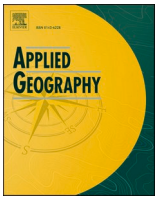
Applied Geography 172 (2024) 103414

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog



Exploring spatiotemporal dynamics, seasonality, and time-of-day trends of $PM_{2.5}$ pollution with a low-cost sensor network: Insights from classic and spatially explicit Markov chains

Michael Biancardi [a], Yuye Zhou [b], Wei Kang [c], Ting Xiao [a,d], Tony Grubesic [e], Jake Nelson [f], Lu Liang [b,*]

[a] Department of Computer Science and Engineering, University of North Texas, Denton, TX, USA
[b] Department of Landscape Architecture and Environmental Planning, University of California, Berkeley, USA
[c] Department of Geography and the Environment, University of North Texas, Denton, TX, 76203, USA
[d] Department of Information Science, University of North Texas, Denton, TX, USA
[e] School of Public Policy, University of California, Riverside, CA, 92521, USA
[f] Department of Geosciences, Auburn University, USA

ARTICLE INFO

Handling Editor: J Peng

Keywords:
Air pollution
Spatial patterns
Spatiotemporal analysis
Markov chains
Particulate matter

ABSTRACT

Fine particulate matter ($PM_{2.5}$) is a major health and environmental concern, with significant spatiotemporal dynamics in urban areas. Low-cost air quality sensor (LCS) networks offer a paradigm-changing opportunity to acquire high spatiotemporal resolution data, revealing the urban pollution landscape with sufficient detail for effective policymaking and health assessment. This study advances geospatial air quality research by using classic and spatial Markov chains to analyze the seasonality and intra-daily variations of $PM_{2.5}$ using LCS data. Results highlight distinctive $PM_{2.5}$ seasonality, with the "Good" state predominating in summer and being least common in winter. Midday is the peak period for the "Good" state, while mornings and nights have poorer conditions, suggesting a need for stricter pollution control during evening traffic rush hours. Notably, the impact of temporal scale on spatial Markov analysis is substantial, showing a broader range of air pollution states, increased stability, and reduced variation between time intervals compared to daily assessments. Site-level analysis reveals that rural sites are more likely to maintain "Good" state and less likely to transition out of it. Overall, this study highlights the effectiveness of high spatiotemporal resolution data and demonstrates the capacity of Markov chains to reveal nuances in phenomena such as air pollution.

1. Introduction

Fine particulate matter less than 2.5 μm in diameter ($PM_{2.5}$) is a prominent air pollutant, exerting substantial impacts on human health and the environment (Brunekreef & Holgate, 2002; Locosselli et al., 2019; Song et al., 2022). As the leading cause of global disease burden (Murray et al., 2020), $PM_{2.5}$ exposure is linked to premature mortality (Brunekreef et al., 2021; Mahler et al., 2023; Wang et al., 2022) and various diseases including heart disease (Alexeeff et al., 2021), lung cancer (Hamra et al., 2014), and other respiratory diseases (Lei et al., 2023).

Given the notable impacts and prevalence of $PM_{2.5}$ in urban environments, city decision-makers and stakeholders need high spatiotemporal air pollution data for policy refinement, evaluating mitigation efforts, and tracking progress toward urban Sustainable Development Goals. Two research questions are particularly important at the city level: How do $PM_{2.5}$ levels evolve over time and how do spatial relationships impact these patterns? Is there discernible evidence of seasonal and intra-daily variability in $PM_{2.5}$ levels? Many studies on spatiotemporal air pollution variations are often conducted on a regional to continental scale using median to coarse resolution data (e.g., Kalisa et al., 2023; Huang et al., 2021). This resolution is inadequate for capturing the fine-grained variations needed to estimate and mitigate risks at the individual or community level (Hart et al., 2020). While the gold standard of air pollution exposure assessments is personal measurements, collecting such measures for large populations remains a

* Corresponding author.
E-mail address: lianglu@berkeley.edu (L. Liang).

https://doi.org/10.1016/j.apgeog.2024.103414
Received 8 May 2024; Received in revised form 7 August 2024; Accepted 5 September 2024
Available online 10 September 2024

major challenge due to cost and logistical constraints (Larkin & Hystad, 2017). Sparsely located government-regulated air quality stations are often found only in populated cities (Badach et al., 2023; Bi et al., 2020). Low-cost sensor (LCS) networks have emerged as a paradigm-shift solution to supplement conventional regulatory stations (Snyder et al., 2013). These dense sensor networks make it feasible to understand localized spatiotemporal pollution patterns (Wang & Brauer, 2014).

One opportunity, yet a challenge, in analyzing high spatiotemporal resolution data is how to effectively distill useful information that traditional datasets cannot reveal. Markov chains, which model state transitions over time, have often been used for analyzing temporal dynamics. However, they neglect the spatial complexities and interactions. Spatial Markov chains (Rey & Franklin, 2022) address this by incorporating geographically adjacent observations, which is crucial for systems like air pollution with both high temporal and spatial dependence (Huang & Kuo, 2018; Yang et al., 2023). Previous studies have examined spatiotemporal patterns of air pollution distribution using Markov chains (Asadollahfardi et al., 2016; Caraka et al., 2019; Holmes & Hassini, 2021) and to a lesser extent, spatial Markov chains (Alyousifi et al., 2020), yet these studies possess limitations. First, seasonality was largely ignored despite evidence of seasonal $PM_{2.5}$ variations (Bodor et al., 2020). Secondly, the intra-daily patterns are rarely investigated due to the complexity and meticulousness required to capture short-term pollution fluctuations, compounded by acquisition and processing limitations of data at high temporal resolutions. Thirdly, most previous Markov chain analyses focus on overall trends, but divergent trends across different sites, influenced by unique site interactions with the urban environment, should be considered.

This study leverages a high spatiotemporal resolution $PM_{2.5}$ dataset obtained from a LCS network deployed across Denton County, Texas (Liang et al., 2023), aiming to gain a deeper understanding of air quality in a region historically affected by pollution (EPA, 2022). We seek to fill the gaps by employing Markov chains to investigate the intra-daily and seasonal trends of $PM_{2.5}$variations, alongside analyzing the spatial patterns and the associated influences from the urban built environment.

2. Theory and calculation

2.1. Classic Markov chains

A (finite) discrete-time Markov chain is a stochastic process that follows the Markov property, which states that the conditional probability of the next state $Y_{t+1} = X_j$ depends solely on the current state $Yt = Xi$ (Hillier & Lieberman, 2010).

$$P\left(Y_{t+1} = X_j \,\middle|\, Y_1 = X_{t_1}, Y_2 = X_{t_2}, \ldots, Y_t = X_i\right) = P\left(Y_{t+1} = X_j \,\middle|\, Y_t = X_i\right) \quad (1)$$

A Markov chain can be estimated using maximum likelihood estimation (MLE). Given the number of transitions *tij* from state *Xi* to state *Xj*, one can estimate transition probabilities by

$$\widehat{p_{ij}} = \frac{t_{ij}}{\sum_{j=1}^{N} t_{ij}}, \quad (2)$$

Transition probabilities are organized into a transition probability matrix $\widehat{P}$, of dimension $k \times k$, where k is the number of Markov states. The transition matrix is the core of the Markov chain model, which governs the dynamics of the system under study.

Assuming a time-homogeneous Markov chain where the transition matrix is consistent over the study period, the system could converge to a steady-state distribution (π) where the probability of being in each state is fixed while the system's dynamics is still governed by the transition matrix. This steady-state distribution serves as a probabilistic prediction of the air pollution in the long run assuming the current transitional dynamics persists. Mathematically, this distribution $\pi = \widehat{P}\pi$ can be calculated using linear algebra methods, since π is an eigenvector of $\widehat{P}$.

Another useful property is the First Mean Passage Time (FMPT), which represents the average number of transitions for the system to reach a specific state *Xj* from the current state *Xi*. The FMPT is given by

$$\pi = P\pi \quad (3)$$

Where fij^n denotes the probability of reaching state j for the first time after n time steps, given that the system begins in state i:

$$f_{ij}^{\,n} = \sum_{k \neq j} p_{ik} f_{kj}^{\,n-1} \quad (4)$$

This results in a system of linear equations that can be solved using standard methods. In general, this expression will only converge when $fij = 1$, meaning that the system will eventually reach state j from state i. However, if $f_{ij} < 1$, the system may never reach state j from state i, making the FMPT infinite.

