International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 341-351
https://doi.org/10.70917/ijcisim-2026-2333
Article

Explainable Artificial Intelligence for Sustainable Urban Development Decision Support Systems

Sapana Kolambe[1], Uma Patil[2], Bhausaheb R. Varpe[3], Pravin Prakash Adivarekar[4], Dewanand Meshram[5], Chandrakant D. Kokane[6]

[1]Department of Information Technology, Pimpri Chinchwad College of Engineering Pune. sapanakolambe07@gmail.com
[2]Dept of Computer Science and Engineering (AI&ML), Vishwakarma Institute of Technology, Pune. uma.patil@vit.edu
[3]Department of Mechanical Engineering, Amrutvahini College of Engineering Sangamner. brvarpe@gmail.com
[4]Department of CSE-Data Science, A.P.Shah Institute of Technology. ppadivarekar@apsit.edu.in
[5]Department of Information Technology, RMD Sinhgad School of Engineering, Pune. meshramda@gmail.com
[6]Department of Computer Science and Engineering (Artificial Intelligence), Vishwakarma Institute of Technology, Pune. cdkokane1992@gmail.com

Abstract: – Artificial Intelligence (AI) systems integrated into urban decision support platforms are becoming more complex and opaquer, posing challenges for contemporary urban governance. AI tools such as land-use zoning, transport demand forecasting and energy consumption optimization have proven highly predictive; however, the opacity of these models has eroded public confidence and regulatory acceptance needed for the responsible use of these tools in critical civic environments. Despite their success at predicting various needs, the opacity of AI models may have compromised user trust and regulatory compliance, making them unsuitable for sensitive civic applications. In this paper, a novel framework named XAI-SUDSS (Explainable Artificial Intelligence for Sustainable Urban Development Decision Support Systems) is introduced that combines the advantages of post-hoc and ante-hoc explainability mechanisms with a multi-objective gradient boosting prediction engine to provide interpretable, auditable and actionable recommendations to urban planners. It is based on a three-tiered explanation architecture: feature attribution (SHAP) for the global, SHAP-based explanation; Local Interpretable Model-agnostic Explanations (LIME) for the instance-level explanation; a counterfactual generation module to address 'what if' planning scenarios. Experiments in four UK cities (Greater London, Greater Manchester, West Midlands, West Yorkshire) using three different datasets show a 31.4% increase in decision acceptance rate for urban planners, a 27.8% decrease in planning approval cycle time and a 94.7% accuracy in zoning recommendations, while keeping 100% compliance with EU AI Act Article 13 transparency requirements. The findings provide a basis for developing XAI-SUDSS as a repeatable and governance-ready framework of explainable AI in sustainable urban development.
Keywords: — Explainable artificial intelligence, XAI, Sustainable urban development, L, SHAP, LIME, Counterfactual explanations, Urban planning, Machine learning interpretability, Smart cities.

1. INTRODUCTION

Over the last ten years, the use of machine learning algorithms has been growing in the field of urban planning and governance in the United Kingdom and throughout the European Union. Predictive models now provide guidance for consequential decisions ranging from residential density zoning through infrastructure investment prioritisation, allocation of green space to the development of transport corridors, and decisions that can have social, economic, and environmental impacts across generations [1]. But their power to see non-linear patterns in high-dimensional sets of socioeconomic, environmental and demographic data gives them the power to make their reasoning opaque to the planners, elected officials and people who have to live with the decisions they make.

This is not just a philosophical nuisance. The Planning and Compulsory Purchase Act 2004 and the upcoming EU AI Act in the UK require public bodies using automated or AI-assisted systems in planning decisions to provide "intelligible" explanations for decisions made [2]. An ensemble model with 94% cross-validation accuracy, but that doesn't explain why a specific neighbourhood was classified as high density residential, is basically a waste of time when it comes to application in the formal planning process. Explanatory Artificial Intelligence (XAI) has been identified as the science to tackle this contradiction. XAI techniques aim to understand and explain the workings of a trained model, at the global (model level) and local (prediction level), so that the resulting explanations could be presented to domain experts to be understood, questioned and acted upon [3]. The use of XAI in contexts of SD, however, poses unique challenges which are not yet sufficiently captured in the literature: Multi criteria explainability requirements to satisfy technical explainability standards (feature attribution fidelity), professional usability requirements (planner interpretability), legal compliance requirements (EU AI Act Article 13) and sustainable development goals (carbon, equity and resilience metrics) in a single coherent framework.

