Community-Level Psychological Stress Assessment in Seoul Using Public Urban Data

Chaehee Park
The Seoul Institute
Seoul, Republic of Korea
chaehee@si.re.kr

Abstract

Identifying communities with elevated psychological stress burden is important for public mental health planning, yet the feasibility of reliable area-level stress assessment from routinely available municipal data remains underexplored. We present a district-level regression benchmark using 2024 public urban indicators across Seoul's 25 administrative districts, evaluating whether air quality, green space, and demographic features can support reliable stress prediction. Across 12 model-feature configurations evaluated under leave-one-out cross-validation, Ridge regression on all 12 features achieves the best performance ($R^2 = 0.19$, MAE = 2.49%p), exceeding shuffled-label performance in a permutation test (uncorrected permutation $p = 0.019$). However, this performance remains insufficient for robust district prioritization given an inter-district SD of 3.3%p. Only a few configurations improve over the mean-prediction dummy baseline, and the gains remain modest. Among 12 bivariate predictors, only ground-level ozone shows a nominally significant association ($r = 0.41$, uncorrected $p = 0.04$), which does not survive Bonferroni correction across the 12 features tested. Spatial diagnostics show no significant clustering in either the target variable (Moran's $I = -0.14$, $p = 0.21$) or Ridge residuals (Moran's $I = -0.02$, $p = 0.41$). These results establish a reproducible benchmark and quantify the limited predictive value of current public municipal indicators for community-level stress assessment.

CCS Concepts

• **Applied computing** → **Health informatics**; • **Information systems** → **Geographic information systems**; • **Computing methodologies** → *Machine learning approaches.*

Keywords

community mental health, area-level assessment, urban computing, public health, regression benchmark, environmental health, spatial analysis

ACM Reference Format:
Chaehee Park. 2026. Community-Level Psychological Stress Assessment in Seoul Using Public Urban Data. In *The International Workshop on AI for Cognitive and Mental Health Support (AI4Mental @ KDD '26), August 9, 2026, Jeju, Republic of Korea.* ACM, New York, NY, USA, 5 pages.

1 Introduction

Mental health is influenced not only by individual and clinical factors, but also by the urban environments in which people live [4, 16]. Urban residents are exposed to multiple environmental conditions, including air pollution, heat stress, noise, and unequal access to restorative spaces [4, 15, 16]. These exposures have been associated with stress and depression, and are increasingly understood as interacting environmental profiles rather than isolated risk factors [15, 16]. Since such conditions are unevenly distributed across urban areas, stress vulnerability may also vary substantially across communities.

This question is especially salient in the Asia-Pacific region, where mental, neurological, substance use disorders, and self-harm account for a larger non-fatal disease burden than communicable diseases, and depressive and anxiety disorders alone constitute roughly 60% of the mental-health-specific burden [9]. Seoul provides a natural testbed: it concentrates one-fifth of South Korea's population under pronounced between-district inequality, and publishes one of the more comprehensive municipal open-data infrastructures in Asia, including a routinely updated district-level mental health indicator.

This motivates a community-level assessment question: can publicly available urban indicators support reliable district-level stress assessment? In public-sector settings, identifying areas with elevated vulnerability to stress is important to prioritize mental health services, preventive interventions, and resource allocation [1, 4, 14]. However, prior studies have largely examined links between specific environmental exposures and individual-level mental health outcomes, leaving the feasibility of reproducible area-level assessment underexplored [15, 16].

We investigate this question using public municipal data from Seoul, South Korea. Seoul provides annual district-level stress perception statistics, together with environmental and demographic indicators across 25 districts. We formulate the task as district-level regression and evaluate whether this limited set of public indicators can support reliable area-level prioritization under interpretable modeling settings.