The last Markov chain property of interest is the sojourn time. The sojourn time for a given state represents the average number of transitions needed to leave that state, thus measuring the stability of each state. Unlike FMPT, the sojourn time only focuses on exiting a state without considering the transitional state for the system. If M is the number of transitions required to leave state i, then:

$$P(M = n) = p_{ii}^{\,n-1} \times (1 - p_{ii}) \quad (5)$$

The sojourn time can then be calculated as

$$ST = E(M) = \frac{1}{(1 - p_{ii})} \quad (6)$$

2.2. Spatial Markov chains

Spatial Markov chains extend the capacities of classic Markov chains by accounting for spatial effects (Rey, 2001; Rey & Franklin, 2022). Spatial Markov modeling starts with defining the spatial relationship between observations. The spatial weight between pairs of observations represents the strength of the relationship between them. Two standard methods are binary weighting (assigning one to observations within a set distance; 0 otherwise) and inverse distance weighting (assigning weights reciprocal to the distance between observations within a certain threshold). After organizing these spatial weights into a matrix W, the spatial lag of an observation i at a given time Li, t is defined as a weighted average of neighbor states, where N is the number of sensors:

$$L_{i,t} = \sum_{j=1}^{M} W_{ij} * y_{j,t}, \quad (7)$$

The spatial lag for each observation at each time point was further discretized into six air pollution states as defined in Table 1. Instead of assuming a single transition matrix governing the air pollution dynamics , the Spatial Markov approach generates k ($k = 6$, in this study) $k \times k$

Table 1
Air quality categories based on EPA $PM_{2.5}$ standards and corresponding sample sizes.

Air quality category	Acronym	Concentration range ($\mu g/m^3$)	Daily sample size	Hourly sample size
Good	G	0–12.0	8147	221,579
Moderate	M	12.1–35.4	2888	64,895
Unhealthy for Sensitive Groups	USG	35.5–55.4	134	5552
Unhealthy	U	55.5–150.4	4	1239
Very Unhealthy	VU	150.5–250.4	1	53
Hazardous	H	250.5+	0	19

Note: Sample sizes denote the total number of occurrences of each state, i.e., the number of days or hours which occupy that state.

transition matrices, each of which is conditioned on one of k spatial lag states. Observations are partitioned by spatial lag, and a Markov chain is fit to each partition, thus enabling analysis of how spatial effects influence transition dynamics.

We used Ordinary Kriging to establish neighbor sets for each sensor based on a distance threshold where spatial autocorrelation becomes insignificant, guiding the calculation of inverse distance weights (additional details in SI). We fitted the Spatial Markov chains using MLE (Equation (2)), analyzing transitions for one spatial lag at a time. We derived three Markov chain properties: the steady-state distribution, FMPT, and the sojourn time.

2.3. Hypothesis tests

To investigate temporal scale, seasonality, and intra-daily effects, we fitted classic and spatial Markov chains to prepared LCS datasets, comprising: 1) hourly and daily data for the entire year; 2) seasonal datasets at both hourly and daily resolutions; and 3) hourly data segmented by time of day. Statistical tests were employed to assess the significance of differences in transition matrices for seasonality and intra-daily patterns, using a series of $X2$ tests at a significance level of 5% to evaluate transition matrix homogeneity (Bickenbach & Bode, 2003). The null hypothesis states identical transition dynamics across systems, as indicated by unchanged transition matrices capturing the system's full dynamics.

3. Data and methods

3.1. Sensor network and study area

Nested within the suburban area of the Dallas-Fort Worth metroplex in Texas (Fig. 1), Denton County spans an area of 2468 km^2 and hosts a densely populated community of nearly one million residents. Renowned as one of the fastest-growing counties in the U.S. (US Census Bureau, 2023), Denton faces persistent and longstanding air pollution issues (EPA, 2022) yet relies solely on a single regulatory $PM_{2.5}$ air quality monitoring station (Fig. 1). In response to this challenge, an initiative was launched to deploy a network of 85 PurpleAir (PA) sensors across Denton County, each equipped with two PMS5003 laser scattering particle counters for cross-checking (Liang et al., 2023). When light collides with the particles in the air, scattering occurs, enabling the estimation of particle counts based on particle diameter (Plantower, 2016). A proprietary count-to-concentration algorithm CF = ATM for outdoor sensors was then used to translate particle count into mass concentration (μg/m^3) by assuming an average particle density (Liang, 2021; Liang & Daniels, 2022). PA sensors continuously record real-time data, providing updates every 2 min and offering the option to download hourly data.

3.2. Low-cost sensor data processing

We collected 1-year hourly PA data from December 2021 to November 2022. Given that LCS devices (e.g., PurpleAir) can be susceptible to errors in particle counts, the conversion to concentration, and the influence of the ambient environment on LCS performance (Liang, 2021), we implemented a series of data pre-processing steps. We first removed data points with temperature values outside the 0–1000 °F range, relative humidity values exceeding 100%, or $PM_{2.5}$ values above 500 $\mu g/m^3$. Additionally, we tested sensor channel agreement by discarding records if the two $PM_{2.5}$ readings from a sensor differed by more than two standard deviations from the mean of the overall dataset or 5 $\mu g/m^3$ (Liang & Daniels, 2022). These two steps resulted in removal of 143,907 and 207,730 (1.8% and 2.7%) of the data, separately. We then applied a neural network-based calibration method, developed using collocated PA and regulatory instruments across the U.S. to calibrate our data (Liang & Daniels, 2022). By comparing with the data collected from the nearby regulatory station, the calibration has improved the R^2 between PA and reference data to 0.94, with an RMSE of 1.5 $\mu g/m^3$.

In addition to the hourly dataset, we compiled a daily aggregated $PM_{2.5}$ dataset for assessing seasonality. Data from days with data completeness below 90% were removed to ensure aggregation accuracy, and L1 outlier detection was applied to eliminate extreme values.

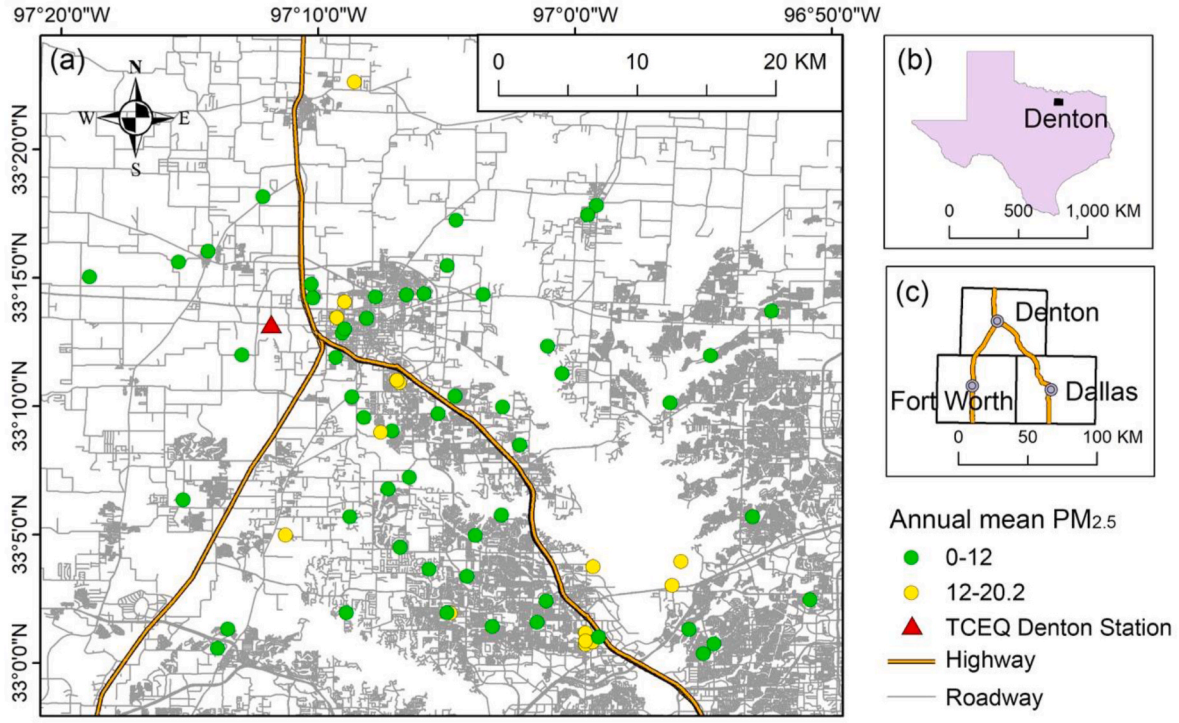


Fig. 1. The spatial distribution of deployed LCS sensors (a) in Denton County, Texas (state and regional location shown in b, c). The annual means of $PM_{2.5}$ collected by the sensors were color-coded according to the EPA standard.