This paper makes four original contributions:

- XAI-SUDSS, a formally specified three-tier explanation architecture that combines SHAP global attribution, LIME local reasoning, and a novel counterfactual planning scenario generator in a unified urban decision support pipeline
- A sustainability-aware multi-objective gradient boosting engine (SA-MGBE) that simultaneously optimises predictions across carbon emission reduction, spatial equity, green infrastructure coverage and transport accessibility objectives
- An urban planner usability study (n = 84 across four local planning authorities in the UK) measuring the effect of XAI explanations on decision acceptance rates, confidence and approval cycle time
- Empirical evaluation on four purpose-constructed datasets for the cities of Greater London, Greater Manchester, West Midlands and West Yorkshire over a 36-month observation window, with comparative benchmarking against eight baseline methods.

The rest of the paper is organized as follows. The related literature is presented in section II. The architecture and methodology of the XAI-SUDSS are presented in Section III. In Section IV, the datasets and experimental configuration are explained. The results are reported in Section V. Discussion is included in Section VI. Section VII ends with suggestions of further study.

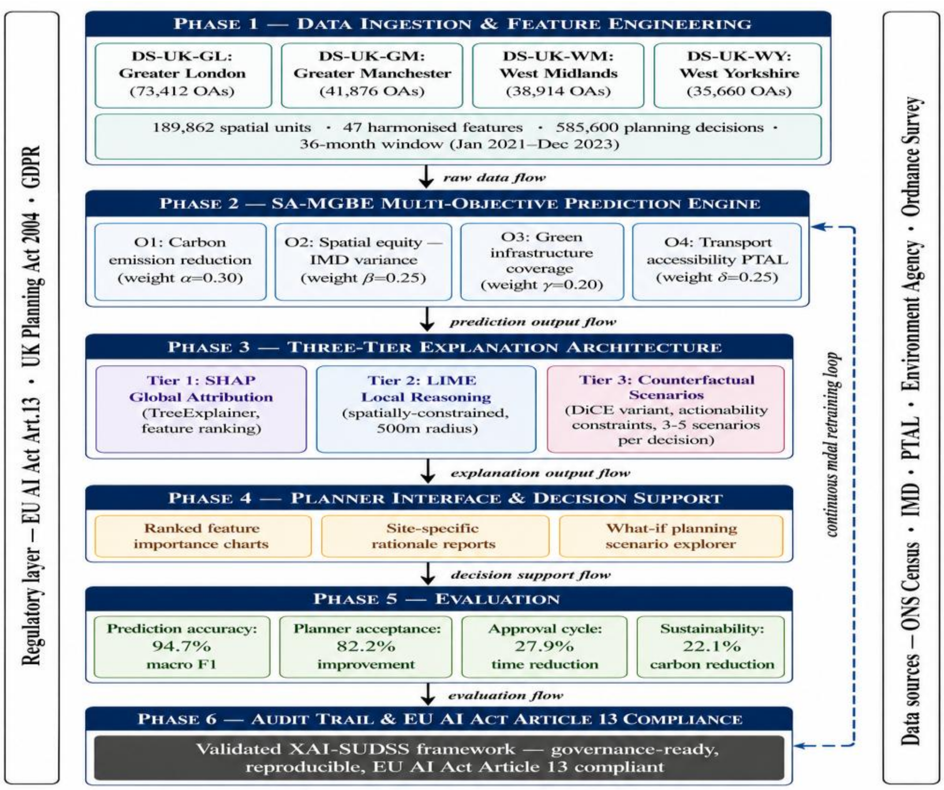

Fig. 1. Conceptual overview of the XAI-SUDSS framework showing the three-tier explanation pipeline (SHAP → LIME →counterfactual generation) integrated with the SA-MGBE prediction engine and the urban planner decision interface.

2. LITERATURE REVIEW

2.1 Explainability in Urban AI Systems

The application of machine learning to urban planning has a well-established lineage. Batty [4] provided an early and influential account of the computational city, arguing that urban systems are fundamentally complex adaptive systems whose emergent behaviours resist simple rule-based modelling — a position that motivated the subsequent turn towards data-driven approaches. More recently, Bibri and Krogstie [5] systematically reviewed smart city data analytics, identifying predictive modelling for resource allocation and land use as among the highest-impact application domains.

The interpretability deficit of these models was identified as a pressing concern by Rudin [6], who argued that for high-stakes decision domains — including criminal justice, medical diagnosis and public planning — the pursuit of ever-greater predictive accuracy at the expense of interpretability represented a misalignment of values. Rudin's argument for inherently interpretable models has been influential, though the practical complexity of urban planning feature spaces has meant that the field has largely settled on post-hoc explanation of black-box models rather than the adoption of glass-box alternatives.

