Environmental Science & Technology

pubs.acs.org/est Article

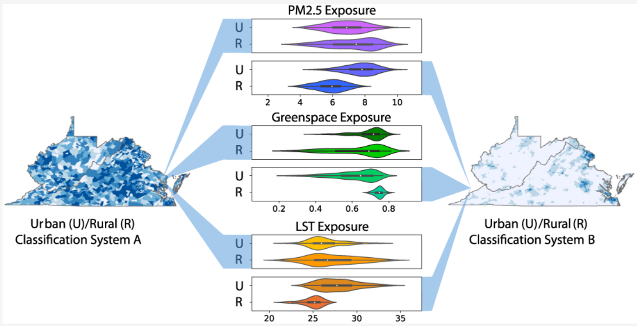
# Delineating Urbanicity and Rurality: Impact on Environmental Exposure Assessment

Yimeng Song,* Nicole C. Deziel, and Michelle L. Bell

Cite This: *Environ. Sci. Technol.* 2024, 58, 19178−19188 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Environmental exposures and their health impacts can vary substantially between urban and rural areas. However, different methods for classifying these areas could lead to inconsistencies in environmental exposure and health studies, which are often overlooked. We constructed different urban/rural classification systems based on multiple population-based (e.g., total population, population density, and commuting) and built-environment-based (e.g., nighttime light intensity, building density, road density, distance to urban centers, point of interest density, and urban area coverage) indicators and various classification schemes. These classification systems were applied to Virginia and West Virginia, United States. We compared differences in urban/rural spatial patterns, demographic compositions, and exposures of particulate matter ($PM_{2.5}$), greenspace, and land surface temperature using these urban/rural classification systems to understand their impacts on environmental exposure and health research. Our findings reveal clear differences in spatial patterns and demographic compositions across various systems. We also observed that different systems can lead to changes in the magnitude and direction of urban/rural disparities in environmental exposure assessment. Addressing the complexities in delineating urbanicity and rurality may include careful consideration of classification systems to reflect those aspects of urbanicity and rurality that are relevant to the research question or the use of multiple, complementary systems.



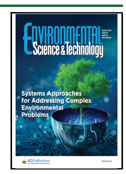

**KEYWORDS:** *urban/rural metrics, urban−rural classification, urban−rural disparities, environmental exposure*

## 1. INTRODUCTION

Urban areas are typically defined by high population density and substantial infrastructure development, including buildings, roads, and other built environments.[1,2] Most environmental health studies have focused on urban areas; however, populations, exposures, and relationships between exposures and health outcomes can differ between urban and rural areas. According to the World Bank,[3] approximately 44% of the global population, or 3.5 billion people, live in rural areas. Yet, due to rapid urbanization, urban areas are expected to house nearly 70% of the global population by 2050.[4] Despite these demographic shifts, large populations in both rural and urban areas continue to face unique environmental and social challenges, potentially leading to distinct health issues. For instance, urban areas can face issues such as overcrowding, traffic pollution, and urban heat island effect.[5] Rural areas often have comparatively limited infrastructure, such as healthcare facilities, grocery stores, and schools, and rely more heavily on agriculture or natural resources, which can introduce pollutants from pesticides or fossil fuel extraction.[5,6] Moreover, rural communities typically have higher proportions of the elderly and children, higher unemployment rates, and inadequate insurance coverage.[7]

Environment and health analyses focusing on urban or rural areas require the classification of areas by the level of urbanicity/rurality. Constructing urban/rural classification systems typically involves two fundamental steps: (1) developing indicators that represent the urban/rural gradient and (2) applying a classification scheme to assign urban/rural levels. Government agencies and researchers often use different indicators and classification schemes depending on the objectives, geography, and data sets,[8] and there is no standard approach. Indicators representing the urban/rural gradient vary, including the population size and density,[8,9] infrastructure density, economic activities, nighttime light intensities, connectivity, and combinations of multiple indicators.[10−12] Even when the same indicators are used, different classification schemes can lead to diverse urban and rural classification systems. In the United States (U.S.), there are more than 15 federal definitions of

Received: July 8, 2024
Revised: October 5, 2024
Accepted: October 7, 2024
Published: October 16, 2024









https://doi.org/10.1021/acs.est.4c06942
*Environ. Sci. Technol.* 2024, 58, 19178−19188

"rural,"[13] leading to varying urban/rural spatial patterns and demographic estimates. For example, the U.S. Department of Agriculture reported 46 million rural Americans (14%) in 2015, compared to the Census Bureau's estimate of nearly 60 million (18%), marking a significant 30% discrepancy.[6]

A growing body of research has focused on identifying environmental and health issues in urban or rural areas as well as the disparities between them. These studies cover topics such as the impacts of air pollution and climate change,[14,15] urbanization and health inequities,[16] health risks from agriculture,[17,18] and lifestyle and health.[19,20] Most of these studies use a single system to classify urban/rural areas, which vary from study to study. A scoping review of urban/rural health services found that among 103 included studies, all from the U.S., 11 different systems were used to assess the urban/rural status,[6] and these variations resulted in differences in rural health accessibility assessments.[21] However, the impact of using different urban/rural classification systems on environmental exposure and health research results has rarely been examined and discussed.

The primary goal of this paper is to understand the differences between various urban and rural classification systems and their impact on environmental exposure and health studies. To achieve this, we pursued three specific objectives: (1) we considered a range of urban/rural classification systems and applied them across the states of Virginia and West Virginia, U.S.; (2) we compared the spatial patterns and demographic compositions resulting from these classification systems; and (3) we evaluated the urban−rural disparities in key environmental exposures using different classification systems.

## 2. METHODOLOGY

**2.1. Study Overview.** In this study, we constructed urban/rural classification systems in two steps: (1) we developed nine urban/rural indicators (Table 1) at the census tract level using diverse geospatial data representing various socio-economic and physical features and (2) we classified all census tracts into five urban/rural levels using three different classification schemes for each indicator. We then compared the spatial patterns and demographic compositions of the constructed urban and rural classification systems to assess the impact of these indicators and classification schemes. Finally, we assessed urban−rural disparities in exposure to fine particulate matter ($PM_{2.5}$), greenspace, and land surface temperature (LST), illustrating the impact of urban/rural classification systems on environmental exposure assessments. These factors were selected to exemplify potential exposure to chemical, built environment, and climate-related factors, each of which may be differentially influenced by urbanity and rurality.

**Table 1. Urban/Rural Indicators**

| Indicator | Data sets and source | Format of original data | Year of original data sets |
|---|---|---|---|
| population density (PopDen) | census data at census tract level | polygon | 2023 |
| population size (TotalPop) | census data at census tract level | polygon | 2023 |
| rural−urban community areas codes (RUCA) | RUCA from the United States Department of Agriculture | polygon | 2010 |
| nighttime light intensity (NTL) | NPP-VIIRS nighttime light data | grid | 2020 |
| point of interest density (POIDen) | POI from OpenStreetMap | point | 2023 |
| urban area coverage (UrbanC) | urban area from census data | polygon | 2020 |
| building density (BldgDen) | building footprint from OpenStreetMap | polygon | 2023 |
| road density (RoadDen) | road network from OpenStreetMap | polyline | 2023 |
| distance to urban center (DisUrban) | urban area from census data | polygon | 2020 |

**2.2. Study Area.** Our study focuses on Virginia and West Virginia in the Southern U.S. (Figure 1a). We chose these states due to their large and diverse populations in both urban and rural settings. According to the 2020 Census, Virginia's urbanization rate is 75.6%, while West Virginia's urbanization rate is 44.6%.[1] This contrast underscores the differing urban and rural landscapes of the two states, facilitating a comparison across varied contexts. Our analysis was conducted at the census tract level with an average area of 70.1 $km^2$ (SD = 128.3 $km^2$).

**2.3. Urban/Rural Indicator Construction.** We selected and constructed the nine indicators (Table 1) based on their established use in prior studies and their ability to capture diverse dimensions of urbanicity and rurality, including population density (PopDen), population size (TotalPop), rural−urban community area codes (RUCA), nighttime light intensity (NTL), point of interest density (POIDen), urban area coverage (UrbanC), building density (BldgDen), road density (RoadDen), and distance to urban center (DisUrban). For each indicator, we considered the most recent data available.