Subsequently, we averaged all hourly data by day. For detailed pre-processing steps, refer to Liang et al. (2023). Subsequently, we discretized the $PM_{2.5}$ concentration data into six distinct states representing varying levels of air quality standards (Table 1), serving as input for the Markov chain models.

Data gaps at a location can stem from prolonged sensor disconnection from Wi-Fi or random missing timesteps caused by inaccuracies, resulting in significant discrepancies between two sensor channels. These data gaps pose challenges in subsequent Markov chain modeling, though the impact of any form of missing data is unclear. To address this, we conducted sensitivity analyses and artificial removal tests, finding that missing data chunks over 20% of the time period notably impact results, while randomly missing timesteps have negligible effects (additional details in SI). Consequently, we excluded sensors with missing chunks exceeding this threshold of the time period in question for each time period we are fitting a Markov chain for (i.e., the entire year and each season).

3.3. Defining seasons and time of day

We defined the seasons in Denton County as March to May (spring), June to August (summer), September to November (fall), and December to February (winter). For the intra-daily analysis, we categorized the hourly data into distinct time-of-day segments based on local traffic patterns and preliminary exploration of $PM_{2.5}$ data. The DFW area's typical morning and evening rush hours are 6:30 a.m.-9:30 a.m. and 3:00 p.m.-7:30 p.m. (Sneed, 2022; TexasView, nd). Moreover, the exploratory data analysis of $PM_{2.5}$ intra-daily patterns reveals a consistent trend across all seasons: an initial rise at 5 a.m., peaking at 8 a.m., descending until 5 p.m., followed by a renewed uptick (Fig. 2). This observation led to the definition of the following time segments: 5 a.m.–10 a.m. as morning, 10 a.m.-3 PM as midday, 3 p.m.–9 p.m. as afternoon/evening, and 10 p.m.-5 a.m. as night. A Markov chain model was then fitted to each time segment, running for the entire year and each season.

3.4. Spatial pattern analysis of site-level Markov chain

To understand the spatial trends and their environmental drivers, we conducted site-specific Markov chain analyses across the LCS network. For each site, we computed four properties: steady probabilities of "Good" and "Moderate" state, and sojourn times in each state. We further quantified nine different built environment variables—percentage of impervious surface, average tree height, and average building height—at three buffer sizes (200 m, 500 m, 1000 m). Impervious surface data was derived from a 2022 land cover map produced using imagery from the National Agriculture Imagery Program and classified with the deep-learning UrbanWatch FLUTE framework (Zhang et al., 2022). Building and tree height models were generated from 1-m airborne LiDAR data, achieving an accuracy of less than 0.286 m by comparing against the ground truth. Further methodological details can be found in Liang et al., 2023. We then correlated each built environment variables with the Markov chain properties to discern potential drivers.

4. Results

4.1. Classic Markov chain results

4.1.1. Diurnal and hourly patterns

Throughout the entire year, we observed daily transitions among the "Good", "Moderate", and "USG" states (Fig. 3a and 4). Specifically, the "Good" state showed a high likelihood of remaining "Good" (0.817), while the "Moderate" state had nearly equal probabilities of persisting as "Moderate" (0.499) or transitioning to "Good" (0.481). The "USG" state showed a strong probability of transitioning to "Moderate" (0.823). The "Good" state had the longest sojourn time (5.5 days), meaning it took an average of 5.5 days to leave the "Good" state. In contrast, it did not take long to leave the "Moderate" state (2.0 days) and "USG" state (1.0 day). These trends align with FMPT results, where the "Good" state requires 1.4 days to revert to itself and 5.6 days to transition to the "Moderate" state. In contrast, the "Moderate" state takes 3.7 days to return to itself and 2.1 days to reach the "Good" state. The steady-state distribution reveals that air quality is predominantly "Good" for most days (71.8%), followed by 26.9% in the "Moderate" state in the long run.

Hourly patterns closely mirror daily patterns, albeit with minor variations. The "Good" state occupies 75.6% of the time, while the "Moderate" state accounts for 22.1% in the long run. A fair degree of stability prevailed throughout the year, with less stability for more severe pollution states. For instance, the sojourn time of the "Good" state notably exceeds that of the other states by four to 13 times (Fig. 4). FMPT indicates that the "Good" state requires significantly less time to

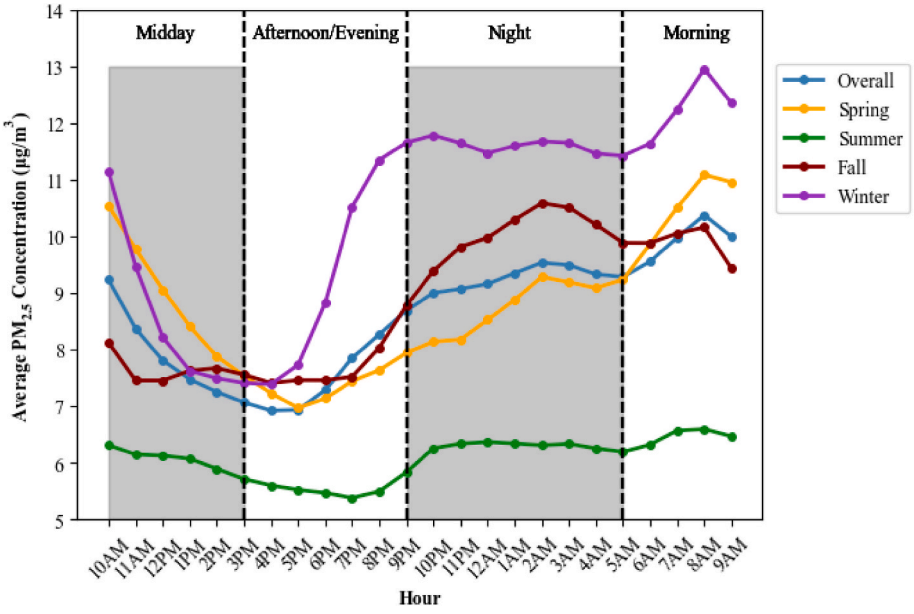


Fig. 2. Hourly average $PM_{2.5}$ concentrations for each season.

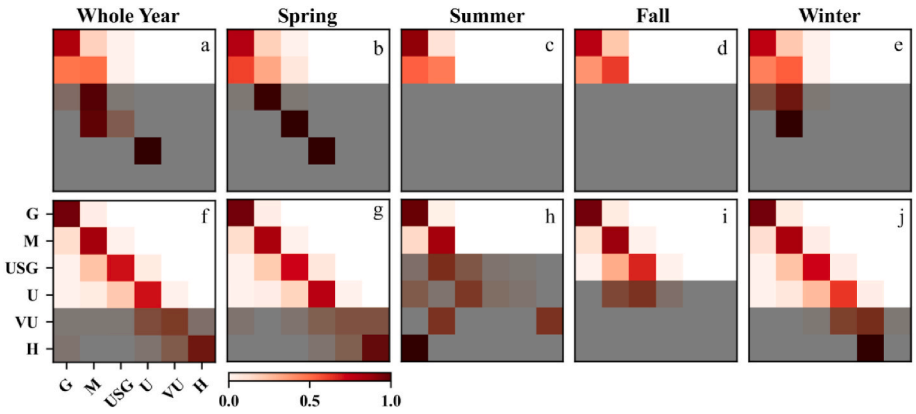


Fig. 3. Transition probabilities: daily (a–e) and hourly (f–j) scales for the whole year, spring, summer, fall, and winter. Grayed boxes indicate small sample sizes (less than 150 transitions). G: Good; M: Moderate; USG: Unhealthy for Sensitive Groups; U: Unhealthy; VU: Very Unhealthy; H: Hazardous.