Our empirical results show that the current indicator set provides statistically detectable but practically limited predictive signal: the best-performing model achieves $R^2 = 0.19$ and MAE = 2.49%p (uncorrected permutation $p = 0.019$), but most configurations are at or below the mean baseline. Only ground-level ozone (O_3) shows a nominally significant bivariate association ($r = 0.41$, uncorrected $p = 0.04$); this does not survive Bonferroni correction and is hypothesis-generating rather than confirmatory. District-level stress also shows no significant spatial clustering (Moran's $I = -0.14$, $p = 0.21$), suggesting proximity-based spatial dependence is not the main source of unexplained variation.

This paper contributes (i) a reproducible district-level benchmark for community-level stress assessment using public municipal data, with publicly released dataset and code; (ii) a quantitative feasibility analysis showing that current indicators yield statistically detectable but practically insufficient signal; and (iii) evidence that stress perception is not strongly spatially clustered, with O_3

identified as a candidate environmental factor warranting further investigation, while not surviving multiple-comparison correction.

2 Related Work

2.1 Urban Environment and Mental Health

Prior research has shown that urban conditions such as green space, air pollution, noise, and socioeconomic deprivation are closely associated with stress and depression, and that their effects operate as interacting environmental profiles rather than isolated factors [15, 16]. Studies in Korea further support this relationship: restorative neighborhood environments have been linked to reduced stress among young adults in Seoul [17], and greenspace exposure has been associated with lower depression in older adults [18]. However, these studies have largely focused on identifying environmental factors associated with mental health, rather than on comparable area-level assessment for identifying vulnerable communities. This leaves a gap in reproducible assessment frameworks that can support public-sector decision-making [1, 4, 14].

2.2 ML-Based Community Mental Health Assessment

Machine learning has been widely applied to mental health assessment using individual-level digital traces such as social media text and images, speech/audio signals, smartphone logs, and wearable sensor data [3, 6, 8, 10], and extended to population-scale psychological and health monitoring via geographic aggregation of online signals [5]. Urban computing studies have further explored street-view images and built-environment features as proxies for mental well-being [2, 7, 13]. However, relatively little work has examined whether routinely available municipal indicators alone can support reliable area-level stress assessment, a setting characterized by low-dimensional, spatially aggregated, and administratively constrained data. Our work addresses this gap by constructing a reproducible district-level benchmark and evaluating whether such indicators provide sufficient predictive signal under a leave-one-district-out(LOOCV) validation protocol.

3 Method

3.1 Problem Formulation

We formulate community-level stress assessment as a district-level regression task over Seoul's 25 *gu* districts. This is a feasibility study: we evaluate whether a sparse set of public municipal indicators (air quality, green space, demographics) contains sufficient signal for reliable district-level stress prediction.

3.2 Data and Features

The target variable is the 2024 district-level *stress perception rate* from the Seoul Mental Health Indicator 2024 [12], defined as the proportion of adults aged 19 years or older who report high daily stress. Across districts, stress perception rates range from 15.7% to 29.6% (mean 23.8%, SD 3.3%). The stress perception rate is a single-item self-report measure rather than a validated psychometric instrument such as the PSS or PHQ-9, and reflects a broad subjective construct that may include everyday strain alongside clinically relevant distress. We adopt it because, to our knowledge,

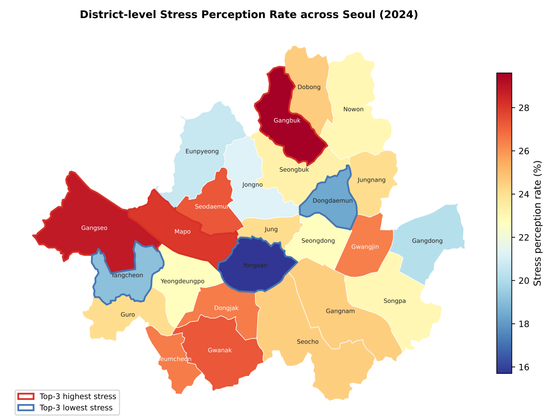

Figure 1: District-level stress perception rates across Seoul's 25 *gu* districts (2024). Red outlines indicate the three highest-stress districts (Gangbuk-gu, Gangseo-gu, Mapo-gu); blue outlines indicate the three lowest-stress districts (Yongsan-gu, Yangcheon-gu, Gangdong-gu).

it is the only mental-health-related outcome routinely published for all 25 *gu* districts of Seoul; as an aggregated survey-derived proportion, it is also subject to sampling uncertainty that imposes a lower bound on achievable predictive precision.

