Prescription Prediction and Vacuums: A GNN-Based Approach to Urban Health Analysis

Minwei Zhao[1], Stephen law[1], Sanja Scepanovic[2], Daniele Quercia[2], Ivica Obadic[3]
[1]University College London [2]Nokia Bell Labs [3]Technical University of Munich

Context & Rationale

Understanding the relationships between urban characteristics and health is critical to the well-being of urban residents. This research can help urban policymakers and health managers make decisions (Obradovich, Migliorini, Paulus, & Rahwan, 2018; British Academy, 2021). The MedSAT dataset compiled by Šćepanović et al. (2023) provides comprehensive support for urban health research.

The concept of border vacuums, as termed by Jane Jacobs, can be defined as area in cities where physical barriers - such as motorways, infrastructure - create dead zones or area of inactivity (Sung&Lee2015). As border vacuum discourage interactions and connectivity, these areas can lead to significant social and health inequalities and differences (Rowan, 2017).

Few studies have captured the links between urban characteristics and health outcomes considering complex spatial relationships. The link between border vacuums and health inequalities is even less discussed.

Method

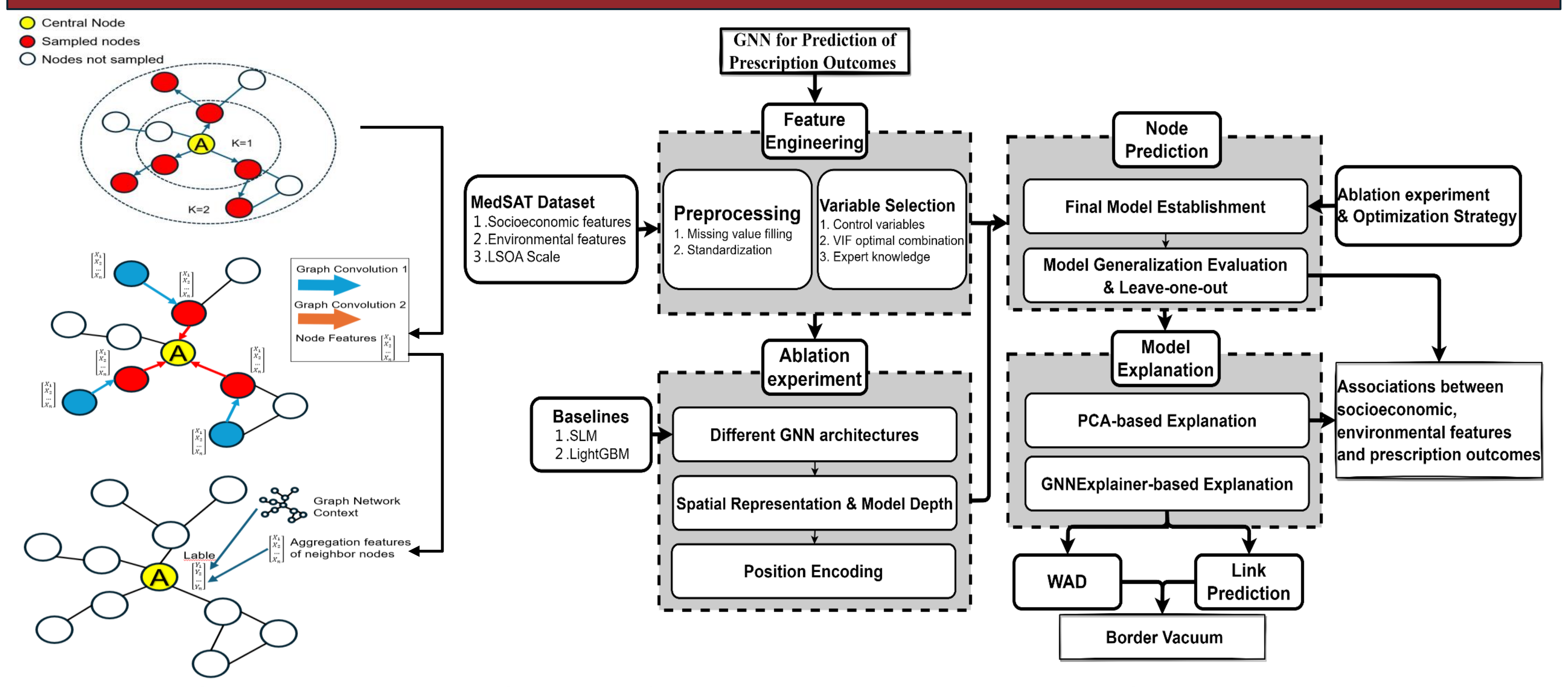


Figure 1 The left figure shows the workflow of establishing a GNN to predict prescription results in the London LSOA districts. The right figure shows the overall method construction of the entire study.

A LSOA spatial network (graph) is constructed from the census output area where the nodes are the LSOA centroids, and the edges are the adjacencies of these tracts.