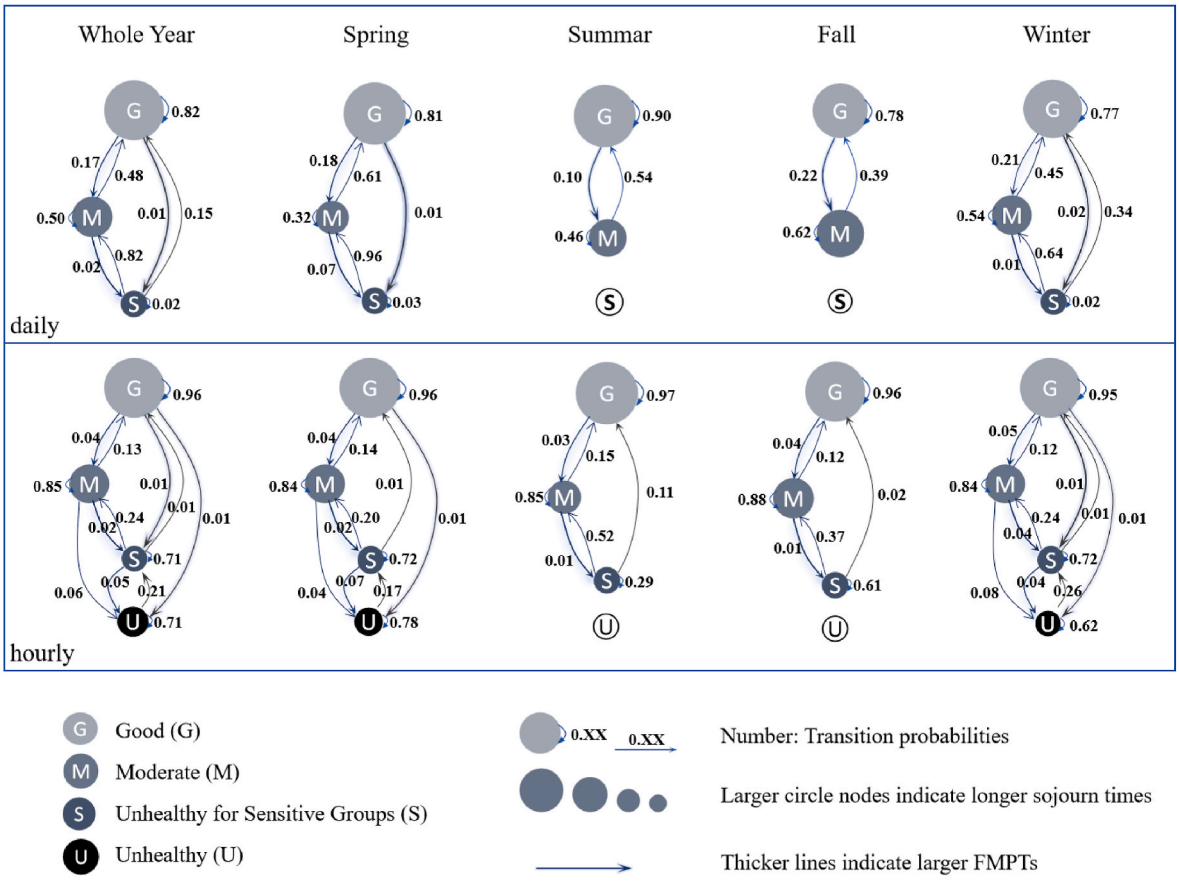


Fig. 4. State diagrams of classic Markov models for daily scale (top row) and hourly scale (bottom row). Larger circle nodes indicate longer sojourn times and thicker lines indicate larger FMPTs. Transition probabilities are labeled. G: Good; M: Moderate; USG: Unhealthy for Sensitive Groups; U: Unhealthy; VU: Very Unhealthy; H: Hazardous.

return to itself (1.3 h) than transition to the "Moderate" state (26.5 h). The "Moderate" state reverts in about 4.5 h, reaches the "Good" state in around 8.5 h, and requires notably more time to transition to worse states. Additionally, the USG and Unhealthy states usually shift to the "Moderate" state after 5.8 h and 9.2 h, respectively.

4.1.2. *Seasonality patterns*

Seasonal differences in transition probabilities are statistically significant, as indicated by X^2 tests for Markov homogeneity (Fig. 3b–e). Sojourn times generally followed consistent seasonal trends when compared to the entire year. However, a notable exception was observed in summer, during which the "Good" state displayed an extended sojourn time of 9.6 days. Steady-state distributions also displayed seasonal variability, with spring resembling the overall pattern. In summer, a greater portion of days (83.8%) were in the "Good" state, which dropped to 63%–66% in fall and winter (Fig. 5).

Unlike the daily-scale results, we did not identify any exceptional cases at the hourly-scale. Although each season's transition probability matrix patterns appear like the overall year's pattern (Fig. 3g–j, Fig. 4), the X^2 tests indicate their noteworthy distinctions (Table S5). Similar

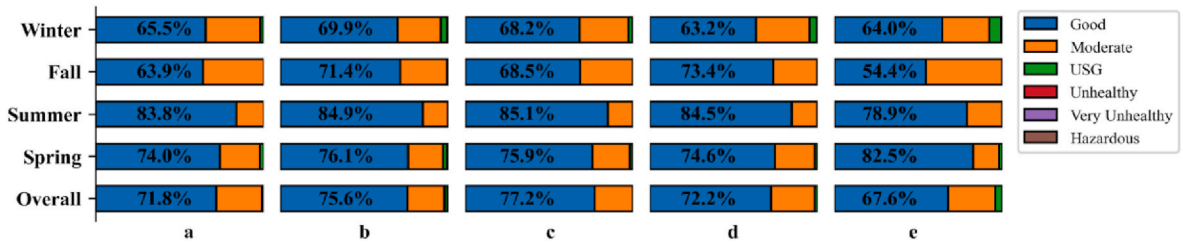


Fig. 5. Steady-state distributions for classic Markov at the daily scale (a), classic Markov at the hourly scale (b), spatial Markov at the daily scale with an average lag of Good (c), spatial Markov at the hourly scale with an average lag of Good (d), and spatial Markov at the hourly scale with an average lag of Moderate (e).

trends are noted in the sojourn times for each season, mirroring the overall year's pattern, with the "Good" state showing notably heightened stability during summer (additional information in SI). Additionally, the trends in steady-state distributions broadly resemble those observed at the daily scale. Summer still commands the largest share of the "Good" state (84.9%), compared to the overall yearly pattern's 75.6%, while winter exhibits the smallest portion of the "Good" state (69.9%).

4.1.3. *Intra-daily pattern of $PM_{2.5}$ variation*

During morning hours, the "Good" state dominates, ranging from 55.5% in spring to 83.7% in summer, with a yearly average of 71.6%. Morning sojourn times often surpass the actual time window, implying high stability. Midday sees higher stability for the "Good" state across all seasons than the morning period, evident through greater transition probabilities and longer sojourn times. For example, in the fall, the Good state's sojourn time is around 27 h compared to 7.3 and 2.2 h for the Moderate and USG states, respectively. Generally, the "Good" state occupies 87%–90.5% of the midday period, with fall and winter displaying a larger proportion than spring and summer. Afternoon/evening is relatively stable in the "Good", "Moderate", and USG states. Seasonal "Good" state proportions vary more dramatically among seasons, ranging from 72.8% yearly average to 86.3% peak in summer and 57.8% trough in winter. While the "Good" state predominates during nighttime on average, it occupies less time than other periods. This is particularly evident in fall and winter, with their lower proportions (59% and 62.8%) than in spring and summer (72.4% and 82.2%).

4.2. *Spatial Markov results*

4.2.1. *Daily and hourly patterns for an entire year*

At the daily scale, spatial Markov chains showed significant spatial dependence among LCS sites, with average lags of Good and Moderate, indicating neighbors' states as "Good" or "Moderate". A χ^2 hypothesis test indicates the statistical significance of spatial effects at the daily scale, meaning a location's $PM_{2.5}$ air quality state depends on the state of its neighbors. The "Good" state shows more substantial stability under an average lag of Good than Moderate: the staying probability decreases from 82.6% to 77.9%, and the sojourn time decreases from 5.7 to 4.5 days. While the "Moderate" state generally exhibits lower stability, its stability improves under a "Moderate" average lag compared to a "Good" average lag, as indicated by an increased probability of remaining from 43.4% to 52.2% and a longer sojourn time from 1.8 to 2.1 days. Furthermore, despite its small sample size (n = 58), the USG state was observed more frequently and demonstrated lower stability, with a high likelihood (84.5%) of transitioning to "Moderate".

At the hourly scale, in addition to lags of Good and Moderate across all seasons, there were also limited instances of average lags for "USG" and "Unhealthy" states. Across all lags, we observed a notable level of hour-to-hour stability, as evidenced by elevated probabilities along the diagonals of the transition matrix. Under an average lag of Good, the "Good" state maintains a predominant role, occupying 72% of the time with a 106-h sojourn time. The "Moderate" state follows with a 26%-time occupancy and a 39-h sojourn time, while the remaining "USG" state has a 25-h sojourn time. In scenarios of "Moderate" average lag, a subtle shift occurs: the dominance of the "Good" state decreases to 68% and a 91-h sojourn time, while the "Moderate" state gains more influence, taking up 28% of the time with a 41-h sojourn time. The "USG" state, now occupying 4% of the time, maintains its 25-h sojourn time (Fig. 6).

Interestingly, the hypothesis test conducted at the hourly scale fails to reject the null hypothesis of spatial homogeneity ($p = 0.128$), likely due to less hour-to-hour variation. As a result, the findings from the hourly spatial Markov analysis align with those from the classic Markov analysis.