We construct 12 predictors from three categories of publicly available municipal data:

- **Air quality (6):** PM2.5, PM10, O_3, NO_2, SO_2, and CO, measured as annual district averages from the Seoul Open Data Plaza.
- **Green space (3):** park area per capita, urban-park ratio, and walkable-park ratio, derived from the Seoul Metropolitan Government park inventory.
- **Demographics (3):** single-person household ratio, elderly ratio (65+), and average household size, obtained from Seoul district-level household statistics.

3.3 Models

Given the small sample size ($n = 25$), we prioritize interpretable and small-sample-appropriate models over high-capacity architectures. We evaluate 12 model-feature configurations spanning four families: (i) regularized linear models, including Ridge and Lasso regression; (ii) kernel methods, including support vector regression and kernel ridge regression; (iii) probabilistic models, including Bayesian Ridge regression and Gaussian process regression with Matérn and RBF kernels; and (iv) tree-based ensembles and instance-based methods, including Random Forest, Gradient Boosting, and k-nearest neighbors ($k = 3$). A mean-prediction dummy regressor is included as a lower bound. All continuous predictors are standardized within each cross-validation fold to prevent data leakage.

3.4 Evaluation

We use leave-one-out cross-validation (LOOCV), suitable for small spatial datasets, training on 24 districts and evaluating on the held-out one in each fold. We report MAE, RMSE, and R^2. We additionally conduct a label-permutation test ($n = 999$) on the best-performing configuration. Because this configuration is selected post-hoc based on cross-validated performance, the resulting p-value does not correct for selection across the 12 configurations; it should be read as a conservative chance-level check on the selected model, not a corrected test over the full model space.

3.5 Spatial Diagnostics

Because nearby districts may share unobserved contextual factors, we examine spatial autocorrelation in both the target variable and model residuals. We compute Global Moran's I using a k-nearest-neighbor weight matrix ($k = 4$) constructed from district centroid coordinates, implemented with `libpysal` and `esda` [11]. Significant spatial autocorrelation would suggest that spatial dependence remains unmodeled and that spatial regression models such as spatial lag or spatial error models may be warranted. Non-significant results would indicate that there is limited evidence of residual spatial structure in this setting.

4 Results

4.1 Feature-level Analysis

Table 1 reports Pearson correlations between each predictor and the district-level stress perception rate. Of the 12 candidate features, only O_3 shows a nominally significant association at the uncorrected 0.05 level ($r = 0.41$, $p = 0.04$). No feature survives Bonferroni correction for the 12 comparisons ($\alpha = 0.0042$), indicating that individual public municipal indicators carry minimal marginal signal for district-level stress prediction.

Table 1: Bivariate Pearson correlations with stress perception rate ($n = 25$). † nominally significant at uncorrected $p < 0.05$; does not survive Bonferroni correction ($\alpha = 0.0042$ for 12 comparisons).

Feature	r	p
O_3	+0.41	0.04†
Single-person HH ratio	+0.33	0.10
Avg. household size	−0.25	0.24
NO_2	−0.24	0.24
PM2.5	−0.23	0.27
PM10	−0.20	0.34
CO	−0.16	0.43
Park area per capita	+0.12	0.56
Urban-park ratio	+0.10	0.63
Elderly ratio (65+)	+0.08	0.71
Walkable-park ratio	−0.07	0.74
SO_2	+0.01	0.95

4.2 Predictive Performance

Table 2 summarizes LOOCV performance across 12 model-feature configurations and a mean-prediction dummy baseline, sorted by MAE. The best-performing configuration is Ridge regression trained on all 12 features (Ridge-all), achieving MAE = 2.49%p, RMSE = 2.99%p, and $R^2 = 0.19$. Ridge-all improves over the LOOCV dummy baseline (MAE = 2.70%p, RMSE = 3.45%p, $R^2 = -0.09$), although the improvement is modest in absolute terms. Lasso-sel yields a near-zero positive R^2 (0.03), while most remaining configurations perform at or below the dummy baseline. Feature selection via SelectKBest ($k = 6$) does not improve performance relative to using all features, and non-linear, ensemble, and instance-based methods generally underperform regularized linear models under LOOCV.