*2.3.1. PopDen.* PopDen captures the concentration of inhabitants within an area and is a widely used indicator for delineating urban/rural levels by scholars and federal agencies, such as the U.S. Census Bureau and the U.S. Department of Agriculture.[8] Low-density regions typically exhibit rural characteristics, such as agricultural land and sparse housing, while high-density areas often feature urban attributes, including dense buildings and more comprehensive public services. We quantified the PopDen as the number of individuals per square kilometer based on census data,[22] as shown in Figure 1b.

*2.3.2. TotalPop.* TotalPop reflects the number of inhabitants within a region, greatly influencing infrastructure needs, resource allocation, and energy consumption. It serves as a fundamental indicator for distinguishing urban/rural settings.[9] For example, the 2020 U.S. Census uses population size to classify census tracts into urbanized areas (>50,000 people), urban clusters (2,500−50,000), and rural areas (<2,500).[1] We selected the TotalPop at the census tract level (Figure 1c) as an indicator, with lower TotalPop values generally indicating more rural areas.

*2.3.3. RUCA.* RUCA is an urban/rural classification scheme developed by the U.S. Department of Agriculture (USDA) that accounts for population density, urbanization, and daily commuting patterns. RUCA codes are categorized into 10 primary and 30 secondary levels, providing a more nuanced classification than traditional binary systems. We obtained RUCA data (Figure 1d) from the USDA's Economic Research Service. We used the primary RUCA categories as an indicator, with values ranging from 1 to 10, where higher values indicate greater rurality.

*2.3.4. NTL.* NTL offers a unique perspective on human activities by observing light use patterns associated with urbanization, economic activity, and population density.[2] Lower nighttime light intensity typically indicates rural areas characterized by lower energy consumption and less infrastructural development. We used NPP-VIIRS NTL images in

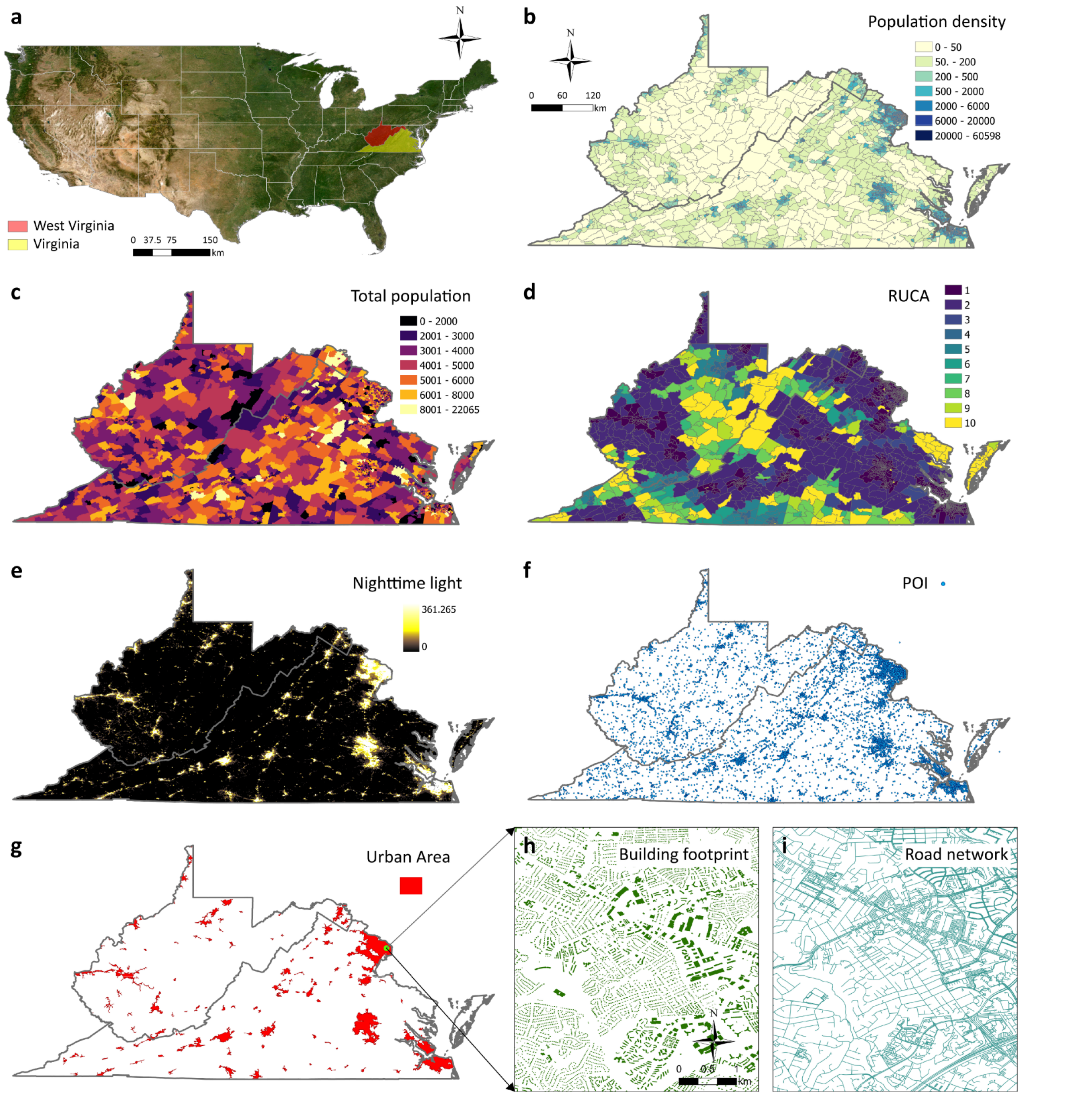



**Figure 1.** Study area and geospatial data used for urban/rural indicator development: (a) geographic location of Virginia and West Virginia, U.S.; (b) population density (people/km$^2$); (c) total population; (d) rural−urban commuting area (RUCA) codes at the census-tract level; (e) nighttime light intensity (nW/cm$^2$/sr); (f) points of interest (POIs); (g) urban area; (h) building footprints; and (i) road network.

this study, which capture Earth's surface light at a fine spatial resolution of 15 arc-seconds.[23] We collected annual cloud-free composites of NPP-VIIRS NTL data (Figure 1e), which eliminate external interference, such as stray light and lunar illumination. Data collection and processing were conducted using the Google Earth Engine (GEE) platform. We quantified the NTL indicator by averaging the intensity values (nW/cm$^2$/sr) of all grid cells within each census tract.

*2.3.5. POIDen.* POIDen reflects the density of points of interest (POIs). POIs refer to specific locations or places, such as grocery stores, schools, parks, and hospitals, that are of significance or interest to people (Figure 1f). The spatial distribution of POIs often correlates with the levels of urbanization.[24] Areas with a high density of POIs typically have more diverse socio-economic activities, indicating urban centers, while areas with a lower POIs density suggest a more rural setting.[25] The POIs data used in this study come from OpenStreetMap (OSM), a collaborative mapping project that allows individuals to contribute, edit, and share geospatial data. We collected the data via the OSM download server of Geofabrik[26] and quantified POIDen as the number of POIs per km$^2$.