4.2.2. *Seasonality of daily and hourly patterns*

At the daily scale, spatial Markov analysis reveals seasonal patterns similar to traditional Markov chains, particularly evident under "Good" average lag conditions. However, under "Moderate" average lag conditions, the "Moderate" state shows distinct seasonal patterns, remaining stable in summer and winter, transitioning more to "Good" state in spring and displaying variability in fall. Notably, under "Moderate" average lag conditions, spring and fall exhibit reduced stability in the "Moderate" state, with shorter sojourn times and increased probabilities of transitioning out of, less pronounced in summer or winter.

We observed seasonality in the hourly-scale spatial Markov results (Fig. 6). In fall, akin to the overall year, the "Good" state prevails with an average lag of "Good" (74% the of time; sojourn time of 127 h). With a "Moderate" average lag, the "Good" state decreases to 54% of the time and a 56-h sojourn time. Summer shows a similar trend, with the "Good" state dominant for both "Good" (85% of the time) and "Moderate" lags (79% of the time). Spring has mixed results. The "Moderate" state becomes slightly more stable but less prominent when the average lag is "Moderate" compared to "Good", with the occupancy time increasing from 15% to 24% and the sojourn time increasing from 35 to 37 h. Winter follows a comparable pattern to spring, albeit with the addition of the "USG" state. Under an average lag of "Good", the "Good", "Moderate", and "USG" states respectively account for 63%, 33%, and 4% of the time. These proportions shift to 64%, 28%, and 7% when the average lag is moderate. The stability of the "Moderate" state remains relatively consistent with sojourn times of 41 h compared to 41.4 h, and an unchanged probability of 97.6% for remaining in the "Moderate" state.

Spatial effects were insignificant at the hourly scale during spring, summer, and winter, with p-values of 0.605, 0.846, and 0.898, respectively. Nonetheless, during fall, spatial effects were found to be significant at the hourly scale with a p-value of 0.031.

4.3. *Site-level Markov chain analysis*

The site-specific Markov chain analysis reveals a clear spatial stratification across the landscape, with a noticeable trend of decreasing steady probabilities of "Good" state observed from rural to suburban to urban areas (Fig. 7a and Fig. 8a). Sojourn times in "Good" state exhibit a similar trend, where urban sites tend to be more likely to transit out of

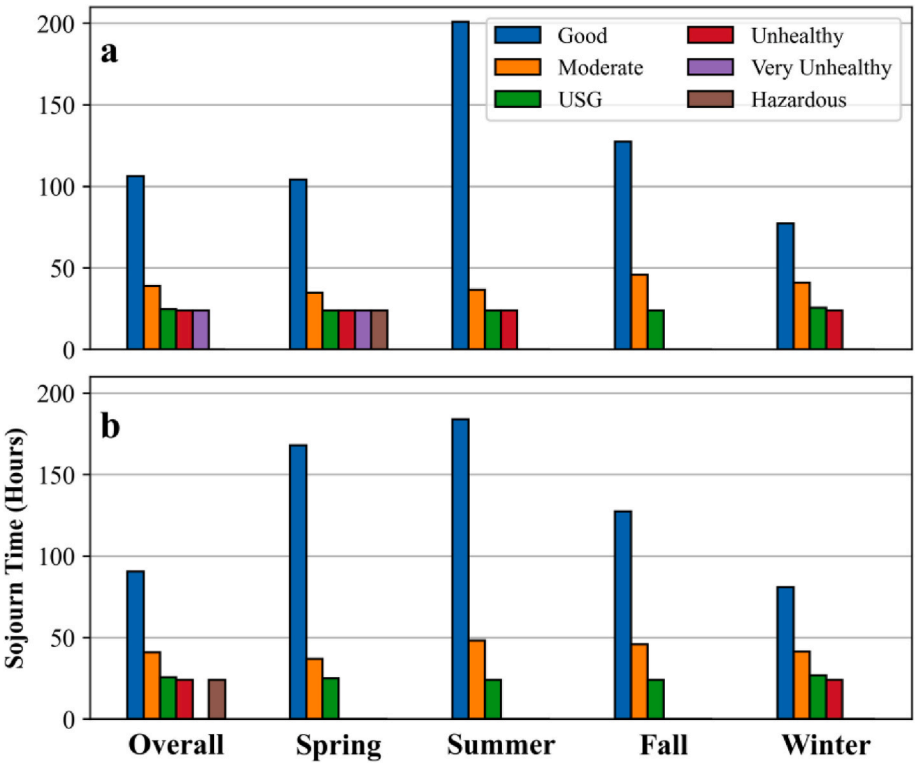


Fig. 6. Sojourn times for hourly-scale Spatial Markov with average lags of Good (a) and Moderate (b).

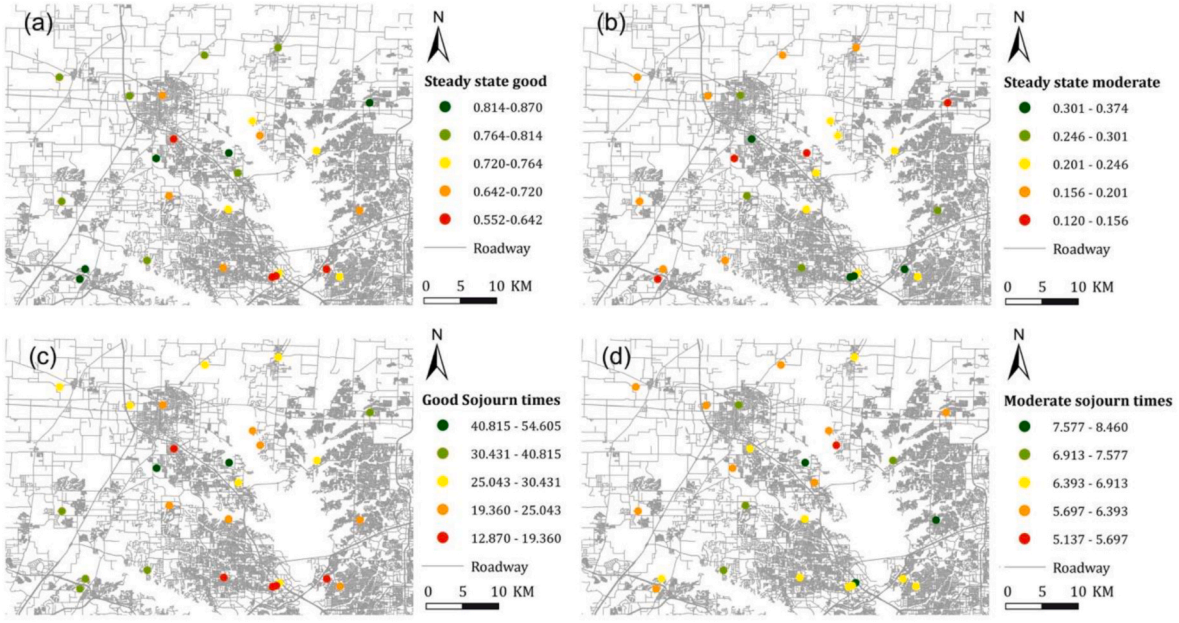


Fig. 7. Spatial distribution of Markov chain results across LCS networks. (a) Steady probabilities of "Good" state; (b) Steady probabilities of "Moderate" state; (c) Sojourn times of "Good" state; (d) Sojourn times of "Moderate" state.

"Good" state (Fig. 8c). This trend shows similar patterns across all seasons, with the summer showing the least variations of steady probability in "Good" state and winter showing the largest variations across different urbanization levels. For "Moderate" state, a reversed spatial pattern was observed, wherein sites at densely urbanized strata, characterized by high impervious surface levels and closer proximity to road networks, exhibited higher probability and longer sojourn times in the "Moderate" state (Fig. 7b–d, Fig. 8d).

The correlation between Markov chain properties and nine built environment variables reveals an interesting pattern (Figs. S8–11). Higher impervious surface percentages, taller trees and taller buildings exhibit negative associations with steady state probabilities in "Good" state but positive correlations with probabilities in "Moderate" state. The sojourn times of "Good" state show a similar pattern with built environment variables. Conversely, the sojourn times of "Moderate" state display essentially no significant relationship with any variables.