Three goals of the methodology:

- To establish a accurate GNN prediction model on health prescription.
- To visualize and explain the model and discover health patterns. This is demonstrated through two explanation methods in Figure 1.
- To further connect border vacuum by prescription differences. We propose WAD (Figure 2), which is built on the local node explanation subgraph. We hope to use this metric to potentially identify border vacuums.

Weighted Average Degree

N is the number of nodes in subgraph.
W_i is the weight of node i.

$$W_i = \begin{cases} 1, & \text{if the importance of node } i \text{ is greater than the threshold,} \\ 0, & \text{if the importance of node } i \text{ is less than or equal to the threshold.} \end{cases}$$

WAD is then calculated as the mean weight of all nodes in the subgraph:

$$\text{WAD} = \frac{1}{N}\sum_{i\in N} w_i$$

Figure 2 The process of calculating WAD based on the local interpretation subgraph of GNNExplainer.

Result A: Health Outcome Prediction

Table 1 Performance of two baselines and optimized GNN models determined by ablation experiments in predicting six prescription outcomes; The ten-fold cross-validation results show that the performance of GNN is significantly improved and has stability.

R2 Evaluation	Chronic		Mental			
	Diabetes	Hypertension	Asthma	Depression	Anxiety	Opioids Use
SLM (Baseline)	0.7112	0.7253	0.6837	0.7138	0.6776	0.6770
LightGBM (Baseline)	0.6448	0.6980	0.6320	0.6603	0.6370	0.5782
GNN with 10-fold CV	**0.8581 ± 0.0079**	**0.8818 ± 0.0094**	**0.8926 ± 0.0053**	**0.9050 ± 0.0034**	**0.8363 ± 0.0223**	**0.8848 ± 0.0097**