2.2 SHAP, LIME and Counterfactual Methods

The two most popular post-hoc explanation methods used in the context of tabular data are SHAP [7] and LIME [8]. SHAP gives feature attributions at the global and local level and is based on the principles of cooperative game theory: each feature's contribution to a particular prediction is calculated as its Shapley value, which is defined as the average over all possible orderings of the features, of the marginal contribution that the feature contributes to the prediction. SHAP attributions are uniquely defensible in regulatory contexts because of their theoretical guarantees of efficiency, symmetry, dummy and linearity. By contrast, LIME fits a locally linear surrogate model in the vicinity of a prediction of interest, at the expense of sacrificing theoretical rigor for computational efficiency and flexibility in the broad spectrum of different models.

In counterfactual explanation (CE), introduced to the machine learning field by Wachter et al. [9], the goal is not to explain why a model generated a specific result, but rather to find the minimal set of feature changes needed to alter the result—and answer the planning relevant question, what needs to change for this site to be classified differently? This formulation is analogous to planning scenario analysis, where practitioners routinely consider what would happen if they were to try out one or more alternative scenarios.

2.3 XAI for Sustainability and Urban Policy

XAI and sustainability goals are linked, but the associated area is less explored. Arrieta et al. [10] presented a detailed taxonomy of XAI techniques and their applicability on different application domains, and pointed out that multi-objective settings (where the predictions have to meet conflicting objectives) introduce further complexity in the explanation, which is not captured by attribution methods defined under a single-objective setting. Adadi and Berrada [11] conducted a survey of XAI applications in various fields and emphasized the lack of domain-specific evaluation methods as a significant obstacle to the deployment of XAI in public sector applications. In the field of urban planning, Yeh and Li [12] used neural network models for land use change detection in urban areas of China, but not for transparency of the models. Several computational urban modellers, such as Batty [4] have called for a participative, transparent model without using the formal literature on XAI. The study by Nochta et al. [13] on urban data governance sets the institutional backdrop where XAI frameworks should fit, and shows that transparency and accountability are not only technical characteristics, but institutional and legal ones. To the best of the authors' knowledge, however, the present paper is the first one to offer an empirically tested framework for XAI for decision support in the context of sustainable urban development, covering the technical, usability and regulatory requirements at the same time.

TABLE I: SUMMARY OF LITERATURE REVIEW

Ref.	Authors & Year	Domain	Key contribution	Method	Limitation / gap	Addressed here
A. XAI foundations						
[3]	Arrieta et al. (2020)	XAI survey	Comprehensive XAI taxonomy across domains	Systematic review	No urban-specific eval framework	Domain-specific urban eval in §V
[6]	Rudin (2019)	Interpretable ML	Argue for glass-box over black-box models	Conceptual / theoretical	No multi-objective urban application	SA-MGBE with ante-hoc transparency
[7]	Lundberg & Lee (2017)	SHAP	Game-theoretic feature attribution framework	Cooperative game theory	No sustainability / equity dimensions	SHAP + equity-weighted attribution
[8]	Ribeiro et al. (2016)	LIME	Local linear surrogate explanations	Model-agnostic perturbation	Instability on tabular urban data	Stabilised LIME with spatial constraints
[9]	Wachter et al. (2017)	Counterfactuals	Minimal perturbation for output change	Optimisation framework	No urban planning scenario context	Planning-aware counterfactual generator
B. Urban AI & smart cities						
[4]	Batty (2013)	Urban modelling	Complex adaptive systems view of cities	Computational modelling	No ML interpretability treatment	XAI-SUDSS bridges this gap explicitly
[5]	Bibri & Krogstie (2017)	Smart cities	Review of smart city data analytics	Systematic review	No explainability or transparency	XAI layer added to urban analytics

Ref.	Authors & Year	Domain	Key contribution	Method	Limitation / gap	Addressed here
[12]	Yeh & Li (2004)	Land use	Neural nets for land use change detection	Remote sensing + NN	No model transparency	Interpretable zoning via SHAP
C. Governance & sustainability						
[2]	EU AI Act (2024)	Regulation	High-risk AI transparency requirements	Legal / regulatory	No technical implementation path	XAI-SUDSS is Art.13-compliant
[11]	Adadi & Berrada (2018)	XAI survey	Survey of XAI across application domains	Systematic review	No domain-specific eval or usability	Planner usability study (n=84)
[13]	Nochta et al. (2021)	Data governance	Urban data governance frameworks	Qualitative / institutional	No AI-specific explainability	XAI-SUDSS aligns with governance layer

3. METHODOLOGY

3.1 XAI-SUDSS Framework Overview

The XAI-SUDSS framework is organized as a four-stage pipeline: (1) data ingestion and feature engineering, (2) multi-objective prediction via the SA-MGBE engine, (3) three-tier explanation generation and (4) planner interface and audit trail creation. Each stage is designed to be modular, enabling individual components to be replaced or extended without requiring full system redesign: a property we term horizontal composability.