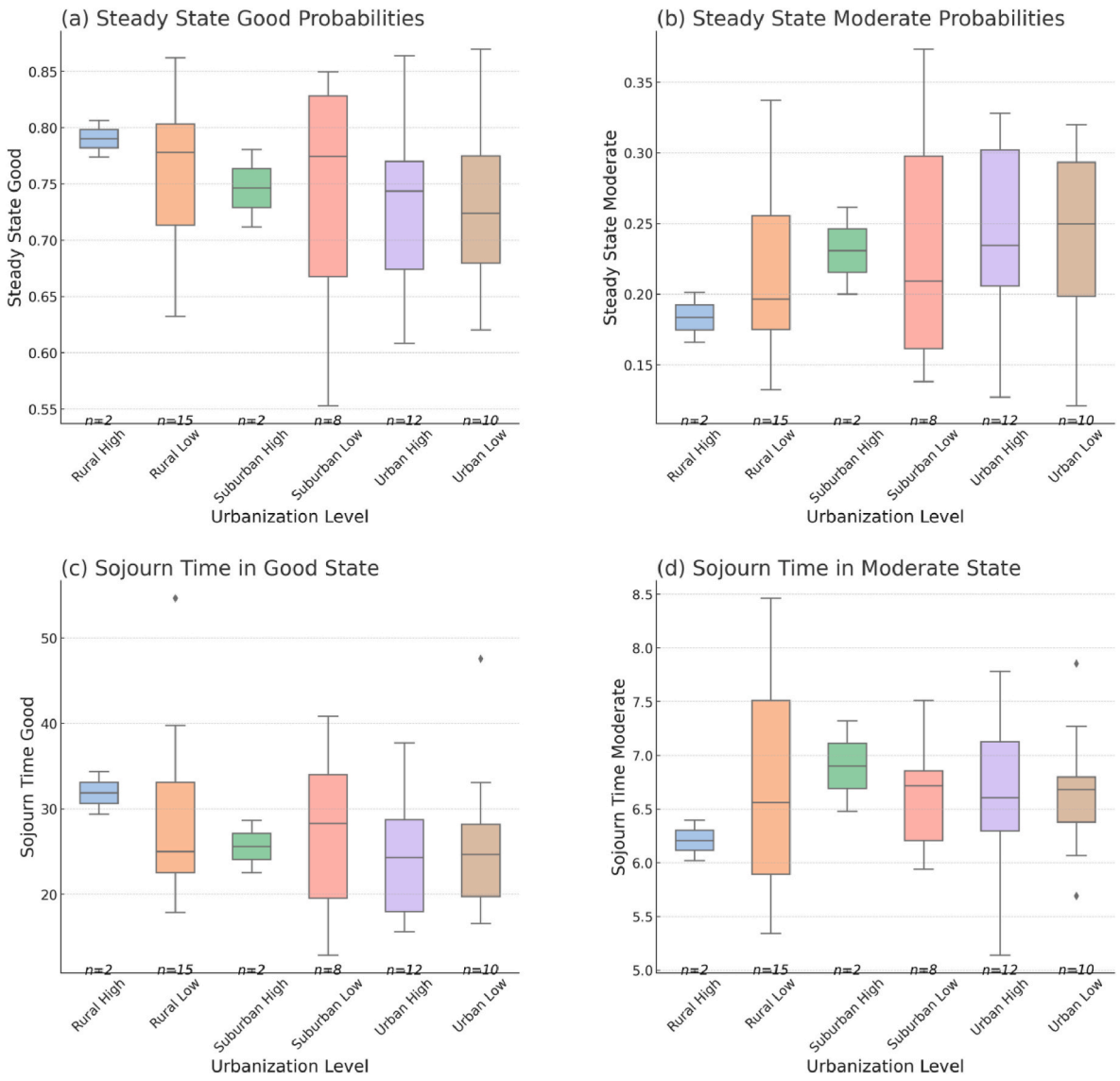


Fig. 8. Boxplot of Markov chain results across different urbanization levels. (a) Steady probabilities of "Good" state; (b) Steady probabilities of "Moderate" state; (c) Sojourn times in "Good" state; (d) Sojourn times in "Moderate" state. "n" represents sample size.

5. Discussion

5.1. Comparison between classic and spatial Markov chains

This study advances the underexplored utilization and comparison of classic and spatial Markov chains in understanding the air quality dynamics. While classic Markov captures dominant state trends, spatial Markov reveals nuanced insights often overlooked. In our study, classic Markov's effectiveness diminishes for less dominant states like "Moderate", contrasting with spatial Markov's comprehensive approach. In essence, classic Markov results are a weighted average across all spatial lags, strongly shaped by the dominant cases, which is evident from result comparison and hypothesis tests.

For scenarios without spatial effects or focusing solely on dominant cases, classic Markov is preferred considering its simpler parameter requirements and explicit results. However, for a holistic view across all scenarios, spatial Markov is superior, particularly in contexts like air pollution, where community and stakeholders concerns often focus on instances of poorer air quality, which may not be dominant cases. These findings align with previous research in various domains, emphasizing the importance of considering spatial effects in Markov processes (Shepero & Munkhammar, 2018; Wang et al., 2023; Wei et al., 2021).

One challenge in spatial Markov, especially at finer scale like hourly, is its computational costs. While fitting the Markov chain is not computationally demanding, defining the necessary neighbor sets can be challenging, as methods like Ordinary Kriging require substantial time and memory resources.

5.2. Effects of time scales on Markov chains

Hourly-scale Markov results reveal a wider range of air pollution states compared to daily-scale results, suggesting potential oversight of short-lived $PM_{2.5}$ spikes. While hourly variation is less than daily variation due to the brief duration of 1-h periods, this scale demonstrates a broader range of states with minimal variation between consecutive timesteps.

Despite short duration of deteriorated air quality states, exposure to air pollution over brief periods can have adverse health impacts (Ai et al., 2019; Cheng et al., 2021; Xu et al., 2023), highlighting the significance of monitoring at hourly intervals. This urges the need to supplement current national ambient air quality standards, which primarily focus on 24-h and annual measurements. While much of the current

research, including deep learning models for $PM_{2.5}$ prediction (e.g., Muthukumar et al., 2022; Qiao et al., 2019; Xiao et al., 2020) and other data-driven models of air quality (e.g., Huang & Kuo, 2018) focus on the daily scale, future research should consider expanding $PM_{2.5}$ modeling to the hourly scale.

5.3. Seasonal and intra-daily trends

Our study reveals distinct seasonal trends in $PM_{2.5}$ air quality, with "Good" state prevailing across all seasons, most prominently and stably in summer, and least in winter, indicating summer as having the most favorable air quality. Our findings align with literature from various regions (Bodor et al., 2020; Schauer et al., 2003), albeit with some discrepancies (Zhao et al., 2018). Lower concentrations were observed during the summer, often below the annual mean. Mornings and nights generally showed higher levels, with midday having the lowest levels except in winter.

During summer, Denton's high surface heating likely creates unstable atmospheric conditions, promoting wind blow and particle dispersion, especially in the absence of high-rise buildings to block the wind. The increase of wintertime $PM_{2.5}$ levels could be attributed to enhanced anthropogenic emissions for large-scale heating and unfavorable meteorological conditions, such as low mixed boundary layer height (Chen et al., 2020). The nighttime worsening of $PM_{2.5}$ levels may be due to stable atmospheric conditions hindering pollutant dispersion. The finding of low midday concentrations implies the role of high temperatures in lowering air pollution, which coincides with the low concentrations observed in summer. The influence of morning and evening rush hour traffic on air quality is evident. Our findings broadly align with existing literature (Javed et al., 2021; Shen et al., 2014; Zhang et al., 2022).

5.4. Spatial patterns and the association with urban form drivers

The site-specific analysis reveals that while rural sites are more likely to maintain "Good" state compared to urban and suburban areas, and they are also less prone to transition out of this state. The trend of decreasing steady probabilities and easier transitions out of "Good" state from urban to suburban to rural areas could be due to their proximity to emission sources. The percentage of impervious surface, a key built environment factors reflecting urbanization levels, positively correlates with lower probability and shorter durations of maintaining in "Good" state. Building height serves as an indicator of urban 3D dimension. In cities like Denton, where most buildings are not high-rise, taller buildings appear to hinder pollution dispersion. The unexpected negative impact of trees on air pollution regulation contrasts with previous findings that highlight their effectiveness in mitigating pollution. Despite their potential to reduce pollutants, trees can also emit particles, particularly when stressed, and act as vertical infrastructure that traps air pollutants. Given Denton's susceptibility to heat stress, further investigation into the role of trees in air pollution mitigation is needed.