*2.3.6. UrbanC.* UrbanC measures the proportion of a region encompassed by urban areas, providing insights into urbanization levels and infrastructural development based on land use patterns.[27] We collected the 2020 urban area layer from the U.S. Census Bureau website[28] (Figure 1g) and quantified UrbanC as the urban area coverage of each tract. An UrbanC value of 100% indicates complete urbanization of a tract, while a value of 0% indicates that the area is entirely rural.


https://doi.org/10.1021/acs.est.4c06942
Environ. Sci. Technol. 2024, 58, 19178−19188

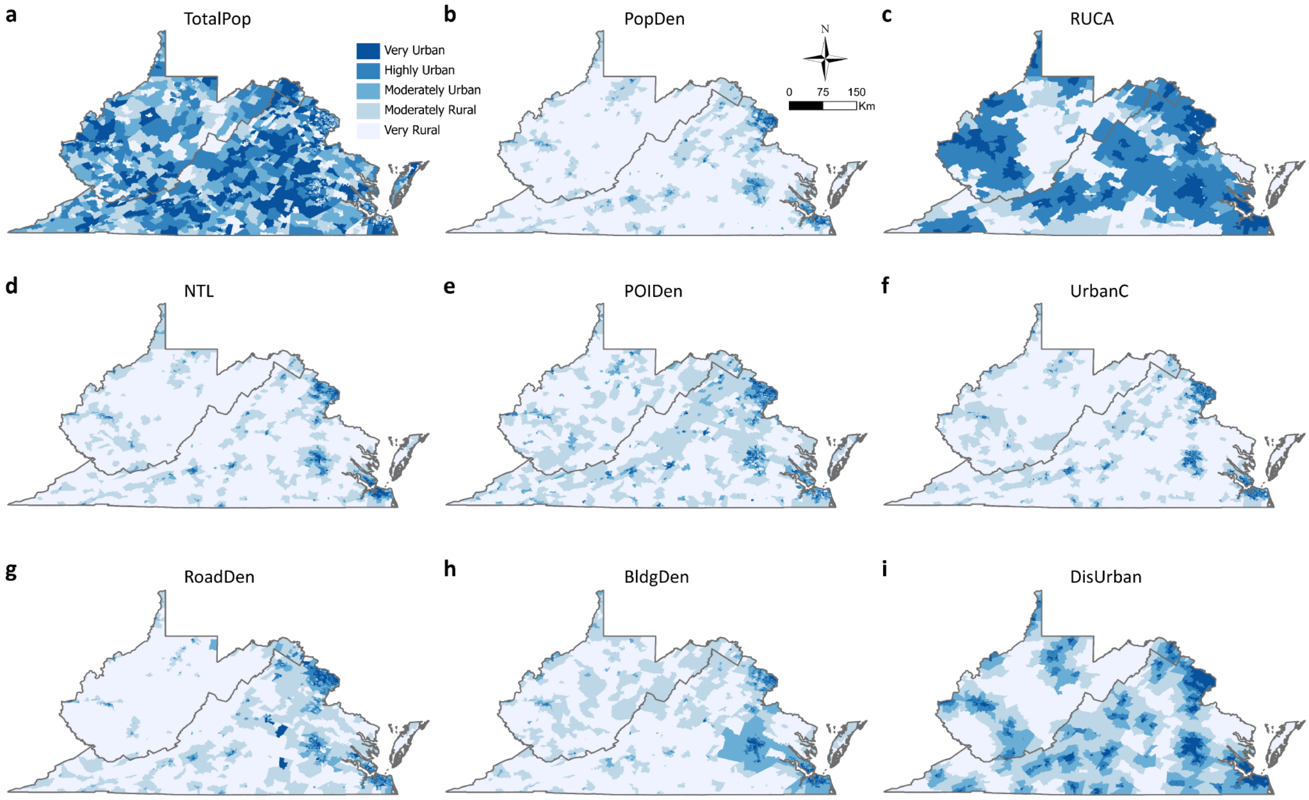



**Figure 2.** Spatial patterns of nine urban/rural classification systems based on different indicators using the global quantile.

*2.3.7. BldgDen.* BldgDen represents the percentage of a region covered by building footprints, which are the ground outlines of buildings' physical boundaries. Building density is crucial for distinguishing urban from rural areas, understanding land use patterns, and measuring the intensity of human activities.[11] We collected the building footprints from OSM (Figure 1h) via Geofabrik[26] and quantified BldgDen by calculating the proportion of each census tract covered by building footprints. Lower BldgDen values indicate more rural areas.

*2.3.8. RoadDen.* RoadDen serves as an indicator of an area's transportation infrastructure, with higher values typically associated with urban settings due to more extensive transportation needs.[29] We used the OSM road network layer (Figure 1i) accessed through Geofabrik[26] to calculate RoadDen. The OSM road network includes various road types, from highways and main thoroughfares to local paths and trails, each represented as a vector polyline. We quantified RoadDen for each census tract by summing the total length of all road types and dividing by the area of the tract.

*2.3.9. DisUrban.* DisUrban measures the proximity of an area to the nearest urban center. Areas closer to urban centers typically have greater access to amenities and services, while regions farther away tend to be more isolated and rural.[30] We quantified DisUrban by calculating the shortest Euclidean distance from the geometric center of each census tract to the boundary of the urban area based on census data. A greater distance indicates a stronger rural character of the tract.

**2.4. Urban/Rural Classification Schemes.** We primarily used quantile classification with a global threshold strategy for the analysis in this paper. Quantile classification ensures an even distribution by creating classes with equal census tract counts. The global threshold was determined using indicator values from census tracts across the entire study area of Virginia and West Virginia. This strategy is widely used in international or nationwide urban/rural classification systems, such as those by federal agencies (e.g., U.S. Census Bureau, U.S. Department of Agriculture).

To investigate the impact of classification schemes on urban/rural classification, we also used two alternative schemes for comparison. One alternative is quantile classification with local thresholds, where thresholds were set separately for Virginia and West Virginia. This approach reflects the spatial heterogeneity of socio-economic or physical characteristics in different regions and mirrors classification systems formulated by local agencies. We used this scheme to explore how threshold strategies can impact urban/rural classification.

The other alternative scheme is the natural break classification with a global threshold. Natural breaks classification groups similar values together, emphasizing disparities between classes.[31] We used this scheme to examine how the classification method can impact urban/rural classification.

In subsequent sections, we refer to "*global quantile*" and "*local quantile*" to denote quantile classification with global and local thresholds, respectively, and "*global natural breaks*" to denote natural break classification with a global threshold. These classification schemes were individually applied to each indicator and used to categorize census tracts into five levels: two rural levels — (1) "very rural" and (2) "moderately rural" — and three urban levels — (3) "moderately urban," (4) "highly urban," and (5) "very urban." The thresholds for each classification scheme are provided in Table S1.

**2.5. Statistical Analysis.** We constructed Pearson correlation coefficient ($r$) matrices and normalized root-mean-square deviation (NRMSD) matrices across urban/rural classification systems to explore their inter-relationships and similarities in



https://doi.org/10.1021/acs.est.4c06942
*Environ. Sci. Technol.* 2024, 58, 19178−19188