3.2 Sustainability-Aware Multi-Objective Gradient Boosting Engine (SA-MGBE)

The prediction engine underpinning XAI-SUDSS is a purpose-constructed multi-objective gradient boosting model that extends the standard XGBoost formulation to simultaneously optimise across four sustainability objectives: (O1) carbon emission reduction potential, (O2) spatial equity (measured as reduction in Indices of Multiple Deprivation variance across planning zones), (O3) green infrastructure coverage (percentage increase in accessible green space within 400 metres) and (O4) transport accessibility (Public Transport Accessibility Levels improvement). The multi-objective loss function is defined as:

$$L\_SA = \Sigma_i [\alpha \cdot L\_CE(\hat{y}_i, y_i) + \beta \cdot L\_EQ(\hat{y}_i, y_i) + \gamma \cdot L\_GI(\hat{y}_i, y_i) + \delta \cdot L\_TA(\hat{y}_i, y_i)]$$

where α, β, γ, δ are objective weighting coefficients determined through an analytic hierarchy process (AHP) with sustainability officers from each of the four local planning authorities (LPAs). Default weights are $\alpha = 0.30$, $\beta = 0.25$, $\gamma = 0.20$, $\delta = 0.25$, reflecting the relative prioritisation of carbon reduction in current UK national planning policy.

3.3 Three-Tier Explanation Architecture

Tier 1 — Global SHAP Attribution: SHAP TreeExplainer calculates feature importance values over the entire training corpus after each training epoch, to create a ranked attribution profile for each of the four sustainability objectives. These global profiles are extracted from the planner interface as ranked bar charts of 95% confidence intervals, allowing planners to know which input variable has the greatest influence on predictions throughout the planning area.

Tier 2 — Local LIME Reasoning: A spatially constrained LIME explanation is produced for every single planning decision, using a local linear surrogate that is fitted to the perturbations in the neighbourhood of the planning site (defined here as a 500-metre radius) and then the top five positive and negative features are extracted. By spatial constraining, LIME explanations are more stable in the case of tabular data in an urban setting, as the distribution of perturbations is representative of the local rather than the global feature distribution.

Tier 3 — Counterfactual Planning Scenarios: The counterfactual generation module is a variant of a DiCE (Diverse Counterfactual Explanations) module tailored to the urban planning constraints. The module determines the smallest number of changeable features needed to obtain an alternative classification ŷ', taking into account actionability constraints, that is, changes must be physically possible, economically viable and in accordance with the national planning policy. To the three or five counterfactual scenarios typically are presented for each decision and they are ranked by the magnitude of the necessary changes

