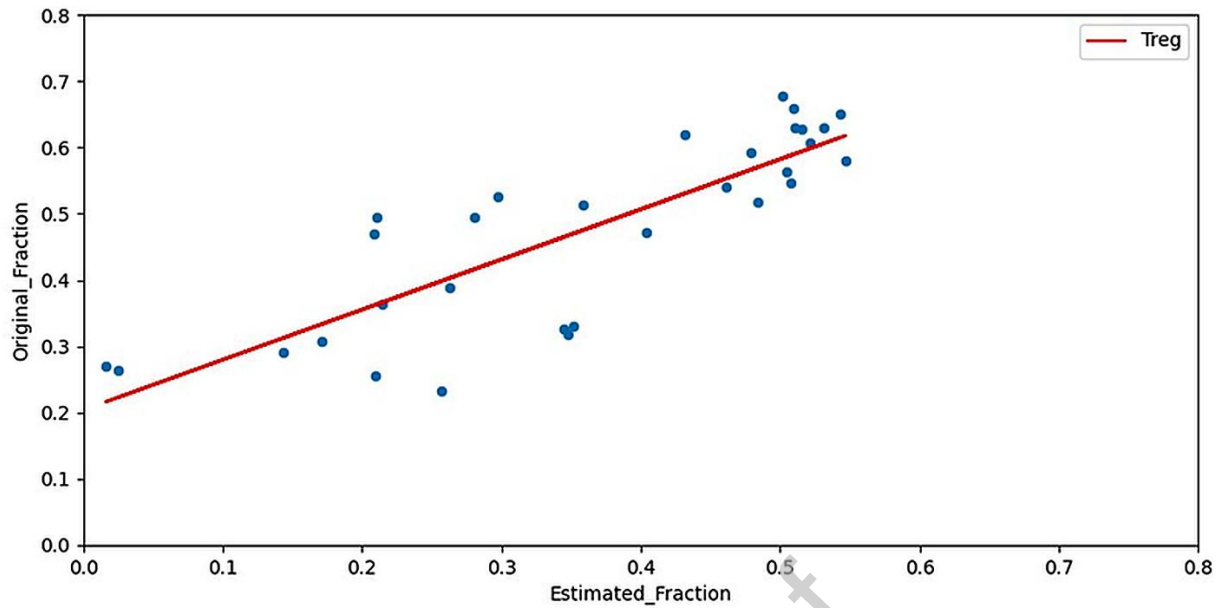


Figure 4: Relationship between the current building cover and the modelled building cover for the baseline year.

2.4.2 Scenario 2 (Traffic control)

This scenario represents a major reduction in heavy-vehicle activity, motivated by long-term traffic management trends and policy direction. We modelled a 50% reduction in heavy-vehicle AADT (HGVs, LGVs, and buses) across the network and updated the traffic predictors (AADT heavy at 50 m and the inverse-distance weighted term) before re-predicting NO_2 . The reduction was applied uniformly across all road segments within the study area by reducing the heavy-vehicle AADT values (HGVs, LGVs and buses) by 50% prior to recalculation of the traffic-related predictor variables. The scenario was intended to represent a strategic reduction in heavy-vehicle activity across the City of London rather than the implementation of a specific traffic management scheme. Consequently, traffic reassignment effects were not explicitly modelled, and it was assumed that the reduction represented an overall decrease in heavy-vehicle movements rather than

redistribution to alternative routes. This approach enables the potential air-quality benefits of reduced heavy-vehicle activity to be explored in a simplified and transparent manner.

2.4.3 Scenario 3: High-rise development (vertical change)

Scenario 3 evaluates the impact of planned tall buildings (>75 m AOD) on NO₂ through changes in local morphology. Based on the Local Plan Monitoring Report (2021), we incorporated 13 proposed tall buildings by updating building footprints and heights in the spatial dataset. Morphological metrics—particularly building height variability (H_std at 100 m)—were recalculated using UMEP and zonal statistics, then propagated through the LUR predictors to estimate NO₂ changes. Where redevelopment replaced existing buildings, footprint area was assumed to remain constant unless revised footprint information was available from the planning documentation. This assumption was adopted to isolate the influence of vertical development and associated changes in urban morphology from the effects of horizontal expansion in building coverage, which were considered separately under Scenario 1. Consequently, Scenario 3 should be interpreted as an assessment of the air-quality implications of increased building height and morphological complexity rather than a comprehensive redevelopment scenario.

In practice, redevelopment schemes may alter both building height and footprint area. The assumption of constant footprint area in Scenario 3 was adopted to isolate the effects of vertical morphological change and therefore represents a simplified redevelopment scenario.

2.4.4 Scenario 4 : Increased tree canopy (targeted vs distributed)

To reflect the Mayor of London's canopy ambitions, we modelled a 10% increase in tree canopy using two implementation logics:

- Scenario 4i (targeted): baseline NO₂ grid cells were ranked from highest to lowest predicted concentration, and additional canopy was allocated preferentially within available non-tree green space in the highest-ranked cells. Allocation proceeded down the NO₂ ranking until sufficient eligible land had been identified to accommodate the prescribed total 10% increase in tree-canopy cover across the study area. Thus, no fixed NO₂ concentration threshold was applied; the allocation was rank-based and prioritised locations with the highest baseline NO₂ concentrations.
- Scenario 4ii (distributed): 10% of the available non-tree green-space area within each eligible polygon was converted to tree canopy. The increase was therefore allocated proportionally according to the area of each eligible green-space polygon across the study area, irrespective of baseline NO₂ concentration, representing a more spatially distributed greening strategy.

For both sub-scenarios, the increase in canopy area was implemented within the existing land-cover dataset and the tree-canopy predictor variable within the optimised 250 m buffer was recalculated prior to re-predicting NO₂ concentrations. The scenarios were intended to evaluate the relative effectiveness of targeted versus distributed canopy expansion and did not explicitly distinguish between individual planting mechanisms (e.g., street trees, park expansion or green-roof installations).

2.4.5 Scenario 5: Integrated strategy

Scenario 5 combines the most effective single interventions into an “air-quality-sensitive urban design” package: heavy-vehicle traffic reduction (Scenario 2), tall-building

development/morphology change (Scenario 3), and targeted canopy increase (Scenario 4i). The spatial allocation of each intervention was retained as defined in its corresponding individual scenario. Thus, the traffic reduction was applied to the same road locations, the tall-building intervention to the same development locations, and the targeted canopy increase retained the same spatial allocation as in Scenario 4i. Where interventions spatially coincided, the corresponding predictor modifications were retained simultaneously, with no additional spatial exclusion or compatibility rules imposed. Predictors were updated simultaneously and the combined NO₂ response was estimated relative to baseline. The combined response therefore represents the RF-LUR model response to the simultaneous modification of these predictors rather than an arithmetic summation of the NO₂ changes predicted under the individual scenarios.

3 Results

3.1 Comparison of different LUR modelling methods

We compared three approaches for predicting annual-average NO₂ at the 80 monitoring locations: MLR, GAM, and RF. For all approaches, predictor buffer sizes were optimised using the ADDRESS procedure (Section 2.2), rather than adopting a broad “include-all-buffers” strategy (e.g., ESCAPE; Beelen et al., 2013). This ensured that each predictor entered the model at the spatial scale most strongly associated with observed NO₂.

Across all performance metrics, RF provided the strongest predictive skill (Table 2). The RF model achieved an adjusted R^2 of 0.88 when fitted to the full dataset, outperforming GAM (0.80) and MLR (0.72). Predictive performance was assessed using LOOCV for which RF produced the lowest prediction errors (RMSE = 7.6 $\mu\text{g m}^{-3}$; MAE = 6.25 $\mu\text{g m}^{-3}$). GAM showed substantially higher error (RMSE = 10; MAE = 7.9), suggesting that although

smooth non-linear terms improved performance relative to MLR, they did not capture the full complexity of relationships between NO₂, traffic intensity, and morphology in this highly heterogeneous urban environment. MLR performed least well in explained variance, consistent with limitations of strictly linear, additive forms for modelling street-canyon-dominated pollution patterns. On the basis of these results, RF was used for subsequent spatial prediction and scenario testing. To improve model interpretability, Random Forest variable importance metrics were examined (Appendix Figure A1). PAI at 500 m had the highest variable-importance score, followed by traffic-related variables and distance to rail stations. Height variability and elevation also contributed meaningfully to model performance, while tree canopy showed a more moderate influence and water area exhibited comparatively low importance. However, variable importance reflects the predictive contribution of each variable within the fitted RF model and should not be interpreted as evidence of causal influence. In particular, the high importance of PAI may reflect both information associated with urban form and information shared with spatially co-varying characteristics, such as traffic intensity, street configuration and other land-use features. PAI may therefore partly act as a proxy for these correlated characteristics rather than representing a direct physical effect of building density on pollutant dispersion. The variable-importance results are consequently interpreted as indicators of predictive importance within the RF-LUR model rather than evidence of causal or mechanistic effects.