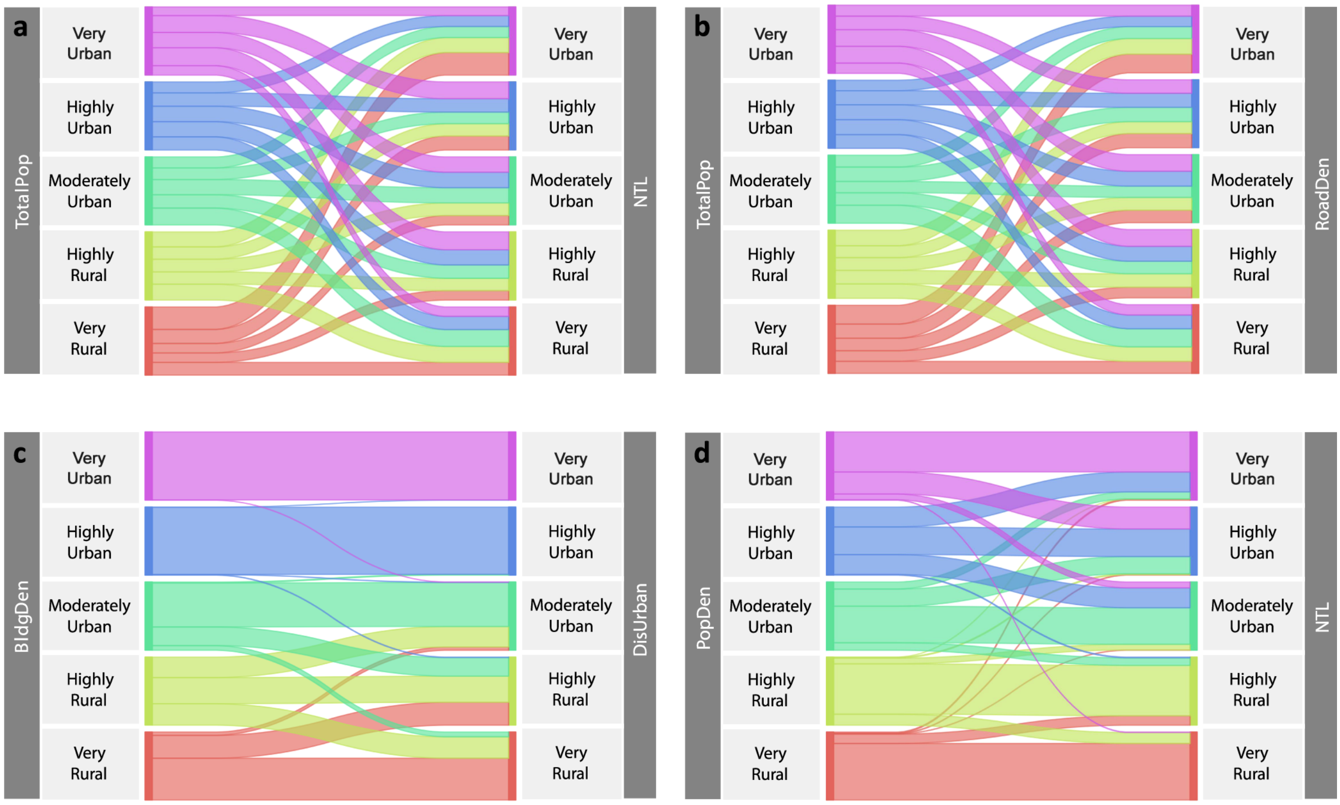



**Figure 3.** Shifts in urban/rural levels of census tracts between two indicator-based urban/rural classification systems using the global quantile: (a) TotalPop vs NTL, (b) TotalPop vs RoadDen, (c) BldgDen vs DisUrban, and (d) PopDen vs NTL.

spatial patterns. A higher $r$ and a lower NRMSD indicate greater similarity between the classification systems. We assigned values 1 to 5 to represent the five levels from "very rural" to "very urban" and calculated these two indices through eqs 1 and 2:

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \quad (1)$$

$$\text{NRMSD} = \frac{\sqrt{\frac{1}{n} \sum_{n=1}^{n} (X_i - Y_i)^2}}{X_{\max}} \quad (2)$$

where $X_i$ and $Y_i$ represent the urban/rural levels (from 1 to 5) of the two classifications systems being compared and $\overline{X}$ and $\overline{Y}$ are the mean values of the classifications.

We calculated the demographic composition as the population percentage belonging to each urban or rural level, although we recognize that this does not fully describe demographics. We also calculated the urbanization rate by aggregating the population percentages belonging to the urban categories of "moderately urban," "highly urban," and "very urban."

**2.6. Environmental Exposure Assessment.** We measured urban–rural disparities in three different environmental exposures, including $PM_{2.5}$, greenspace, and LST, to illustrate the impact of various urban/rural classification systems on disparity assessment. Satellite-derived $PM_{2.5}$ concentration data from 2000 to 2016, developed by Di et al. (2019),[32] with a spatial resolution of 1 km were used. We calculated the $PM_{2.5}$ exposure as the mean $PM_{2.5}$ concentration for each census tract over this period. Greenspace exposure was assessed using the enhanced vegetation index (EVI) as the indicator, calculated as the mean EVI value for each tract. The EVI data, derived from specific spectral bands (near-IR, red, and blue) of 2020 MODIS/006/MOD09GA surface reflectance, were processed using the method established by Huete et al. (2002).[33] LST data were obtained from the MOD11A1.061 Terra LST and Emissivity Daily Global 1-km product,[34] with exposure assessed by calculating the average daytime LST for the years 2000 to 2020 for each census tract. Data collection and processing for both EVI and LST were conducted through the GEE platform.

We compared disparities using global quantile-based urban and rural classification systems. Urban areas include tracts classified as "moderately urban," "highly urban," and "very urban," and rural areas include tracts classified as "moderately rural" and "very rural."

## 3. RESULTS

**3.1. Impact of Indicators on Urban/Rural Spatial Pattern.** We observed distinct spatial patterns in various indicator-based urban/rural classification systems using the global quantile (Figures 2 and 4a,d). Among them, the spatial pattern of the TotalPop-based system is significantly different from the others ($r < 0.02$, NRMSD > 0.4). The urban/rural levels of census tracts in the TotalPop-based system shifted significantly when other indicator-based systems, such as NTL and RoadDen, were applied (Figure 3a,b). In contrast, relatively similar spatial patterns were observed between systems based on indicators such as BldgDen and DisUrban (NRMSD = 0.12, $r$ = 0.91), as well as NTL and PopDen (NRMSD = 0.15, $r$ = 0.85). Notably, shifts in urban/rural levels between the BldgDen- and DisUrban-based systems predominantly occurred within the "rural" categories (Figure 3c), while shifts between the NTL-


https://doi.org/10.1021/acs.est.4c06942
Environ. Sci. Technol. 2024, 58, 19178−19188

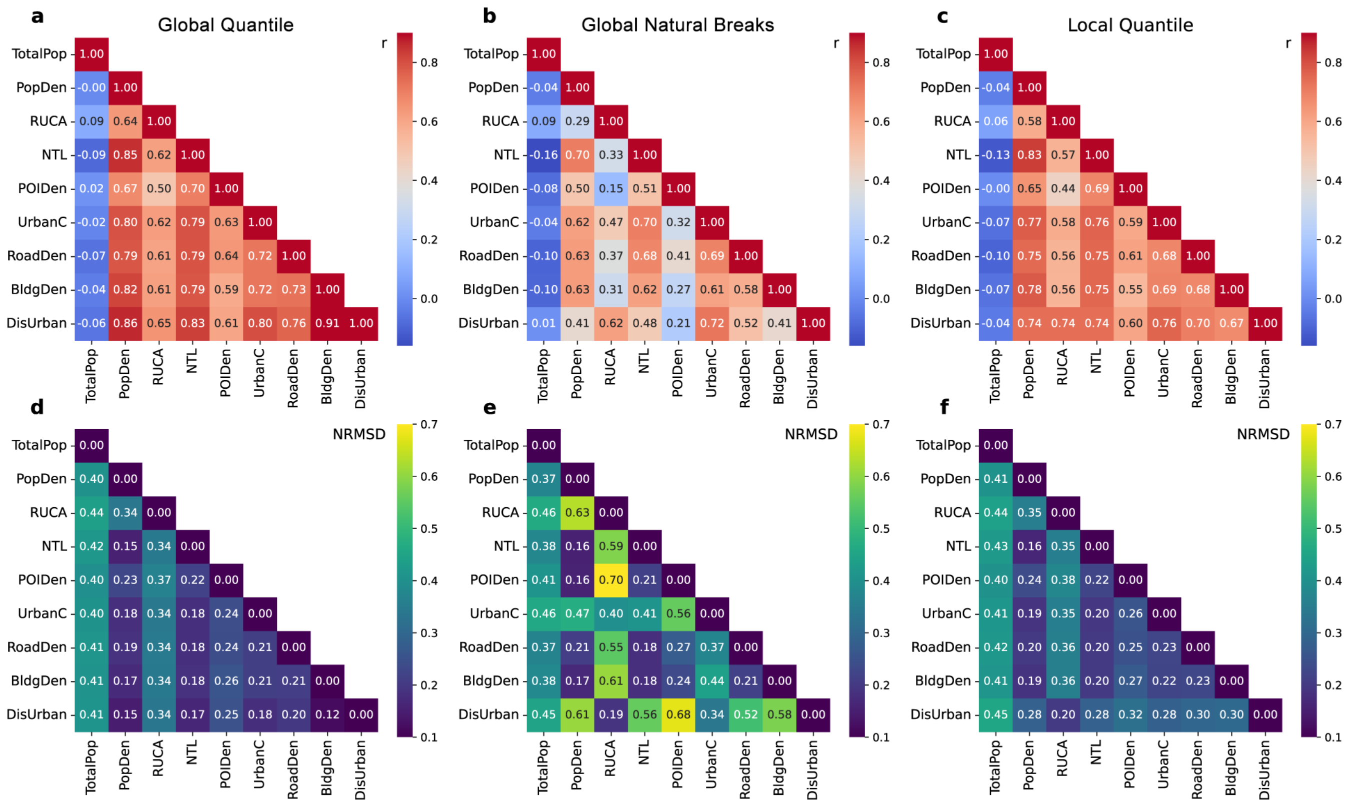



**Figure 4.** Pearson correlation coefficient ($r$) matrices and normalized root mean squared deviation (NRMSD) matrices across urban/rural classification systems using (a, d) global quantile, (b, e) global natural breaks, and (c, f) local quantile, respectively.

and PopDen-based systems predominantly occurred within the "urban" categories (Figure 3d).