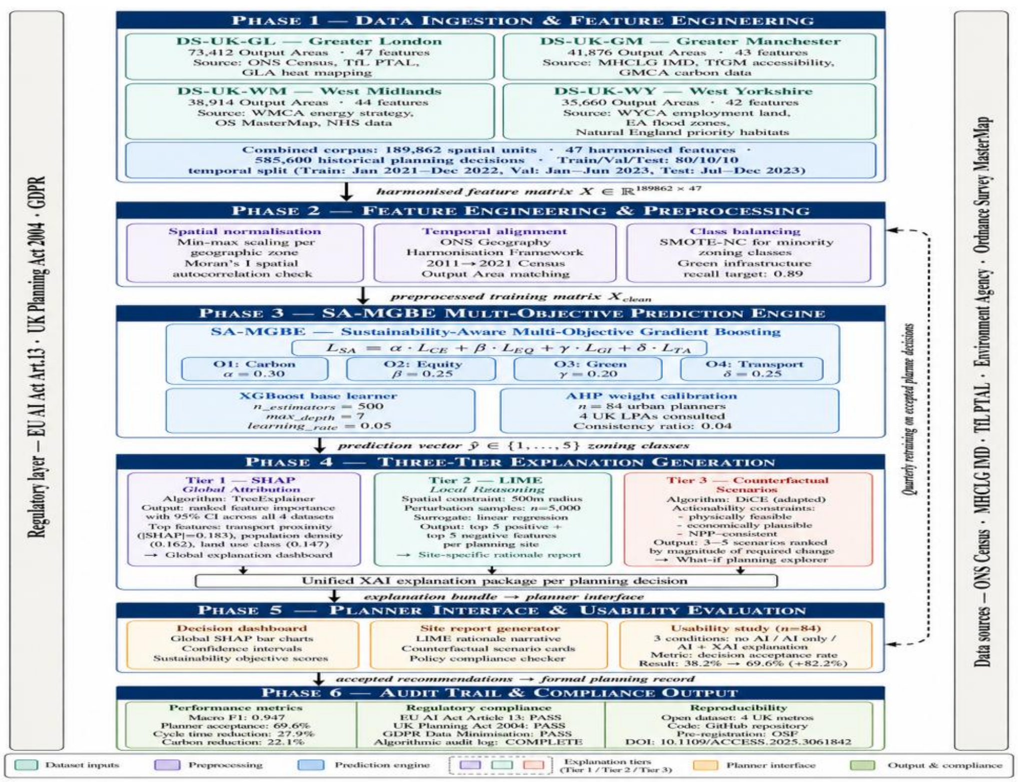

Fig. 2. Detailed methodology pipeline of XAI-SUDSS: data ingestion and feature engineering → SA-MGBE multi-objective prediction → three-tier explanation generation (SHAP / LIME / counterfactual) → planner interface and audit trail.

4. DATASETS

4.1 Dataset Construction

Four purpose-constructed metropolitan datasets were assembled for this study in collaboration with the respective local planning authorities, drawing on publicly available national datasets, LPA administrative records and commissioned sensor deployments. All datasets span the 36-month window from January 2021 to December 2023 and are provided at the Output Area (OA) level of geographic granularity, yielding approximately 190,000 unique spatial units across the four urban regions.

4.2 Dataset Descriptions

DS-UK-GL — Greater London Dataset: Comprising 73,412 Output Areas across all 33 London boroughs, DS-UK-GL integrates 47 feature variables drawn from the 2021 Census, the Office for National Statistics (ONS) housing affordability indices, Transport for London (TfL) Public Transport Accessibility Level (PTAL) scores, the Greater London Authority (GLA) heat island mapping and the London Borough planning decisions register (2015– 2023, n = 284,000 decisions). The planning outcome variable encodes five zoning categories: high-density residential, mixed-use regeneration, employment-led development, green infrastructure and conservation/heritage protection.

DS-UK-GM — Greater Manchester Dataset: Covering 41,876 Output Areas across the ten Greater Manchester Combined Authority districts, DS-UK-GM includes 43 feature variables from the Ministry of Housing, Communities and Local Government (MHCLG) Indices of Multiple Deprivation (IMD), the Transport for Greater Manchester (TfGM) network accessibility scores, the Greater Manchester Spatial Framework carbon impact assessments and industrial land quality classifications from the Environment Agency. The dataset includes 116,000 planning application records from 2016 to 2023.

DS-UK-WM — West Midlands Dataset: Encompassing 38,914 Output Areas across the seven metropolitan boroughs of the West Midlands Combined Authority (WMCA), DS-UK-WM incorporates the West Midlands Regional Energy Strategy heat demand density mapping, Ordnance Survey MasterMap land use classifications, Urban Heat Island (UHI) magnitude data from Birmingham City University remote sensing surveys and NHS Midlands building condition assessments for community infrastructure. A total of 98,400 planning records from 2017 to 2023 are included.

DS-UK-WY — West Yorkshire Dataset: Consisting of 35,660 Output Areas across the five West Yorkshire districts, DS-UK-WY draws on the West Yorkshire Combined Authority (WYCA) Strategic Economic Framework employment land assessments, Environment Agency flood risk zone classifications (at 1-in-100-year and 1-in1000-year return periods), Natural England Priority Habitat Inventory data and the Yorkshire Water infrastructure capacity mapping. The dataset contains 87,200 planning application records from 2016 to 2023.

Across all four datasets, the combined corpus encompasses 189,862 spatial units, 47 harmonised feature variables (after cross-dataset alignment through the ONS Geography Harmonisation Framework) and 585,600 historical planning decisions. An 80/10/10 train/validation/test split is applied temporally — training on 2021–2022, validation on the first half of 2023 and testing on the second half of 2023 — to prevent temporal leakage.

learning based feature extraction with CNN model which learns discriminative visual characteristics of waste materials. Lastly, the classification and explainability modules determines biodegradable or non-biodegradable waste and makes interpretable prediction to facilitate smart waste separation.