Table 2: Performance comparison of modelling methods.

Model	Adjusted R^2 (full model fit)	LOOCV RMSE	LOOCV MAE
-------	---------------------------------	------------	-----------

RF	0.88	7.6	6.25
GAM	0.80	10	7.9
MLR	0.72	8.3	6.4

3.2 Impact of air quality sensitive urban design interventions

Scenario results are reported as point-estimate changes in predicted NO₂ relative to the baseline model and should be interpreted primarily in terms of the magnitude and spatial pattern of the modelled response rather than as statistically bounded estimates of concentration change.

3.2.1 Scenario 1: Population growth

Figure 5 maps baseline NO₂ concentrations across the City of London, while Figure 6 shows Scenario 1 changes relative to baseline. Under the population growth scenario, the assumed increase in built fraction was associated with higher predicted NO₂ concentrations in areas where urban density increased most.

Overall, the modelled response indicated higher NO₂ concentrations across much of the study area relative to baseline. The strongest predicted increases occurred in the north and north-east, commonly ranging from less than 1 $\mu\text{g m}^{-3}$ to approximately 5.5 $\mu\text{g m}^{-3}$. These areas include streets characterised by constrained geometry, where the modelled response to additional densification is consistent with reduced ventilation and pollutant dilution. Only a small number of locations showed slight predicted decreases (up to approximately 0.7 $\mu\text{g m}^{-3}$). Collectively, these results indicate an association between the assumed increase in built fraction and higher predicted NO₂ under the conditions represented in Scenario 1.

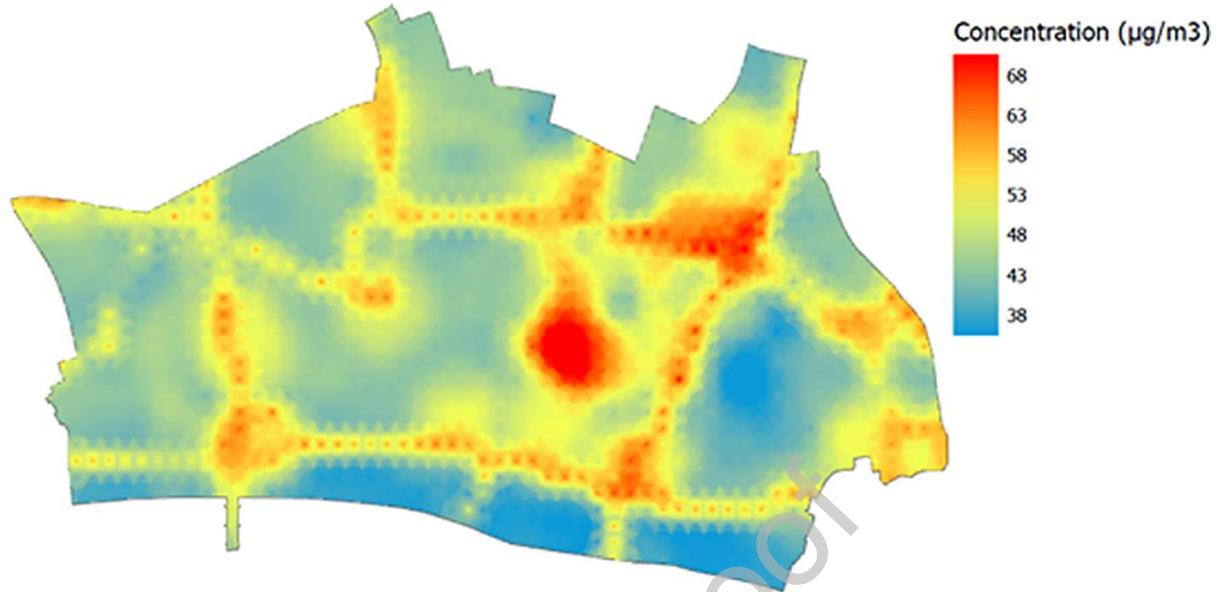


Figure 5: Spatial distribution of Baseline NO₂ concentration in the City of London.

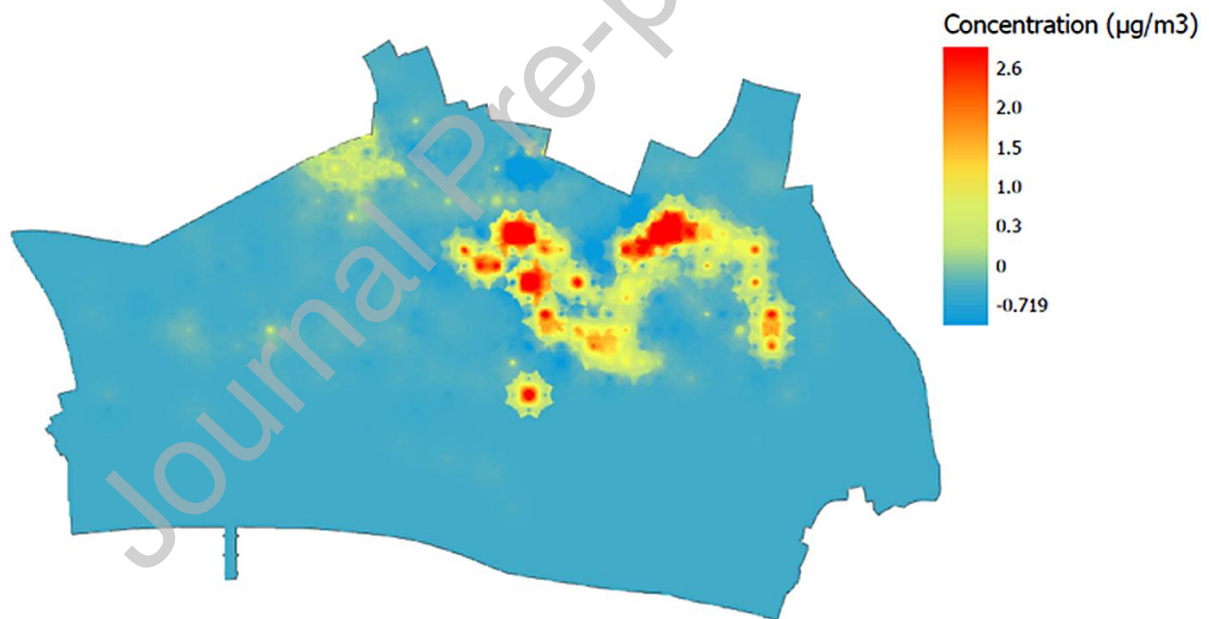


Figure 6: Change in predicted annual mean NO₂ concentration ($\mu\text{g m}^{-3}$) across the City of London under Scenario 1 relative to the baseline scenario.

3.2.2 Scenario 2: traffic control

Scenario 2 represents a substantial reduction in heavy-vehicle activity, implemented by halving relevant AADT inputs, affecting two of the eight key predictors in the final model.

The resulting NO₂ surface indicates that predicted decreases are concentrated along major traffic corridors and at key junctions.

Figure 7 shows that coherent predicted decreases along principal routes, with decreases typically in the 1–6 $\mu\text{g m}^{-3}$ range. The largest improvements occur along the northern arterial road and around the south-western major intersection, consistent with areas where baseline emissions intensity is highest. These results indicate that the assumed reduction in heavy-vehicle activity was associated with spatially targeted decreases in predicted NO₂, particularly where heavy vehicles contribute disproportionately to near-road concentrations.