**3.2. Impact of Classification Schemes on Urban/Rural Spatial Patterns.** Figure 4 illustrates the impact of classification schemes on the spatial patterns of urban/rural classification systems. Compared to the global quantile, global natural breaks increase the differences in spatial patterns between any two urban/rural classification systems (Figure 4a,b). For example, BldgDen- and DisUrban-based systems are most similar when using global quantile but exhibit significant changes under global natural breaks: $r$ decreases from 0.91 to 0.41 and NRMSD increases from 0.12 to 0.58. When using the local quantile, increased differences are observed between most systems compared to the global quantile (Figure 4c), though the impact is less pronounced than that of global natural breaks. Meanwhile, the similarity between certain systems is enhanced. For example, the $r$ between RUCA- and DisUrban-based systems increased from 0.65 to 0.74, while NRMSD decreased from 0.34 to 0.20.

When using the local quantile, increased differences are observed between most systems compared to the global quantile (Figure 4c), though the impact is less pronounced than that of global natural breaks. Meanwhile, the similarity between certain systems is enhanced. For example, the $r$ between RUCA- and DisUrban-based systems increased from 0.65 to 0.74, while NRMSD decreased from 0.34 to 0.20.

**3.3. Impact of Classification Schemes on Urban/Rural Demographic Compositions.** Figure 5 illustrates the compositions of the urban and rural population in the whole study area, Virginia, and West Virginia, respectively, using different classification schemes.

For the whole study area (Figure 5a−c), when using the global quantile (Figure 5a), the RUCA-based system has the highest urbanization rate at 89.43%, followed by DisUrban- (79.73%) and TotalPop-based (75.54%) systems. All other systems show urbanization rates of around 59.5%. When using global natural breaks (Figure 5b), notable variabilities in urban/rural demographic compositions are observed compared to the global quantile. Specifically, the DisUrban-based system turns out to be the one with the highest urbanization rate at 93.76%, followed by the RUCA- (89.43%) and TotalPop-based (73.61%) systems. Additionally, very low urbanization rates are observed in systems based on indicators such as BldgDen (20.65%), PopDen (11.49%), and POIDen (2.87%). When the local quantile was used (Figure 5c), demographic compositions of the whole study area remain generally consistent with the global quantile.

For Virginia (Figure 5d−f), the impacts of classification schemes were similar to those observed in the whole study area. Compared to global quantile (Figure 5d), global natural breaks (Figure 5e) result in notable variabilities in demographic compositions, but the local quantile (Figure 5f) has a very limited impact. In contrast, for West Virginia (Figure 5g−i), notable variations are observed with both global natural breaks (Figure 5h) and the local quantile (Figure 5i) compared to the global quantile (Figure 5g). Particularly, the local quantile results in a noticeable increase in urbanization rates, especially when using systems based on indicators such as RoadDen, BldgDen, and NTL.

**3.4. Impacts of Urban/Rural Classification Systems on Environmental Exposure Assessment.** The distribution patterns of the census tract with different environmental exposure levels are illustrated in Figures 6, S1, and S2 using violin plots. For the urban−rural disparity in $PM_{2.5}$ exposure in

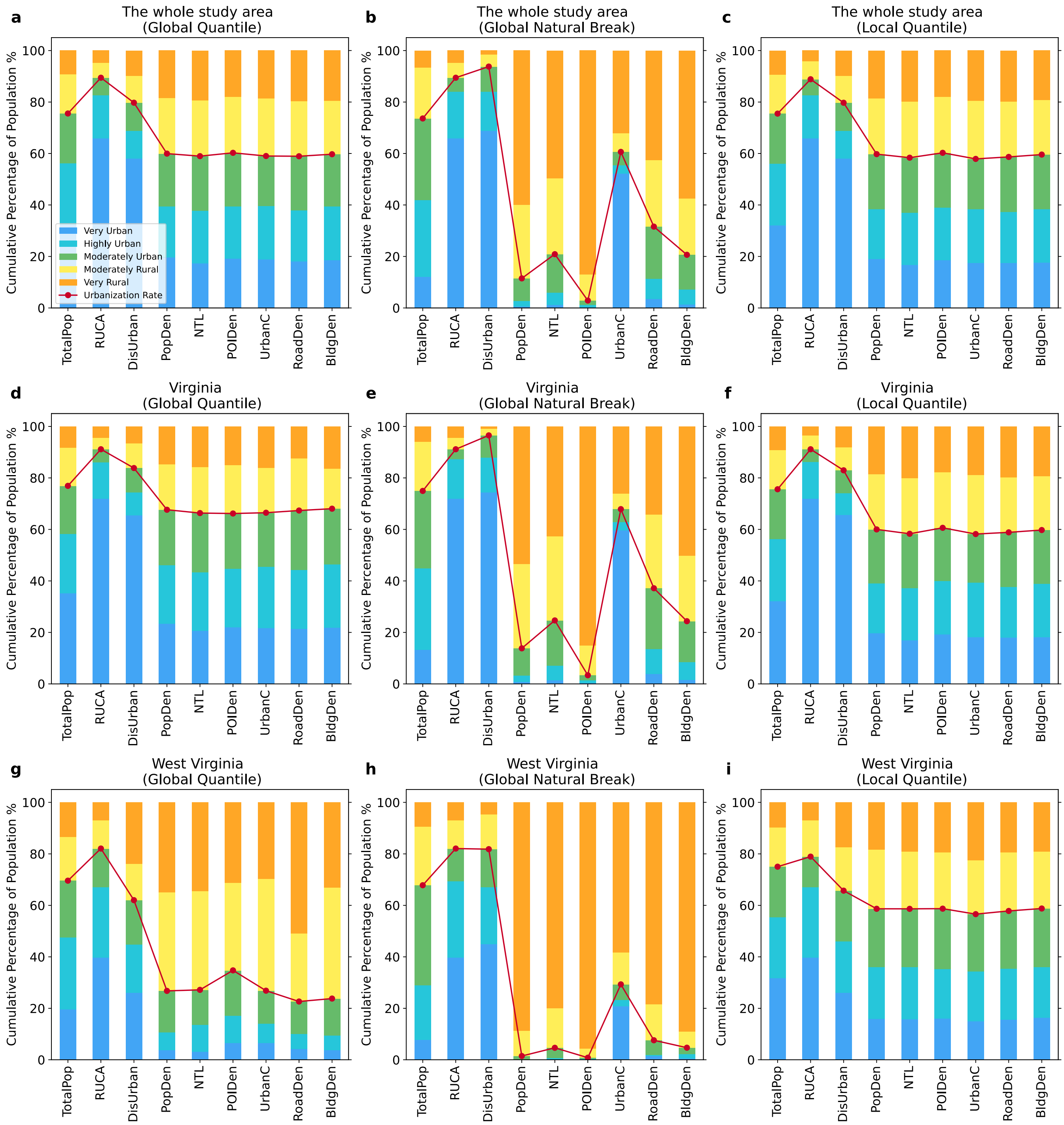



**Figure 5.** Demographic compositions of the whole study area (a−c), Virginia (d−f), and West Virginia (g−i), respectively, using various urban/rural classification systems.