Table 2: LOOCV performance across 12 model-feature configurations, a univariate O_3 baseline, and a mean-prediction dummy baseline ($n = 25$), sorted by MAE. "all" = all 12 features; "sel" = SelectKBest top-6; "base" = original 4 features.

Model	Features	MAE	RMSE	R^2
Ridge	all	2.49	2.99	+0.19
Lasso	sel	2.55	3.27	+0.03
RF	base	2.60	3.76	−0.28
O3-only (LinReg)	1	2.66	3.30	+0.01
Dummy	mean	2.70	3.45	−0.09
BayesianRidge	sel	2.70	3.33	−0.01
GPR-Matérn	sel	2.71	3.45	−0.09
GPR-RBF	sel	2.71	3.45	−0.09
SVR	sel	2.76	3.66	−0.22
RF	sel	2.78	3.61	−0.18
Ridge	base	2.97	3.73	−0.27
KNN-3	sel	3.04	3.96	−0.43
GradBoost	sel	3.09	3.86	−0.36
KernelRidge	sel	5.21	7.09	−3.57

The district-level residuals of Ridge-all exhibit a wide spread (Figure 2), with the largest errors observed in Gangseo-gu (+6.01%p), Gangbuk-gu (+5.23%p), and Dongdaemun-gu (−4.61%p). The magnitude of these residuals relative to the overall range of stress rates (15.7%–29.6%, SD = 3.3%) suggests that meaningful district-specific variance remains unexplained by the current indicator set.

To probe whether the apparent signal arises from a single dominant feature, we evaluate a univariate baseline using only O_3—the feature with the strongest bivariate correlation—as a single LinReg predictor under the same LOOCV protocol (Table 2). This baseline achieves $R^2 = 0.01$ and MAE = 2.66%p, statistically indistinguishable from the mean-prediction baseline despite O_3's nominally significant in-sample correlation. In contrast, Ridge-all achieves $R^2 = 0.19$ ($\Delta R^2 = 0.18$), indicating that out-of-sample performance derives from the joint contribution of multiple weak features rather than any single dominant predictor.

To assess whether the observed performance exceeds chance-level associations, we conduct a label-permutation test with $n = 999$ shuffled targets on Ridge-all. The observed $R^2 = 0.19$ lies above the 95th percentile of the null R^2 distribution (−0.04) and yields a permutation p-value of 0.019. Because Ridge-all is selected post-hoc as the best configuration among the 12 evaluated, this p-value does not account for the implicit model search; the reported value should therefore be interpreted as a chance-level check on the selected configuration rather than a corrected significance test over the model space. With this caveat, the result suggests that the current indicator set carries statistically detectable predictive signal, albeit

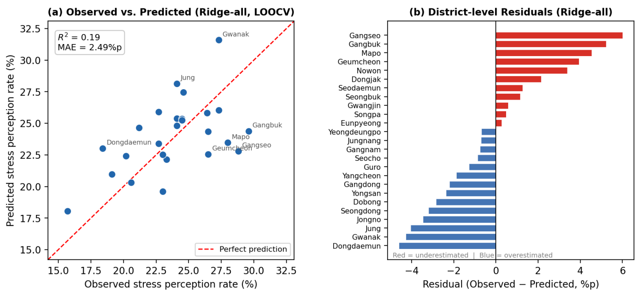

Figure 2: LOOCV results for Ridge-all (12 features). (a) Observed vs. predicted stress perception rates across 25 Seoul districts. (b) District-level residuals sorted by magnitude; red bars indicate under-prediction (observed > predicted), blue bars indicate over-prediction. Labeled districts in (a) have residuals exceeding 3.5%p.

modest in absolute terms (MAE = 2.49%p against an inter-district SD of 3.3%p).