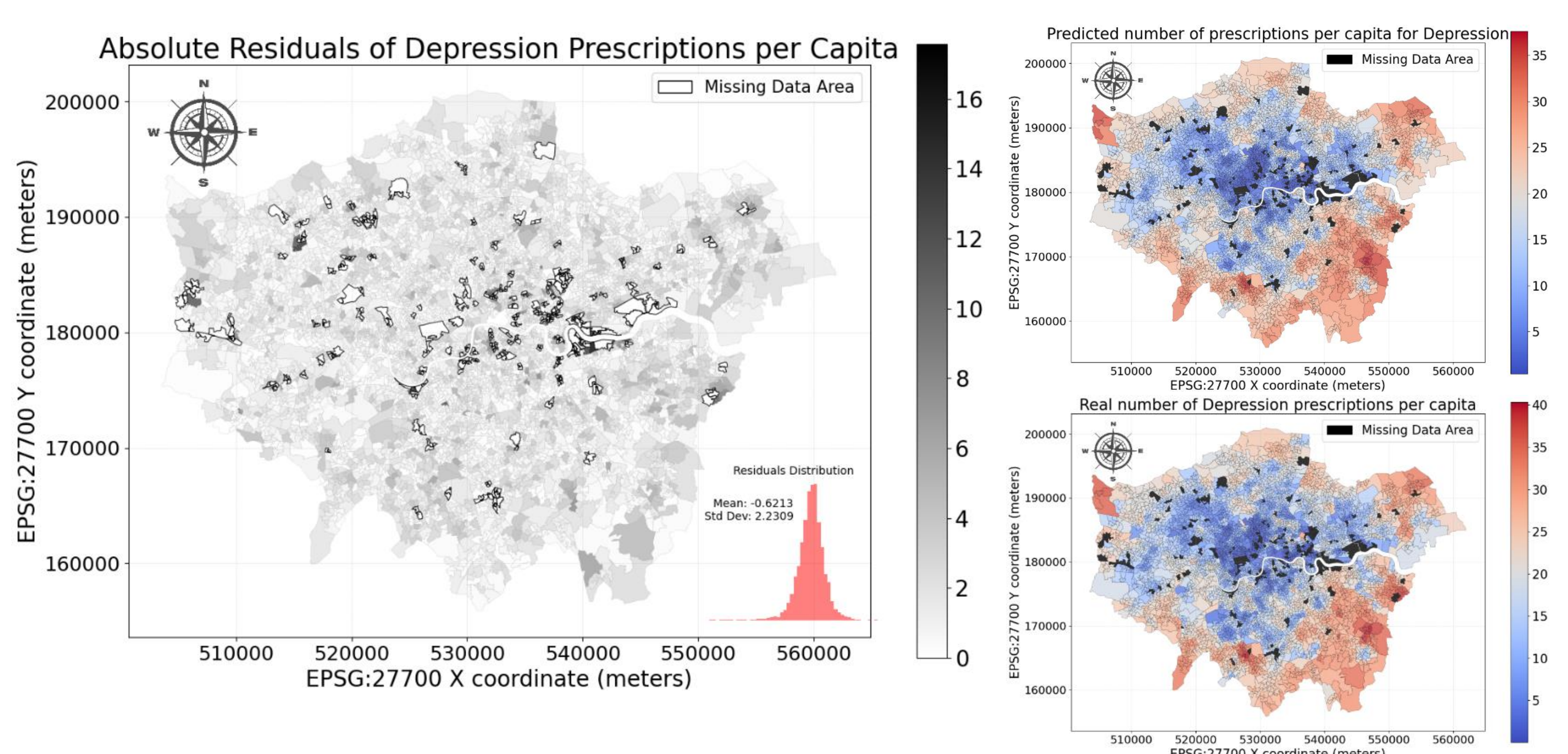


Figure 3 Distribution maps of true value, predicted value and residual of per capita prescription for depression.

- The prediction performance of the model established after feature engineering and ablation experiments is far better than the previous baseline.
- GNN-PCA based explanations demonstrate the correlation of community characteristics with specific prescriptions (Figure 4)