Virginia, the impact of urban/rural classification systems is slight (Figures 6a, S1a, and S2a). Across all systems, urban and rural areas have similar shapes of violin plots, with a mean value around 7.10 $\mu g/m^3$. In contrast, a noticeable impact is observed in West Virginia (Figures 6b, S1b, and S2b), varying according to the classification system used. For instance, when the PopTotal-based system is used, rural areas have a higher mean $PM_{2.5}$ exposure (7.14 $\mu g/m^3$) compared to urban areas (6.86 $\mu g/m^3$). Conversely, the RUCA-, PopDen-, and NTL-based systems show reversed disparities, with rural areas having lower $PM_{2.5}$ exposure than urban areas. For example, using the PopDen-based system, rural areas have a $PM_{2.5}$ exposure of 6.41 $\mu g/m^3$ compared to 8.21 $\mu g/m^3$ in urban areas. Additionally, the shapes of the $PM_{2.5}$ exposure violin plots differ significantly across various urban and rural classification systems.

For greenspace exposure measured by EVI (Figures 6c,d, S1c,d, and S2c,d), the impact of urban/rural classification systems is noticeable in both Virginia and West Virginia. When the PopTotal-based system is used, urban areas consistently have slightly higher EVI values than rural areas. For example, in Virginia, urban areas have a greenspace exposure of 0.54 compared to 0.51 in rural areas. Other indicator-based systems,



https://doi.org/10.1021/acs.est.4c06942
Environ. Sci. Technol. 2024, 58, 19178−19188

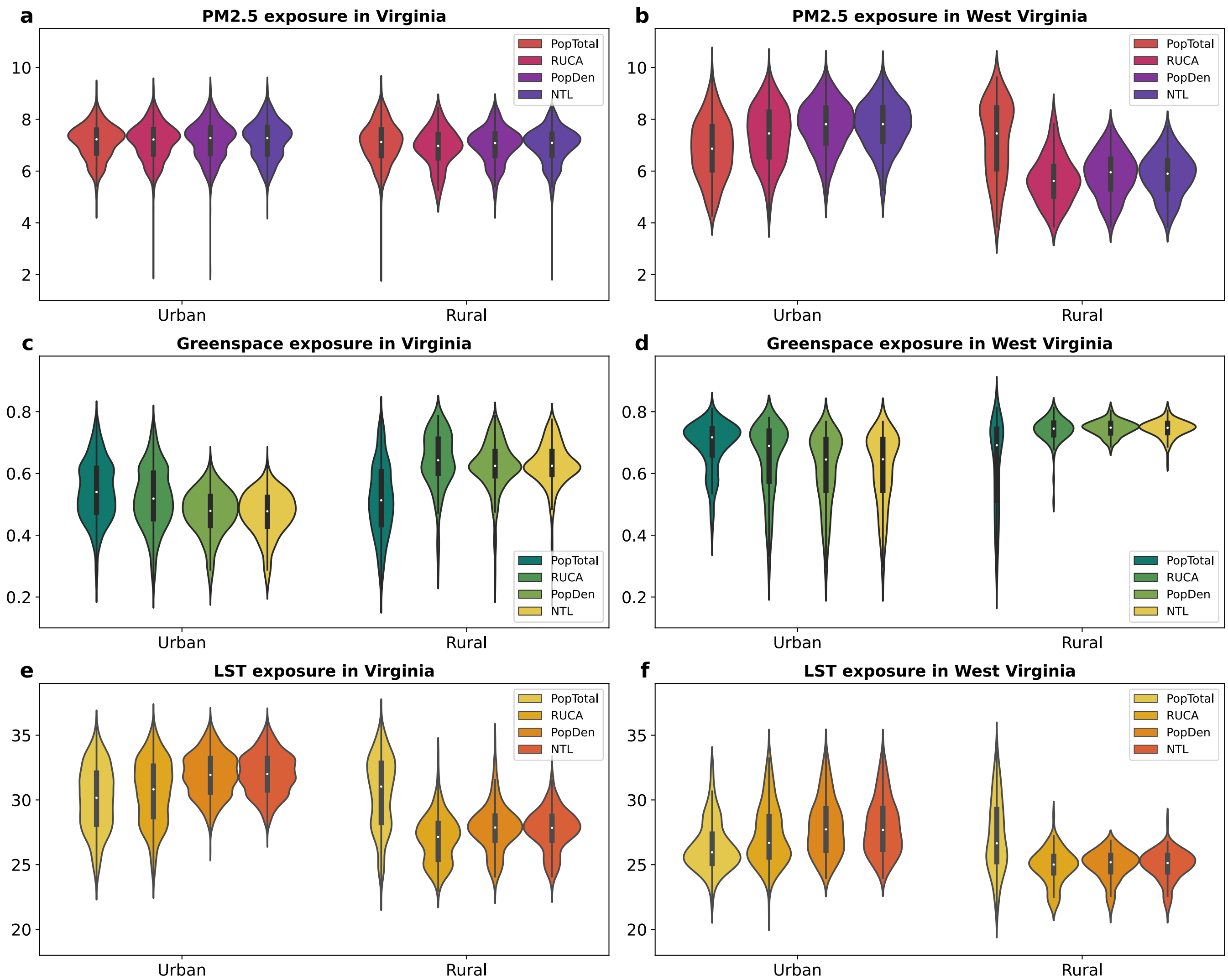



**Figure 6.** Violin plots of urban−rural disparities in environmental exposure in Virginia and West Virginia using various urban/rural classification systems based on PopTotal, RUCA, PopDen, and NTL indicators: (a, b) $PM_{2.5}$ exposure, (c, d) greenspace exposure, and (e, f) land surface temperature (LST) exposure.

such as RUCA, PopDen, and NTL, show a lower exposure level in urban areas than in rural areas. Moreover, distinct shapes of greenspace exposure violin plots are observed in urban and rural areas of the same classification system and across different classification systems.

For LST exposure (Figures 6e,f, S1e,f, and S2e,f), the impact of urban/rural classification systems is also noticeable, similar to the greenspace. Rural areas experience higher LST exposures when using the PopTotal-based system, whereas urban areas show higher LST exposures when other indicator-based systems are applied. Furthermore, significantly different shapes of LST exposure violin plots are observed both within the same classification system (between urban and rural areas) and across different classification systems.

## 4. DISCUSSION

In this study, we investigated how different urban and rural classification systems impact spatial patterns, demographic compositions, and environmental exposure assessments using data from two U.S. states. The disparities observed across various indicator-based classification systems highlight the fact that each indicator captures a unique aspect of the urban/rural gradient. However, each has its limitations and cannot fully encapsulate the complexity of urbanity and rurality. For example, census-derived indicators may suffer from demographic biases and delays in data updates.[35] An NTL-based indicator can be skewed by external light sources such as wildfires, ships, and large industrial areas.[30] Similarly, geospatial data sets like road network and building footprints may misclassify nonurban areas with high industrial activity, such as mining sites and power stations.[36]

To mitigate the limitations of single-factor indicators, advanced multidimensional approaches have been developed. For instance, the U.S. Department of Agriculture's Rural−Urban Continuum Codes (RUCCs), Urban Influence Codes (UICs), and RUCA codes incorporate factors like population density, population size, commuting patterns, and proximity to urbanized areas.[37] Additionally, some geospatial indicators integrate data from diverse sources, including optical remote sensing imagery, mobile phone data, geo-tagged social media data, POI, and NTL.[25,38] Although these multidimensional indicators provide a more comprehensive picture of the urban−rural gradient, they still carry the inherent limitations of the


https://doi.org/10.1021/acs.est.4c06942
Environ. Sci. Technol. 2024, 58, 19178−19188

original data. Given the complexity of urban/rural classification, no indicator can capture all facets of urban and rural areas.