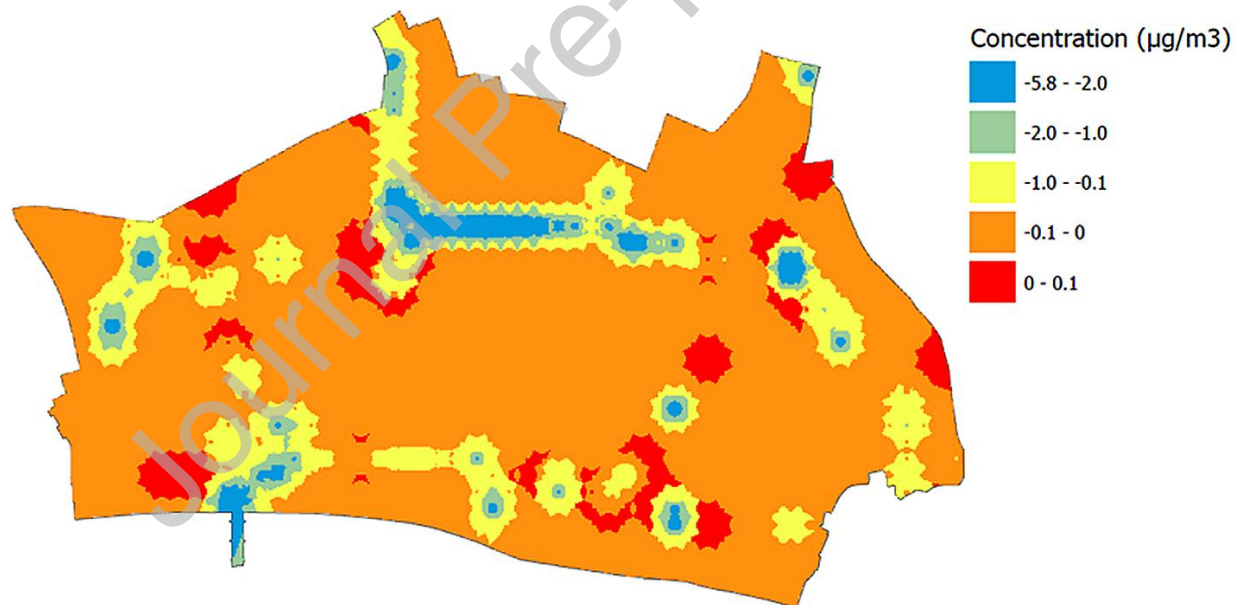


Figure 7: Concentration difference between predicted NO₂ across the City of London (Scenario 2 - Baseline)

3.2.3 Scenario 3: High-rise development (vertical change)

Scenario 3 evaluates the effect of adding 13 new tall buildings, implemented through updates to building height and footprint layers and subsequent recalculation of morphology predictors (notably height standard deviation at the 100 m buffer). Figure 8 maps the resulting changes.

The model predicts a generally lower NO₂ concentrations across impacted areas, with the most pronounced decreases in the eastern cluster. In the north-east and east, predicted decreases reach up to $\sim 7 \mu\text{g m}^{-3}$. These modelled responses are consistent with aerodynamic evidence that greater building-height variability may, under some configurations, enhance local mixing and dispersion in dense urban blocks. Small local predicted increases (generally 0–0.5 $\mu\text{g m}^{-3}$) occur in a limited number of locations. These may reflect the non-linear statistical relationships captured by the RF-LUR model and the sensitivity of predicted NO₂ to local urban characteristics. Overall, Scenario 3 indicates that increased vertical variability was associated with lower predicted NO₂ concentrations in parts of the City, particularly where baseline concentrations were highest. This pattern is consistent with findings that height heterogeneity can influence pollutant dispersion in dense districts (e.g., Huang et al., 2022), although the underlying aerodynamic processes are not explicitly simulated by the RF-LUR model.

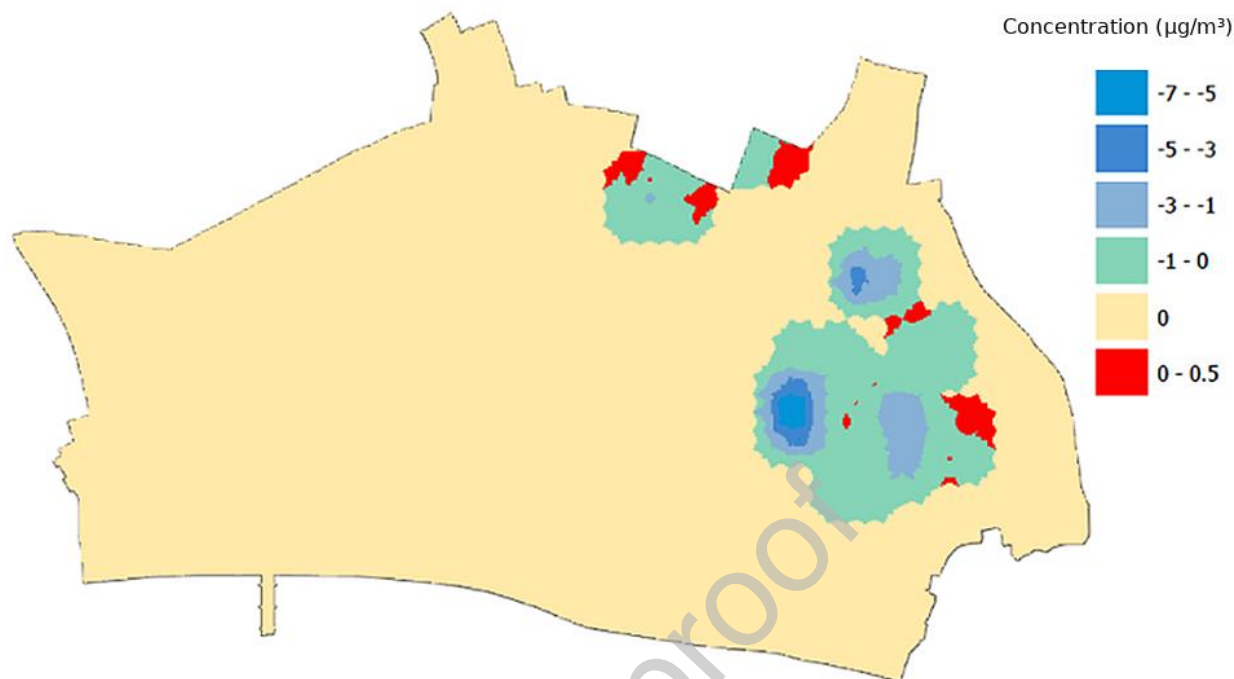


Figure 8: Concentration difference between predicted NO₂ across the City of London (Scenario 3 - Baseline).

3.2.4 Scenario 4: Tree canopy increase

Both tree-canopy scenarios were associated with modest but spatially variable changes in predicted NO₂, with the clearest predicted decreases occurring in pollution hotspots. Under Scenario 4i, predicted NO₂ decreases reached up to $\sim 4 \mu\text{g m}^{-3}$ in targeted areas (Figure 9). However, localised predicted increases also occurred, particularly in parts of the east and north-east, reaching up to $\sim 2 \mu\text{g m}^{-3}$. Scenario 4ii was associated with broadly similar magnitudes of predicted change but with a different spatial pattern (Figure 10): predicted decreases were more widely distributed, while small increases remained in some central/eastern locations. These mixed responses indicate that the modelled response to additional tree canopy was spatially non-uniform, with predicted changes

varying according to street geometry, existing morphology and the other predictors represented in the RF-LUR model.

Although Scenario 4ii increases canopy over a much larger area (19,284 m²) than Scenario 4i (6,670 m²), Scenario 4i was associated with more extensive predicted NO₂ decreases across the study area. Figure 11 indicates that the targeted strategy concentrates canopy increases in locations with high baseline NO₂ and greater model sensitivity to the canopy predictor. Within the modelled scenarios, this suggests that the spatial placement of additional canopy may be more influential than the total canopy area added in determining predicted NO₂ responses in dense urban cores.

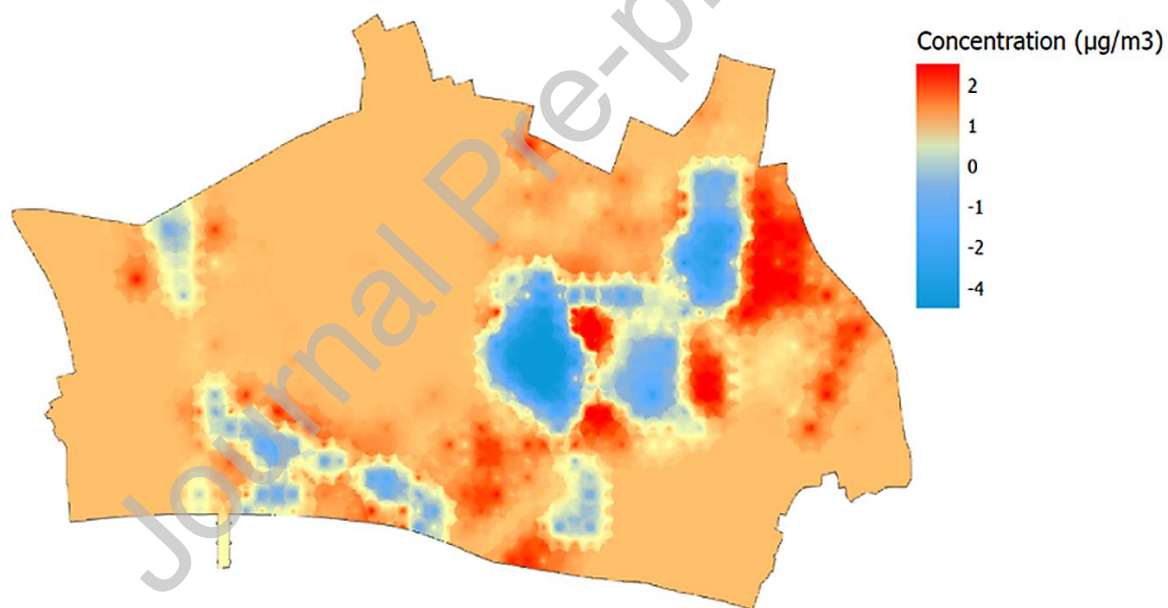


Figure 9: Concentration difference between predicted NO₂ across the City of London (Scenario 4i – Baseline).