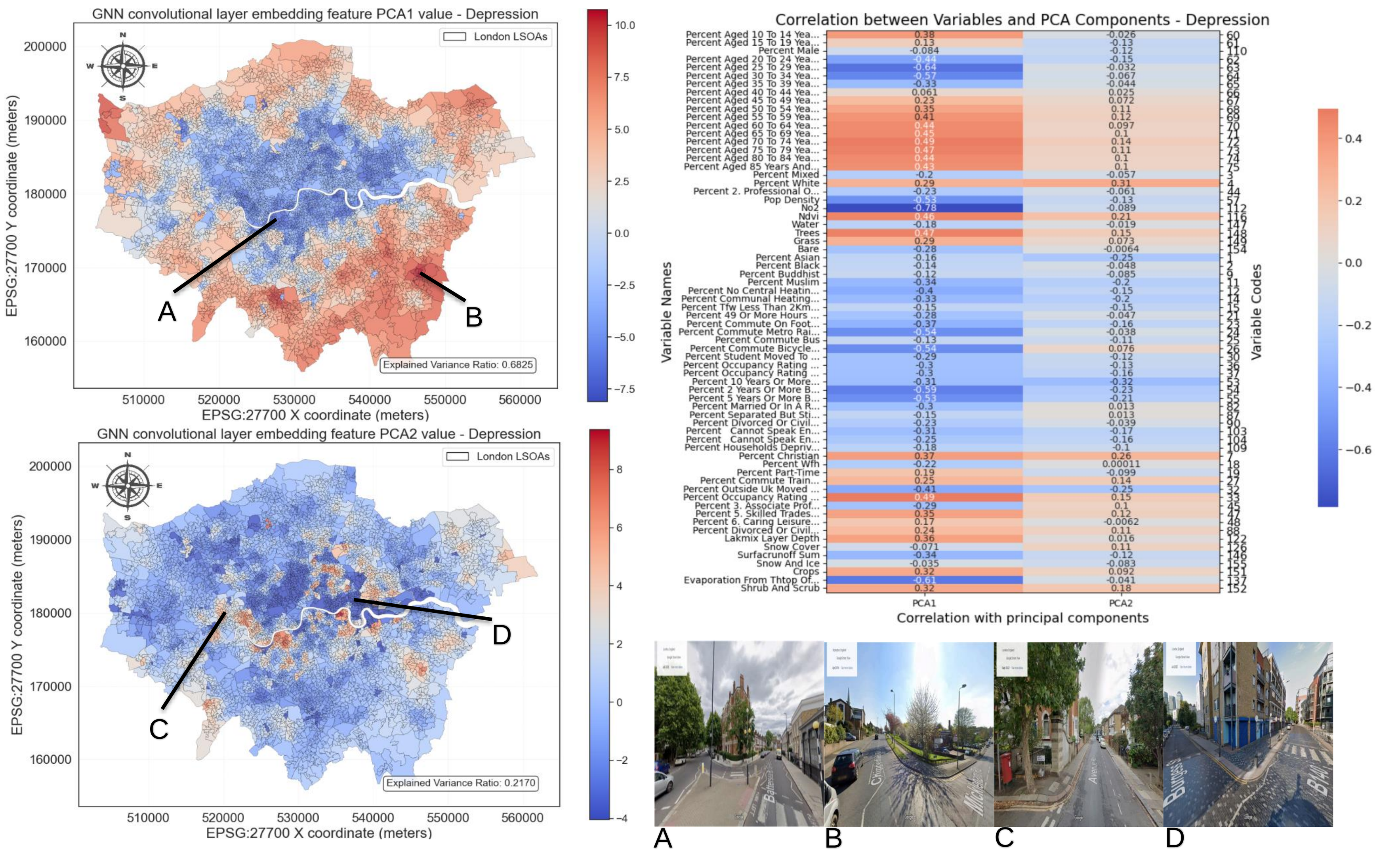


Figure 4 GNN-PCA based explanations for the depression prescription prediction model (the first principal components of the graph embeddings exhibit urbanity.).