Beyond urban/rural indicators, our findings also emphasize that the choice of classification scheme can significantly influence disparities between urban/rural classification systems. This complexity is further compounded in practical applications by additional factors not extensively covered in this paper. For example, in addition to the five-level classification strategy used in this study, other strategies including binary categories (e.g., metropolitan vs nonmetropolitan), three-tier systems (e.g., urbanized areas, urban clusters, and rural areas), continuous variables (e.g., population density), and multilevel categories (e.g., RUCCs with 9 levels and UICs with 12 levels) are also widely employed in urban/rural classification systems.[6,21] Additionally, the inconsistent use of geographic units is common across different urban/rural classification systems. While our study adopted the census tract level, other commonly used spatial scales include county, ZIP code, and multiscale pixels.[6,39,40] Any changes in the number of categories or the spatial scale can also result in distinct urban/rural classifications.[21]

The use of different urban/rural classification systems has led to varying, and sometimes contrasting, conclusions regarding urban/rural disparities in environmental exposure, highlighting the sensitivity of such assessments to the chosen classification systems. Moreover, the impact of these systems can vary across different regions (see Figure 6). Despite its importance, this issue has received a limited amount of attention. Most studies select an urban/rural classification system based on its widespread use or data availability, with few considering how different systems might impact research outcomes. We also observed that some urban and rural classification systems produce similar environmental exposure assessment outcomes. For instance, the PopDen, UrbanC, and RoadDen-based systems showed comparable results (see Figure S1). Previous studies have used such observations to conclude that 'similar' urban/rural classification systems have little impact on findings.[37] However, as shown in Figures 3 and S3, although the overall results may seem similar, the urban/rural systems are actually spatially distinct. More importantly, these observed differences in classification systems could have broader implications for environmental and health-related research.

Given the complexity of urban/rural classification systems and their impacts on applications such as environmental exposure assessment, there have been numerous calls for standardization.[5,8] However, it is important to recognize that each system has its own validity, effectively capturing certain aspects of urban/rural identity, but none is entirely comprehensive or flawless. Therefore, rather than striving for harmonized standards in every case, it is crucial to develop a deeper understanding of the distinctions between urban and rural classification systems and their implications across various application scenarios. This understanding would enable a more effective use of the appropriate system in specific contexts and ensure an accurate interpretation of the results.

Researchers should consider using multiple urban/rural classification systems in their analyses and conducting sensitivity analyses to better reveal urban−rural disparities and assess the robustness of findings across different classification approaches, rather than relying solely on a single system. For policymakers, it is essential to understand the limitations and implications of different classification systems to formulate policies that accurately address the needs of both urban and rural populations. We recommend that policymakers work closely with researchers to ensure that the chosen classification systems align with the policy goals and the characteristics of the target population. Collaborative and coordinated approaches to managing urban−rural differences can also help mitigate the limitations of individual classification systems and promote more integrated solutions.

In this study, we incorporated nine indicators and three classification schemes to construct urban−rural classification systems. However, due to the complexity of urban-rural characteristics, many other indicators could represent the urban/rural gradient, such as 3D information (e.g., building heights), economic activities, healthcare accessibility, or educational facilities.[41,42] Additionally, beyond the three classification schemes that we employed, various other methods can be applied for urban−rural classification. These include, but are not limited to, univariate classification (e.g., geometrical interval), cluster-based classification (e.g., hierarchical clustering), multicriteria classification (e.g., *K*-means), and supervised classification (e.g., decision trees).[43] Although we did not include these additional indicators and classification schemes in our analysis, this does not diminish the main finding of our study: different urban/rural classification systems can lead to significant changes in the magnitude and direction of urban/rural disparities in environmental exposure and health assessment.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c06942.

Thresholds of urban/rural levels for different urban/rural classification systems (Table S1); violin plots of urban−rural disparities in environmental exposure in Virginia and West Virginia (Figures S1 and S2); comparison of urban/rural spatial patterns between urban/rural classification systems (Figure S3) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Yimeng Song** − *School of the Environment, Yale University, New Haven, Connecticut 06511, United States;* orcid.org/0000-0001-9558-1220; Email: yimeng.song@yale.edu

### Authors

**Nicole C. Deziel** − *School of Public Health, Yale University, New Haven, Connecticut 06511, United States;* orcid.org/0000-0002-5751-9191

**Michelle L. Bell** − *School of the Environment, Yale University, New Haven, Connecticut 06511, United States; School of Health Policy and Management, College of Health Sciences, Korea University, Seoul 02841, Republic of Korea*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.4c06942

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under the award number

R01MD016054. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## ■ REFERENCES

(1) U.S. Census Bureau. *Urban and Rural*. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html. accessed 17 July 2024.

(2) Zhao, M.; Zhou, Y.; Li, X.; Cheng, W.; Zhou, C.; Ma, T.; Li, M.; Huang, K. Mapping urban dynamics (1992−2018) in Southeast Asia using consistent nighttime light data from DMSP and VIIRS. *Remote Sens. Environ.* **2020**, *248*, 111980.

(3) World Bank Group. *Rural population (% of total population)*. https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS. accessed 29 June 2024.

(4) Hannah, R.; Samborska, V.; Roser, M. *Urbanization*. https://ourworldindata.org/urbanization. accessed 29 June 2024.

(5) Hart, L. G.; Larson, E. H.; Lishner, D. M. Rural Definitions for Health Policy and Research. *Am. J. Public Health* **2005**, *95* (7), 1149−1155.

(6) Danek, R.; Blackburn, J.; Greene, M.; Mazurenko, O.; Menachemi, N. Measuring rurality in health services research: a scoping review. *BMC Health Services Res.* **2022**, *22* (1), 1340.

(7) Singh, G. K.; Siahpush, M. Widening Rural−Urban Disparities in Life Expectancy, U.S., 1969−2009. *Am. J. Prevent. Med.* **2014**, *46* (2), No. e19−e29.

(8) Nelson, K. S.; Nguyen, T. D.; Brownstein, N. A.; Garcia, D.; Walker, H. C.; Watson, J. T.; Xin, A. Definitions, measures, and uses of rurality: A systematic review of the empirical and quantitative literature. *J. Rural Stud.* **2021**, *82*, 351−365.

(9) Wang, S.; Kou, C.; Liu, Y.; Li, B.; Tao, Y.; D'Arcy, C.; Shi, J.; Wu, Y.; Liu, J.; Zhu, Y.; Yu, Y. Rural−Urban Differences in the Prevalence of Chronic Disease in Northeast China. *Asia Pacific J. Public Health* **2015**, *27* (4), 394−406.

(10) Seto, K. C.; Güneralp, B.; Hutyra, L. R. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Int. Acad. Sci* **2012**, *109* (40), 16083−16088.

(11) Qin, Y.; Xiao, X.; Dong, J.; Chen, B.; Liu, F.; Zhang, G.; Zhang, Y.; Wang, J.; Wu, X. Quantifying annual changes in built-up area in complex urban-rural landscapes from analyses of PALSAR and Landsat images. *ISPRS J. Photogramm. Remote Sens.* **2017**, *124*, 89−105.

(12) Gong, P.; Li, X.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Liu, X.; Xu, B.; Yang, J.; Zhang, W.; Zhou, Y. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sen. Environ.* **2020**, *236*, 111510.

(13) Blake, K. D.; Moss, J. L.; Gaysynsky, A.; Srinivasan, S.; Croyle, R. T. Making the Case for Investment in Rural Cancer Control: An Analysis of Rural Cancer Incidence, Mortality, and Funding Trends. *Cancer Epidemiol. Biomarkers Prevent.* **2017**, *26* (7), 992−997.

(14) Tuholske, C.; Caylor, K.; Funk, C.; Verdin, A.; Sweeney, S.; Grace, K.; Peterson, P.; Evans, T. Global urban population exposure to extreme heat. *Proc. Int. Acad. Sci.* **2021**, *118* (41), No. e2024792118.

(15) Mohajeri, N.; Hsu, S. C.; Milner, J.; Taylor, J.; Kiesewetter, G.; Gudmundsson, A.; Kennard, H.; Hamilton, I.; Davies, M. Urban-rural disparity in global estimation of PM2.5 household air pollution and its attributable health burden. *Lancet Planet. Health* **2023**, *7* (8), No. e660−e672.

(16) Kuddus, M. A.; Tynan, E.; McBryde, E. Urbanization: a problem for the rich and the poor? *Public Health Rev.* **2020**, *41* (1), 1.

(17) Son, J. Y.; Bell, M. L. Exposure to animal feeding operations including concentrated animal feeding operations (CAFOs) and environmental justice in Iowa, USA. *Environmental Res.* **2023**, *1* (1), 015004.