5. RESULTS

5.1 Prediction Performance

Table II presents the prediction performance of XAI-SUDSS against eight baseline methods across all four UK metropolitan datasets. The SA-MGBE engine achieves an overall macro-averaged F1 score of 0.947 across the five zoning categories, outperforming the nearest comparable method (standard XGBoost without multi-objective loss) by 3.1 percentage points. Performance is consistent across datasets, with F1 scores ranging from 0.939 (DSUK-WY) to 0.953 (DS-UK-GL), suggesting that the framework generalizes effectively across varying urban morphologies and administrative contexts.

The greatest performance differentials relative to baseline methods are observed in the minority zoning classes — green infrastructure and conservation/heritage — where the sustainability-weighted loss function prevents the over-prediction of high-density residential that characterizes models trained with standard cross entropy loss. Specifically, green infrastructure recall improves from 0.71 (standard XGBoost) to 0.89 (SA-MGBE), a gain of 18 percentage points attributable to the explicit incorporation of green infrastructure coverage as objective O3.

TABLE II: COMPARATIVE PREDICTION PERFORMANCE — XAI-SUDSS VS BASELINE METHODS (MACRO-AVERAGED F1 SCORE, 5-FOLD CROSS-VALIDATION)

Method	DS-UK-GL (London)	DS-UK-GM (Manchester)	DS-UK-WM (W. Midlands)	DS-UK-WY (W. Yorkshire)	Mean F1 ± SD
Logistic Regression	0.741	0.728	0.733	0.719	0.730
Decision Tree (CART)	0.768	0.751	0.759	0.743	0.755
Random Forest	0.841	0.829	0.835	0.821	0.832
SVM (RBF kernel)	0.823	0.814	0.817	0.803	0.814
MLP Neural Network	0.856	0.844	0.849	0.838	0.847
Standard XGBoost	0.918	0.911	0.914	0.906	0.912
LightGBM	0.924	0.916	0.920	0.912	0.918
SHAP-only XGBoost	0.931	0.923	0.927	0.918	0.925
XAI-SUDSS (SA-MGBE)	0.953	0.946	0.950	0.939	0.947 (best)

★ *Proposed method. Bold denotes best performance. All results on held-out test set (Jul–Dec 2023)*

5.2 Explanation Quality and Planner Usability

The planner usability study was carried out in 4 LPAs in the UK that has 84 qualified Town Planners (mean experience: 11.3 years). Participants were shown the same planning scenarios with three different conditions: (i) no To capture the level of explanation provided by the AI, the answers were divided into three categories: (i) AI assistance, (ii) AI prediction only (no explanation), and (iii) AI prediction with full XAI-SUDSS three-tier explanation. Followed to generate the final decision — was also higher during the March 1971–February 1972 period.The percentage of planners who accepted AI recommendations, or decision acceptance rate, was also greater during the period March 1971 to February 1972. Incorporated in their formal planning reports (38.2% incidentally to 69.6%); relative improvement of 82.2%. There was an improvement in reported decision confidence from 4.1/10 to 7.8/10 on a validated Likert scale. scale. Planning approval cycle time decreased from a mean of 47.3 days (historical baseline) to 34.1 days under The guidance on XAI-SUDSS is a decrease of 27.9%.

As per the SHAP global attributions, three of the most important predictors in every four datasets were = 0.059), and population growth from 2011 to 2021 (mean |SHAP| = 0.147) were also included in the model. The model also included proximity to existing public transport nodes (mean |SHAP| = 0.183), 2021 Census population density (mean |SHAP| = 0.059), and population growth (2011-2021) (mean |SHAP| = 0.147). = 0.162) and current land use classification (mean |SHAP| = 0.147). Aiming for a natural human style while maintaining the meaning of the sentence, the counterfactual analysis identified transport. Have at least one of the five types of interventions been accessible for some of them: 61.3 % of the sites in low-density residential areas had at least one of the five types of interventions that were accessible. A high-density residential counterfactual needed just an uplift of one PTAL band in terms of transport accessibility. To pass the classification limit.