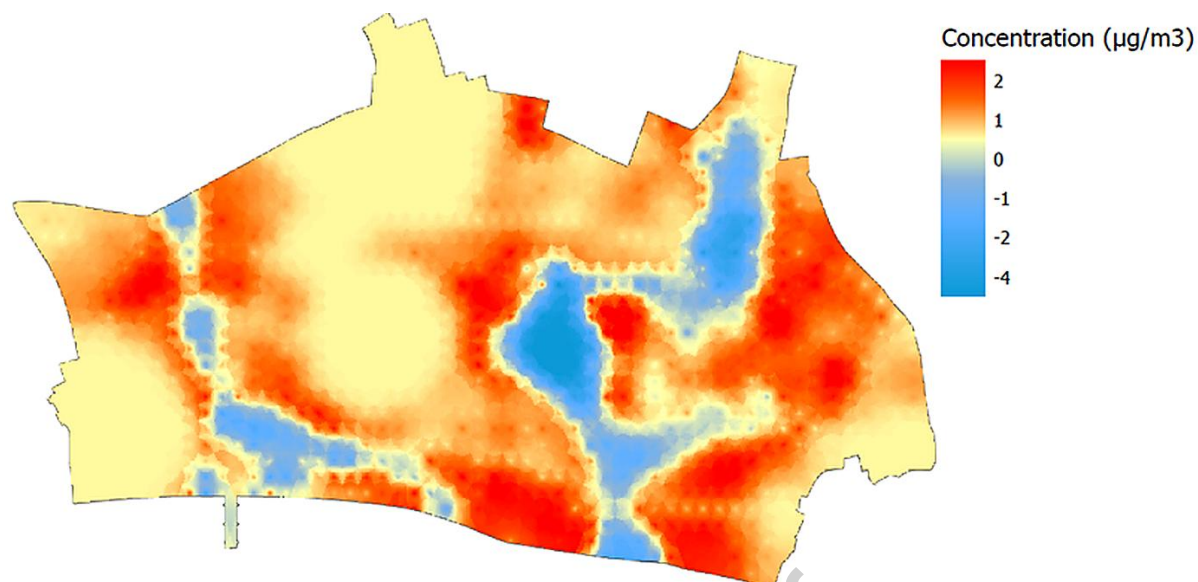


Figure 10: Concentration difference between predicted NO₂ across the City of London (Scenario 4ii – Baseline).

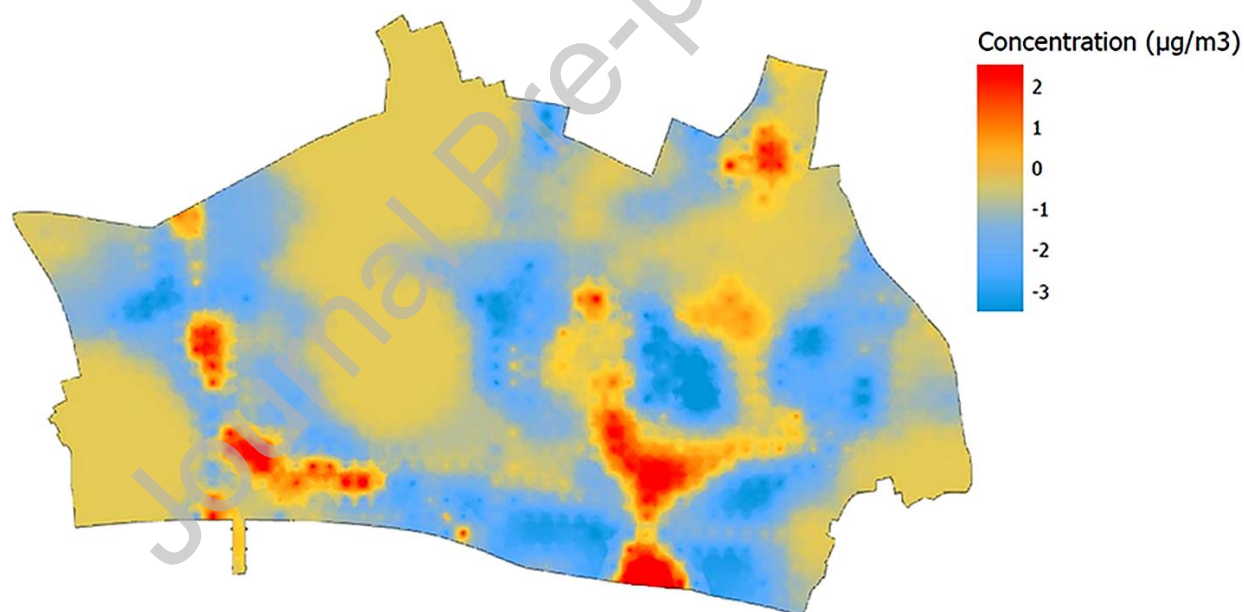


Figure 11: Concentration difference between predicted NO₂ across the City of London (Scenario 4i - Scenario 4ii).

3.2.5 Scenario 5: Integrated design

The combined package produces the largest overall improvements, particularly in the eastern and north-eastern parts of the City where baseline concentrations are elevated

(Figure 12). Maximum reductions reach $\sim 7 \mu\text{g m}^{-3}$, and most of the study area experiences decreases between 0 and $7 \mu\text{g m}^{-3}$. As with individual interventions, small local increases remain (generally $0\text{--}1.5 \mu\text{g m}^{-3}$) in a limited number of adjacent areas, suggesting that even integrated strategies can have uneven micro-scale outcomes due to highly local street and building configurations. Overall, Scenario 5 was associated with greater predicted NO_2 decreases than the individual interventions in several areas, particularly where traffic intensity and constrained dispersion coincide.

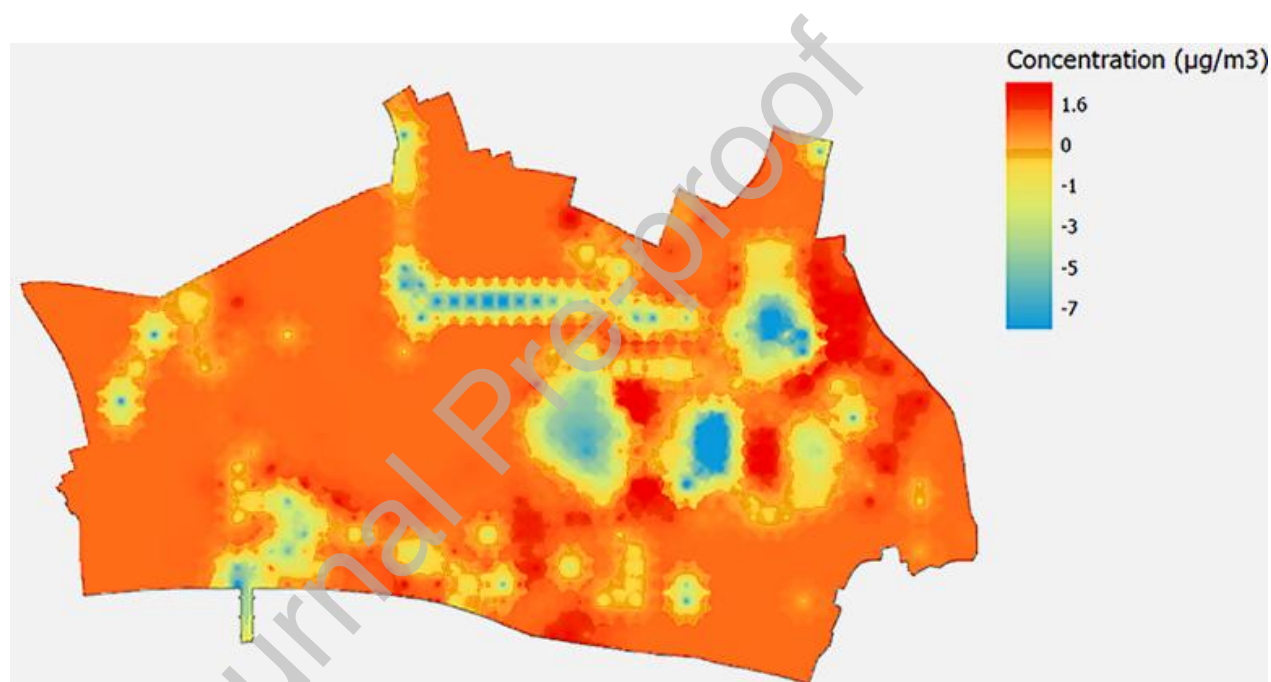


Figure 12: Concentration difference between predicted NO_2 across the City of London (Scenario 5 - Baseline)

4 Discussion

This study developed high-resolution intra-urban NO_2 land-use regression (LUR) models for the City of London using a machine-learning framework that explicitly incorporates urban morphology (e.g., building height metrics and density). Among the evaluated approaches, the Random Forest LUR (RF-LUR) achieved the best predictive performance.