5.5. Key findings and implications for urban planning

Spatial Markov results indicate strong evidence for spatial effects across all seasons. If neighbors have "Good" air quality, a location is more likely to remain or transition to "Good", while "Moderate" neighbors imply stability in the "Moderate" state. This spatial dependence aligns with the First Law of Geography, suggesting that improving air quality in one area may require improvements in surrounding areas. Although our study focuses on the intra-county scale, evidence indicates similar effects at larger scales (Liang & Gong, 2020). It is worth noting that within our dataset, no location exhibited spatial lags in more severe air quality states such as "Unhealthy", limiting our ability to determine spatial effects at worse pollution levels. Nonetheless, future research in cities with pronounced air quality issues may shed light on this phenomenon.

The evident seasonality and intra-daily variability also have important implications for policymakers and urban planners. It emphasizes the necessity of considering seasonality when assessing policy effects to ensure credible outcomes. For example, attributing a drop in average $PM_{2.5}$ concentration from May to August solely to a new policy enacted in May requires statistical separation from expected seasonal shifts.

The low concentration during summer and midday highlights the strong influence of meteorological factors on air pollution dispersion. While meteorological conditions cannot be changed, urban planners should consider environmental engineering projects like wind-corridors to facilitate $PM_{2.5}$ dispersion (Beijing Municipal Government, 2017). For traffic management, since morning rush hours are typically followed by favorable meteorological conditions and evening rush hours by unfavorable condition, it is crucial to implement better traffic control during the evening to reduce emissions and prevent worsening nighttime pollution.

6. Conclusions

This study examines the spatiotemporal patterns of $PM_{2.5}$ across various time scales using classic and spatial Markov chain models. The findings reveal distinctive seasonality in $PM_{2.5}$, with the "Good" state prevailing in summer and least in winter. Midday is the time when "Good" state is most pronounced, while mornings and nights exhibit lower prevalence. While our study confirms the role of grey infrastructure in increasing the air pollution levels, it highlights the need for urban planners to consider the adverse impacts of green infrastructure more thoroughly. This study, pioneering the use of spatial Markov chain methods with high spatiotemporal resolution data from LCSs, highlights the potential value of such data and novel approaches for uncovering the nuances aspects of the urban pollution landscape that traditional observational methods often miss. This methodology allows us to identify and analyze fine-scale variations and patterns that are crucial for effective policymaking and health assessments.

Code availability statement

We wrote code for all analyses in Python, leveraging several open-source Python packages. All Markov chain calculations were implemented using the PySAL-giddy Python package (Kang et al., 2019). Kriging computations used the SciKit-GStat library (Mälicke, 2022). Our code is available at https://github.com/lu-liang-geo/MarkovModel.

CRediT authorship contribution statement

Michael Biancardi: Writing – original draft, Validation, Software, Methodology. **Yuye Zhou:** Writing – review & editing, Validation, Visualization, Software. **Wei Kang:** Methodology, Conceptualization. **Ting Xiao:** Investigation. **Tony Grubesic:** Writing – review & editing. **Jake Nelson:** Writing – review & editing. **Lu Liang:** Writing – review & editing, Supervision, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

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

Lu Liang reports financial support and article publishing charges were provided by National Science Foundation. Michael Biancardi reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project is supported by the National Science Foundation (BCS-2117505). The authors would like to thank many student volunteers at the University of North Texas for setting up the calibration system at the Texas site, deploying sensors, and collecting data. Special thanks to Daniel Jacobs for his contributions to this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apgeog.2024.103414.

References

Ai, S. Q., Wang, C. K., Qian, Z., Cui, Y., Liu, Y., Acharya, B. K., Sun, X. Y., Hinyard, L., Jansson, D. R., Qin, L., & Lin, H. L. (2019). Hourly associations between ambient air pollution and emergency ambulance calls in one central Chinese city: Implications for hourly air quality standards. *The Science of the Total Environment, 696*, Article 133956. https://doi.org/10.1016/j.scitotenv.2019.133956

Alexeeff, S. E., Liao, N. S., Liu, X., Van Den Eeden, S. K., & Sidney, S. (2021). Long-term PM2.5 exposure and risks of ischemic heart disease and stroke events: Review and meta-analysis. *Journal of the American Heart Association, 10*(1), Article e016890. https://doi.org/10.1161/JAHA.120.016890

Alyousifi, Y., Ibrahim, K., Kang, W., & Zin, W. Z. W. (2020). Modeling the spatio-temporal dynamics of air pollution index based on spatial Markov chain model. *Environmental Monitoring and Assessment, 192*(11), 719. https://doi.org/10.1007/s10661-020-08666-8

Asadollahfardi, G., Zangooei, H., & Aria, S. H. (2016). Predicting PM2.5 concentrations using artificial neural networks and Markov chain, a case study karaj city. *Asian Journal of Atmospheric Environment, 10*(2), 67–79. https://doi.org/10.5572/ajae.2016.10.2.067

Badach, J., Wojnowski, W., & Gębicki, J. (2023). Spatial aspects of urban air quality management: Estimating the impact of micro-scale urban form on pollution dispersion. *Computers, Environment and Urban Systems, 99*, Article 101890.

Beijing Municipal Government. (2017). Beijing master plan 2016-2035. http://www.bjdch.gov.cn/n1727355/n2669874/n2669876/c6642749/content.htmldoi:10.1016/j.compenvurbsys.2022.101890.

Bi, J., Wildani, A., Chang, H. H., & Liu, Y. (2020). Incorporating low-cost sensor measurements into high-resolution $PM_{2.5}$ modeling at a large spatial scale. *Environmental Science and Technology, 54*(4), 2152–2162. https://doi.org/10.1021/acs.est.9b06046

Bickenbach, F., & Bode, E. (2003). Evaluating the markov property in studies of economic convergence. *International Regional Science Review, 26*(3), 363–392. https://doi.org/10.1177/0160017603253789

Bodor, Z., Bodor, K., Keresztesi, Á., & Szép, R. (2020). Major air pollutants seasonal variation analysis and long-range transport of PM10 in an urban environment with specific climate condition in transylvania (Romania). *Environmental Science and Pollution Research International, 27*(30), 38181–38199. https://doi.org/10.1007/s11356-020-09838-2

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *Lancet (London, England), 360*(9341), 1233–1242. https://doi.org/10.1016/S0140-6736(02)11274-8

Brunekreef, B., Strak, M., Chen, J., Andersen, Z. J., Atkinson, R., Bauwelinck, M., et al. (2021). *Mortality and morbidity effects of long-term exposure to low-level $PM_{2.5}$, BC, NO_2, and O_3: An analysis of european cohorts in the ELAPSE project* (Vol. 2021, p. 208). Research Reports: Health Effects Institute. eCollection 2021.

Caraka, R. E., Chen, R. C., Toharudin, T., Pardamean, B., Yasin, H., & Wu, S. H. (2019). Prediction of status particulate matter 2.5 using state Markov chain stochastic process and HYBRID VAR-NN-PSO. *IEEE Access, 7*, 161654–161665. https://doi.org/10.1109/ACCESS.2019.2950439

Chen, Z., Chen, D., Zhao, C., Kwan, M. P., Cai, J., Zhuang, Y., Zhao, B., Wang, X., Chen, B., Yang, J., & Li, R. (2020). Influence of meteorological conditions on PM2. 5 concentrations across China: A review of methodology and mechanism. *Environment International, 139*, Article 105558.

Cheng, J., Tong, S., Su, H., & Xu, Z. (2021). Hourly air pollution exposure and emergency department visit for acute myocardial infarction: Vulnerable populations and susceptible time window. *Environmental Pollution, 288*, Article 117806. https://doi.org/10.1016/j.envpol.2021.117806

Environmental Protection Agency. (2022). *Determinations of attainment by the attainment date, extensions of the attainment date, and reclassification of areas classified as marginal for the 2015 ozone national ambient air quality standards*.

Hamra, G. B., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J. M., et al. (2014). Outdoor particulate matter exposure and lung cancer: A systematic review and meta-analysis. *Environmental Health Perspectives, 122*(9), 906–911. https://doi.org/10.1289/ehp/1408092

Hillier, F. S., & Lieberman, G. J. (2010). *Introduction to operations research* (9th ed.). McGraw-Hill.

Holmes, J., & Hassini, S. (2021). Discrete-time Markov chain modelling of the ontario air quality health index. *Water, Air, and Soil Pollution, 232*(4), 158. https://doi.org/10.1007/s11270-021-05096-1, 1.