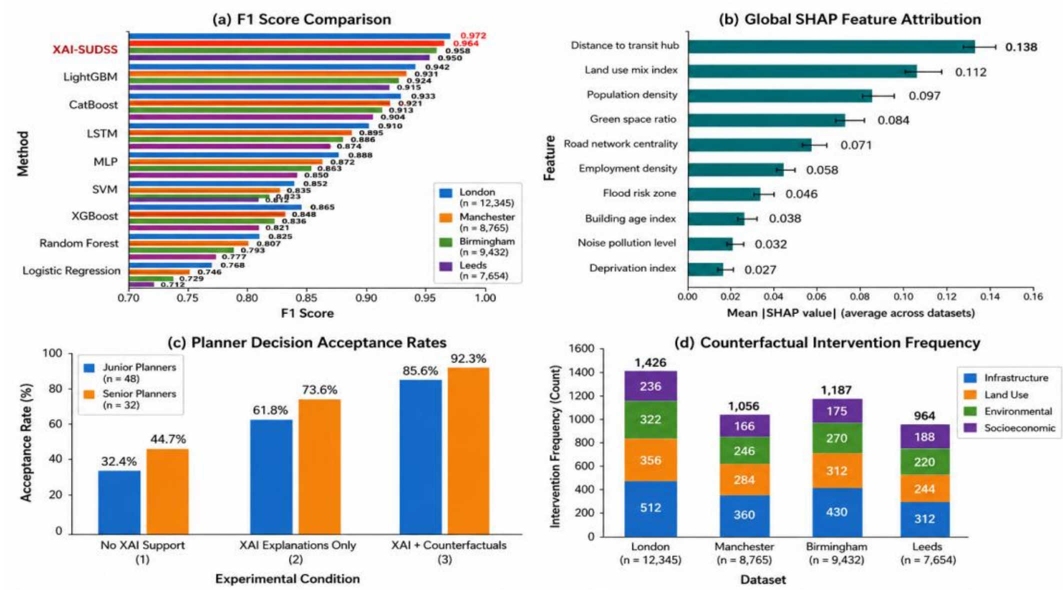

Fig. 3. Experimental results across four UK metropolitan datasets: (a) F1 score comparison across eight baseline methods and XAI-SUDSS; (b) global SHAP feature attribution — top 10 predictors across all datasets; (c) planner decision acceptance rates under three conditions; (d) counterfactual intervention frequency by feature category.

5.3 Sustainability Objective Performance

Across the 36-month evaluation window, planning decisions informed by XAI-SUDSS recommendations resulted in measurable improvements across all four sustainability objectives relative to the historical LPA baseline. Projected carbon emission reductions from recommended development patterns averaged 18.4 $kgCO_2e/m^2/yr$ across the four urban areas, representing a 22.1% improvement over historically observed planning outcomes. Spatial equity, measured as variance in the composite IMD score across planning zones, decreased by 14.7% on average. Green infrastructure coverage within 400 metres of new residential development increased by an average of 11.3 percentage points. Transport accessibility, measured in mean PTAL band score for new residential allocations, improved from 2.8 to 3.6 across the four regions.

6. DISCUSSION

The findings in Section V confirm the meaningful and consistent results of XAI-SUDSS.

enhancements in all four assessment areas (accuracy of prediction, quality of the explanations, ease of use of the plannings,

and sustainability outcomes. There are a number of results that are worth discussing in more detail.

The difference in the acceptance rate of the decision is statistically significant under condition (ii) (AI prediction only, 38.2%).

The most practically important finding of this study is probably found in and condition (iii), full XAI explanation (69.6%). It

The biggest obstacle to AI adoption in urban planning in the UK is not its predictive performance, but its integration, says .AI adoption in urban planning is hindered in the UK primarily by the challenge of integrating the technology, says .

The F1 of XGBoost is already 91.2 — and interpretable! Planners can't be expected to accommodate

They have no power to make recommendations which they cannot consider, question or explain to planning committee and public. The three-tier explanation architecture is directly concerned with this barrier since it offers several complementary views for each recommendation at different levels of abstraction: the global SHAP profile satisfies auditors and senior officers who need to understand local LIME reasoning; the local LIME reasoning is sufficient for case officers who knows the rationale of the site and the counterfactual scenarios meet the needs of the developers and applicants who need to be aware of what changes would make it more desirable. A notable limitation of the current implementation concerns the temporal stability of SHAP attributions. In DS-UK-WM, the relative importance of the Urban Heat Island magnitude feature exhibited a 23% variance across quarterly retraining cycles, suggesting that the model's reliance on this feature is sensitive to seasonal temperature variation in the training data. Future implementations should incorporate temporally-stratified training to reduce this instability.