(18) Weichenthal, S.; Moase, C.; Chan, P. A Review of Pesticide Exposure and Cancer Incidence in the Agricultural Health Study Cohort. *Environ. Health Perspect* **2010**, *118* (8), 1117−1125.

(19) Aggarwal, R.; Chiu, N.; Loccoh, E. C.; Kazi, D. S.; Yeh, R. W.; Wadhera, R. K. Rural-Urban Disparities: Diabetes, Hypertension, Heart Disease, and Stroke Mortality Among Black and White Adults, 1999−2018. *J. Am. College Cardiol.* **2021**, *77* (11), 1480−1481.

(20) Cross, S. H.; Mehra, M. R.; Bhatt, D. L.; Nasir, K.; O'Donnell, C. J.; Califf, R. M.; Warraich, H. J. Rural-Urban Differences in Cardiovascular Mortality in the US, 1999−2017. *JAMA* **2020**, *323* (18), 1852−1854.

(21) Zahnd, W. E.; Del Vecchio, N.; Askelson, N.; Eberth, J. M.; Vanderpool, R. C.; Overholser, L.; Madhivanan, P.; Hirschey, R.; Edward, J. Definition and categorization of rural and assessment of realized access to care. *Health Services Res.* **2022**, *57* (3), 693−702.

(22) U.S. Census Bureau. *Decennial Census Redistricting PL 94−171*. https://www.census.gov/programs-surveys/decennial-census/about/rdo/summary-files.html. accessed 29 June 2024.

(23) Zheng, Y.; Tang, L.; Wang, H. An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI. *J. Cleaner Prod.* **2021**, *328*, 129488.

(24) Gong, P.; Chen, B.; Li, X.; Liu, H.; Wang, J.; Bai, Y.; Chen, J.; Chen, X.; Fang, L.; Feng, S.; Feng, Y.; Gong, Y.; Gu, H.; Huang, H.; Huang, X.; Jiao, H.; Kang, Y.; Lei, G.; Li, A.; Li, X.; Li, X.; Li, Y.; Li, Z.; Li, Z.; Liu, C.; Liu, C.; Liu, M.; Liu, S.; Mao, W.; Miao, C.; Ni, H.; Pan, Q.; Qi, S.; Ren, Z.; Shan, Z.; Shen, S.; Shi, M.; Song, Y.; Su, M.; Suen, H. P.; Sun, B.; Sun, F.; Sun, J.; Sun, L.; Sun, W.; Tian, T.; Tong, X.; Tseng, Y.; Tu, Y.; Wang, H.; Wang, L.; Wang, X.; Wang, Z.; Wu, T.; Xie, Y.; Yang, J.; Yang, J.; Yuan, M.; Yue, W.; Zeng, H.; Zhang, K.; Zhang, N.; Zhang, T.; Zhang, Y.; Zhao, F.; Zheng, Y.; Zhou, Q.; Clinton, N.; Zhu, Z.; Xu, B. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. *Sci. Bull.* **2020**, *65* (3), 182−187.

(25) Song, Y.; Chen, B.; Ho, H. C.; Kwan, M. P.; Liu, D.; Wang, F.; Wang, J.; Cai, J.; Li, X.; Xu, Y.; He, Q.; Wang, H.; Xu, Q.; Song, Y. Observed inequality in urban greenspace exposure in China. *Environ. Int.* **2021**, *156*, 106778.

(26) OpenStreetMap Data Extracts. https://download.geofabrik.de/north-america/us.html. accessed 29 June 2024.

(27) Balk, D.; Leyk, S.; Jones, B.; Montgomery, M. R.; Clark, A. Understanding urbanization: A study of census and satellite-derived urban classes in the United States, 1990−2010. *PLoS One* **2018**, *13* (12), No. e0208487.

(28) U.S. Census Bureau. *TIGER/Line Shapefiles*. https://www2.census.gov/geo/tiger/TIGER2020/UAC. accessed 2 June 2024.

(29) Levin, N.; Zhang, Q. A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas. *Remote Sen. Environ.* **2017**, *190*, 366−382.

(30) Cai, J.; Huang, B.; Song, Y. Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* **2017**, *202*, 210−221.

(31) De Smith, M. J.; Goodchild, M. F.; Longley, P. *Geospatial analysis: A comprehensive guide to principles, techniques and software tools*; Troubador publishing ltd, 2007.

(32) Di, Q.; Amini, H.; Shi, L.; Kloog, I.; Silvern, R.; Kelly, J.; Sabath, M. B.; Choirat, C.; Koutrakis, P.; Lyapustin, A.; Wang, Y.; Mickley, L. J.; Schwartz, J. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **2019**, *130*, 104909.

(33) Huete, A.; Didan, K.; Miura, T.; Rodriguez, E. P.; Gao, X.; Ferreira, L. G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83* (1), 195−213.

(34) Wan, Z.; Hook, S.; Hulley, G. *MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid*. https://lpdaac.usgs.gov/products/mod11a1v061. accessed 2 March 2024.

(35) Xu, Y.; Song, Y.; Cai, J.; Zhu, H. Population mapping in China with Tencent social user and remote sensing data. *Appl. Geograp.* **2021**, *130*, 102450.

(36) Sun, K.; Hu, Y.; Ma, Y.; Zhou, R. Z.; Zhu, Y. Conflating point of interest (POI) data: A systematic review of matching methods. *Comput, Environ. Urban Syst.* **2023**, *103*, 101977.

(37) Mulrooney, T.; Liang, C. L.; Kurkalova, L. A.; McGinn, C.; Okoli, C. Quantitatively defining and mapping rural: A case study of North Carolina. *J. Rural Stud.* **2023**, *97*, 47−56.

(38) Huang, X.; Hu, T.; Li, J.; Wang, Q.; Benediktsson, J. A. Mapping Urban Areas in China Using Multisource Data With a Novel Ensemble SVM Method. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56* (8), 4258−4273.

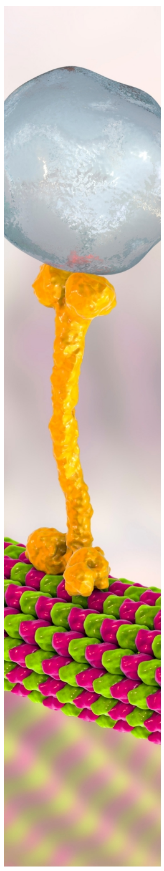
(39) Onega, T.; Weiss, J. E.; Alford-Teaster, J.; Goodrich, M.; Eliassen, M. S.; Kim, S. J. Concordance of Rural-Urban Self-identity and ZIP Code-Derived Rural-Urban Commuting Area (RUCA) Designation. *J. Rural Health* **2020**, *36* (2), 274−280.

(40) Liu, X.; Huang, Y.; Xu, X.; Li, X.; Li, X.; Ciais, P.; Lin, P.; Gong, K.; Ziegler, A. D.; Chen, A.; Gong, P.; Chen, J.; Hu, G.; Chen, Y.; Wang, S.; Wu, Q.; Huang, K.; Estes, L.; Zeng, Z. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nat Sustainabil.* **2020**, *3* (7), 564−570.

(41) Che, Y.; Li, X.; Liu, X.; Xu, X.; Huang, K.; Zhu, P.; Shi, Q.; Chen, Y.; Wu, Q.; Arehart, J. H.; Yuan, W.; Li, X. Mapping of individual building heights reveals the large gap of urban-rural living spaces in the contiguous US. *Innovat. Geosci.* **2024**, *2* (2), 100069.

(42) Song, Y.; Tan, Y.; Song, Y.; Wu, P.; Cheng, J. C.; Kim, M. J.; Wang, X. Spatial and temporal variations of spatial population accessibility to public hospitals: a case study of rural−urban comparison. *GISci. Remote Sens.* **2018**, *55* (5), 718−744.

(43) Guo, J.; Wang, J.; Xu, C.; Song, Y. Modeling of spatial stratified heterogeneity. *GISc. Remote Sens.* **2022**, *59* (1), 1660−1677.







https://doi.org/10.1021/acs.est.4c06942
*Environ. Sci. Technol.* 2024, 58, 19178−19188