Result B: Border Vacuum

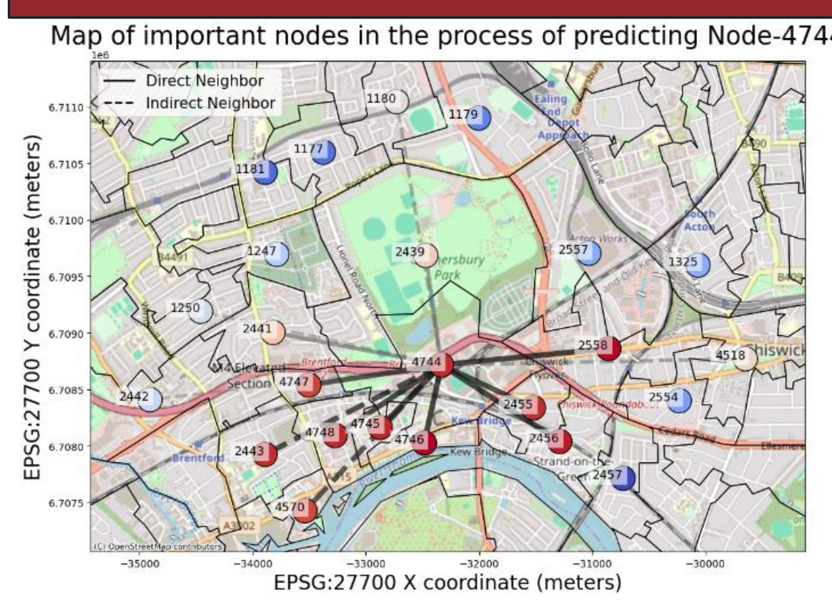
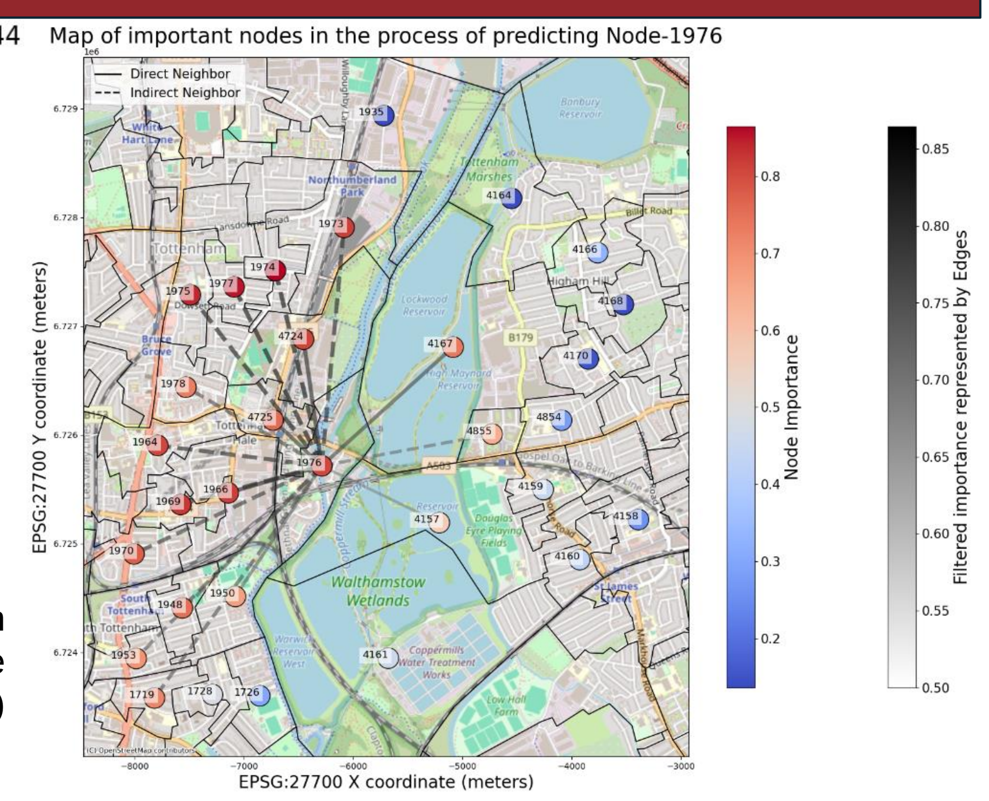


Figure 5 GNNExplainer results embedded in geospatial space, showing local node importance graphs when predicting node 4744 (WAD = 0.5) and 1976 (WAD = 0.55).