The improvement is consistent with evidence that RF better captures non-linearities and higher-order interactions common in urban air pollution processes, and has previously shown lower errors in NO₂ prediction (e.g., Champendal et al., 2014; Brokamp et al., 2017; Sakshi et al., 2021). Importantly, our results indicate that including morphology predictors adds explanatory value in a dense city-centre context where dispersion constraints and canyon effects are key determinants of spatial NO₂ variability.

While the RF-LUR framework was subsequently applied to evaluate alternative planning scenarios, these analyses should be interpreted as exploratory assessments of how the statistical relationships captured by the model respond to changes in key predictor variables. The scenario results therefore provide insight into the potential direction and relative magnitude of air-quality responses rather than direct mechanistic evidence of how future developments would alter pollutant dispersion in practice.

4.1 Interventions and the role of urban morphology

4.1.1 Implications for urban planning

The scenario experiments indicated spatially heterogeneous modelled NO₂ responses that varied with on local street geometry, building form, and proximity to traffic corridors. The magnitude of the predicted changes we observe (up to $\sim 7 \mu\text{g m}^{-3}$ in some areas under morphology-focused or combined interventions) is comparable to reported impacts of urban design on dispersion and near-road concentrations (e.g., Metrangolo et al., 2024; Ling et al., 2025). These results suggest the value of integrating both emission-focused measures (e.g., traffic management) and form-focused measures (e.g., built massing, vertical heterogeneity and targeted greening) when considering air-quality planning in compact urban centres, rather than treating land-use and transport as independent levers.

Our findings also complement broader-scale LUR evidence. For example, the global spatiotemporal NO₂ model of Larkin et al. (2023) identifies major roads as robust predictors across regions, highlighting the persistent importance of traffic and road structure. At the city scale, Ge et al. (2022) similarly emphasize traffic and building density as key determinants of exposure (for UFP), supporting our conclusion that built environment and activity intensity jointly structure pollution fields in dense urban areas. Together, these studies suggest that high-resolution urban interventions should be assessed not only by total area modified (e.g., canopy added) but by whether they target locations where concentrations and sensitivities are highest.

4.1.2 Urban morphology: density versus vertical variation

A key insight from our scenarios is the contrasting role of different morphology dimensions. PAI also had the highest variable-importance score in the fitted Random Forest model, indicating that it provided substantial predictive information within the retained predictor set. As discussed in Section 3.1, however, this importance should not be interpreted as demonstrating a direct causal effect of building density on NO₂:

- Higher building density / built fraction generally increased NO₂, consistent with reduced ventilation and dispersion in compact blocks and deeper canyons.
- Greater height variability (e.g., higher standard deviation of building heights) was associated with decreases in NO₂ in several high-burden areas, suggesting enhanced mixing and dispersion under certain configurations.

From an urban-aerodynamics perspective, the association between greater building-height variability and lower predicted NO₂ in some locations is consistent with

mechanisms identified in previous numerical modelling studies. Non-uniform building heights can modify within- and above-canopy airflow and turbulent exchange and, under some configurations, enhance vertical ventilation and pollutant removal relative to more uniform building arrangements (Kim et al., 2023; Hang et al., 2012; Gu et al., 2011). These aerodynamic mechanisms provide a physically plausible interpretation of the modelled association observed here. However, because the RF-LUR framework does not explicitly resolve airflow, turbulence or pollutant transport, these mechanisms were not directly simulated or demonstrated in the present study. The Scenario 3 results should therefore be interpreted as a statistical model response consistent with established aerodynamic mechanisms rather than evidence that enhanced vertical mixing caused the predicted NO₂ reductions.

This “density worsens, heterogeneity may help” pattern aligns with previous studies reporting that building density and compact urban form are generally associated with higher pollutant concentrations, whereas greater variation in building heights can enhance ventilation and dispersion under certain conditions (Metrangolo et al., 2024; Wan et al., 2022). Our results extend these observations for the City of London by indicating that *where* height variability is introduced matters: the largest predicted decreases occurred in locations with higher baseline concentrations and constrained ventilation.

At the same time, the relationship is not monotonic. As Yang et al. (2019) note, tall buildings can enhance circulation in some wind regimes yet also create complex recirculation and stagnant pockets in others. The small local predicted increases observed in some locations are consistent with this non-linearity, although the RF-LUR model cannot resolve the micro-scale airflow processes that may contribute to such patterns.

4.1.3 Traffic control: localised modelled responses near major roads

Traffic interventions produced the clearest and most spatially targeted improvements, with reductions concentrated along primary routes and junctions. The estimated ~14% reduction near heavily trafficked roads following a 50% reduction in heavy-vehicle emissions is consistent with evidence that traffic management tends to yield modest citywide changes but meaningful local benefits when applied in congested corridors (Tang et al., 2018 & 2013; Bigazzi et al., 2017). This is also consistent with the broader literature that highlights larger and more reliable impacts when measures are targeted to high-emission links and exposure hotspots (e.g., Boogaard et al., 2022; Castro et al., 2017; Duque et al., 2016; Ferreira et al., 2015).

However, the predicted NO₂ decreases should not be interpreted as necessarily implying equivalent improvements across all air-quality metrics. In NO_x-rich urban environments, reductions in NO emissions from road traffic can reduce the titration of O₃ and may therefore be associated with local increases in O₃ concentrations (Macintyre et al., 2023). As O₃ chemistry was not represented in the RF-LUR framework, the traffic scenario should be interpreted specifically in terms of its predicted NO₂ response rather than as a comprehensive assessment of the overall air-quality effect of reducing heavy-vehicle activity.

The modelled spatial pattern suggests that traffic controls may warrant prioritisation in high-emission corridors and exposure hotspots (e.g., through freight restrictions on selected corridors, junction redesign or time-windowed access), particularly when considered alongside interventions addressing dispersion constraints in canyon-like streets.

4.1.4 Tree canopy: spatially variable modelled responses

Tree-canopy expansion was associated with predicted NO₂ decreases of ~2–4 µg m⁻³ in targeted hotspots, supporting the role of green infrastructure as a complementary mitigation strategy (e.g., Grylls & van Reeuwijk, 2022; Abhijith et al., 2017). However, we also observe small local increases in some areas, reinforcing that vegetation effects are not universally beneficial across the study area.

The wider atmospheric-chemistry implications of additional tree canopy should also be considered. Although vegetation can contribute to pollutant removal through deposition, some tree species emit biogenic volatile organic compounds (BVOCs), which can participate in photochemical O₃ formation under suitable atmospheric conditions. The present RF-LUR framework does not represent BVOC emissions or NO₂–O₃ chemistry, and the greening scenarios should therefore be interpreted specifically in terms of their predicted NO₂ response rather than as an assessment of the net multi-pollutant air-quality effect of increased canopy.

Mechanistically, this mixed behaviour is well documented: deposition and surface uptake can reduce pollutant loads, but dense canopies can also reduce wind speed and inhibit dispersion, particularly in street-canyons. In these environments, tree crowns may partially obstruct airflow and reduce the exchange of air between street level and the atmosphere above, limiting the dilution of traffic-related emissions and resulting in localised pollutant accumulation. Buccolieri et al. (2018) show that under perpendicular winds, trees can increase near-tree concentrations substantially due to aerodynamic blockage. Yli-Pelkonen et al. (2017) similarly report that canopy presence does not necessarily translate to lower NO₂ under the canopy relative to nearby open areas. These

studies provide plausible mechanistic context for the spatially variable modelled responses observed here and suggest that the air-quality implications of additional canopy may depend on siting, canyon geometry, prevailing winds, spacing and canopy porosity.