Huang, C., Hu, J., Xue, T., Xu, H., & Wang, M. (2021). High-resolution spatiotemporal modeling for ambient PM2.5 exposure assessment in China from 2013 to 2019. *Environmental Science and Technology, 55*(3), 2152–2162. https://doi.org/10.1021/acs.est.0c05815

Huang, C., & Kuo, P. (2018). A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities. *Sensors, 18*(7), 2220. https://doi.org/10.3390/s18072220

Javed, M., Bashir, M., & Zaineb, S. (2021). Analysis of daily and seasonal variation of fine particulate matter (PM2.5) for five cities of China. *Environment, Development and Sustainability, 23*(8), 12095–12123. https://doi.org/10.1007/s10668-020-01159-1

Kalisa, W., Zhang, J., Igbawua, T., Henchiri, M., Mulinga, N., Nibagwire, D., et al. (2023). Spatial and temporal heterogeneity of air pollution in east africa. *The Science of the Total Environment, 886*, Article 163734. https://doi.org/10.1016/j.scitotenv.2023.163734

Kang, W., Rey, S., Stephens, P., Malizia, N., Wolf, L. J., Lumnitz, S., et al. (2019). *Pysal/giddy: Giddy 2.2.1*.

Larkin, A., & Hystad, P. (2017). Towards personal exposures: How technology is changing air pollution and health research. *Current environmental health reports, 4*, 463–471.

Lei, J., Chen, R., Liu, C., Zhu, Y., Xue, X., Jiang, Y., et al. (2023). Fine and coarse particulate air pollution and hospital admissions for a wide range of respiratory diseases: A nationwide case-crossover study. *International Journal of Epidemiology*. , Article dyad056. https://doi.org/10.1093/ije/dyad056

Liang, L. (2021). Calibrating low-cost sensors for ambient air monitoring: Techniques, trends, and challenges. *Environmental Research, 197*, Article 111163. https://doi.org/10.1016/j.envres.2021.111163

Liang, L., & Daniels, J. (2022). What influences low-cost sensor data calibration? - a systematic assessment of algorithms, duration, and predictor selection. *Aerosol and Air Quality Research, 22*(9), Article 220076. https://doi.org/10.4209/aaqr.220076

Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., & South, J. (2023). Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions. *Environmental Pollution, 331*(Pt 1), Article 121832. https://doi.org/10.1016/j.envpol.2023.121832, 1987.

Liang, L., & Gong, P. (2020). Urban and air pollution: A multi-city study of long-term effects of urban landscape patterns on air quality trends. *Scientific Reports, 10*(1), Article 18618. https://doi.org/10.1038/s41598-020-74524-9

Locosselli, G. M., Camargo, E. P.d., Moreira, T. C. L., Todesco, E., Andrade, M. F., André, C. D., et al. (2019). The role of air pollution and climate on the growth of urban trees. *Science of the Total Environment, 666*, 652–661. https://doi.org/10.1016/j.scitotenv.2019.02.291

Mahler, B., Băiceanu, D., Panciu, T. C., Florea, R. M., Iorga, A. L., Gnat, M., et al. (2023). Air pollutants and their impact on chronic Diseases—a retrospective study in bucharest, Romania. *Atmosphere, 14*(5), 867. https://doi.org/10.3390/atmos14050867

Mälicke, M. (2022). SciKit-GStat 1.0: A SciPy-flavored geostatistical variogram estimation toolbox written in python. *Geoscientific Model Development, 15*(6), 2505–2532. https://doi.org/10.5194/gmd-15-2505-2022

Murray, C. J. L., Aravkin, A. Y., Zheng, P., Abbafati, C., Abbas, K. M., Abbasi-Kangevari, M., et al. (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet, 396*(10258), 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2

Muthukumar, P., Cocom, E., Nagrecha, K., Comer, D., Burga, I., Taub, J., et al. (2022). Predicting PM2.5 atmospheric air pollution using deep learning with meteorological data and ground-based observations and remote-sensing satellite big data. *Air Quality, Atmosphere and Health, 15*(7), 1221–1234. https://doi.org/10.1007/s11869-021-01126-3

Plantower. (2016). *Digital universal particle concentration sensor*.

Qiao, W., Tian, W., Tian, Y., Yang, Q., Wang, Y., & Zhang, J. (2019). The forecasting of PM2.5 using a hybrid model based on wavelet transform and an improved deep learning algorithm. *IEEE Access, 7*, 142814–142825. https://doi.org/10.1109/ACCESS.2019.2944755

Rey, S. J. (2001). Spatial empirics for economic growth and convergence. *Geographical Analysis, 33*(3), 195–214. https://doi.org/10.1111/j.1538-4632.2001.tb00444.x

Rey, S. J., & Franklin, R. S. (2022). *Handbook of spatial analysis in the social sciences*. Cheltenham, UK ; Northhampton, MA, USA: Edward Elgar Publishing.

Schauer, C., Niessner, R., & Pöschl, U. (2003). Polycyclic aromatic hydrocarbons in urban air particulate matter: Decadal and seasonal trends, chemical degradation, and sampling artifacts. *Environmental Science and Technology, 37*(13), 2861–2868. https://doi.org/10.1021/es034059s

Shen, Z., Cao, J., Zhang, L., Liu, L., Zhang, Q., Li, J., et al. (2014). Day–night differences and seasonal variations of chemical species in PM_{10} over xi'an, northwest China. *Environmental Science and Pollution Research International, 21*(5), 3697–3705. https://doi.org/10.1007/s11356-013-2352-z

Shepero, M., & Munkhammar, J. (2018). Spatial Markov chain model for electric vehicle charging in cities using geographical information system (GIS) data. *Applied Energy, 231*, 1089–1099. https://doi.org/10.1016/j.apenergy.2018.09.175

Sneed, M. (2022). 14 ways to beat Dallas traffic. https://www.sneedmitchell.com/post/beat-dallas-traffic. (Accessed 5 August 2024).

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., et al. (2013). The changing paradigm of air pollution monitoring. *Environmental Science and Technology, 47*(20), 11369–11377. https://doi.org/10.1021/es4022602

Song, J., Ding, Z., Zheng, H., Xu, Z., Cheng, J., Pan, R., et al. (2022). Short-term PM1 and PM2.5 exposure and asthma mortality in jiangsu province, China: What's the role of neighborhood characteristics? *Ecotoxicology and Environmental Safety, 241*, Article 113765. https://doi.org/10.1016/j.ecoenv.2022.113765

TexasView. (n. d). How bad is fort worth traffic? URL: https://texasview.org/how-bad-is-fort-worth-traffic/(accessed on August 5, 2024).

US Census Bureau. (2023). Growth in the nation's largest counties rebounds in 2022. https://www.census.gov/newsroom/press-releases/2023/population-estimates-counties.html.

Wang, A., & Brauer, M. (2014). Review of next generation air monitors for air pollution. Available from: https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0132725.

Wang, Y., Chen, F., Wei, F., Yang, M., Gu, X., Sun, Q., et al. (2023). Spatial and temporal characteristics and evolutionary prediction of urban health development efficiency in China: Based on super-efficiency SBM model and spatial Markov chain model. *Ecological Indicators, 147*, Article 109985. https://doi.org/10.1016/j.ecolind.2023.109985

Wang, Y., Xiao, S., Zhang, Y., Chang, H., Martin, R. V., Van Donkelaar, A., et al. (2022). Long-term exposure to PM2.5 major components and mortality in the southeastern United States. *Environment International, 158*, Article 106969. https://doi.org/10.1016/j.envint.2021.106969

Xiao, F., Yang, M., Fan, H., Fan, G., & Al-Qaness, M. A. A. (2020). An improved deep learning model for predicting daily PM2.5 concentration. *Scientific Reports, 10*(1), Article 20988. https://doi.org/10.1038/s41598-020-77757-w

Xu, Q., Guan, Q., Lu, Y., Xu, J., Deng, S., Dong, C., et al. (2023). Effect of short-term ambient air pollution exposure on early miscarriage and pregnancy hormones with critical window identification. *Journal of Hazardous Materials.* , Article 132328. https://doi.org/10.1016/j.jhazmat.2023.132328

Yang, X., Feng, Z., & Chen, Y. (2023). Study on the spatio-temporal patterns of urban air pollution and its spatial mismatch with air pollutant emissions in the yellow river basin, China. *Environmental Research Communications, 5*(5), Article 55008. https://doi.org/10.1088/2515-7620/acd0f4

Zhang, M., Jia, J., Wang, B., Zhang, W., Gu, C., Zhang, X., et al. (2022). Source apportionment of fine particulate matter during the day and night in lanzhou, NW China. *International Journal of Environmental Research and Public Health, 19*(12), 7091. https://doi.org/10.3390/ijerph19127091

Zhao, N., Liu, Y., Vanos, J. K., & Cao, G. (2018). Day-of-week and seasonal patterns of PM2.5 concentrations over the United States: Time-series analyses using the prophet procedure. *Atmospheric Environment, 192*, 116–127. https://doi.org/10.1016/j.atmosenv.2018.08.050