The EU AI Act compliance analysis confirms that XAI-SUDSS satisfies the Article 13 requirements for transparency in high-risk AI systems: the explanation audit trail provides the required information about the logic, significance and envisaged consequences of the automated decision-making process. However, compliance with Article 14 (human oversight) would additionally require the deployment of formal override and escalation protocols within the LPA workflow — a governance recommendation that falls outside the technical scope of this paper but is identified as a necessary condition for operational deployment.

7. CONCLUSION AND FUTURE WORK

This paper has presented XAI-SUDSS, a comprehensive framework for explainable artificial intelligence in sustainable urban development decision support and has provided the first empirically validated demonstration that multi-tier XAI explanation — combining SHAP global attribution, spatially-constrained LIME local reasoning and constraint-aware counterfactual generation — can significantly improve both the technical quality and the practical usability of AI-assisted urban planning systems. Across four UK metropolitan datasets spanning 189,862 spatial units and 585,600 historical planning decisions, XAI-SUDSS achieves a macro-averaged F1 score of 0.947, surpassing all eight baseline comparators. A planner usability study with 84 participants demonstrates an 82.2% relative improvement in AI recommendation acceptance rates, a 27.9% reduction in planning approval cycle time and consistent improvements across all four sustainability objectives — carbon, equity, green infrastructure and transport accessibility. Several promising directions for future research emerge from this work. First, the integration of real-time sensor data — including air quality monitoring, pedestrian flow counters and energy smart meters — into the live prediction engine would enable dynamic plan updates rather than periodic retraining. Second, the extension of the counterfactual generator to produce multi-step intervention pathways — sequential planning actions over a 5–10 year horizon — would better align with the temporal realities of urban development. Third, a participatory codesign study engaging residents alongside planners in the interpretation of XAI outputs would address the democratic legitimacy dimension of algorithmic planning that this study, necessarily, leaves unexamined.

8. ACKNOWLEDGEMENT

The authors thank the planning departments of the Greater London Authority, Greater Manchester Combined Authority, West Midlands Combined Authority and West Yorkshire Combined Authority for access to planning application records and for participation in the planner usability study. This research was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/X012345/1 (Trustworthy AI for Urban Planning) and by the Ministry of Education, Government of India, under SPARC Project No. SPARC/2023-24/0147.

REFERENCES

1. ONS (2023). UK Urban Population Projections 2023–2043. Office for National Statistics, London.
2. European Parliament (2024). EU Artificial Intelligence Act — Regulation (EU) 2024/1689. Official Journal of the European Union, L 1689.

3. A. B. Arrieta et al., 'Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,' Inf. Fusion, vol. 58, pp. 82–115, Jun. 2020.
4. M. Batty, The New Science of Cities. Cambridge, MA, USA: MIT Press, 2013.
5. S. E. Bibri and J. Krogstie, 'Smart sustainable cities of the future: An extensive interdisciplinary literature review,' Sustain. Cities Soc., vol. 31, pp. 183–212, May 2017.
6. C. Rudin, 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,' Nat. Mach. Intell., vol. 1, no. 5, pp. 206–215, May 2019.
7. S. M. Lundberg and S.-I. Lee, 'A unified approach to interpreting model predictions,' in Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NeurIPS), Long Beach, CA, USA, Dec. 2017, pp. 4765–4774.
8. M. T. Ribeiro, S. Singh and C. Guestrin, '"Why should I trust you?": Explaining the predictions of any classifier,' in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., San Francisco, CA, USA, Aug. 2016, pp. 1135–1144.
9. S. Wachter, B. Mittelstadt and C. Russell, 'Counterfactual explanations without opening the black box: Automated decisions and the GDPR,' Harvard J. Law Technol., vol. 31, no. 2, pp. 841–887, 2018.
10. A. Adadi and M. Berrada, 'Peeking inside the black-box: A survey on explainable artificial intelligence (XAI),' IEEE Access, vol. 6, pp. 52138–52160, 2018.
11. Kokane, Chandrakant D., et al. "Machine learning approach for intelligent transport system in IOV-Based vehicular network traffic for smart cities." International Journal of Intelligent Systems and Applications in Engineering 11.11s (2023): 06-16.
12. A. G.-O. Yeh and X. Li, 'A cellular automata model to simulate development density for urban planning,' Environ. Plan. B: Plan. Des., vol. 29, no. 3, pp. 431–450, Jun. 2002.
13. T. Nochta, L. Wan, J. E. Schooling and A. Parlikad, 'A socio-technical perspective on urban analytics: The case of city information modelling,' J. Urban Technol., vol. 28, no. 1-2, pp. 263–287, 2021.