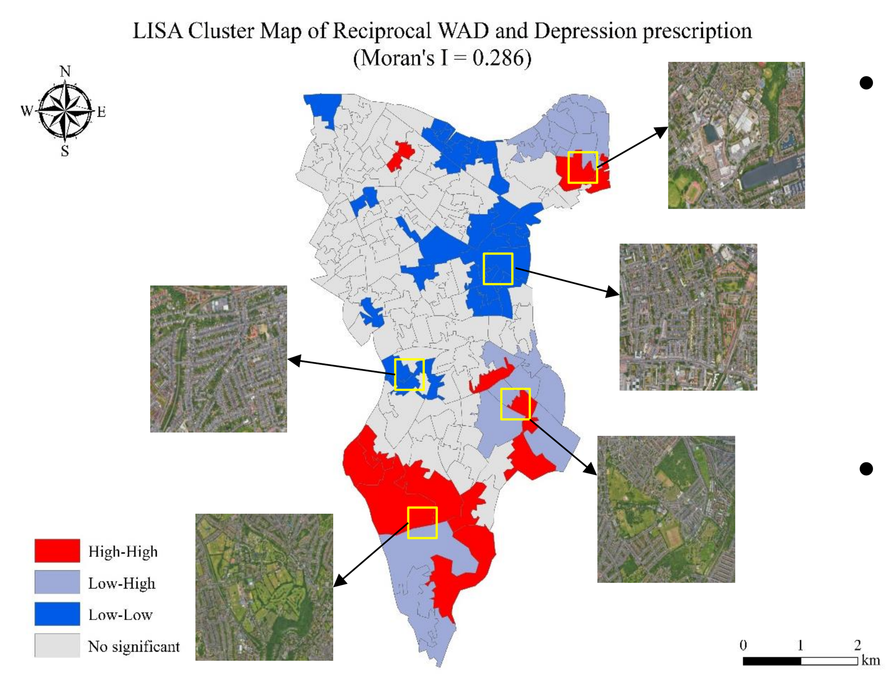


Figure 6 Results of bivariate Moran's Index analysis of depression prescriptions and corresponding WAD in Southwark.

- In Figure 5, edges are built by Contribution to the central node, which more clearly shows the uneven information aggregation during prediction.
- The results in Figures 5 and 6 provide visual evidence of links between health outcomes and border vacuums.

Discussion & Conclusion

- Leveraging geospatial artificial intelligence (GeoAI), this study proposes an interdisciplinary data-driven approach to advancing urban health research.
- Our findings have important implications for urban planning and public health policy, capturing the strong interrelationships between community-level prescription outcomes and urban characteristics.
- highlighting the potential links between border vacuums and health disparities.
- While explanatory methods can offer valuable insights, they must be grounded in reliable evaluation and validation to be effective. Further evaluation of model reliability may be needed.
- Further quantitative research using high-resolution remote sensing imagery could be employed to identify border vacuums more effectively, rather than relying solely on visual evidence.

Key References

Obradovich, N., Migliorini, R., Paulus, M. P., & Rahwan, I. (2018). Empirical evidence of mental health risks posed by climate change. Proceedings of the National Academy of Sciences, 115(43), 10953-10958.
Office for National Statistics (ONS). (n.d.). Statistical geographies.
Scepanovic, S., Obadic, I., Joglekar, S., Giustarini, L., Nattero, C., Quercia, D., & Zhu, X. (2024). MedSat: A Public Health Dataset for England Featuring Medical Prescriptions and Satellite Imagery. Advances in Neural Information Processing Systems, 36.
Sung, H., & Lee, S. (2015). Residential built environment and walking activity: Empirical evidence of Jane Jacobs' urban vitality. Transportation Research Part D: Transport and Environment, 41, 318-329.
Rowan, J. C. (2017). The literary craft of Jane Jacobs. In Reconsidering Jane Jacobs (pp. 43-168). Routledge.