Consistent with this, our targeted canopy strategy produced broader NO₂ improvements than a larger but more diffuse canopy increase, suggesting that prioritising hotspots and sensitive corridors can deliver greater exposure benefits than maximising total canopy area. The results should also be interpreted in the context of canopy structure, which was not explicitly represented in the scenario modelling. Previous studies have shown that vegetation characteristics such as crown density, porosity and species type can substantially influence airflow and pollutant dispersion. In street-canyon environments, dense and continuous canopies may restrict ventilation and contribute to local pollutant accumulation, whereas more porous or strategically spaced tree planting can provide deposition benefits while maintaining air exchange. Consequently, the effectiveness of urban greening interventions depends not only on the amount of canopy added but also on the design and placement of the vegetation.

4.1.5 Combined interventions: integrated modelled responses and residual local trade-offs

The combined scenario (traffic control + canopy increase + morphology modification) was associated with the largest overall improvement (reductions up to $\sim 7 \mu\text{g m}^{-3}$). The integrated scenario was modelled by simultaneously applying the traffic-control, canopy-expansion and morphology-modification interventions and re-evaluating the LUR predictors. Consequently, the reported changes do not represent a simple arithmetic sum of the individual scenario outcomes but instead reflect the RF-LUR model response to

concurrent changes in emissions-related, land-cover and morphology predictors. The present framework was not designed to formally distinguish additive, synergistic or antagonistic interactions between measures. The larger predicted decreases under the combined scenario are consistent with the potential value of integrating measures acting on different urban characteristics, although the underlying physical interactions are not explicitly simulated by the RF-LUR model. This supports the idea that measures act through different pathways—emission reduction, deposition/near-road shielding, and dispersion modification—so combining them can yield larger net gains than any single measure (cf. Shukla et al., 2022). It also aligns with work highlighting the value of integrating diverse datasets and drivers for exposure assessment (Dimakopoulou et al., 2022).

Nevertheless, small local increases persisted even under the combined package, illustrating a recurring theme: in compact centres, interventions can shift pollution patterns as well as reduce totals. This reinforces the need for site-specific design and evaluation, particularly around canyon streets and complex junctions. While combined intervention assessments are increasingly common (e.g., Nieuwenhuijsen et al., 2016; Khomenko et al., 2021; Tsoulou et al., 2021; Rawat et al., 2022), the explicit inclusion of urban morphology alongside emission-control and greening measures remains relatively limited; our work contributes evidence that vertical form and density metrics can materially change predicted outcomes.

4.2 Limitations and future research

This study, while pioneering in its approach to modelling NO₂ variability in urban environments using RF-LUR, is subject to several limitations that offer avenues for future

investigation. First, the scenario analysis retains the broader air-quality and emissions context represented by the 2019 calibration dataset, except where individual predictors were explicitly modified. It therefore does not account for expected future changes in vehicle emission factors, fleet composition, regulatory policy, atmospheric chemistry or regional background NO₂ concentrations. These factors are likely to change over the period to 2036 and could alter both baseline concentrations and the magnitude of the modelled responses to the interventions considered. The scenarios should therefore be interpreted as exploratory sensitivity tests of selected urban-form, vegetation and traffic-related predictor changes relative to the 2019 baseline, rather than forecasts of absolute future air quality. Future work should integrate projected fleet composition, emission factors and background concentrations to evaluate the interventions under alternative future baseline conditions.

A further limitation relates to model validation and potential spatial dependence. Model performance was evaluated using LOOCV, supplemented by a global Moran's I test of model residuals. The analysis indicated no statistically significant spatial autocorrelation in the Random Forest residuals (Moran's I = -0.019, p = 0.525), suggesting that substantial residual global spatial structure was not detected in the monitoring dataset. However, no spatially blocked cross-validation or independent external validation was undertaken, and the model's predictive performance under spatial or scenario extrapolation was therefore not directly evaluated. This is particularly relevant to the scenario analysis, in which predictor values were modified from those represented in the 2019 calibration dataset. Consequently, the reported LOOCV performance should be interpreted as an assessment of predictive performance within the available monitoring dataset rather than evidence that equivalent predictive accuracy would be maintained under scenario conditions. A

quantitative comparison of the monitoring network with the full prediction domain further showed that 83.1%, 87.3% and 76.6% of prediction locations fell within the ranges represented by the monitoring sites for PAI at 500 m, heavy-vehicle AADT at 50 m and building-height variability at 100 m, respectively. Thus, although the monitoring network covered a substantial proportion of the predictor conditions across the study domain, some locations—particularly those characterised by lower PAI or building-height variability and very high heavy-vehicle traffic intensity—were outside the ranges represented by the monitoring data. Predictions for these locations therefore involve greater extrapolation from the conditions represented in the calibration dataset and should be interpreted with additional caution. Future work could therefore incorporate local spatial diagnostics (e.g., Local Moran's I or LISA) and spatially explicit validation approaches to further assess model robustness.

Furthermore, focusing exclusively on the City of London may limit the generalizability of our findings to other urban settings with distinct designs and policy frameworks. The inherent "black box" nature of machine-learning models like RF-LUR also poses challenges in fully understanding the mechanistic relationships between urban morphology and NO₂ concentrations, suggesting a need for more interpretable modelling approaches.

A further limitation relates to potential feature-selection bias introduced by the ADDRESS procedure. Candidate predictors were pre-screened partly on the basis of their bivariate relationships with NO₂ before development of the Random Forest model. While this approach reduced the number of candidate variables and potential multicollinearity, it may have excluded variables with relatively weak individual associations that could nevertheless become influential through non-linear interactions with other predictors. For

example, urban-morphology variables such as SVF and FAI may contain predictive information through interactions that was not evaluated once these variables were excluded during pre-selection. Consequently, the Random Forest variable-importance analysis presented in Appendix Figure A1 reflects the relative importance of the eight predictors retained by ADDRESS and should not be interpreted as demonstrating that excluded variables have no predictive relevance. Future research could evaluate the wider candidate predictor set using interaction-aware interpretability approaches, such as SHAP, to investigate these potential effects.

In addition, although the inclusion of morphology indicators improves representation of urban form, the RF-LUR framework does not explicitly resolve airflow dynamics, turbulence or street-canyon circulation in the manner of CFD, LES or detailed dispersion models. Consequently, the model should be interpreted as capturing statistical associations between urban characteristics and NO_2 concentrations rather than directly simulating the underlying physical processes governing pollutant transport and dispersion. Future studies could combine morphology-sensitive LUR models with CFD or LES simulations to investigate the physical mechanisms underlying the statistical relationships identified here.

The study focused specifically on NO_2 and did not explicitly model atmospheric chemistry or multi-pollutant responses. As discussed in Sections 4.1.3 and 4.1.4, reductions in traffic-related NO_x may affect O_3 concentrations through changes in NO titration, while additional vegetation may influence O_3 formation through biogenic volatile organic compound (BVOC) emissions. Consequently, predicted reductions in NO_2 should not be assumed to represent equivalent improvements in overall air quality. Future work should

therefore adopt a multi-pollutant perspective when evaluating urban-planning and transport interventions. In addition, the findings are derived from the highly urbanised City of London and should be extrapolated to other urban contexts with caution. Lastly, the potential impact of climate change on urban air quality and intervention outcomes cannot be overlooked. Future changes in temperature, atmospheric stability and wind patterns may alter pollutant dispersion and ventilation within urban canopies. Such changes could influence the effectiveness of morphology-based interventions by either enhancing or reducing the ventilation and pollutant-purging effects associated with increased height variability and modified urban form. Similarly, the performance of green infrastructure may be affected by climate-driven changes in vegetation characteristics and local microclimates. Future studies should therefore seek to integrate climate projections with urban-morphology and air-quality modelling frameworks.

An additional source of uncertainty relates to the population-growth scenario. The relationship between population density and building cover adopted from Ward and Grimmond (2017) explained approximately 69% of the variance in observed building cover and was originally developed using Greater London borough-scale data. Its application at Output Area level assumes that a similar population–building-cover relationship remains applicable at a substantially finer spatial resolution, an assumption that was not independently validated in this study. No sensitivity analysis was undertaken to quantify the uncertainty associated with alternative building-growth assumptions. Consequently, uncertainty in the magnitude and spatial allocation of the estimated building-footprint changes propagates into the Scenario 1 NO₂ predictions. Scenario 1 should therefore be interpreted as an exploratory representation of the model response to assumed

horizontal densification rather than a quantitative forecast of future development patterns or NO₂ concentrations.

Additional uncertainty arises from both the fitted RF-LUR model and the scenario assumptions. The scenario outputs are presented as point estimates and no formal prediction intervals were derived. Accordingly, small differences between baseline and scenario predictions, particularly the localised changes close to zero, should not necessarily be interpreted as meaningful changes in real-world NO₂ concentrations, as they may fall within the predictive uncertainty of the model. Sources of uncertainty include measurement error in the monitoring data, predictor uncertainty, model-fitting and sampling variability, and the magnitude and spatial allocation of the assumed scenario modifications. These uncertainties were not formally propagated through the scenario analysis. The predicted changes should therefore be interpreted primarily in terms of their relative magnitude and spatial pattern, with greater confidence placed on broad, spatially coherent modelled responses than on small changes at individual locations. Future work could quantify uncertainty through resampling approaches such as bootstrap-based model ensembles or other methods capable of estimating prediction intervals under both baseline and scenario conditions.