4.3 Spatial Diagnostics

Global Moran's I computed with a KNN weight matrix ($k = 4$) yields non-significant results for all variables of interest (Table 3). Residuals from Ridge-all show $I = -0.02$ ($p = 0.41$), providing no evidence of residual spatial structure. These results indicate that spatial regression models are not warranted in this setting.

Table 3: Global Moran's I results (KNN $k = 4$, permutation $n = 9{,}999$).

Variable	I	z	p
Stress perception rate	−0.136	−0.810	0.214
O_3	−0.119	−0.662	0.273
Ridge-all residuals	−0.023	+0.164	0.409

Local Moran's I (LISA) identifies one uncorrected local outlier pattern: Jung-gu exhibits a high-low (HL) configuration ($I_i = -0.096$, $p = 0.004$, uncorrected), indicating that its stress perception rate is above average while neighboring districts are below average. This pattern does not survive Bonferroni correction ($\alpha = 0.05/25 = 0.002$), and no high-high (HH) or low-low (LL) clusters are detected, consistent with the absence of global spatial autocorrelation.

5 Discussion

5.1 Weak but Detectable Predictive Signal

The results suggest that public municipal indicators contain a weak but detectable signal for district-level stress assessment. Ridge-all achieves the best performance ($R^2 = 0.19$, MAE = 2.49%p), outperforming the mean dummy baseline; the permutation test indicates that this is unlikely to arise by chance ($p = 0.019$), with the caveat in Section 4.2 that this does not formally correct for model selection. However, the MAE remains close to the inter-district SD of 3.3%p, limiting usefulness for fine-grained district prioritization.

5.2 Ozone as a Candidate Environmental Factor

O_3 is the only feature with a nominally significant bivariate association with stress perception ($r = 0.41$, uncorrected $p = 0.04$). This does not survive Bonferroni correction; combined with the cross-sectional ecological design, the result is hypothesis-generating rather than confirmatory. Moreover, O_3 alone does not generalize out-of-sample ($R^2 = 0.01$; Section 4.2), so ozone should be treated as one component of a multivariate exposure profile rather than a stand-alone marker.

5.3 Limits of Static District-Level Indicators

Most green space, demographic, and air quality variables show weak marginal associations. A likely reason is aggregation: annual district-level averages obscure finer variation in exposure, mobility, income, social isolation, healthcare access, and neighborhood quality. The poor performance of non-linear, ensemble, and instance-based models suggests that the main bottleneck is not model capacity, but the limited information content of the available indicators.

A more consequential limitation is the absence of direct SES measures (e.g., median household income, public-assistance enrollment), which are well-established structural determinants of area-level mental health; the household-composition variables here serve only as partial proxies. Observed environmental associations, including the O_3 signal, may therefore partially reflect SES-related confounding, and the apparent predictive ceiling may be derived from this missing axis of variation. Incorporating SES is a primary direction of extension.

6 Conclusion

We introduced a reproducible district-level benchmark for community-level stress assessment using 2024 public urban data across Seoul's 25 districts. Ridge regression on all 12 features achieved the best performance ($R^2 = 0.19$, MAE = 2.49%p, uncorrected permutation $p = 0.019$), but most configurations performed at or below the mean-prediction baseline. Two findings stand out: (i) predictive performance arises from the joint contribution of multiple weak features rather than any single dominant predictor, as a univariate O_3 baseline matches the mean-prediction floor; and (ii) O_3 is the only feature with a nominally significant bivariate association ($r = 0.41$), although this does not survive Bonferroni correction. Stress perception is not significantly spatially clustered (Moran's $I = -0.14$). Future work should extend this benchmark with richer predictors, SES indicators, longitudinal data, and broader mental health outcomes.

References

[1] Somayyeh Azimi and Nasir Uddin. 2025. Workforce Allocation in Urban Community Mental Health Services: GIS-Based Analytical Insights for Policy and Planning. In *Healthcare*, Vol. 13. MDPI, 2092.