A further limitation relates to the use of an empirically derived LUR model for scenario evaluation. The RF-LUR framework was calibrated using observed 2019 relationships between NO₂ concentrations and urban characteristics. When predictor variables are modified under future scenarios, the model assumes that these relationships remain valid beyond the range of conditions represented in the calibration dataset. Consequently, the scenario results should be interpreted as indicative model responses rather than direct

predictions of real-world outcomes. Future studies could combine LUR approaches with dispersion modelling or CFD-based analyses to provide a more mechanistic assessment of urban-design interventions.

5 Conclusions

This study developed a neighbourhood-scale NO₂ land-use regression framework for the City of London that explicitly incorporates urban morphology (e.g., building density and height-related metrics) alongside conventional land-use and traffic predictors. Using an ADDRESS-based buffer optimisation strategy, the Random Forest LUR model provided the strongest predictive performance (adjusted $R^2 = 0.88$), supporting the value of machine-learning approaches for capturing the non-linear and interaction-driven behaviour of NO₂ in dense city-centre environments.

Scenario testing indicated associations between changes in urban-form predictors and intra-urban NO₂ patterns. Changes that increased built density generally associated with higher predicted NO₂ concentrations, whereas interventions that increased vertical heterogeneity and/or improved local ventilation were associated with reductions in NO₂, reaching up to $\sim 7 \mu\text{g m}^{-3}$ in the most responsive locations. Traffic controls was associated with corridor-focused reductions in predicted NO₂ near major roads, and tree-canopy expansion was associated with modest but spatially variable changes in predicted NO₂, indicating that the effectiveness of greening depends strongly on placement and surrounding morphology.

Finally, combining scenario—incorporating traffic management, targeted greening, and building-form adjustments, was associated with the largest overall reductions in predicted NO₂ among the scenarios evaluated, indicating that multi-lever strategies may offer

greater potential for NO₂ mitigation than individual interventions, while still requiring careful local design to avoid small adverse effects in specific micro-environments.

Overall, the work provides an evidence base for integrating morphology-sensitive analytics into air-quality management and urban design appraisal. For policymakers and planners in high-density cities, the results suggest the value of considering urban morphology alongside emission-reduction measures and of prioritising interventions in locations where the model indicates the greatest potential for NO₂ reductions.

6 References

- Augusto, B., Coelho, S., Rafael, S., Coelho, M. C., & Ferreira, J. (2025). How does urban morphology impact cities air quality? A modelling study. *Science of The Total Environment*, 973, 179138. <https://doi.org/10.1016/j.scitotenv.2025.179138>
- Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K. T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., ... de Hoogh, K. (2013). Development of NO₂ and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmospheric Environment*, 72, 10–23. <https://doi.org/10.1016/j.atmosenv.2013.02.037>
- Boogaard, H., Patton, A. P., Atkinson, R. W., Brook, J. R., Chang, H. H., Crouse, D. L., Fussell, J. C., Hoek, G., Hoffmann, B., Kappeler, R., Kutlar Joss, M., Ondras, M., Sagiv, S. K., Samoli, E., Shaikh, R., Smargiassi, A., Szpiro, A. A., Van Vliet, E. D. S., Vienneau, D., ... Forastiere, F. (2022). Long-term exposure to traffic-related air pollution and selected health outcomes: A systematic review and meta-analysis. *Environ Int*, 164, 107262. <https://doi.org/10.1016/j.envint.2022.107262>
- Brokamp, C., Jandarov, R., Rao, M. B., LeMasters, G., & Ryan, P. (2017). Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches. *Atmos Environ* (1994), 151, 1-11. <https://doi.org/10.1016/j.atmosenv.2016.11.066>
- Buccolieri, R., Jeanjean, A. P. R., Gatto, E., & Leigh, R. J. (2018). The impact of trees on street ventilation, NO_x and PM_{2.5} concentrations across heights in Marylebone Rd street canyon, central London. *Sustainable Cities and Society*, 41, 227-241. <https://doi.org/https://doi.org/10.1016/j.scs.2018.05.030>
- Castro, A., Kunzli, N., & Gotschi, T. (2017). Health benefits of a reduction of PM(10) and NO(2) exposure after implementing a clean air plan in the Agglomeration Lausanne-

- Morges. *Int J Hyg Environ Health*, 220(5), 829-839.
<https://doi.org/10.1016/j.ijheh.2017.03.012>
- CERC, (2016). Calculation of Urban Morphology Parameterisations for London for use with the ADMS-Urban Dispersion Model. Cambridge, UK. <http://technical.cloud-journals.com/index.php/IJARSG/article/view/Tech-602>
- Champendal, A., Kanevski, M., Huguenot, PE. (2014). Air Pollution Mapping Using Nonlinear Land Use Regression Models. *Computational Science and Its Applications*.
https://doi.org/10.1007/978-3-319-09150-1_50
- Chen, Y.-H., Tu, Y.-P., Sung, S.-Y., Weng, W.-C., Huang, H.-L., & Tsai, Y. I. (2022). A comprehensive analysis of the intervention of a fresh air ventilation system on indoor air quality in classrooms. *Atmospheric Pollution Research*, 13(4).
<https://doi.org/10.1016/j.apr.2022.101373>
- City of London Local Plan (2021). KEY AREAS OF CHANGE. City Plan 2036.
- City of London (2019). City of London Transport Strategy.
- City of London (2021). City's Plan 2036: Shaping the Future City. City Plan 2036 March 2021 (cityoflondon.gov.uk)
- Criado, A., Armengol, J. M., Petetin, H., Rodriguez-Rey, D., Benavides, J., Guevara, M., Pérez García-Pando, C., Soret, A., & Jorba, O. (2023). Data fusion uncertainty-enabled methods to map street-scale hourly NO₂ in Barcelona: a case study with CALIOPE-Urban v1.0. *Geoscientific Model Development*, 16(8), 2193-2213.
<https://doi.org/10.5194/gmd-16-2193-2023>
- Department for Transport (2019). Annual Report and Accounts.
<https://assets.publishing.service.gov.uk/media/5d30514fed915d2fe47af3f1/2018-2019-dft-annual-report-print.pdf>
- Dimakopoulou, K., Samoli, E., Analitis, A., Schwartz, J., Beevers, S., Kitwiroon, N., Beddows, A., Barratt, B., Rodopoulou, S., Zafeiratou, S., Gulliver, J., & Katsouyanni, K. (2022). Development and Evaluation of Spatio-Temporal Air Pollution Exposure Models and Their Combinations in the Greater London Area, UK. *Int J Environ Res Public Health*, 19(9). <https://doi.org/10.3390/ijerph19095401>
- European Environment Agency, (2022). Air quality in Europe.
<https://www.eea.europa.eu/publications/air-quality-in-europe-2022>
- Fecht, D. et al. (2016). Spatial and temporal associations of road traffic noise and air pollution in London. *Environment International*, 88, 235–242.
- Futcher, J. (2021). Urban Lab City' Investigating the Role of Urban Built Form on levels of Air Pollution at a Microscale level - The City of London Pilot Study.
<https://www.plea2020.org/>
- Futcher, J., & Mills, G. (2020). 'Urban Lab City': Investigating the Role of Built Form on Air Quality and Urban Microclimates– City of London Case Study.
- Gao, Y., Wang, Z. Y., Liu, C., & Peng, Z. R. (2019). Assessing neighborhood air pollution exposure and its relationship with the urban form. *Building and Environment*, 155, 15-24. <https://doi.org/10.1016/j.buildenv.2018.12.044>
- Ge, Y., Fu, Q., Yi M., Chao, Y., Lei, X., Xu, X., Yang, Z., Hu, Z., Kan, H., Cai, J. (2022). High spatial resolution land-use regression model for urban ultrafine particle exposure assessment in Shanghai, China. *Science of The Total Environment*.
<https://doi.org/10.1016/j.scitotenv.2021.151633>
- Greater London Authority (GLA) (2019). London Atmospheric Emissions Inventory 2016. Greater London Authority, London, UK.