[2] Chongxian Chen, Haiwei Li, Weijing Luo, Jiehang Xie, Jing Yao, Longfeng Wu, and Yu Xia. 2022. Predicting the effect of street environment on residents' mood states in large urban areas using machine learning and street view images. *Science of the Total Environment* 816 (2022), 151605.

[3] Prerna Chikersal, Afsaneh Doryab, Michael Tumminia, Daniella K Villalba, Janine M Dutcher, Xinwen Liu, Sheldon Cohen, Kasey G Creswell, Jennifer Mankoff, J David Creswell, et al. 2021. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: a machine learning approach

with robust feature selection. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 1 (2021), 1–41.

[4] Pamela Y Collins, Moitreyee Sinha, Tessa Concepcion, George Patton, Thaisa Way, Layla McCay, Augustina Mensa-Kwao, Helen Herrman, Evelyne de Leeuw, Nalini Anand, et al. 2024. Making cities mental health friendly for adolescents and young adults. *Nature* 627, 8002 (2024), 137–148.

[5] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social Media as a Measurement Tool of Depression in Populations. In *Proceedings of the 5th Annual ACM Web Science Conference*. 47–56. doi:10.1145/2464464.2464480

[6] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, Vol. 7. 128–137.

[7] Xin Han, Lei Wang, Seong Hyeok Seo, Jie He, and Taeyeol Jung. 2022. Measuring perceived psychological stress in urban built environments using google street view and deep learning. *Frontiers in public health* 10 (2022), 891736.

[8] Lin Sze Khoo, Mei Kuan Lim, Chun Yong Chong, and Roisin McNaney. 2024. Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches. *Sensors* 24, 2 (2024), 348.

[9] OECD and World Health Organization. 2024. Health at a Glance: Asia/Pacific 2024. doi:10.1787/51fed7e9-en

[10] Andrew G Reece and Christopher M Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Science* 6, 1 (2017), 15.

[11] Sergio J Rey and Luc Anselin. 2009. PySAL: A Python library of spatial analytical methods. In *Handbook of applied spatial analysis: Software tools, methods and applications*. Springer, 175–193.

[12] Seoul Mental Health Welfare Center. 2024. Seoul Mental Health Indicator 2024. Report. Seoul Mental Health Statistics Database.

[13] Musab Wedyan and Fatemeh Saeidi-Rizi. 2024. Assessing the impact of urban environments on mental health and perception using deep learning: A review and text mining analysis. *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 101, 2 (2024), 327.

[14] Pui-Hang Wong, Karima Kourtit, and Peter Nijkamp. 2025. A Meta-Analysis of Neighbourhood Interventions on Subjective Wellbeing and Mental Health. *Applied Research in Quality of Life* (2025), 1–37.

[15] Zheng Xian, Tomoki Nakaya, Kun Liu, Bing Zhao, Junhua Zhang, Jiao Zhang, Yuxuan Lin, and Jinguang Zhang. 2024. The effects of neighbourhood green spaces on mental health of disadvantaged groups: a systematic review. *Humanities and social sciences communications* 11, 1 (2024), 1–19.

[16] Jiayuan Xu, Nana Liu, Elli Polemiti, Liliana Garcia-Mondragon, Jie Tang, Xiaoxuan Liu, Tristram Lett, Le Yu, Markus M Nöthen, Jianfeng Feng, et al. 2023. Effects of urban living environments on mental health in adults. *Nature medicine* 29, 6 (2023), 1456–1467.

[17] Da-Hye Yim and Youngsang Kwon. 2025. How residential environments affect depressive mood and stress among young adults in South Korea: The role of restorativeness. *Cities* 166 (2025), 106165.

[18] Eun-Hye Yoo, Jin-Young Min, Baek-Yong Choi, Seung-Woo Ryoo, Kyoung-Bok Min, and John E Roberts. 2024. Spatiotemporal variability of the association between greenspace exposure and depression in older adults in South Korea. *BMC public health* 24, 1 (2024), 2556.