- Gu, Z.L., Zhang, Y.W., Cheng, Y., Lee, S.C. (2011). Effect of uneven building layout on air flow and pollutant dispersion in non-uniform street canyons. *Building and Environment*, 46(12), 2657–2665. <https://doi.org/10.1016/j.buildenv.2011.06.028>
- Hang, J., Li, Y., Sandberg, M., Buccolieri, R., Di Sabatino, S. (2012). The influence of building height variability on pollutant dispersion and pedestrian ventilation in idealized high-rise urban areas. *Building and Environment*, 56, 346–360. <https://doi.org/10.1016/j.buildenv.2012.03.023>
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33), 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>
- Jeanjean, A.P.R., Buccolieri, R., Eddy, J., Monks, P.S., Leigh, R.J. (2017). Air quality affected by trees in real street canyons: The case of Marylebone neighbourhood in central London. *Urban Forestry & Urban Greening*. <https://doi.org/10.1016/j.ufug.2017.01.009>
- Jerrett, M., Finkelstein, M. M., Brook, J. R., Arain, M. A., Kanaroglou, P., Stieb, D. M., Gilbert, N. L., Verma, D., Finkelstein, N., Chapman, K. R., & Sears, M. R. (2009). A Cohort Study of Traffic-Related Air Pollution and Mortality in Toronto, Ontario, Canada. *Environmental Health Perspectives*, 117(5), 772–777. <https://doi.org/doi:10.1289/ehp.11533>
- Kim, J.W., Baik, J.J., Park, S.B., Han, B.S. (2023). Impacts of building-height variability on turbulent coherent structures and pollutant dispersion: Large-eddy simulations. *Atmospheric Pollution Research*, 14(5), 101736. <https://doi.org/10.1016/j.apr.2023.101736>
- Kunzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., Jr., Puybonnieux-Textier, V., Quenel, P., Schneider, J., Seethaler, R., Vergnaud, J. C., & Sommer, H. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lancet*, 356(9232), 795–801. [https://doi.org/10.1016/S0140-6736\(00\)02653-2](https://doi.org/10.1016/S0140-6736(00)02653-2)
- Larkin, A., Anenberg, S., Goldberg, D. L., Mohegh, A., Brauer, M., & Hystad, P. (2023). A global spatial-temporal land use regression model for nitrogen dioxide air pollution. *Frontiers in Environmental Science*, 11. <https://doi.org/10.3389/fenvs.2023.1125979>
- Lee, M., Brauer, M., Wong, P., Tang, R., Tsui, T. H., Choi, C., Cheng, W., Lai, P.-C., Tian, L., Thach, T.-Q., Allen, R., & Barratt, B. (2017). Land use regression modelling of air pollution in high density high rise cities: A case study in Hong Kong. *Science of The Total Environment*, 592, 306–315. <https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.03.094>
- Ling, C., Zhang, Y., Shen, Q., Dai, R., Lou, B., Kang, Y., & He, S. (2025). Urban form and environmental characteristics as drivers of air pollution exposure variability and inequality in Fujian Province, China. *GeoHealth*, 9, e2025GH001393. <https://doi.org/10.1029/2025GH001393>
- Lindberg, F., Grimmond, C.S.B., Gabey, A., Huang, B., Kent, C.W., Sun, T., et al. (2018). Urban Multi-scale Environmental Predictor (UMEP): An integrated tool for city-based climate services. *Environmental Modelling & Software*, 99, 70–87.
- Liu, M., Chen, H., Wei, D., Wu, Y., & Li, C. (2021). Nonlinear relationship between urban form and street-level PM_{2.5} and CO based on mobile measurements and gradient boosting

- decision tree models. *Building and Environment*, 205, 108265.
<https://doi.org/10.1016/j.buildenv.2021.108265>
- LPMR, (2021). Local Plan Core Strategic Policy CS14: Tall Buildings. City of London.
- Ma, X., Longley, I., Salmond, J., Gao, J. (2020). PyLUR: Efficient software for land use regression modeling the spatial distribution of air pollutants using GDAL/OGR library in Python. *Front. Environ. Sci. Eng.*, 14(3): 44 <https://doi.org/10.1007/s11783-020-1221-5>
- Macintyre, H. L., Heaviside, C., Neal, L. S., Agnew, P., Thornes, J., Vardoulakis, S., & Milojevic, A. (2023). Future urban air quality and public health impacts under climate change and emissions reduction scenarios in the United Kingdom. *Environment International*, 181, 108046. <https://doi.org/10.1016/j.envint.2023.108046>
- Metrangolo, C., Dinoi, A., Esposito, A., Pappacogli, G., Donato, A., Santiago, J. L., & Buccolieri, R. (2024). Assessing urban air pollution dynamics: The impact of traffic emissions and urban morphology in Lecce and Bari, Italy. *Bulletin of Atmospheric Science and Technology*, 5, 13. <https://doi.org/10.1007/s42865-024-00076-9>
- Miskell, G., Salmond, J. A., Longley, I. D., & Dirks, K. N. (2015). A novel approach in quantifying the effect of urban design features on local-scale air pollution in central urban areas. *Environmental Science & Technology*, 49(15), 9004–9011. <https://doi.org/10.1021/acs.est.5b00947>
- Molter, A., Lindley, S. (2021). Developing land use regression models for environmental science research using the XLUR tool – More than a one-trick pony. *Environmental Modelling & Software*, 146. <https://doi.org/10.1016/j.envsoft.2021.105108>
- Munir, S., Mayfield, M., Coca, D., Mihaylova, L. S. (2020). A Nonlinear Land Use Regression Approach for Modelling NO₂ Concentrations in Urban Areas from Low-Cost Sensors and Diffusion Tubes. *Atmosphere*, 11, 736. doi:10.3390/atmos11070736
- Samoli, E., Atkinson, R.W., Analitis, A., Fuller, G.W., Green, D.C., Mudway, I., Anderson, H.R., Kelly, F.J. (2016). Associations of short-term exposure to traffic-related air pollution with cardiovascular and respiratory hospital admissions in London, UK. *Occupational and Environmental Medicine*, 73, 300–307.
- Sabaliauskas, K., Jeong, C.-H., Yao, X., Reali, C., Sun, T., & Evans, G. J. (2015). Development of a land-use regression model for ultrafine particles in Toronto, Canada. *Atmospheric Environment*, 110, 84-92. <https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.02.018>
- Su, J. G., Jerrett, M., & Beckerman, B. (2009). A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. *Sci Total Environ*, 407(12), 3890-3898. <https://doi.org/10.1016/j.scitotenv.2009.01.061>
- Tang, R., Tian, L. W., Thach, T. Q., Tsui, T. H., Brauer, M., Lee, M., Allen, R., Yuchi, W., Lai, P. C., Wong, P., & Barratt, B. (2018). Integrating travel behavior with land use regression to estimate dynamic air pollution exposure in Hong Kong. *Environment International*, 113, 100-108. <https://doi.org/10.1016/j.envint.2018.01.009>
- Wan, Y., Yuan, L. & Xu, X. (2022). Adopting urban morphological indicators to land use regression modeling of seasonal mean PM_{2.5} concentrations for a high-density city. *Air Qual Atmos Health* 15, 559–573. <https://doi.org/10.1007/s11869-021-01134-3>
- Wang, B., Sun, S., & Duan, M. (2018). Impact of building form on the wake flow wind potential. *Energy Procedia*, 153, 383-388. <https://doi.org/https://doi.org/10.1016/j.egypro.2018.10.070>

- Ward, H. C., & Grimmond, C. S. B. (2017). Assessing the impact of changes in surface cover, human behaviour and climate on energy partitioning across Greater London. *Landscape and Urban Planning*, 165, 142-161. <https://doi.org/10.1016/j.landurbplan.2017.04.001>
- Yao, M., Smith, M., & Peng, C. (2025). Modelling the effects of vegetation and urban form on air quality in real urban environments: A systematic review of measurements, methods, and predictions. *Urban Forestry & Urban Greening*, 105, 128693. <https://doi.org/10.1016/j.ufug.2025.128693>
- Yang, J., Jin, S., Xiao, X., Jin, C., Xia, J., Li, X., & Wang, S. (2019). Local climate zone ventilation and urban land surface temperatures: Towards a performance-based and wind-sensitive planning proposal in megacities. *Sustainable Cities and Society*, 47. <https://doi.org/10.1016/j.scs.2019.101487>
- Yli-Pelkonen, V., Scott, A. A., Viippola, V., & Setälä, H. (2017). Trees in urban parks and forests reduce O₃, but not NO₂ concentrations in Baltimore, MD, USA. *Atmospheric Environment*, 167, 73-80. <https://doi.org/10.1016/j.atmosenv.2017.08.020>
- Yuan, Y., Luo, Z. W., Liu, J., Wang, Y. W., & Lin, Y. Y. (2018). Health and economic benefits of building ventilation interventions for reducing indoor PM_{2.5} exposure from both indoor and outdoor origins in urban Beijing, China. *Science of The Total Environment*, 626, 546-554. <https://doi.org/10.1016/j.scitotenv.2018.01.119>

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: